



**HAL**  
open science

# Méthodes hiérarchiques pour l'optimisation géométrique de structures rayonnantes

Benoît Chaigne

► **To cite this version:**

Benoît Chaigne. Méthodes hiérarchiques pour l'optimisation géométrique de structures rayonnantes. Sciences de l'ingénieur [physics]. Université Nice Sophia Antipolis, 2009. Français. NNT: . tel-00429366

**HAL Id: tel-00429366**

**<https://theses.hal.science/tel-00429366v1>**

Submitted on 2 Nov 2009

**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.

UNIVERSITÉ DE NICE-SOPHIA ANTIPOLIS – UFR SCIENCES  
ÉCOLE DOCTORALE SCIENCES FONDAMENTALES ET APPLIQUÉES

# THÈSE

présentée en vue d'obtenir le grade de Docteur en Sciences,  
spécialité « Mathématiques appliquées »

par

Benoît CHAIGNE

## MÉTHODES HIÉRARCHIQUES POUR L'OPTIMISATION GÉOMÉTRIQUE DE STRUCTURES RAYONNANTES

Thèse dirigée par Jean-Antoine DÉSIDÉRI  
et codirigée par Claude DEDEBAN

soutenue le 27 octobre 2009 devant le jury composé de :

M.	CHRISTIAN PICHOT	Directeur de recherche CNRS	(Président)
M.	MOHAMED MASMOUDI	Prof. Université Paul Sabatier	(Rapporteur)
M.	GIUSEPPE PELOSI	Prof. Université de Florence	(Rapporteur)
M.	CLAUDE DEDEBAN	Ingénieur de recherche Orange Labs, La Turbie	(Examineur)
M.	JEAN-ANTOINE DÉSIDÉRI	Directeur de recherche INRIA	(Examineur)



# REMERCIEMENTS

Avant toute chose je souhaiterais préciser que cette thèse a été possible grâce au soutien financier de la région Provence-Alpes-Côte d'Azur et de Orange Labs. Pour cela je les en remercie.

Au-delà de l'aspect financier, je voudrais remercier toute l'équipe d'Orange Labs du Fort de la Tête de Chien à La Turbie. Je pense en particulier à Raymond Bills qui a pris le temps de m'initier aux joies des Surfaces Rayonnantes à Symétrie de Révolution avant de rejoindre les terrasses ensoleillées de Èze. Je remercie Claude Dedebean pour ses conseils avisés en matière d'électromagnétisme et son enthousiasme. Je mentionne également Jean-Louis Desvilles et Patrice Brachat pour leur soutien technique et leur expertise.

Je remercie les professeurs Mohamed Masmoudi et Giuseppe Pelosi d'avoir accepté de rapporter cette thèse et de m'avoir fait part de leurs commentaires précieux. Je les remercie également de s'être déplacés de loin afin de participer à la soutenance. Ma gratitude va aussi au professeur Christian Pichot qui est venu présider le jury.

Je remercie Antoine Désidéri d'avoir bien voulu diriger cette thèse et de m'avoir apporté sa rigueur scientifique, en particulier grâce à son exigence en matière de rédaction et de communication. Je souhaite exprimer toute ma gratitude à l'ensemble de l'équipe Opale ; je citerais Régis Duvigneau dont la ponctualité des pauses caféinées ont été l'occasion d'autant de discussions enrichissantes ; Louis Blanchard, pour sa connaissance étendue du répertoire de Carlos et sa recette du vin chaud ; et Montserrat Argente, tant pour son aide indéfectible que pour ses talents d'organisatrice de randonnées en tout genre dans l'arrière-pays.

Bien entendu, je suis reconnaissant envers tous ceux qui m'ont soutenu, parfois sans le savoir. En vrac : Julien et ses styles  $\LaTeX$  qui m'ont épargné de longues heures perdues en quête d'esthétisme typographique ; Louis, Julien, Joseph, Clément et Thibaud pour leur médiocre sens de l'humour ; Birgitta, Patrick, Olivier et tous les fous de Vallis Bonna pour avoir supporté le mien (certes encore moins bon) ; Sylvain, Christian ( $\times 2$ ), Marc, Germain, Bruno, et autres souffleurs de biniou, gratteurs de cordes molles, saltimbanques et troubadours ; Cécile, Philippe, Yannick et Thomas, la fabuleuse équipe percussive d'Antibes ; enfin les bucois (et assimilés) dont certains ont fait le déplacement.

Pour finir, une pensée toute particulière à ma famille.

# TABLE DES MATIÈRES

<b>Table des matières</b>	<b>iv</b>
<b>Introduction</b>	<b>1</b>
<b>1 Formulation de problèmes de conception optimale d'antenne</b>	<b>5</b>
1.1 Équations de Maxwell	7
1.1.1 Formulation générale	7
1.1.2 Dépendance temporelle harmonique : système de Helmholtz	7
1.1.3 Conditions aux limites et conditions d'interface	8
1.2 Rayonnement des antennes à réflecteur	9
1.2.1 Modélisation d'une antenne en émission	9
1.2.2 Simulation de la propagation d'ondes en espace libre	11
1.2.3 Le logiciel SRSR	15
1.3 Caractéristiques typiques des antennes	16
1.3.1 Énergie et puissance électromagnétique en espace libre diélectrique	16
1.3.2 Intensité et directivité	17
1.3.3 Diagramme de rayonnement	18
1.3.4 Coefficient de réflexion d'une source guidée	19
1.4 Fonctions objectifs pour la conception optimale d'antenne	20
1.4.1 Problèmes inverses sur la puissance rayonnée en zone lointaine	21
1.4.2 Minimisation des pertes dans la source guidée	23
1.4.3 Optimisation multipoint sur une bande de fréquence	23
1.4.4 Optimisation bicritère	23
<b>2 Optimisation numérique de forme : géométrie et algorithmes</b>	<b>25</b>
2.1 Généralités	27
2.1.1 Formulation d'un problème d'optimisation de forme	27
2.1.2 Conditions d'optimalité	27
2.1.3 Méthodologie générale	29
2.2 Représentation de la géométrie	30
2.2.1 Approche continue	30
2.2.2 Approche topologique	30
2.2.3 Approche <i>CAD-free</i> (non paramétrique)	30
2.2.4 Approche paramétrique classique	31
2.2.5 Approche paramétrique <i>Free-Form</i>	35
2.2.6 Déformation de la surface par un champ normal	37
2.2.7 Contraintes sur les variables de conception	39
2.3 Algorithmes d'optimisation numérique	41
2.3.1 Enjeux	41

2.3.2	Algorithmes de descente . . . . .	41
2.3.3	Algorithmes sans gradient . . . . .	44
2.3.4	Hybridation . . . . .	47
2.4	Diagnostic de la convergence . . . . .	47
2.4.1	Taux et vitesse de convergence . . . . .	47
2.4.2	Critère d'arrêt . . . . .	48
2.5	Validation des algorithmes et illustration de cas pathologiques . . . . .	49
2.5.1	Algorithmes de descente et illustration de la raideur . . . . .	49
2.5.2	Algorithme PSO et illustration de la robustesse . . . . .	53
2.5.3	Problème inverse sur la puissance rayonnée . . . . .	55
	Conclusion . . . . .	56
<b>3</b>	<b>Algorithme bigrille idéal pour l'optimisation de forme</b>	<b>57</b>
3.1	État de l'art et enjeux . . . . .	59
3.2	Définition d'un problème modèle géométrique . . . . .	59
3.2.1	Problème de meilleure approximation . . . . .	60
3.2.2	Fonctionnelle paramétrique . . . . .	60
3.2.3	Application à la paramétrisation P1 . . . . .	61
3.3	Méthodes itératives de base . . . . .	66
3.3.1	Rappels sur la convergence d'une itération linéaire . . . . .	66
3.3.2	Itération de Jacobi . . . . .	67
3.3.3	Accélération de Tchebychev, méthode de Richardson . . . . .	67
3.3.4	Méthode de plus grande descente . . . . .	67
3.3.5	Préconditionnement par une méthode de type quasi-Newton . . . . .	68
3.3.6	Fonction d'amortissement . . . . .	68
3.4	Algorithmes multigrilles . . . . .	71
3.4.1	Opérateurs de prolongement et de restriction . . . . .	72
3.4.2	Enrichissement progressif : <i>Nested Iteration</i> . . . . .	73
3.4.3	Algorithme bigrille idéal . . . . .	74
3.4.4	MG, FMG, FAS . . . . .	75
3.4.5	Application des algorithmes classiques pour l'optimisation de forme . . . . .	76
3.4.6	Redéfinition des opérateurs de transfert . . . . .	78
3.4.7	Révision de l'algorithme bigrille idéal . . . . .	80
3.4.8	Expériences numériques . . . . .	82
3.5	Influence de la paramétrisation sur la raideur . . . . .	89
3.5.1	Paramétrisation de Bézier . . . . .	89
3.5.2	Paramétrisation B-splines . . . . .	92
3.5.3	Polynômes orthogonaux . . . . .	94
3.5.4	Influence de la dimension sur la raideur . . . . .	94
3.5.5	Expériences numériques . . . . .	95
	Conclusion . . . . .	101
<b>4</b>	<b>Méthodes hiérarchiques pour l'optimisation paramétrique d'une antenne</b>	<b>103</b>
4.1	Étude du hessien paramétrique pour un problème inverse . . . . .	105
4.1.1	Modèle approché d' <i>Optique Physique</i> . . . . .	105
4.1.2	Modèle électromagnétique . . . . .	108
4.1.3	Conclusion sur l'étude des hessiens paramétriques . . . . .	111
4.2	Hiérarchisation des paramètres de conception . . . . .	111
4.2.1	Prolongement classique : élévation du degré . . . . .	111
4.2.2	Prolongement basé sur le hessien paramétrique . . . . .	111
4.2.3	Approximation par le modèle géométrique : filtre passe-haut . . . . .	112

4.3	Algorithmes hiérarchiques en optimisation globale . . . . .	112
4.3.1	Extension de l'algorithme PSO . . . . .	112
4.3.2	Expériences numériques de validation en <i>Optique Physique</i> . . . . .	114
4.3.3	Application à l'optimisation d'un diagramme en directivité d'une antenne . . . . .	117
4.4	Algorithmes hiérarchiques en optimisation locale . . . . .	119
4.4.1	Algorithme multiniveau non-linéaire . . . . .	119
4.4.2	Expériences numériques . . . . .	121
	Conclusion . . . . .	132
<b>5</b>	<b>Optimisation multipoint sur une bande de fréquence</b>	<b>135</b>
5.1	Généralités sur l'optimisation multicritère . . . . .	137
5.1.1	Formulation d'un problème d'optimisation multicritère . . . . .	137
5.1.2	Optimalité au sens de Pareto . . . . .	137
5.2	Méthodologies d'optimisation multicritère . . . . .	139
5.2.1	Optimisation d'une combinaison convexe des objectifs . . . . .	139
5.2.2	Minimisation du pire des cas . . . . .	139
5.2.3	Recherche de front de Pareto . . . . .	139
5.2.4	Cas particulier de l'optimisation multipoint . . . . .	140
5.3	Méthode hiérarchique sur les points de fonctionnement . . . . .	141
5.3.1	Description du cas-test . . . . .	141
5.3.2	Diagramme initial sur une bande de fréquence . . . . .	142
5.3.3	Conception optimale à une fréquence . . . . .	142
5.3.4	Conception optimale sur une bande de fréquence . . . . .	147
5.3.5	Méthode hiérarchique sur les fréquences . . . . .	148
	Conclusion . . . . .	158
<b>6</b>	<b>Stratégie de moindre sensibilité pour l'optimisation bicritère</b>	<b>159</b>
6.1	Approche d'optimisation bicritère par stratégie de jeux . . . . .	161
6.1.1	Généralités . . . . .	161
6.1.2	Introduction aux jeux de Nash . . . . .	161
6.1.3	Résultat de partage optimal de territoire . . . . .	162
6.1.4	Méthodologie pour l'optimisation d'une antenne . . . . .	163
6.2	Optimisation d'une antenne compacte à une fréquence . . . . .	163
6.2.1	Description du cas-test . . . . .	164
6.2.2	Optimisation du coefficient de réflexion . . . . .	165
6.2.3	Optimisation du diagramme . . . . .	166
6.2.4	Étude numérique de la sensibilité des fonctionnelles au second ordre . . . . .	167
6.2.5	Optimisation dans le sous-espace de moindre sensibilité . . . . .	170
6.3	Application pour l'optimisation d'une antenne sur une bande de fréquence . . . . .	171
6.3.1	Coefficient de réflexion comme critère principal . . . . .	171
6.3.2	Diagramme de rayonnement comme critère principal . . . . .	175
6.3.3	Synthèse . . . . .	179
	Conclusion . . . . .	180
	<b>Conclusion et perspectives</b>	<b>181</b>
<b>A</b>	<b>Annexes</b>	<b>185</b>
A.1	Rayon spectral des algorithmes bigrille idéaux . . . . .	187
A.1.1	Transformation de similarité des matrices d'amplification . . . . .	187
A.1.2	Rayon spectral du cycle idéal classique . . . . .	191

A.1.3	Rayon spectral du cycle idéal en optimisation . . . . .	193
A.1.4	Rayon spectral du cycle idéal avec réorganisation des valeurs propres . . . . .	195
A.1.5	Rayon spectral du cycle idéal algébrique . . . . .	199
A.2	Développement logiciel : plate-forme d'optimisation . . . . .	200
A.2.1	Structure logique des composants de la plate-forme . . . . .	200
A.2.2	Librairies externes . . . . .	200
A.2.3	Extension de la plate-forme . . . . .	201

<b>Bibliographie</b>		<b>203</b>
----------------------	--	------------





# INTRODUCTION

**Introduction au contexte physique** Une antenne est un dispositif destiné à recevoir ou émettre une information (texte, image, son, etc.) sous la forme d'une onde électromagnétique (EM). Typiquement, la transmission d'une information met en jeu une antenne émettrice et une ou plusieurs antennes réceptrices. On distingue généralement deux domaines d'étude d'un système de transmission : d'une part le codage de l'information et d'autre part la propagation de l'onde.

La discipline dite du *traitement du signal* examine les aspects du codage et du décodage de l'information. Le *signal* désigne ainsi l'onde électromagnétique en tant que support de l'information. Sans entrer dans les détails, précisons simplement que ce signal est caractérisé par une onde porteuse simple (monochrome). L'information à transmettre se traduit alors par des (petites) variations de fréquences et/ou d'amplitude de ce signal monochrome (modulations).

Dans un autre cadre d'étude, celui qui nous intéresse, on se préoccupe du phénomène de la propagation des ondes électromagnétiques en laissant de côté l'information codée. Il s'agit essentiellement de déterminer l'énergie transmise d'un système à un autre : en réception, le rôle d'une antenne consiste à « récupérer » l'énergie EM de l'espace libre ; en émission il s'agit de répartir cette énergie dans l'espace libre de telle sorte que les récepteurs soient en mesure de la récupérer. Du point de vue énergétique, lorsque l'on s'intéresse à la propagation d'un signal de l'émetteur vers le récepteur, on néglige les modulations de l'onde porteuse et on se restreint à l'étude de la propagation d'une onde monochrome dont la *fréquence de fonctionnement* correspond à un *canal de transmission*.

Pour être opérationnelle, une antenne émettrice est soumise à des contraintes de qualité généralement exprimées dans un cahier des charges. Par exemple, la puissance transportée par l'onde émise dans l'espace libre doit respecter certains seuils : d'une part le signal doit être suffisamment fort dans la zone où des antennes réceptrices sont susceptibles de se trouver et d'autre part il doit rester faible dans la zone complémentaire afin d'éviter les interférences avec d'autres dispositifs de transmission. Plus précisément, la qualité d'une antenne se mesure à l'aide de grandeurs physiques caractéristiques. La conception d'une antenne revient donc à chercher des configurations (forme des composants de l'antenne, choix des matériaux, type d'alimentation, etc.) qui satisfont ces critères physiques pertinents.

En vue du développement de techniques d'aide à la conception, la qualité d'un système est modélisée par une fonction numérique qui dépend de ces grandeurs physiques. Cette fonction est communément appelée *fonction coût* ou *fonction objectif* ou encore *fonctionnelle* (puisque'elle dépend d'une fonction qui est le champ électromagnétique). Ainsi, on peut définir la conception optimale d'une antenne comme la minimisation (ou la maximisation) d'une fonction objectif.

**Méthodes multiniveaux pour l'optimisation de forme de systèmes soumis à des EDP** Les méthodes numériques de résolution d'équations aux dérivées partielles (EDP) ont atteint une maturité qui leur permet de participer à des procédés d'aide à la conception en ingénierie. Parallèlement, le domaine de l'optimisation s'est considérablement développé avec les infrastructures de calcul. On dispose aujourd'hui d'un certain nombre d'algorithmes pour la résolution de problèmes de *programmation non-linéaire*, déterministes (méthodes de descente, simplexe) comme stochastiques (recuit simulé, algorithmes évolutionnaires, algorithmes par essaim de particules, etc.). Il paraît donc naturel de combiner ces deux domaines afin de proposer des stratégies automatiques d'aide à la conception basées sur la simulation. Pour cela, en vue de son optimisation, le système étudié est avant tout décrit au travers de paramètres de contrôle, discrets ou continus : composants, matériaux, topologie, géométrie, etc. Ce vaste domaine pose de nombreuses questions, de la théorie aux architectures informatiques en passant par la méthodologie de résolution.

On distingue deux grandes difficultés majeures qui peuvent mettre en défaut les algorithmes d'optimisation. La première est liée à la formulation du problème. En effet, la notion de conception optimale est basée sur la définition d'une fonction coût qui constitue un modèle de la qualité du système. S'il peut arriver que cette fonction possède un minimum global unique, il existe souvent de nombreux minima locaux vers lesquels les algorithmes sont susceptibles de converger prématurément (multimodalité). La deuxième est liée au processus itératif des algorithmes, qui sont soumis aux difficultés numériques classiques dues à la raideur (convergence lente).

C'est dans ce contexte que s'inscrit la présente thèse. Elle vise à contribuer à la méthodologie d'aide à la conception optimale d'une d'antenne émettrice, ou plus généralement d'une « structure rayonnante ». En particulier on s'attachera à formuler et examiner des stratégies hiérarchiques pour l'optimisation de la forme de réflecteurs. Sur la base de représentations hiérarchiques de la géométrie, ces méthodes s'inspirent en partie des méthodes multigrilles. D'une part elles cherchent à rendre les algorithmes moins sensibles à la multimodalité (robustesse), d'autre part elles cherchent à améliorer la vitesse de convergence.

**Optimisation multicritère** La conception de systèmes complexes en ingénierie s'inscrit naturellement dans le cadre de l'optimisation multicritère : l'enjeu n'est généralement pas exprimé sous la forme d'un critère unique mais au contraire par un ensemble de requêtes mesurables, pertinentes selon le contexte physique, et qui constitue le cahier des charges. D'un point de vue mathématique, chaque requête est une fonction objectif à minimiser. Intuitivement, on comprend que la difficulté d'un tel problème repose sur l'antagonisme des objectifs : l'optimisation d'un des critères est susceptible de dégrader la performance des autres. On est donc contraint de réaliser un compromis entre les critères.

Parmi les objectifs classiquement considérés on trouve (par ordre croissant de difficulté) : (1) l'optimisation multipoint : les objectifs ne diffèrent que par un paramètre de fonctionnement qui agit sur l'état ; (2) l'optimisation monodisciplinaire : on considère plusieurs objectifs de nature différente mais dont la physique sous-jacente reste la même ; (3) l'optimisation multidisciplinaire : les objectifs diffèrent par leur contexte physique (électromagnétisme, mécanique du solide, mécanique des fluides, etc.).

En pratique, une antenne émettrice n'est pas conçue pour transmettre un seul signal à une fréquence unique de fonctionnement. Au contraire une antenne efficace doit être capable d'émettre plusieurs signaux distincts, chacun caractérisé par une fréquence de fonctionnement donnée. Idéalement, le cahier des charges doit être respecté sur une bande de fréquence. On est donc confronté à un problème d'optimisation multipoint où le paramètre de fonctionnement est la fréquence du régime harmonique. Dans le cadre de cette thèse on propose également une stratégie hiérarchique qui vise à réduire le temps de calcul et augmenter la robustesse des algorithmes de base.

Enfin, nous serons amenés à évaluer la qualité d'une antenne relativement à deux critères. Le premier critère est lié à la puissance rayonnée dans l'espace libre, le deuxième est lié à la puissance émise à la source ou plutôt aux pertes de puissance dans une source guidée. Il peut arriver que ces deux critères soient couplés dans des configurations dites « compactes » où la source et les surfaces rayonnantes sont proches. Les techniques largement utilisées pour la résolution d'un tel problème font appel aux notions d'optimalité de Pareto pour fournir un ensemble de compromis. Ces techniques s'avèrent toutefois très coûteuses. On conduit ici une étude numérique qui défend le point de vue des jeux dynamiques pour la recherche d'un compromis.

**Plan de la thèse** Le chapitre 1 est dédié à la simulation de la propagation d'ondes EM et à la définition de fonctions objectifs. Dans le chapitre 2 on pose le problème d'optimisation à proprement parler, en définissant d'une part le contrôle au moyen de paramètres géométriques et en décrivant les principaux algorithmes de résolution. On en profitera pour rappeler les résultats de caractérisation d'un minimum. Le chapitre 3 est dédié à l'étude théorique d'un algorithme bigrille idéal pour l'optimisation de forme. Par analogie aux algorithmes multigrilles dans le contexte classique de la résolution d'un système linéaire issu de la discrétisation d'une EDP linéaire, une étude de la convergence sera conduite. Les résultats de ce chapitre motivera le développement des algorithmes multiniveaux décrits au chapitre 4 et appliqués à l'optimisation d'une antenne. Enfin les chapitres 5 et 6 traitent de problèmes multicritères : dans le cha-

---

pitre 5 on propose une stratégie multiniveau sur le calcul de la quadrature d'une fonctionnelle dépendant d'un paramètre de fonctionnement. Dans le chapitre 6 on conduit une expérimentation numérique d'un problème bicritère qui défend l'approche par des stratégies de jeux.



# FORMULATION DE PROBLÈMES DE CONCEPTION OPTIMALE D'ANTENNE

---

<b>1.1</b>	<b>Équations de Maxwell</b>	<b>7</b>
1.1.1	Formulation générale	7
1.1.2	Dépendance temporelle harmonique : système de Helmholtz	7
1.1.3	Conditions aux limites et conditions d'interface	8
<b>1.2</b>	<b>Rayonnement des antennes à réflecteur</b>	<b>9</b>
1.2.1	Modélisation d'une antenne en émission	9
1.2.2	Simulation de la propagation d'ondes en espace libre	11
1.2.3	Le logiciel SRSR	15
<b>1.3</b>	<b>Caractéristiques typiques des antennes</b>	<b>16</b>
1.3.1	Énergie et puissance électromagnétique en espace libre diélectrique	16
1.3.2	Intensité et directivité	17
1.3.3	Diagramme de rayonnement	18
1.3.4	Coefficient de réflexion d'une source guidée	19
<b>1.4</b>	<b>Fonctions objectifs pour la conception optimale d'antenne</b>	<b>20</b>
1.4.1	Problèmes inverses sur la puissance rayonnée en zone lointaine	21
1.4.2	Minimisation des pertes dans la source guidée	23
1.4.3	Optimisation multipoint sur une bande de fréquence	23
1.4.4	Optimisation bicritère	23

---

Ce premier chapitre a pour but de familiariser le lecteur avec les notions relatives aux antennes et en particulier aux antennes à réflecteurs. Après avoir brièvement rappelé les notions fondamentales de l'électromagnétisme on énonce les hypothèses de travail du système considéré. L'équation d'état est un problème classique de diffraction en régime harmonique dont on aborde les aspects de la simulation numérique. Enfin, on présente des fonctionnelles du champ électrique rayonné qui caractérisent la qualité des antennes.



## 1.1 Équations de Maxwell

On renvoie le lecteur aux ouvrages de référence [93, 52, 13] pour une présentation détaillée de l'électromagnétisme classique et des équations de Maxwell ainsi qu'aux ouvrages [26, 75] pour l'étude des antennes en particulier.

### 1.1.1 Formulation générale

Les phénomènes électromagnétiques sont décrits par les champs suivants :

$\vec{e}$	champ électrique
$\vec{h}$	champ magnétique
$\vec{d}$	induction électrique
$\vec{b}$	induction magnétique
$\vec{j}$	densité de courant électrique
$\rho$	densité de charge électrique

Ils sont liés entre eux par les équations de Maxwell :

$$\begin{cases} \overrightarrow{\text{rot}} \vec{e} = -\frac{\partial}{\partial t} \vec{b} \\ \overrightarrow{\text{rot}} \vec{h} = \frac{\partial}{\partial t} \vec{d} + \vec{j} \\ \text{div} \vec{d} = \rho \\ \text{div} \vec{b} = 0 \end{cases} \quad (1.1)$$

Le domaine d'étude est composé de milieux (diélectriques) dont la matière est caractérisée par la permittivité électrique et la perméabilité magnétique, respectivement notées  $\varepsilon$  et  $\mu$ . Elles interviennent dans les relations constitutives :

$$\vec{d} = \varepsilon \vec{e} \quad \vec{b} = \mu \vec{h}. \quad (1.2)$$

En général, dans un milieu inhomogène et anisotrope, les grandeurs  $\varepsilon$  et  $\mu$  sont des tenseurs d'ordre 2. Par la suite, en dehors de certains obstacles, seul un milieu diélectrique homogène et isotrope sera considéré. Par conséquent  $\varepsilon$  et  $\mu$  sont équivalents à des constantes réelles positives. Plus précisément, on considérera que le milieu d'étude a les propriétés du vide dont les caractéristiques  $\varepsilon_0$  et  $\mu_0$  sont bien connus.

Par ailleurs, sachant que pour tout champ de vecteur  $\vec{v}$  régulier on a

$$\text{div} \overrightarrow{\text{rot}} \vec{v} = 0$$

et d'après les relations 2 et 3 de (1.1), on obtient la loi de conservation

$$\frac{\partial \rho}{\partial t} + \text{div} \vec{j} = 0. \quad (1.3)$$

### 1.1.2 Dépendance temporelle harmonique : système de Helmholtz

Les champs du système de Maxwell sont des fonctions de l'espace et du temps. On suppose ici que la dépendance temporelle est harmonique. Autrement dit, pour chaque grandeur, la dépendance temporelle est de la forme  $e^{i\omega t}$ . Par convention on note en majuscule la dépendance spatiale, c'est-à-dire si  $\vec{v}$  désigne un champ de vecteurs on a

$$\vec{v}(\mathbf{x}, t) = \vec{V}(\mathbf{x})e^{i\omega t}$$

où  $\mathbf{x}$  désigne un point de l'espace,  $t$  le temps, et  $\omega$  la pulsation. Par conséquent, sachant que  $\frac{d}{dt}(e^{i\omega t}) = i\omega e^{i\omega t}$ , en supprimant par linéarité la dépendance temporelle et en utilisant les relations constitu-



tives (1.2), le système de Maxwell (1.1) devient

$$\begin{cases} \overrightarrow{\text{rot}} \vec{E} &= -i\omega\mu\vec{H} \\ \overrightarrow{\text{rot}} \vec{H} &= i\omega\varepsilon\vec{E} + \vec{J} \\ \text{div} \vec{D} &= \rho \\ \text{div} \vec{B} &= 0 \end{cases} \quad (1.4)$$

On obtient alors le système de Helmholtz en appliquant l'opérateur  $\overrightarrow{\text{rot}}$  aux deux premières relations de (1.4)

$$\begin{cases} \overrightarrow{\text{rot}}(\overrightarrow{\text{rot}} \vec{E}) - k^2 \vec{E} &= -i\omega\mu\vec{J} \\ \overrightarrow{\text{rot}}(\overrightarrow{\text{rot}} \vec{H}) - k^2 \vec{H} &= \overrightarrow{\text{rot}} \vec{J} \end{cases} \quad (1.5)$$

où on a noté  $k^2 = \omega^2 \mu \varepsilon$  le nombre d'onde caractéristique du milieu  $(\varepsilon, \mu)$ . La vitesse de propagation de l'onde  $c$  est définie par  $c^2 = \frac{1}{\mu\varepsilon}$ . On a donc aussi  $k = \frac{\omega}{c} = \frac{2\pi}{\lambda}$ .

Dans le cadre de l'étude d'antennes, la pulsation de l'alimentation est une donnée du problème comme nous le verrons par la suite. On parlera plutôt de fréquence d'alimentation  $f = \frac{\omega}{2\pi}$  (ou encore fréquence de fonctionnement).

### 1.1.3 Conditions aux limites et conditions d'interface

Pour une large classe de problèmes le domaine d'étude est homogène par morceaux : il est constitué de plusieurs milieux dont les conditions à l'interface dépendent du type de matériau.

**Condition d'interface diélectrique** Soit  $\Omega_1$  et  $\Omega_2$  deux milieux diélectriques caractérisés respectivement par  $(\varepsilon_1, \mu_1)$  et  $(\varepsilon_2, \mu_2)$ . À l'interface il y a continuité des composantes tangentielles du champ, soit

$$\vec{n} \wedge \vec{E}_1 = \vec{n} \wedge \vec{E}_2 \quad \text{et} \quad \vec{n} \wedge \vec{H}_1 = \vec{n} \wedge \vec{H}_2 \quad (1.6)$$

lorsque cette interface n'est pas support de densité de courant. Par la suite on ne considérera pas ce type d'interface. En dehors des obstacles (qui ne sont pas des milieux diélectriques) l'espace de propagation est homogène.

**Condition aux limites métallique** Soit un objet métallique parfaitement conducteur de surface  $\mathcal{S}$  régulière (de classe  $C^1$ ) dont la normale extérieure est notée  $\vec{n}$ . Il s'agit d'un cas limite non physique où l'on considère que la conductivité  $\sigma$  de cet objet est infinie. Ce modèle est valide si la différence de conductivité entre deux milieux est telle que  $\sigma \gg \omega\varepsilon$ . Cette propriété s'exprime comme une condition appliquée au champ électrique sur le bord  $\mathcal{S}$  du conducteur parfait : les composantes tangentielles sont nulles, soit

$$\vec{n} \wedge \vec{E} = 0. \quad (1.7)$$

Un conducteur parfait se distingue par le fait que le champ EM n'existe pas à l'intérieur d'un tel objet : ce n'est pas un milieu diélectrique mais un obstacle.

**Condition de rayonnement en espace libre non borné** Pour un problème de propagation en domaine non borné, une condition de rayonnement à l'infini est nécessaire pour assurer l'unicité de la solution. La condition de Sommerfeld (1.8) caractérise la condition de « non retour des ondes » ou la notion de solution sortante<sup>1</sup>.

$$\lim_{r \rightarrow \infty} r \left( \frac{\partial \vec{E}}{\partial r} + ik\vec{E} \right) = 0 \quad (1.8)$$

où  $r$  représente la distance à l'origine d'un système de coordonnées ( $r = |\mathbf{x}|$ ).

<sup>1</sup>Dans la littérature on trouve également  $\lim_{r \rightarrow \infty} r \left( \frac{\partial \vec{E}}{\partial r} - ik\vec{E} \right) = 0$ . Le signe dépend de la convention adoptée pour le régime harmonique :  $e^{i\omega t}$  ou  $e^{-i\omega t}$ . On adopte ici le formalisme de [8] dont est issu le code de simulation.

## 1.2 Rayonnement des antennes à réflecteur

Nous sommes désormais en mesure de modéliser un problème de diffraction électromagnétique. Il s'agit d'écrire les équations dans un domaine qui met en jeu les composants typiques d'une antenne. Une fois les équations posées on abordera les aspects de la résolution numérique.

### 1.2.1 Modélisation d'une antenne en émission

Une antenne à réflecteur en émission est schématiquement composée d'une alimentation (élément actif) et d'un ensemble de surfaces rayonnantes (éléments passifs). L'alimentation se traduit par la connaissance d'une des variables du problème dans un domaine dit intérieur (le complémentaire est appelé domaine extérieur) : elle peut être définie soit comme une densité de courant sur un domaine conducteur (générateur), soit par un champ EM incident au bord du domaine intérieur. Dans le premier cas, l'inconnue est le champ EM dans le domaine extérieur, dans le deuxième cas l'inconnue est le courant induit à la surface des conducteurs du domaine extérieur ; on en déduit alors le champ EM rayonné dans un second temps. On se place dans ce dernier cas.

Soit les domaines ouverts de  $\mathbb{R}^3$  suivants (voir la Figure 1.1) :

- $D$  : le domaine intérieur de bord  $\Sigma$  (élément actif qui sera en général un guide d'onde).
- $A$  : un domaine de bord régulier métallique  $S$  (éléments passifs, ensemble des obstacles rayonnants tels le prolongement du guide d'onde et les réflecteurs).
- $\Omega$  : le domaine extérieur, espace de propagation libre et non borné qui a les caractéristiques du vide (on appelle *espace libre* un espace homogène, isotrope et sans charge). Il est délimité de manière fictive par la *sphère à l'infini*  $S^2$  qui représente les directions de rayonnement de l'antenne.

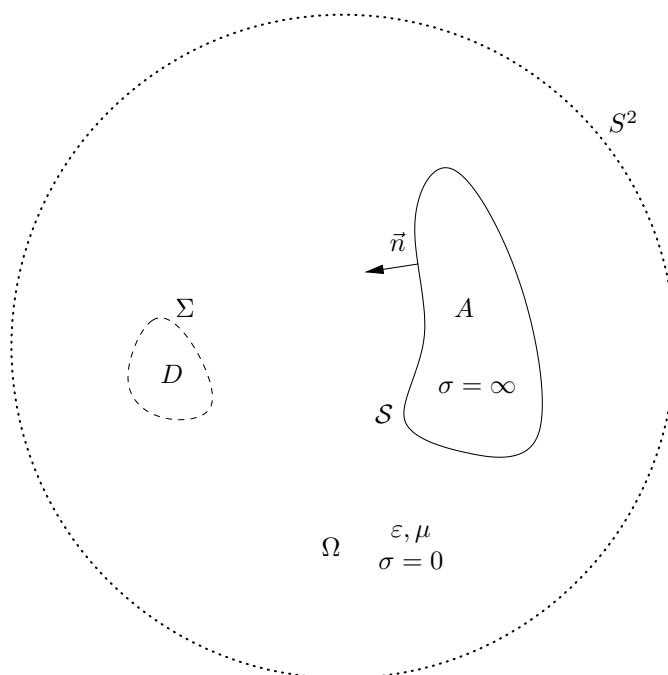


FIG. 1.1 – *Domaine d'étude de la diffraction d'une onde EM émise par un guide d'onde (domaine intérieur  $D$  de bord  $\Sigma$ ) par un réflecteur (domaine  $A$  de surface métallique régulière  $S$ ) dans l'espace libre  $\Omega$ . L'espace de propagation non borné  $\Omega$  est délimité par la surface fictive  $S^2$ , sphère des directions de rayonnement ou sphère à l'infini.*

Puisque  $A$  est un conducteur parfait on considère que le champ est nul à l'intérieur. De même, on ne cherche pas à déterminer le champ dans le domaine  $D$ , toute l'information de l'alimentation est contenue à sa surface  $\Sigma$ . On reviendra sur ce dernier point à la section 1.3.4 lorsqu'on abordera les caractéristiques du guide d'onde.

Dans ce contexte, d'après (1.5), (1.7) et (1.8), l'équation qui régit le champ électrique  $\vec{E}$  en régime harmonique dans le domaine extérieur est

$$\begin{cases} \vec{\text{rot}} \vec{\text{rot}} \vec{E} - k^2 \vec{E} & = 0 & \mathbf{x} \in \Omega \\ \vec{n} \wedge \vec{E} & = 0 & \mathbf{x} \in \mathcal{S} \\ \lim_{r \rightarrow \infty} r \left( \frac{\partial \vec{E}}{\partial r} + ik \vec{E} \right) & = 0 & \mathcal{S}^2 \end{cases} \quad (1.9)$$

**Remarque** Le champ magnétique  $\vec{H}$  peut être obtenu grâce à la relation  $\vec{\text{rot}} \vec{E} = -i\omega\mu\vec{H}$  de l'équation (1.4).

Sauf cas particuliers, les antennes considérées par la suite comprennent deux éléments passifs : le prolongement du guide d'onde (ou cornet) de surface  $\mathcal{S}_g$  et un réflecteur de surface  $\mathcal{S}_r$  ( $\mathcal{S} = \mathcal{S}_g \cup \mathcal{S}_r$ ). On suppose que la géométrie du système présente une symétrie de révolution :  $\Sigma$  est une section du domaine  $D$  cylindrique ; les surfaces  $\mathcal{S}$  sont les surfaces de révolution définies par les courbes génératrices ou *méridiennes*  $\mathcal{C}$ ,  $\mathcal{C}_g$  pour le guide et  $\mathcal{C}_r$  pour le réflecteur (c.f. Figure 1.2 et Figure 1.3, les domaines intérieurs  $D$  et  $A$  ne sont plus représentés). L'alimentation est équivalent au champ électromagnétique donné sur la section amont  $\Sigma$  du guide d'onde.

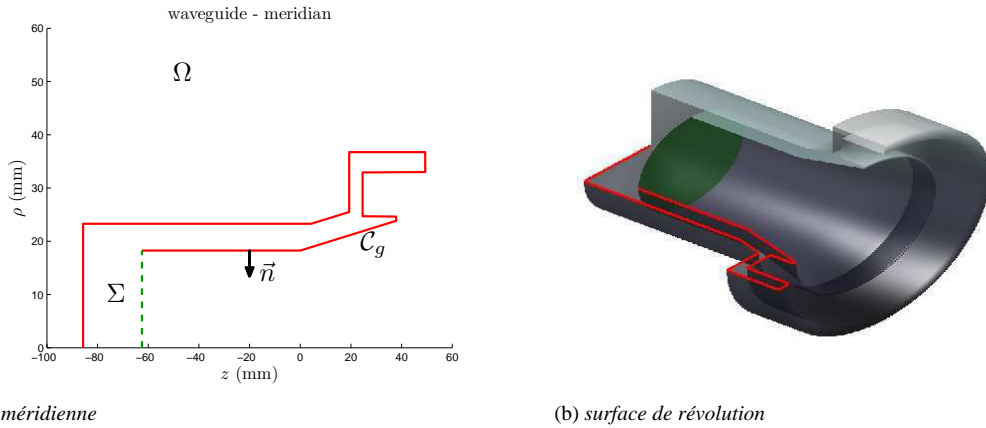


FIG. 1.2 – Guide d'onde de section circulaire uniforme. On suppose que le champ transverse est connu dans une section amont  $\Sigma$  au delà duquel se trouve le domaine intérieur  $D$  (non représenté ici). Le guide est prolongé par la surface parfaitement conductrice  $\mathcal{S}_g$ , surface de révolution générée par la courbe  $\mathcal{C}_g$ .

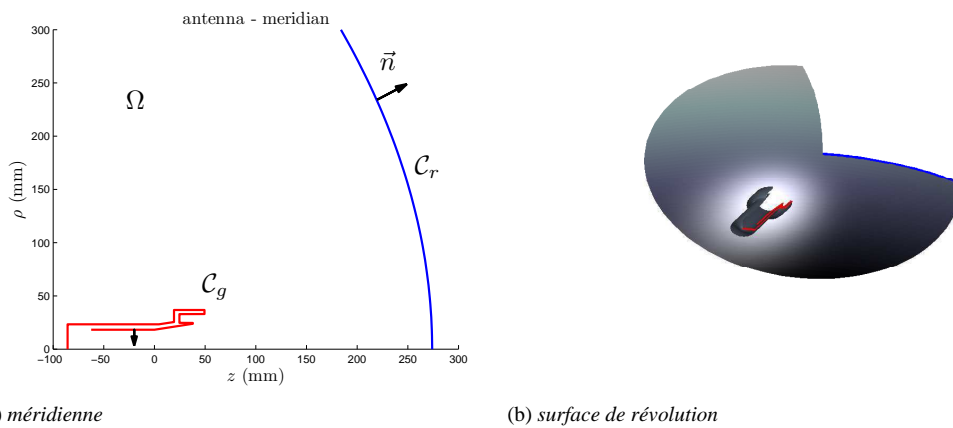


FIG. 1.3 – Éléments passifs d'une antenne à symétrie de révolution : prolongement d'un guide d'onde ( $\mathcal{C}_g$ ) et réflecteur (paraboloïde,  $\mathcal{C}_r$ ).

**Remarque** Dans certains cas l'antenne sera alimentée par un dipôle, source ponctuelle idéale, dont le rayonnement est connu analytiquement. On ne considérera alors qu'un réflecteur comme élément passif.

### 1.2.2 Simulation de la propagation d'ondes en espace libre

Pour pouvoir évaluer les grandeurs utiles à l'estimation de la qualité de l'antenne, il est nécessaire de déterminer le champ EM rayonné par l'antenne en un point de l'espace  $\Omega$  et en particulier en zone lointaine, là où des antennes réceptrices sont susceptibles de se trouver.

D'un point de vue qualitatif la démarche est la suivante :

1. le champ électromagnétique incident connu en amont du guide (supposé parfait dans un premier temps, c'est-à-dire de longueur infinie) induit des courants à la surface des supports conducteurs du domaine extérieur (le cornet et/ou le(s) réflecteur(s)) ; ce champ est harmonique de fréquence  $f$  ;
2. le système de Helmholtz (1.9) exprimé sous la forme d'une équation intégrale (principe de Huygens) est résolu par une méthode de collocation (éléments finis de surface) pour déterminer la densité de courant surfacique sur  $\mathcal{S}$  ; ce problème est bien posé [75, 13] ;
3. une partie du champ rayonné est réfléchi à l'intérieur du domaine  $D$ , c'est-à-dire dans le guide parfait ; un coefficient de réflexion modélise cette perte qui est déduite des solutions obtenues précédemment (par linéarité des équations) ; on obtient alors le rayonnement réel du guide ;
4. ces courants sont alors eux-mêmes sources d'ondes électromagnétiques : le champ électromagnétique rayonné dans l'espace libre résulte de la somme de tous ces champs (propriété due à la linéarité des équations de Maxwell) ;
5. enfin, le champ électrique en zone lointaine est déterminé à partir d'approximations asymptotiques lorsque la distance d'observation tend vers l'infini. Ces approximations sont valides dans la zone dite de Fraunhofer. Dans cette zone le champ EM peut s'écrire sous la forme d'une onde sphérique : elle est le produit d'une fonction dépendant des variables angulaires (direction de rayonnement) et d'une fonction dépendant de la variable radiale  $r$  de décroissance en  $\frac{1}{r}$ .

#### Fonction de Green

Le système (1.9) peut être reformulé plus simplement à l'aide de la relation

$$\overrightarrow{\text{rot}} \overrightarrow{\text{rot}} = \nabla \text{div} - \Delta \quad (1.10)$$

Sachant par ailleurs que le domaine extérieur  $\Omega$  est libre de charge ( $\rho = 0$ ), on a également la relation  $\text{div} \vec{E} = 0$ . Ainsi, pour tout  $\mathbf{x}$  de  $\Omega$  on écrit de manière équivalente  $\overrightarrow{\text{rot}} \overrightarrow{\text{rot}} \equiv -\Delta$ . Il en découle l'équation de Helmholtz vectorielle (1.11) par analogie à l'équation de Helmholtz scalaire.

$$(\Delta + k^2)\vec{E} = 0. \quad (1.11)$$

La fonction de Green  $G$  de l'opérateur scalaire  $\Delta + k^2$  est la solution du problème fondamental

$$\left\{ \begin{array}{ll} (\Delta + k^2)G(\mathbf{x}) & = -\delta \quad \mathbf{x} \in \Omega \\ \lim_{r \rightarrow \infty} r \left( \frac{\partial G}{\partial r} + ikG \right) & = 0 \quad S^2 \end{array} \right. \quad (1.12)$$

où  $\delta$  est la fonction/mesure de Dirac. Le domaine  $\Omega$  étant non borné, la condition de Sommerfeld est nécessaire pour déterminer  $G$  de manière unique. Dans  $\mathbb{R}^3 \setminus \{0\}$  on a

$$G(\mathbf{x}) = \frac{e^{-ik|\mathbf{x}|}}{4\pi|\mathbf{x}|} = \frac{e^{-ikr}}{4\pi r}. \quad (1.13)$$

#### Équations intégrales : principe de Huygens

On note  $M$  un point de  $\Omega$ , appelé point d'observation, et  $M'$  un point du bord  $\partial\Omega$  ( $\Sigma, \mathcal{S}, S^2$ ), appelé point source.  $M$  et  $M'$  sont respectivement de coordonnées  $\mathbf{x}$  et  $\mathbf{x}'$ . On note également  $\mathbf{R} = \mathbf{x} - \mathbf{x}'$  le

vecteur  $\overrightarrow{M'M}$  et  $R = |\mathbf{R}|$  la distance qui sépare ces deux points. Le principe de Huygens exprime le fait que le champ rayonné en un point  $M$  résulte de la contribution des champs et courants sur tout le bord (i.e. de tous les points source  $M'$ ). On retrouve ce principe dans les formules de Stratton-Chu ou encore dans les formules de Kottler.

D'un point de vue mathématique  $\vec{E}$  et  $\vec{J}$  sont à prendre au sens faible (au sens physique on aurait  $\vec{J} = \sigma \vec{E}$ , ce qui n'a pas de sens dans notre cas limite  $\sigma = \infty$ ; cette propriété est traduite par la condition  $\vec{n} \wedge \vec{E} = 0$  sur  $S$ ).  $\vec{J}$  réapparaît dans la formulation variationnelle sous la forme de courant équivalent  $\vec{n} \wedge \vec{H}$ . En adoptant le formalisme des distributions on peut retrouver les formules de Kottler [8].

Sachant que  $\Omega$  est libre, le champ rayonné en tout point  $M$  est donné par [26, 75]

$$\vec{E}(\mathbf{x}) = \frac{1}{i\omega\varepsilon} \int_{\partial\Omega} \left[ k^2 G(\mathbf{R})(\vec{n} \wedge \vec{H}) + ((\vec{n} \wedge \vec{H}) \cdot \nabla_{M'}) \nabla_{M'} G(\mathbf{R}) + i\omega\varepsilon(\vec{n} \wedge \vec{E}) \wedge \nabla_{M'} G(\mathbf{R}) \right] d\Gamma \quad (1.14)$$

On distingue les trois types de la frontière  $\partial\Omega$  :

1. Sur  $\Sigma$  les quantités  $\vec{H}_0^t = \vec{n} \wedge \vec{H}$  et  $\vec{E}_0^t = \vec{n} \wedge \vec{E}$  sont les composantes transverses du champ incident, données du problème (cf. section 1.3.4).
2. Sur  $S$  la condition métallique impose  $\vec{n} \wedge \vec{E} = 0$  et on note  $\vec{J} \stackrel{\text{def.}}{=} \vec{n} \wedge \vec{H}$ , c'est-à-dire le courant équivalent dans le cas d'un conducteur parfait, l'inconnue du problème.
3. Sur  $S^2$  l'intégrale s'annule d'après les conditions aux limites (1.8).

Le champ électrique est donc donné en tout point  $M \in \Omega$  par

$$\vec{E}(\mathbf{x}) = \frac{1}{i\omega\varepsilon} \int_{\Sigma} \left[ k^2 G(\mathbf{R}) \vec{H}_0^t + (\vec{H}_0^t \cdot \nabla_{M'}) \nabla_{M'} G(\mathbf{R}) + i\omega\varepsilon \vec{E}_0^t \wedge \nabla_{M'} G(\mathbf{R}) \right] d\Gamma + \frac{1}{i\omega\varepsilon} \int_S \left[ k^2 G(\mathbf{R}) \vec{J} + (\vec{J} \cdot \nabla_{M'}) \nabla_{M'} G(\mathbf{R}) \right] d\Gamma \quad (1.15)$$

que l'on écrit de manière condensée

$$\vec{E}(\mathbf{x}) = I_{\Sigma}(\vec{E}_0^t, \vec{H}_0^t, \mathbf{x}) + I_S(\vec{J}, \mathbf{x}). \quad (1.16)$$

où  $I_{\Sigma}$ , l'intégrale sur  $\Sigma$ , est linéaire en  $\vec{E}_0^t$  et  $\vec{H}_0^t$ , et  $I_S$ , l'intégrale sur  $S$ , est linéaire en  $\vec{J}$ .

### Méthode de collocation point-segment

On cherche à approcher  $\vec{J}$  dans un espace de dimension finie et choisi *a priori*. Soit un espace de dimension  $N$  généré par la famille libre  $\{\vec{j}_k\}_{k=1}^N$ . La densité de courant se décompose dans cette base comme

$$\vec{J}(\mathbf{x}') = \sum_{k=1}^N J_k \vec{j}_k(\mathbf{x}'). \quad (1.17)$$

Les conditions aux limites de type métallique (1.7) appliquées de part et d'autre de (1.16) nous donne l'équation

$$\vec{n} \wedge I_S(\vec{J}, \mathbf{x}) = -\vec{n} \wedge I_{\Sigma}(\vec{E}_0^t, \vec{H}_0^t, \mathbf{x}), \quad \forall \mathbf{x} \in S \quad (1.18)$$

Le courant (1.17) doit satisfaire (1.18), soit

$$\sum_{k=1}^N J_k [\vec{n} \wedge I_S(\vec{j}_k, \mathbf{x})] = -\vec{n} \wedge I_{\Sigma}(\vec{E}_0^t, \vec{H}_0^t, \mathbf{x}), \quad \forall \mathbf{x} \in S \quad (1.19)$$

Cette équation est vérifiée en  $N$  points  $M_j \in S$  de coordonnées  $\mathbf{x}_j$ , ce qui conduit au système linéaire

$$AJ = b \quad (1.20)$$

où  $A$  est une matrice pleine dont les éléments sont

$$a_{jk} = \vec{n} \wedge I_{\mathcal{S}}(\vec{J}_k, \mathbf{x}_j) \quad (1.21)$$

et  $b$  est le second membre de composantes

$$b_j = -\vec{n} \wedge I_{\Sigma}(\vec{E}_0^t, \vec{H}_0^t, \mathbf{x}_j). \quad (1.22)$$

Cette méthode revient à choisir comme fonctions test les mesures de dirac  $\delta_{M_j}$ . En effet, l'intégrale sur  $\mathcal{S}$  du produit de l'équation (1.19) par  $\delta_{M_j}$  revient au calcul de (1.19) en ce point (d'après la définition de  $\delta$ ).

**Remarque** Lors de l'assemblage de la matrice  $A$  on s'aperçoit que les intégrales deviennent impropres lorsque  $R = 0$ , c'est-à-dire lorsque le point d'observation  $M$  se confond avec le point d'intégration  $M'$ . On montre cependant que les composantes tangentielles de  $I_{\mathcal{S}}$  et  $I_{\Sigma}$  ( $\vec{n} \wedge I_{\mathcal{S}}$  et  $\vec{n} \wedge I_{\Sigma}$ ) sont définies [8].

### Géométrie à symétrie de révolution

La méthode de collocation se simplifie sous l'hypothèse de symétrie axiale des structures rayonnantes. Soit  $\mathbf{x}'$  un point de  $\mathcal{S}$  dont la méridienne  $\mathcal{C}$  est donnée par la courbe paramétrée  $(\rho(s), z(s))$  :

$$\mathbf{x}'(s, \varphi) = \begin{pmatrix} \rho(s) \cos \varphi \\ \rho(s) \sin \varphi \\ z(s) \end{pmatrix} \quad (1.23)$$

où  $\varphi \in [0, 2\pi]$  est l'angle azimutal et  $s$  l'abscisse curviligne de la méridienne.

La densité de courant surfacique est nécessairement périodique suivant  $\varphi$ . Les équations étant linéaires on peut décomposer la dépendance azimutale en modes de Fourier

$$\vec{J}(\mathbf{x}') = \sum_{m \geq 0} \vec{J}(s, m) e^{im\varphi} \quad (1.24)$$

et résoudre le système mode par mode. En pratique, les alimentations considérées [8] seront telles que  $\vec{J}(s, m)$  est non nul pour le seul mode  $m = 1$ , soit (voir la section 1.3.4)

$$\vec{J}(\mathbf{x}') = \vec{J}(s) e^{i\varphi}. \quad (1.25)$$

L'intégration de (1.15) selon  $\varphi$  se fait de manière analytique. Dans ce cas la relation (1.18) peut se réécrire en deux intégrales curvilignes sur  $\mathcal{C}$ . En notant  $(\vec{e}_1, \vec{e}_2)$  un repère local du plan tangent à  $\mathcal{S}$  en  $\mathbf{x}'$ , la méthode de collocation revient à choisir des fonctions de bases de la forme

$$\vec{J}_{k,j}(\mathbf{x}') = \vec{J}_{k,j}(s, \varphi) = f_k(s) e^{i\varphi} \vec{e}_j(s, \varphi) \quad j = 1, 2 \quad k = 1, \dots, N \quad (1.26)$$

ce qui conduit à un système à  $2N$  inconnues ( $N$  fonctions réelles  $f_k$  de la variable curviligne  $s$  et 2 composantes tangentielles par fonction de base).

### Modèle approché d'Optique Physique

Dans certains cas on utilisera une approche simplifiée connue sous le nom d'*Optique Physique*. Ce modèle fournit un cadre idéal pour la validation des méthodes d'optimisation. En effet, le champ rayonné devient alors une fonction explicite de la forme de la structure rayonnante : d'une part la simulation est rapide, d'autre part le gradient ainsi que le hessien de la fonction objectif par rapport à la géométrie sont exacts. Il est utilisé uniquement dans le contexte d'une source ponctuelle (dipôle).

Soit  $(\vec{E}^i, \vec{H}^i)$  le champ incident d'un dipôle. On cherche à approcher la densité de courant sur les conducteurs parfaits. L'approximation d'*Optique Physique* consiste à négliger la courbure du réflecteur :

la densité de courant en chaque point de la surface est approchée par la densité de courant qu'il y aurait sur le plan tangent. En pratique cette approche est justifiée lorsque le rayon de courbure est beaucoup plus grand que la longueur d'onde. Il en résulte une relation explicite entre le champ magnétique incident et la densité de courant [6]

$$\vec{J}_{OP}(\mathbf{x}') = -2\vec{n} \wedge \vec{H}^i(\mathbf{x}'), \quad \mathbf{x}' \in \mathcal{S}. \quad (1.27)$$

On obtient ainsi de manière explicite le champ rayonné dans l'espace libre en sommant le champ incident et le champ diffracté :

$$\vec{E}_{OP}(\mathbf{x}) = \vec{E}^i(\mathbf{x}) + I_{\mathcal{S}}(\vec{J}_{OP}, \mathbf{x}), \quad \mathbf{x} \in \Omega \quad (1.28)$$

où  $I_{\mathcal{S}}$  est l'intégrale sur  $\mathcal{S}$  de l'équation (1.15).

### Approximation en zone lointaine

Une fois les courants équivalents déterminés, on est en mesure de calculer le champ rayonné en tout point de l'espace libre  $\Omega$  grâce à la relation (1.15). On se place dans le cas où le point d'observation  $M$  est situé « loin » de l'antenne, c'est-à-dire loin des zones sources  $\Sigma \cup \mathcal{S}$ . Le sens de « loin » est déterminé par les approximations qui suivent.

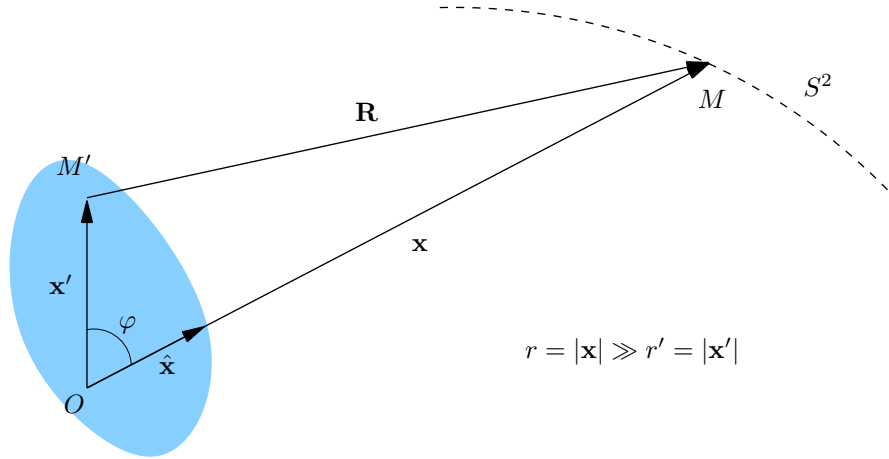


FIG. 1.4 – Configuration géométrique de l'observation du champ EM en zone lointaine.  $M'$  parcourt les zones sources de l'antenne (zone bleue).

Soit  $O$  le centre du repère situé au niveau de l'antenne. On a  $\mathbf{x} = \overrightarrow{OM}$ ,  $\mathbf{x}' = \overrightarrow{OM'}$ , et on note  $\varphi$  l'angle  $(\overrightarrow{OM}, \overrightarrow{OM'})$  (voir Figure 1.4). La distance  $R = |\overrightarrow{MM'}|$  s'écrit donc

$$R = |\mathbf{x} - \mathbf{x}'| = \sqrt{|\mathbf{x}|^2 - 2\mathbf{x} \cdot \mathbf{x}' + |\mathbf{x}'|^2} = r \sqrt{1 - 2\frac{r'}{r} \cos \varphi + \frac{r'^2}{r^2}} \quad (1.29)$$

où on a noté  $r = |\mathbf{x}|$  et  $r' = |\mathbf{x}'|$ . D'après le développement limité

$$\sqrt{1+x} = 1 + \frac{x}{2} - \frac{x^2}{8} + O(x^3) \quad (1.30)$$

on peut approcher  $R$  par

$$R = \underbrace{r}_{O(r)} - \underbrace{\mathbf{x}' \cdot \hat{\mathbf{x}}}_{O(1)} + \underbrace{\frac{1}{2r}(r'^2 - (\mathbf{x}' \cdot \hat{\mathbf{x}})^2)}_{O(\frac{1}{r^2})} + O\left(\frac{1}{r^2}\right) \quad (1.31)$$

où  $\hat{\mathbf{x}} = \frac{\mathbf{x}}{r}$  est le vecteur unitaire de la direction d'observation  $\overrightarrow{OM}$ .

On utilise ce développement pour approcher la fonction de Green  $G(\mathbf{R})$  dans (1.15). Au dénominateur (approximation de l'amplitude) on se contente de l'approximation  $R = r + O(1)$  qui est valide lorsque

$r \gg r'$  pour tout  $M'$ . Si on note  $l$  le diamètre de l'antenne cette condition se traduit par  $r \gg l$ . Au numérateur (approximation de la phase) on ne peut pas négliger le terme en  $O(1)$ . En revanche le terme en  $O(\frac{1}{r})$  peut être négligé si  $r \gg \frac{2l^2}{\lambda}$ . On considère alors l'approximation  $R = r - \mathbf{x}' \cdot \hat{\mathbf{x}} + O(\frac{1}{r})$ . En résumé, les conditions de champ lointain qui définissent la zone dite de Fraunhofer (voir Figure 1.5) sont

$$r \gg l \quad \text{et} \quad r \gg \frac{2l^2}{\lambda}. \quad (1.32)$$

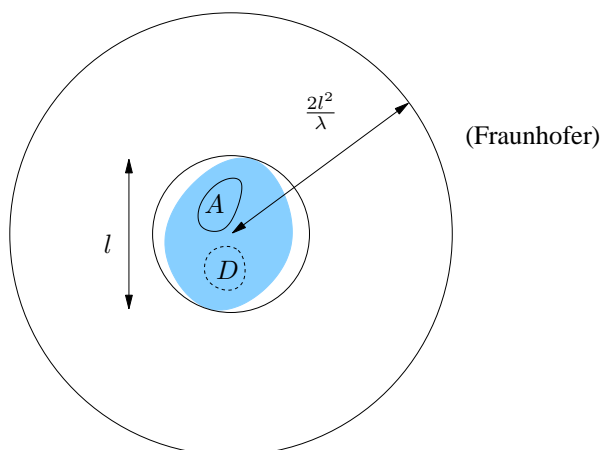


FIG. 1.5 – Zone de Fraunhofer.

D'après (1.15) et en développant les expressions  $\nabla_{M'} G$  et  $(\vec{u} \cdot \nabla_{M'}) \nabla_{M'} G$ , le champ électrique a la forme d'une onde sphérique, c'est-à-dire est tel que la dépendance radiale  $r$  est de la forme  $G(\mathbf{x})$  uniquement, donc découplée de la dépendance angulaire  $\hat{\mathbf{x}}$ , soit

$$\vec{E}(\mathbf{x}) \approx \vec{E}_\infty(\hat{\mathbf{x}}) \frac{e^{-ikr}}{4\pi r}. \quad (1.33)$$

### 1.2.3 Le logiciel SRSR

Le code d'analyse utilisé pour réaliser les simulations est le code SRSR développé à Orange Labs, La Turbie. Il est issu des travaux de thèse [8, 9] pour la simulation des antennes à symétrie de révolution (calcul du champ lointain et du coefficient de réflexion). On en présente les principales caractéristiques.

#### Discrétisation de la géométrie

La géométrie des antennes à symétrie de révolution est définie par leur méridienne  $\mathcal{C}$ . Dans SRSR, sa définition est hiérarchisée en trois niveaux :

- Les *parties* définissent globalement les composants logiques de l'antenne qui sont souvent non connexes : par exemple le guide, le cornet et le(s) réflecteur(s). Chaque partie, notée  $\mathcal{C}_i$ , est une courbe paramétrée :

$$\mathcal{C} = \bigcup_i \mathcal{C}_i \quad \mathcal{C}_i(s) = \begin{cases} \rho(s) \\ z(s) \end{cases} \quad (1.34)$$

- Une partie est continue par morceaux, composée d'un ensemble de *portions* (linéaires, coniques, polynomiales, etc.).
- Enfin, les portions sont discrétisées en *segments*  $\tau_k$ . Ils sont déduits des portions et d'une densité de maillage qui dépend de la longueur d'onde (en général on requiert 5 à 10 segments par longueur d'onde).



En vue de l'optimisation géométrique des antennes nous serons nécessairement amenés à manipuler ces entités. Seules les parties et portions sont accessibles via un fichier de description de la géométrie, les segments étant calculés en dernière instance par une routine interne pour l'assemblage de la matrice et du second membre.

D'un point de vue technique les parties serviront à identifier les zones à optimiser. Ces parties seront composées de portions linéaires uniquement : elles constituent la discrétisation pour l'optimisation. Il y a donc deux niveaux de discrétisation : un pour l'optimisation, et un autre, plus fin, pour l'analyse.

### Fonctions de base de la collocation point-segment

Les fonctions de bases de la collocation sont des fonctions à support local basées sur les segments  $\tau_k$  :

$$\text{supp}(f_k) \stackrel{\text{def.}}{=} \tau_{k-1} \cup \tau_k \cup \tau_{k+1} \quad (1.35)$$

Elles sont définies par

$$f_k(s) = \begin{cases} a_k + b_k \sin(k(s - s_k)) + c_k \sin^2\left(k \frac{s-s_k}{2}\right) & s \in \text{supp}(f_k) \\ 0 & \text{sinon} \end{cases} \quad (1.36)$$

où  $s$  est l'abscisse curviligne,  $s_k$  le milieu du segment  $\tau_k$  et  $(a_k \ b_k \ c_k)$  des paramètres liés au support.

Les points d'observation  $M_j$  sont les points d'abscisses  $s_j$  (i.e. les milieux des segments).

### Résolution du système linéaire

Le système linéaire (1.20) est résolu par la méthode directe de Gauss. La taille de la matrice est de l'ordre de 300 pour les problèmes les plus petits et 2000 pour les plus grands. Le coût total de la simulation en terme de temps CPU est de 0.5 s et 15 s respectivement (assemblage + résolution + calcul du champ lointain). Il s'agit donc de problèmes assez petits, grâce à l'hypothèse de symétrie qui réduit considérablement la taille du problème. Toutefois, dans le cadre d'un algorithme d'optimisation, le coût d'une itération qui requiert une évaluation de fonctionnelle reste dominé par le coût de l'analyse. En comparaison, les opérations réalisées lors d'une itération d'un algorithme d'optimisation sont négligeables.

## 1.3 Caractéristiques typiques des antennes

Les grandeurs qui vont nous permettre d'estimer la qualité d'une antenne sont liées à la puissance rayonnée. On s'intéresse donc à l'énergie portée par une onde EM qui se propage dans l'espace libre.

### 1.3.1 Énergie et puissance électromagnétique en espace libre diélectrique

Soit une onde EM se propageant en espace libre en régime harmonique, soit  $\vec{e}(\mathbf{x}, t) = \vec{E}(\mathbf{x})e^{i\omega t}$  et  $\vec{h}(\mathbf{x}, t) = \vec{H}(\mathbf{x})e^{i\omega t}$ . L'impédance du milieu est donnée par  $\eta = \sqrt{\mu/\varepsilon}$ .

**Définition 1.1.** L'énergie électromagnétique instantanée  $\mathcal{W}(t)$  dans un volume  $V$  de l'espace libre est donnée par

$$\mathcal{W}(t) = \frac{1}{2} \int_V \varepsilon |\vec{e}(\mathbf{x}, t)|^2 + \mu |\vec{h}(\mathbf{x}, t)|^2 dV. \quad (1.37)$$

**Définition 1.2.** La puissance instantanée  $p(t)$  est donnée par

$$p(t) = \frac{d\mathcal{W}}{dt}(t) = - \int_S \vec{p} \cdot \vec{n} dS \quad (1.38)$$

où  $S$  représente le bord de  $V$  et  $\vec{n}$  est la normale extérieure à  $S$ . On appelle  $\vec{p}$  la densité de puissance instantanée, ou vecteur de Poynting. La puissance fournie à l'extérieur de  $V$  est donc égale au flux de la densité de puissance à travers son bord  $S$ .

**Définition 1.3.** En régime harmonique le vecteur de Poynting complexe  $\vec{\mathcal{P}}$  représente la valeur moyenne de la densité de puissance au cours d'une période de durée  $T$ , soit

$$\vec{\mathcal{P}} = \frac{1}{T} \int_0^T \vec{p}(t) dt = \frac{1}{2} \vec{E} \wedge \vec{H}^*. \quad (1.39)$$

La densité de puissance active est la partie réelle du vecteur de Poynting. Si  $\vec{E}$  et  $\vec{H}$  sont en phase alors le vecteur de Poynting est purement réel.

### 1.3.2 Intensité et directivité

Lorsque l'on étudie une antenne à réflecteur on s'intéresse principalement à son rayonnement en champ lointain, où se trouve potentiellement un récepteur. On a défini à la section 1.2.2 la zone lointaine ou *zone de Fraunhofer* par les conditions (1.32). En pratique, pour les antennes considérées (micro-ondes), la condition la plus contraignante est la condition sur la phase  $r \gg \frac{2l^2}{\lambda}$  où  $\lambda$  est la longueur d'onde de la source (monochromatique) et  $l$  est le diamètre du système (encombrement de l'antenne).

Du point de vue de la zone d'observation, l'antenne peut-être assimilée au point  $O$ , centre du repère sphérique. Une direction de propagation  $\hat{\mathbf{x}}$  est équivalente à un vecteur unitaire d'origine  $O$ . L'ensemble des directions de propagation est donc caractérisé par la sphère  $S^2$  centrée en  $O$ . Toute direction  $\hat{\mathbf{x}}$  s'exprime en fonction des variables  $\theta$  (colatitute) et  $\varphi$  (azimut).

On a vu que les formules de Kottler se simplifiaient sous les hypothèses de champ lointain. Le champ EM prend alors la forme d'une onde sphérique

$$\vec{E}(\mathbf{x}) = \vec{E}(r, \theta, \varphi) = \vec{E}_\infty(\theta, \varphi) \frac{e^{-ikr}}{4\pi r}. \quad (1.40)$$

Puisque la dépendance radiale du champ est séparée de la dépendance angulaire, on appelle par abus de langage « champ lointain » la grandeur  $\vec{E}_\infty(\theta, \varphi)$  qui n'a pourtant pas la dimension d'un champ électrique. Elle représente simplement la répartition du champ électrique dans toutes les directions de propagation, sans considérer son atténuation radiale qui est uniforme en  $\frac{1}{r}$ . D'autre part les champs électrique et magnétique sont perpendiculaires, en phase, et leurs amplitudes sont liées par la relation  $\|\vec{E}\| = \eta \|\vec{H}\|$ . Le vecteur de Poynting prend donc la forme suivante :

$$\vec{\mathcal{P}} = \frac{\|\vec{E}(r, \theta, \varphi)\|^2}{2\eta} \hat{\mathbf{x}} = \frac{\|\vec{E}_\infty(\theta, \varphi)\|^2}{2\eta(4\pi r)^2} \hat{\mathbf{x}}. \quad (1.41)$$

On définit une grandeur de puissance indépendante de la distance radiale :

**Définition 1.4.** La puissance rayonnée par unité d'angle solide ou intensité de rayonnement  $U$  est

$$U(\theta, \varphi) = r^2 \|\vec{\mathcal{P}}\| = \frac{\|\vec{E}_\infty(\theta, \varphi)\|^2}{2\eta(4\pi)^2}. \quad (1.42)$$

La puissance totale de rayonnement  $\mathbf{P}_r$  est l'intégrale de l'intensité sur  $S^2$ , soit

$$\mathbf{P}_r = \int_{S^2} U(\theta, \varphi) dS^2 = \int_0^{2\pi} \int_0^\pi U(\theta, \varphi) \sin \theta d\theta d\varphi. \quad (1.43)$$

Il est parfois plus adéquat de manipuler des grandeurs adimensionnalisées :

**Définition 1.5.** La directivité est l'intensité de rayonnement normalisée par la puissance moyenne rayonnée

$$D(\theta, \varphi) = \frac{U(\theta, \varphi)}{\frac{1}{4\pi} \mathbf{P}_r} = \frac{\|\vec{E}(\theta, \varphi)\|^2}{\frac{1}{4\pi} \iint_{S^2} \|\vec{E}(\theta, \varphi)\|^2 \sin \theta d\theta d\varphi}. \quad (1.44)$$

Pour  $\varphi = \varphi_0$  fixé on définit également la directivité comme l'intensité de rayonnement normalisée par la puissance moyenne rayonnée sur ce plan, soit

$$D(\theta) = \frac{U(\theta, \varphi_0)}{\frac{1}{2} \mathbf{P}_r} = \frac{\|\vec{E}(\theta, \varphi_0)\|^2}{\frac{1}{2} \int_0^\pi \|\vec{E}(\theta, \varphi_0)\|^2 \sin \theta d\theta}. \quad (1.45)$$

### 1.3.3 Diagramme de rayonnement

L'argument de symétrie ainsi que des hypothèses sur la nature de l'alimentation (mode TE<sub>11</sub>, cf. section 1.3.4) nous permet de restreindre la mesure de la puissance rayonnée à deux « plans » fixés par les angles azimutaux  $\varphi = 0$  et  $\varphi = \frac{\pi}{2}$  (voir Figure 1.6). Ces plans sont également appelés plans E et plans H car ils contiennent le champ électrique et le champ magnétique respectivement (pour l'alimentation considérée).

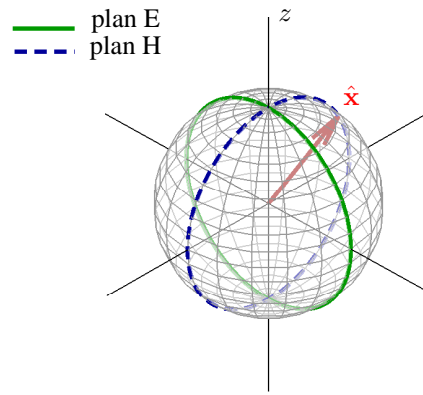


FIG. 1.6 – Sphère  $S^2$  des directions de rayonnement. Le système rayonnant est situé au centre de la sphère et a pour axe de symétrie l'axe  $z$ . Les plans  $\varphi = 0$  (bleu, pointillés) et  $\varphi = \frac{\pi}{2}$  (vert, trait plein) correspondent aux directions de rayonnement que nous considérons par la suite.

La puissance rayonnée sur les plans E et H est représentée dans un diagramme de rayonnement. Un diagramme de rayonnement est la représentation en décibels (db) de l'intensité. L'échelle de puissance en décibels est une échelle relative logarithmique. On définit donc avant tout une intensité de référence  $U_{ref}$ . Il est courant de prendre  $U_{ref}$  comme le maximum de l'intensité, ou encore comme l'intensité de la source isotrope équivalente :

**Définition 1.6.** Une source isotrope est une source idéale non physique telle que l'intensité  $U_i$  est constante dans toutes les directions. La source isotrope équivalente à un système rayonnant est la source isotrope telle que  $U_i = \frac{\mathbf{P}_r}{4\pi}$  (puissance moyenne).

On a donc

$$db(\theta, \varphi) = 10 \log_{10} \frac{U(\theta, \varphi)}{U_{ref}} \quad (1.46)$$

avec dans le cas particulier où  $U_{ref} = U_i$

$$db_i(\theta, \varphi) = 10 \log_{10} D(\theta, \varphi) \quad (1.47)$$

Pour donner un ordre de grandeur, une différence de 3 db correspond à un facteur 2 en puissance.

S'il est naturel de représenter le diagramme en coordonnées polaires, il est d'usage de le représenter en coordonnées cartésiennes. Par exemple, on représente à la Figure 1.7 le diagramme de rayonnement de l'antenne illustrée sur la Figure 1.3 pour une alimentation de fréquence 7,133333 GHz.

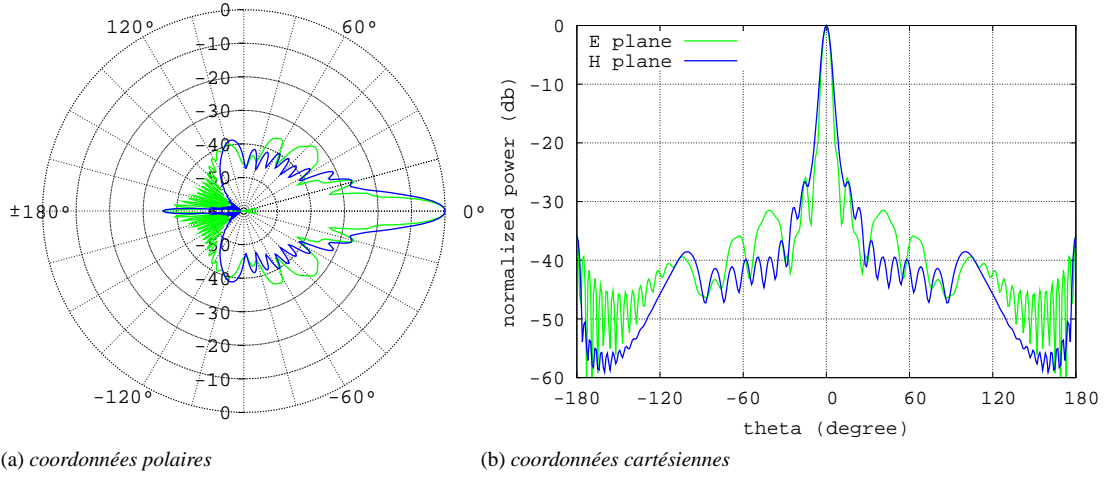


FIG. 1.7 – Diagramme de rayonnement d'un paraboloïde alimenté par un guide d'onde optimisé.  $U_{ref} = U_{max}$ .

### 1.3.4 Coefficient de réflexion d'une source guidée

Le guide d'onde est l'élément source de l'antenne. Il est composé d'un circuit d'alimentation prolongé par un guide de section uniforme  $\Sigma$ . Dans notre cas la section est nécessairement circulaire pour des raisons de symétrie (voir Figure 1.8). Lorsque ce guide est un *guide d'onde parfait* (de longueur infinie), le champ EM ( $\vec{E}_0, \vec{H}_0$ ) sur une section d'un tel guide s'obtient analytiquement [26, 75].

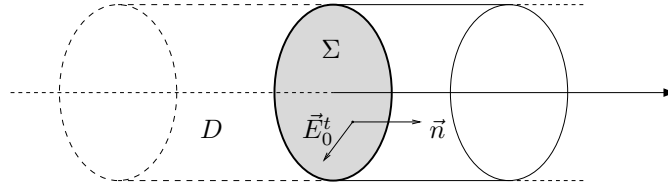


FIG. 1.8 – Schéma d'un guide d'onde parfait de section circulaire.  $\vec{E}_0^t$  représente le champ transverse et  $\vec{n}$  la direction de propagation (longitudinale, normale à la section  $\Sigma$ ).  $D$  est le domaine intérieur qui représente l'alimentation (électronique) à l'intérieur duquel nous ignorons le comportement du champ.

Le mode du guide d'onde circulaire principalement considéré est le mode dit *Transverse Electric* TE<sub>11</sub>. Ce champ est de la forme

$$\begin{aligned}\vec{E}_0^t &= \vec{E}_1 \sin \varphi + \vec{E}_2 \cos \varphi \\ \vec{H}_0^t &= \vec{H}_1 \cos \varphi + \vec{H}_2 \sin \varphi\end{aligned}\quad (1.48)$$

ce qui justifie la représentation du courant (1.25). Les lignes de champ des composantes tangentielles sur une section  $\Sigma$  sont représentées à la Figure 1.9.

Cependant, il n'existe pas de guide d'onde parfait. Le champ rayonné réel est modélisé comme étant la somme d'une onde incidente (parfaite) et d'une onde réfléchie. Une caractéristique essentielle d'un guide d'onde est le coefficient de réflexion :

**Définition 1.7.** Le coefficient de réflexion  $\Gamma$  au niveau d'une section du guide est le rapport des amplitudes complexes des champs électriques transverses réfléchi et incident respectivement. Le champ EM transverse ( $\vec{E}^t, \vec{H}^t$ ) dans une section du guide devient donc

$$\begin{aligned}\vec{E}^t &= (1 + \Gamma) \vec{E}_0^t \\ \vec{H}^t &= (1 - \Gamma) \vec{H}_0^t\end{aligned}\quad (1.49)$$

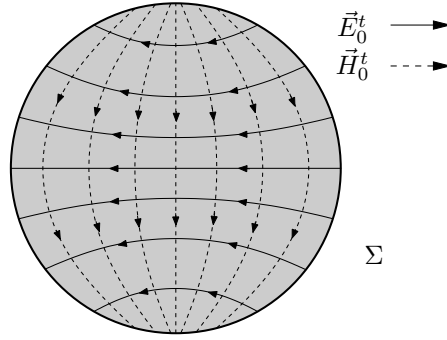


FIG. 1.9 – Lignes de champ des composantes transverse  $\vec{E}_0^t$  (trait plein) et  $\vec{H}_0^t$  (pointillées) du mode principal TE11 d'un guide d'onde de section circulaire  $\Sigma$ .

Du point de vue de la puissance, si on note  $\mathbf{P}_a$  la puissance d'alimentation, la puissance active émise  $\mathbf{P}_e$  est

$$\mathbf{P}_e = (1 - |\Gamma|^2)\mathbf{P}_a \quad (1.50)$$

puisque le vecteur de Poynting s'écrit

$$\begin{aligned} \vec{\mathcal{P}}_e &= \frac{1}{2}\vec{E} \wedge \vec{H}^* \\ &= \frac{1}{2}\vec{E}^t \wedge \vec{H}^{t*} \\ &= \frac{1}{2}(1 + \Gamma)(1 - \bar{\Gamma})\vec{E}_0^t \wedge \vec{H}_0^{t*} \\ &= (1 - |\Gamma|^2 + 2\Im(\Gamma))\vec{\mathcal{P}}_a \end{aligned}$$

On en déduit la relation entre la puissance active d'alimentation et la puissance active émise en prenant la partie réelle de chaque côté de l'égalité.

**Remarque** Dans un milieu sans pertes on s'attend à ce que  $\mathbf{P}_e = \mathbf{P}_r$ . Cependant, en partie à cause de l'approximation en zone lointaine, ce n'est pas toujours vrai. Dans le contexte d'un problème d'optimisation numérique, il est donc plus judicieux de manipuler des grandeurs adimensionnalisées comme la directivité.

Le coefficient de réflexion dépend non seulement du guide mais aussi de la configuration proche de l'antenne (c'est-à-dire des surfaces rayonnantes du cornet  $\mathcal{S}_g$  et éventuellement des réflecteurs  $\mathcal{S}_r$ ). Le calcul de  $\Gamma$  est détaillé dans [8].

En pratique on représente souvent le *Rapport d'Ondes Stationnaires*, ou *impédance relative* au lieu de la norme du coefficient de réflexion :

**Définition 1.8.** Le Rapport d'Ondes Stationnaire ROS (Voltage Standing Wave Ratio VSWR) ou impédance relative  $Z$  est le rapport

$$Z = \frac{1 + |\Gamma|}{1 - |\Gamma|}. \quad (1.51)$$

## 1.4 Fonctions objectifs pour la conception optimale d'antenne

Dans les sections précédentes on a défini des grandeurs d'observation d'une antenne. On également détaillé la méthode de calcul de ces grandeurs basée sur la simulation d'ondes EM en espace libre. On définit désormais des fonctionnelles dont le rôle est d'estimer la qualité d'une antenne et que l'on cherchera à optimiser.

Du point de vue du champ lointain, le principe général de la conception optimale d'antenne est de se rapprocher le plus près possible d'une configuration idéale (champ lointain, puissance, directivité). On

parle alors de problèmes de synthèse ou de reconstruction. En termes mathématiques il s'agit d'un problème inverse : on cherche à déterminer s'il existe un système (une forme) dont le rayonnement conduit à une solution idéale. Ces problèmes sont généralement mal posés et requièrent des méthodes robustes pour s'approcher au mieux de la configuration souhaitée. On répondra en partie à cette question à l'aide de méthodes hiérarchiques sur la géométrie.

Du point de vue de la source il s'agit d'un problème de maximisation de la puissance émise (ou de manière équivalente de minimisation des pertes, c-à-d du coefficient de réflexion). Comme nous le verrons par la suite, il est facile d'optimiser une source à une fréquence donnée. En revanche, il est plus délicat d'obtenir les mêmes performances sur une bande de fréquence, surtout si l'antenne est compacte (c'est-à-dire lorsqu'un réflecteur est proche de l'ouverture du guide et contribue ainsi à un fort coefficient de réflexion).

### 1.4.1 Problèmes inverses sur la puissance rayonnée en zone lointaine

Soit  $\mathcal{O}(\vec{E})$  une observation de la variable d'état, fonction  $L_2$  sur les directions de propagation et  $\tilde{\mathcal{O}}$  une fonction cible idéale. Le problème inverse consiste à minimiser la norme  $L_2$  de la différence entre l'observation et la valeur cible, soit

$$F(\vec{E}) = \frac{1}{2} \int_{S^2} \left\| \mathcal{O}(\vec{E}) - \tilde{\mathcal{O}} \right\|^2 dS^2. \quad (1.52)$$

L'observation  $\mathcal{O}$  peut être le champ lui-même, l'intensité, la directivité, ou d'autres mesures encore (voir [4]).

Si  $\mathcal{O}$  est le champ cela suppose que l'on connaisse toutes ses composantes complexes, c'est-à-dire à la fois la phase et l'amplitude. Il est difficile et peu intéressant de construire une telle cible, sauf si celle-ci vient du calcul d'une structure connue. Il s'agit alors d'un problème de reconstruction qui n'a d'intérêt que pour des exercices de validation.

Un cas plus intéressant consiste à se donner une répartition en puissance. Une des difficultés de ce genre de problème est qu'il n'existe probablement pas de solution. Pour des raisons de normalisation il est plus adéquat de se donner une cible en directivité plutôt qu'en puissance. Un cas largement traité dans cette thèse est le cas d'une cible où la directivité est uniforme sur un sous ensemble de  $S^2$ .

#### Cas d'une cible en directivité uniforme

On cherche à résoudre le problème de minimisation suivant

$$\min_{\vec{E}} F(\vec{E}) = \frac{1}{2} \int_{S^2} \left( D(\vec{E}) - \tilde{D} \right)^2 dS^2. \quad (1.53)$$

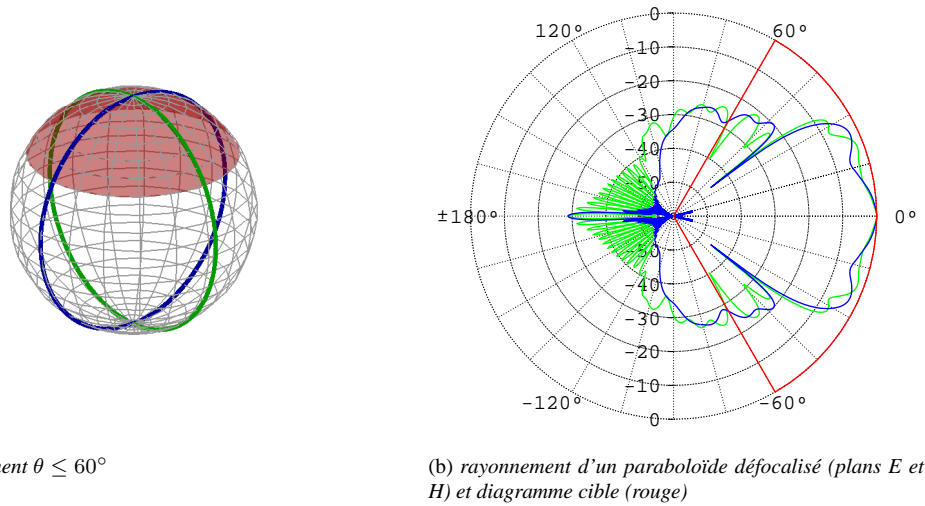
Tout d'abord, sur les plans E et H, notons que

$$\int_0^\pi D(\theta) \sin \theta d\theta = \int_0^\pi \frac{U(\theta)}{\frac{1}{2} \mathbf{P}_r} \sin \theta d\theta = \frac{2}{\mathbf{P}_r} \int_0^\pi U(\theta) \sin \theta d\theta = 2 \quad (1.54)$$

par définition de  $\mathbf{P}_r$ . Donc une cible  $\tilde{\mathcal{O}}$  en directivité doit nécessairement satisfaire cette condition. Soit maintenant l'intervalle  $[\theta_1, \theta_2]$ ,  $0 \leq \theta_1 < \theta_2 \leq \pi$ , pour lequel on souhaite que la directivité soit uniforme et maximale. Pour qu'elle soit uniforme on note trivialement  $\tilde{D}(\theta) = \tilde{D}$  sur  $[\theta_1, \theta_2]$ . Pour que la directivité  $\tilde{D}$  soit maximale il suffit que  $\tilde{D}(\theta)$  soit nulle en dehors de l'intervalle souhaité. La valeur cible  $\tilde{D}$  est donc donnée par la condition (1.54) :

$$\int_0^\pi \tilde{D}(\theta) \sin \theta d\theta = 2 \quad \Rightarrow \quad \int_{\theta_1}^{\theta_2} \tilde{D} \sin \theta d\theta = 2 \quad \Rightarrow \quad \tilde{D} = \frac{2}{\cos \theta_1 - \cos \theta_2} \quad (1.55)$$

On exclut évidemment le cas où  $\theta_1 = \theta_2$ . On donne un exemple à la Figure 1.10 où  $\theta_1 = 0^\circ$  et  $\theta_2 = 60^\circ$ .

(a) Zone de rayonnement  $\theta \leq 60^\circ$ 

(b) rayonnement d'un paraboloïde défocalisé (plans E et H) et diagramme cible (rouge)

FIG. 1.10 – Illustration d'un cas-test d'optimisation d'un diagramme uniforme

### Quelques résultats de calcul différentiel

Considérons la fonctionnelle (1.53) dépendant du champ  $u$ . On calcul sa différentielle par rapport à une variation du champ  $\delta u$ . Par la suite, connaissant les dérivées partielles de  $u$  par rapport à des déformations de la forme d'un réflecteur, par composition de fonctions, en en déduit les dérivées partielles de la fonction coût par rapport à ces déformations.

On adopte les notations suivantes

$$u, \delta u, \delta v \in \mathbf{X} = \mathbb{L}_2(S^2, \mathbb{C}^3) \quad \|z\|^2 = z \cdot z = z^T \bar{z} \quad \Theta \subset S^2 \quad |\Theta| = \int_{\Theta} dS^2$$

où  $u$  est non identiquement nul et  $\Theta$  non vide.

### fonctionnelles

$$\begin{aligned} \text{(puissance moyenne)} \quad \bar{\mathbf{P}}_r : \mathbf{X} &\longrightarrow \mathbb{R}_+ \\ u &\longmapsto \frac{1}{|\Theta|} \int_{\Theta} \|u\|^2 dS^2 \end{aligned}$$

$$\begin{aligned} \text{(directivité)} \quad D : \mathbf{X} &\longrightarrow L_2(\Theta, \mathbb{R}_+) \\ u &\longmapsto \frac{\|u(\cdot)\|^2}{\bar{\mathbf{P}}_r} \end{aligned}$$

$$\begin{aligned} \text{(fonction coût)} \quad F : \mathbf{X} &\longrightarrow \mathbb{R}_+ \\ u &\longmapsto \frac{1}{2} \int_{\Theta} (D(u) - \tilde{D})^2 dS^2 \end{aligned}$$

**dérivées premières**

$$\langle \bar{\mathbf{P}}'_r(u), \delta u \rangle = \frac{2}{|\Theta|} \int_{\Theta} \Re(u \cdot \delta u) dS^2$$

$$D'(u)\delta u = \frac{2\Re(u \cdot \delta u)\bar{\mathbf{P}}_r(u) - \|u\|^2 \langle \bar{\mathbf{P}}'_r(u), \delta u \rangle}{\bar{\mathbf{P}}_r^2(u)}$$

$$\langle F'(u), \delta u \rangle = \int_{\Theta} D'(u)\delta u (D(u) - \tilde{D}) dS^2$$

**dérivées secondes**

$$\langle \bar{\mathbf{P}}''_r(u)\delta v, \delta u \rangle = \frac{2}{|\Theta|} \int_{\Theta} \Re(\delta v \cdot \delta u) dS^2$$

$$D''(u)(\delta v, \delta u) = \frac{1}{\bar{\mathbf{P}}_r^3(u)} \left\{ \bar{\mathbf{P}}_r(u) \left[ 2\Re(\delta v \cdot \delta u)\bar{\mathbf{P}}_r(u) + 2\Re(u \cdot \delta u) \langle \bar{\mathbf{P}}'_r(u), \delta v \rangle \right. \right. \\ \left. \left. - 2\Re(u \cdot \delta v) \langle \bar{\mathbf{P}}'_r(u), \delta u \rangle - \|u\|^2 \langle \bar{\mathbf{P}}''_r(u)\delta v, \delta u \rangle \right] \right. \\ \left. - 2 \left[ 2\Re(u \cdot \delta u)\bar{\mathbf{P}}_r(u) - \|u\|^2 \langle \bar{\mathbf{P}}'_r(u), \delta u \rangle \right] \langle \bar{\mathbf{P}}'_r(u), \delta v \rangle \right\}$$

$$\langle F''(u)\delta v, \delta u \rangle = \int_{\Theta} D'(u)\delta v D'(u)\delta u + (D(u) - \tilde{D}) D''(u)(\delta v, \delta u) dS^2$$

**1.4.2 Minimisation des pertes dans la source guidée**

La formulation d'une fonction objectif pour l'optimisation de la source est triviale. Afin de minimiser le coefficient de réflexion on cherche tout simplement à minimiser

$$F(\Gamma) = \frac{1}{2} |\Gamma|^2 \quad (1.56)$$

où le carré a été rajouté pour des raisons de différentiabilité.

**1.4.3 Optimisation multipoint sur une bande de fréquence**

On désigne désormais par  $u$  la variable d'état qui désigne  $\vec{E}$  ou  $\Gamma$ . Pour une fréquence de fonctionnement  $f$  donnée, on note  $u(f)$  la solution unique de l'équation d'état. Les fonctionnelles précédemment considérées étaient de la forme  $F(u) = F(u(f))$ . Soit désormais une bande de fréquence  $I_f = [f_- f_+]$  et  $\|\cdot\|$  une norme de fonctions de la variable réelle définies sur  $I_f$ . D'une manière générale, l'objectif multifréquence s'écrit sous la forme

$$F_{I_f}(u) = \|F(u(\cdot))\| \quad (1.57)$$

Différentes stratégies peuvent être envisagées, différentiables ou non selon la norme ( $L_p$ , max, moyenne, etc.). On abordera ces problèmes au chapitre 5.

**1.4.4 Optimisation bicritère**

L'optimisation d'une antenne requiert de manière quasi-systématique l'optimisation du rayonnement et de la source. Lorsque le(s) réflecteur(s) et le guide sont faiblement couplés ces deux processus peuvent



être réalisés indépendamment. Dans des configurations compactes d'antennes, c'est-à-dire où la source et le réflecteur sont proches, il est plus délicat d'améliorer un critère sans dégrader l'autre.

L'optimisation multicritère (bicritère ici) n'est pas un problème d'optimisation au sens strict. En effet il n'y a pas de relation d'ordre lorsque plusieurs critères entrent en jeu. On abordera au chapitre 5 les notions de dominance et d'optimalité au sens de Pareto.

# OPTIMISATION NUMÉRIQUE DE FORME : GÉOMÉTRIE ET ALGORITHMES

# 2

---

<b>2.1 Généralités</b> . . . . .	<b>27</b>
2.1.1 Formulation d'un problème d'optimisation de forme . . . . .	27
2.1.2 Conditions d'optimalité . . . . .	27
2.1.3 Méthodologie générale . . . . .	29
<b>2.2 Représentation de la géométrie</b> . . . . .	<b>30</b>
2.2.1 Approche continue . . . . .	30
2.2.2 Approche topologique . . . . .	30
2.2.3 Approche <i>CAD-free</i> (non paramétrique) . . . . .	30
2.2.4 Approche paramétrique classique . . . . .	31
2.2.5 Approche paramétrique <i>Free-Form</i> . . . . .	35
2.2.6 Déformation de la surface par un champ normal . . . . .	37
2.2.7 Contraintes sur les variables de conception . . . . .	39
<b>2.3 Algorithmes d'optimisation numérique</b> . . . . .	<b>41</b>
2.3.1 Enjeux . . . . .	41
2.3.2 Algorithmes de descente . . . . .	41
2.3.3 Algorithmes sans gradient . . . . .	44
2.3.4 Hybridation . . . . .	47
<b>2.4 Diagnostic de la convergence</b> . . . . .	<b>47</b>
2.4.1 Taux et vitesse de convergence . . . . .	47
2.4.2 Critère d'arrêt . . . . .	48
<b>2.5 Validation des algorithmes et illustration de cas pathologiques</b> . . . . .	<b>49</b>
2.5.1 Algorithmes de descente et illustration de la raideur . . . . .	49
2.5.2 Algorithme PSO et illustration de la robustesse . . . . .	53
2.5.3 Problème inverse sur la puissance rayonnée . . . . .	55
<b>Conclusion</b> . . . . .	<b>56</b>

---

Dans ce chapitre on aborde les notions générales liées à l'optimisation numérique de forme, du traitement de la géométrie aux algorithmes. On insiste en particulier sur les représentations hiérarchiques des formes en vue de l'étude de stratégies multiniveaux. On en profite également pour rappeler les grandes classes de méthodes d'optimisation numérique en détaillant les algorithmes que nous utiliserons et en validant les codes correspondants.



## 2.1 Généralités

### 2.1.1 Formulation d'un problème d'optimisation de forme

On se place dans le cadre de l'optimisation d'un système gouverné par une EDP. Dans une formulation typique on désigne par  $\mathcal{S}$  la forme qui définit la géométrie du système étudié. L'équation d'état (1.9) qui régit le système est notée sous sa forme implicite

$$\mathcal{M}(\mathcal{S}, f, u) = 0 \quad (2.1)$$

où  $u$  désigne la variable d'état (champ et/ou coefficient de réflexion) et  $f$  un point de fonctionnement (la fréquence d'alimentation). En admettant que (2.1) soit bien posé, il existe une solution unique que l'on désigne également sous sa forme explicite  $u = u(\mathcal{S}, f)$ .

On cherche à minimiser la fonctionnelle de la variable d'état  $F(u)$  (une parmi celles définies au chapitre 1 pourvu que  $u$  soit physique, c'est-à-dire pourvu que  $u$  satisfasse l'équation d'état (2.1). Formellement on cherche à résoudre le problème de minimisation sous contrainte suivant :

$$\begin{cases} \min_u & F(u) \\ \text{t.q.} & \mathcal{M}(\mathcal{S}, f, u) = 0 \end{cases} \quad (2.2)$$

En utilisant la formulation explicite de  $u$  on considère de manière équivalente la fonctionnelle de forme  $\mathcal{J}$

$$\min_{\mathcal{S}} \mathcal{J}(\mathcal{S}) = F(u(\mathcal{S}, f)). \quad (2.3)$$

Pour minimiser  $\mathcal{J}$  on agit sur la forme  $\mathcal{S}$ . Dans certains problèmes inverses de détection, on peut également considérer la fréquence comme variable d'optimisation [34]. Dans notre cas il s'agit d'un paramètre de fonctionnement imposé.

Pour l'instant le problème n'est pas correctement formulé puisqu'on n'a pas encore donné de sens au contrôle, c'est-à-dire à la manière dont on manipule la géométrie. Un des objectifs de ce chapitre est de préciser ce que l'on entend par « agir sur la forme » et définir un *espace de recherche*. En pratique on considérera une représentation paramétrique de la forme du réflecteur. La paramétrisation consiste à définir un espace de formes de dimension finie. En d'autres termes on considérera des représentations de  $\mathcal{S}$  sous la forme

$$\mathcal{S} = \mathcal{S}(x) \quad x \in \mathbb{R}^n \quad (2.4)$$

où  $x$  est le vecteur des *paramètres de conception*. Le problème intrinsèque (2.3) devient un problème d'optimisation paramétrique de dimension finie

$$\min_{x \in \mathbb{R}^n} J(x) \stackrel{\text{def}}{=} \mathcal{J}(\mathcal{S}(x)). \quad (2.5)$$

### 2.1.2 Conditions d'optimalité

Le problème d'optimisation en dimension finie (2.5) est éventuellement soumis à des contraintes additionnelles d'égalité et/ou d'inégalité sur les paramètres de conception. Il s'écrit alors comme un problème classique d'optimisation :

$$\begin{cases} \min_{x \in \mathbb{R}^n} & J(x) \\ \text{t.q.} & g(x) = 0 \\ & h(x) \leq 0 \end{cases} \quad (2.6)$$

où  $g : \mathbb{R}^n \rightarrow \mathbb{R}^{m_e}$  et  $h : \mathbb{R}^n \rightarrow \mathbb{R}^{m_i}$ <sup>1</sup>.

On introduit quelques définitions utiles.

<sup>1</sup>Les relations d'égalité et d'inégalité entre deux vecteurs sont définies ici terme à terme.

**Définition 2.1.** L'espace admissible est le sous-ensemble de  $\mathbb{R}^n$  dont les éléments satisfont les contraintes, soit

$$\mathcal{F} \equiv \{x \in \mathbb{R}^n \mid g(x) = 0, h(x) \leq 0\} \quad (2.7)$$

**Définition 2.2.** Une contrainte d'inégalité  $i$  est dite active en  $x$  si  $h_i(x) = 0$ . On note  $I_0$  l'ensemble des contraintes actives en  $x$ , soit tel que  $h_{I_0}(x) = 0$ .

Une notion importante mais délicate à aborder est la notion de *qualification des contraintes* lorsque  $g$  et  $h$  sont dérivables (voir [74] par exemple). Il s'agit d'une notion locale qui caractérise le fait que la linéarisation des contraintes représente bien les directions admissibles. Une condition suffisante pour assurer la qualification des contraintes est l'indépendance linéaire des gradients de l'ensemble des contraintes d'égalité et des contraintes actives. On dit alors que les contraintes sont qualifiées, ce que l'on supposera toujours par la suite.

On suppose désormais que  $J$ ,  $g$  et  $h$  sont de classe  $C^2$  sur  $\mathbb{R}^n$ . On note  $\nabla J(x)$  le gradient de  $J$ ,  $H(x)$  le hessien, et  $g'$  et  $h'$  les jacobiennes des contraintes. On a les conditions nécessaires et suffisantes d'optimalité suivantes (voir par exemple [45, 11]) :

### Conditions nécessaires d'optimalité du premier ordre (Karush-Kuhn-Tucker)

Soit  $x$  un minimum local de (2.6) dont  $m_a$  contraintes sont actives. Alors il existe  $\lambda \in \mathbb{R}^{m_a}$  et  $\mu \in \mathbb{R}^{m_e}$  (appelés multiplicateurs de Lagrange) tels que

$$\begin{cases} (a) & \nabla J(x) + h'_{I_0}(x)^T \lambda + g'(x)^T \mu = 0 \\ (b) & g(x) = 0 \\ (c) & h(x) \leq 0 \\ (d) & h_{I_0}(x) = 0 \\ (e) & \lambda \geq 0 \end{cases} \quad (2.8)$$

La condition (2.8)e exprime le fait que les directions de descente de  $J$  vont à l'encontre de l'espace admissible. En introduisant le lagrangien

$$\begin{aligned} L : \mathbb{R}^n \times \mathbb{R}^{m_i} \times \mathbb{R}^{m_e} &\longrightarrow \mathbb{R} \\ (x, \lambda, \mu) &\longmapsto L(x, \lambda, \mu) = J(x) + \lambda^T h(x) + \mu^T g(x) \end{aligned} \quad (2.9)$$

la condition (2.8)a s'écrit sous la forme

$$\nabla_x L(x, \lambda, \mu) = 0 \quad (2.10)$$

où on a noté  $\lambda_i = 0$  pour  $i \notin I_0$ .

Dans le cas particulier sans contraintes on a la condition

$$\nabla J(x) = 0 \quad (2.11)$$

que l'on peut retrouver avec un simple développement de Taylor.

La condition (2.8)a revient à dire que le gradient de la fonction coût s'exprime comme une combinaison linéaire des gradients des contraintes actives. Si  $A^T = (h'_{I_0}(x)^T \ g'(x)^T) \in \mathbb{R}^{n \times (m_e + m_i)}$  désigne le transposé de la matrice de linéarisation des contraintes en  $x$ , alors  $\nabla J(x) \in \mathcal{R}(A^T)$ . Or comme  $\mathcal{R}(A^T) = \mathcal{N}(A)^\perp$  et si  $Z$  est une base orthogonale de  $\mathcal{N}(A)$  alors on a

$$Z^T \nabla J(x) = 0 \quad (2.12)$$

qui est souvent utilisé dans les algorithmes en tant que « gradient projeté ». Dans le cas où seules des contraintes d'égalité sont considérées et si  $x$  est admissible ( $g(x) = 0$ ), (2.12) est équivalent à (2.8), le signe des multiplicateurs de Lagrange  $\mu$  n'important pas.

Un point  $x$  qui satisfait (2.8) est appelé *point stationnaire* puisque les variations du premier ordre de  $J$  dans les directions admissibles sont nulles ou positives. Sous des hypothèses de convexité du domaine admissible et de la fonction coût, les conditions (2.8) sont suffisantes pour caractériser le minimum global. Dans le cas général ce point peut toutefois ne pas être un minimum local.

### Conditions suffisantes d'optimalité du second ordre

Des conditions suffisantes s'obtiennent lorsque la fonction coût est localement convexe dans les directions admissibles. Cela se déduit du comportement local au second ordre. On se restreint au cas où un minimum local est associé à un unique couple de multiplicateurs  $(\lambda, \mu)$ , ce qui s'obtient sous des hypothèses de qualification de contraintes.

On introduit la notion de cône critique :

**Définition 2.3.** Le cône critique en un point  $x$  qui vérifie (2.8) est l'ensemble  $C$  des directions

$$C = \{d \in \mathbb{R}^n \mid g'(x)^T d = 0, \quad h'_{I_0}(x)^T d = 0, \quad h'(x)^T d \leq 0\} \quad (2.13)$$

Le point  $x$  est un minimum local dont  $m_a$  contraintes sont actives ssi il existe  $\lambda \in \mathbb{R}^{m_a}$  et  $\mu \in \mathbb{R}^{m_e}$  tels que (2.8) soit satisfait et

$$\forall d \in C(x); d \neq 0 \quad d^T \nabla_{xx}^2 L(x) d > 0 \quad (2.14)$$

où  $\nabla_{xx}^2 L(x)$  désigne le hessien du lagrangien (2.9) par rapport à  $x$ .

Dans le cas sans contraintes,  $x$  est un minimum local ssi

$$\begin{aligned} \nabla J(x) &= 0 \\ H(x) &\text{ est définie positive} \end{aligned} \quad (2.15)$$

Comme pour les conditions nécessaires, si  $Z$  est une base orthogonale du noyau de la jacobienne des contraintes d'égalité et des contraintes actives au point  $x$  alors on a la relation

$$\begin{aligned} Z^T \nabla J(x) &= 0 \\ Z^T H(x) Z &\text{ est définie positive} \end{aligned} \quad (2.16)$$

qui est équivalent à (2.14) dans le cas des seules contraintes d'égalité si  $x$  est admissible.

### 2.1.3 Méthodologie générale

À proprement parler, le problème de minimisation (2.6) consiste en la recherche d'un (du) minimum global. Sans résultat d'existence et d'unicité préalable nous n'avons aucun critère numérique pour caractériser un tel minimum. Alternativement, si l'on est muni d'une routine du calcul du gradient, on peut caractériser un point stationnaire, voire un minimum local si le calcul du hessien (exact ou approché) est accessible, ce qui est rarement le cas en pratique.

L'enjeu de l'optimisation numérique est donc de trouver un « bon » minimum local. La réalisation de cette recherche passe par la construction itérative d'une suite minimisante. On décrit dans les grandes lignes la procédure d'optimisation d'une antenne :

1. Dans un premier temps on identifie les composants de l'antenne que l'on souhaite considérer comme contrôle ; les parties choisies sont remaillées en portions linéaires puis on définit un espace paramétrique du contrôle. La ou les fréquence(s) d'alimentation est/sont définie(s).
2. Étant donné un vecteur de paramètres de conception  $x$  on construit la géométrie  $\mathcal{S}$  pour laquelle on souhaite évaluer la fonctionnelle ; un maillage de  $\mathcal{S}$  est construit (d'un point de vue technique on génère un fichier de description de la méridienne  $\mathcal{C}$  lisible par SRSR).
3. L'équation d'état est résolue par SRSR.
4. La variable d'état  $u$  (champ + coefficient de réflexion) est renvoyée ; on calcule alors la fonction coût ainsi que les contraintes et éventuellement le gradient.
5. Selon la valeur de la fonction coût, le meilleur candidat est mis à jour, on modifie les paramètres de conception et on recommence le processus en 2.

On s'attache maintenant à préciser ce que l'on entend par espace paramétrique, c-à-d le traitement de la géométrie. Puis on discutera des méthodes numériques.

## 2.2 Représentation de la géométrie

### 2.2.1 Approche continue

Cette approche est en premier lieu théorique. Elle consiste à définir les objets mathématiques concernant les formes et leur évolution/déformation. On donne ainsi un sens à une classe de formes pour laquelle la fonction  $\mathcal{J}(\mathcal{S})$  est bien définie [90]. Sous certaines hypothèses on peut formuler un gradient et un hessien de forme. Le calcul du gradient de forme se fait alors par une méthode adjointe dont on réfère l'origine à [17]. La formulation du gradient de forme et son utilisation en vue d'une optimisation numérique a fait l'objet des thèses [68, 34] dans la cadre de problèmes de diffraction électromagnétique. Le gradient numérique est alors l'expression discrète du gradient de forme continue. Cette méthode est cependant délicate à mettre en œuvre et n'est pas toujours adaptée aux codes existants de résolution du problème direct.

### 2.2.2 Approche topologique

Dans l'approche précédente on cherche à optimiser un domaine par déformation de sa frontière. En général cela suppose que la topologie de ce domaine ait été fixée. Dans une autre approche on cherche à optimiser la topologie du domaine pour laquelle on trouve une formulation rigoureuse du gradient dans [64, 89, 18] (gradient topologique). Schématiquement, elle est la formulation asymptotique de l'insertion d'un « trou » dans le domaine à optimiser. Dans le cadre de problèmes d'optimisation en électromagnétisme et de problèmes inverses de rayonnement, ce gradient topologique a été formulé et utilisé numériquement dans des cas concrets (détection et imagerie sous-sol [65, 66], optimisation de guide d'onde [82]). En pratique, son calcul repose aussi sur une méthode adjointe.

On mentionne également la technique des lignes de niveaux (*level set*) qui permet des changements faciles de topologie [76]. Dans cette méthode on représente la frontière d'un domaine par une formulation implicite  $\phi = 0$  (c-à-d la ligne de niveau 0 de la fonction  $\phi$ ). On cherche alors à faire varier cette frontière en agissant sur la fonction  $\phi$ . Si cette méthode permet la séparation d'un domaine originellement connexe ou l'agrégation d'un domaine non connexe, il semble toutefois difficile de faire apparaître des « trous » [2]. Dans le cadre de problèmes inverses de diffraction électromagnétique, la méthode de *level set* a été utilisée dans [42, 79]. De plus, à l'aide d'un gradient de forme, elle a également été appliquée dans [34, 10, 43].

Par la suite, on ne cherchera pas à changer la topologie des parties à optimiser. Par ailleurs, on considérera des représentations discrètes des formes.

### 2.2.3 Approche *CAD-free* (non paramétrique)

Le calcul d'analyse de la densité de courant repose sur la discrétisation des surfaces métalliques. L'ensemble des éléments (triangles, nœuds, segments, etc.) caractérise d'une part le domaine de définition de la variable d'état et d'autre part la structure discrète du domaine à optimiser. Chacun de ces éléments peut donc être considéré comme un degré de liberté du problème d'optimisation. Le maillage constitue donc à la fois le support des inconnues discrètes de l'état et les variables de conception.

Cependant, les caractéristiques du maillage ( finesse, anisotropie, etc.) sont avant tout déterminées par des contraintes liées au calcul d'analyse. Il n'y a pas lieu de penser que cette structure soit adaptée à l'optimisation. En particulier cette approche conduit à des problèmes de grande taille nécessaire pour la précision du calcul de l'état mais potentiellement surdimensionnés du point de vue de l'optimisation. Pour ces raisons on préférera une approche qui sépare la représentation discrète de la géométrie pour le calcul d'analyse de celle en vue de son optimisation.

### 2.2.4 Approche paramétrique classique

De manière générale, une courbe paramétrée  $\mathcal{C}$  de  $\mathbb{R}^2$  est décrite par

$$\mathcal{C}(t) = \begin{cases} x(t) \\ y(t) \end{cases} \quad (2.17)$$

où  $t$  est le paramètre de la représentation de  $\mathcal{C}$  (une abscisse curviligne). Lorsque la paramétrisation s'écrit sous forme polynomiale, c-à-d  $x(t) = \sum_k a_k t^k$  et  $y(t) = \sum_k b_k t^k$  dans la base canonique, alors les coefficients  $a_k$  et  $b_k$  décrivent entièrement  $\mathcal{C}$ . On est donc en mesure de définir des espaces de dimension finie décrivant des formes continues.

Parmi les représentations possibles pour une même courbe  $\mathcal{C}$ , certaines s'avèrent plus adéquates du fait que les coefficients  $a_k$  et  $b_k$  ont un sens géométrique. Ces représentations constituent la base de la *Conception Assistée par Ordinateur* (CAO) ou *Computer Aided Design* (CAD) largement utilisée en ingénierie. De la plus simple à la plus générale, on trouve les paramétrisations suivantes :

- Bézier : représentation polynomiale dans la base de Bernstein ; les coefficients ou points de contrôle forment un polygone qui donne une idée grossière de la forme ; une propriété d'élévation du degré permet une définition hiérarchique ( $p$ -enrichissement) [39].
- B-splines : représentation polynomiale par morceaux ; il s'agit d'une généralisation de la représentation de Bézier ; le support des fonctions de bases est local ; on peut contrôler la continuité de la courbe aux points de jonction des patches ; elles possèdent des propriétés d'enrichissement de type  $p$  (élévation du degré)  $h$  (diminution du support et insertion de fonctions de même degré) ou encore  $k$  ( $h$  et  $p$ ) [27].
- NURBS (Non Uniform Rational B-Splines) : il s'agit d'une représentation par des fonctions rationnelles de B-splines et permet en particulier la représentation exacte de sections coniques [80].
- T-splines : ces dernières ont été introduites récemment ; elles sont une généralisation des NURBS et permettent des raffinements locaux non conformes [87, 85].

Dans notre cas seules les deux premières techniques seront suffisantes pour les configurations considérées.

Bien qu'elle constitue une représentation exacte du point de vue de l'optimisation, l'approche paramétrique classique ne nous affranchit pas d'une procédure de discrétisation des formes ainsi définies dans le but de réaliser le calcul d'analyse.

#### Représentation de Bézier

Les polynômes de Bernstein sont définis pour  $t \in [0, 1]$  par

$$B_n^k(t) = C_n^k t^k (1-t)^{n-k} \quad (2.18)$$

où  $C_n^k = \frac{n!}{k!(n-k)!}$  sont les coefficients binomiaux (voir Figure 2.1).

Une courbe de Bézier de degré  $n$  dans  $\mathbb{R}^p$  est une courbe paramétrée définie par

$$P(t) = \sum_{k=0}^n B_n^k(t) P_k \quad t \in [0, 1] \quad (2.19)$$

où  $P_k \in \mathbb{R}^p$  sont les *points de contrôle* de la courbe. Elle a les propriétés élémentaires suivantes :

- la courbe passe par les points  $P_0$  et  $P_n$  ;
- les vecteurs  $\overrightarrow{P_0 P_1}$  et  $\overrightarrow{P_{n-1} P_n}$  sont tangents à la courbe ;
- la courbe se trouve dans l'enveloppe convexe du polygone de contrôle formé par les points de contrôle ;
- la dérivée et l'intégrale d'une courbe de Bézier est une courbe de Bézier ;

Par ailleurs, on a la propriété essentielle des courbes de Bézier dite d'*élévation du degré* :



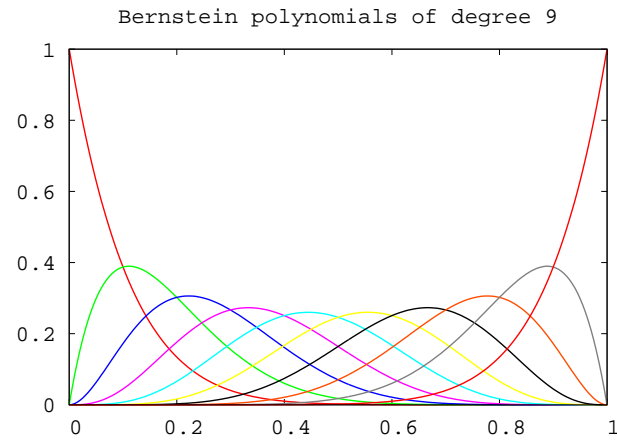
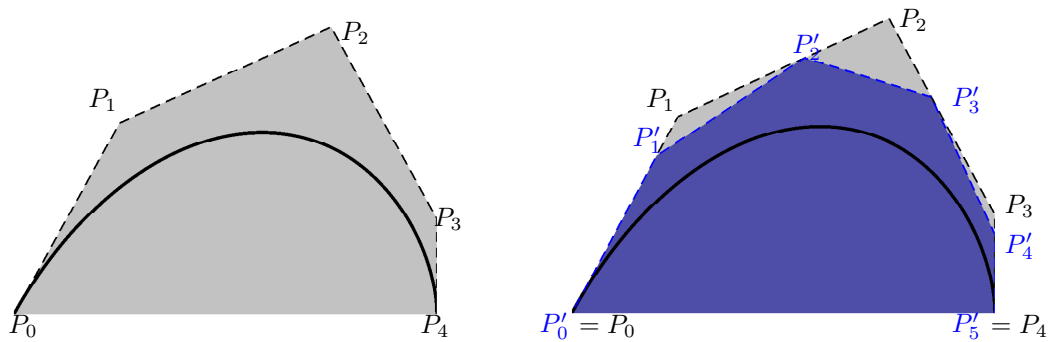


FIG. 2.1 – Polynômes de Bernstein de degré 9.



(a) courbe de Bézier de degré 4

(b) courbe identique exprimée comme une courbe de Bézier de degré 5

FIG. 2.2 – Illustration d'une courbe de Bézier avec ses propriétés élémentaires (polygone de contrôle, enveloppe convexe, et élévation du degré).

**Propriété 2.1.** Soit  $P(t)$  une courbe de Bézier de degré  $n$  définie par les points de contrôle  $P_k$ . Alors  $P'(t)$  est la courbe de Bézier de degré  $n + 1$  définie par les points de contrôle  $P'_k$

$$P'(t) = \sum_{k=0}^{n+1} B_k^{n+1}(t) P'_k$$

tels que

$$\begin{cases} P'_0 &= P_0 \\ P'_k &= \frac{k}{n+1} P_{k-1} + \left(1 - \frac{k}{n+1}\right) P_k \quad k = 1 \dots n \\ P'_{n+1} &= P_n \end{cases} \quad (2.20)$$

En d'autres termes, toute courbe de Bézier de degré  $n$  peut être exprimée comme une courbe de Bézier de degré  $n + 1$  dont les points de contrôle sont des combinaisons convexes des points de contrôle initiaux (c-à-d de la courbe de degré  $n$ ). En appliquant récursivement la propriété d'élévation du degré, une courbe de Bézier de degré  $n$  peut être exprimée comme une courbe de Bézier de degré  $m > n$  en appliquant  $m - n$  fois la propriété 2.1. La courbe de Bézier est la limite lorsque le degré tend vers l'infini du polygone de contrôle. La propriété d'élévation du degré constitue la base des algorithmes hiérarchiques.

Une surface de Bézier est définie par un produit tensoriel de courbes de Bézier, soit la surface paramétrée  $\mathcal{S}(s, t)$

$$\mathcal{S}(s, t) = \sum_{j=0}^{n_j} \sum_{k=0}^{n_k} B_j^{n_j}(s) B_k^{n_k}(t) P_{jk} \quad (2.21)$$

La propriété d'élévation du degré s'étend trivialement aux surfaces de Bézier : l'élévation du degré s'applique alors sur chaque dimension d'espace.

## B-splines

Les fonctions B-splines sont entièrement définies par récurrence à partir des données suivantes :

- un degré  $d$  (ou ordre  $d + 1$ ) ;
- une suite de  $m + 1$  réels non décroissants  $T = \{t_k\}_{k=1}^{m+1}$  appelés nœuds ( $T$  est le vecteur des nœuds).

Soit  $N_k^d$  la  $k$ -ème fonction B-spline de degré  $d$ . Les fonctions B-splines de degré 0, constantes par morceaux, sont

$$N_k^0(t) = \chi_{[t_k, t_{k+1}[} = \begin{cases} 1 & \text{si } t \in [t_k, t_{k+1}[ \\ 0 & \text{sinon} \end{cases} \quad (2.22)$$

On a donc  $m$  fonctions  $N_k^0$  telles que

$$\text{Supp}(N_k^0) = [t_k, t_{k+1}[ \quad (2.23)$$

Puis, pour  $d > 0$  et  $1 \leq k \leq m - (d + 1)$ ,  $N_k^d$  est définie par la formule de récurrence

$$N_k^d(t) = \omega_k^d(t) N_k^{d-1} + (1 - \omega_{k+1}^d(t)) N_{k+1}^{d-1}(t) \quad (2.24)$$

$$\omega_k^d(t) = \begin{cases} \frac{t-t_k}{t_{k+d}-t_k} & \text{si } t_k \neq t_{k+d} \\ 0 & \text{sinon} \end{cases} \quad (2.25)$$

On a donc  $m - (d + 1)$  fonctions telles que

$$\text{Supp}(N_k^d) = \text{Supp}(N_k^{d-1}) \cup \text{Supp}(N_{k+1}^{d-1}) = [t_k, t_{k+d+1}[ \quad (2.26)$$

Ainsi  $n$  fonctions B-splines de degré  $d$  sont nécessairement définies par un vecteur de  $m = n + d + 1$  nœuds.

Une courbe B-spline s'exprime comme une combinaison linéaire des fonctions B-splines, soit

$$P(t) = \sum_{k=0}^n N_k^d(t) P_k \quad t \in [t_0 \ t_{m+1}] \quad . \quad (2.27)$$

La régularité de la courbe au niveau d'un nœud est déterminée par la multiplicité de ce nœud : si  $t_k$  est répété  $m_k$  fois dans le vecteur de nœuds alors  $P(t)$  est de classe  $C^{d-m_k}$  en  $t = t_k$ . Pour que la courbe coïncide avec les points de contrôle aux extrémités, c-à-d aux bornes  $t_0$  et  $t_{m+1}$ , il est nécessaire de copier  $d + 1$  fois les valeurs extrêmes du vecteur de nœuds, soit

$$0 = t_1 = \dots = t_{d+1} < t_{d+2} < \dots < t_{n+1} = \dots = t_{n+d+1} = 1. \quad (2.28)$$

**Remarque** Les courbes de Bézier de degré  $n$  correspondent au cas particulier des courbes B-splines où le vecteur des nœuds est donné par  $T = \underbrace{[0 \ \dots \ 0]_{n+1}}_{n+1} \underbrace{[1 \ \dots \ 1]_{n+1}}_{n+1}$ .

En général, afin d'assurer une continuité  $C^2$ , on manipulera des splines cubiques ( $d = 3$ ) de support uniforme ( $t_{k+1} - t_k = \frac{1}{n-2d}$ ,  $d + 1 \leq k \leq n$ ). On représente de telles fonctions à la Figure 2.3 pour  $n = 10$  (donc  $m = n + d + 1 = 14$ ).

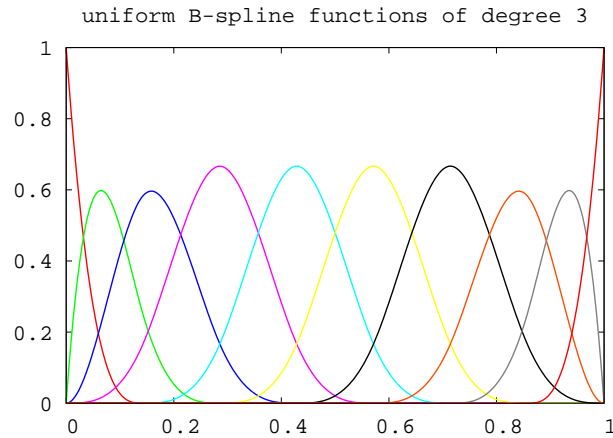


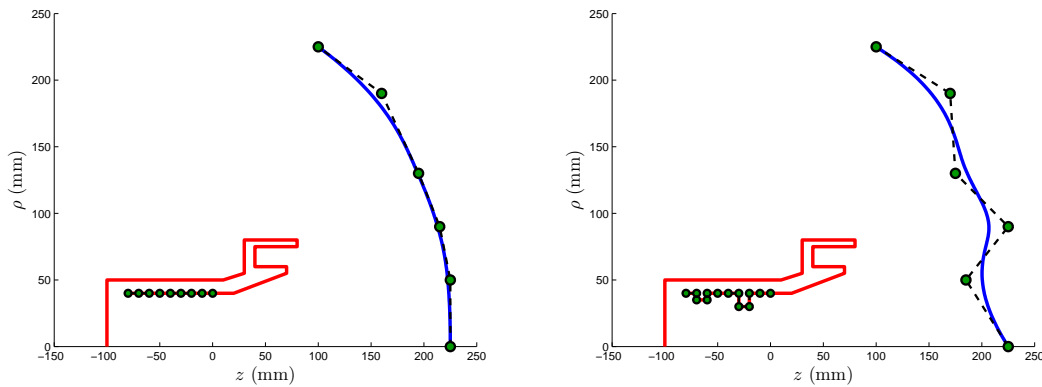
FIG. 2.3 – Fonctions B-splines de degré 3 ( $d$ 'ordre 4) définies par un vecteur de nœuds uniformément répartis sur l'intervalle  $[0 \ 1]$ .

Toutefois, il est d'usage pour l'optimisation d'un guide d'onde de rajouter des « iris » : pour un guide circulaire cela revient à déformer l'intérieur du guide par des patches constants qui sont bien représentés par des B-splines de degré 0.

Bien évidemment, si plusieurs composants d'une antenne doivent être optimisés dans un même processus, les paramétrisations de chacun d'entre eux sont indépendantes (la nature et le degré peuvent varier d'un composant à l'autre). On donne à la Figure 2.4 un exemple d'antenne composée d'un guide et d'un réflecteur dont les méridiennes sont modélisées par une courbe B-spline d'ordre 1 et une courbe B-spline d'ordre 4 respectivement.

Deux types d'enrichissement sont possibles :

- élévation du degré : comme pour les courbes de Bézier, on peut passer d'une base de polynômes de degré 3 par morceaux à une base de polynômes de degré 4 par morceaux ; le support des fonctions de la base est alors élargi.
- insertion d'un nœud : en insérant un nœud dans le vecteur  $T$  on ajoute une fonction de base ; on diminue le support des fonctions voisines et on diminue la régularité en un point si le nouveau nœud augmente la multiplicité d'un nœud déjà existant.



(a) avant déplacement des points de contrôle

(b) après déplacement des points de contrôle

FIG. 2.4 – Modélisation de la méridienne d'une antenne par courbes B-splines. Le déplacement des points de contrôle déforme les parties à optimiser selon la régularité requise :  $C^0$  pour le guide d'onde (rouge) et  $C^2$  pour le réflecteur (bleu).

### À propos d'autres paramétrisations polynomiales

Les représentations de Bézier et B-splines sont adéquates pour la description de géométries complexes. Elles ont également une structure hiérarchique intrinsèque via la propriété 2.1 qui font d'elles des paramétrisations de choix pour les méthodes multiniveaux.

Toutefois, cette propriété s'étend pour toutes les autres bases des espaces polynômiaux  $\mathcal{P}^n$ . En effet, on a  $\mathcal{P}^n \subset \mathcal{P}^m$  pour  $m > n$ , c-à-d tout polynôme de degré au plus  $n$  est également un polynôme de degré au plus  $m$ . Si en plus on munit  $\mathcal{P}^n$  et  $\mathcal{P}^m$  des bases  $\mathcal{B}^n$  et  $\mathcal{B}^m$  respectivement, alors il existe une application linéaire de  $\mathbb{R}^n$  dans  $\mathbb{R}^m$  qui donne les coefficients dans la base  $\mathcal{B}^m$  de l'espace  $\mathcal{P}^m$  du même polynôme exprimé dans la base  $\mathcal{B}^n$ . En particulier pour la base canonique ou les bases de polynômes orthogonaux, cette application est triviale.

Certaines bases de paramétrisation s'avèrent numériquement mieux adaptées mais d'un point de vue géométrique n'ont pas les bonnes propriétés de la CAO. Il est en effet difficile de modéliser une antenne à partir des polynômes de Legendre par exemple. Cette difficulté est levée avec l'approche de déformation *Free-Form*.

### 2.2.5 Approche paramétrique *Free-Form*

L'antenne de la Figure 2.4 a été modélisée « à la main » par des courbes B-splines. En pratique il est rare de concevoir une antenne à partir de rien. Le contexte impose quasi-systématiquement une configuration initiale déjà fournie. Selon la manière dont cette géométrie est décrite, il est nécessaire de réaliser une procédure préliminaire d'approximation en accord avec la représentation paramétrique. On distingue les cas suivants :

- l'antenne est décrite sous forme continue, par des portions de sections coniques, polynômes, etc. : on recherche les paramètres de contrôle qui décrivent exactement ou au mieux la géométrie (au sens d'une interpolation par exemple).
- l'antenne est décrite par son maillage : on approche la géométrie par une procédure d'interpolation.

Alternativement, l'approche de déformation *Free-Form* (FFD) propose une gestion élégante de la paramétrisation qui évite toute approximation initiale. Elle consiste à utiliser les techniques de la CAO pour la paramétrisation de la déformation au lieu de la forme elle-même. Cette technique a été introduite dans [86].

Soit  $S^0$  une géométrie initiale contenue dans un domaine  $D$ . Le déplacement  $\Delta S$  de  $S^0$  résulte de la déformation du domaine  $D$  tout entier. Formellement la courbe déformée s'exprime comme une fonction

des paramètres de conception *Free-Form x* de la manière suivante :

$$\mathcal{S}(x) = \mathcal{S}_0 + \Delta\mathcal{S}(x) \quad (2.29)$$

où  $\Delta\mathcal{S}(x)$  correspond à une paramétrisation classique.

Illustrons le principe de la *FFD* pour la déformation de la méridienne d'une antenne, c-à-d un ensemble de courbes dans un espace à 2 dimensions  $\Omega$  muni d'un système de coordonnées dit *réel*. Soit  $D$  une boîte (domaine rectangulaire) contenant  $\mathcal{S}_0$  et défini comme le domaine de déformation. On munit ce domaine d'un système de coordonnées local  $(\xi, \zeta) \in [0, 1]^2$ . Quel que soit le point du domaine  $D$ , son déplacement est défini continûment comme une combinaison linéaire de produits tensoriels de polynômes de Bernstein (la déformation prend alors la forme d'une surface de Bézier) dont les coefficients  $P_{jk}$  sont aussi appelés *points de contrôle*. Soit  $M$  un point de  $D$  de coordonnées locales  $(\xi, \zeta)$  ; la déformation de  $M$  en coordonnées réelles s'écrit

$$\Delta M = \begin{pmatrix} \Delta z \\ \Delta \rho \end{pmatrix} = \sum_{j=0}^{n_j} \sum_{k=0}^{n_k} B_j^{n_j}(\xi) B_k^{n_k}(\zeta) P_{jk} \quad (2.30)$$

Puisque la déformation est définie sur  $D$  tout entier alors elle l'est pour tout point de  $\mathcal{S}^0$ . En particulier si  $\mathcal{S}^0$  est définie par un maillage, la courbe déformée est obtenue par déformation de chacun des points du maillage. Avec cette approche, de la déformation d'un maillage résulte un maillage de la géométrie déformée.

Puisque  $\Delta\mathcal{S}$  a la forme d'une surface de Bézier, alors on a une formule d'élévation du degré pour la *FFD* (2.30) : une déformation de degré  $(n_j, n_k)$  peut être exprimée comme une déformation de degré  $(m_j, m_k)$  si  $m_j \geq n_j$  et  $m_k \geq n_k$ .

Historiquement, la paramétrisation *FFD* (2.30) est un produit tensoriel de courbes de Bézier. On peut toutefois étendre le concept de la *FFD* en remplaçant les polynômes de Bernstein par d'autres fonctions de base (par ex. à support local comme les *B-splines*).

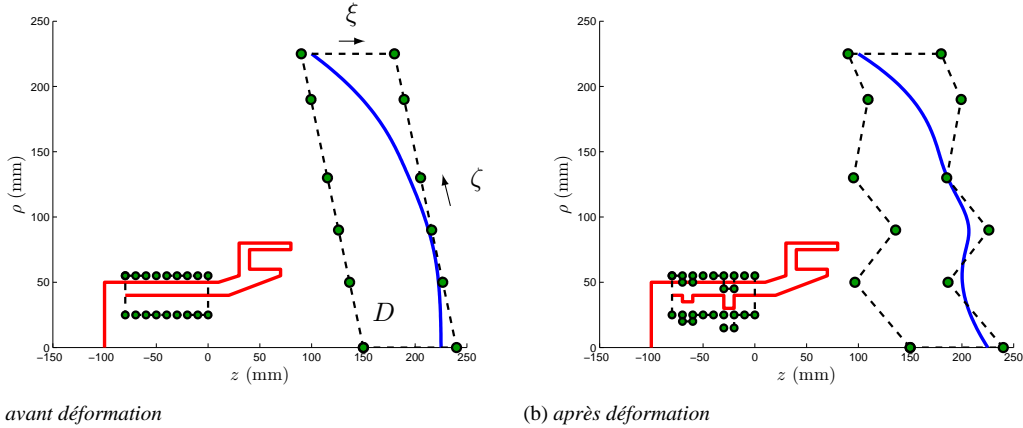


FIG. 2.5 – Application de la *FFD* pour la déformation d'une antenne. Deux boîtes sont définies autour des parties à optimiser. La déformation des boîtes par les points de contrôle déforme le maillage ; seules les parties logiquement incluses sont déformées. La boîte du réflecteur a pour degré  $(n_j, n_k) = (1, 5)$  et celle du guide  $(n_j, n_k) = (16, 1)$ .

**Remarque** Le cas présent est particulièrement simple puisque la méridienne est une simple courbe de  $\mathbb{R}^2$ . Grâce à la propriété de partition unité des polynômes de Bernstein ( $\sum_{k=0}^n B_k^n(t) = 1$ ) et puisque  $n_j = 1$  ou  $n_k = 1$ , si on requiert la condition  $P_{0k} = P_{1k}$  pour tout  $k$  alors la déformation *FFD* s'exprime comme une simple courbe de Bézier et non comme un produit tensoriel. Cela revient à réduire l'épaisseur de la boîte à 0, c-à-d à exprimer la déformation comme une fonction de l'abscisse curviligne seule.

Par ailleurs on ne déforme que selon l'axe  $z$  ou l'axe  $\rho$  selon la partie (guide ou réflecteur). Ainsi on a

$$\Delta z = \sum_{k=0}^{n_k} B_k^{n_k}(\zeta) P_k \quad \text{ou} \quad \Delta \rho = \sum_{j=0}^{n_j} B_j^{n_j}(\xi) P_j.$$

### 2.2.6 Déformation de la surface par un champ normal

Un résultat fondamental de [90] est que le gradient de forme est normal à la surface. Intuitivement on comprend que les déplacements tangentiels ne modifient pas la forme et donc l'état. On propose donc une paramétrisation de type FFD pour la déformation de la forme par un champ normal à la méridienne. La différence majeure avec cette méthode est que l'espace de recherche est redéfini à chaque itération puisque la paramétrisation dépend de la forme elle-même. Les paramètres de conception n'appartiennent plus à une structure d'espace vectoriel.

Formellement la déformation de la méridienne à l'itération  $i$  s'exprime comme une vitesse  $\vec{V}$  :

$$\mathcal{C}^{i+1}(x^{i+1}, \mathcal{C}^i) = \mathcal{C}^i + \vec{V}(\mathcal{C}^i, x^{i+1}). \quad (2.31)$$

On construit la paramétrisation de la manière suivante : soit  $t \in [0, 1]$  l'abscisse curviligne de  $\mathcal{C}$  ; on désigne par  $\vec{n}^i(t)$  le champ normal unitaire à  $\mathcal{C}^i$  (voir Figure 2.6) ; une base de fonctions de déformation normales à  $\mathcal{C}^i$  est définie à partir d'une base de fonctions  $\{f_k\}$  (Bernstein, B-splines, etc. voir la section précédente) :

$$\vec{V}_k^i(t) = f_k(t) \vec{n}^i(t) \quad k = 0 \dots n \quad (2.32)$$

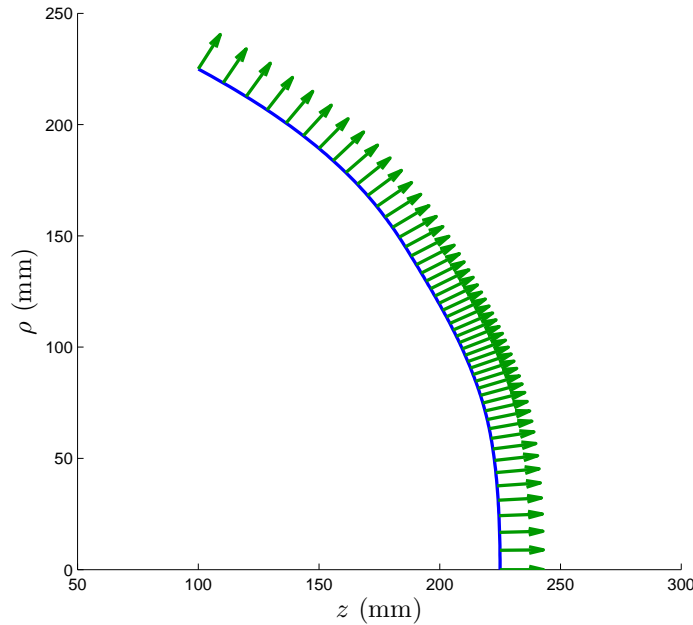


FIG. 2.6 – Champ normal unitaire à la méridienne d'un réflecteur.

On représente une telle base à la Figure 2.7 lorsque les fonction  $f_k$  sont des fonctions B-splines de degré 3. Ainsi la déformation totale  $\vec{V}$  est donnée en fonction du vecteur de conception  $x^i \in \mathbb{R}^{n+1}$  comme une combinaison de ces vitesses de base :

$$\vec{V}(\mathcal{C}^i, x^i) = \sum_{k=0}^n x_k^i \vec{V}_k^i \quad (2.33)$$

Pour exprimer  $\mathcal{C}^i$  en fonction de la forme initiale  $\mathcal{C}^0$  il faut retracer toute l'histoire de la convergence :  $\mathcal{C}^1(x^1, \mathcal{C}^0), \mathcal{C}^2(x^2, \mathcal{C}^1), \dots, \mathcal{C}^i(x^i, \mathcal{C}^{i-1})$ . Les variables de conception  $x^i$  n'appartiennent pas à une structure d'espace vectoriel. Pour cette raison, on verra que seul l'algorithme de plus grande descente a un sens dans cette paramétrisation.

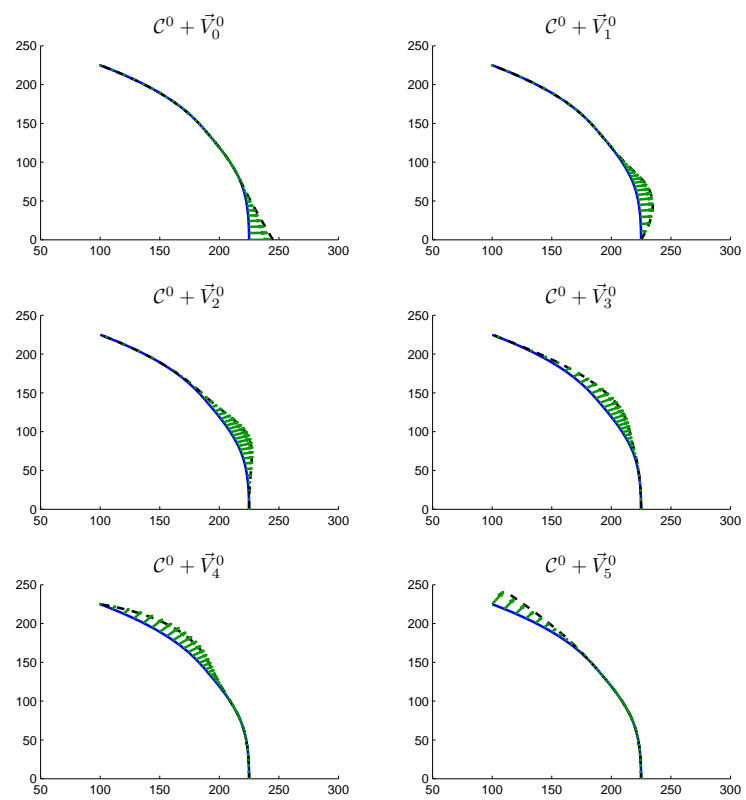


FIG. 2.7 – Base de vitesses normales au réflecteur.

Illustrons cette approche par un exemple de déformation d'un réflecteur en simulant deux phases d'optimisation à la Figure 2.8. On verra que les stratégies hiérarchiques sont possibles dans cette paramétrisation. En effet, la finesse de la paramétrisation n'est en aucune manière imposée dans le processus d'optimisation. À chaque itération le degré peut être modifié.

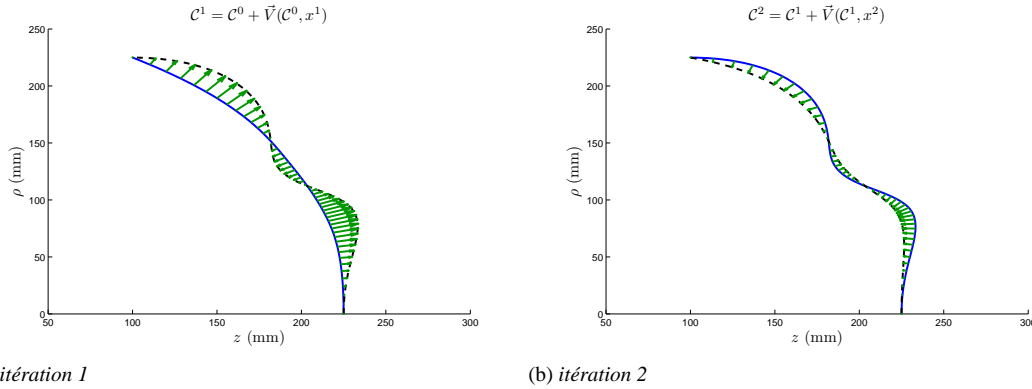


FIG. 2.8 – Simulation de deux itérations d'optimisation d'un réflecteur dans une base de déformation par un champ normal. À l'itération 1 (a), la courbe initiale  $C^0$  (bleu) est déformée par le champ normal (vert) qui est une combinaison linéaire des vitesses de base de la Figure 2.7 et donne la nouvelle méridienne  $C^1$  (pointillée). À l'itération 2 (b),  $C^1$  devient la courbe de référence (bleu) et est déformée à nouveau par un champ normal pour donner  $C^2$  (pointillée).

### 2.2.7 Contraintes sur les variables de conception

Afin de compléter la mise en équation du problème d'optimisation (2.6) on pose les contraintes sur les variables de conception. D'une manière générale peu de contraintes seront considérées. Pour la plupart des cas seules des contraintes linéaires d'égalité seront utiles.

Dans le cas de surfaces à symétrie de révolution il est (souvent) nécessaire d'assurer la continuité  $C^1$  de la surface au niveau de l'axe de symétrie (abscisse  $s = 0$ ). Cela revient à imposer que le plan tangent au réflecteur soit normale à l'axe de symétrie. Similairement, la tangente à l'extrémité du réflecteur (abscisse  $s = 1$ ) peut être imposée. Si on note  $\vec{T}(s)$  le vecteur tangent à la méridienne, ces conditions sont

$$\begin{aligned}\vec{T}(0) &= (0 \ 1)\mathcal{T} \\ \vec{T}(1) &= \vec{T}_1\end{aligned}\tag{2.34}$$

Par ailleurs, les extrémités sont parfois fixées, soit si  $M(s)$  désigne un point de la méridienne,

$$\begin{aligned}M(0) &= M_0 \\ M(1) &= M_1\end{aligned}\tag{2.35}$$

Ces contraintes s'expriment trivialement comme des contraintes linéaires d'égalité sur les paramètres de conception. On illustre ces quatre contraintes à la Figure 2.9.



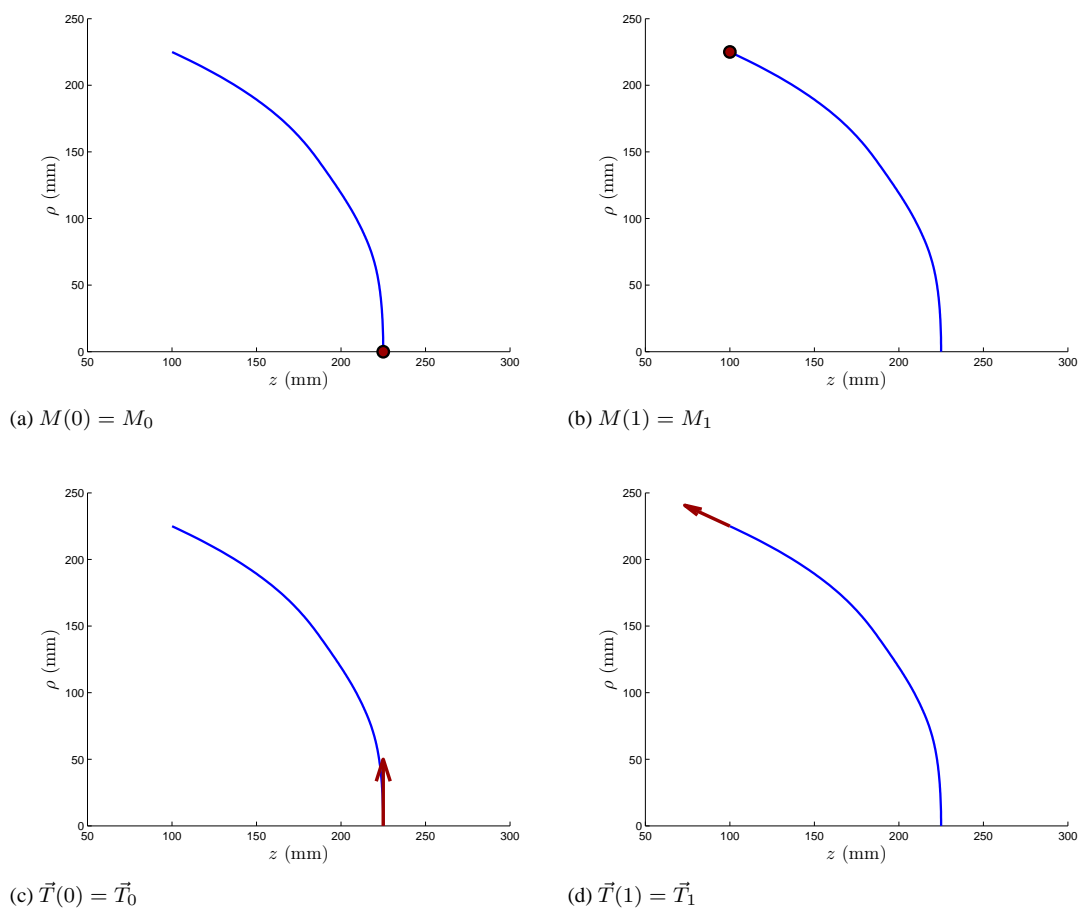


FIG. 2.9 – Contraintes d'égalité sur le réflecteur.

## 2.3 Algorithmes d'optimisation numérique

Dans cette section on présente les grandes classes de méthodes d'optimisation numérique. Les algorithmes utilisés par la suite sont détaillés.

### 2.3.1 Enjeux

L'enjeu principal d'un algorithme d'optimisation numérique est de construire une suite minimisante.

**Définition 2.4.** Une suite minimisante  $\{x^i\}$  est une suite telle que  $\{J(x^i)\}$  satisfait

$$J(x^1) > J(x^2) > \dots > J(x^i) > \dots \quad (2.36)$$

Dans l'idéal on souhaite que cette suite converge vers le minimum global de la fonction objectif. En pratique, à moins d'avoir des résultats préalables d'unicité, on ne sait pas caractériser un tel minimum. Si  $J$  est une fonction différentiable dont on sait évaluer le gradient, on souhaite obtenir au moins les conditions nécessaires d'optimalité. On recherche donc un point stationnaire dont on pourra vérifier éventuellement s'il satisfait les conditions suffisantes.

Les principales difficultés de l'optimisation numérique sont

- la multimodalité : il existe plusieurs minima locaux qui sont autant de points d'attraction vers lesquels les algorithmes risquent de converger ; un algorithme est dit *robuste* s'il est plus capable de déterminer le minimum global parmi les minima locaux.
- la raideur : la suite minimisante converge lentement à cause de la raideur du problème posé ; une convergence (numérique) acceptable peut être hors de portée si un trop grand nombre d'itérations est nécessaire.

On distingue deux grandes classes d'algorithmes : les algorithmes basés sur le gradient (parfois appelés algorithmes de descente) et les algorithmes sans gradient. Les premiers sont efficaces pour converger vers un minimum local mais sont peu robustes ; les derniers sont plus robustes mais sont plus coûteux en temps de calcul.

### 2.3.2 Algorithmes de descente

Soit  $J : \mathbb{R}^n \rightarrow \mathbb{R}$  une fonction différentiable. Dans un souci de clarté on considère ici un problème de minimisation sans contraintes.

**Définition 2.5.** Une direction de descente pour la fonction  $J$  au point  $x$  est un vecteur  $d \in \mathbb{R}^n$  qui satisfait

$$\nabla J(x) \cdot d < 0. \quad (2.37)$$

Un développement de Taylor d'ordre 1 montre que si  $\nabla J$  n'est pas nul alors il existe une telle direction. La plus évidente est  $d = -\nabla J(x)$ . Étant donnée une direction de descente il existe un réel positif  $\tau$  tel que  $J(x + \tau d) < J(x)$ . Par conséquent, si  $x^0$  désigne un vecteur initial, on est capable de construire une suite minimisante  $x^i$  de manière récursive en calculant une direction  $d^i$  et un pas  $\tau^i$  en chaque point de la suite

$$x^{i+1} = x^i + \tau^i d^i \quad (2.38)$$

jusqu'à ce que les conditions de stationnarité soient satisfaites à une tolérance numérique donnée. Il reste à déterminer le paramètre  $\tau$  qui est appelé *pas de descente* le long de  $d$ . La méthode à pas fixe est la plus simple mais converge sous des hypothèses particulières sur  $J$  (convexité). Dans l'idéal, un algorithme à pas optimal résout le sous-problème

$$\tau^* = \arg \min_{\tau} J(x + \tau d) \quad (2.39)$$

à chaque itération. En pratique il existe plusieurs méthodes dites de *recherche linéaire* ou encore de *globalisation* pour approcher le pas optimal (dichotomie, interpolation quadratique, etc.). Lors de cette procédure il est d'usage d'assurer que le pas  $\tau^i$  vérifie les conditions de Wolfe (2.40) et (2.41) :

$$J(x^i + \tau^i d^i) \leq J(x^i) + \omega_1 \tau^i \nabla J(x^i) \cdot d^i \quad (2.40)$$

$$\nabla J(x^i + \tau^i d^i) \cdot d^i \geq \omega_2 \nabla J(x^i) \cdot d^i \quad (2.41)$$

où  $0 < \omega_1 < \omega_2 < 1$ . Ces conditions assurent que

1. la fonction coût diminue suffisamment ; on évite ainsi que l'algorithme converge vers un point non stationnaire à cause d'une suite de pas trop petits – condition (2.40) appelée également condition d'Armijo.
2. la pente en  $x^i + \tau^i d^i$  (le long de  $d^i$ ) est plus grande que celle en  $x^i$  (soit négative et de plus petite amplitude, soit positive) – condition (2.41).

Des valeurs typiques des paramètres sont  $\omega_1 = 0.001$  et  $\omega_2 = 0.9$ . La condition (2.41) peut être remplacée par la condition plus restrictive

$$|\nabla J(x^i + \tau^i d^i) \cdot d^i| \leq |\omega_2 \nabla J(x^i) \cdot d^i| \quad (2.42)$$

qui assure une réduction de la pente le long de  $d^i$ . L'ensemble (2.40) et (2.42) constitue les conditions fortes de Wolfe.

Si les conditions (2.41) et (2.42) sont utiles pour les preuves de convergence, elles sont en revanche coûteuses à évaluer car elles nécessitent l'évaluation du gradient  $\nabla J(x^i + \tau^i d^i)$  pour chaque valeur de  $\tau$  testée dans le processus de globalisation. On ne vérifiera que la condition d'Armijo.

Un prototype d'algorithme de descente à pas optimal est donné dans l'Algorithme 1.

---

**Algorithme 1** : Algorithme de descente à pas optimal

---

**Entrées** :  $x^0, \varepsilon$

descente : procédure de calcul d'une direction de descente (2.37)

// Initialisation

$J^0 \leftarrow J(x^0)$

$d^0 \leftarrow -\nabla J(x^0)$

$i \leftarrow 0$

// Itérations de descente

**tant que**  $\|\nabla J(x^i)\| > \varepsilon$  **faire**

$\tau^i \leftarrow \arg \min_{\tau} J(x^i + \tau d^i)$

$x^{i+1} \leftarrow x^i + \tau^i d^i$

$J^{i+1} \leftarrow J(x^{i+1})$

$d^{i+1} \leftarrow \text{descente}(\nabla J(x^{i+1}), \nabla J(x^i), d^i, x^i, x^{i+1})$

$i \leftarrow i + 1$

**fin**

**Sorties** :  $i, x^i, J^i, \nabla J(x^i)$

---

Il reste à déterminer la direction de descente. La méthode la plus simple et la plus intuitive est la méthode de *plus grande descente* :

$$d^i = -\nabla J(x^i) \quad (2.43)$$

appelée ainsi puisqu'elle satisfait

$$-\frac{\nabla J(x)}{\|\nabla J(x)\|} = \arg \min_{\|d\|=1} \nabla J(x) \cdot d \quad (2.44)$$

Cette méthode est rarement utilisée car elle converge très lentement (voir la section 2.4.1). On présente deux autres classes de méthodes de descente : le Gradient Conjugué non-linéaire et les méthodes de Newton et quasi-Newton.

### Gradient Conjugué non-linéaire

L'algorithme du Gradient Conjugué non-linéaire est une extension directe de la méthode du Gradient Conjugué pour la résolution de systèmes linéaires symétriques définis positifs. Toutefois, dans le cas non-linéaire, la notion de directions conjuguées n'a plus de sens.

La direction de descente est donnée par

$$\begin{cases} d^0 = -\nabla J(x^0) \\ d^i = -\nabla J(x^i) + \beta d^{i-1} \quad i > 0 \end{cases} \quad (2.45)$$

Il existe deux versions du Gradient Conjugué :  $\beta$  est soit le coefficient de Fletcher-Reeves

$$\beta_{FR} = \frac{\|\nabla J(x^i)\|^2}{\|\nabla J(x^{i-1})\|^2} \quad (2.46)$$

ou le coefficient de Polak-Ribière

$$\beta_{PR} = \frac{\nabla J(x^i) \cdot (\nabla J(x^i) - \nabla J(x^{i-1}))}{\|\nabla J(x^{i-1})\|^2}. \quad (2.47)$$

### Newton et quasi-Newton

On suppose que  $J$  est de classe  $C^2$ . Une direction de descente d'une méthode de Newton ou de quasi-Newton a la forme

$$d^i = -(B^i)^{-1} \nabla J(x^i) \quad (2.48)$$

où  $B^i$  est le hessien exact (Newton) ou approché (quasi-Newton) au point  $x^i$  agissant comme préconditionneur du gradient. Si  $B^i$  est définie positive alors  $d^i$  est effectivement une direction de descente.

La stratégie quasi-Newton la plus utilisée est la méthode Broyden-Fletcher-Goldfarb-Shanno (BFGS). Cette stratégie consiste à raffiner l'approximation du hessien par différences finies à partir des itérations précédentes : soit  $y^i = \nabla J(x^i) - \nabla J(x^{i-1})$  et  $s^i = x^i - x^{i-1}$ . La mise à jour BFGS du hessien s'écrit

$$B^{i+1} = B^i + \frac{y^i y^{iT}}{y^{iT} s^i} - \frac{B^i s^i s^{iT} B^i}{s^{iT} B^i s^i}. \quad (2.49)$$

Les termes de correction constituent une mise à jour de rang 2 et sont symétriques. Il est donc évident que  $B^i$  est symétrique pour tout  $i$  si  $B^0$  l'est. La matrice  $B^{i+1}$  est garantie définie positive à chaque itération pourvu que la matrice initiale  $B^0$  le soit et que  $x^i$  vérifie les conditions de Wolfe. En pratique on pose  $B^0 = I$ . Si la technique de globalisation ne prend pas en compte les conditions de Wolfe, il est possible que  $B^i$  ne soit plus définie positive. Dans ce cas il est alors courant de réinitialiser la matrice  $B^i$  à l'identité.

Il existe une variante des méthodes de quasi-Newton où l'inverse du hessien est approché plutôt que le hessien lui-même. Cela nous permet résoudre le problème (2.48) directement. Cependant, lorsque le problème est soumis à des contraintes supplémentaires, il est plus adéquat de résoudre un problème de « programmation quadratique ». Si l'algorithme de quasi-Newton est plus coûteux que celui de plus grande descente ou du Gradient Conjugué étant donné que le système linéaire (2.48) doit être résolu à chaque itération, dans le contexte d'un problème d'optimisation soumis à la résolution d'une EDP, ce coût devient négligeable devant l'évaluation de la fonction objectif.

### Évaluation du gradient

Le gradient peut être évalué de plusieurs manières :

- par des méthodes adjointes : la formulation d’un gradient de forme pour un problème de diffraction électromagnétique a fait l’objet des thèses [68, 34] basée sur [90] ; en vue d’une optimisation numérique, le gradient est obtenu par discrétisation du gradient de forme continu. On n’utilisera pas cette méthode puisque certaines géométries considérées ne satisfont les hypothèses de régularité.
- par différentiation automatique (DA) : le code de calcul de la fonction coût est regardé comme une composition de fonctions élémentaires ; la technique de DA analyse le code pour générer un code de calcul du gradient numérique exacte. Le code SRSR ne se prête pas à la différentiation automatique ; en revanche, pour motiver les approches multiniveaux proposées, nous seront amenés à étudier le hessien de la fonction coût lorsque l’état est obtenu par le modèle d’OP. Comme nous connaissons le gradient analytiquement, le hessien est obtenu par DA avec le logiciel TAPE-NADE [50]<sup>2</sup>.
- par différences finies : le schéma centré

$$\frac{\partial}{\partial x_k} J(x) = \frac{J(x - h\mathbf{e}_k) - 2J(x) + J(x + h\mathbf{e}_k)}{h^2} + O(h^2) \quad (2.50)$$

est utilisé pour le calcul du gradient avec le modèle de Maxwell. Il est certes coûteux en temps de calcul ( $2n$  évaluations par itération) mais permet toutefois d’étudier la convergence des stratégies de descente. Pour des raisons de stabilité on calcule le gradient de l’état  $u$  (champ électrique) par différences finies plutôt que celui de  $J$  directement, soit  $\nabla_{FD} u(x)$ . On en déduit alors  $\nabla J(x) = F'(u(x))\nabla_{FD} u(x)$  où  $F$  est une fonctionnelle de l’état et  $F'$  sa différentielle par rapport à  $u$ .

### 2.3.3 Algorithmes sans gradient

#### Simplexe

Les méthodes dites du simplexe sont des algorithmes déterministes sans gradient. Elles peuvent être appliquées pour la minimisation de fonctions non différentiables. La version originelle du simplexe pour la minimisation d’une fonction numérique dans  $\mathbb{R}^n$  a été introduite par Spendley *et al.* [91] puis étendu par Nelder et Mead [72]. Elle consiste à explorer l’espace de recherche en faisant évoluer les sommets d’un simplexe de degré  $n + 1$ . Chaque sommet du complexe est un vecteur de paramètres de conception qui correspond à une forme candidate pour laquelle on évalue la fonction objectif. En fonction des valeurs obtenues pour un simplexe donné, des règles de transformation sont appliquées afin de parcourir l’espace de recherche de telle sorte que l’on se dirige vers un meilleur candidat (réflexion par rapport à un sommet, expansion ou contraction du simplexe).

Une version introduite par Torczon (*Multidirectional Search Algorithm*, MSA) [96] est particulièrement appréciée pour ses résultats de convergence, contrairement aux algorithmes de base. Elle est originellement dédiée aux architectures parallèles.

Des algorithmes hiérarchiques basés sur la méthode du simplexe ont été développés et appliqués avec succès dans le cadre de la conception optimale de forme en aérodynamique [1].

#### Algorithmes Évolutionnaires

Les Algorithmes Évolutionnaires (EAs) sont des algorithmes sans gradient à population dont les Algorithmes Génétiques (GAs) sont les plus connus [46, 7]. Ils sont réputés pour leur robustesse mais sont coûteux en terme de temps de calcul puisque le coût d’une itération est dominé par le temps de calcul des évaluations de la fonction objectif par tous les individus de la population.

<sup>2</sup><http://www-sop.inria.fr/tropics/>

### Algorithme par Essaim de Particules

Enfin, on présente un algorithme introduit par Kennedy et Eberhart [55] appelé *Optimisation par Essaim de Particules* (OEP) ou *Particle Swarm Optimization* (PSO)<sup>3</sup>. On verra par la suite que cette approche est particulièrement bien adaptée aux stratégies hiérarchiques.

L'algorithme *PSO* s'inspire de la dynamique du mouvement d'un essaim dans l'espace ambiant. Plus précisément, il a été développé en premier lieu pour l'étude et la simulation d'essaims d'oiseaux. Dans un second temps, par analogie avec la recherche commune d'un lieu propice (par exemple riche en nourriture), il a été utilisé pour la minimisation d'une fonction dans  $\mathbb{R}^n$ . Chaque particule de l'essaim est un vecteur de paramètres de conception. L'espace de recherche est exploré par l'essaim de manière « intelligente » en suivant une dynamique (position et vitesse) guidées par l'expérience de chaque particule ainsi que par des échanges d'information entre les individus. Cette dynamique est régie en partie par des règles déterministes et en partie par des termes stochastiques. On parle d'algorithme de population semi-stochastique.

Soit un essaim de  $N$  particules, c-à-d  $N$  vecteurs de  $\mathbb{R}^n$ . À l'itération  $i$ , chaque particule  $j$  est caractérisée par une position  $x_j^i$  et une vitesse  $v_j^i$ . L'évolution de l'essaim est régie par la simple relation de dynamique

$$x_j^{i+1} = x_j^i + \Delta t v_j^i \quad (2.51)$$

où  $\Delta t$  est un « pas de temps » que l'on prend égal à 1 en général (le temps est une notion fictive dans ce contexte, une unité de temps correspond à une itération d'optimisation). La fonction objectif est évaluée en chacune des positions, enrichissant ainsi l'expérience de l'essaim. En particulier on garde en mémoire les valeurs suivantes :

- la meilleure position trouvée par l'essaim (mémoire globale) :  $x^* = \arg \min_{i,j} J(x_j^i)$  ;
- la meilleure position trouvée par la particule  $j$  (mémoire locale) :  $x_j^* = \arg \min_i J(x_j^i)$  ;

Cette expérience intervient dans l'expression de la nouvelle vitesse affectée à chaque particule. Plus précisément, la vitesse est une pondération de confiance locale et globale :

$$v_j^{i+1} = c_1 r_1 (x_j^* - x_j^i) + c_2 r_2 (x^* - x_j^i) \quad (2.52)$$

où  $c_1$  et  $c_2$  sont deux coefficients fixés, et  $r_1$  et  $r_2$  sont deux nombres aléatoires suivant une loi uniforme sur  $[0 \ 1]$ , ce afin d'accroître la diversité de l'essaim. Chaque terme peut être vu comme une force de rappel vers les meilleurs points à la connaissance de  $j$ . Une version légèrement modifiée de la vitesse intègre un terme d'inertie qui en pratique permet de contrôler également la diversité de l'essaim au cours des itérations :

$$v_j^{i+1} = w^i v_j^i + c_1 r_1 (x_j^* - x_j^i) + c_2 r_2 (x^* - x_j^i) \quad (2.53)$$

où  $w^i$  est le coefficient d'inertie à l'itération  $i$ . Si  $w^i > 1$  alors l'essaim a tendance à s'étendre, si  $w^i < 1$  il a tendance à se regrouper. On contrôle la valeur de ce coefficient de sorte que  $w^i > 1$  au cours des premières itérations et  $w^i < 1$  en fin de convergence à l'aide de la réduction géométrique

$$w^{i+1} = w_c w^i \quad (2.54)$$

pour  $w^0 > 1$  et  $0 < w_c < 1$  donnés. On résume l'ensemble de la procédure dans l'Algorithme 2.

Il existe quelques résultats de convergence et de stabilité dans des cas particuliers [25]. Ici la réduction du coefficient d'inertie impose une convergence de l'essaim en un point qui n'est pas nécessairement un point stationnaire. En pratique un nombre maximal d'itérations est fixé. Alternativement, un critère de convergence peut être défini à partir de données statistiques de l'essaim comme la variance de la fonction coût ou des positions (utilisée comme une mesure de la diversité de l'essaim).

La version historique de l'algorithme PSO décrit dans [55] a donné naissance à de nombreuses extensions. Parmi les contributions proposées on trouve en particulier

- une borne supérieure  $v_{\max}$  sur la norme des vitesses.

<sup>3</sup>Le site <http://www.swarmintelligence.org/> référence les projets, articles, tutoriels, etc. en relation avec les algorithmes PSO et leurs applications.

- une condition de décroissance de l'inertie plutôt qu'une décroissance géométrique : on décide par exemple d'appliquer la réduction (2.54) si la meilleure position n'a pas changée pendant  $n_c$  itérations.
- un critère de « folie » : il existe une (faible) probabilité  $p_c$  pour qu'une particule ne suive pas la règle (2.53) mais prenne une direction purement aléatoire.
- une mémoire intermédiaire entre la mémoire locale et la mémoire globale. La version de l'algorithme telle qu'elle a été utilisée dans ce travail comporte une modification de la définition de la vitesse : il s'agit de l'algorithme appelé *Unified PSO* (voir [77, 78]). Dans un premier temps on définit des sous-groupes de particules au moyen d'une notion de voisinage différente de celle donnée par une métrique de l'espace de recherche, afin d'éviter l'agglomération de particules. Typiquement, une topologie circulaire basée sur l'indice des particules est utilisée : on note  $g_j$  le sous ensemble de particules défini par les  $n_u$  premiers éléments de  $\{j, j+1, \dots, N-1, N, 1, \dots, j-1\}$ . Soit  $x_{g_j}^*$  la meilleure position trouvée par le groupe  $g_j$ , la vitesse est définie par

$$\begin{aligned} \mathcal{G}_j^{i+1} &= c_1 r_1 (x_j^* - x_j^i) + c_2 r_2 (x^* - x_j^i) \\ \mathcal{L}_j^{i+1} &= c_1 r_1' (x_j^* - x_j^i) + c_2 r_2' (x_{g_j}^* - x_j^i) \\ v_j^{i+1} &= w^i v_j^i + u \mathcal{G}_j^{i+1} + (1-u) \mathcal{L}_j^{i+1} \end{aligned} \quad (2.55)$$

---

**Algorithme 2** : Optimisation par Essaim de Particules

---

**Données** :  $N, c_1, c_2, w^0, w_c, \text{imax}, \rho$

**Entrées** :  $x^0$

rand : routine de calcul d'un nombre pseudo-aléatoire uniforme dans  $[0, 1]$

// Initialisation de la mémoire globale

$J^* \leftarrow J(x^0)$

$x^* \leftarrow x^0$

// Initialisation aléatoire de l'essaim (positions et vitesses)

Distribution uniforme des positions dans  $D = B_\infty(x^0, \rho)$

$x_j^0 \sim \mathcal{U}(D) \quad j = 1 \dots N$

Distribution uniforme des vitesses dans  $D' = B_\infty(\mathbf{0}, v_{\max})$

$v_j^0 \sim \mathcal{U}(D') \quad j = 1 \dots N$

// Évolution de l'essaim

**pour**  $i \leftarrow 0$  à imax **faire**

$x^* \leftarrow \arg \min_{i,j} J(x_j^i)$

$x_j^* \leftarrow \arg \min_j J(x_j^i)$

$r_1 \leftarrow \text{rand}()$

$r_2 \leftarrow \text{rand}()$

$v_j^{i+1} \leftarrow \underbrace{w^i v_j^i}_{\text{inertie}} + \underbrace{c_1 r_1 (x_j^* - x_j^i)}_{\text{mémoire locale}} + \underbrace{c_2 r_2 (x^* - x_j^i)}_{\text{mémoire globale}}$

$x_j^{i+1} \leftarrow x_j^i + v_j^{i+1}$

$w^{i+1} \leftarrow w^i w_c$

**fin**

**Sorties** :  $x^*, J(x^*)$

---

Notons que la dynamique de l'algorithme PSO est soumise à certains paramètres qu'il faut régler. On trouve dans la littérature des valeurs communément utilisées. Par la suite, sauf explicitement modifiées, on utilisera les valeurs données dans le Tableau 2.1.

Paramètre	Valeur typique
$c_1$	2
$c_2$	2
$w^0$	1.2
$w_c$	0.8
$n_c$	3
$v_{\max}$	$\frac{1}{4} \text{diam}(D)$
$p_c$	0.05
$u$	0.4
$n_u$	$\text{round}(\frac{1}{5}N)$

TAB. 2.1 – Valeurs standards des paramètres de l’algorithme PSO.

### 2.3.4 Hybridation

À partir des méthodes sans gradient et des méthodes de descente on peut construire des algorithmes hybrides afin de bénéficier à la fois des performances en recherche globale des unes et des performances en recherche locale des autres. On adopte ici la plus simple des stratégies qui consiste à réaliser successivement une optimisation globale puis un raffinement par une méthode locale.

## 2.4 Diagnostic de la convergence

Une manière de caractériser la performance d’un algorithme d’optimisation numérique est d’étudier la convergence de la suite minimisante qu’il produit.

### 2.4.1 Taux et vitesse de convergence

**Définition 2.6.** On appelle  $e^i = x^i - x^*$  l’erreur itérative d’une suite  $\{x^i\}$  convergeant vers  $x^*$ .

**Définition 2.7.** Soit  $\{x^i\}$  une suite de  $\mathcal{F}$  convergeant vers  $x^*$ . S’il existe  $\mu \in ]0, 1[$  tel que

$$\lim_{i \rightarrow \infty} \frac{\|x^{i+1} - x^*\|}{\|x^i - x^*\|} = \lim_{i \rightarrow \infty} \frac{\|e^{i+1}\|}{\|e^i\|} = \mu \quad (2.56)$$

alors  $\mu$  est le taux de convergence linéaire : l’erreur est réduite asymptotiquement d’un facteur  $\mu$  à chaque itération. Si

$$\lim_{i \rightarrow \infty} \frac{\|x^{i+1} - x^*\|}{\|x^i - x^*\|} = 0 \quad (2.57)$$

alors on dit que le taux de convergence est super-linéaire. En particulier si on a (2.57) et si il existe  $\mu > 0$  tel que

$$\lim_{i \rightarrow \infty} \frac{\|x^{i+1} - x^*\|}{\|x^i - x^*\|^\alpha} = \mu \quad (2.58)$$

pour  $\alpha > 1$ , alors la convergence est dite super-linéaire d’ordre  $\alpha$  (quadratique si  $\alpha = 2$ ). En revanche si

$$\lim_{i \rightarrow \infty} \frac{\|x^{i+1} - x^*\|}{\|x^i - x^*\|} = 1 \quad (2.59)$$

alors le taux de convergence est sous-linéaire.

En optimisation on représente plutôt la convergence de la suite  $J^i$  vers  $J^* = J(x^*)$ . La représentation de la suite  $J^i$  en fonction des itérations comme sur la Figure 2.10 nous donne une idée de la performance itérative d’un algorithme. Cependant, cela ne nous donne pas d’indication sur le coût itératif. Il est donc



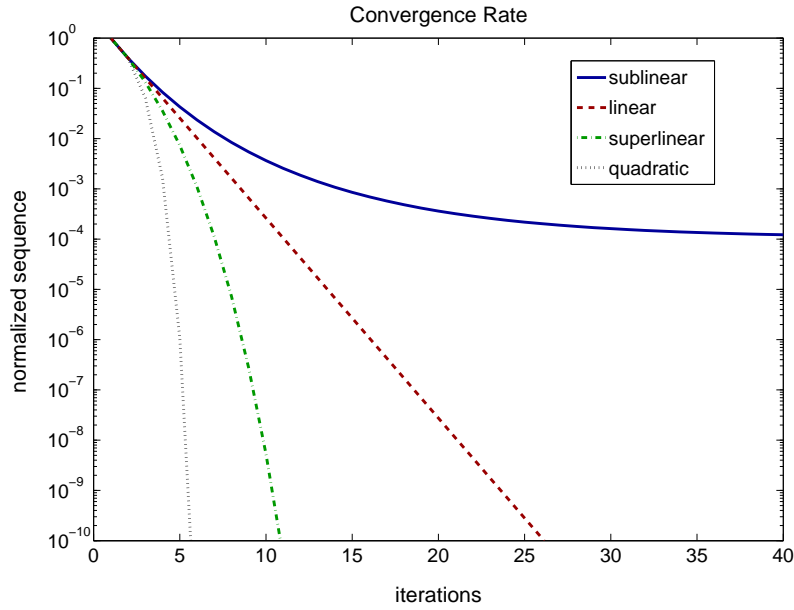


FIG. 2.10 – Illustration du taux de convergence de suites.

parfois plus judicieux de représenter la valeur de la fonctionnelle en fonction du temps CPU. Dans le contexte de l'optimisation de systèmes gouvernés par des EDP, c'est le calcul d'analyse qui domine le coût d'une itération. Cela revient à représenter l'évolution de la suite minimisante en fonction du nombre d'évaluations de fonctionnelle. Pour un algorithme de gradient cela revient à dire que l'on prend en compte les évaluations effectuées au cours de la recherche de pas optimal.

On montre dans [54] que le taux de convergence locale de la méthode de plus grande descente est linéaire, celui des Gradient Conjugué non-linéaire et de la méthode BFGS est superlinéaire, et celui de la méthode de Newton est quadratique.

## 2.4.2 Critère d'arrêt

Un algorithme de gradient est un algorithme de recherche de point stationnaire. Un critère de convergence s'exprime donc naturellement en fonction de la norme euclidienne du gradient : l'algorithme s'arrête dès que cette norme est inférieure à une tolérance  $\varepsilon$  donnée (voir l'Algorithme 1).

Cependant, dans le cadre d'un problème d'optimisation de forme, la précision du calcul du gradient peut varier selon la paramétrisation adoptée. Par ailleurs, si un schéma de différences finies est appliqué, il est difficile d'établir une tolérance en accord avec la précision du calcul. On préfère utiliser un critère qui s'exprime en fonction du résidu de la fonction coût : l'algorithme s'arrête dès que la fonctionnelle « ne descend plus », c-à-d tant que le résidu absolu

$$r^i = J^i - J^{i-1} \quad (2.60)$$

est plus petit qu'une tolérance fixée. On peut vérifier la norme du gradient *a posteriori*. Si la fonction coût ne tend pas vers 0 on mesure de préférence le résidu relatif

$$r^i = \frac{J^i - J^{i-1}}{J^{i-1}} \quad (2.61)$$

et la tolérance est choisie de l'ordre de la précision machine. Le critère (2.62) est un compromis entre les deux notions de résidu relatif et absolu en fonction de la valeur de la fonction coût [45].

$$r^i = \frac{J^i - J^{i-1}}{1 + J^{i-1}} \quad (2.62)$$

## 2.5 Validation des algorithmes et illustration de cas pathologiques

Cette section est dédiée à valider les codes d'optimisation numérique et à illustrer leur convergence sur des problèmes standards. On cherche à minimiser une fonction  $F$  dans  $\mathbb{R}^2$ . On note  $x = (x_1 \ x_2)^T$ .

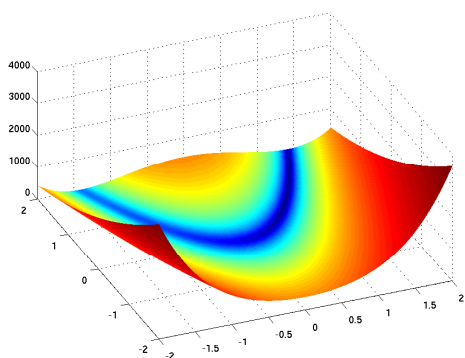
### 2.5.1 Algorithmes de descente et illustration de la raideur

La fonction de *Rosenbrock* et la fonction de *Powell mal conditionnée* conduisent à des problèmes raides.

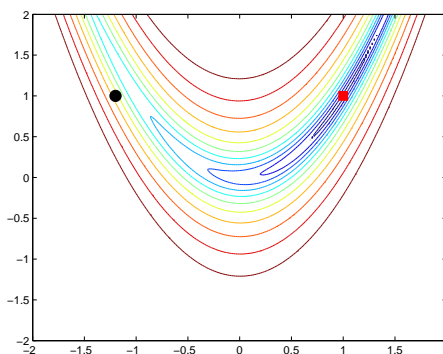
#### Fonction de Rosenbrock

$$F(x) = (1 - x_1)^2 + 100(x_2 - x_1^2)^2 \quad (2.63)$$

- domaine de définition :  $[-2 \ 2]^2$
- point initial :  $x^0 = (-1.2 \ 1)$
- solution :  $x^* = (1 \ 1)$
- valeur de la fonction coût à la solution :  $J^* = 0$
- nombre max d'itérations : 200



(a) surface  $z = F(x_1, x_2)$



(b) courbes de niveau, point initial ●, solution ■

FIG. 2.11 – Fonction de Rosenbrock.

On présente les résultats à la Figure 2.12 pour les méthodes de plus grande descente (1e ligne), gradient conjugué (2e ligne) et quasi-Newton BFGS (3e ligne), en terme de convergence de la fonctionnelle (1e colonne), convergence de la norme du gradient (2e colonne) et chemin parcouru par la suite minimisante (3e colonne).

Au bout de 200 itérations l'approximation obtenue par la méthode SD paraît proche de la solution sans l'atteindre : la convergence est linéaire avec un taux proche de 1. En outre, la norme du gradient peine à diminuer et oscille, comme les positions successives des  $x^i$ . Ces observations sont symptomatiques d'un problème raide.

La convergence de la méthode CG est bien meilleure : elle est clairement superlinéaire. La solution est atteinte en 60 itérations. La norme du gradient converge plus régulièrement et le chemin parcouru par les  $x^i$  est beaucoup moins oscillant.

Enfin la convergence de la méthode de QN est très satisfaisante : on observe bien le comportement d'une suite qui converge superlinéairement. La solution est atteinte en 40 itérations et la norme du gradient converge de manière encore plus régulière.

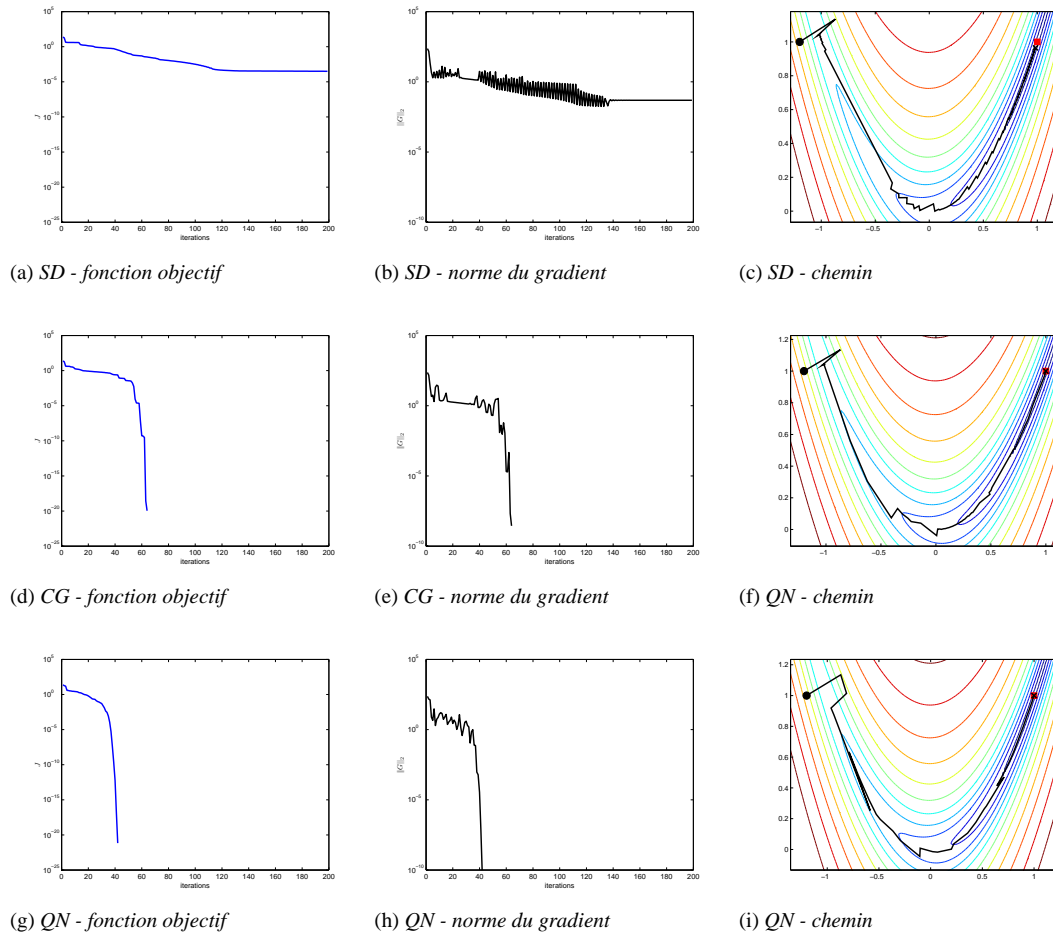
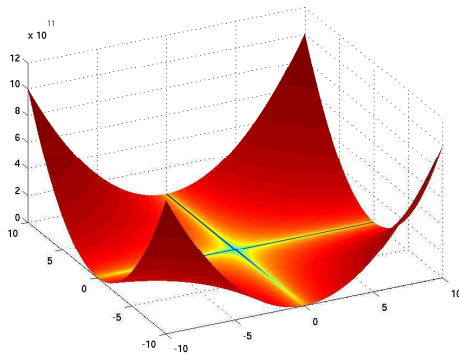
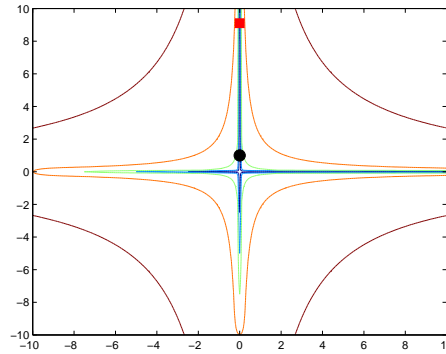


FIG. 2.12 – Convergence des méthodes de descente : plus grande descente (*SD*), gradient conjugué de Polak-Ribière (*CG*), quasi-Newton *BFGS* (*QN*).

### Fonction de Powell mal conditionnée

$$F(x) = (10000x_1x_2 - 1)^2 + (e^{-x_1} + e^{-x_2} - 1.0001)^2 \quad (2.64)$$

- domaine de définition :  $[-10 \ 10]^2$
- point initial :  $x^0 = (0 \ 1)$
- solution :  $x^* = (1.098159 \cdot 10^{-5} \ 9.106146)$
- valeur de la fonction coût à la solution :  $J^* = 0$
- nombre max d'itérations : 200

(a) surface  $z = F(x_1, x_2)$ 

(b) courbes de niveau, point initial ●, solution ■

FIG. 2.13 – « Badly Scaled Powell Function ».

On présente les résultats à la Figure 2.14 pour les méthodes de plus grande descente (1e ligne), gradient conjugué (2e ligne) et quasi-Newton BFGS (3e ligne), en terme de convergence de la fonctionnelle (1e colonne), convergence de la norme du gradient (2e colonne) et chemin parcouru par la suite minimisante (3e colonne).

Le même comportement que pour la minimisation de la fonction de « Rosenbrock » est observé : la raideur s'observe par la convergence (très) lente de l'algorithme SD : au bout de 200 itérations l'approximation est toujours loin de la solution. La convergence est linéaire avec un taux proche de 1. En outre la norme du gradient peine à diminuer et oscille, comme les positions successives des  $x^i$ .

Ce problème s'avère encore plus raide que le précédent : le taux de convergence linéaire de la méthode CG est plus petit et la solution est quasiment atteinte mais la convergence de la norme du gradient est toujours très oscillante. L'algorithme s'est arrêté au bout de 120 itérations (à satisfaction du critère du résidu absolu).

Enfin la convergence de la méthode de QN est encore la meilleure : on observe bien le comportement d'une suite qui converge superlinéairement, la norme du gradient converge de manière plus régulière vers la précision machine.

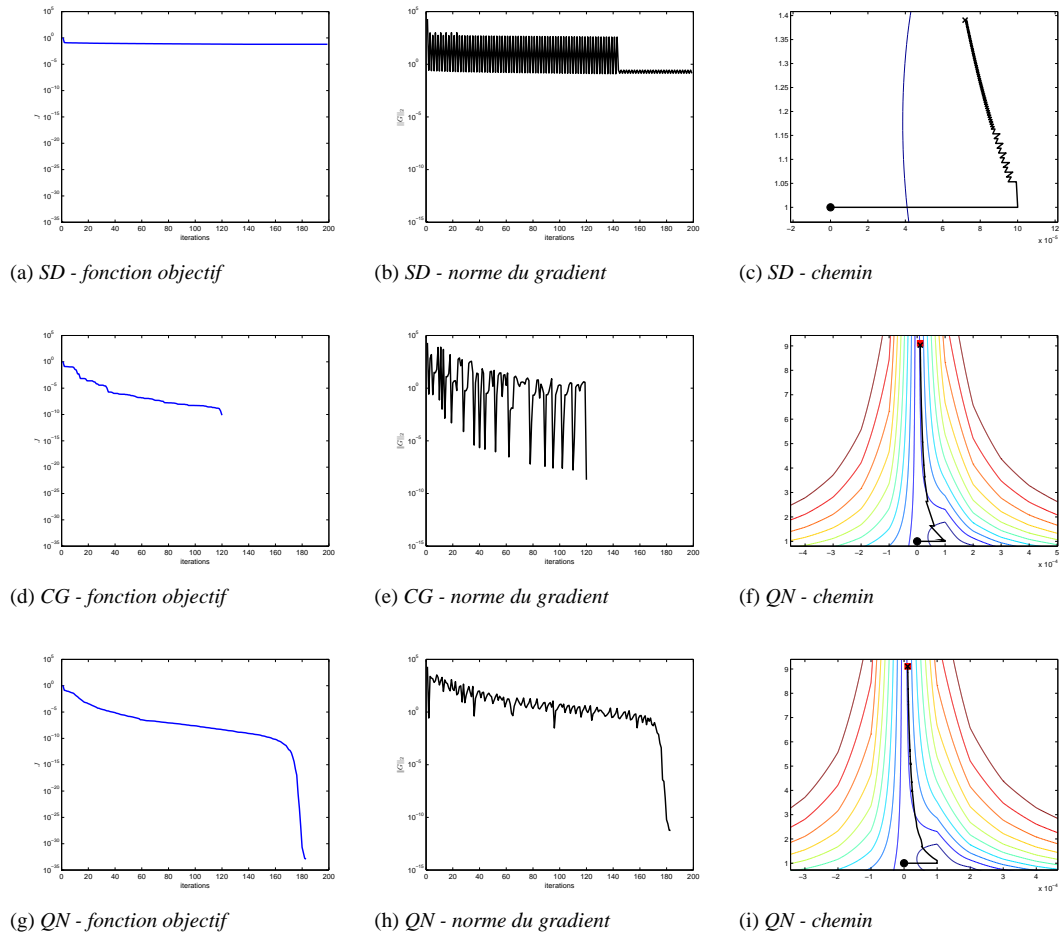


FIG. 2.14 – Convergence des méthodes de descente : plus grande descente (*SD*), gradient conjugué de Polak-Ribière (*CG*), quasi-Newton *BFGS* (*QN*).

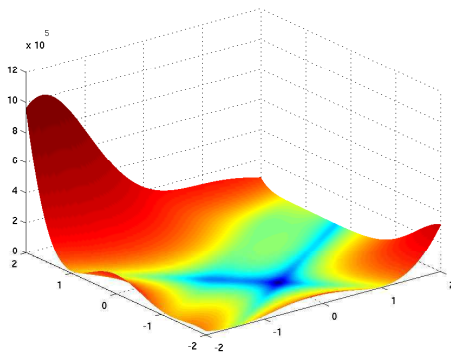
## 2.5.2 Algorithme PSO et illustration de la robustesse

Les fonctions de *Goldstein & Price* et de *Schwefel* sont des fonctions multimodales (celle de *Schwefel* est fortement multimodale).

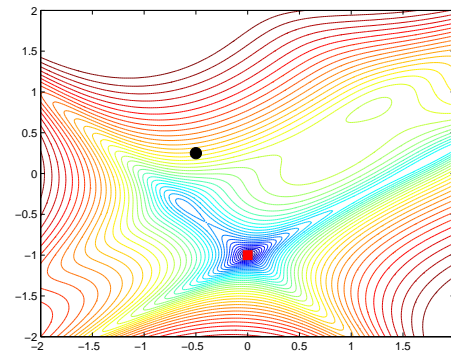
### Fonction de Goldstein & Price

$$\begin{aligned}
 a &= (x_1 + x_2 + 1) \\
 b &= (19 - 14(x_1 + x_2) + 6x_1x_2 + 3(x_1^2 + x_2^2)) \\
 c &= (2x_1 - 3x_2) \\
 d &= (18 - 32x_1 + 48x_2 - 36x_1x_2 + 12x_1^2 + 27x_2^2) \\
 F(x) &= (1 + a^2b) (30 + c^2d)
 \end{aligned} \tag{2.65}$$

- domaine de définition :  $[-2 \ 2]^2$
- point initial :  $x^0 = (-0.5 \ 0.25)$
- solution :  $x^* = (0 \ -1)$
- valeur de la fonction coût à la solution :  $J^* = 3$
- nombre de particules : 10
- nombre d'itérations : 100



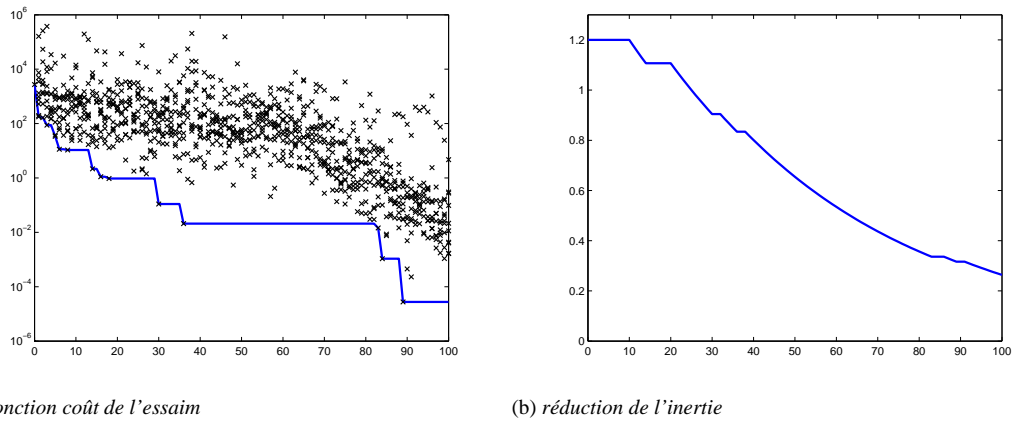
(a) surface  $z = F(x_1, x_2)$



(b) courbes de niveau, point initial ●, solution ■

FIG. 2.15 – Fonction de Goldstein & Price.

On illustre les résultats à la Figure 2.16. On représente la convergence de  $J^i - J^*$  pour tout l'essaim. La convergence par sauts de la meilleure position est typique de l'algorithme PSO. Chaque saut est un changement probable de bassin d'attraction. Contrairement aux algorithmes de descente, l'algorithme PSO ne produit pas de suite strictement minimisante. On remarque que lorsque la valeur de la fonctionnelle stagne, l'inertie est réduite géométriquement. Cette fonctionnelle est peu multimodale : l'essaim converge rapidement vers la solution (90 itérations). Cet exemple montre l'efficacité de l'algorithme sur un problème simple.



(a) fonction coût de l'essaim

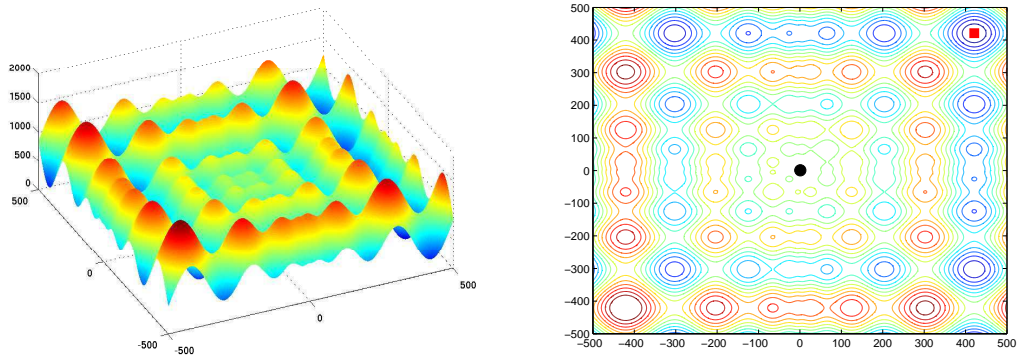
(b) réduction de l'inertie

FIG. 2.16 – Convergence de la méthode PSO : évolution des valeurs de la fonction objectif de l'essaim et réduction conditionnelle de l'inertie.

### Fonction de Schwefel

$$F(x) = 418,9829n - \sum_{i=1}^n \left( x_i \sin \sqrt{|x_i|} \right) \quad (2.66)$$

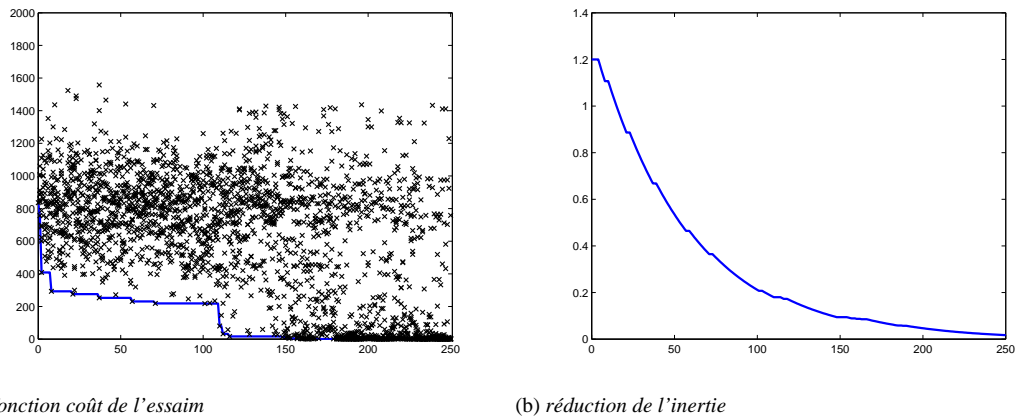
- domaine de définition :  $[-500 \ 500]^2$
- point initial :  $x^0 = (0 \ 0)$
- solution :  $x^* = (420.9687 \ 420.9687)$
- valeur de la fonction coût à la solution :  $J^* = 0$
- nombre de particules : 10
- nombre d'itérations : 250

(a) surface  $z = F(x_1, x_2)$ 

(b) courbes de niveau, point initial ●, solution ■

FIG. 2.17 – Fonction de Schwefel.

On illustre les résultats à la Figure 2.18. Ce problème est beaucoup plus difficile car il possède de nombreux minima locaux. Lorsque les valeurs de la fonction coût de l'essaim sont réparties, l'espace est exploré, lorsqu'il tend vers la meilleure valeur c'est que tout l'essaim a tendance à converger vers la meilleure particule dont la localisation a peu changé. On observe qu'il faut plus de 150 itérations pour que l'essaim converge. Le saut vers la 110e itération est typique d'un changement de bassin d'attraction.



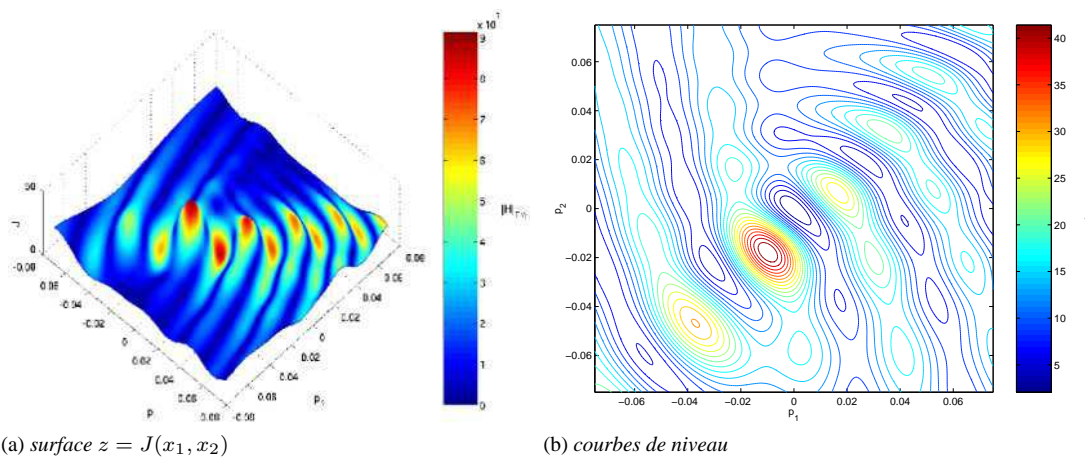
(a) fonction coût de l'essaim (b) réduction de l'inertie  
 FIG. 2.18 – Convergence de la méthode PSO : évolution des valeurs de la fonction objectif de l'essaim et réduction conditionnelle de l'inertie.

### 2.5.3 Problème inverse sur la puissance rayonnée

Les problèmes inverses sont des exercices pratiques de validation des méthodes pour le traitement de problèmes concrets. Par exemple, on montre ici un cas de conformation d'une antenne par rapport à un diagramme en puissance

$$\mathcal{J}(S) = \frac{1}{2} \int_{S^2} \left( \|u(S, f)\|^2 - \|u(S^t, f)\|^2 \right)^2 dS^2 \quad (2.67)$$

où  $S^t$  est la forme cible donnée par la géométrie d'une antenne connue. Pour illustrer le fait que ces problèmes sont à la fois soumis à des problèmes de raideur et de multimodalité on représente la fonction coût en fonction de deux paramètres de déformation à la Figure 2.19.



(a) surface  $z = J(x_1, x_2)$  (b) courbes de niveau  
 FIG. 2.19 – Représentation de la fonctionnelle paramétrique de forme autour de la solution du problème inverse en fonction des points de contrôle FFD  $P_1$  et  $P_2$ . Le degré de la paramétrisation est  $n = 2$ , on a fixé  $P_0 = 0$ . La forme initiale est la forme cible  $S^0 = S^t$  de telle sorte que la solution est obtenue pour  $P = (0 \ 0 \ 0)^T$ .



## Conclusion

Dans ce chapitre on a exposé les généralités concernant l'optimisation numérique de forme. Dans un premier temps on a rappelé les résultats fondamentaux de caractérisation d'un minimum local dans le cadre d'un problème d'optimisation en dimension finie. Ayant adopté une approche paramétrique pour la représentation de la géométrie, le problème posé se réduit effectivement à un problème en dimension finie. On a alors introduit les paramètres d'optimisation dans le contexte de l'optimisation de la forme d'une antenne à réflecteur. En particulier on a introduit une paramétrisation *Free-Form* ainsi qu'une approche de déformation des surfaces régulières par un champ normal.

Dans un second temps on a présenté les méthodes d'optimisation numérique. Les codes développés ont été validés sur des fonctions modèles. Dans le cas des méthodes de descente, on a pu observer que les taux de convergence numériques sont en accord avec les taux de convergence théoriques.

Dans ce contexte on est en mesure de résoudre numériquement des problèmes de conception optimale de la forme d'une antenne à réflecteur. On s'attache dans les trois chapitres suivants à développer et expérimenter des approches multiniveaux dans le but de proposer d'une part des méthodes robustes et d'autre part des méthodes dont la convergence est plus rapide qu'avec les méthodes classiques.

# ALGORITHME BIGRILLE IDÉAL POUR L'OPTIMISATION DE FORME

# 3

---

<b>3.1</b>	<b>État de l'art et enjeux</b>	<b>59</b>
<b>3.2</b>	<b>Définition d'un problème modèle géométrique</b>	<b>59</b>
3.2.1	Problème de meilleure approximation	60
3.2.2	Fonctionnelle paramétrique	60
3.2.3	Application à la paramétrisation P1	61
<b>3.3</b>	<b>Méthodes itératives de base</b>	<b>66</b>
3.3.1	Rappels sur la convergence d'une itération linéaire	66
3.3.2	Itération de Jacobi	67
3.3.3	Accélération de Tchebychev, méthode de Richardson	67
3.3.4	Méthode de plus grande descente	67
3.3.5	Préconditionnement par une méthode de type quasi-Newton	68
3.3.6	Fonction d'amortissement	68
<b>3.4</b>	<b>Algorithmes multigrilles</b>	<b>71</b>
3.4.1	Opérateurs de prolongement et de restriction	72
3.4.2	Enrichissement progressif : <i>Nested Iteration</i>	73
3.4.3	Algorithme bigrille idéal	74
3.4.4	MG, FMG, FAS	75
3.4.5	Application des algorithmes classiques pour l'optimisation de forme	76
3.4.6	Redéfinition des opérateurs de transfert	78
3.4.7	Révision de l'algorithme bigrille idéal	80
3.4.8	Expériences numériques	82
<b>3.5</b>	<b>Influence de la paramétrisation sur la raideur</b>	<b>89</b>
3.5.1	Paramétrisation de Bézier	89
3.5.2	Paramétrisation B-splines	92
3.5.3	Polynômes orthogonaux	94
3.5.4	Influence de la dimension sur la raideur	94
3.5.5	Expériences numériques	95
	<b>Conclusion</b>	<b>101</b>

---

Ce chapitre a deux objectifs : rappeler les caractéristiques essentielles des algorithmes multigrilles ; en parallèle, à l'aide d'un problème modèle, montrer comment ces algorithmes peuvent être étendus à des problèmes d'optimisation de forme. Dans le chapitre suivant on justifiera en quoi le problème modèle est représentatif des problèmes raides pour la conception d'antenne.



### 3.1 État de l'art et enjeux

Les méthodes multigrilles (MG) pour la résolution d'EDP ont été largement étudiées par le passé. Introduites par Fedorenko [40, 41] pour la résolution de problèmes linéaires comme l'équation de Poisson, elles sont utilisées aujourd'hui pour la résolution de grands systèmes non-linéaires avec des approches de type *Full Approximation Scheme (FAS)* ou Newton [14, 49, 101, 29, 16]. À l'origine basées sur une hiérarchie de représentations géométriques, elles ont été étendues à une approche purement algébrique (AMG) [81]. Plus récemment, des méthodes ont été proposées pour élargir les multigrilles au contexte de l'optimisation de forme appliquée à des systèmes gouvernés par des EDP. Parmi celles-ci on trouve les méthodes dites *one-shot*, l'algorithme *MG/Opt*, ainsi que des méthodes d'optimisation paramétrique [63]. Dans un cadre plus général d'optimisation multiniveau, on note également les travaux [53]. On trouvera une bibliographie complète dans [12].

La stratégie *one-shot* consiste à rassembler dans un même système les équations d'état et les conditions de stationnarité de la fonction coût. Le tout étant non-linéaire, une stratégie MG de type *FAS* ou Newton est appliquée pour résoudre le système enrichi [61, 62, 5, 84, 33, 98]. Lorsque l'équation d'état est linéaire et résolue par une méthode directe, comme c'est ici le cas, cette stratégie ne s'avère plus pertinente.

L'algorithme *MG/Opt* [70, 56, 57] est un algorithme d'optimisation de forme développé par Nash *et al* inspiré par l'algorithme *FAS*. Il est basé sur la hiérarchie des espaces d'approximation de la variable d'état. Ces espaces d'approximation reposent généralement sur un maillage (comme par exemple les fonctions de base d'une méthode d'éléments finis). Puisque la fonction coût dépend de l'état, à chaque finesse de maillage  $h$  correspond une approximation de la fonction coût  $J_h$ . En d'autres termes, une hiérarchie d'espaces d'approximation de l'état implique une hiérarchie d'approximation de la fonction coût, et ce quel que soit le type de contrôle adopté (approche *CAD-free* ou approche paramétrique indépendante du maillage). Afin d'assurer la consistance de la fonction à minimiser entre deux niveaux d'approximation, une stratégie de type *FAS* est adoptée : une direction de descente sur un niveau grossier est alors une direction de descente sur le niveau fin.

Dans ce chapitre on s'intéresse aux stratégies qui portent sur la hiérarchie des variables de conception dans un contexte d'optimisation paramétrique. Ici la finesse du maillage pour le calcul de l'état reste fixée au cours des itérations. Ainsi, un espace grossier paramétrique ne change pas le calcul de la fonction coût. On verra dans le chapitre suivant que la raideur numérique de certains problèmes concrets peut être due à la dimension de l'espace de recherche. Par conséquent des méthodes itératives telles que les méthodes de descente deviennent moins efficaces. Pour illustrer ce phénomène et construire des algorithmes efficaces on considère ici un problème modèle géométrique. Par analogie aux méthodes MG dites *idéales* on propose différentes stratégies à deux niveaux dont on étudie le rayon spectral afin de montrer si le cycle améliore la convergence ou non (si le taux de convergence du cycle est indépendant de la dimension de l'espace de recherche ou non). Chacune des stratégies se résume à la définition d'un opérateur de transfert qui généralise l'élévation du degré.

### 3.2 Définition d'un problème modèle géométrique

On considère le problème géométrique de meilleure approximation « en moyenne » d'une fonction dans un sous-espace de dimension finie. On verra que ce problème s'apparente étroitement au problème de Poisson  $-\Delta u = f$  qui est le problème modèle typique des méthodes MG linéaires.

Dans les sections qui suivent on note  $H_0$  l'espace de Hilbert des fonctions  $L_2$  sur l'intervalle  $[0, 1]$  telles que  $f(0) = f(1) = 0$  muni du produit scalaire usuel dans  $L_2$

$$(f, g) = \int_0^1 f(t)g(t)dt \quad (3.1)$$

et de sa norme associée

$$\|f\| = \sqrt{(f, f)} = \left( \int_0^1 |f(t)|^2 dt \right)^{1/2}. \quad (3.2)$$

### 3.2.1 Problème de meilleure approximation

Soit  $\tilde{u}$  une fonction de  $H_0$  et  $F$  un sous-espace de  $H_0$ . La meilleure approximation de  $\tilde{u}$  dans  $F$  est l'élément  $u \in F$  qui minimise la norme  $L_2$  de la différence entre  $u$  et la cible  $\tilde{u}$ , soit

$$\min_{u \in F} \mathcal{J}(u) = \frac{1}{2} \|u - \tilde{u}\|^2 = \frac{1}{2} \int_0^1 |u(t) - \tilde{u}(t)|^2 dt. \quad (3.3)$$

$\mathcal{J}$  est évidemment continue, différentiable et quadratique (donc convexe). Soit  $\mathcal{G}$  sa différentielle

$$\langle \mathcal{G}(u), \delta u \rangle = (u - \tilde{u}, \delta u) = \int_0^1 (u(t) - \tilde{u}(t)) \delta u(t) dt, \quad \forall \delta u \in F \quad (3.4)$$

et  $\mathcal{H}$  le hessien

$$\langle \mathcal{H}(u) \delta u, \delta v \rangle = (\delta u, \delta v) = \int_0^1 \delta u(t) \delta v(t) dt, \quad \forall \delta u, \delta v \in F \quad (3.5)$$

Puisque  $\mathcal{J}$  est quadratique  $\mathcal{H}$  ne dépend pas de  $u$  et est défini positif : pour tout  $u$  on a  $\langle \mathcal{H}(u) \delta u, \delta u \rangle = \|\delta u\|^2 > 0$ .

### 3.2.2 Fonctionnelle paramétrique

Comme détaillée au chapitre 2, l'approche paramétrique consiste à définir  $u$  dans un espace de dimension finie. Soit  $\{u_k\}_{k=1}^n$  une famille libre de  $H_0$ . L'espace  $F = \text{vect}\{\dots, u_k, \dots\}$  est un sous-espace de  $H_0$  de dimension  $n$  (ou de degré  $n - 1$  pour les espaces polynômiaux).

On note  $u[x]$  la fonction de  $F$  dont  $x \in \mathbb{R}^n$  est le vecteur des coefficients dans la base  $\{u_k\}_{k=1}^n$ , c-à-d

$$u[x](t) = \sum_{k=1}^n x_k u_k(t). \quad (3.6)$$

Si on injecte cette représentation dans (3.3) on obtient

$$\forall u \in F, \quad \mathcal{J}(u) = \frac{1}{2} \int_0^1 \left| \sum_{k=1}^n x_k u_k(t) - \tilde{u}(t) \right|^2 dt. \quad (3.7)$$

On en déduit la fonction coût paramétrique définie sur  $\mathbb{R}^n$  :

$$J(x) = \mathcal{J}(u[x]) \quad (3.8)$$

qui est simplement l'expression de la fonctionnelle (3.3) définie sur le sous-espace  $F$  en terme des variables de conception  $x$ . Étant donné un sous-espace  $F$ , il existe une infinité de bases ; un même problème peut donc conduire à des formulations différentes de la fonctionnelle paramétrique. On verra que le choix de cette base peut être critique du point de vue des propriétés numériques qui en découlent. Plus précisément, la convergence d'une méthode de descente peut devenir extrêmement lente à cause d'un mauvais conditionnement de la matrice hessienne.

Le gradient  $G$  de  $J$  est

$$G(x) = \begin{pmatrix} \vdots \\ \langle \mathcal{G}(u[x]), u_k \rangle \\ \vdots \end{pmatrix}, \quad (3.9)$$

et le hessien  $H$

$$H(x) = \begin{pmatrix} \cdots & \vdots & \langle \mathcal{H}(u[x])u_k, u_j \rangle \end{pmatrix}. \quad (3.10)$$

On peut reformuler plus simplement la fonctionnelle paramétrique sous la forme classique

$$J(x) = \frac{1}{2}x^T Hx - b^T x + c \quad (3.11)$$

où  $b_k = (\tilde{u}, u_k)$  et  $c = \frac{1}{2} \|\tilde{u}\|^2$ . Notons que  $H$  est toujours définie positive pour les mêmes raisons que  $\mathcal{H}$  est définie positive. Dans ce cas, la condition nécessaire d'optimalité  $G(x) = 0 \Leftrightarrow Hx - b = 0$  est une condition suffisante pour caractériser le minimum global. Résoudre le problème (3.3) est donc équivalent à la résolution du système linéaire  $Hx = b$ . On étudie désormais les propriétés numériques du hessien  $H$  dont les éléments  $h_{kj}$  sont

$$h_{kj} = \int_0^1 u_k(t)u_j(t)dt \quad (3.12)$$

dans une base  $\{u_k\}_{k=1}^n$  particulière. Par la suite, on réfère ce problème en tant que problème modèle d'optimisation de forme.

### 3.2.3 Application à la paramétrisation P1

On considère un espace d'approximation linéaire par morceaux (éléments P1) comme espace d'approximation de  $\tilde{u}$ . Cette paramétrisation dépend de la discrétisation du support  $[0, 1]$  de la fonction  $f$ . Soit  $\mathcal{T}_h$  une discrétisation uniforme de l'intervalle  $[0, 1]$  :  $t_k = kh$ ,  $h = \frac{1}{n+1}$ ,  $k = 0, \dots, n+1$ . Les fonctions P1 (voir Figure 3.1) sont définies par

$$u_k(t) = \begin{cases} \frac{t-t_{k-1}}{h} & t \in [t_{k-1}, t_k] & k > 0 \\ \frac{t_{k+1}-t}{h} & t \in [t_k, t_{k+1}] & k < n+1 \\ 0 & t \notin [t_{k-1}, t_{k+1}] \end{cases}. \quad (3.13)$$

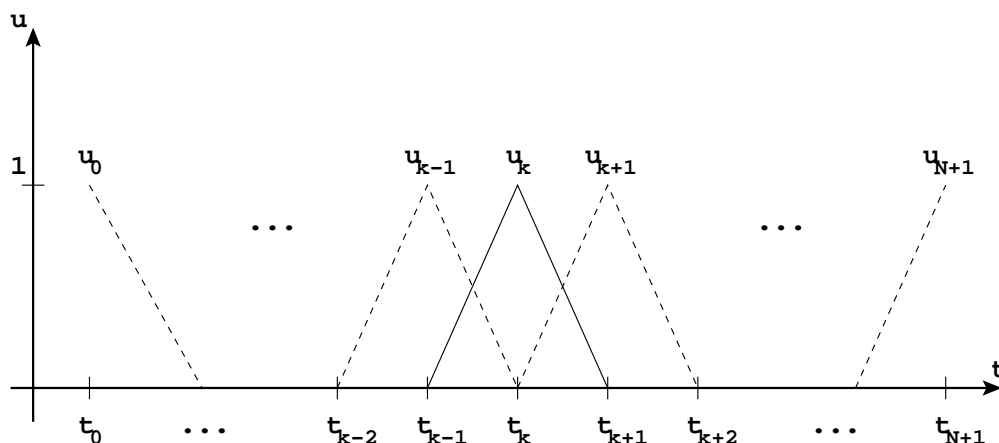


FIG. 3.1 – Éléments P1 sur l'intervalle  $[0, 1]$ .

Le support local des fonctions  $u_k$  implique que la matrice hessienne, notée  $H_h$ , a une structure bande. Avec une approximation linéaire la largeur de bande est de 3 :

$$h_{jk} = \begin{cases} \int_0^h \left(\frac{t}{h}\right)^2 dt & = \frac{h}{3} & j = k = 0, j = k = n+1 \\ 2 \int_0^h \left(\frac{t}{h}\right)^2 dt & = \frac{2h}{3} & 0 < j = k < n+1 \\ \int_0^h \frac{t}{h} \left(1 - \frac{t}{h}\right) dt & = \frac{h}{6} & j = k+1, j = k-1 \end{cases}. \quad (3.14)$$

On a donc  $H_h = \frac{h}{6}B$  avec

$$B = \begin{pmatrix} 4 & 1 & & & \\ 1 & 4 & & & \\ & & \ddots & \ddots & \ddots \\ & & & 1 & 4 & 1 \\ & & & & 1 & 4 \end{pmatrix} \in \mathbb{R}^{n \times n} \quad (3.15)$$

où on a appliqué les conditions aux limites de Dirichlet (c-à-d, on a ignoré les fonctions  $u_0$  et  $u_{n+1}$ ).

$B$  est une matrice réelle symétrique. Elle admet donc une diagonalisation orthogonale  $B = \Omega \Lambda \Omega^T$  avec des valeurs propres réelles. En outre  $B$  est strictement à diagonale dominante. D'après le théorème de Gershgorin le spectre de  $B$  est tel que  $\sigma(B) \subset [2 \ 6]$ . Par conséquent le nombre de conditionnement de  $B$  est borné par  $\kappa_2 \leq 3$ , quelle que soit la dimension de la discrétisation.

On étudie désormais le spectre de  $B$ . Pour cela on rappelle l'équation bien connue de Poisson en 1D.

### Relation avec l'équation de Poisson

L'équation de Poisson 1D sur l'intervalle fermé  $[0 \ 1]$  avec conditions homogènes de Dirichlet est

$$\begin{cases} -u''(t) = f(t) & t \in ]0 \ 1[ \\ u(0) = u(1) = 0 \end{cases} \quad (3.16)$$

On considère le schéma centré de différences finies (3.17) pour approcher la dérivée seconde.

$$-u''(t) = \frac{-u(t-h) + 2u(t) - u(t+h)}{h^2} + O(h^2). \quad (3.17)$$

En appliquant cette approximation aux nœuds  $t_k$  du maillage uniforme  $\mathcal{T}_h$  on obtient le système linéaire  $A_h u_h = f_h$  où

$$A_h = \frac{1}{h^2} A = \frac{1}{h^2} \begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & & \ddots & \ddots & \ddots \\ & & & -1 & 2 & -1 \\ & & & & -1 & 2 \end{pmatrix} \in \mathbb{R}^{n \times n}, \quad (3.18)$$

$(u_h)_k = u(t_k)$  et  $(f_h)_k = f(t_k)$ .

Comme pour  $B$ ,  $A$  est réelle et symétrique et admet donc une décomposition orthogonale  $A = S \Pi S^T$  avec des valeurs propres réelles. De plus, toujours d'après le théorème de Gershgorin, le spectre de  $A$  est tel que  $\sigma(A) \in [0 \ 4]$ .

Les deux problèmes modèles (3.3) et (3.16) conduisent aux systèmes linéaires  $H_h x = b_h$  et  $A_h u_h = f_h$  respectivement. On a vu que  $H_h = \frac{h}{6}B$  et  $A_h = \frac{1}{h^2}A$  où  $A$  et  $B$  ne dépendent pas de la discrétisation. Il est trivial de vérifier que les deux problèmes sont liés par la relation

$$B = 6I - A. \quad (3.19)$$

Avant d'aborder le traitement numérique de la résolution des systèmes, discutons de la structure spectrale de  $A$  et  $B$ .

### Analyse spectrale

D'après (3.19), la décomposition spectrale de  $B$  est déduite de celle de  $A$  :

$$\Omega = S \quad (3.20)$$

$$\Lambda = 6I - \Pi \quad (3.21)$$

Les deux matrices ont la même base orthogonale de vecteurs propres. Supposons sans perte de généralité que les valeurs propres de  $A$ , notées  $\mu_k$ , sont ordonnées de manière croissante

$$0 < \mu_1 < \dots < \mu_n < 4, \quad (3.22)$$

alors les valeurs propres de  $B$ , notées  $\lambda_k$ , sont

$$\lambda_k = 6 - \mu_k \quad \Rightarrow \quad 6 > \lambda_1 > \dots > \lambda_n > 2, \quad (3.23)$$

c-à-d, les valeurs propres  $\lambda_k$  sont ordonnées en sens inverse des  $\mu_k$  et sont translatées.

La diagonalisation de  $A$  est bien connue (voir [29, 28] par exemple) : si  $S_k$  est le vecteur propre associé à  $\mu_k$  on a

$$S_k = \sqrt{2h} \begin{pmatrix} \vdots \\ \sin(jk\pi h) \\ \vdots \end{pmatrix}, \quad \mu_k = 2 - 2 \cos(k\pi h). \quad (3.24)$$

**Remarque** On remarque que les spectres de  $A$  et  $B$  sont bien dans les intervalles donnés par le théorème de Gershgorin. Par ailleurs, les inégalités (3.22) et (3.23) étant strictes, on a  $\kappa_2(B) < 3$ .

**Remarque**  $S$  est à la fois symétrique et orthogonal, donc  $S = S^T = S^{-1}$ .

Les vecteurs propres sont les modes de Fourier discrets sur  $\mathcal{T}_h$  (voir Figure 3.2). Le coefficient  $\sqrt{2h}$  est un coefficient de normalisation. À chaque vecteur propre on associe le paramètre de fréquence  $\theta_k = k\pi h$ . On appelle les modes de *haute fréquence* (High Frequency, HF) les vecteurs propres  $S_k$  tels que  $\theta_k \geq \frac{\pi}{2}$  et les modes de *basse fréquence* (Low Frequency, LF) ceux qui vérifient  $\theta_k < \frac{\pi}{2}$ .

En utilisant la terminologie introduite ci-dessus, les deux problèmes se distinguent par leur structure spectrale de la manière suivante : les modes LF sont associés aux valeurs propres les plus petites dans le cas du problème de Poisson ; ils sont associés aux valeurs propres les plus grandes dans le cas du problème de meilleure approximation. Symétriquement, les modes HF sont associés aux valeurs propres les plus grandes dans le cas du problème de Poisson et aux valeurs propres les plus petites dans le cas du problème de meilleure approximation (voir Figure 3.3).

Par conséquent, les opérateurs linéaires  $A_h$  et  $H_h$  ont des propriétés de lissage opposées : le produit d'un vecteur par l'une de ces matrices amplifie les modes correspondant aux plus grandes valeurs propres, c-à-d les modes LF ou HF selon le problème. À la manière de [57], une méthode simple pour illustrer numériquement qu'un opérateur linéaire  $M$  est un lisseur (ou un « anti-lisseur ») est d'observer la transformée de Fourier discrète (DFT) des vecteurs de Krylov  $q^i = M^i x$  pour  $x$  quelconque (non nul dans la direction de chaque vecteur propre). Un exemple est donné à la Figure 3.4 avec les matrices  $A$  et  $B$  : en accord avec leur structure spectrale  $A_h$  (resp.  $H_h$ ) amplifie les composantes HF (resp. LF) de  $x$ .

La connaissance analytique du spectre de  $B$  nous sera utile par la suite pour l'étude de la convergence de schémas bigrilles idéaux. En anticipant sur la suite, on réalise déjà qu'une stratégie multigrille classique pour la résolution du problème modèle d'optimisation de forme ne fonctionnera pas. En effet nous manquons d'un opérateur de lissage. Pour préciser ce que l'on entend par *lisseur* on revient sur les méthodes itératives de base en faisant un parallèle entre le point de vue de l'optimisation (voir section 2.3) et le point de vue de la résolution de systèmes linéaires.

**Remarque** Les méthodes MG sont efficaces pour la résolution de systèmes raides ; or le problème de meilleure approximation dans la paramétrisation P1 est très bien conditionné ( $\kappa_2 < 3$ ). À la section 3.5 on généralise ce problème à d'autres paramétrisations dont on étudie numériquement la structure spectrale. On verra alors que le conditionnement explose avec la dimension.



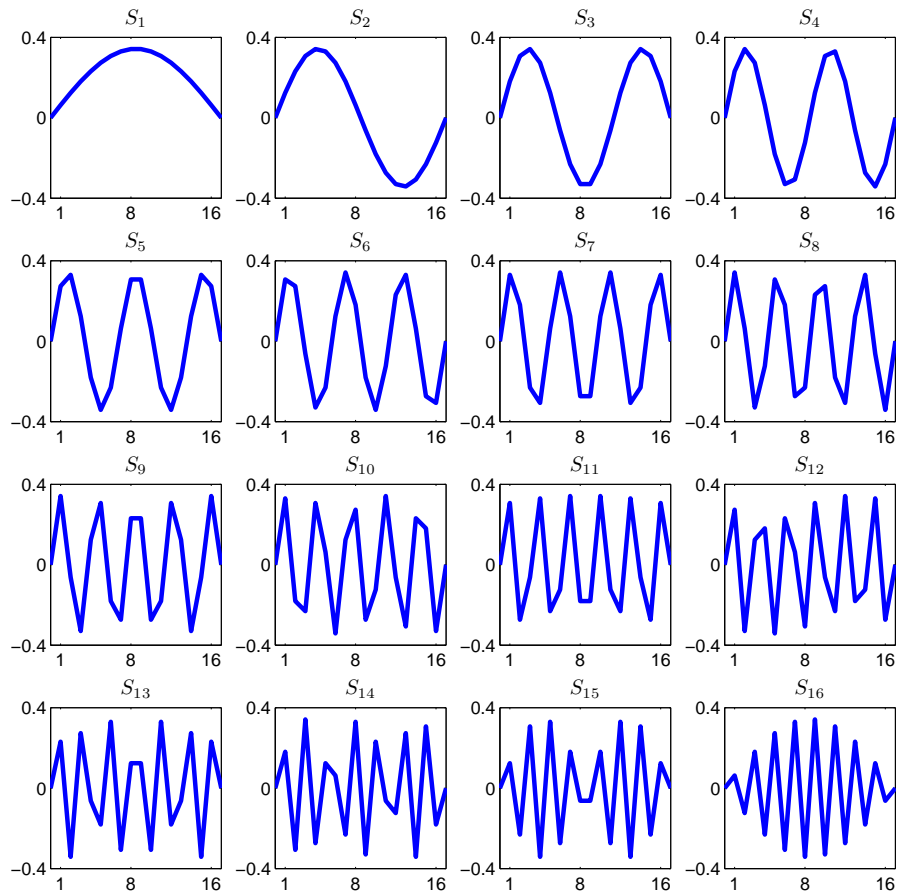
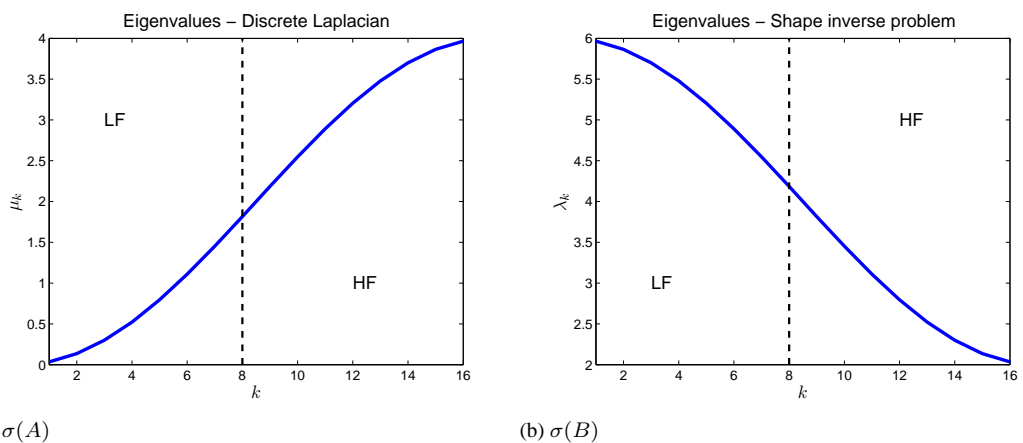


FIG. 3.2 – Vecteurs propres de  $A$  et  $B$  ( $n = 16$ ) : modes de Fourier discrets.



(a)  $\sigma(A)$

(b)  $\sigma(B)$

FIG. 3.3 – Spectre de  $A$  et  $B$  pour  $n = 16$ . Les valeurs propres sont représentées en fonction des valeurs croissantes de fréquence (à gauche les modes basse fréquence et à droite les modes haute fréquence).

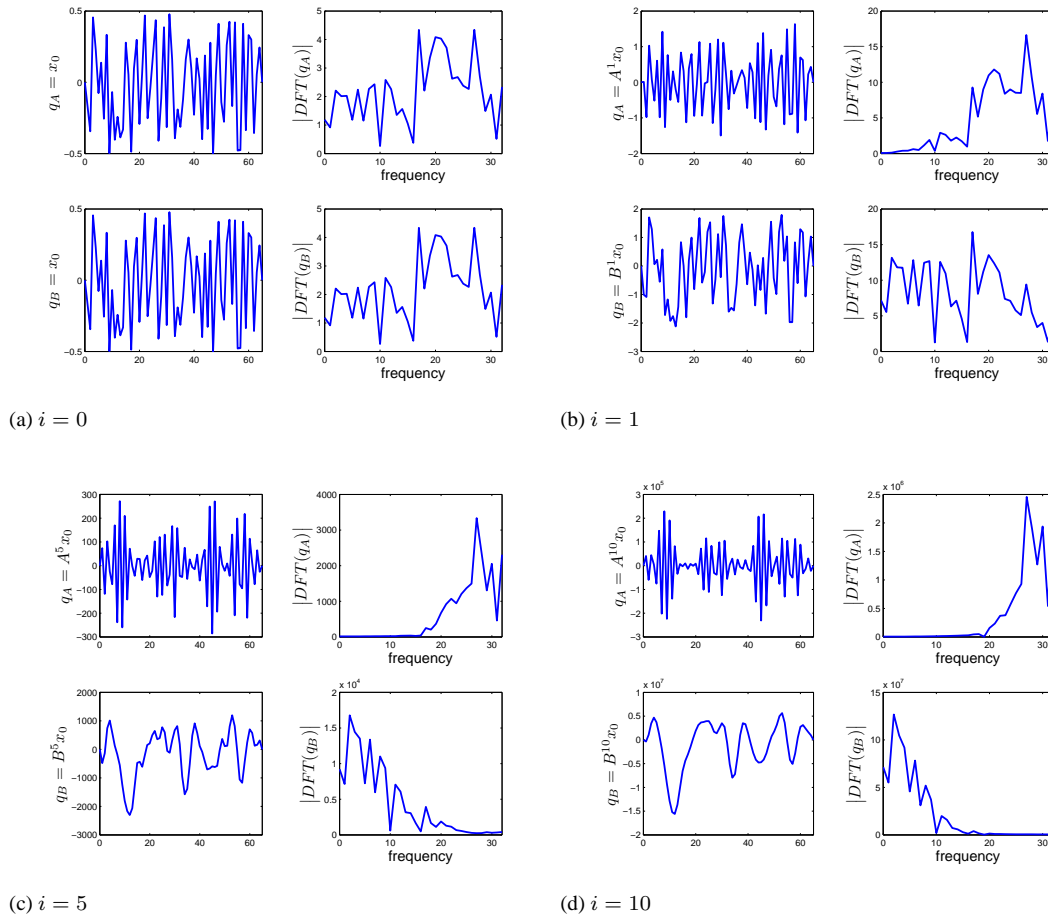


FIG. 3.4 – Vecteurs de Krylov  $q_A^i$  et  $q_B^i$  où  $q_A^0 = q_B^0 = v_h = b_h$  est un vecteur aléatoire et  $n = 64$ . La DFT du vecteur initial ne montre aucune répartition en fréquence particulière. Après une itération les modes HF de  $q_B^1$  semblent moins amplifiés que les autres (de 20 à 30); similairement, les modes LF de  $q_A^1$  sont (légèrement) moins amplifiés. Après 5 ou 10 itérations, seule la moitié des modes reste observable : les modes LF de  $q_B^i$  et les modes HF de  $q_A^i$ .

### 3.3 Méthodes itératives de base

Soit  $M$  une matrice réelle s.d.p. de dimension  $n$  et  $b$  un vecteur quelconque de  $\mathbb{R}^n$ . Dans cette section on considère le système linéaire

$$Mx = b \quad (3.25)$$

qui est équivalent au problème de minimisation sans contraintes

$$\min_{x \in \mathbb{R}^n} J(x) = \frac{1}{2}x^T Mx - b^T x. \quad (3.26)$$

#### 3.3.1 Rappels sur la convergence d'une itération linéaire

Soit la suite  $\{x^i\}$  générée par l'itération linéaire

$$x^{i+1} = Gx^i + c \quad (3.27)$$

pour  $x^0$  donné.  $G$  est appelée la *matrice d'amplification*. Le *rayon spectral*  $\rho$  d'une matrice carrée  $G$  est

$$\rho(G) = \max_i |g_i| \quad (3.28)$$

où les  $g_i$  sont les valeurs propres de  $G$ . Une condition nécessaire et suffisante pour que la suite (3.27) converge est que  $\rho(G) < 1$ . Par ailleurs si  $G$  est inversible (donc  $\forall i, g_i \neq 0$  nécessairement), alors  $I - G$  est inversible.

On suppose désormais que  $G$  est inversible et que  $\rho(G) < 1$ . La limite  $x^*$  de la suite (3.27) est le point fixe (théorème du point fixe de Banach) :

$$x^* = Gx^* + c \Leftrightarrow x^* = (I - G)^{-1}c \quad (3.29)$$

L'erreur itérative  $e^i = x^i - x^*$  satisfait donc

$$e^{i+1} = Ge^i \quad (3.30)$$

ce qui signifie que l'itération linéaire (3.27) est équivalente à l'amplification (ou amortissement) de l'erreur par  $G$ .

Supposons que  $G$  soit diagonalisable. On note  $G = T\Gamma T^{-1}$  où  $\Gamma$  est la matrice diagonale avec  $0 < |g_1| \leq \dots \leq |g_n| < 1$ . Les colonnes de  $T$  sont des vecteurs propres associés aux valeurs propres  $g_k$ . On a donc

$$\begin{aligned} T^{-1}e^{i+1} &= \Gamma T^{-1}e^i \\ \epsilon^{i+1} &= \Gamma \epsilon^i \end{aligned} \quad (3.31)$$

où  $T\epsilon^i \stackrel{\text{def.}}{=} e^i$ , c-à-d  $\epsilon^i$  est l'erreur itérative dans la base modale  $T$ . Une itération correspond donc à l'atténuation de l'erreur d'un facteur  $|g_k|$  dans la direction du mode correspondant

$$\epsilon_k^{i+1} = g_k \epsilon_k^i. \quad (3.32)$$

Les coefficients  $g_k$  sont appelés *coefficients d'amortissement*. Leur valeur absolue peuvent être interprétée comme des taux de convergence linéaire par composante modale. Le taux de convergence global d'une itération est donné par le rayon spectral  $\rho$ , c-à-d le coefficient d'amortissement de plus grande amplitude.

On examine désormais des méthodes itératives pour résoudre notre problème de deux points de vue différents : comme un système linéaire (3.25) (méthode de point-Jacobi) ou comme la minimisation d'une fonctionnelle quadratique (3.26) (méthode de plus grande descente). On explicite leur matrice d'amplification.

### 3.3.2 Itération de Jacobi

La méthode de point-Jacobi (ou de Jacobi) est la méthode itérative

$$x^{i+1} = (I - D^{-1}M)x^i + D^{-1}b \quad (3.33)$$

où  $D$  est la partie diagonale de  $M$ . Soit  $\tau > 0$  un paramètre de relaxation. L'itération de Jacobi généralisée est

$$x^{i+1} = (I - \tau D^{-1}M)x^i + \tau D^{-1}b \quad (3.34)$$

soit en terme d'erreur itérative

$$e^{i+1} = (I - \tau D^{-1}M)e^i. \quad (3.35)$$

En outre, on suppose sans perte de généralité que le problème a été préconditionné au préalable par l'inverse de la diagonale de  $M$  de telle sorte que  $D = I$  (procédé de « mise à l'échelle »). Alors la matrice d'amplification d'une itération de Jacobi est

$$G_\tau = I - \tau M. \quad (3.36)$$

Plus généralement, après  $k$  itérations de Jacobi pour les coefficients de relaxation  $\tau_j$ ,  $j = 1 \dots k$ , la matrice d'amplification s'écrit

$$G_k = (I - \tau_k M) \dots (I - \tau_1 M). \quad (3.37)$$

Les coefficients  $\tau_j$  peuvent être choisis de manière optimale afin de réduire au mieux l'erreur dans la direction des modes correspondant à un sous-ensemble du spectre. Cette méthode est connue comme la méthode de Richardson ou encore comme la méthode d'accélération de Tchebychev.

### 3.3.3 Accélération de Tchebychev, méthode de Richardson

Soit  $\lambda$  une valeur propre de  $M$  (on rappelle que  $\lambda > 0$ ) ; si le paramètre de relaxation de (3.34) est tel que  $\tau = \frac{1}{\lambda}$  alors la composante de l'erreur dans la direction du mode correspondant à cette valeur propre est annihilée en une itération. Par conséquent, si on prend  $\tau = \lambda_i$ ,  $i = 1 \dots n$  successivement alors la solution est atteinte en  $n$  itérations.

En pratique le spectre de  $M$  n'est pas connu (sinon le problème est résolu !). De manière alternative, supposons qu'un sous-ensemble de  $\sigma(M)$  est inclus dans l'intervalle  $[a, b]$ ,  $0 < a < b$ . On considère  $k$  itérations de relaxation ( $k < n$ ), soit

$$e^{i+k} = (I - \tau_k M) \dots (I - \tau_1 M)e^i. \quad (3.38)$$

Les paramètres de relaxation optimaux pour l'intervalle  $[a, b]$  sont ceux qui minimisent le rayon spectral de  $G = \prod_{i=1}^k (I - \tau_i M)$ . Cela revient à trouver les  $\tau_i$  qui minimisent la norme infinie du polynôme  $P(\lambda) = \prod_{i=1}^k (1 - \tau_i \lambda)$  où  $\lambda \in [a, b]$ .

Soit  $T_k$  le  $k^e$  polynôme de Chebychev et  $\xi_i$ ,  $i = 1 \dots k$ , ses racines. On peut montrer que le polynôme optimal a pour racines  $\mu_i = \frac{b+a}{2} + \frac{b-a}{2}\xi_i$  et que les paramètres optimaux sont leurs inverses  $\tau_i = \frac{1}{\mu_i}$ . Une preuve est donnée dans [29].

### 3.3.4 Méthode de plus grande descente

Le gradient de la fonction (3.26) est

$$\nabla J(x) = Mx - b. \quad (3.39)$$

L'application de la méthode de plus grande descente pour la minimisation de  $J$  donne l'itération linéaire suivante

$$\begin{aligned} x^{i+1} &= x^i - \tau \nabla J(x^i) \\ &= x^i - \tau(Mx^i - b) \\ &= (I - \tau M)x^i + \tau b \end{aligned} \quad (3.40)$$

où le pas de recherche linéaire  $\tau$  est soit fixé soit donné par une procédure de minimisation le long de la direction de descente  $Mx^i - b$ .

Ainsi, une itération de la méthode de plus grande descente peut être vue comme une itération de Jacobi avec paramètre de relaxation  $\tau$  si le problème a été mis à l'échelle au préalable (c-à-d la diagonale de  $M$  est l'identité). Par conséquent la matrice d'amplification est équivalente à (3.36) pour une itération et à (3.37) pour  $k$  itérations.

### 3.3.5 Préconditionnement par une méthode de type quasi-Newton

On a vu qu'une itération de quasi-Newton s'écrit

$$x^{i+1} = x^i - \tau^i (B^i)^{-1} \nabla J(x^i) \quad (3.41)$$

où  $B^i$  est le hessien approché par la formule BFGS par exemple. Il s'agit donc d'une itération de Jacobi preconditionnée.

On suppose que  $\tau^i = 1$  (cette condition est nécessaire pour obtenir un taux de convergence superlinéaire). Appliquée au problème quadratique (3.26), on a

$$\begin{aligned} x^{i+1} &= x^i - (B^i)^{-1} (Mx^i - b) \\ &= (I - (B^i)^{-1} M)x^i + (B^i)^{-1} b \end{aligned} \quad (3.42)$$

En terme d'erreur itérative, pour  $k$  itérations, cela s'écrit

$$e^{i+k} = (I - (B^{i+k-1})^{-1} M) \dots (I - (B^i)^{-1} M) e^i. \quad (3.43)$$

Autrement dit, une itération de la méthode de Richardson (3.38) peut être vue comme une itération de QN où le hessien est approché par la matrice diagonale  $B^{i-1+j} = \mu_j I$ ,  $j = 1 \dots k$  telle que  $(B^{i-1+j})^{-1} = \frac{1}{\mu_j} I = \tau_j I$ .

### 3.3.6 Fonction d'amortissement

Revenons aux problèmes modèles initiaux. On a vu l'équivalence d'une itération de Jacobi et d'une itération de plus grande descente pour lesquelles on aboutit à la même matrice d'amplification. Appliquée aux problèmes modèles, on en déduit les coefficients d'amortissement <sup>1</sup>.

Dans un premier temps on met à l'échelle les systèmes linéaires, c-à-d

$$D_A^{-1} A_h f_h = D_A^{-1} v_h \quad (3.44)$$

$$D_B^{-1} H_h x_h = D_B^{-1} b_h \quad (3.45)$$

On en déduit les matrices d'amplification

$$G_\tau^A = I - \frac{\tau}{2} A \quad (3.46)$$

$$G_\tau^B = I - \frac{\tau}{4} B \quad (3.47)$$

Les deux matrices sont diagonalisables et les coefficients d'amortissement sont

$$g_k^A = 1 - \frac{\tau}{2} \mu_k \quad (3.48)$$

$$g_k^B = 1 - \frac{\tau}{4} \lambda_k \quad (3.49)$$

En rappelant que les vecteurs propres sont les modes de Fourier discret, on représente à la Figure 3.5 la fonction d'amortissement, c-à-d les coefficients d'amortissement en fonction des paramètres de fréquence, pour plusieurs valeurs de  $\tau$  (on observe donc les taux de convergence en fonction de la fréquence du mode). Les fonctions d'amortissement sont des fonctions monotones du paramètre de fréquence : elle

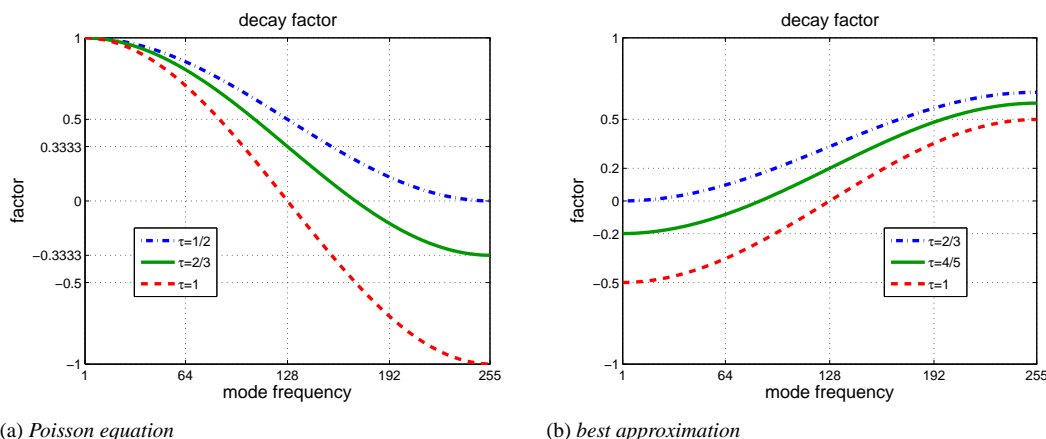


FIG. 3.5 – Fonction d'amortissement d'une itération de Jacobi (ou de plus grande descente) pour différentes valeurs du paramètre de relaxation  $\tau$  et pour  $n = 255$ .

est décroissante dans le cas de l'équation de Poisson et croissante dans le cas du problème d'optimisation.

Pourvu que les coefficients soient compris dans l'intervalle  $] - 1 \ 1[$ , le paramètre de relaxation  $\tau$  peut modifier le comportement de la convergence. Idéalement il faudrait optimiser  $\tau$  de telle sorte que le rayon spectral soit minimisé. Cependant, un problème raide telle que l'équation de Poisson discrète est caractérisé par un rayon spectral proche de 1 quelle que soit la valeur de  $\tau$ . Il est donc plus pertinent d'optimiser  $\tau$  sur un sous-ensemble de l'espace plutôt que l'espace tout entier. Par la suite on distingue l'espace d'approximation (ou espace de recherche) en deux sous-espaces supplémentaires : l'espace généré par les modes LF et l'espace généré par les modes HF, identifié respectivement par les sous-ensembles d'indices  $I_{LF}$  et  $I_{HF}$  (en supposant que  $n$  est impair)

$$I_{LF} = \left\{ 1, \dots, \frac{n-1}{2} \right\} \quad I_{HF} = \left\{ \frac{n+1}{2}, \dots, n \right\}. \quad (3.50)$$

Dans les sections qui suivent on fournit les valeurs optimales de  $\tau$  qui minimisent l'un des critères suivant

$$\rho = \max_{k \in I_{LF} \cup I_{HF}} |g_k| \quad \rho_{LF} = \max_{k \in I_{LF}} |g_k| \quad \rho_{HF} = \max_{k \in I_{HF}} |g_k| \quad (3.51)$$

en accord avec le problème considéré.

### Paramètre de relaxation optimal pour l'équation de Poisson

Tout d'abord notons que la méthode converge ssi  $\tau$  est compris dans  $]0 \ 1[$ . On assure ainsi que les coefficients d'amortissement satisfont  $|g_k^A| < 1$ .

Pour tout  $\tau$  dans  $]0 \ 1[$ , les coefficients LF sont positifs. Puisque la fonction d'amortissement est monotone décroissante par rapport à  $k$ , le coefficient de plus grande magnitude est  $g_1^A = 1 - \tau + \tau \cos \pi h = 1 - O(h)$  qui est minimum pour  $\tau = 1$ , impliquant  $\rho_{LF} = \cos \pi h$ . En d'autres termes l'itération ne peut pas être efficace sur l'espace LF à moins que  $h$  soit assez grand (c-à-d une paramétrisation avec peu de degrés de liberté). Puisque  $\rho$  est nécessairement supérieur ou égal à  $\rho_{LF}$  (ici  $\rho = \rho_{LF}$ ) il n'est pas judicieux non plus d'optimiser  $\tau$  globalement.

En revanche,  $\tau$  peut être optimisé pour minimiser  $\rho_{HF}$ . En ce sens la valeur optimale est  $\tau = \frac{2}{3}$  pour laquelle le plus grand coefficient est obtenu en  $k = \frac{n+1}{2}$  et égal à  $\frac{1}{3}$ . Il est important de noter que cette valeur est indépendante de la taille du maillage  $h$ . Dans la terminologie multigrille cet opérateur est

<sup>1</sup>l'étude de la méthode de Jacobi appliquée à l'équation de Poisson est un problème bien connu ; on rappelle les résultats dans un souci de comparaison avec le problème d'optimisation.

appelé *lisseur* puisque qu'il réduit efficacement (optimalement) les modes HF de l'erreur. Le lisseur est également appelé *solution operator* dans la littérature anglophone.

Les coefficients d'amortissement sont alors

$$g_k^A = \frac{1}{3} (1 + 2 \cos(k\pi h)) \quad (3.52)$$

Le rayon spectral correspondant à  $n$  itérations de Jacobi est donc donné par

$$\begin{aligned} \rho(G_n^A) &= \frac{1}{3^n} \max_k |(1 + 2 \cos(k\pi h))^n| \\ &= \frac{1}{3^n} \left( \max_k |1 + 2 \cos(k\pi h)| \right)^n \\ &= \frac{1}{3^n} (1 + 2 \cos(\pi h))^n \end{aligned} \quad (3.53)$$

### Paramètre de relaxation optimal pour le problème inverse d'optimisation de forme

Dans ce cas la convergence de la méthode impose que  $\tau$  soit compris dans  $]0, \frac{4}{3}]$  (autorisant ainsi  $\tau$  à être un paramètre de sur-relaxation). La principale différence avec le problème précédent est que  $\tau$  peut être optimisé pour l'un ou l'autre des sous-espaces considérés, LF ou HF. Cette propriété est due au fait qu'aucune valeur propre de  $B$  n'est proche de 0.

D'un point de vue global, le paramètre optimal qui minimise  $\rho$  est obtenu pour  $\tau = 1$  et donne  $\rho = \frac{1}{2} \cos(\pi h)$ , qui est dépendant du maillage bien que borné supérieurement par  $\frac{1}{2}$ . Cela définit l'itération qui a le plus petit taux de convergence linéaire. Sur le sous-espace HF, la valeur optimale est  $\rho_{HF} = \frac{1}{3}(2 \cos(\pi h) - 1)$  et atteint pour  $\tau = \frac{4}{3}$ , qui est aussi dépendant du maillage et borné supérieurement par  $\frac{1}{3}$ .

On supposera toutefois qu'aucune de ces configurations n'est envisageable. On justifie ce choix pour deux raisons : d'une part on cherche à développer une méthode idéale dont le taux de convergence est indépendant du maillage, ce qui n'est pas possible si notre « lisseur » l'est ; d'autre part, dans d'autres paramétrisations, telle que celle de Bézier-Bernstein, le problème est raide (la plus petite valeur propre tend vers 0, voir [30]) et pour lequel aucun paramètre de relaxation ne peut améliorer de manière significative la convergence sur l'espace HF. Cela est bien illustré numériquement à la section 3.5.1. Cependant, puisque nous n'avons pas de forme analytique du spectre, il ne nous est pas possible de conduire une étude rigoureuse de la convergence des schémas idéaux proposés.

Dans ce contexte on cherche à minimiser  $\rho_{LF}$  (le *solution operator* est alors un bon *anti-lisseur*). La valeur optimale est  $\rho_{LF} = \frac{1}{5}$  (qui est bien indépendant du maillage) et obtenue pour  $\tau = \frac{4}{5}$ . On a alors

$$g_k^B = \frac{1}{5} (1 - 2 \cos(k\pi h)) \quad (3.54)$$

Ainsi, le rayon spectral de la matrice d'amplification équivalente à  $n$  itérations de Jacobi est

$$\begin{aligned} \rho(G_n^B) &= \frac{1}{5^n} \max_k |(1 - 2 \cos(k\pi h))^n| \\ &= \frac{1}{5^n} \left( \max_k |1 - 2 \cos(k\pi h)| \right)^n \\ &= \frac{1}{5^n} (1 + 2 \cos(\pi h))^n \end{aligned} \quad (3.55)$$

### Influence de la taille du maillage sur le rayon spectral

Il est intéressant d'observer le comportement du rayon spectral de la méthode itérative de base lorsque la taille du maillage diminue afin de comparer ce résultat avec les méthodes hiérarchiques des sections suivantes.

Lorsque  $h \rightarrow 0$  on a  $\cos(\pi h) \rightarrow 1$ . Donc, le rayon spectral de chacun des problèmes a la propriété suivante

$$\rho(G_n^A) \rightarrow \frac{1}{3^n} (1+2)^n = 1 \quad (3.56)$$

ce qui signifie que la méthode tend vers une méthode qui ne converge pas, alors que

$$\rho(G_n^B) \rightarrow \left(\frac{3}{5}\right)^n \quad (3.57)$$

ce qui signifie que dans le pire des cas la méthode convergera toujours.

À la section suivante notre but est de montrer comment ce résultat peut être amélioré dans le contexte de l'optimisation de forme en s'appuyant sur les démonstrations du problème de Poisson. Par améliorer on entend trouver un cycle idéal dont le taux de convergence est plus petit que la valeur limite (3.57) et indépendant de la taille du maillage  $h$ . Les résultats nous permettront de construire des stratégies multiniveaux non-linéaire pour l'optimisation d'antenne.

### 3.4 Algorithmes multigrilles

La popularité des méthodes MG repose sur deux propriétés : (1) le taux de convergence est indépendant du maillage ; (2) la complexité de l'algorithme est optimalement proportionnelle à la taille  $n$  du maillage.

Dressons ici brièvement les concepts de base des MG (voir [29, 101] pour une présentation approfondie). Premièrement admettons que nous sommes pourvus d'un maillage que nous appelons *maillage fin*. L'espace d'approximation qui repose sur ce maillage peut être vu comme la somme directe de deux sous-espaces en somme directe : un espace basse fréquence et un espace haute fréquence. L'efficacité des MG pour la résolution numérique d'une EDP repose sur la complémentarité de deux méthodes :

1. une méthode itérative simple (Jacobi, Gauss-Seidel, SOR, etc.) réduit efficacement les composantes HF de l'erreur : le *lisseur* (voir section 3.3) ;
2. des opérateurs de transfert entre le maillage fin et des maillages grossiers (à définir) sont utilisés pour représenter ce qu'il reste de l'erreur dans le sous-espace LF ; on cherche alors à réduire ou annihiler les composantes LF de l'erreur : c'est la *correction de grille grossière* (ou *Coarse Grid Correction CGC*).

Les cycles MG sont composés de phases de relaxation (lissage) et de phases de correction de grille grossière. Sur le maillage fin on a vu que le rayon spectral du lisseur est dépendant du maillage, à cause de son inefficacité à réduire les modes LF de l'erreur. En revanche il est indépendant du maillage sur la partie HF. Dans un second temps les modes LF de l'erreur sont bien représentés sur un maillage grossier. Relativement à ce maillage grossier les modes de plus hautes fréquences deviennent les modes HF du nouveau problème. De nouveau, le lisseur peut être appliqué sur ce nouveau maillage afin de réduire efficacement (c-à-d avec un taux de convergence indépendant du maillage) l'erreur dans la direction des modes HF nouvellement définis. En outre, le coût de calcul du lissage est plus faible puisque la taille du maillage est plus petite. Ceci peut être répété récursivement sur des maillages de plus en plus grossiers. On suppose que le nombre de d.d.l. sur le dernier niveau est assez petit pour que l'on puisse résoudre exactement le problème (par exemple à l'aide d'une méthode directe). Dans ce cas l'algorithme est dit *idéal*.

Un cycle MG peut être vu comme une itération linéaire. Rigoureusement, on étudie l'efficacité du cycle en analysant le spectre de la matrice d'amplification équivalente au cycle. Un cycle bigrille idéal est suffisant pour montrer que le taux de convergence est dépendant ou non du maillage. C'est le but de cette section.

L'équation de Poisson est le problème standard typique des MG linéaires. En effet, la structure spectrale du problème discret est telle que le transfert du résidu à une grille grossière est exact (car les modes propres sont les modes de Fourier discret) ; en 1D, la méthode de Jacobi est un bon lisseur pour la moitié



supérieure du spectre (HF). Dans le cadre de l'optimisation de forme ces propriétés ne sont pas immédiates. Dans le cas particulier du problème de meilleure approximation en norme  $L_2$  dans l'espace des fonctions linéaires par morceaux, le transfert est également exact (les modes propres sont également les modes de Fourier discret). Malheureusement la méthode de Jacobi (ou de plus grande descente) n'a pas la propriété requise de lissage. Au contraire on a vu qu'elle a de bonne propriété d'« anti-lissage ». On peut donc douter de la pertinence d'une méthode MG classique. Intuitivement la restriction du résidu sur une grille grossière ne semble pas utile puisque le problème aura déjà été bien résolu sur l'espace grossier. On démontre cette assertion de manière rigoureuse en étudiant le rayon spectral d'un cycle bigrille idéal. Par la suite on examine deux opérateurs de transfert alternatifs pour pallier ce problème, pour lesquels on conduit une étude similaire des cycles idéaux nouvellement définis.

### 3.4.1 Opérateurs de prolongement et de restriction

Soit deux grilles  $\mathcal{T}_h$  et  $\mathcal{T}_{2h}$  comme définies à la section 3.2.3 avec respectivement  $n$  et  $n'$  nœuds intérieurs (voir Figure 3.6). On les appelle grille fine et grille grossière. Supposons que la dimension du problème est  $n = 2^p - 1$  sur la grille fine et  $n' = 2^{p-1} - 1$  sur la grille grossière pour  $p > 1$ . Dans ce cas on a la relation  $2n' + 1 = n$ , ce qui revient à dire que la taille de la discrétisation sur la grille grossière  $h'$  est deux fois plus grande  $h' = 2h$ . On cherche à résoudre les systèmes linéaires  $A_h u_h = f_h$  et  $H_h x = b_h$  sur la grille fine  $\mathcal{T}_h$ .

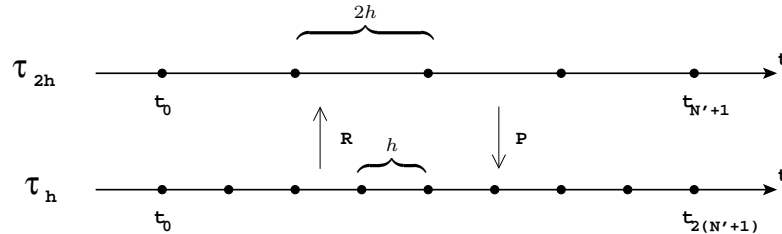


FIG. 3.6 – Discrétisations fine  $\mathcal{T}_h$  et grossière  $\mathcal{T}_{2h}$ .

Tout d'abord on définit des opérateurs de transfert entre ces grilles : un opérateur de prolongement  $P : \mathcal{T}_{2h} \rightarrow \mathcal{T}_h$  et un opérateur de restriction  $R : \mathcal{T}_h \rightarrow \mathcal{T}_{2h}$ . Puisqu'on a considéré des conditions aux limites de Dirichlet, on ne considère que les nœuds intérieurs. Une fois que ces opérateurs sont définis, on distingue deux définitions d'un sous-problème sur la grille grossière :

- Approximation de Galerkin (Galerkin Coarse grid Approximation, GCA) : la matrice  $A'$  de la grille grossière est obtenue par projection de la matrice de la grille fine en utilisant les opérateurs de transfert, soit  $A' = RA_h P$  ;
- Approximation discrète (Discrete Coarse grid Approximation, DCA) : la matrice  $A'$  de la grille grossière est obtenue par discrétisation sur le maillage  $\mathcal{T}_{2h}$  du problème original, soit  $A' = A_{2h}$ .

Si les opérateurs sont bien définis, alors ces deux définitions sont équivalentes. Cette propriété est adéquate pour les preuves de convergence comme on peut le constater dans l'Annexe A.1<sup>2</sup>.

<sup>2</sup>On parle de *propriétés variationnelles* lorsque les opérateurs de transfert vérifient  $A_{2h} = RA_h P$  et qu'il existe  $\alpha > 0$  tel que  $R = \alpha P^T$ .

Soit  $P$  l'opérateur d'interpolation linéaire :

$$P = \begin{pmatrix} \frac{1}{2} & & & & & & & & & \\ 1 & & & & & & & & & \\ \frac{1}{2} & \frac{1}{2} & & & & & & & & \\ & 1 & & & & & & & & \\ & & \frac{1}{2} & & & & & & & \\ & & & \ddots & & & & & & \\ & & & & & & & & & \\ & & & & & & \frac{1}{2} & & & \\ & & & & & & 1 & & & \\ & & & & & & \frac{1}{2} & & & \\ & & & & & & & \frac{1}{2} & & \\ & & & & & & & & 1 & \\ & & & & & & & & & \frac{1}{2} \end{pmatrix}. \quad (3.58)$$

En ce qui concerne l'équation de Poisson, si l'opérateur de restriction est l'opérateur  $R \stackrel{\text{def.}}{=} \frac{1}{2}P^T$  (opérateur de moyenne), alors les définitions GCA et DCA sont équivalentes ( $A_{2h} = \frac{1}{2}P^T A_h P$ ). En ce qui concerne le problème d'optimisation de forme, l'équivalence est obtenue si  $R \stackrel{\text{def.}}{=} P^T$  ( $H_{2h} = P^T H_h P$ ).

Par la suite, dans les schémas, on note  $\searrow$  l'opérateur de restriction,  $\nearrow$  ou  $\uparrow$  l'opérateur de prolongement,  $\square$  une phase de relaxation et  $\blacksquare$  une correction idéale. On note  $r_h = A_h u_h - f_h$  (resp.  $r_h = H_h x - b_h$ ) le résidu. Pour tout  $u_h$  (resp.  $x$ ) on a l'égalité  $Ae_h = Au_h - f_h = r$  (resp.  $H_h e_h = H_h x - b_h = r_h$ ) où  $u^*$  (resp.  $x^*$ ) est la solution exacte et  $e_h = u_h - u^*$  (resp.  $e_h = x - x^*$ ) l'erreur.

**Remarque** Dans les sections qui suivent, le contexte sera suffisant pour déterminer si l'on traite du problème de Poisson ou d'optimisation de forme. On enlève donc les exposants  $A$  et  $B$  des matrices d'amplification  $G$  pour alléger les notations. Ainsi  $G_k$  fait référence à une relaxation de  $k$  itérations de Jacobi, dont le paramètre est optimisé selon le problème.

### 3.4.2 Enrichissement progressif : *Nested Iteration*

Supposons que l'on dispose de  $n + 1$  grilles  $\mathcal{T}_{2^k h}$ ,  $k = 0 \dots n$  (un niveau fin et  $n$  niveaux intermédiaires et grossiers) construites de la même manière que  $\mathcal{T}_{2h}$ . On note  $P_{k+1}^k$  l'opérateur de prolongement entre deux niveaux successifs.

La stratégie la plus intuitive est la stratégie dite d'*enrichissement progressif* ou *Nested Iteration*. Elle consiste à fournir au niveau fin une valeur initiale  $u_h^{(0)}$  obtenue par prolongement d'une approximation du système sur les grilles grossières. Dans un premier temps on considère le problème sur la grille la plus grossière, soit

$$A_{2^n h} u_{2^n h} = f_{2^n h} \quad (3.59)$$

que l'on traite par relaxation. Vu le faible nombre de degrés de liberté, ce problème est supposé simple. Le prolongement de cette première approximation sert de préconditionneur pour le niveau suivant

$$u_{2^{n-1} h}^{(0)} = P_n^{n-1} u_{2^n h} \quad (3.60)$$

On résout alors le problème de nouveau par relaxation, ce qui donne, pour  $\nu$  itérations de Jacobi

$$u_{2^{n-1} h} = G_\nu u_{2^{n-1} h}^{(0)} + f'_{2^{n-1} h} \quad (3.61)$$

où  $f'_h$  est un terme qui dépend de  $f_{2^{n-1} h}$ ,  $\tau$ ,  $G_\tau$ , et  $\nu$ . Ce problème est plus coûteux et plus raide mais converge plus vite à partir de l'approximation de la grille précédente qu'avec une approximation quelconque.

La nouvelle approximation est alors prolongée sur le niveau suivant et constitue de nouveau un préconditionneur pour ce problème plus fin. Ainsi de suite, on répète ces opérations jusqu'au dernier niveau

pour lequel on effectue autant d'itérations que nécessaire pour atteindre une tolérance numérique donnée.

$$\begin{aligned} u_{2^{k-1}h}^{(0)} &= P_k^{k-1} u_{2^k h} \\ u_{2^{k-1}h} &= G_\nu u_{2^{k-1}h}^{(0)} + f'_{2^{k-1}h} \end{aligned} \quad (3.62)$$

Schématiquement, l'algorithme d'enrichissement progressif est illustré à la Figure 3.7.

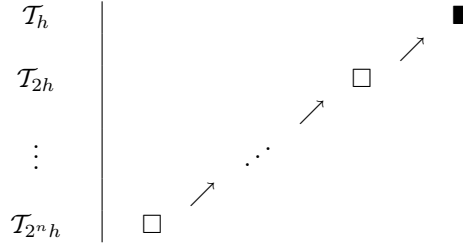


FIG. 3.7 – Schéma de l'algorithme d'enrichissement progressif.

On reviendra sur cette stratégie au chapitre suivant. En effet, l'algorithme d'optimisation PSO associé à une stratégie d'enrichissement progressif s'avère plus robuste que sa version originelle. Ce résultat concerne l'optimisation globale. En matière d'optimisation locale on s'intéresse à des schémas plus élaborés dont l'algorithme de base est l'algorithme bigrille idéal.

### 3.4.3 Algorithme bigrille idéal

Soit deux grilles  $\mathcal{T}_h$  et  $\mathcal{T}_{2h}$ . L'approximation initiale est notée  $u_h^{(0)}$ . Un algorithme bigrille idéal met en jeu trois phases :

1. une phase de pré-relaxation sur la grille fine : lissage de l'erreur grâce à quelques itérations de Jacobi ; on obtient une première approximation  $u_h^{(1)}$  :

$$u_h^{(1)} = G_k u_h^{(0)} + f'_h \quad (3.63)$$

2. une phase de correction de grille grossière : le résidu

$$r = A_h u_h^{(1)} - f_h \quad (3.64)$$

vérifie sur la grille grossière l'équation

$$R A_h P e_{2h} = R r \quad (3.65)$$

résolue exactement. On en déduit une deuxième approximation de la solution sur la grille fine par prolongement de l'erreur  $e_{2h}$  sous la forme d'une correction

$$u_h^{(2)} = u_h^{(1)} - P e_{2h} = u_h^{(1)} - P (R A_h P)^{-1} R r \quad (3.66)$$

3. une phase de post-relaxation : une dernière approximation est donnée par relaxation de  $u_h^{(2)}$ , soit

$$u_h^{(3)} = G_k u_h^{(2)} + f'_h \quad (3.67)$$

Ces trois phases sont illustrées à la Figure 3.8.

Pour l'étude de la convergence d'un tel algorithme on considérera sans perte de généralité une seule itération de Jacobi en guise de relaxation, comme indiqué dans l'Algorithme 3. On en déduit la matrice d'amplification équivalente à un cycle idéal :

$$G = G_\tau \left[ I - P (R A_h P)^{-1} R A_h \right] G_\tau. \quad (3.68)$$

On montre l'efficacité de l'algorithme en étudiant le rayon spectral de  $G$ .

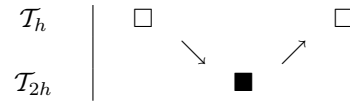


FIG. 3.8 – Schéma de l'algorithme bigrille idéal.

**Algorithme 3** : Algorithme bigrille idéal.

// Relaxation

$$u_h^{(1)} = G_\tau u_h^{(0)} + \tau f_h$$

// Correction de Grille Grossière

$$r = A_h u_h^{(1)} - f_h$$

$$R A_h P e_{2h} = R r$$

$$u_h^{(2)} = u_h^{(1)} - P e_{2h} = u_h^{(1)} - P (R A_h P)^{-1} R r$$

// Relaxation

$$u_h^{(3)} = G_\tau u_h^{(2)} + \tau f_h$$

**Rayon spectral du cycle idéal**

La preuve de la convergence est donnée dans l'Annexe A.1 à des fins de comparaison avec le problème d'optimisation. La trame de la preuve consiste dans un premier temps à montrer que  $G$  est similaire à une matrice de la forme  $\Sigma D^2$  (voir). On en déduit alors le spectre de  $G$  d'après celui de  $\Sigma D^2$  :

- $n'$  valeurs propres sont nulles ;
- $n' + 1$  valeurs propres sont égales à  $\frac{1}{9}$  ;

On a donc  $\rho(G) = \frac{1}{9}$ . D'une part la méthode est convergente ( $\rho < 1$ ), d'autre part elle est indépendante du maillage. La convergence est donc efficace quelle que soit la taille de la grille fine.

Pour donner un ordre d'idée, sur une grille fine de  $n = 31$ , soit  $h = \frac{1}{32}$ , on peut donner le nombre d'itérations de Jacobi qu'il faut pour atteindre le même taux de convergence. On cherche donc  $n$  tel que

$$\rho(G_n) \leq \frac{1}{9} \quad (3.69)$$

Soit

$$n \geq \frac{\log(9)}{\log(3) - \log(1 + 2 \cos \pi h)} \approx 684 \quad (3.70)$$

Sachant qu'on a considéré 2 itérations de Jacobi pour les phases de pré- et post-relaxation, il faut donc que le travail effectué pour la CGC soit inférieur à 682 itérations. Pour les résultats de complexité on renvoie par exemple à [29, 101].

**3.4.4 MG, FMG, FAS**

Dans l'Algorithme 3, la phase de correction est supposée résolue exactement. En considérant d'autres grilles plus grossières on peut appliquer récursivement l'algorithme idéal pour résoudre cette phase. Un cycle bigrille devient un cycle multigrille. Sur le niveau le plus bas on suppose également que le problème est résolu exactement. Plus la grille est grossière, moins la relaxation est coûteuse et plus il est facile de résoudre complètement le problème puisque la raideur diminue avec la dimension.

Un cycle symétrique composé de phases de pré- et post-relaxation est appelé *cycle en V* ; un cycle composé d'une phase de pré-relaxation uniquement est dit *cycle en dent de scie*. ces schéma MG sont illustrés à la Figure 3.9.

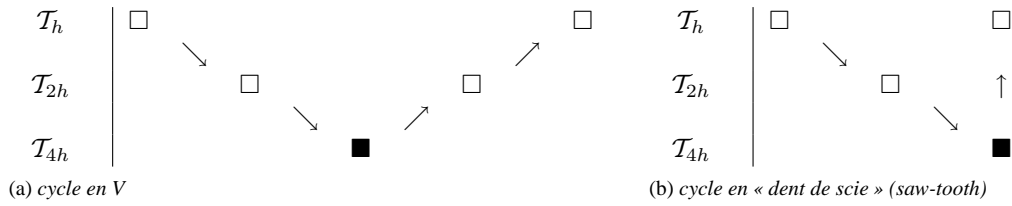


FIG. 3.9 – Schémas de cycles MG sur trois niveaux

Enfin, le schéma le plus élaboré est le schéma dit *Full Multigrid* (FMG) : il rassemble dans un même processus la technique d'enrichissement progressif et les cycles MG. Plus précisément, il s'agit d'une stratégie d'enrichissement où chaque niveau est résolu par un ou plusieurs cycles MG (le nombre de cycles peut être déterminé par une condition sur le résidu). Brièvement, sur trois niveaux (grossier-moyen-fin), il se décompose de la manière suivante :

1. Le problème est résolu complètement sur la grille grossière.
2. La solution est prolongée sur la grille suivante.
3. Un cycle à deux niveaux est utilisé pour résoudre le problème sur le niveau moyen.
4. La solution est prolongée sur la grille fine.
5. Un cycle à trois niveaux est utilisé pour résoudre le problème.

Ce schéma est illustré à la Figure 3.10. Pour plusieurs grilles intermédiaires on répète les étapes 2 et 3 autant de fois que nécessaire. On montre que la complexité de cet algorithme est optimalement proportionnelle au nombre de nœuds de la grille fine.

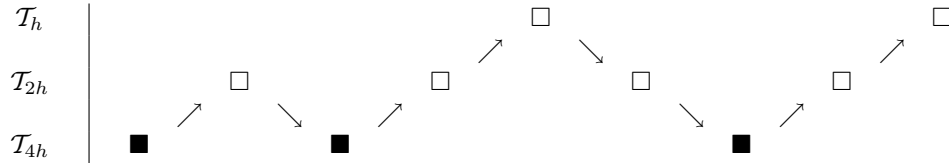


FIG. 3.10 – Schéma de l'algorithme FMG sur trois niveaux : nested + MG.

Enfin, ces schémas sont valides pour les problèmes linéaires mais ne sont pas adaptés à la résolution d'un système non linéaire du type

$$\mathcal{A}(u) = f \quad (3.71)$$

En effet, dans ce cas le résidu non linéaire  $r = \mathcal{A}(u) - f$  ne vérifie pas la relation  $\mathcal{A}(e) = r$ . Dans ce cas la consistance du problème sur les grilles grossières n'est pas évidente. L'algorithme *FAS* prend en compte un terme de correction qui assure cette consistance. On verra par la suite, dans l'extension des méthodes MG à l'optimisation de forme d'antenne où le gradient de la fonctionnelle est non-linéaire par rapport aux paramètres de conception, qu'une formulation de correction est suffisante pour traiter la non-linéarité (sachant que les contraintes, elles, restent linéaires).

### 3.4.5 Application des algorithmes classiques pour l'optimisation de forme

On s'intéresse désormais au problème d'optimisation. Dans un premier temps on applique le schéma bigrille idéal tel qu'il est décrit à la section 3.4.3. On considère toujours les paramétrisations P1 sur  $\mathcal{T}_h$  et  $\mathcal{T}_{2h}$  ainsi que l'opérateur d'interpolation linéaire  $P$  comme opérateur de prolongement.

Soit  $x^{(0)}$  l'approximation initiale des paramètres de conception. On décrit les phases de l'algorithme idéal en terme d'optimisation :

- Les phases de relaxation sont regardées comme une itération de plus grande descente avec pas  $\tau$ .
- La CGC est regardé comme la minimisation exacte sur la paramétrisation grossière du problème de correction

$$\min_{y \in \mathcal{T}_{2h}} J(x + Py) \quad (3.72)$$

pour  $x$  donné par une phase de relaxation. La fonction objectif sur la grille grossière est toujours quadratique et le domaine est convexe, le minimum global est donc caractérisé uniquement par les conditions de stationnarité.

Avec une phase de pré-relaxation et une phase de post-relaxation on a l'Algorithme 4.

---

**Algorithme 4** : Algorithme bigrille idéal en optimisation.

---

// Relaxation

$$x^{(1)} = G_\tau x^{(0)} + \tau b_h$$

// Correction de Grille Grossière

$$\begin{aligned} \Delta y &= \arg \min_{y \in \mathcal{T}_{2h}} j(y) = J(x^{(1)} + Py) \\ x^{(2)} &= x^{(1)} + P\Delta y \end{aligned}$$

// Relaxation

$$x^{(3)} = G_\tau x^{(2)} + \tau b_h$$


---

Explicitons la condition de stationnarité sur la grille grossière. Puisque le gradient est linéaire par rapport à  $x$  on a

$$\begin{aligned} g(y) &= P^T \nabla J(x^{(1)} + Py) \\ &= P^T (H_h(x^{(1)} + Py) - b_h) \\ &= P^T H_h P y + P^T (H_h x^{(1)} - b_h) \\ &= P^T H_h P y + P^T r \end{aligned}$$

où  $r = H_h x^{(1)} - b_h$  est le résidu linéaire. La condition de stationnarité s'écrit alors  $g(\Delta y) = 0$ , soit

$$\Delta y = - (P^T H_h P)^{-1} P^T r \quad (3.73)$$

donc

$$x^{(2)} = x^{(1)} - P (P^T H_h P)^{-1} P^T r. \quad (3.74)$$

Similairement à l'Algorithme 3 on en déduit la matrice d'amplification équivalente au cycle

$$G = G_\tau \left[ I - P (P^T H_h P)^{-1} P^T H_h \right] G_\tau \quad (3.75)$$

qui a exactement la même structure que celle de (3.68). Notons toutefois qu'ici, l'opérateur de restriction doit être le transposé de l'opérateur de prolongement :

$$R \stackrel{\text{def.}}{=} P^T. \quad (3.76)$$

### Rayon spectral du cycle idéal

On donne le détail du calcul du rayon spectral en annexe à la section A.1.3. Comme pour le problème de Poisson on commence par effectuer des transformations similaires (toujours en annexe, à la

section A.1.1). On montre que  $G$  est également similaire à une matrice de la forme  $\Sigma D^2$  où  $D$  est diagonale et  $\Sigma$  a la structure particulière illustrée à la Figure A.1.

Les valeurs propres sont cependant plus délicates à obtenir. En effet, si la structure de  $\Sigma$  est la même, il est difficile d'en simplifier les valeurs. On adopte alors la stratégie suivante :

- a) une hypothèse sur la structure des vecteurs propres est effectuée ;
- b) des systèmes linéaires sont déduits de cette hypothèse ;
- c) les systèmes sont résolus en utilisant le logiciel de calcul symbolique Maple ;
- d) les solutions sont vérifiées ainsi que l'indépendance linéaire des vecteurs propres.

On obtient le résultat suivant :

- $n'$  valeurs propres  $\alpha_k$  sont nulles.
- $\alpha_{n'+1} = \frac{1}{25}$  est une valeur propre évidente.
- les  $n'$  autres valeurs propres données par Maple sont

$$\alpha_{n'+1+k} = \frac{1}{25} \frac{27 - 144 \cos^2\left(\frac{\theta_k}{2}\right) + 304 \cos^4\left(\frac{\theta_k}{2}\right) - 320 \cos^6\left(\frac{\theta_k}{2}\right) + 160 \cos^8\left(\frac{\theta_k}{2}\right)}{3 - 8 \cos^2\left(\frac{\theta_k}{2}\right) + 8 \cos^4\left(\frac{\theta_k}{2}\right)}.$$

Le rayon spectral du cycle idéal s'écrit donc

$$\begin{aligned} \rho(\Sigma D^2) &= \max_{k=1..n} |\alpha_k| \\ &= \max_{k=n'+1..n} |\alpha_k| \end{aligned}$$

On peut montrer que  $\alpha_{n'+1+k} > \frac{1}{25}$  pour  $k = 1 \dots n'$  et que  $\alpha$  est une fonction monotone décroissante sur  $\theta \in ]0, \frac{\pi}{2}[$ . Alors le maximum est atteint en  $\theta = \theta_1$ , c-à-d  $\rho(\Sigma D^2) = \alpha_{n'+2}$  qui est dépendant du maillage. Lorsque l'on fait tendre la taille du maillage vers 0 ( $h \rightarrow 0$ ) on a  $\alpha_{n'+2} \rightarrow \frac{9}{25} = \left(\frac{3}{5}\right)^2$ . C'est exactement le comportement de 2 itérations de Jacobi (voir la section 3.3.6). En d'autres termes la correction de grille grossière est inutile.

### 3.4.6 Redéfinition des opérateurs de transfert

On a bien vu que la stratégie classique est inefficace pour le problème d'optimisation de forme. En effet, il nous manque la complémentarité entre le lisseur et la correction de grille grossière : la phase de relaxation agissant comme un anti-lisseur et la restriction du résidu sur la grille grossière comme un filtre passe-bas, seuls les modes déjà bien atténués (LF) sont représentés. Les valeurs propres nulles de la matrice d'amplification du cycle idéal correspondent aux modes annihilés : les modes de basse fréquence.

Dans cette section on s'attache à la définition de nouveaux opérateurs de transfert pour le problème d'optimisation. Dans ce cadre, les notions de prolongement et de restriction sont étendues à d'autres paramétrisations, bien que l'étude de la convergence se fasse toujours sur la base des fonctions P1. On garantit toutefois que la nouvelle définition est cohérente avec la formulation classique.

Soit  $F$  l'espace de recherche de référence (la paramétrisation fine) et  $V$  un sous-espace de  $F$  (la paramétrisation grossière).  $F$  est isomorphe à  $\mathbb{R}^n$  et  $V$  à  $\mathbb{R}^{n'}$  ( $n' < n$ ). Comme on l'a vu à la section 3.2.3, l'espace fin repose sur la paramétrisation  $\{u_k\}_{k=1}^n$ . En tant qu'élément de  $F$ , tout élément  $v$  de  $V \subset F$  s'écrit dans cette base de manière unique. Dans une formulation générale, un sous-espace  $V$  est défini par

$$V \equiv \left\{ v = \sum_{k=1}^n x_k u_k \in F \mid x = Qy, \quad \forall y \in \mathbb{R}^{n'} \right\} \quad (3.77)$$

où  $Q$  est la matrice de l'application linéaire de  $\mathbb{R}^{n'}$  dans  $\mathbb{R}^n$  qui exprime tout  $v$  de  $V$  dans la base de paramétrisation de  $F$  (les colonnes de  $Q$  constituent une base de  $V$  exprimée dans  $F$ ) :  $Q$  peut être vu comme un opérateur de prolongement.

Selon le type de transfert, la notion de « grille », fine ou grossière, n'a plus lieu d'être. Puisqu'il s'agit d'un problème d'optimisation, on parle encore d'espace de recherche de grande et petite dimension. Par ailleurs, par « grossier » on entend peu de degrés de liberté mais pas nécessairement un espace d'approximation plus lisse (basse fréquence).

Trois cas sont considérés :

1. les sous-espaces sont des espaces de paramétrisations emboîtées ; cette formulation équivaut à la formulation classique étendue à d'autres paramétrisations.
2. les sous-espaces sont des espaces de paramétrisations emboîtées préconditionnés par une matrice anti-diagonale de permutation (matrice d'échange) afin de définir un filtre passe-haut ; c'est la méthode proposée dans [30].
3. les sous-espaces sont des sous-espaces propres ; dans cette dernière formulation on s'affranchit des espaces de paramétrisations emboîtées et on adopte une stratégie de type algébrique. Sur la base de l'analyse de Fourier des vecteurs propres on est en mesure de définir n'importe quel type de filtre comme opérateur de restriction.

Pour chacune de ces formulations on examine le rayon spectral du cycle idéal lorsque la paramétrisation est celle des éléments P1.

### Paramétrisations emboîtées

Soit  $n$  espaces de paramétrisations emboîtées  $V_1 \subset V_2 \subset \dots \subset V_n$ . Chaque espace est muni d'une base  $\{u\}_{k=1}^{n'}$  (par ex. les espaces polynômiaux de degré croissant dans la base de Bernstein, les fonctions B-splines de degré constant avec insertion de nœuds, etc.). Soit  $F = V_n$  l'espace fin. On suppose que pour tous ces espaces il existe une application linéaire de  $\mathbb{R}^{n'}$  dans  $\mathbb{R}^n$  notée  $E_{n'}^n$  telle que

$$\forall y \in \mathbb{R}^{n'}, \quad x = E_{n'}^n y \quad \Rightarrow \quad v = \sum_{k=1}^{n'} y_k u_k^{n'} = \sum_{k=1}^n x_k u_k^n \in V_{n'}. \quad (3.78)$$

Autrement dit  $E_{n'}^n$  est l'application qui transforme les composantes de  $v$  dans  $V_{n'}$  dans la base de  $F$ . Pour les espaces polynômiaux  $\mathcal{P}^{n'}$  et  $\mathcal{P}^n$  (de dimension  $n' + 1$  et  $n + 1$  resp.) dans la base de Bernstein, cette application est l'*élévation du degré* (cf. section 2.2.4). Formellement les sous-espaces sont

$$V_{n'} \equiv \text{span}\{u_k^{n'}\} \equiv \left\{ v = \sum_{k=1}^n x_k u_k^n \in F \quad \middle| \quad x = E_{n'}^n y, \quad \forall y \in \mathbb{R}^{n'} \right\} \iff Q = E_{n'}^n. \quad (3.79)$$

**Remarque** L'opérateur de prolongement  $Q = E_{n'}^n$  exhibe une structure triviale lorsque la base des sous-espaces  $V_{n'}$  est composée d'un sous-ensemble de la base de l'espace fin  $F$ . C'est le cas de la base canonique et des polynômes orthogonaux pour les espaces polynômiaux. Précisément cette opérateur devient

$$E_{n'}^n = \begin{pmatrix} I_{n' \times n'} \\ 0_{(n-n') \times n} \end{pmatrix}.$$

où  $I_{n' \times n'}$  est la matrice identité et  $0_{(n-n') \times n}$  une matrice nulle.

Si l'espace de paramétrisation est défini avec les éléments P1, alors  $Q$  est l'opérateur classique d'interpolation linéaire  $Q = P$ . On a déjà montré que dans ce cas, l'Algorithme 4 n'est pas efficace. Pour d'autres paramétrisations telles que la paramétrisation de Bézier-Bernstein, des expériences numériques corroborent ce résultat [102, 20].

### Réorganisation de la structure spectrale par permutation du spectre

Dans cette deuxième section on examine la méthode proposée dans [30]. On garde l'hypothèse selon laquelle on est muni d'espaces de paramétrisations emboîtées et de leurs bases ainsi que d'opérateurs



d'élévation du degré  $E_{n'}^n$  (prolongement classique). Un nouvel opérateur de transfert est défini de telle sorte qu'il agisse comme un filtre passe-haut. Les modes de haute fréquence sont projetés dans un espace « grossier ».

Soit  $H$  la matrice hessienne. Comme elle est réelle s.p.d., alors elle est diagonalisable, ses vecteurs propres constituent une base orthonormée de  $\mathbb{R}^n$ , et ses valeurs propres sont strictement positives :  $H = \Omega \Lambda \Omega^T$ ,  $\Omega^T \Omega = \Omega \Omega^T = I_n$ . On suppose (sans perte de généralité) que les valeurs propres  $\lambda_k$  sont triées par ordre croissant. Soit les sous-espaces suivants :

$$V_{n'} \equiv \left\{ v = \sum_{k=1}^n x_k u_k \in F \mid x = \Omega \mathbb{P} \Omega^T E_{n'}^n y, \quad \forall y \in \mathbb{R}^{n'} \right\} \iff Q = \Omega \mathbb{P} \Omega^T E_{n'}^n \quad (3.80)$$

où  $\mathbb{P}$  est la matrice anti-diagonale de permutation

$$\mathbb{P} = \begin{pmatrix} & & & 1 \\ & & & \\ & & \dots & \\ & & & \\ 1 & & & \end{pmatrix}. \quad (3.81)$$

Le rôle d'un tel opérateur est de réorganiser les couples de vecteurs et valeurs propres de sorte que les modes de haute fréquence soient associés aux valeurs propres les plus grandes.

### Stratégie algébrique : sous-espaces propres

Enfin, dans cette dernière formulation on s'affranchit des espaces de paramétrisations emboîtées. Elle peut être vue comme une version algébrique du transfert. On ne se préoccupe plus de l'analyse de Fourier des modes propres, seule la relation entre la grandeur des valeurs propres et la vitesse de convergence nous intéresse.

Soit l'espace de recherche fin  $F$  muni de la base  $\{u_k\}$ . Les sous-espaces sont directement déduits de la diagonalisation du hessien : le sous-espace de dimension  $n'$  est généré par les  $n'$  derniers vecteurs propres (on a supposé au préalable que les vecteurs propres sont ordonnés par ordre décroissant, de manière consistante avec la section 3.2.3), c-à-d

$$H = \Omega \Lambda \Omega^T = \begin{pmatrix} \Omega_1 & \Omega_2 \end{pmatrix} \begin{pmatrix} \Lambda_1 & 0 \\ 0 & \Lambda_2 \end{pmatrix} \begin{pmatrix} \Omega_1^T \\ \Omega_2^T \end{pmatrix} \quad (3.82)$$

où

$$\Omega_1 = \begin{pmatrix} \omega_1 & \dots & \omega_{n-n'} \end{pmatrix}, \quad \Omega_2 = \begin{pmatrix} \omega_{n+1-n'} & \dots & \omega_n \end{pmatrix} \quad (3.83)$$

et

$$V_{n'} \equiv \text{span}\{\Omega_2\} \equiv \left\{ v = \sum_{k=1}^n x_k u_k \in F \mid x = \Omega_2 y, \quad \forall y \in \mathbb{R}^{n'} \right\} \iff Q = \Omega_2. \quad (3.84)$$

**Remarque** Ce formalisme prend en compte simplement les contraintes linéaires d'égalité. Dans ce cas  $H$  est pris comme le hessien projeté sur l'espace admissible. Alors, les vecteurs propres  $\Omega$  appartiennent également à l'espace admissible. Une correction grossière est prolongée sur la paramétrisation fine en utilisant une base orthonormée  $Z$  du noyau de la jacobienne des contraintes :  $x = Z \Omega_2 y$  et  $Q = Z \Omega_2$ . Ainsi, les espaces de recherche grossiers ne sont pas soumis à ces contraintes d'égalité.

### 3.4.7 Révision de l'algorithme bigrille idéal

Les sous-espaces  $V$  sont considérés comme des espaces de correction. En d'autres termes, pour  $x^* \in F$  donné, le problème grossier consiste en un problème de minimisation sur l'espace affine  $x^* + V$ , soit  $x = x^* + \Delta x$  où  $\Delta x = Q y$  quel que soit  $y \in \mathbb{R}^{n'}$ . La fonction objectif du problème grossier est

$$j(y) = J(x^* + Q y). \quad (3.85)$$

Donc le gradient de  $j$  est donné par

$$g(y) = Q^T G(x^* + Qy) \quad (3.86)$$

et la matrice hessienne

$$h(y) = Q^T H(x^* + Qy) Q. \quad (3.87)$$

La phase CGC de l'Algorithme 4 peut être réécrite comme

$$\begin{aligned} \Delta y &= \arg \min_{y \in V} j(y) = J(x^{(1)} + Qy) \\ x^{(2)} &= x^{(1)} + Q\Delta y \end{aligned}$$

Et la matrice d'amplification équivalent à un cycle est donné par

$$G = G_\tau \left[ I - Q (Q^T H_h Q)^{-1} Q^T H_h \right] G_\tau. \quad (3.88)$$

La formulation est sensiblement la même si ce n'est que les sous-espaces  $V$  sont généralisés à une définition plus large des espaces de paramétrisations et  $Q$  est un des opérateurs de prolongement proposés à la section précédente.

Notons que l'opérateur de restriction est automatiquement le transposé de l'opérateur de prolongement. Par ailleurs on ne sait rien *a priori* de la matrice du problème grossier  $Q^T H_h Q$ . C'est pourquoi les sous-problèmes sont définis au sens d'une approximation GCA. Autrement dit le choix de l'opérateur  $Q$  définit le sens de la paramétrisation grossière.

### Rayon spectral du cycle idéal avec réorganisation des valeurs propres

On considère toujours la paramétrisation P1 sur chacun des niveau et  $E_{n'}^n = P$ . On conduit une analyse spectrale de la matrice d'amplification (3.88) avec  $Q = \Omega P \Omega^T P$ . Comme pour le calcul du rayon de la méthode classique on adopte la méthodologie suivante :

- a) L'expression du problème grossier  $H' = Q^T H_h Q$  est simplifié.
- b) On en déduit une forme plus simple de  $I - Q H'^{-1} Q^T H_h$  et des transformations de similarité sont appliquées à  $G$  ; Il s'ensuit que  $G$  est similaire à une matrice de la forme  $\Sigma D^2$  où  $D$  est diagonale.
- c) Le calcul des entrées de  $\Sigma$  exhibe une structure illustrée à la Figure A.1.
- d) On propose une hypothèse sur la structure des vecteurs propres de  $\Sigma D^2$  et le système linéaire résultant est résolu avec l'aide de Maple. On en déduit une formule analytique des vecteurs propres et par conséquent le rayon spectral.

La démonstration est faite en annexe (voir la section A.1.4). On trouve  $\rho(G) = \frac{1}{25}$ .

### Rayon spectral du cycle idéal algébrique

La preuve est beaucoup plus simple que pour la méthode précédente étant donné que la base est constituée des vecteurs propres. Comme ces derniers sont orthonormaux alors  $G$  est simplifiée trivialement. On trouve également  $\rho(G) = \frac{1}{25}$  (voir la section A.1.5).

### Comparaison des algorithmes idéaux avec plusieurs pas d'une méthode de gradient

D'après l'étude du rayon spectral des algorithmes idéaux étendus à l'optimisation de forme, le rayon spectral du cycle est indépendant du maillage. Avec deux itérations de Jacobi/gradient (pré- et/ou post-relaxation) avec pas  $\tau = \frac{4}{5}$  le rayon spectral est

$$\rho_{MG} = \frac{1}{25}. \quad (3.89)$$

Avec  $n$  itérations de plus grande descente on a (voir 3.3.6)

$$\rho_J < \left(\frac{3}{5}\right)^n \quad (3.90)$$

quelle que soit la taille de maille. Il faudrait donc seulement  $n \geq \frac{\log(25)}{\log(5)-\log(3)} \approx 6.30$  itérations pour atteindre le même taux de convergence (c-à-d  $n \geq 7$ ).

En d'autres termes, le travail effectué pour résoudre le problème grossier ne doit pas excéder le travail équivalent à 5 itérations de Jacobi pour que la stratégie MG soit plus efficace. Encore une fois, on rappelle que cela est dû au fait que le problème d'optimisation dans la base P1 est bien conditionné. Le conditionnement peut devenir sévèrement mauvais dans d'autres paramétrisations et/ou lorsque la physique entre en jeu.

### 3.4.8 Expériences numériques

Afin d'illustrer et de confirmer les résultats théoriques obtenus précédemment, on conduit quelques expériences numériques d'un algorithme multigrille sur un problème de reconstruction de forme avec les différents opérateurs de prolongement proposés.

La paramétrisation est linéaire par morceaux sur les grilles  $\mathcal{T}_{2^p h}$  où  $p = 0, 1, 2, \dots$  et  $h = \frac{1}{n+1}$  est la taille de maille de la paramétrisation fine. Le membre de droite est donné par le vecteur illustré à la Figure 3.11. On réalise l'expérience pour les maillages fins  $n = 127$  et  $n = 512$ .

On représente la convergence en terme de norme  $L_2$  discrète du résidu  $r(x) \stackrel{\text{def.}}{=} \|H_h x - b_h\|_{L_2}$ . Le seuil de convergence est fixé à  $\varepsilon = 10^{-14} r^0$  où  $r^0 = r(x^{(0)})$  est le résidu de l'approximation initiale  $x^{(0)} = \mathbf{0}$ .

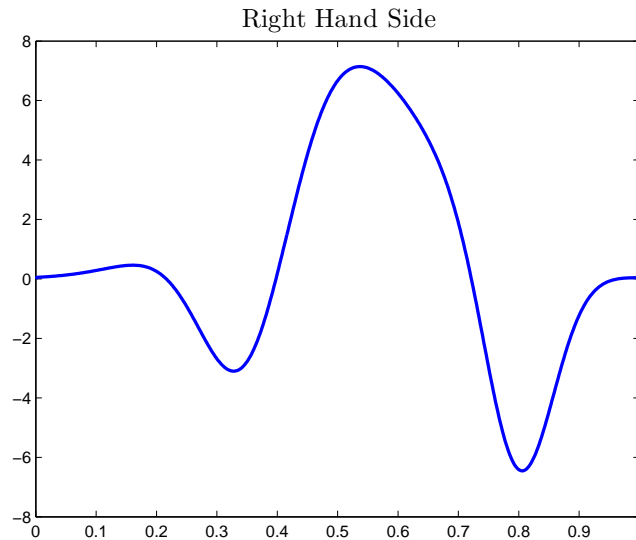
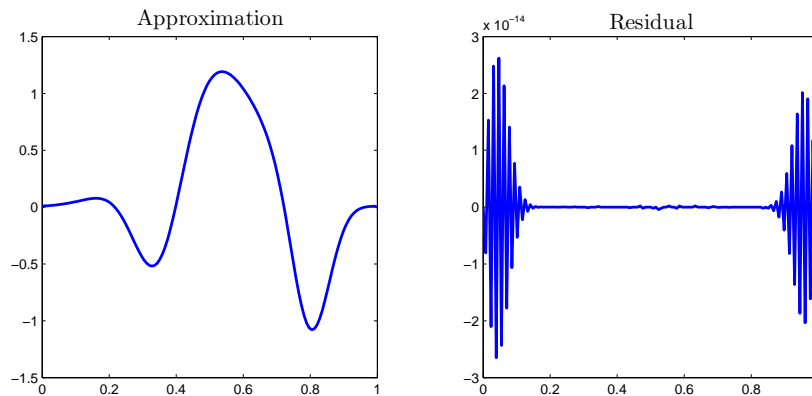


FIG. 3.11 – Cas-test : membre de droite.

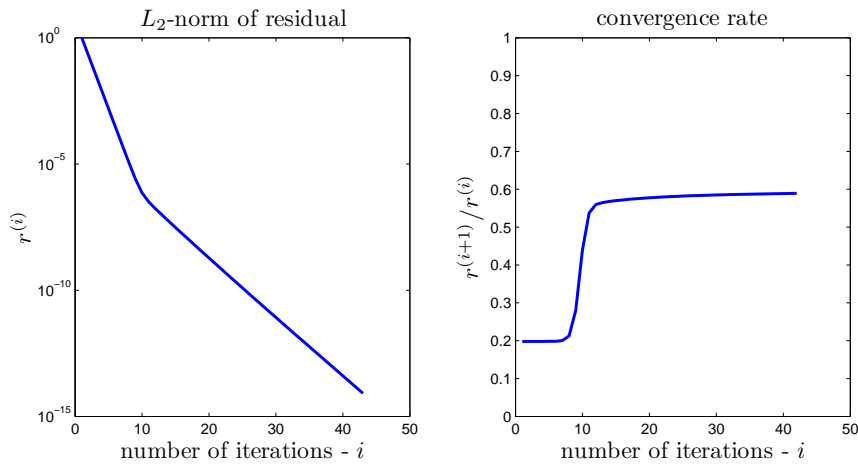
#### Paramétrisation $n = 127$

- niveaux :  $p = 0, \dots, 6$
- paramètre de relaxation :  $\tau = \frac{4}{5}$
- pré-relaxation :  $\nu_1 = 1$  itération de Jacobi
- post-relaxation :  $\nu_2 = 1$  itération de Jacobi

Les résultats obtenus sont illustrés sur les Figures 3.12 (Jacobi), 3.13 (MGV classique), et 3.14 (MGV optimisation). Comme prévu, le taux de convergence de la méthode de Jacobi tend vers  $\frac{3}{5} = 0.6$ . Le seuil de convergence est atteint en un peu plus de 40 itérations. Le résidu est alors de haute fréquence. Le taux de convergence de la méthode MGV classique tend vers  $(\frac{3}{5})^2 = 0.36$ , soit l'équivalent de deux itérations de Jacobi (pré- et post-relaxation). Le seuil est atteint en une vingtaine de cycles, soit une quarantaine d'itérations de Jacobi. Le résidu est également de haute fréquence. Enfin, le taux de convergence de la méthode MGV modifié est de  $\frac{1}{25} = 0.04$  comme prédit. La convergence est atteinte en une dizaine de cycles, soit 20 itérations de Jacobi sur la grille fine. Le résidu est de basse fréquence.

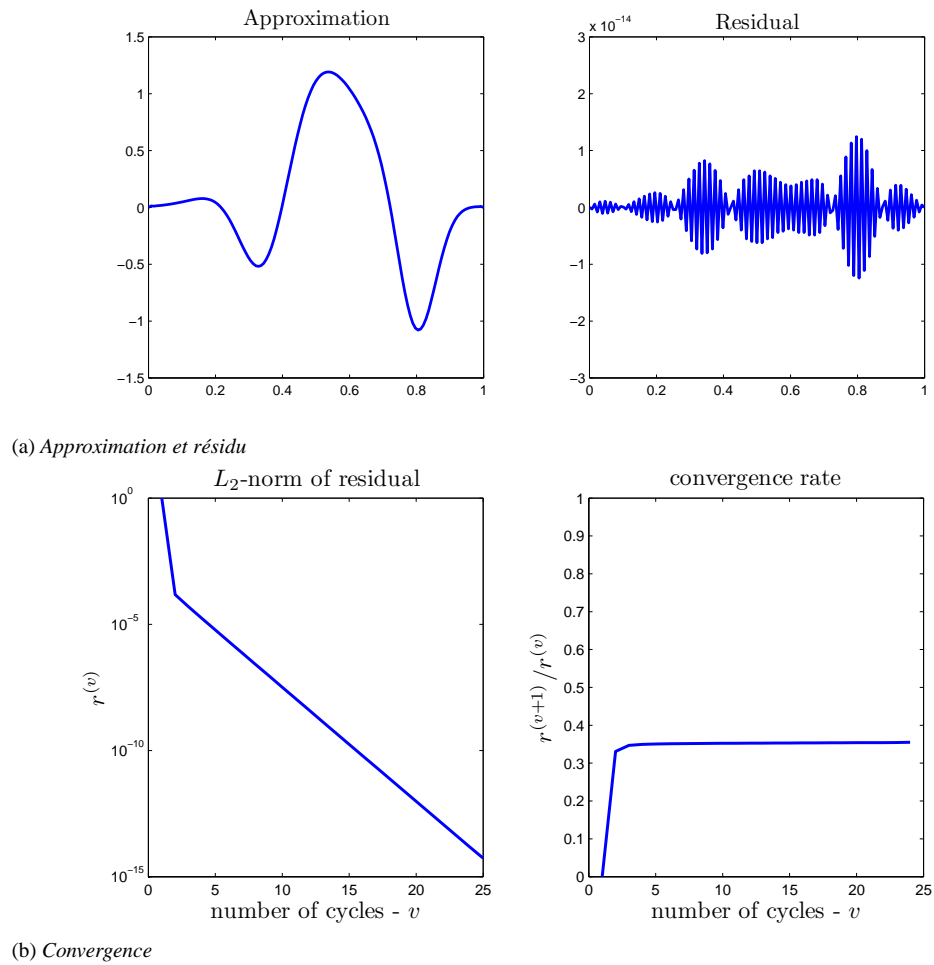


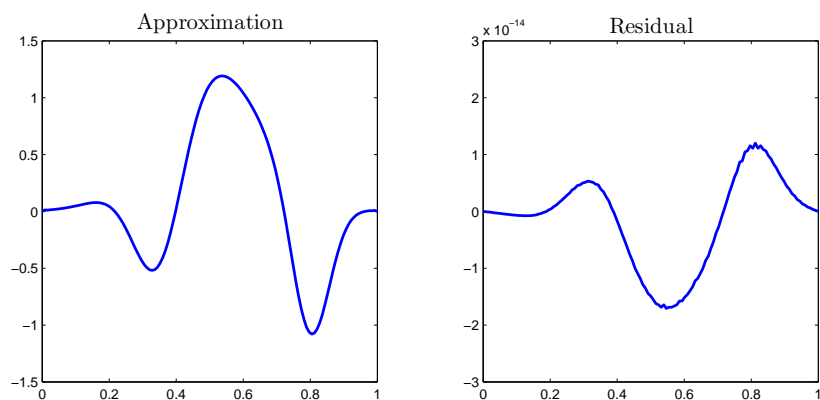
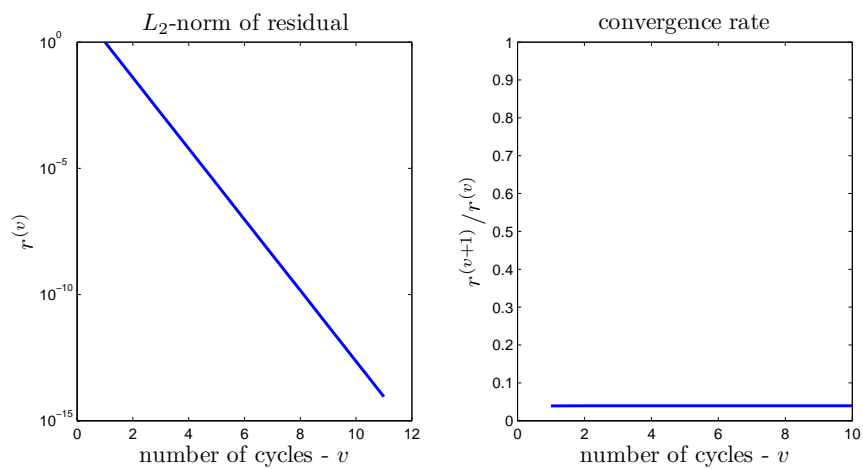
(a) Approximation et résidu



(b) Convergence

FIG. 3.12 – Méthode de Jacobi,  $n = 127$ .

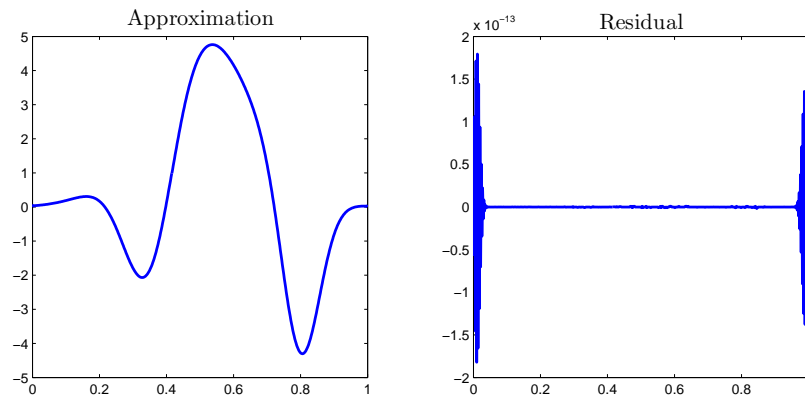
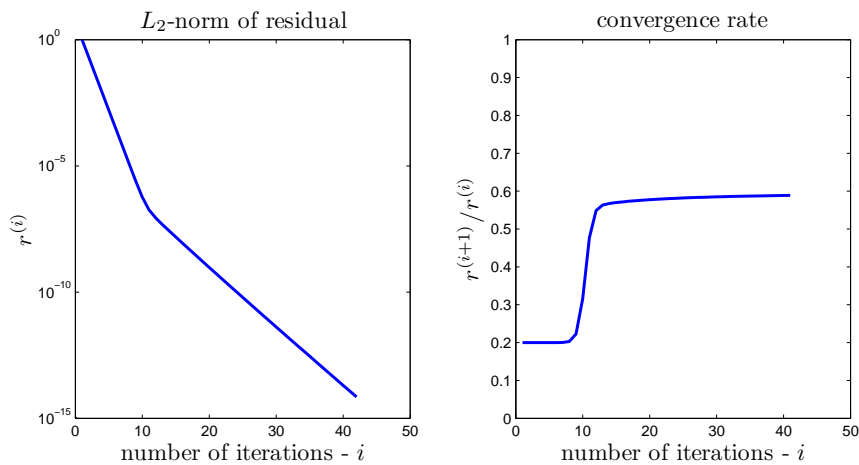
FIG. 3.13 – Méthode MGV classique,  $n = 127$ .

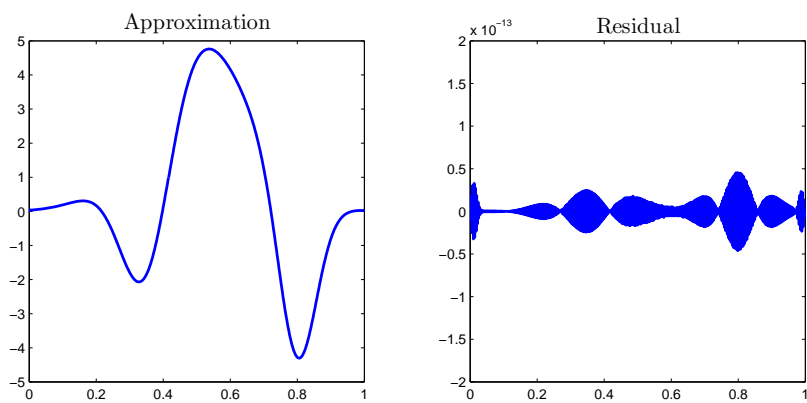
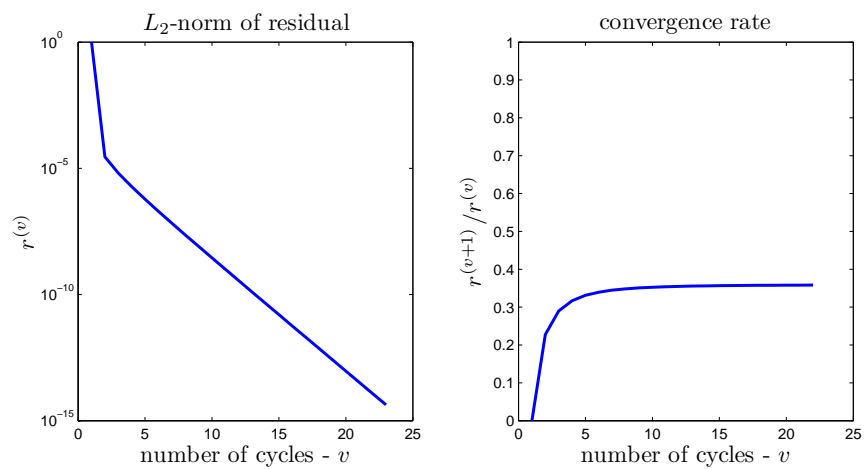
(a) *Approximation et résidu*(b) *Convergence*FIG. 3.14 – *Méthode MG*V optimisation,  $n = 127$ .

**Paramétrisation**  $n = 511$ 

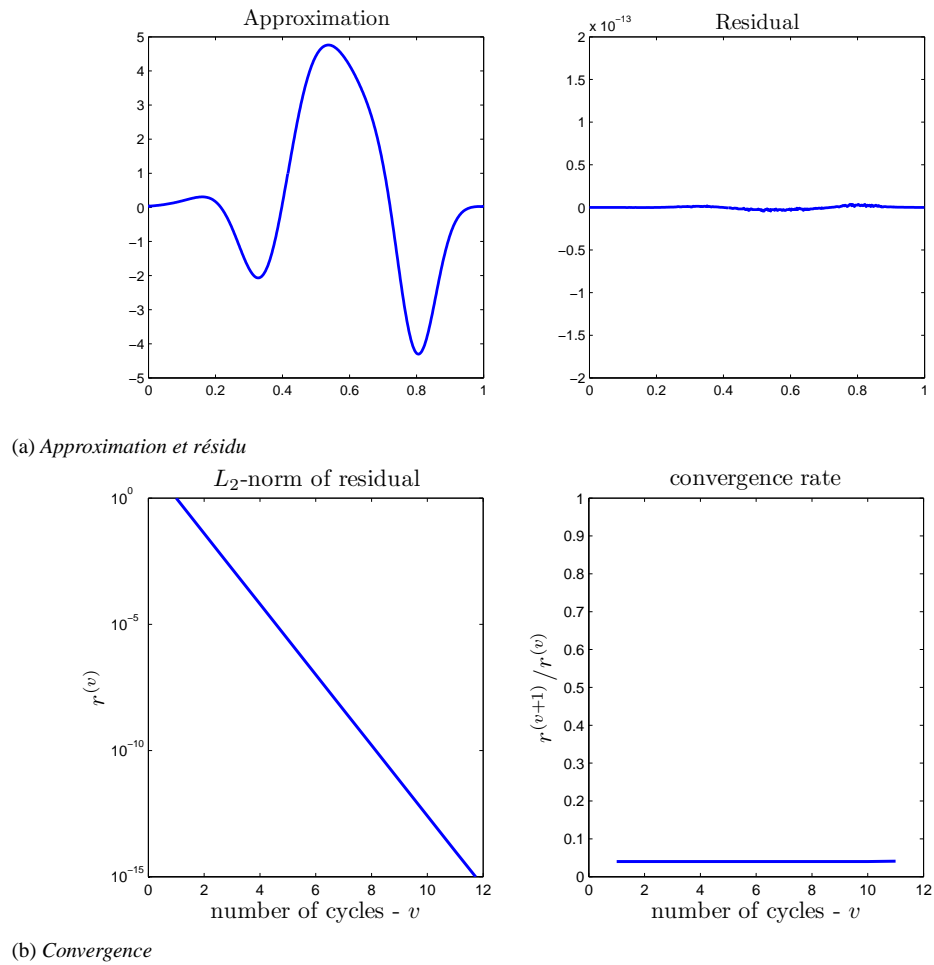
- niveaux :  $p = 0, \dots, 8$
- paramètre de relaxation :  $\tau = \frac{4}{5}$
- pré-relaxation :  $\nu_1 = 1$  itération de Jacobi
- post-relaxation :  $\nu_2 = 1$  itération de Jacobi

Les résultats obtenus sont illustrés sur les Figures 3.15 (Jacobi), 3.16 (MGV classique), et 3.17 (MGV optimisation). Pour les trois méthodes, des taux de convergence identiques à ceux de la paramétrisation  $n = 127$  sont obtenus. Il faut donc un même nombre de cycles pour atteindre le seuil de convergence fixé.

(a) *Approximation et résidu*(b) *Convergence*FIG. 3.15 – *Méthode de Jacobi*,  $n = 511$ .

(a) *Approximation et résidu*(b) *Convergence*FIG. 3.16 – *Méthode MGV classique,  $n = 511$ .*



FIG. 3.17 – Méthode MGV optimisation,  $n = 511$ .

## 3.5 Influence de la paramétrisation sur la raideur

La raideur se traduit par un mauvais conditionnement de la matrice hessienne. La norme matricielle induite par la norme vectorielle  $\|\cdot\|_p$  est définie par

$$\|A\|_p = \sup_{x \neq 0} \frac{\|Ax\|_p}{\|x\|_p}. \quad (3.91)$$

Le nombre de conditionnement  $\kappa_p$  d'une matrice inversible  $A$  est

$$\kappa_p(A) = \|A\|_p \|A^{-1}\|_p. \quad (3.92)$$

En particulier, pour  $p = 2$  on a

$$\kappa_2(A) = \frac{\lambda_n}{\lambda_1}. \quad (3.93)$$

### 3.5.1 Paramétrisation de Bézier

On reformule le problème de meilleure approximation dans un espace polynômial. Dans la base de Bernstein cela revient à considérer une représentation de Bézier du graphe de  $u$ . En comparaison avec la paramétrisation P1, une forme plus régulière avec peu de paramètres est obtenue.

On définit  $n + 2$  fonctions de bases  $u_k$  par  $u_k = B_{n+1}^k$  (voir la section 2.2.4). Seul  $B_{n+1}^0$  est non nul en  $t = 0$  et  $B_{n+1}^{n+1}$  non nul en  $t = 1$ , les conditions de Dirichlet s'écrivent alors simplement

$$x_0 = 0, \quad x_{n+1} = 0.$$

Le hessien projeté sur l'espace admissible est immédiatement obtenu en supprimant les premières et dernières lignes et colonnes du hessien. Les éléments de la matrice hessienne sont donnés par [39] :

$$h_{kj} = \int_0^1 u_k(t)u_j(t)dt = \frac{C_{n+1}^k C_{n+1}^j}{C_{2(n+1)}^{k+j}} \frac{1}{2(n+1)+1}. \quad (3.94)$$

On ne connaît pas de forme analytique des vecteurs et valeurs propres d'une telle matrice. On les approche numériquement.

Pour être consistant avec l'analyse conduite avec les éléments linéaires par morceaux, on suppose que le problème a été mis à l'échelle (c-à-d, préconditionné par l'inverse de la diagonale). Les facteurs d'amplification sont donc obtenus par le calcul des paires de vecteurs et valeurs propres de  $D_H^{-1}H$  où  $D_H$  est la diagonale de  $H$ , le hessien projeté. Les Figures 3.18 et 3.19 illustrent les vecteurs propres et valeurs propres respectivement.

On observe une structure spectrale similaire à ceux des éléments P1 : les vecteurs propres ressemblent à des modes de Fourier dans le mesure où chaque mode est caractérisé par un nombre d'alternances de signe ; les modes LF sont associés aux plus grandes valeurs propres et les modes HF aux plus petites. La matrice est cependant très mal conditionnée : pour  $n = 16$  on a  $\kappa_2(D_H^{-1}H) \approx 10^9$ . Il en résulte une matrice d'amplification de l'erreur telle qu'aucun paramètre de relaxation  $\tau$  ne peut définir un lisseur. Pour illustrer que  $H$  amplifie les modes LF, on représente les vecteurs de Krylov  $q^i$  pour  $x_0$  quelconque et  $n = 32$  à la Figure 3.20

À la Figure 3.21 on montre les facteurs d'amplification d'une itération de Jacobi dont la matrice est

$$G_\tau = I - \tau D_H^{-1}H.$$

La convergence est obtenue ssi  $\tau \in ]0, \frac{2}{\lambda_{\max}}[$  où  $\lambda_{\max} \approx 6$ . Dans cet intervalle, le taux de convergence des modes HF reste proche de 1. Ici, si on se réfère à la définition des ensembles LF et HF de la section 3.3.6, même  $\rho_{LF}$  est proche de 1. Autrement dit, aucune des méthodes de bases considérées n'est capable de réduire efficacement l'erreur sur la partie LF du spectre. Par conséquent, la paramétrisation grossière d'un cycle idéal doit être redéfinie en tant qu'espace de dimension supérieure à la moitié de la dimension de la paramétrisation fine. Pour un algorithme à plusieurs niveaux, cela revient à considérer beaucoup de niveaux intermédiaires.

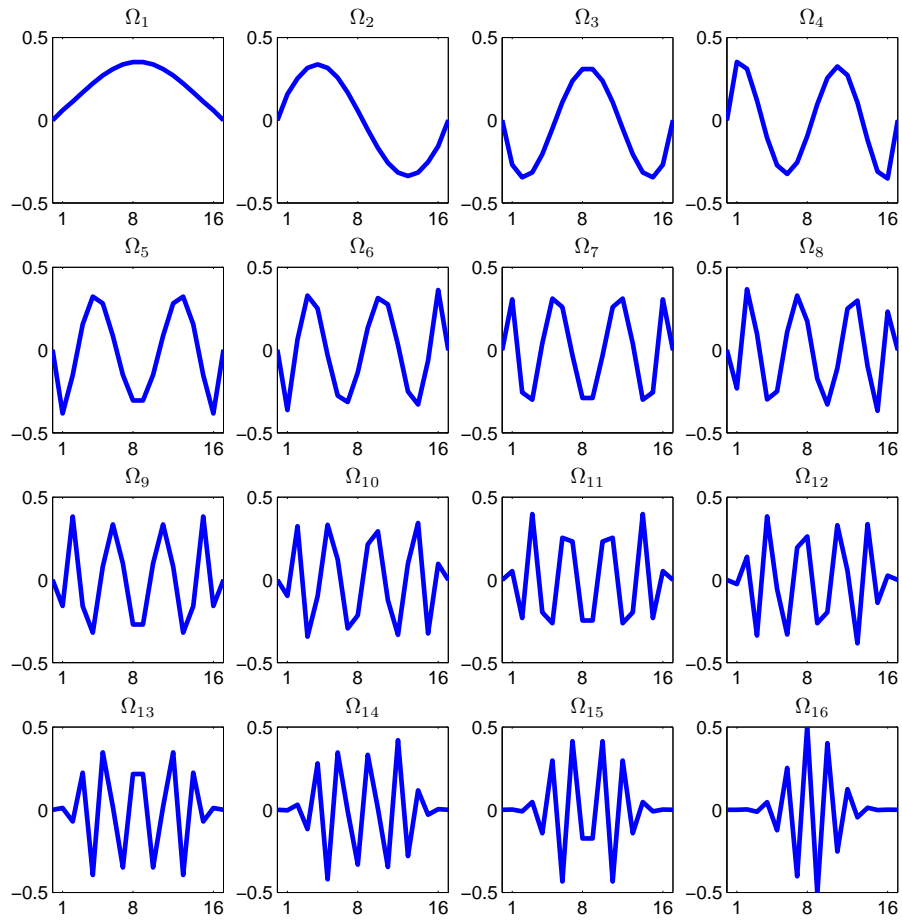


FIG. 3.18 – Bernstein - vecteurs propres -  $n = 16$

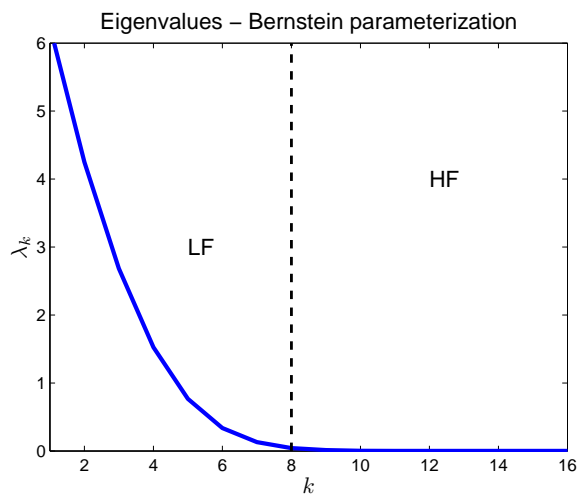


FIG. 3.19 – Bernstein - valeurs propres -  $n = 16$

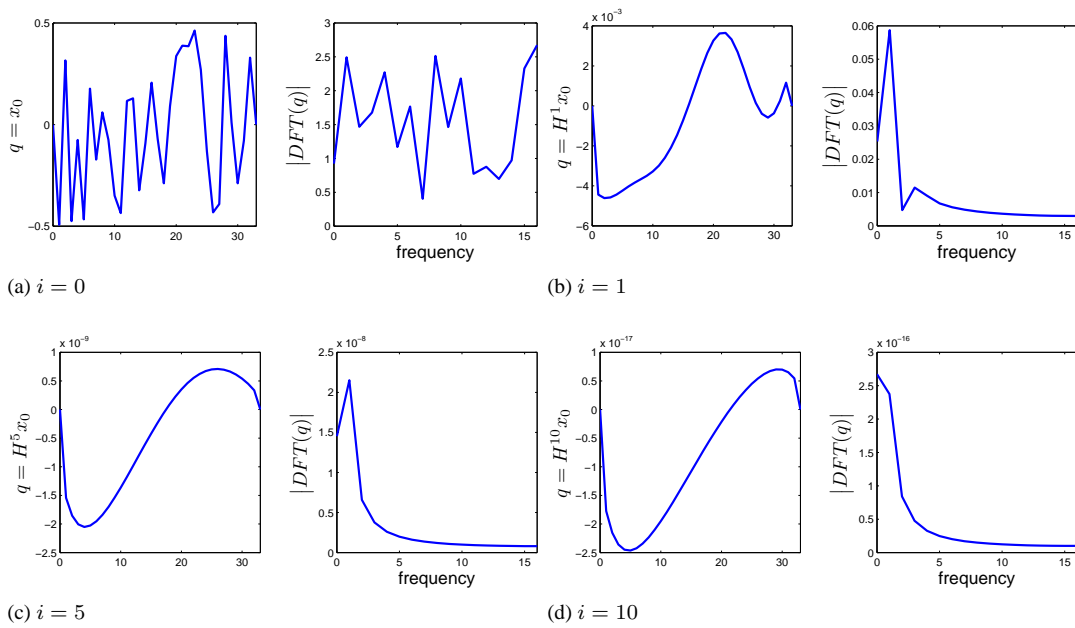


FIG. 3.20 – Vecteurs de Krylov pour la paramétrisation de Bernstein -  $n = 32$

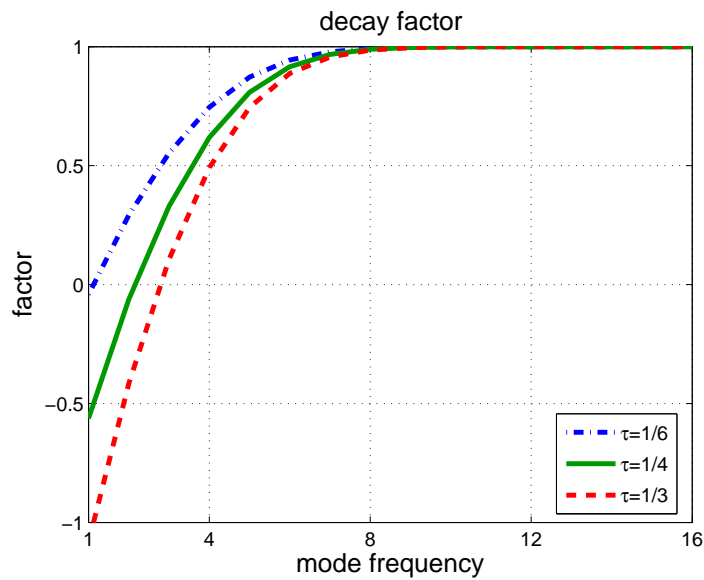


FIG. 3.21 – Coefficients d'amortissement pour le problème d'optimisation dans la base de Bernstein -  $n = 16$ .

### 3.5.2 Paramétrisation B-splines

À notre connaissance on ne connaît pas de forme analytique pour l'intégrale du produit de deux fonctions B-splines de degré  $d$ . Cependant, comme ce produit de fonctions est un polynôme de degré  $2d$ , pour calculer un hessien numérique exact, on peut utiliser la quadrature de Gauss avec  $2d + 1$  points.

Comme pour la paramétrisation de Bézier, les contraintes sur les paramètres de contrôle s'écrivent

$$x_0 = 0 \quad \text{et} \quad x_n = 0 \quad (3.95)$$

On suppose toujours que le problème a été mis à l'échelle (c-à-d, préconditionné par l'inverse de la diagonale). Les facteurs d'amplification sont obtenus par le calcul des paires de vecteurs et valeurs propres de  $D_H^{-1}H$  où  $D_H$  est la diagonale de  $H$ , le hessien projeté. Les Figures 3.22 et 3.23 illustrent les vecteurs propres et valeurs propres respectivement. Les résultats s'apparentent à ceux de la représentation de Bernstein.

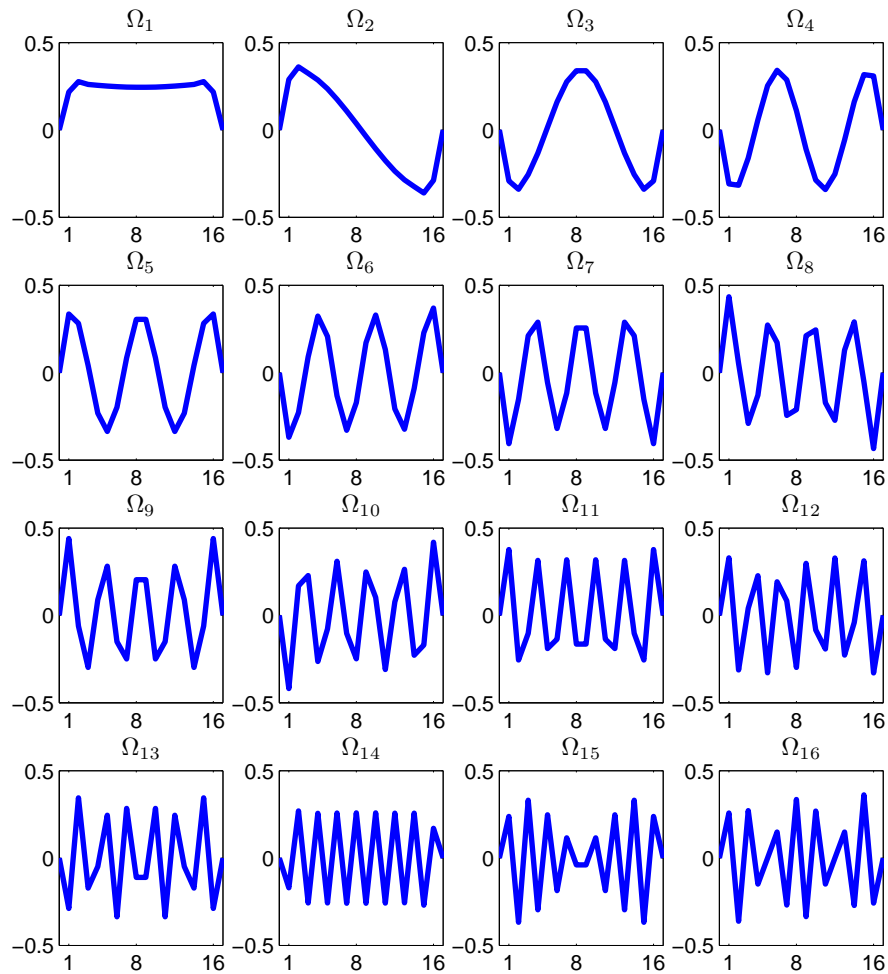


FIG. 3.22 – B-spline - vecteurs propres -  $n = 16$

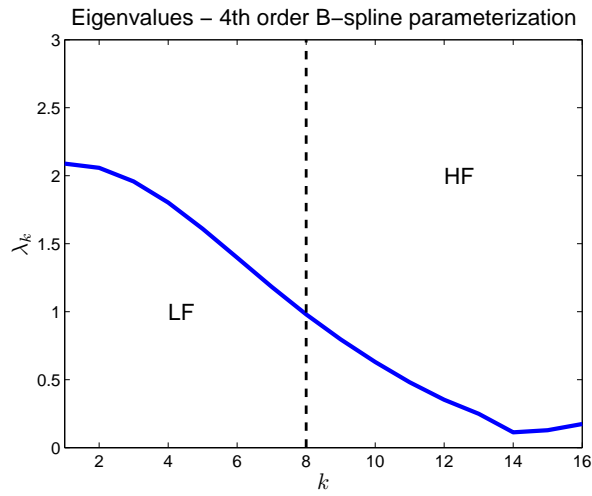


FIG. 3.23 – B-spline - valeurs propres -  $n = 16$

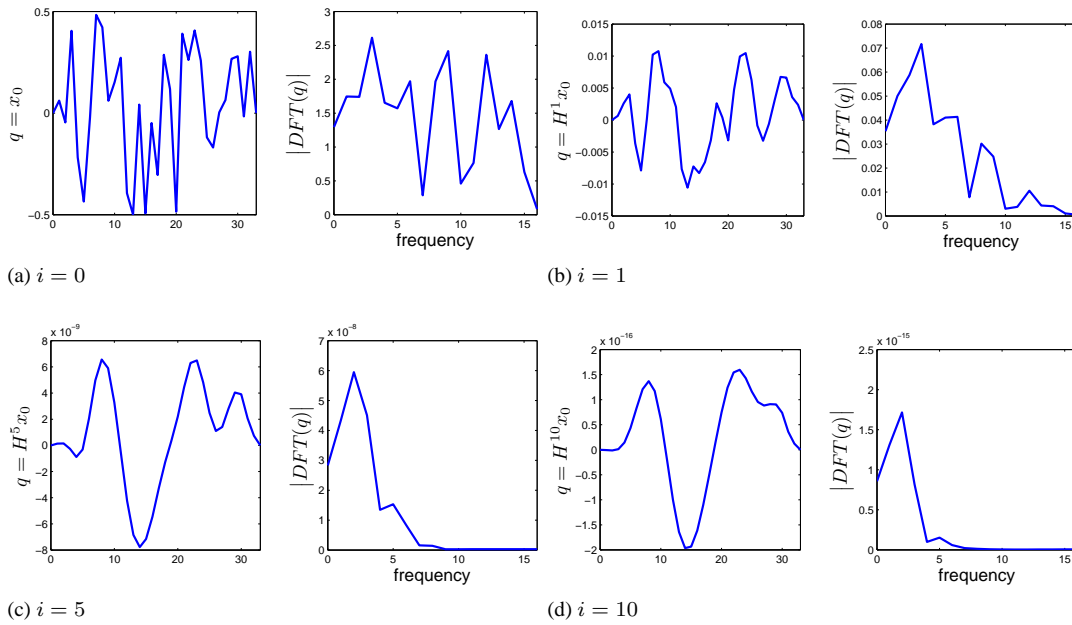


FIG. 3.24 – Vecteurs de Krylov pour la paramétrisation de B-spline -  $n = 32$

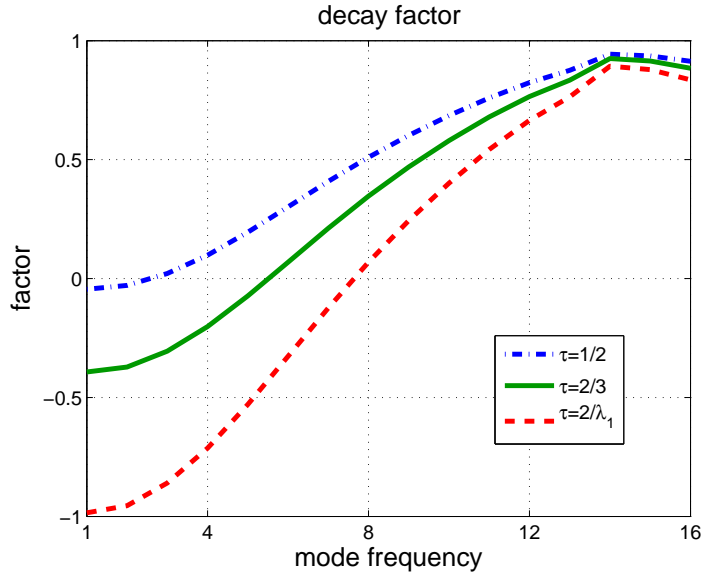


FIG. 3.25 – Coefficients d'amortissement pour le problème d'optimisation dans la base de B-spline -  $n = 16$ .

### 3.5.3 Polynômes orthogonaux

On rappelle un résultat de meilleure approximation d'une fonction  $\tilde{u}$  de  $H_0$  dans l'espace des polynômes de degré  $n$  sur  $[0, 1]$  noté  $\mathcal{P}^n$ . Tout d'abord on note que  $\mathcal{P}^n$  est bien un sous-espace de  $H_0$  puisque  $[0, 1]$  est borné et fermé (et dont les fonctions sont donc sommables).

Une base orthogonale de  $\mathcal{P}^n$  est donnée par les polynômes de Legendre  $P_k$  pour le produit scalaire (3.1). On a

$$\int_{-1}^1 P_k(t)P_j(t)dt = a_k\delta_{kj} \quad (3.96)$$

où  $\delta_{kj}$  est le symbole de Kronecker et  $a_k = \frac{2}{2k+1}$ .

Le problème de meilleure approximation devient trivial puisque la matrice hessienne à inverser  $H$  est diagonale avec  $h_{kk} = a_k$ . Les coefficients du second membre  $b_k = (\tilde{u}, p_k)$  sont la projection orthogonale de  $\tilde{u}$  sur la base  $P_k$  et l'inversion correspond à la normalisation par  $a_k = \|P_k\|^2$  de chacun des coefficients de la projection.

Le spectre et le conditionnement de ce système est évident puisque qu'il est diagonal. Les vecteurs propres (en terme de paramètres de conception) sont les vecteurs de la base canonique de  $\mathbb{R}^{n+1}$  (constituant l'identité), les fonctions propres sont les fonctions  $p_k$ , et les valeurs propres sont les éléments diagonaux  $a_k$ . Le conditionnement de ce système est alors  $\kappa_2(H) = \frac{a_0}{a_n} = 2n + 1$ .

Les conditions au bord  $u(0) = u(1) = 0$  s'écrivent

$$\sum_{k=0}^n x_k = 0 \quad \text{et} \quad \sum_{k=0}^n (-1)^k x_k = 0. \quad (3.97)$$

### 3.5.4 Influence de la dimension sur la raideur

L'inverse du nombre de conditionnement donne une distance à la matrice non inversible la plus proche [28]. En ce sens, plus le nombre de conditionnement est élevé, plus la matrice s'apparente à une matrice non inversible, ce qui conduit à un problème raide. En arithmétique flottante, si cette distance est de l'ordre de la précision machine, la matrice est numériquement non inversible.

On représente à la Figure 3.26 une estimation du nombre de conditionnement des matrices hessiennes

pour plusieurs paramétrisations (Bernstein, B-spline, Legendre, Tchebychev) en fonction du nombre de ddl.

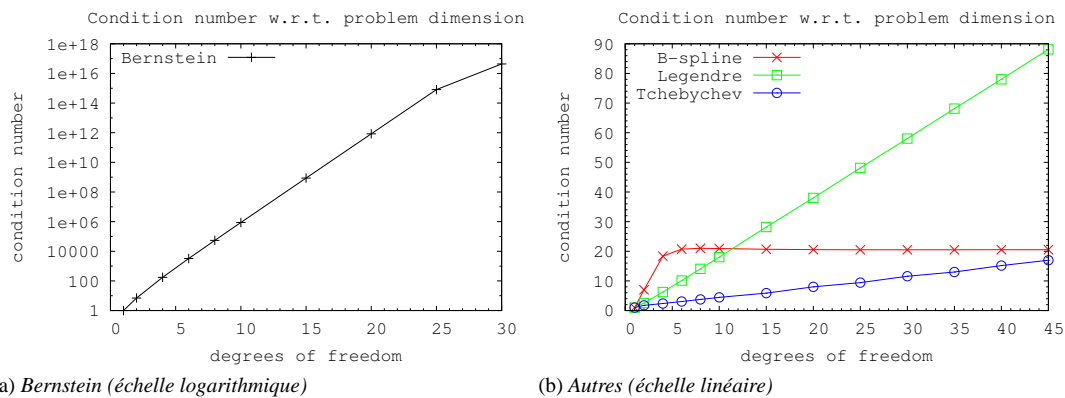


FIG. 3.26 – Comparaison du conditionnement de la matrice hessienne du problème géométrique pour différentes paramétrisations.

### 3.5.5 Expériences numériques

On cherche à approcher le graphe de la fonction cible décrite à la Figure 3.27 par une courbe de Bézier de degré  $n$ . La fonction cible est une somme de fonctions gaussiennes (exponentielles), elle n'appartient donc pas à un espace polynômial. La valeur optimale  $J^*$  de la fonction objectif est calculée à partir de la forme obtenue par projection orthogonale de la fonction cible dans l'espace des polynômes de degré  $n$  (dans la base des polynômes de Legendre). On est donc en mesure de représenter la convergence des méthodes numériques en terme d'erreur sur la fonction objectif  $e = J - J^*$ .

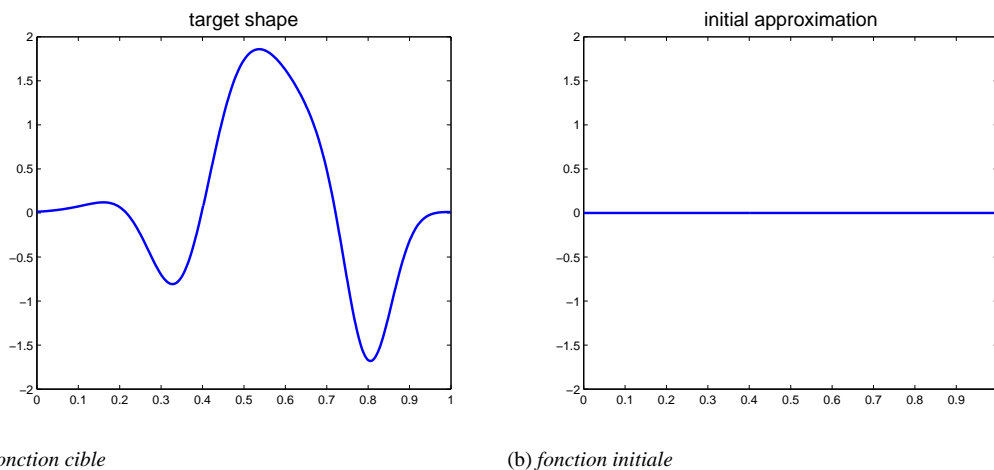


FIG. 3.27 – Description du cas-test de meilleure approximation de forme par une courbe de Bézier. La fonction cible est une somme de gaussiennes. L'approximation initiale est nulle.

La paramétrisation fine est de degré  $n = 24$ . On réalise dans un premier temps une optimisation numérique classique sur le seul niveau fin. Les résultats sont illustrés à la Figure 3.28. Ce problème est effectivement très raide : la convergence est linéaire dont le taux de convergence est proche de 1.

Les stratégies multiniveaux sont testées avec deux types de cycles : un cycle en « dent de scie » et un cycle en V. Par ailleurs, on se place désormais dans le contexte d'un problème d'optimisation, on ne supposera pas que le problème a été préconditionné par l'inverse de la diagonale du hessien. Une



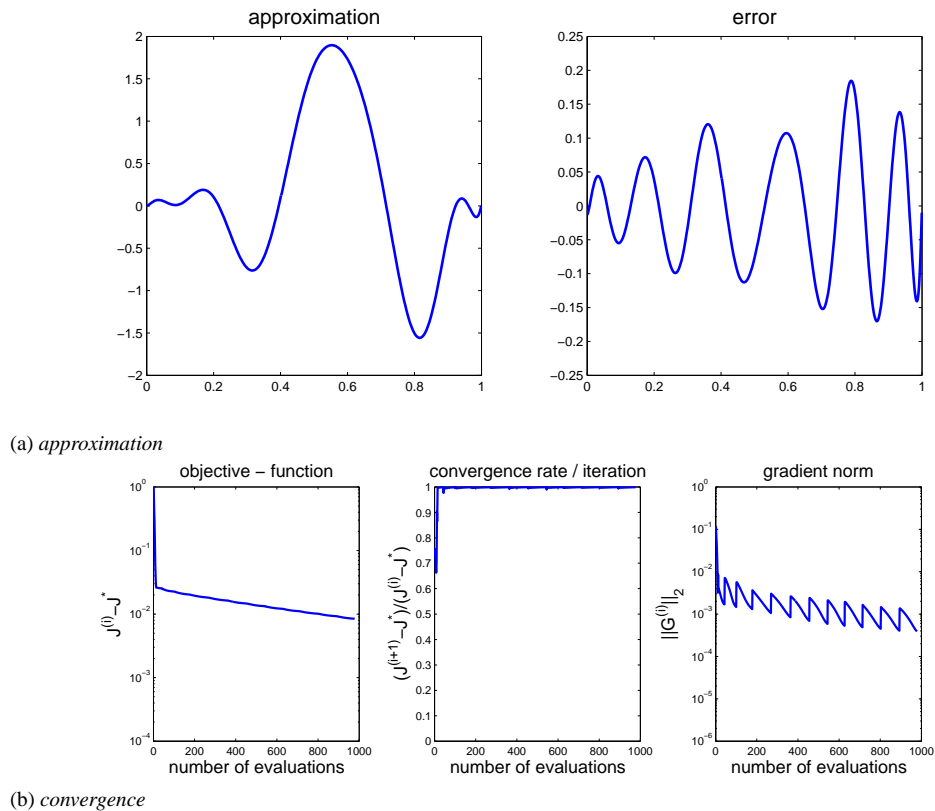


FIG. 3.28 – Méthode du Gradient Conjugué de Polak-Ribière sur le niveau fin.

approximation du pas optimal est obtenue par une procédure de recherche linéaire à chaque itération. On fait en sorte que le nombre d'évaluations de la fonction objectif soit environ égal à 1000.

Les paramètres des schémas multiniveaux sont :

- degrés des paramétrisations :  $n = 3, 6, 9, 12, 15, 18, 21, 24$
- pré-relaxation :  $\nu_1 = 2$  itérations de Gradient Conjugué (Polak-Ribière)
- post-relaxation :  $\nu_2 = 2$  itérations de Gradient Conjugué (Polak-Ribière) (V-cycle uniquement)

### Cycle en « dent de scie »

Les résultats sont présentés sur les Figures 3.29, 3.30 et 3.31 pour les transferts : a) classique, b) avec permutation du spectre, c) algébrique. On constate que la stratégie classique est inefficace : l'erreur reste du même ordre de grandeur qu'avec la stratégie à un seul niveau et le taux de convergence tend vers 1 également. En revanche les autres stratégies semblent plus efficaces : la fonction coût descend plus bas avec un taux de convergence linéaire plus petit ( $\approx 0.70$ ). L'erreur uniforme est environ 25 fois plus petite. La méthode algébrique ( $J - J^* \approx 10^{-8}$ ) semble cependant meilleure que la méthode de réorganisation du spectre ( $J - J^* \approx 10^{-6}$ ). Pour cette dernière le taux de convergence tend vers 0.9 sur les derniers cycles. On résume à la Figure 3.32 la convergence de l'ensemble des méthodes en terme d'évolution de la fonction objectif.

### Cycle en V

Les résultats sont présentés sur les Figures 3.33, 3.34 et 3.35 pour les transferts : a) classique, b) avec permutation du spectre, c) algébrique. On constate les mêmes résultats qu'avec les cycles en dent de scie. On résume à la Figure 3.36 la convergence de l'ensemble des méthodes en terme d'évolution de la fonction objectif.

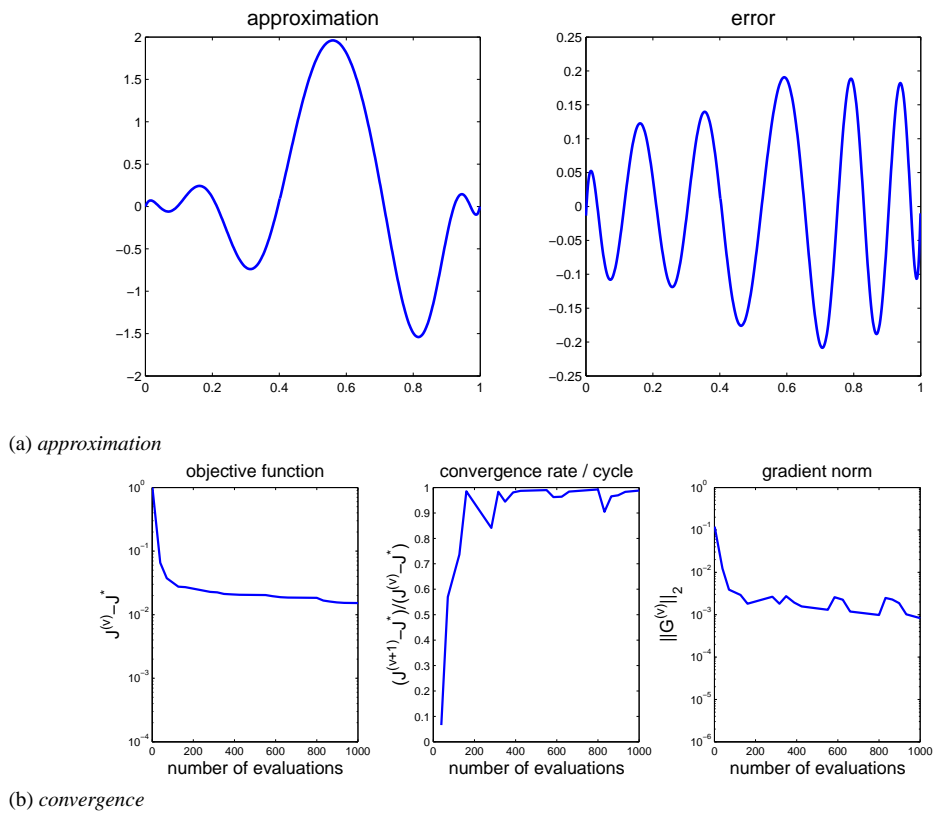


FIG. 3.29 – Transfert classique.

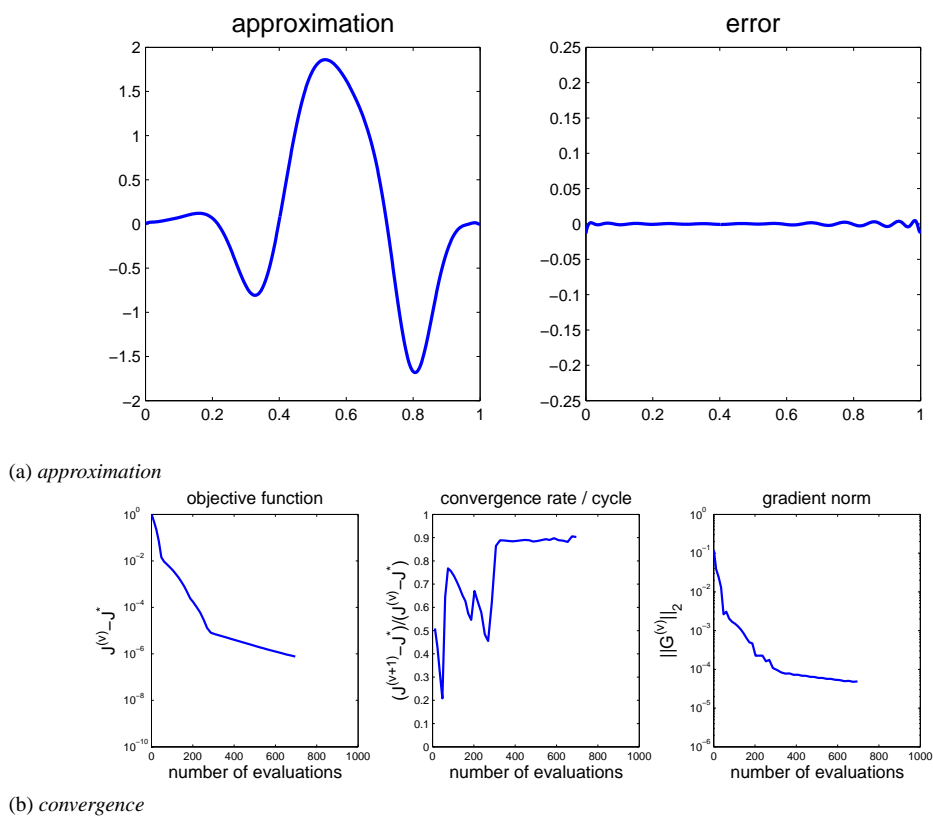


FIG. 3.30 – Transfert avec permutation du spectre.

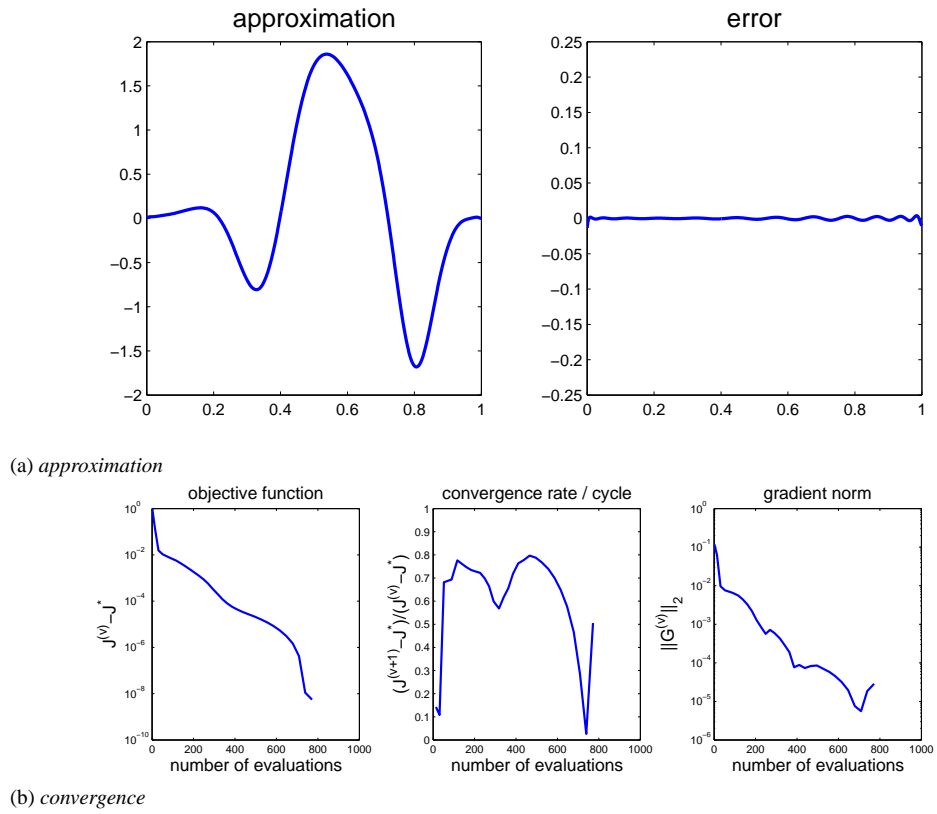


FIG. 3.31 – Transfert algébrique.

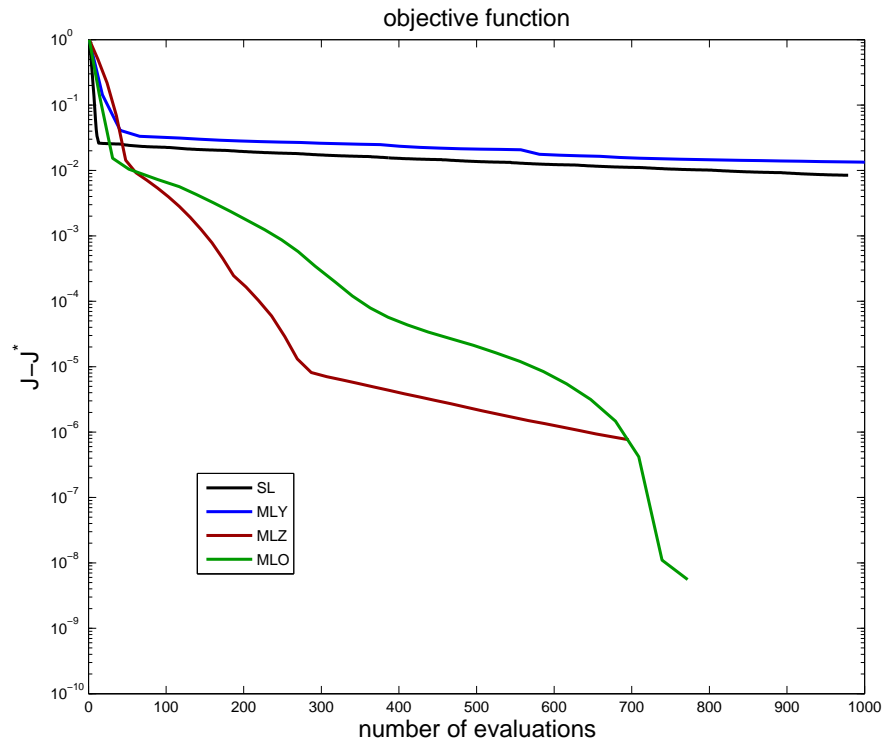
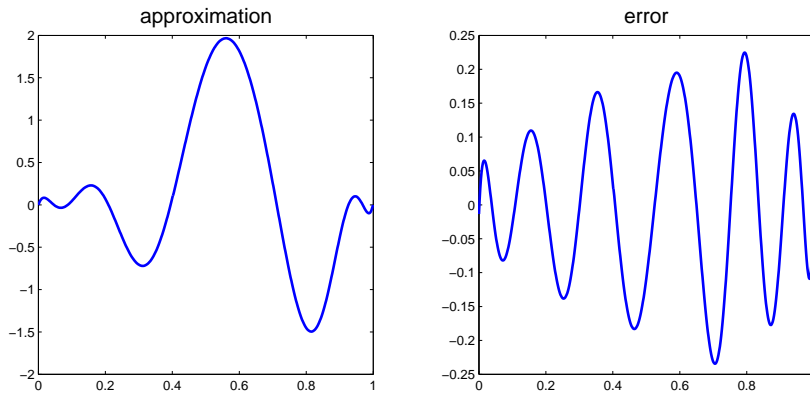
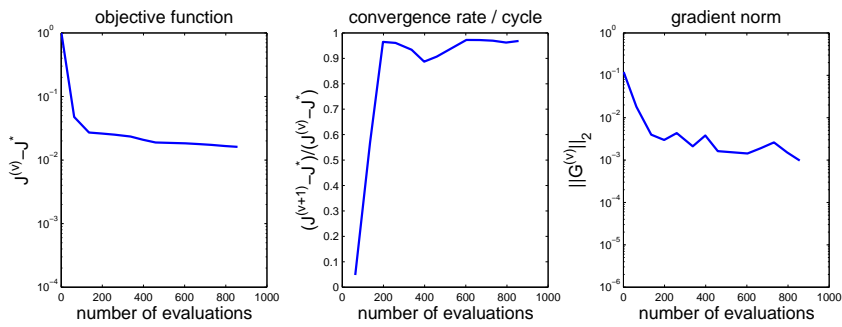


FIG. 3.32 – Convergence de la fonction objectif pour l'ensemble des stratégies multiniveaux. Cycle en dent de scie.

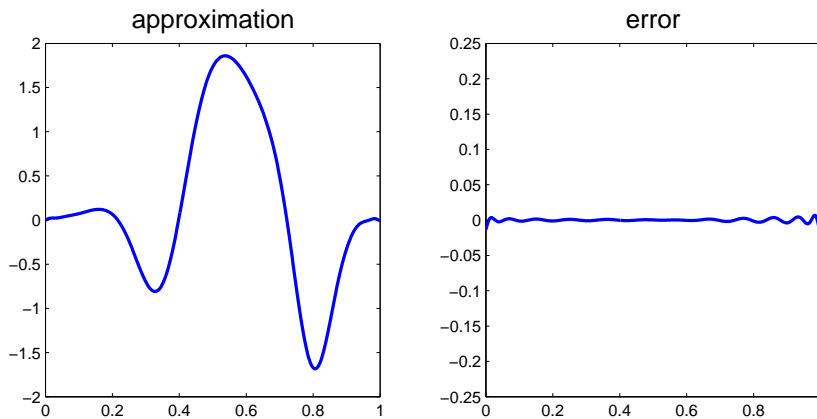


(a) approximation

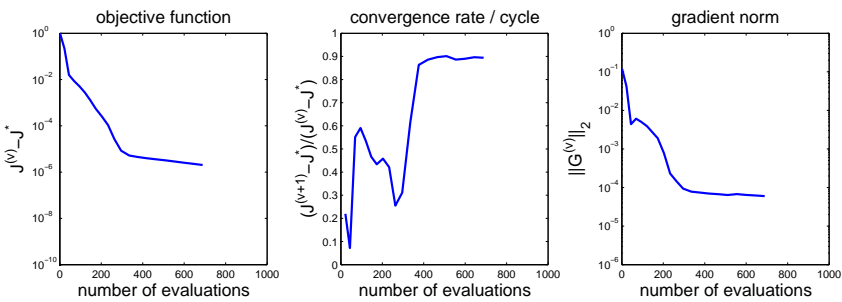


(b) convergence

FIG. 3.33 – Transfert classique.



(a) approximation



(b) convergence

FIG. 3.34 – Transfert avec permutation du spectre.

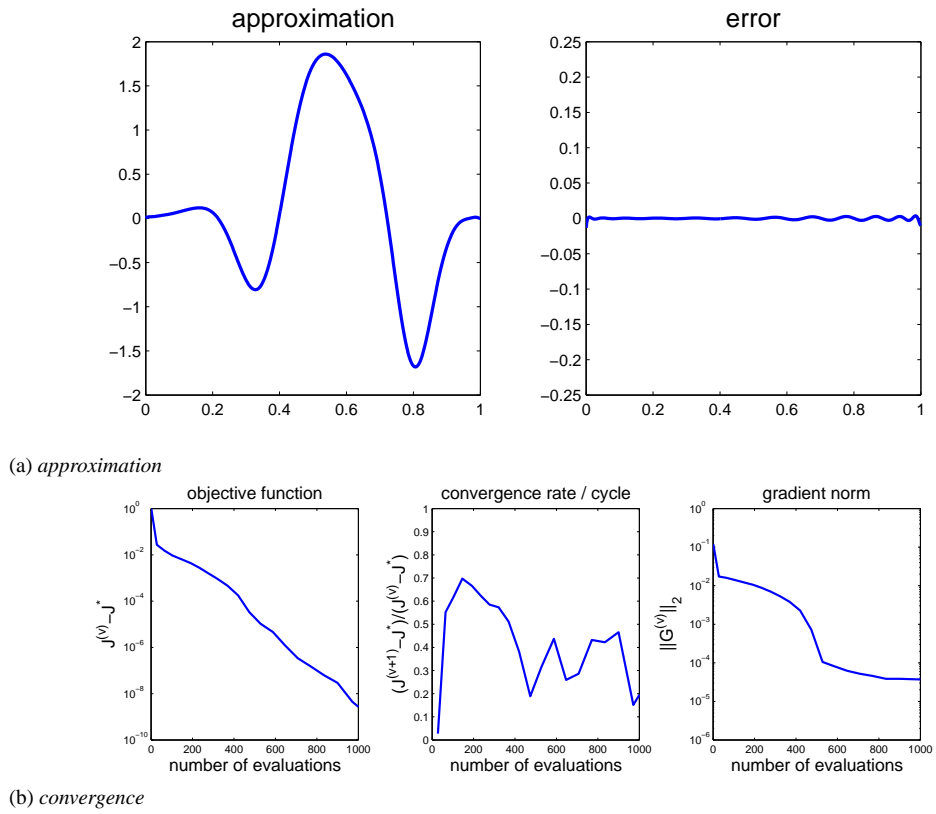


FIG. 3.35 – Transfert algébrique.

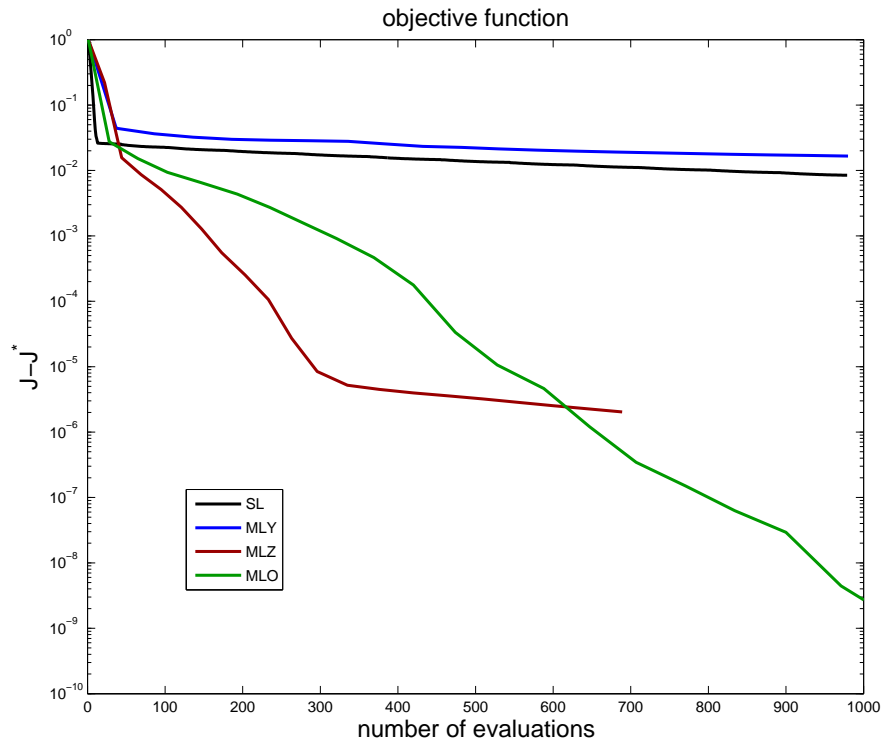


FIG. 3.36 – Convergence de la fonction objectif pour l'ensemble des stratégies multiniveaux. Cycle en V.

## Conclusion

Dans ce chapitre on a formalisé un ensemble de stratégies multiniveaux pour l'optimisation de forme dans le cadre modèle d'une fonction objectif quadratique. Dans la mesure du possible on s'est efforcé de rattacher les concepts liés à l'optimisation aux concepts liés aux méthodes multigrilles linéaires. Cela nous a conduit à la formulation d'un algorithme bigrille idéal dont on a explicité la matrice d'amplification équivalente.

D'après l'analyse de Fourier des applications linéaires correspondant aux méthodes itératives de base (gradient, Jacobi), les opérateurs classiques de transfert ont été revus. Il en découle plusieurs stratégies hiérarchiques. Pour chaque stratégie, on a étudié le rayon spectral de la matrice d'amplification afin de montrer la convergence d'un cycle idéal. Ce faisant, on a mis en évidence les stratégies susceptibles d'améliorer la vitesse de convergence. En particulier, les stratégies efficaces sont celles dont les niveaux « grossiers » correspondent à des sous-espaces de haute fréquence pour lesquels l'erreur itérative est la plus grande. Les résultats théoriques ont été confirmés par des expériences numériques.



# MÉTHODES HIÉRARCHIQUES POUR L'OPTIMISATION PARAMÉTRIQUE D'UNE ANTENNE

---

<b>4.1 Étude du hessien paramétrique pour un problème inverse</b>	<b>105</b>
4.1.1 Modèle approché d' <i>Optique Physique</i>	105
4.1.2 Modèle électromagnétique	108
4.1.3 Conclusion sur l'étude des hessiens paramétriques	111
<b>4.2 Hiérarchisation des paramètres de conception</b>	<b>111</b>
4.2.1 Prolongement classique : élévation du degré	111
4.2.2 Prolongement basé sur le hessien paramétrique	111
4.2.3 Approximation par le modèle géométrique : filtre passe-haut	112
<b>4.3 Algorithmes hiérarchiques en optimisation globale</b>	<b>112</b>
4.3.1 Extension de l'algorithme PSO	112
4.3.2 Expériences numériques de validation en <i>Optique Physique</i>	114
4.3.3 Application à l'optimisation d'un diagramme en directivité d'une antenne	117
<b>4.4 Algorithmes hiérarchiques en optimisation locale</b>	<b>119</b>
4.4.1 Algorithme multiniveau non-linéaire	119
4.4.2 Expériences numériques	121
<b>Conclusion</b>	<b>132</b>

---

Ce chapitre est dédié à l'expérimentation numérique d'algorithmes hiérarchiques pour la conception optimale d'antenne. On distingue les stratégies globales (robustes) des stratégies locales (descente). Dans un premier temps on conduit une étude spectrale du hessien paramétrique pour mettre en évidence le comportement de la convergence itérative au voisinage d'un minimum.





## 4.1 Étude du hessien paramétrique pour un problème inverse

On considère le problème de minimisation

$$\min_{\mathcal{S}} \mathcal{J}(\mathcal{S}) = \frac{1}{2} \int_{S^2} \left( \|u(\mathcal{S}, f)\|^2 - \|u(\mathcal{S}^t, f)\|^2 \right)^2 dS^2. \quad (4.1)$$

où  $\mathcal{S}^t$  est une forme connue qui appartient à l'espace admissible de la déformation d'une forme donnée  $\mathcal{S}^0$ . En particulier, en choisissant  $\mathcal{S}^0 = \mathcal{S}^t$ , une (la) solution est donnée par le vecteur de conception *Free-Form*  $x = \mathbf{0}$ .

Cet exercice nous permet d'étudier le hessien en un point de conception qui est par construction une solution de (4.1). Pour une paramétrisation donnée, on vérifie que les conditions suffisantes d'optimalité sont satisfaites, c-à-d que le hessien paramétrique est strictement défini positif (le minimum est alors un minimum strict). Soumis à des contraintes d'égalité, dans un voisinage du minimum, les variations sont dominées par les termes du second ordre dans les directions admissibles. Ces termes sont donnés par le hessien projeté (voir chapitre 2) dont on étudie le spectre par analogie avec le problème géométrique du chapitre 3.

### 4.1.1 Modèle approché d'Optique Physique

Soit la géométrie cible  $\mathcal{S}^t$  représentée à la Figure 4.1 dont le rayonnement est calculé par le modèle approché d'Optique Physique.

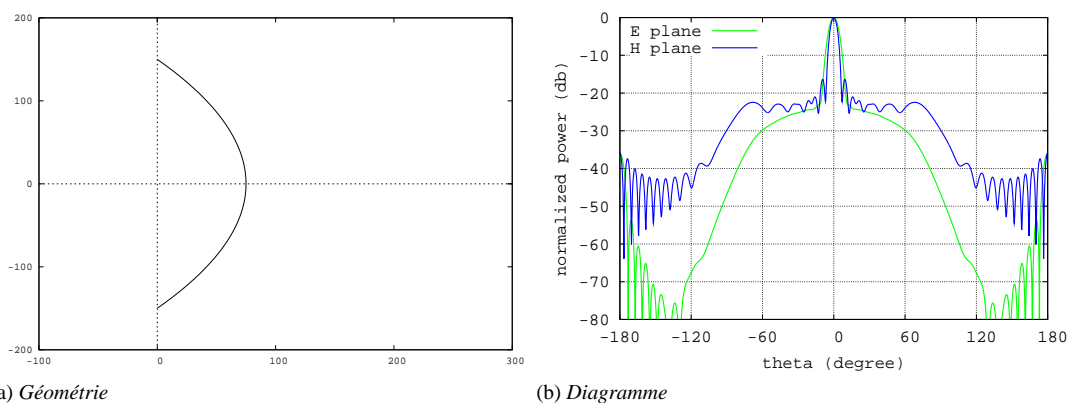


FIG. 4.1 – Configuration de l'antenne cible : paraboloïde alimenté par un dipôle situé à la distance focale en (0, 0).

Le champ incident est donné par un dipôle. Le champ EM diffracté, et donc le champ total rayonné, est une fonction explicite de la géométrie. Étant donnée une surface métallique dont la méridienne est discrétisée en segments, on connaît les dérivées partielles exactes du champ par rapport à la position des nœuds des segments. On dispose également du hessien numérique exact obtenu par différenciation automatique du code de calcul des dérivées [19]. On en déduit le gradient et le hessien de la fonctionnelle par rapport aux variables de conception paramétriques (voir section 1.4.1).

La décomposition spectrale est calculée numériquement<sup>1</sup> pour des espaces de recherche de dimension 6 et 10 avec conditions aux limites de Dirichlet (soit des espaces admissibles de dimension 4 et 8 resp.) dans les paramétrisations de Bézier-Bernstein et B-spline.

<sup>1</sup>Le calcul des spectres est réalisé à l'aide de la librairie d'algèbre linéaire LAPACK disponible à l'adresse [www.netlib.org/lapack](http://www.netlib.org/lapack).

### Paramétrisation de Bézier-Bernstein

Dans la paramétrisation de Bernstein on représente les vecteurs propres à la Figure 4.2 et les valeurs propres à la Figure 4.3. Les modes propres présentent une structure spectrale similaire aux modes de Fourier discrets et sont proches des modes propres du problème géométrique (voir la Figure 3.18); les modes de basse fréquence sont associés aux valeurs propres les plus grandes et ceux de haute fréquence aux valeurs propres les plus petites.

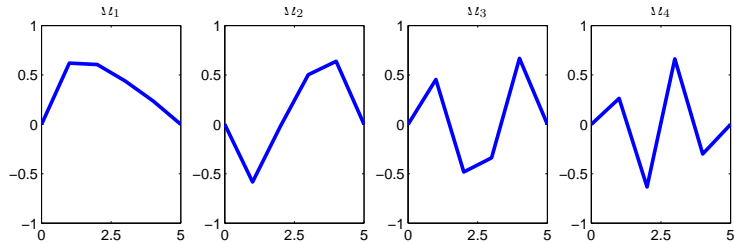
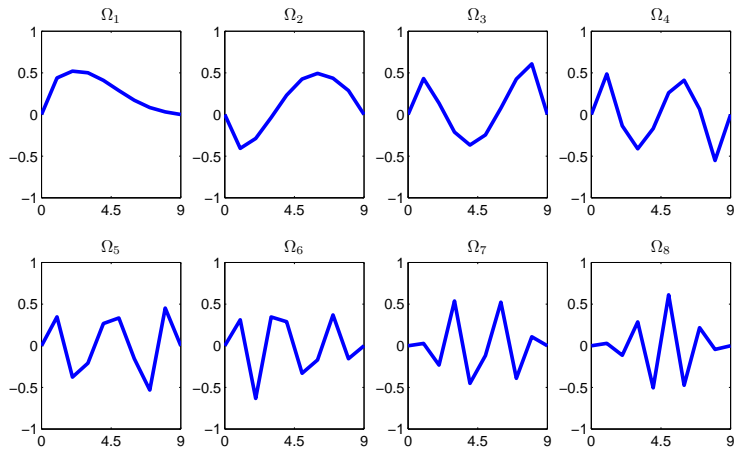
(a)  $n = 4$ (b)  $n = 8$ 

FIG. 4.2 – Vecteurs propres du hessien paramétrique (Bézier-Bernstein) de la fonctionnelle (4.1) au point  $x = \mathbf{0}$  tels que  $x_0 = x_{n+1} = 0$  pour les degrés  $n = 4$  (a) et  $n = 8$  (b). Les modes propres pour la paramétrisation de degré  $n = 4$  correspondent aux modes de basse fréquence de la paramétrisation de degré  $n = 8$ .

### Paramétrisation B-spline

Dans la paramétrisation B-spline d'ordre 4 on représente les vecteurs propres à la Figure 4.4 et les valeurs propres à la Figure 4.5. Comme pour la base de Bernstein, les modes propres sont proches des modes du problème géométrique (voir Figure 3.22) et ont donc une structure spectrale similaire aux modes de Fourier discrets. Les modes de basse fréquence sont associés aux valeurs propres les plus grandes et inversement ceux de haute fréquence aux valeurs propres les plus petites.

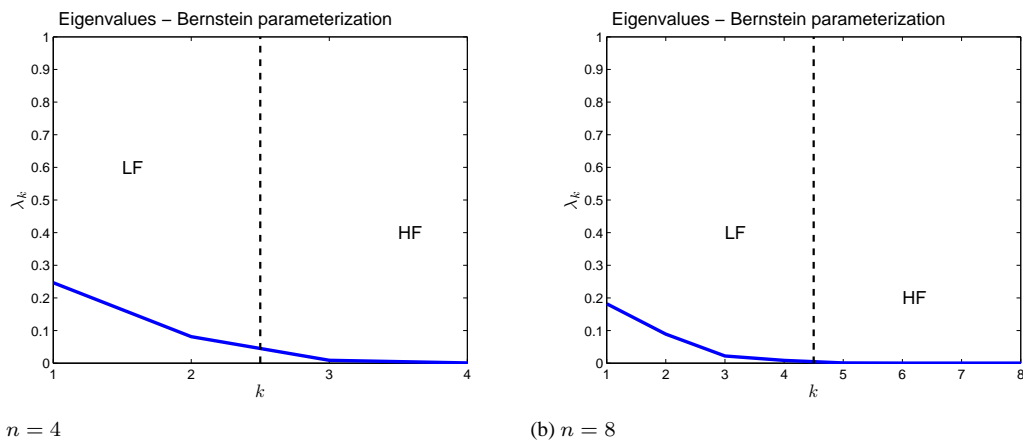


FIG. 4.3 – Valeurs propres du hessien paramétrique (Bézier-Bernstein) de la fonctionnelle (4.1) au point  $x = \mathbf{0}$ ; les valeurs propres sont toutes positives (le minimum est bien un minimum strict) et monotones décroissantes en fonction du mode  $k$ ; on distingue le spectre en deux sous-ensembles correspondant aux modes LF et HF.

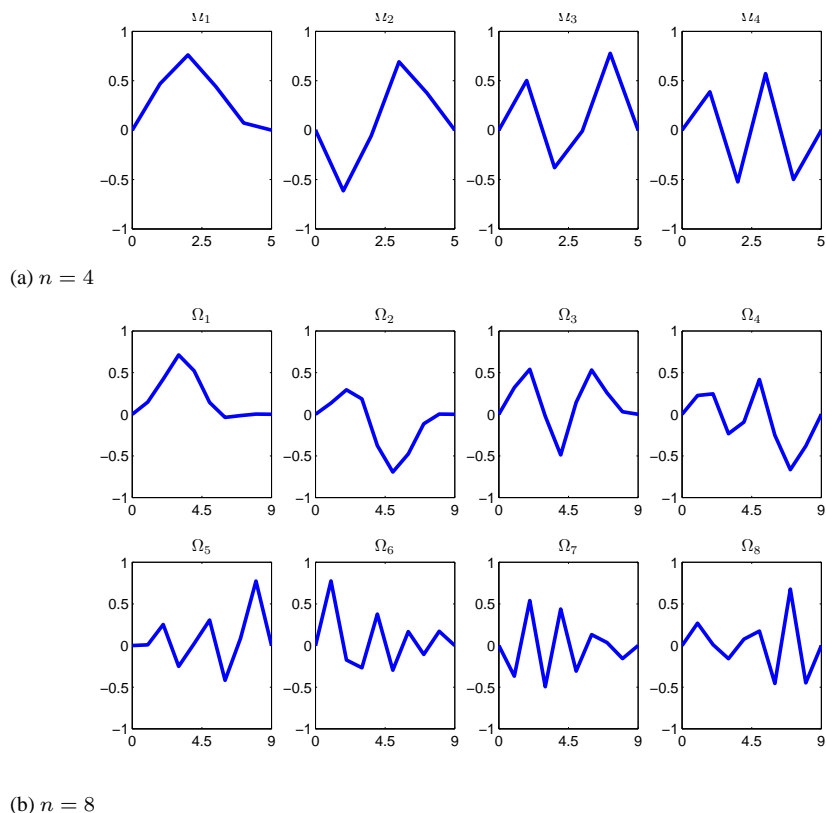


FIG. 4.4 – Vecteurs propres du hessien paramétrique (B-splines d'ordre 4) de la fonctionnelle (4.1) au point  $x = \mathbf{0}$ .

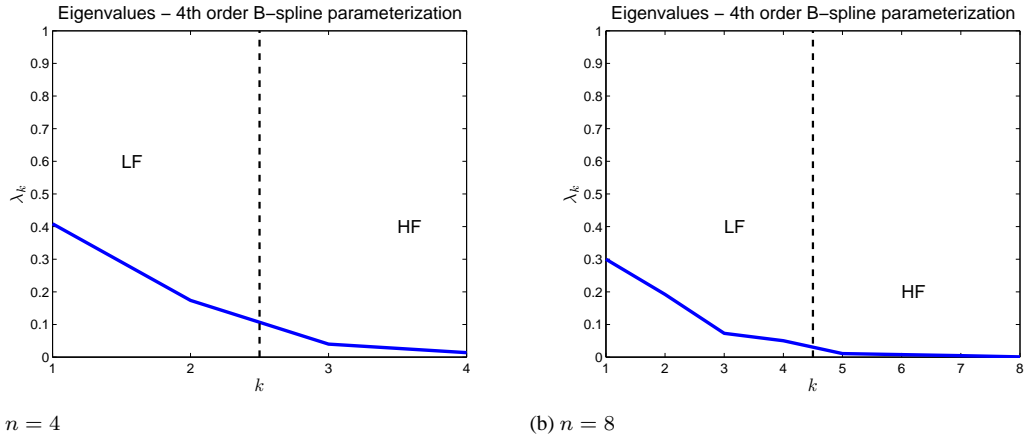


FIG. 4.5 – Valeurs propres du hessien paramétrique (B-splines d'ordre 4) de la fonctionnelle (4.1) au point  $x = 0$ .

### 4.1.2 Modèle électromagnétique

Soit la géométrie cible  $S^t$  représentée à la Figure 4.6 dont le rayonnement est donné par le système de Helmholtz et résolu par SRSR.

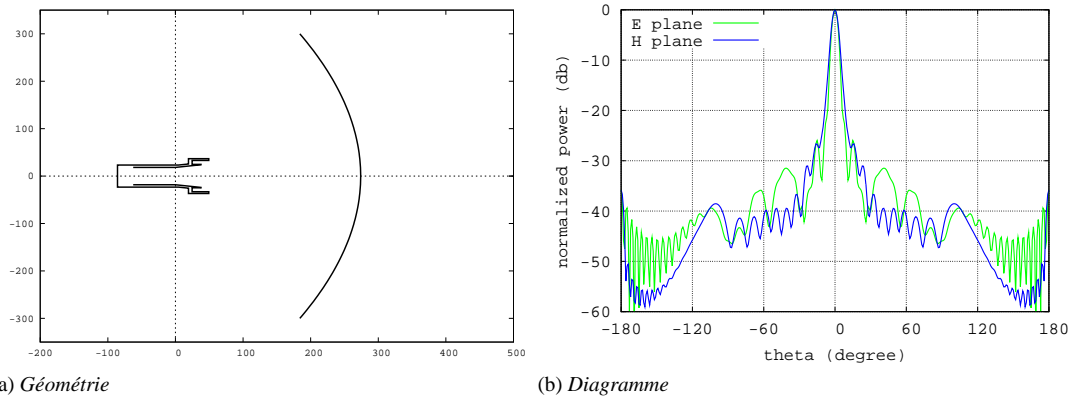


FIG. 4.6 – Configuration de l'antenne cible : réflecteur conformé à un diagramme et alimenté par un guide d'onde.

Comme nous n'avons pas de formule exacte (discrète ou continue) du hessien, on en calcule une approximation à l'aide du schéma centré de différences finies (4.2). Le pas  $h$  est choisi expérimentalement pour assurer la stabilité du calcul.

$$\frac{\partial J}{\partial x_j \partial x_k}(x) \approx \frac{J(x + h(\mathbf{e}_j + \mathbf{e}_k)) + J(x - h(\mathbf{e}_j + \mathbf{e}_k)) - J(x + h(\mathbf{e}_j - \mathbf{e}_k)) - J(x + h(-\mathbf{e}_j + \mathbf{e}_k))}{4h^2} \quad (4.2)$$

Comme pour le modèle d'Optique Physique, on conduit une analyse spectrale numérique dans les deux bases de paramétrisation : Bézier-Bernstein et B-spline d'ordre 4.

**Remarque** Les résultats présentés dans les sections qui suivent sont doublement soumis à des erreurs numériques, d'une part dues au calcul du hessien par différences finies et d'autre part au calcul du spectre. Cela peut expliquer quelques anomalies dans la structure des vecteurs propres (nombre d'oscillations, etc.). À ces variations près, ces résultats semblent globalement pertinents.

### Paramétrisation de Bézier-Bernstein

On représente les vecteurs propres à la Figure 4.7 et les valeurs propres à la Figure 4.8. Hormis le vecteur propre unimodal qui est absent, la structure spectrale semble conservée : lorsque les valeurs propres sont triées par ordre décroissant, la fréquence des modes associés augmente.

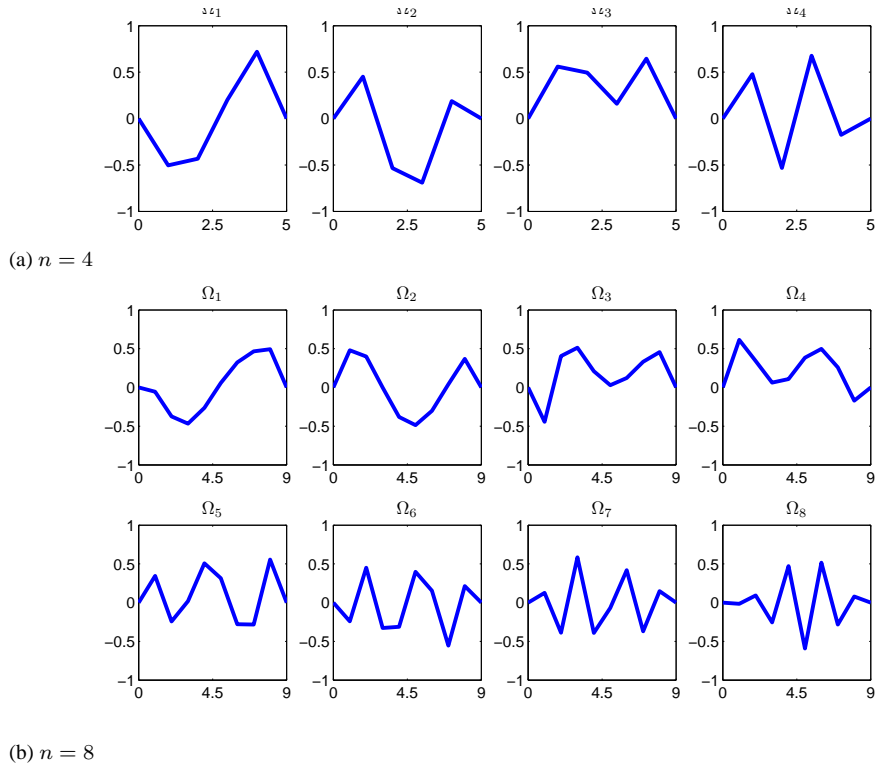


FIG. 4.7 – Vecteurs propres du hessien paramétrique approché (Bézier-Bernstein) de la fonctionnelle (4.1) au point  $x = \mathbf{0}$ .

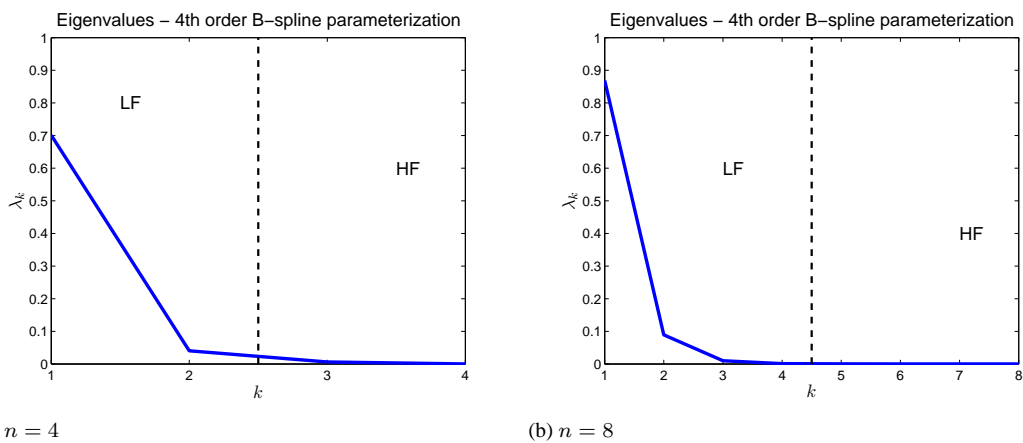


FIG. 4.8 – Valeurs propres du hessien paramétrique approché (Bézier-Bernstein) de la fonctionnelle (4.1) au point  $x = \mathbf{0}$ .

**Paramétrisation B-splines**

On représente les vecteurs propres à la Figure 4.9 et les valeurs propres à la Figure 4.10. Comme pour la paramétrisation de Bernstein, le premier mode est absent mais le reste de la structure est conservée.

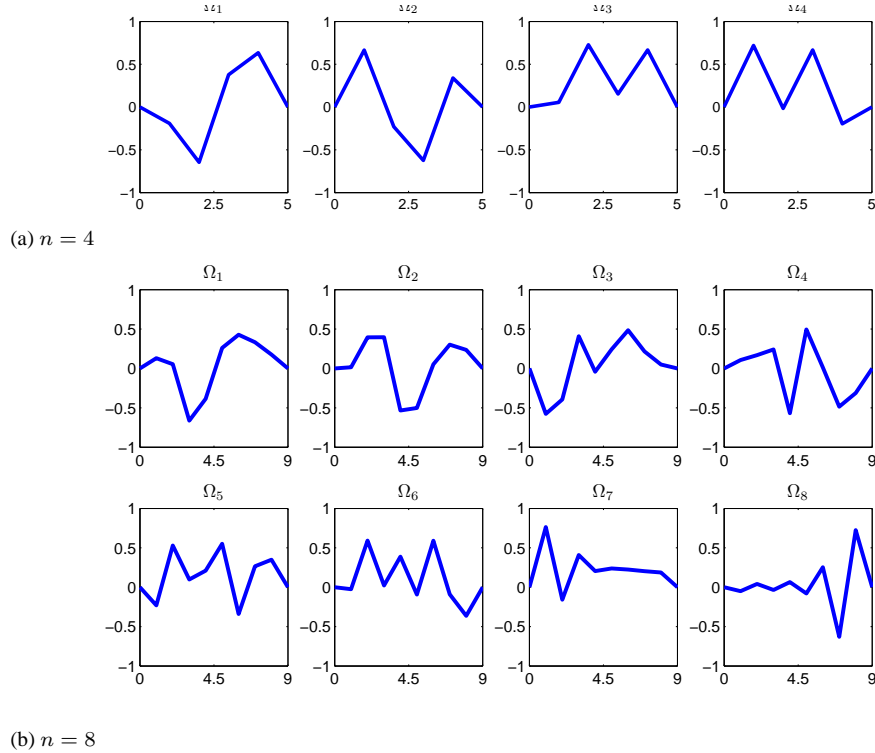


FIG. 4.9 – Vecteurs propres du hessien paramétrique approché (B-spline) de la fonctionnelle (4.1) au point  $x = \mathbf{0}$ .

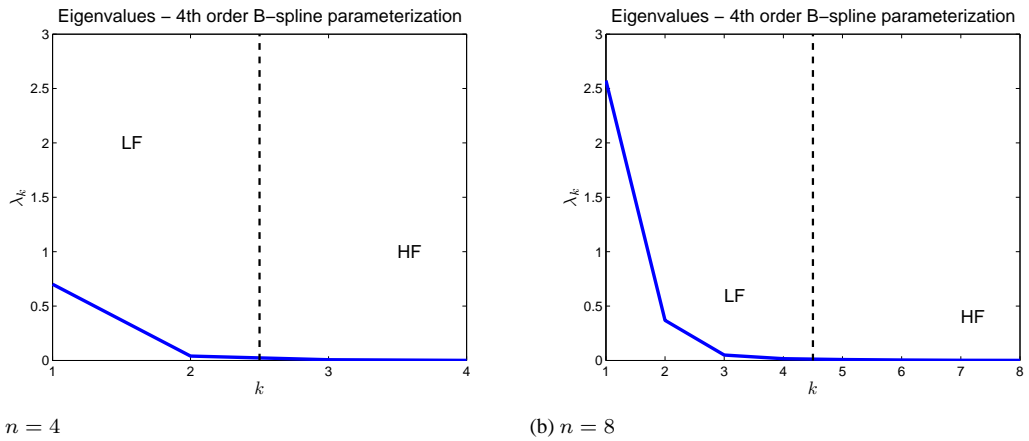


FIG. 4.10 – Valeurs propres du hessien paramétrique approché (B-spline) de la fonctionnelle (4.1) au point  $x = \mathbf{0}$ .

### 4.1.3 Conclusion sur l'étude des hessiens paramétriques

D'après l'étude spectrale réalisée dans cette section, on remarque que le hessien du problème physique ressemble au hessien du problème géométrique du chapitre précédent : d'une part on peut associer à chaque vecteur propre un paramètre de fréquence lié au nombre d'alternances de signe de ses composantes, d'autre part, une fois cette structure établie, on remarque que les valeurs propres sont une fonction décroissante de ce paramètre.

Dans la mesure où le hessien représente un modèle quadratique local de la fonctionnelle au voisinage de la solution, on s'attend à ce que les méthodes de descente appliquées à des problèmes de conception optimale d'antenne se comportent comme pour le problème géométrique, en particulier si le problème devient de plus en plus raide avec la dimension du problème.

## 4.2 Hiérarchisation des paramètres de conception

Désormais on s'attache à formuler et expérimenter des stratégies multiniveaux appliquées à l'optimisation de la géométrie d'antennes à réflecteurs. Pour cela on a vu que l'on disposait de plusieurs hiérarchies de paramétrisation. Selon le contexte, qu'il s'agisse d'une optimisation globale ou locale, l'utilisation d'une hiérarchie plutôt qu'une autre s'avèrera plus adéquate. On revient sur l'ensemble des stratégies dont nous disposons.

### 4.2.1 Prolongement classique : élévation du degré

Cette stratégie correspond à la généralisation de la technique d'*élévation du degré*. Typiquement, les espaces de recherche sont des espaces polynomiaux dont on utilise la propriété d'inclusion (tout polynôme de degré  $m$  est également un polynôme de degré  $n$ )

$$\forall m < n, \quad \mathcal{P}^m \subset \mathcal{P}^n \quad (4.3)$$

pour établir une hiérarchie de paramétrisations. Plus le degré est élevé, plus la représentation de la méridienne est fine et plus l'espace de recherche est grand.

Considérons deux espaces polynomiaux de degré  $m$  et  $n$  tels que  $m < n$  (paramétrisation grossière et fine respectivement). Chaque espace est muni d'une base  $\{u_k^{m+1}\}$  et  $\{u_k^{n+1}\}$ . Soit  $x \in \mathbb{R}^{m+1}$  un vecteur de déformation dans  $\mathcal{P}^m$ , alors il existe toujours un vecteur  $y \in \mathbb{R}^{n+1}$  qui définit la même déformation dans  $\mathcal{P}^n$ . On note  $P \in \mathbb{R}^{(n+1) \times (m+1)}$  la matrice de prolongement entre les espaces de recherche grossier et fin, par analogie aux méthodes multigrilles (voir la section 3.4.1). Ainsi les vecteurs  $x$  et  $y$  tels que

$$y = Px \quad (4.4)$$

définissent la même déformation de la méridienne

$$\Delta C(t) = \sum_{k=1}^{m+1} x_k u_k^{m+1}(t) = \sum_{k=1}^{n+1} y_k u_k^{n+1}(t). \quad (4.5)$$

On réfère l'opérateur  $P$  comme l'opérateur d'*élévation du degré*.

### 4.2.2 Prolongement basé sur le hessien paramétrique

On se rapporte ici aux opérateurs de la section 3.4.6. On considère une paramétrisation fine de dimension  $n$  et des paramétrisations grossières munies d'opérateurs d'élévation du degré  $P$  correspondants. Soit  $x \in \mathbb{R}^n$  un vecteur de conception de la paramétrisation fine et  $H = Z^T H(x) Z$  le hessien projeté sur l'espace admissible en  $x$ . L'opérateur de prolongement  $Q$  est issu de l'alternative suivante :

1.  $Q = P$  : prolongement classique



2.  $Q = \Omega \mathbb{P} \Omega^T P$  : réorganisation du spectre
3.  $Q = \Omega_2$  : sous-espace modal

L'opérateur de restriction est simplement le transposé de  $Q$  (cf. chapitre 3). Notons que les opérateurs 2 et 3 de l'alternative peuvent changer au cours des cycles puisqu'ils dépendent de la décomposition spectrale du hessien en un point  $x$  donné. Le problème n'étant pas quadratique, on peut mettre à jour l'opérateur  $Q$  au cours de la convergence de  $x$ . Cependant, ce calcul est à la fois coûteux et imprécis. En se basant sur l'étude de la section 4.1, on propose de remplacer  $H$  par une approximation purement géométrique.

### 4.2.3 Approximation par le modèle géométrique : filtre passe-haut

Soit  $H_g$  la matrice hessienne donnée par (3.10), le hessien du problème géométrique de meilleure approximation. Puisque ce problème est quadratique, ce hessien ne dépend pas du vecteur de conception. On modélise en quelque sorte le comportement quadratique de la fonctionnelle par son comportement géométrique. L'avantage de cette approche est que les opérateurs de transfert sont obtenus sans coût significatif supplémentaire. Sachant que  $H_g = \Omega_g \Lambda_g \Omega_g^T$  on a :

2.  $Q = \Omega_g \mathbb{P} \Omega_g^T P$  : réorganisation du spectre géométrique
3.  $Q = \Omega_{g,2}$  : sous-espace modal géométrique

L'enjeu des expériences numériques qui vont suivre est de vérifier que cette approche améliore toutefois le taux de convergence bien que les espaces grossiers ne correspondent pas exactement aux modes dans les directions desquelles les algorithmes convergent le plus lentement.

**Remarque** On pourrait également approcher  $H$  par le hessien exact d'un métamodèle de  $J$ , issu par exemple d'une base de données pourvue par les évaluations réalisées au cours d'une procédure d'optimisation par PSO. Cette approche est utilisée par exemple dans [22, 21, 35, 36, 38].

## 4.3 Algorithmes hiérarchiques en optimisation globale

Les algorithmes hiérarchiques sont étudiés en deux phases : d'abord dans le contexte de l'optimisation globale, puis dans le contexte de l'optimisation locale (section 4.4), afin de respecter l'ordre de la stratégie hybride d'optimisation successive (globale+locale) utilisée par la suite.

### 4.3.1 Extension de l'algorithme PSO

En s'inspirant de l'algorithme d'enrichissement progressif (voir section 3.4.2) on construit un algorithme hiérarchique où chaque niveau est résolu (entièrement ou par relaxation) sur la base d'un algorithme PSO. On utilise la mémoire globale pour transmettre la connaissance d'un bon candidat d'un niveau à un autre au moyen de l'élévation du degré. Cette stratégie est également décrite dans [37].

En partant d'une paramétrisation grossière, par exemple de degré  $n = 2$ , vu la faible dimension du problème, on est en mesure de converger rapidement vers un bon candidat que l'on note  $x^{*2}$ . En utilisant la propriété d'élévation du degré de la paramétrisation choisie (voir la section 4.2) cette même déformation est exprimée dans un espace de dimension supérieure, par exemple  $n = 4$ , et on note  $x^{*4}$  le vecteur de conception correspondant. On utilise alors  $x^{*4}$  comme initialisation de la mémoire globale dans le voisinage de laquelle les particules sont distribuées aléatoirement selon une loi uniforme. On conduit alors quelques itérations de PSO. Ce nouveau problème est *a priori* plus difficile car les risques de convergence vers un minimum local augmentent avec la dimension. Cependant, en partant d'un point « intelligent », on espère accroître la robustesse. En outre, le diamètre du domaine d'initialisation est réduit. Plus précisément, si on note  $\forall \rho > 0 \quad B_\infty(x^*, \rho) = \{x \in \mathbb{R}^n ; \|x - x^*\|_\infty \leq \rho\}$  la boule fermée centrée en  $x^*$  pour la norme infinie, alors l'initialisation du domaine est définie par  $D_0 = B(x_0, \rho_0)$  sur le premier niveau. Puis  $\rho$  est réduit à chaque changement de niveau de paramétrisation par un coefficient de contraction  $\beta$ . L'ensemble de la procédure est décrite dans l'Algorithme 5.

---

**Algorithme 5** : Optimisation Multiniveau par Essaim de Particules (MPSO)

---

**Données** : PSO –  $N, c_1, c_2, w_0, w_c, \text{imax}$ **Données** : MPSO –  $k_1 < k_2 < \dots < k_n$  : degrés des paramétrisations**Données** : MPSO –  $\rho_0, \beta$  : paramètres d'initialisation**Entrées** :  $x_0$ 

PSO : Algorithme 2, PSO classique

elevation : routine d'élévation du degré pour la paramétrisation considérée

// Géométrie initiale sur le niveau grossier

calcul initial de la fonction coût :  $J^* \leftarrow J(x_0)$  ( $x_0 \in \mathbb{R}^{k_1}$ )initialisation de la meilleure particule :  $x^{*k_1} \leftarrow x_0$ domaine d'initialisation centré en  $x_0$  :  $D_0 \leftarrow B_\infty(x_0, \rho_0)$ 

// Enrichissement progressif

**pour**  $i \leftarrow 1$  à  $n - 1$  **faire**    PSO classique :  $x^{*k_i} \leftarrow \text{PSO}(N, c_1, c_2, w_0, w_c, \text{imax}_i, x^{*k_i})$     transfert de la mémoire globale :  $x^{*k_{i+1}} \leftarrow \text{elevation}(x^{*k_i})$     réduction du domaine d'initialisation :  $\rho_{i+1} \leftarrow \beta \rho_i$     nouveau domaine :  $D_{i+1} \leftarrow B_\infty(x^{*k_{i+1}}, \rho_{i+1})$ **fin**

// Paramétrisation fine

PSO classique :  $x^{*k_n} \leftarrow \text{PSO}(N, c_1, c_2, w_0, w_c, \text{imax}_n, x^{*k_n})$ solution approchée :  $x^* \leftarrow x^{*k_n}$ fonction coût :  $J^* \leftarrow J(x^*)$ **Sorties** :  $x^*, J^*$ 

---

### 4.3.2 Expériences numériques de validation en *Optique Physique*

#### Validation de la méthode par un problème de conception inverse

On commence par un cas-test modèle de reconstruction de forme. Le champ électrique rayonné est approché par le modèle d'*Optique Physique*. Ce cas-test a deux avantages pour valider la méthode : (1) le problème inverse ainsi défini a une solution exacte connue ; (2) le temps de calcul d'une évaluation de la fonction objectif est très court, on peut donc se permettre de réaliser un grand nombre d'expériences. En faisant varier le paramètre aléatoire de l'algorithme PSO on peut analyser les résultats par rapport à l'échantillon statistique obtenu. Le cas-test est décrit à la Figure 4.11.

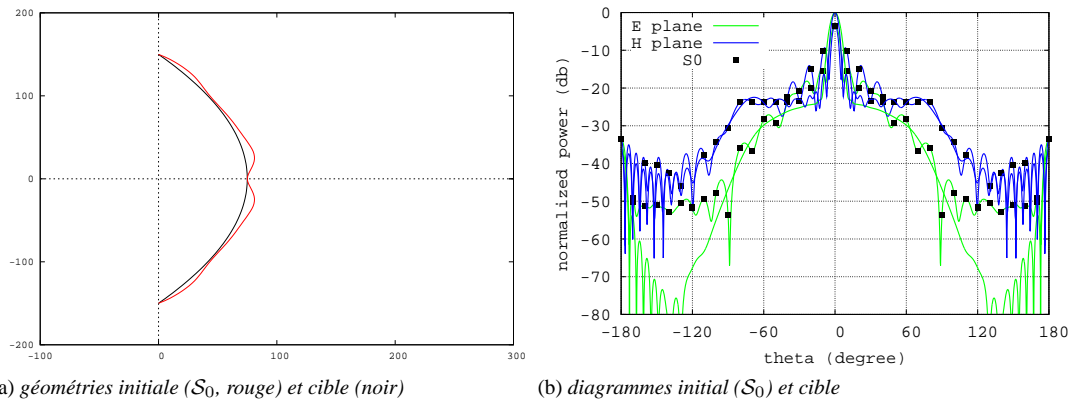


FIG. 4.11 – Description d'un cas-test de reconstruction de forme. Le champ électrique est calculé par le modèle approché d'Optique Physique.

On utilise trois niveaux qui sont  $5 \rightarrow 10 \rightarrow 15$ . Le nombre de particules est fixé à  $N = 40$  et le nombre maximal d'itérations à 300. Le coefficient de contraction est fixé à  $\beta = 0.25$  et on considère un critère de convergence sur la diversité de l'essaim pour le passage à un niveau supérieur : si la variance des positions est réduite de moitié, on change de niveau. Sur le dernier niveau on réalise 300 itérations. Les paramètres de l'algorithme PSO sont ceux fixés dans le Tableau 2.1. Enfin, on réalise 50 expériences indépendantes par rapport aux séquences aléatoires  $r_1$  et  $r_2$ . Les résultats globaux de l'échantillon sont représentés à la Figure 4.12.

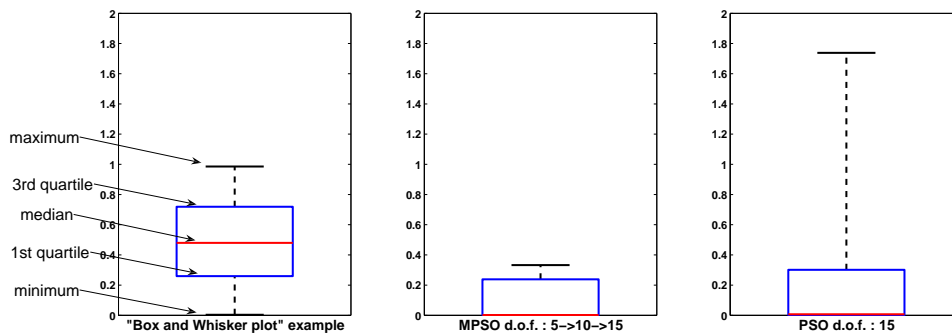


FIG. 4.12 – Représentation des résultats de l'échantillon sous forme de « boîtes à moustache ».

Algorithme	moyenne	médiane	écart type
PSO	0.238	$6.9 \cdot 10^{-3}$	0.434
MPSO	0.081	$1.7 \cdot 10^{-3}$	0.124

TAB. 4.1 – Indicateurs statistiques sur la convergence des approximations du minimum.

D'après les résultats obtenus il apparaît que l'algorithme MPSO est environ 3.5 fois moins dispersif que l'algorithme classique. Il est moins sensible aux variations de la séquence aléatoire. En ce sens il est plus *robuste*.

**Remarque** Le terme de *robustesse* apparaît dans le domaine de l'optimisation à plusieurs titres. Ici on parle de robustesse lorsque qu'un algorithme est peu sensible aux minima locaux. Il s'agit de la robustesse par rapport à une configuration initiale. Dans le cadre d'un problème d'optimisation multipoint on parle également de robustesse par rapport à un paramètre de fonctionnement.

### Optimisation d'un diagramme en directivité

Désormais, toujours dans le cadre du modèle d'*Optique Physique*, on cherche à uniformiser le diagramme de rayonnement sur l'intervalle  $[-10^\circ \ 10^\circ]$  à partir d'un paraboloïde alimenté par un dipôle. Comme le dipôle est placé à la distance focale du paraboloïde, le diagramme est directif (la puissance rayonnée est concentrée dans une direction), il s'agit donc d'un mauvais point de départ (cf. Figure 4.13).

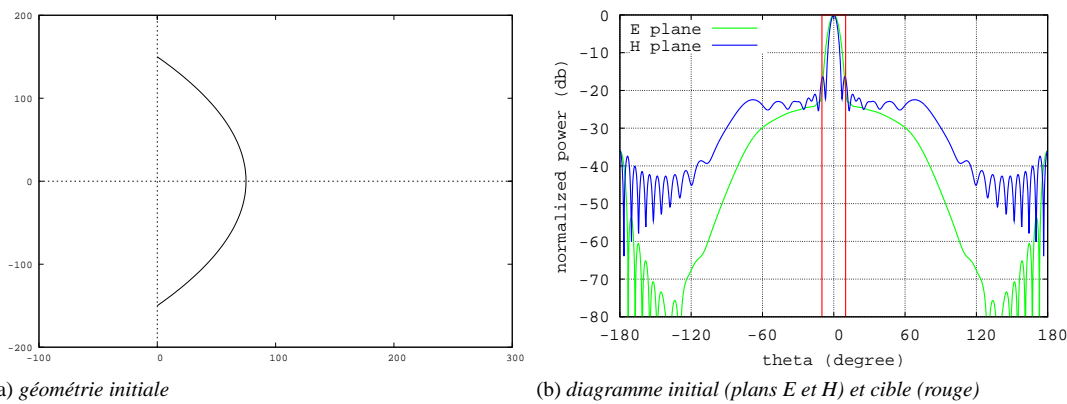


FIG. 4.13 – Description du cas-test d'optimisation d'un diagramme en directivité par la méthode MPSO.

Ce problème est très multimodal comme le montre la Figure 4.14 où on a représenté la fonction objectif en fonction de la déformation par rapport à deux paramètres de conception. Le point  $x = \mathbf{0}$  correspond à une déformation nulle, c'est-à-dire au paraboloïde.

Dans un premier temps on compare les algorithmes PSO et MPSO en terme de convergence de la fonction objectif pour un nombre de particules donné, à savoir  $N = 30$ . Les niveaux sont  $5 \rightarrow 10 \rightarrow 15$  (Figure 4.15a). Puis on augmente le nombre de particules de l'algorithme PSO jusqu'à atteindre une

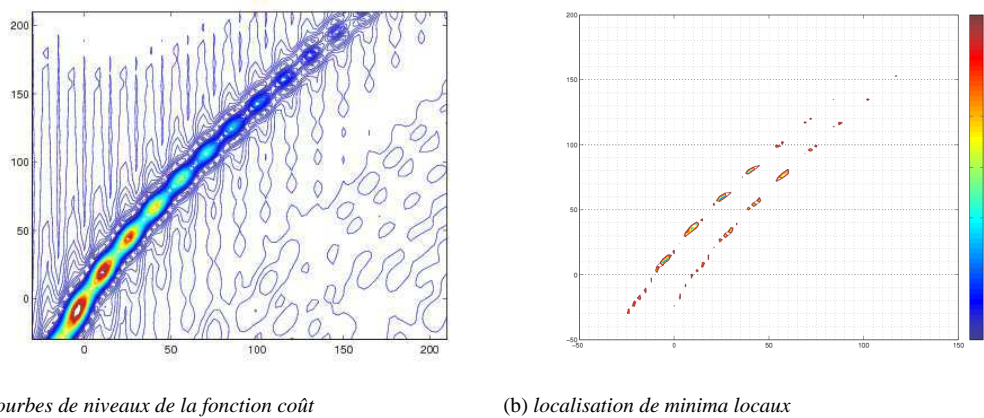
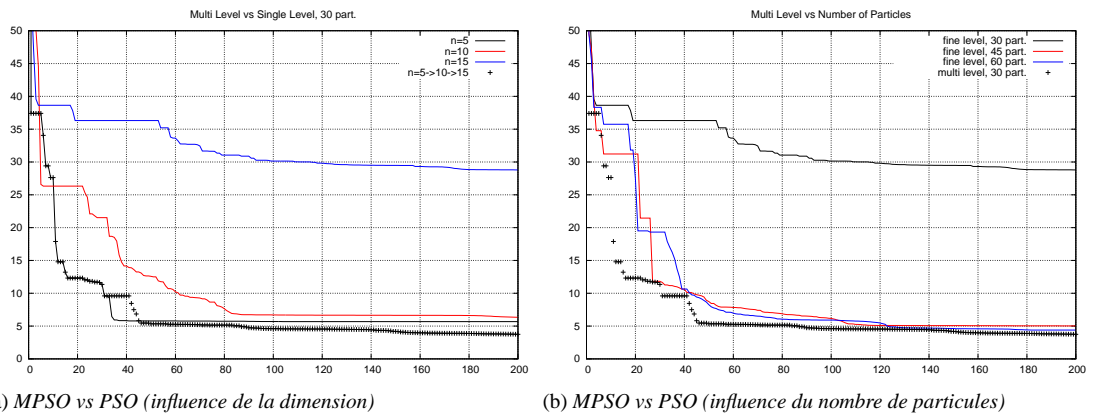


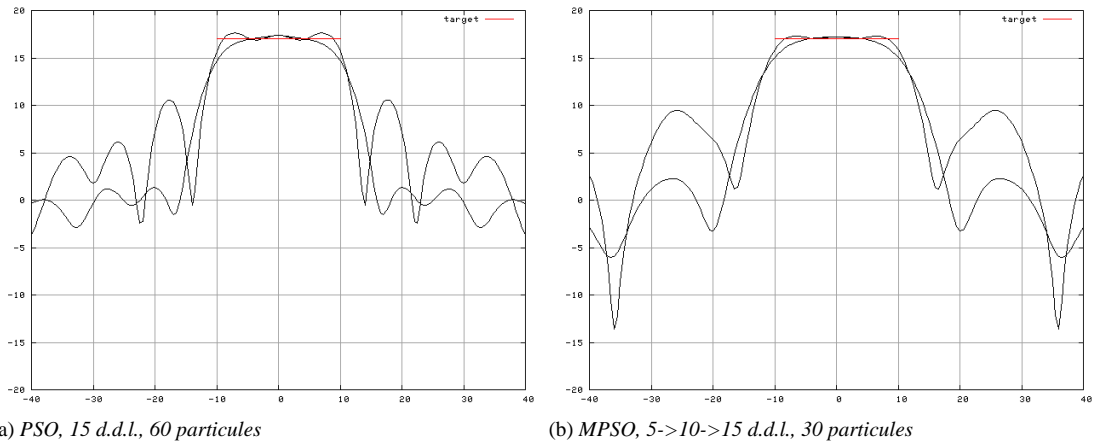
FIG. 4.14 – Valeur de la fonction objectif pour 2 degrés de liberté dans un voisinage de la forme initiale (paraboloïde).



(a) MPSO vs PSO (influence de la dimension)

(b) MPSO vs PSO (influence du nombre de particules)

FIG. 4.15 – Illustration de la convergence de la fonction coût en fonction des itérations. À gauche (a) on compare la convergence de l'algorithme MPSO pour la hiérarchie de degrés  $n = 5 \rightarrow 10 \rightarrow 15$  avec l'algorithme PSO pour les degrés 5, 10, et 15 indépendamment ; plus la dimension augmente, moins l'algorithme PSO est robuste ; la convergence sur le seul niveau fin est très mauvaise : la valeur de la fonction coût est 6 fois supérieure à celle obtenue avec l'algorithme MPSO au bout de 200 itérations. À la Figure (b) on augmente progressivement le nombre de particules de la méthode PSO : il faut 60 particules pour atteindre la performance de l'algorithme MPSO.



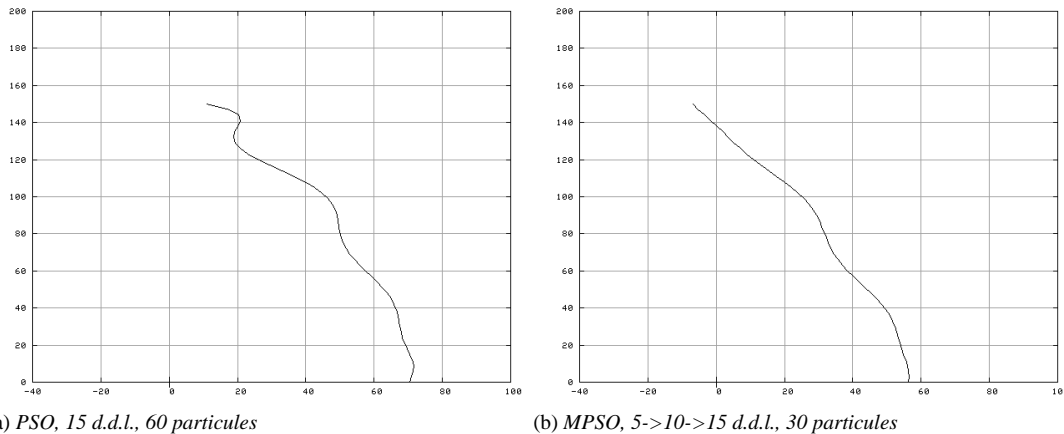
(a) PSO, 15 d.d.l., 60 particules

(b) MPSO, 5-&gt;10-&gt;15 d.d.l., 30 particules

FIG. 4.16 – Diagrammes de rayonnement des formes convergées de l'algorithme PSO avec 60 particules (a) et MPSO avec 30 particules (b) : les deux résultats sont satisfaisants du point de vue du cahier des charges mais la méthode à un seul niveau est deux fois plus coûteuse en temps de calcul.

précision (en terme de valeur de la fonction objectif) de l'ordre de celle obtenue avec l'algorithme MPSO (Figure 4.15b). À nombre égal de particules, l'algorithme MPSO est plus robuste. Il faut 60 particules à l'algorithme PSO pour atteindre la même performance, soit 2 fois plus d'évaluations de fonction coût.

Si on compare les diagrammes des formes obtenues à convergence des algorithmes PSO et MPSO (avec 60 et 30 particules resp.) on s'aperçoit que dans les deux cas la cible est atteinte de manière satisfaisante (cf. Figure 4.16). Cependant, si on compare la géométrie des réflecteurs (Figure 4.17), celle obtenue à convergence de la stratégie multiniveau est plus lisse. Ce résultat est confirmé par l'expérience suivante.



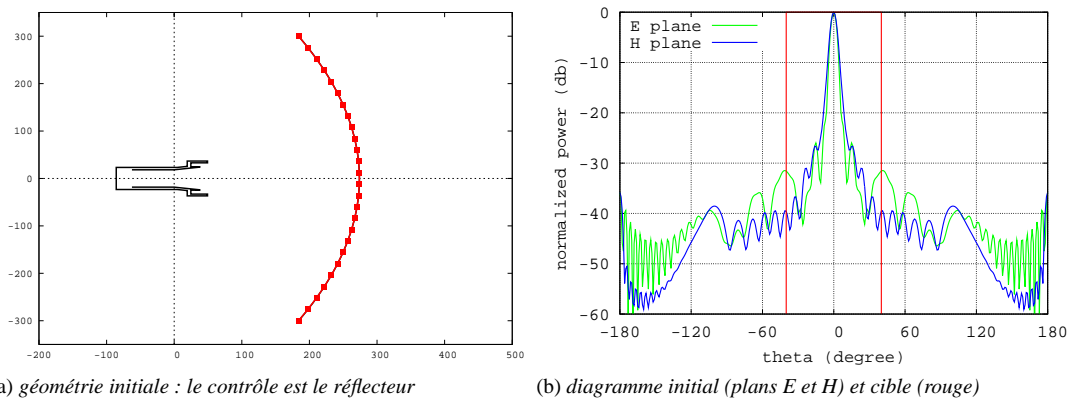
(a) PSO, 15 d.d.l., 60 particules (b) MPSO, 5->10->15 d.d.l., 30 particules  
 FIG. 4.17 – Formes convergées : la géométrie obtenue avec la méthode multiniveau est plus régulière.

### 4.3.3 Application à l'optimisation d'un diagramme en directivité d'une antenne

Considérons désormais le modèle de Maxwell. On cherche à optimiser un réflecteur alimenté par un guide d'onde pour que la puissance rayonnée soit uniforme sur l'intervalle  $[-40^\circ \ 40^\circ]$  et la plus faible possible ailleurs. Les paramètres du cas-test sont indiqués dans le Tableau 4.2 et la géométrie initiale à la Figure 4.18.

<b>Fréquence</b>	$f = 7.13333 \text{ GHz}$
<b>Paramétrisation</b>	Bernstein
<b>Hiérarchie</b>	$n = 5 \rightarrow 10 \rightarrow 15 \rightarrow 20 \rightarrow 25 \rightarrow 30$
<b>Nombre d'itérations (MPSO)</b>	$i_{\max} = 20 \rightarrow 30 \rightarrow 40 \rightarrow 60 \rightarrow 80 \rightarrow 170$
<b>Nombre d'itérations (PSO)</b>	$i_{\max} = 400$
<b>Nombre de particules</b>	$N = 100$

TAB. 4.2 – Paramètres du cas-test.



(a) géométrie initiale : le contrôle est le réflecteur (b) diagramme initial (plans E et H) et cible (rouge)

FIG. 4.18 – Description du cas-test d'optimisation de la géométrie d'un réflecteur.

La Figure 4.19 montre la convergence des algorithmes en terme d'évolution de la performance au cours des itérations. Les Figures 4.20 et 4.21 illustrent les diagrammes et les formes convergées respectivement. La performance en terme de fonction coût est meilleure avec la stratégie multiniveau : la plus petite valeur atteinte est de 34% plus petite que celle obtenue avec la stratégie classique. À la vue des diagrammes de rayonnement dans la zone  $[-40^\circ \ 40^\circ]$ , la cible est atteinte par les deux stratégies. La directivité maximale est alors de 9.75 dbi et 9.87 dbi respectivement. Le diagramme de l'antenne obtenue avec la stratégie multiniveau est toutefois meilleur dans la zone complémentaire où le maximum de la

directivité est de -18 db (par rapport à -12 db avec l'algorithme classique). La différence la plus évidente réside dans la géométrie des réflecteurs : la forme obtenue avec la stratégie multiniveau est beaucoup plus régulière.

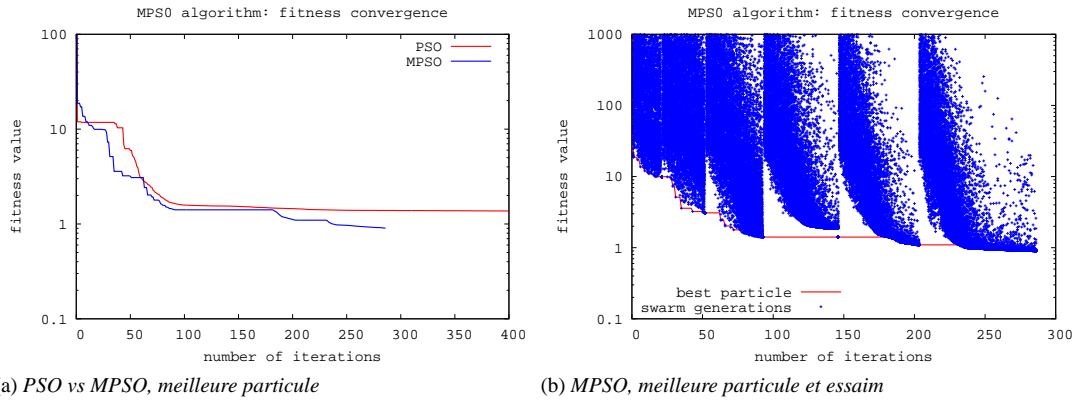


FIG. 4.19 – Convergence de la fonction coût.

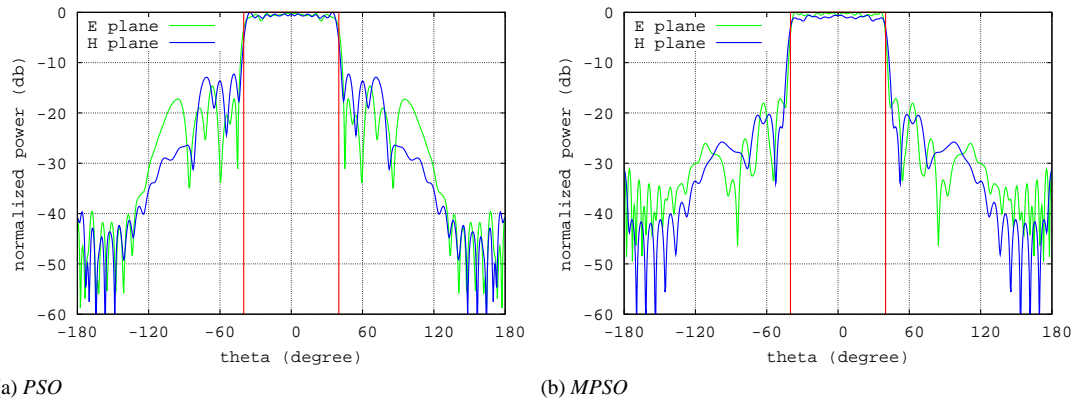


FIG. 4.20 – Diagrammes de rayonnement des formes convergées.

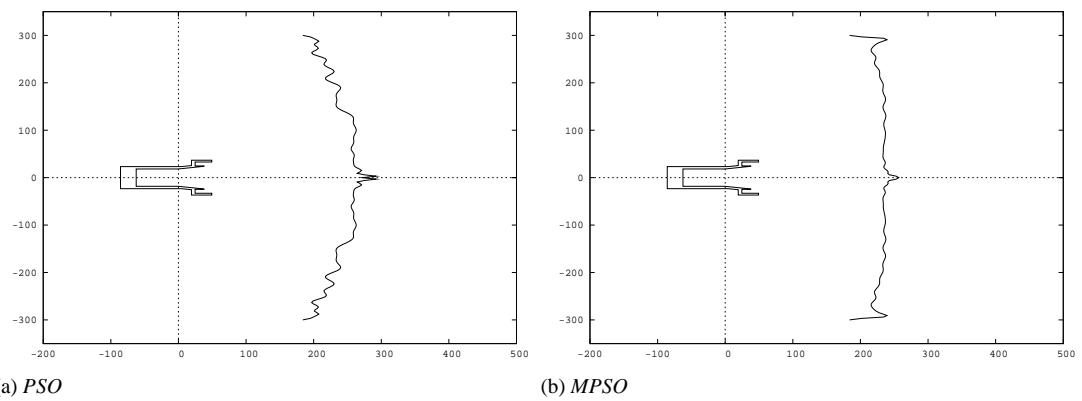


FIG. 4.21 – Géométrie des antennes à convergence des algorithmes.

## 4.4 Algorithmes hiérarchiques en optimisation locale

### 4.4.1 Algorithme multiniveau non-linéaire

On cherche à étendre les stratégies multiniveaux du chapitre 3 à l'optimisation de forme des réflecteurs. Clairement, les fonctions objectifs ne sont pas quadratiques par rapport aux paramètres de conception. La recherche d'un point stationnaire est donc un problème non-linéaire.

Comme pour tout algorithme de descente on dispose d'un vecteur de conception initial  $x^0$ . Les sous-espaces de recherche, y compris le niveau fin  $F$ , sont considérés comme des espaces de correction. Plus précisément, pour  $x \in F$  donné, on cherche à résoudre le problème de minimisation

$$\min_{y \in \mathbb{R}^m} j(y) = J(x + Qy) \quad (4.6)$$

pour la correction de  $x$  sur un sous-espace  $V$  de dimension  $m$  (3.77) où  $Q$  est un des opérateurs de prolongement proposés à la section 4.2. Il est fondamental de remarquer que pour tout niveau (c-à-d quelle que soit la dimension du sous-espace), la fonction coût est toujours évaluée sur l'espace fin  $F$ . Les sous-problèmes (4.6) sont donc consistants avec le problème initial : la minimisation de  $J$ .

Supposons que  $x$  soit un vecteur de conception du niveau fin obtenu au bout d'un certain temps du processus de minimisation et posons  $y^0 = \mathbf{0}$  comme vecteur initial de correction. On a

$$j(y^0) = J(x) \quad (4.7)$$

$$\nabla_y j(y^0) = Q^T \nabla_x J(x) \quad (4.8)$$

de telle sorte que si  $x$  est stationnaire

$$\nabla_x J(x) = \mathbf{0} \quad \Rightarrow \quad \nabla_y j(y^0) = \mathbf{0} \quad (4.9)$$

quel que soit  $Q$ , et sinon

$$-Q \nabla_y j(y^0) \cdot \nabla_x J(x) = -\nabla_x J(x)^T Q Q^T \nabla_x J(x) = -\|Q^T \nabla_x J(x)\|^2 < 0 \quad (4.10)$$

où  $Q \nabla_y j(y^0)$  est le prolongement sur le niveau fin du gradient de  $j$  en  $y^0$ . En d'autres termes, d'une part il y a continuité de la fonction coût entre deux niveaux (4.7), et d'autre part le prolongement de  $-\nabla_y j(y^0)$ , qui est la direction de plus grande descente pour  $j$  en  $y^0$ , est une direction de descente pour  $J$  en  $x$  (4.10).

Au même titre que l'Algorithme 4 pour la résolution du problème modèle linéaire, les sous-problèmes (4.6) sont soumis à une phase de relaxation équivalente à quelques itérations d'une méthode de descente (SD, CG). Il s'agit alors d'un « anti-lisseur » si l'on se réfère à l'analyse spectrale de la section 3.5.1. Par analogie avec les méthodes idéales, on considérera une convergence complète sur le niveau de plus faible dimension, soit à l'aide d'une méthode de Newton-Raphson (c-à-d par inversion du hessien exact, pourvu qu'il soit défini positif) soit avec une méthode de quasi-Newton (jusqu'à satisfaction d'un critère de convergence).

On appelle *cycle* une séquence de sous-espaces de recherche pour la correction successive des paramètres de conception. Un tel cycle est entièrement défini par la dimension des sous-espaces :  $m_1, m_2, m_3$ , etc. et le choix de l'opérateur de prolongement. Ce cycle est répété autant de fois que nécessaire pour atteindre une tolérance numérique donnée. Si les sous-espaces reposent sur la décomposition spectrale du hessien en un point  $x$  donné, alors ils ne sont pas définis de manière unique. La stratégie non-linéaire consiste à mettre à jour les sous-espaces à l'issue de chaque cycle. Un exemple typique de cycle en « dent de scie » à  $K$  niveaux ( $m_K = n$ ) est illustré à la Figure 4.22 où on a noté  $\boxplus$  une phase de relaxation sur le niveau fin précédée par un calcul des sous-espaces propres du hessien en  $x^{(v)}$ , approximation obtenue après le cycle  $v$ . L'algorithme 6 résume l'ensemble de la procédure.



**Algorithme 6** : Algorithme d'optimisation multiniveau

**Données** : paramétrisation fine  $F$  ( $\dim(F) = n$ ),  $K$  sous-espaces  $V_{m_k}$  ( $F = V_{m_K} = V_n$ )

**Données** : tolérance  $\varepsilon$ , nombre d'itérations de relaxation  $\nu$

**Entrées** :  $x^{(0)} \in F$

descente : routine de calcul d'une direction de descente

spec : routine de calcul de la décomposition spectrale

update : routine de calcul de l'opérateur de prolongement

critere : routine de calcul d'un critère de convergence (norme du gradient, résidu, etc.)

$v \leftarrow 0$

// Tant que le critère de convergence n'est pas satisfait

**tant que**  $r^{(v)} > \varepsilon$  **faire**

    // Mise à jour des sous-espaces propres

$H \leftarrow \nabla^2 J(x^{(v)})$

$\Omega \Lambda \Omega^T \leftarrow \text{spec}(H)$

    // Sur chaque sous-espace du cycle

**pour chaque**  $V_m$  **faire**

        // Opérateur de prolongement

$Q \leftarrow \text{update}(\Omega, m)$

        // Initialisation de la correction

$y^0 \leftarrow \mathbf{0} \in \mathbb{R}^m$

        // Résolution complète OU Relaxation

**si**  $m = m_1$  **alors**

$x^{(v+1)} \leftarrow x^{(v)} + Q \arg \min_y J(x^{(v)} + Qy)$

**sinon**

**pour**  $j \leftarrow 1$  à  $\nu$  **faire**

$d^{j-1} \leftarrow \text{descente}(y^{j-1})$

$y^j \leftarrow y^{j-1} + \tau^{j-1} d^{j-1}$

$x^{(v+1)} \leftarrow x^{(v)} + Qy^j$

**fin**

**fin**

**fin**

$v \leftarrow v + 1$

$r^{(v)} \leftarrow \text{critere}(\nabla J(x^{(v)}), J(x^{(v)}), J(x^{(v-1)}))$

**fin**

**Sorties** :  $v, x^{(v)}, J(x^{(v)})$

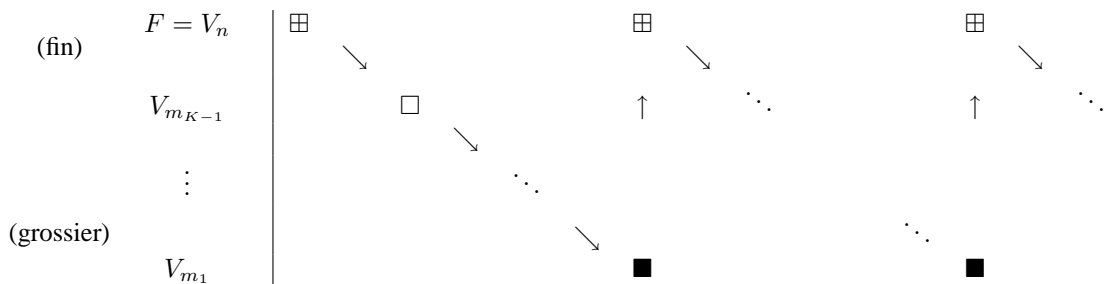


FIG. 4.22 – Schéma de cycles en dents de scie avec mise à jour des sous-espaces de correction;  $\square$ =relaxation,  $\boxplus$ =mise à jour des sous-espaces et relaxation,  $\blacksquare$ =correction idéale.

## 4.4.2 Expériences numériques

### Cas-test modèle *Optique Physique*

On considère un réflecteur alimenté par un dipôle. On cherche à uniformiser le diagramme sur l'intervalle  $[-60^\circ 60^\circ]$ . La forme initiale représentée à la Figure 4.23 est obtenue avec l'algorithme MPSO dans la paramétrisation de Bézier de degré 15 sur le niveau fin. À partir de cette solution on cherche à raffiner la forme obtenue avec une méthode de descente.

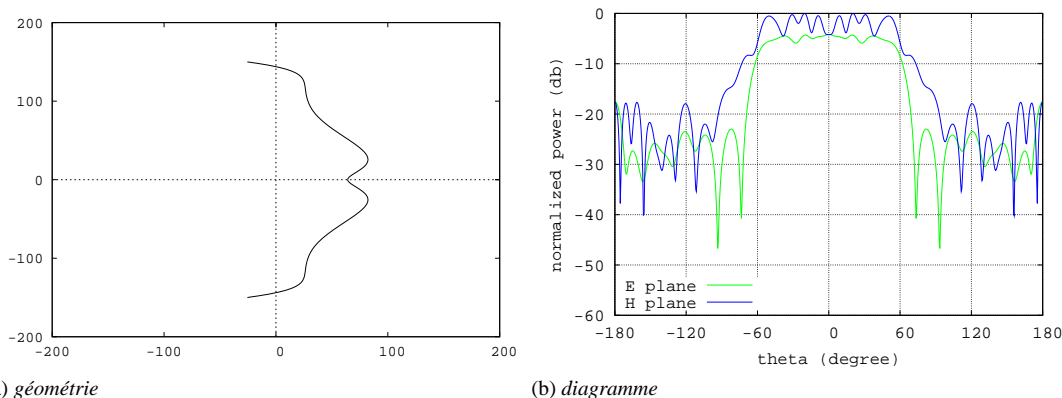


FIG. 4.23 – Configuration initiale du cas-test d'optimisation multiniveau.

L'intérêt de réaliser un cas-test avec une analyse par le modèle d'*Optique Physique* vient du fait que l'on dispose à la fois du gradient exact et du hessien numérique exact obtenu par différentiation automatique. On peut donc comparer les stratégies multiniveaux lorsque le transfert est basé sur le hessien exact (section 4.2.2) ou lorsqu'il est basé sur le hessien géométrique (section 4.2.3).

Dans les expérimentations qui suivent on adopte les notations suivantes pour identifier les méthodes :

- simple niveau : SL
- multiniveau avec transfert « élévation du degré » : MLY
- multiniveau avec transfert « réorganisation du spectre » : MLZ
- multiniveau avec transfert « algébrique » : MLO

Pour les stratégies MLZ et MLO on précisera si le hessien est exact (*Optique Physique*, noté  $H$ ) ou géométrique (noté  $H_g$ ). Les paramètres sont donnés dans le Tableau 4.3.

<b>Fréquence</b>	$f = 7.0$ GHz
<b>Paramétrisation</b>	Bernstein
<b>Hiérarchie</b>	$n = 4 - 7 - 10 - 13 - 16 - 19 - 21 - 22 - 25$
<b>Cycle</b>	dents de scie (Figure 4.22)
<b>Relaxation</b>	CG, $\nu = 2$
<b>Nombre maximal de cycles</b>	100

TAB. 4.3 – Paramètres de l'algorithme multiniveau.

Les résultats en terme de convergence de la fonction coût sont représentés à la Figure 4.24 : à gauche (a) les sous-espaces de correction sont mis à jour par le calcul des vecteurs propres du hessien exact ; à droite (b) les sous-espaces de correction sont donnés par les vecteurs propres du problème géométrique. On donne également dans le Tableau 4.4 les valeurs de la fonction coût normalisée obtenus au bout de 100 cycles, ce qui correspond environ à 2000 évaluations de la fonction coût pour la stratégie à un seul niveau.

Globalement il apparaît que les stratégies multiniveaux Z et O sont plus efficaces que la stratégie à un seul niveau ou que la stratégie multiniveau Y. Ces résultats sont en accord avec ceux obtenus pour le problème modèle géométrique. Par ailleurs, lorsque les sous-espaces de correction sont définis par les espaces propres du hessien exact, les stratégies Z et O sont les plus performantes en terme de convergence

itérative. Cependant on n'a pas représenté le coût du calcul du hessien. En ce sens, si la version géométrique des stratégies Z et O est moins efficace en terme de taux de convergence, elle est moins coûteuse en temps de calcul. On garde toutefois en tête que si le temps de calcul est un facteur important, c'est bien la convergence qui nous intéresse,

SL	ML Y	ML Z		ML O	
		$H$	$H_g$	$H$	$H_g$
0.92919	0.93672	0.84471	0.86765	0.85038	0.86959

TAB. 4.4 – Performance au bout de 100 cycles, soit environ 2000 évaluations de la fonction coût.

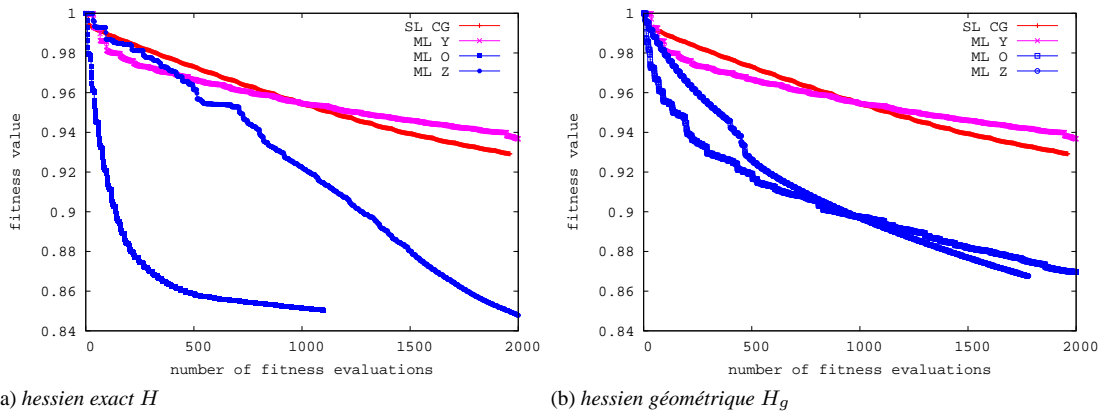


FIG. 4.24 – Évolution de la performance en fonction du nombre d'évaluations de la fonction coût. La relaxation est réalisée par une méthode de CG.

On modifie désormais la méthode de relaxation (voir Tableau 4.5) : on utilise une méthode de quasi-Newton (QN) BFGS. On compare la stratégie SL avec la stratégie ML Z. D'après la convergence de la fonction coût (cf. Figure 4.25a), les stratégies SL et ML Z basée sur le hessien exact semblent aboutir à la même valeur optimale au bout de 500 évaluations de fonction coût ( $J^* \approx 0.70$ ). La version géométrique de la stratégie multiniveau est moins performante. En d'autres termes la stratégie multiniveau n'apporte pas d'amélioration supplémentaire au préconditionnement BFGS (on ne représente pas de résultats avec la méthode MLO, les résultats étant sensiblement les mêmes).

Enfin, si on change la paramétrisation de Bernstein par celle de Legendre, la convergence est encore plus rapide (convergence en moins de 300 itérations) et plus précise ( $J^* \approx 0.63$ , cf. Figure 4.25b). On rappelle que la base de Legendre est mieux conditionnée du point de vue du problème purement géométrique (voir la Figure 3.26).

<b>Fréquence</b>	$f = 7.0$ GHz
<b>Paramétrisation</b>	Bernstein / Legendre
<b>Hiérarchie</b>	$n = 4 - 7 - 10 - 13 - 16 - 19 - 21 - 22 - 25$
<b>Cycle</b>	dents de scie (Figure 4.22)
<b>Relaxation</b>	QN, $\nu = 2$
<b>Nombre maximal de cycles</b>	100

TAB. 4.5 – Paramètres de l'algorithme multiniveau.

En résumé, on compare à la Figure 4.26 la convergence des principales stratégies : simple niveau avec la méthode du Gradient Conjugué dans la paramétrisation de Bernstein, multiniveau avec relaxation par Gradient Conjugué, et simple niveau avec la méthode de quasi-Newton BFGS dans les paramétrisations de Bernstein et Legendre. Il apparaît que dans le contexte d'une stratégie itérative du Gradient Conjugué, les méthodes multiniveaux améliorent le taux de convergence pourvu que les sous-espaces soient correctement choisis (MLZ ou MLO). Cependant, la méthode de quasi-Newton est un bon préconditionneur pour lequel les stratégies multiniveaux n'ont pas d'influence supplémentaire.

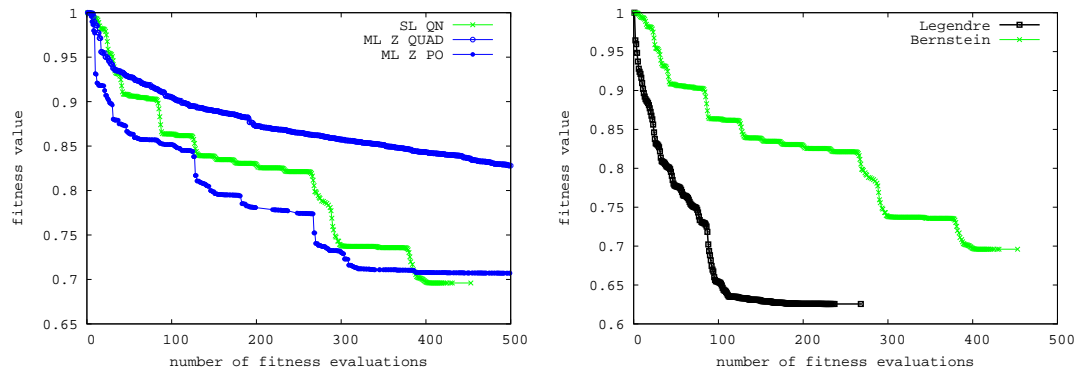
(a) *ML Z, Bernstein*(b) *SL, Bernstein vs Legendre*

FIG. 4.25 – Évolution de la performance en fonction du nombre d'évaluations de la fonction coût. La relaxation est réalisée par la méthode de QN BFGS.

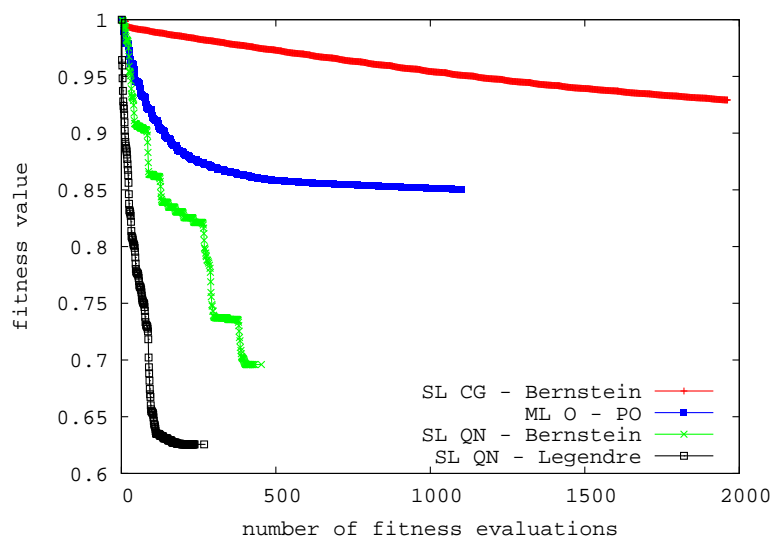


FIG. 4.26 – Convergence des principales stratégies.

On représente les diagrammes et les formes obtenus à convergence des méthodes simple niveau QN pour chaque paramétrisation sur les Figures 4.27 et 4.28 respectivement. La différence maximale de la puissance rayonnée dans la zone  $[-60^\circ \ 60^\circ]$  par la forme initiale est de -4.5 db (cf. Figure 4.23b). Après optimisation de la forme du réflecteur elle est de -3 db (Bernstein) et de -2.5 db (Legendre). Ce résultat est cependant à relativiser du point de vue de la physique étant donné que les formes, très irrégulières, ne satisfont plus les hypothèses d'approximation du champ par le modèle d'*Optique Physique*. Les résultats pertinents de cette expérience se rapportent essentiellement à la convergence numérique des stratégies multiniveaux proposées dans le contexte d'un problème modèle complexe non-linéaire.

On cherche désormais à confirmer ces résultats dans le cadre physique d'une analyse par le modèle de Maxwell. Seule la stratégie « géométrique » est considérée puisque nous n'avons pas d'autre approximation que celle fournie par un calcul de différences finies et qui serait trop coûteuse en temps de calcul.

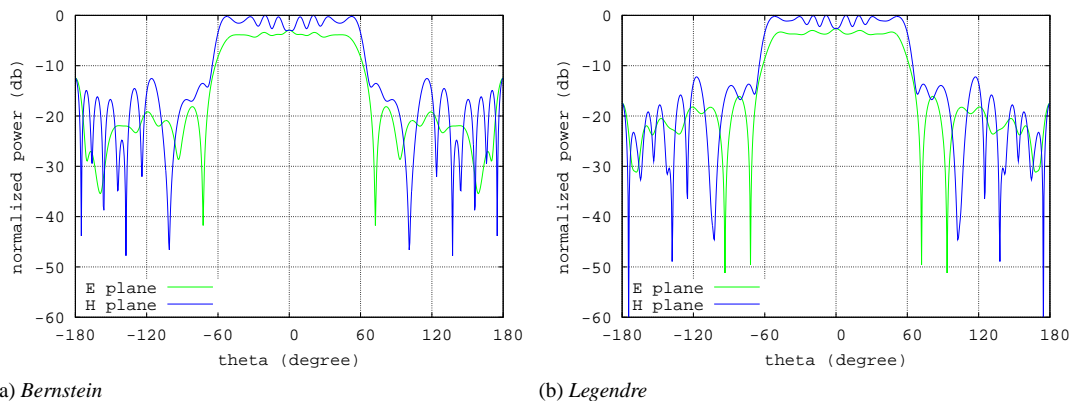


FIG. 4.27 – Diagrammes à convergence des algorithmes SL QN dans les paramétrisations de Bernstein et Legendre.

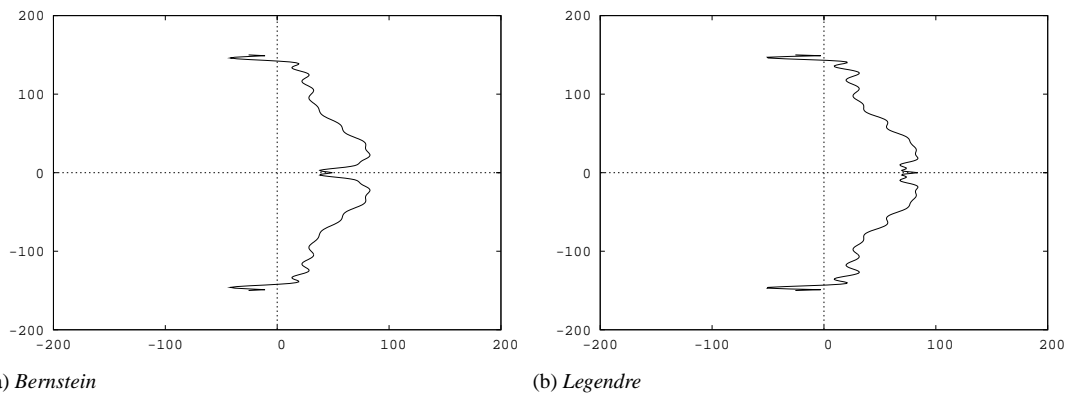


FIG. 4.28 – Formes convergées des algorithmes SL QN dans les paramétrisations de Bernstein et Legendre.

### Optimisation d'une d'antenne en directivité

On reprend le problème d'optimisation de la section 4.3.3. Le forme initiale est donnée par l'algorithme MPSO dans la paramétrisation de Bézier de degré 15 sur le niveau fin. La forme est régulière et le diagramme proche de la cible mais pas encore satisfaisant (cf. Figure 4.29). On conduit une expérimentation numérique de l'ensemble des stratégies multiniveaux locales.

Trois paramétrisations sont considérées : Bézier-Bernstein et B-spline d'ordre 4 dans une formulation FFD classique (cf. section 2.2.5), et B-spline d'ordre 4 dans une formulation de déformation normale (cf. section 2.2.6).

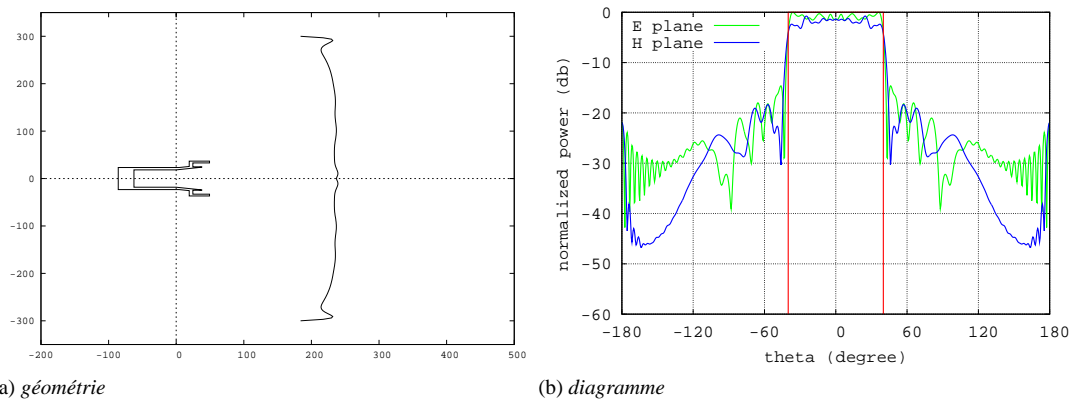


FIG. 4.29 – Configuration initiale de l'antenne en vue de l'optimisation du diagramme par une méthode multiniveau.

Dans la base des fonctions B-splines, ni l'opérateur d'élévation du degré ni celui d'insertion de nœuds n'a été considéré. Ainsi, seule la stratégie multiniveau de type « algébrique » (MLO) est expérimentée. L'avantage de cette méthode est de créer une hiérarchie de paramétrisations emboîtées à partir d'une paramétrisation quelconque. Toutefois, le terme « algébrique » peut paraître inapproprié puisqu'on ne dispose pas du hessien exact. Les sous-espaces de correction sont les sous-espaces propres du problème géométrique, utilisant de ce fait l'argument géométrique de la stratégie, à savoir que ce sont les modes de haute fréquence dont la convergence est la plus lente.

On considère un cycle en dents de scie comme illustré sur la Figure 4.22, sans mise à jour du hessien ni convergence complète sur le niveau le plus bas. Les dimensions des sous-espaces sont données dans le Tableau 4.6. En général, la relaxation est réalisée par la méthode du Gradient Conjugué. Cependant, dans la paramétrisation de déformation normale, seule une relaxation par la méthode de plus grande descente a un sens (cf. section 2.2.6). Le nombre d'itérations de relaxation est fixé à  $\nu = 2$ . Enfin, lorsque cela sera possible, on comparera les résultats des stratégies multiniveaux avec celui d'une stratégie à un seul niveau et la méthode de quasi-Newton (c-à-d dans le cas des paramétrisations classiques).

	Hiérarchie de la dimension des sous-espaces
<b>Bernstein</b>	6 → 9 → 12 → 15 → 16 → 18 → 21 → 24
<b>B-spline</b>	3 → 6 → 9 → 12 → 15 → 16 → 18 → 21 → 24 → 27 → 30

TAB. 4.6 – Dimensions des hiérarchies de paramétrisations.

En résumé, on trouve dans le Tableau 4.7 l'ensemble des stratégies expérimentées selon la paramétrisation et les méthodes de relaxation correspondantes.

	SL	ML Y	ML Z	ML O
<b>Bernstein</b>	CG QN	CG	CG	CG
<b>B-spline</b>	CG QN			CG
<b>B-spline, normal</b>	SD			SD

TAB. 4.7 – Résumé des expériences numériques réalisées.

**Paramétrisation classique – Bernstein** La convergence de la fonction objectif est représentée à la Figure 4.30. L'allure des courbes de convergence est conforme aux résultats précédents : les stratégies multiniveaux O et Z améliorent le taux de convergence par rapport à la stratégie sur le seul niveau fin (au bout de 400 évaluations de la fonction coût, la performance est meilleure) ; la stratégie multiniveau Y converge plus lentement. Enfin, la stratégie QN est la plus efficace.

Les améliorations relatives sont toutefois assez faibles : par rapport à la forme initiale, le critère est réduit de 2.0% par la méthode simple niveau, de 1.5% par la méthode Y, et d'environ 2.5% par les méthodes Z et O. Ce résultat se retrouve dans les diagrammes (cf. Figure 4.31) et les formes (cf. Figure 4.32) : les

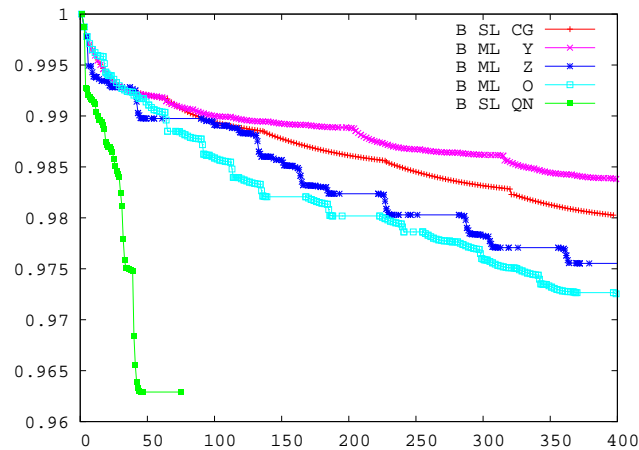


FIG. 4.30 – Évolution de la fonction coût – Bernstein

diagrammes ne sont pas discernables entre eux, ni avec le diagramme initial ; les déformations de la stratégie à un seul niveau et de la méthode Y sont identiques, celles des méthodes Z et O comportent plus de modes de haute fréquence.

Le problème est rendu extrêmement raide par une paramétrisation mal conditionnée. Même la méthode QN semble peu efficace, qui, à convergence au bout de 75 itérations, améliore de moins de 4% le critère initial. Il est probable que la solution obtenue soit un minimum local ou que le gradient ne soit pas assez précis dans cette paramétrisation. Pour le vérifier, on réalise la même expérience avec la paramétrisation B-spline d'ordre 4, mieux conditionnée d'un point de vue géométrique.

**Paramétrisation classique – B-spline** La convergence de la fonction objectif est représentée sur la Figure 4.33, les diagrammes sur la Figure 4.34 et les formes sur la Figure 4.35. Au bout de 400 évaluations la stratégie à un seul niveau améliore de 30.0% le critère initial avec la méthode du Gradient Conjugué. Encore une fois, la stratégie multiniveau ML O est plus efficace : elle atteint la même performance en 200 évaluations (soit deux fois moins d'effort de calcul) et elle améliore de 37.5% le critère initial en 400 évaluations. Enfin, la méthode de quasi-Newton est la plus efficace : elle converge en 200 évaluations et améliore le critère initial de 50.0%.

Même avec une paramétrisation bien conditionnée la méthode multiniveau s'avère avantageuse dans le contexte d'un optimiseur de type CG. Dans l'expérience suivante où la déformation est normale à la méridienne, la méthode de quasi-Newton ou du Gradient Conjugué n'a pas de sens. La relaxation est effectuée par une méthode de plus grande descente.

**Paramétrisation normale – B-spline** La convergence de la fonction objectif est représentée sur la Figure 4.36, les diagrammes sur la Figure 4.37 et les formes sur la Figure 4.38. Des courbes de convergence similaires à celle de la paramétrisation classique sont obtenues. Au bout de 400 évaluations la stratégie à un seul niveau améliore de 31.5% le critère initial. La stratégie multiniveau ML O est également efficace : elle permet d'atteindre cette performance en 160 évaluations et elle améliore de 38.0% le critère initial en 400 évaluations.

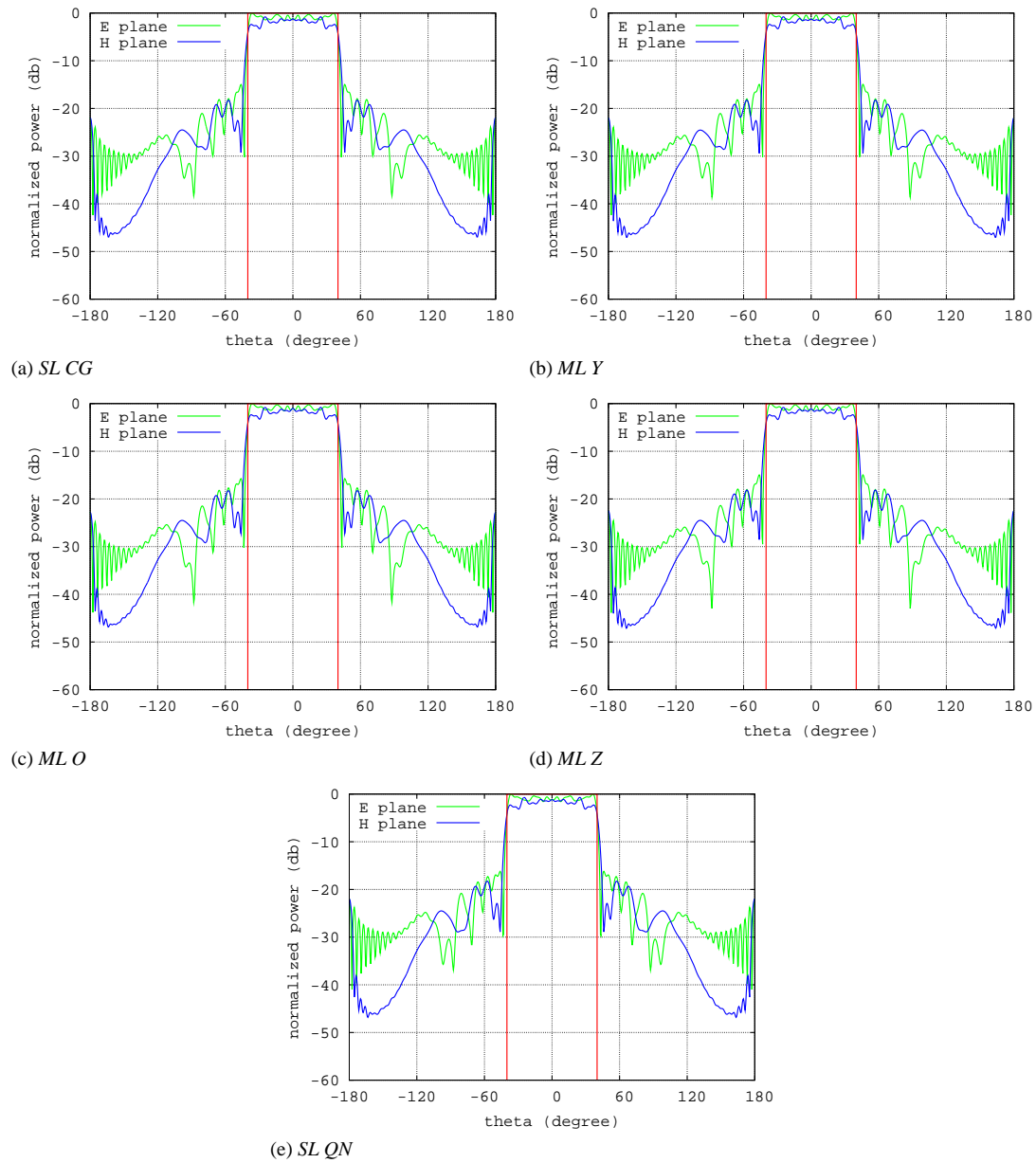


FIG. 4.31 – Diagrammes obtenus à convergence ou au bout de 400 évaluations de la fonction coût – Bernstein



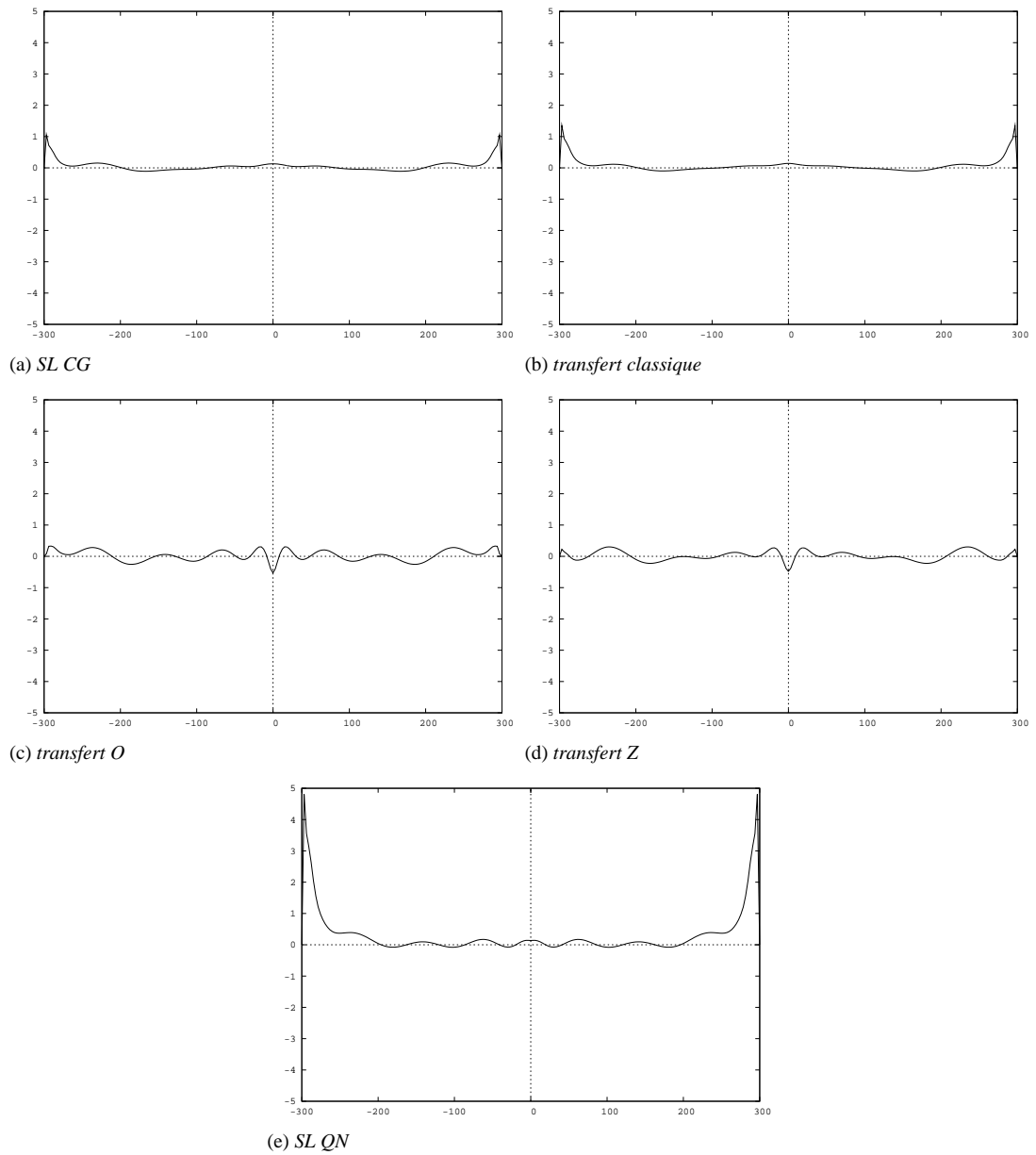


FIG. 4.32 – Formes convergées : représentation de la déformation du réflecteur initial – Bernstein

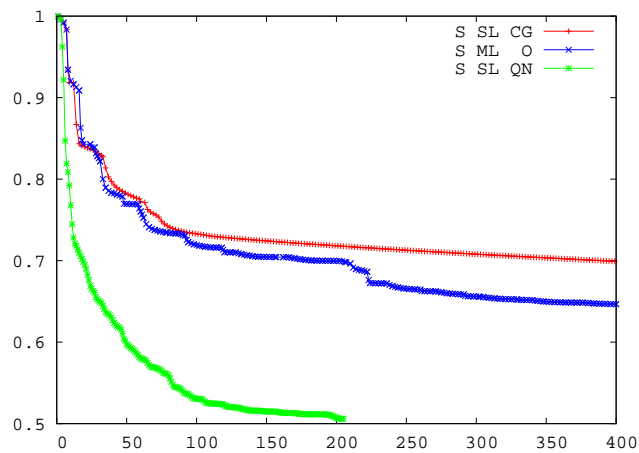


FIG. 4.33 – Évolution de la fonction coût – B-spline

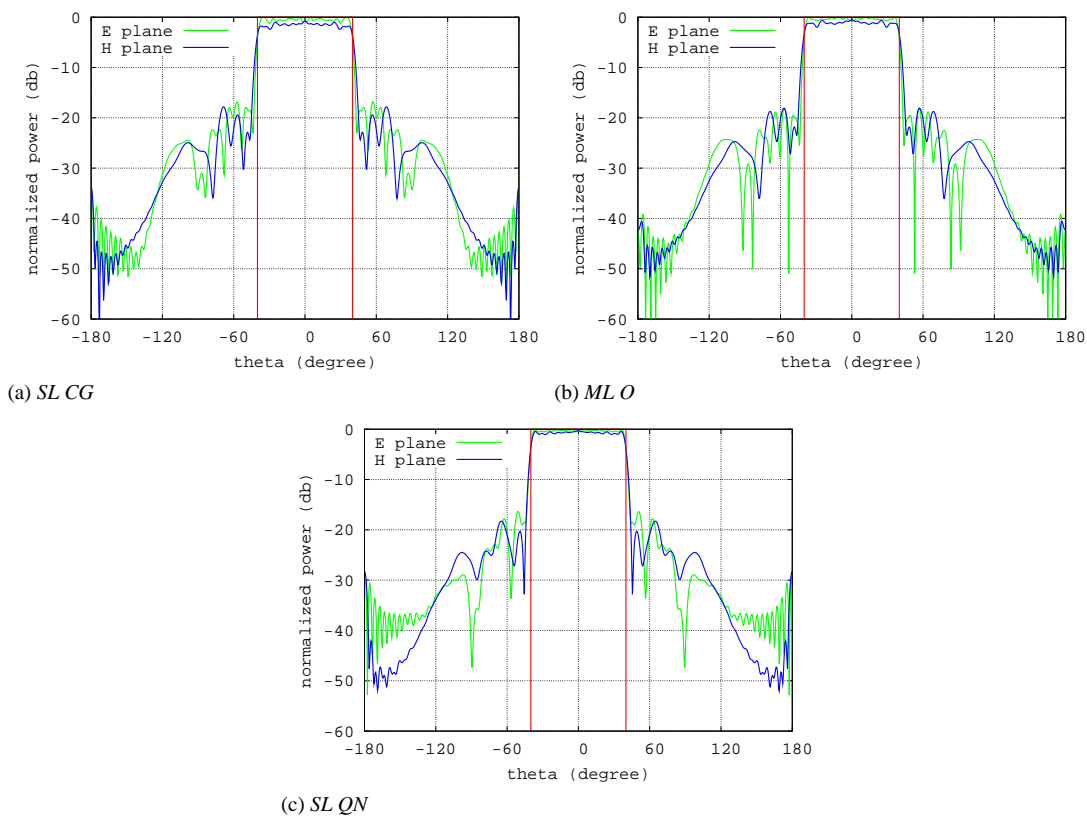


FIG. 4.34 – Diagrammes obtenus à convergence ou au bout de 400 évaluations de la fonction coût – B-spline

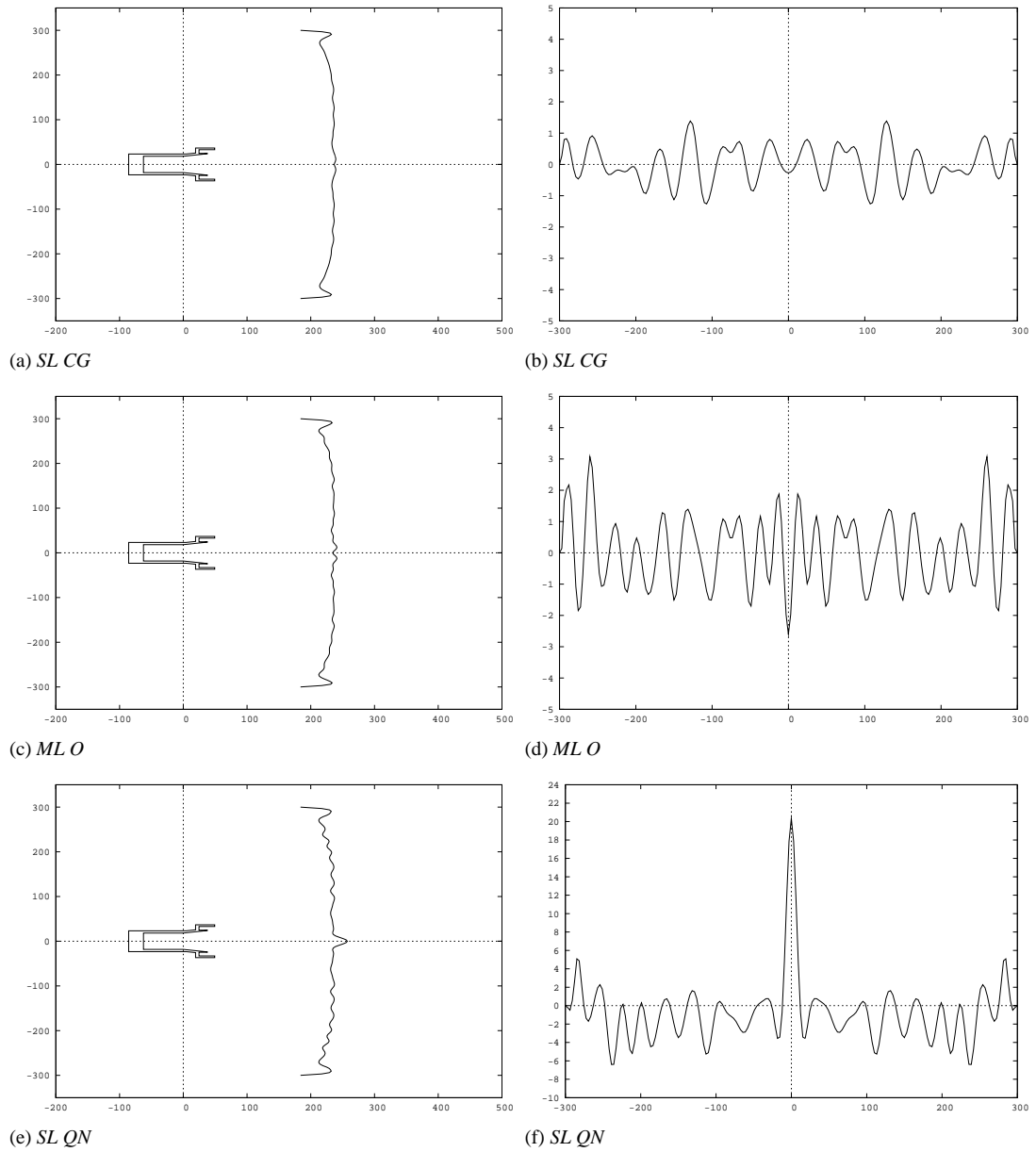


FIG. 4.35 – Formes convergées : représentation de l'antenne (a-c-e) et de la déformation du réflecteur initial (b-d-f) – B-spline

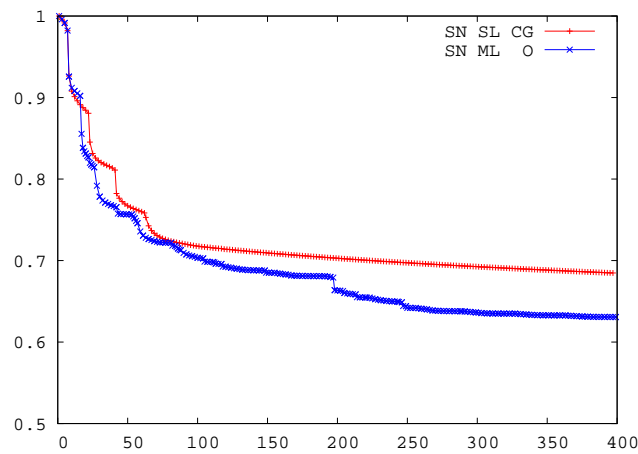


FIG. 4.36 – Évolution de la fonction coût – B-spline, FFD normale

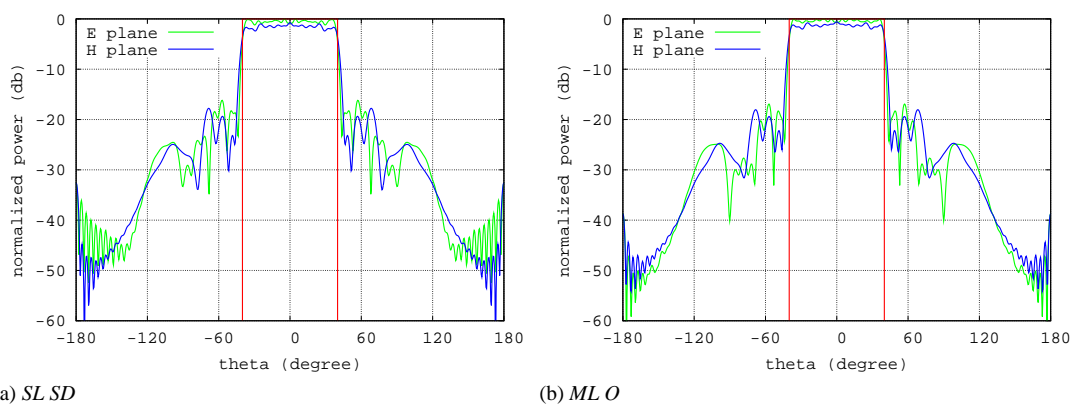


FIG. 4.37 – Diagrammes obtenus au bout de 400 évaluations de la fonction coût – B-spline, FFD normale

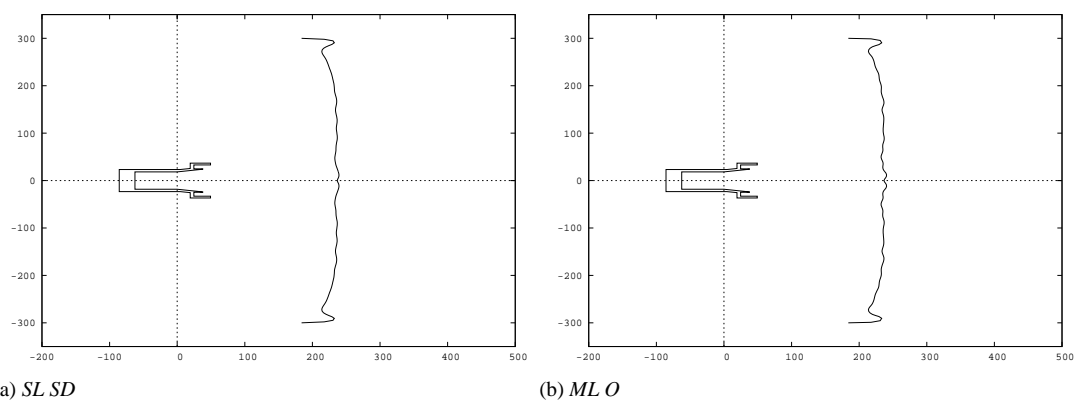


FIG. 4.38 – Formes convergées : représentation de l'antenne – B-spline, FFD normale

## Conclusion

Dans le chapitre précédent on a développé des stratégies multiniveaux pour traiter un problème modèle d'optimisation de forme. Dans cette section on a étendu ces stratégies pour le traitement de problèmes complexes. Des applications à la conception optimale d'antenne ont été réalisées. On a motivé nos choix algorithmiques en mettant en relation le problème modèle géométrique et le problème physique. Cela nous a permis de proposer une simplification des opérateurs de transfert basée sur une notion purement géométrique et non plus algébrique. Plus précisément, on a exploité le fait que les modes de haute fréquence convergent lentement. Puis on a abordé la résolution du problème en deux étapes : l'approche globale suivie de l'approche locale.

Dans la première phase on a vu qu'une stratégie d'enrichissement progressif de la paramétrisation permettait d'accroître la robustesse de l'algorithme semi-stochastique PSO et que les formes convergées étaient plus régulières. La spécificité de l'algorithme PSO dans cette stratégie vient du fait que l'information transmise par élévation du degré de la mémoire globale (c'est-à-dire du meilleur candidat sur un espace donné) est conservée et reste intacte sur le niveau supérieur ; au contraire, dans le cadre d'un algorithme génétique, il semble que l'information transmise par élévation du degré du meilleur individu est altérée ou « diluée » au cours des itérations (*générations*) par *croisement* et *mutation* (voir par exemple les tests effectués dans [3]).

Dans la deuxième phase on a vu qu'une stratégie s'inspirant des méthodes multigrilles et prenant en compte la non-linéarité du gradient de la fonction objectif permettait d'accélérer la convergence de méthodes de descente, pourvu qu'un transfert adéquat soit utilisé. Toutefois, l'intérêt de l'approche multiniveau locale ne paraît pas significatif devant l'efficacité du préconditionnement par une méthode de quasi-Newton.

Cela étant dit, les remarques suivantes peuvent être faites :

- dans les cas traités, la dimension de l'espace de recherche n'excède pas 30 d.d.l., ce qui reste raisonnable pour une stratégie quasi-Newton ; or, lorsque le nombre de d.d.l. augmente, les stratégies quasi-Newton s'avèrent coûteuses en mémoire et en temps (la matrice quasi-hessienne doit être gardée en mémoire et un système linéaire par itération doit être résolu) ; alternativement il existe des versions à mémoire limitée (LM-BFGS) [73, 71]. Il faudrait donc comparer ces méthodes aux stratégies multiniveaux sur des problèmes où la géométrie est complexe (3D).
- dans la mesure où la géométrie est déformée par un champ normal, les stratégies multiniveaux restent pertinentes puisque les méthodes de quasi-Newton ou du Gradient Conjugué n'ont pas de sens ; par analogie aux méthodes quasi-Newton on peut également envisager de formuler un préconditionnement du gradient en prenant en compte les variations des champs de vitesse de déformation entre deux itérations.
- si on ne connaît pas le gradient, ce qui est souvent le cas dans le cadre de la conception optimale de systèmes gouvernés par des EDP puisque son calcul n'est pas trivial, alors les stratégies multiniveaux pourraient améliorer les algorithmes de type simplexe qui s'apparentent à des méthodes de descente sans gradient.

Par ailleurs, on rappelle que l'on a traité les problèmes de multimodalité et de conditionnement dans deux contextes séparés. Il serait judicieux de traiter ces deux problèmes dans un même processus, hybride, tout en exploitant les différentes représentations hiérarchiques de la géométrie. À la vue des résultats précédents, on propose un schéma multiniveau de type FMG pour l'optimisation de forme paramétrique où l'on combinerait les techniques développées précédemment dans un même algorithme :

- en utilisant l'élévation du degré classique pour la robustesse dans un schéma d'enrichissement progressif ;
- en utilisant les opérateurs de transfert comme filtre passe-haut et un schéma en V ou en dent de scie pour la vitesse de convergence.

Dans un premier temps on résout complètement le problème sur le niveau grossier. Puis l'espace de recherche est enrichi pour raffiner l'approximation obtenue. Pour améliorer la convergence sur ce niveau on applique un schéma de correction de telle sorte que la grille grossière soit un espace « haute fréquence »

(méthode MLZ ou MLO). De nouveau l'espace de recherche est enrichi (au moyen d'un opérateur classique de prolongement). Enfin, un schéma de correction haute fréquence est appliqué. On illustre cet algorithme à la Figure 4.39 où on note  $[ \nearrow ]$  l'opérateur d'élévation classique et  $[ \searrow ]$ ,  $[ \swarrow ]$  les opérateurs de restriction et de prolongement haute fréquence respectivement.

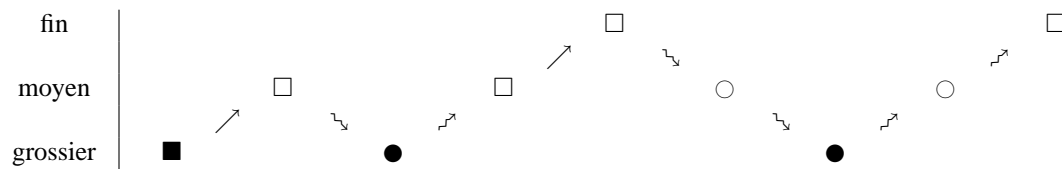


FIG. 4.39 – schéma FMG à trois niveaux de paramétrisation pour l'optimisation : les symboles pleins (■ et ●) correspondent à des phases de correction complète, les symboles vides (□ et ○) à des phases de relaxation. Les (□ et ■) sont des espaces d'enrichissement et les (○ et ●) des sous-espaces « haute fréquence ».



# OPTIMISATION MULTIPOINT SUR UNE BANDE DE FRÉQUENCE

---

<b>5.1</b>	<b>Généralités sur l'optimisation multicritère</b>	<b>137</b>
5.1.1	Formulation d'un problème d'optimisation multicritère	137
5.1.2	Optimalité au sens de Pareto	137
<b>5.2</b>	<b>Méthodologies d'optimisation multicritère</b>	<b>139</b>
5.2.1	Optimisation d'une combinaison convexe des objectifs	139
5.2.2	Minimisation du pire des cas	139
5.2.3	Recherche de front de Pareto	139
5.2.4	Cas particulier de l'optimisation multipoint	140
<b>5.3</b>	<b>Méthode hiérarchique sur les points de fonctionnement</b>	<b>141</b>
5.3.1	Description du cas-test	141
5.3.2	Diagramme initial sur une bande de fréquence	142
5.3.3	Conception optimale à une fréquence	142
5.3.4	Conception optimale sur une bande de fréquence	147
5.3.5	Méthode hiérarchique sur les fréquences	148
<b>Conclusion</b>		<b>158</b>

---

À une fréquence donnée correspond un canal de transmission. Le canal ou régime de fonctionnement d'une antenne est imposé par le contexte d'utilisation (télévision, radio, application militaire, etc.). En général une antenne n'est pas conçue pour transmettre des informations à un régime unique ; au contraire on cherche à la rendre efficace pour un ensemble de fréquences. On cherche alors à résoudre un problème d'optimisation dit multipoint qui est un cas particulier de l'optimisation multicritère. On propose dans ce chapitre une méthode hiérarchique sur les fréquences de fonctionnement pour l'optimisation d'une antenne sur une bande de fréquence.





## 5.1 Généralités sur l'optimisation multicritère

Dans cette section on rappelle les concepts fondamentaux de l'optimisation *multicritère* ou *multiobjectif* (MOP pour *Multiobjective Optimization Problem* dans la littérature anglophone) ainsi que les principales stratégies de résolution. Pour une étude rigoureuse de la formulation et des résultats de caractérisation dans le cas différentiable on renvoie à l'ouvrage [67]. On se concentre ici sur les méthodes qui se rapportent à la notion d'optimalité au sens de Pareto. Dans le chapitre suivant on abordera une autre manière de traiter un problème multicritère en faisant appel aux stratégies de jeux.

### 5.1.1 Formulation d'un problème d'optimisation multicritère

Soit  $N$  fonctions objectifs  $J_i : \mathcal{F} \subset \mathbb{R}^n \rightarrow \mathbb{R}$ . On considère le problème

$$\begin{aligned} \min_x \quad & \{J_1(x), J_2(x), \dots, J_N(x)\} \\ \text{t.q.} \quad & x \in \mathcal{F} \end{aligned} \tag{5.1}$$

où  $x$  est le vecteur de conception et  $\mathcal{F} \subset \mathbb{R}^n$  désigne l'espace admissible (éventuellement défini par un ensemble de contraintes d'égalité et/ou d'inégalité). Le sens du problème (5.1) est de minimiser simultanément tous les critères. Lorsque cela n'est pas possible, c'est qu'il existe un conflit entre au moins deux critères, on cherche alors à réaliser un « bon » compromis. Pour formaliser ce compromis, on introduit les concepts de Pareto qui reposent sur la relation d'ordre multicritère de *dominance*.

### 5.1.2 Optimalité au sens de Pareto

On désigne par  $x$  et  $y$  deux points de l'espace admissible de conception. On note  $J(x) = (J_1, \dots, J_N)^T$  le vecteur de performance.

**Définition 5.1.** On dit que  $J(x)$  domine  $J(y)$  (noté  $J(x) \preceq J(y)$ ) si  $x$  est au moins aussi performant que  $y$  pour tous les critères et strictement plus performant pour au moins un critère, c-à-d si

$$\begin{aligned} \forall j \in \{1, \dots, N\} \quad & J_j(x) \leq J_j(y) \\ \text{et} \quad & \\ \exists i \in \{1, \dots, N\} \quad & J_i(x) < J_i(y) \end{aligned} \tag{5.2}$$

**Définition 5.2.** L'ensemble de Pareto est formé des vecteurs de conception  $x$  tels que  $J(x)$  est non-dominé. Autrement dit,  $x$  appartient à l'ensemble de Pareto si pour tout  $y \in \mathcal{F}$  on a  $J(y) \not\preceq J(x)$ , c-à-d

$$\begin{aligned} \exists i \in \{1, \dots, N\} \quad & J_i(x) < J_i(y) \\ \text{ou} \quad & \\ \forall j \in \{1, \dots, N\} \quad & J_j(x) \leq J_j(y) \end{aligned} \tag{5.3}$$

On dit alors que  $x$  est Pareto-optimal.

**Définition 5.3.** Le front de Pareto est l'ensemble des vecteurs  $J(x)$  donné pour tout  $x$  de l'ensemble de Pareto (l'image par  $J$  de l'ensemble de Pareto).

On illustre à la Figure 5.1 la notion de dominance et d'ensemble non dominé pour un problème bicritère. À la Figure 5.2 on illustre le fait que le front de Pareto est borné par les minima de chacun des critères lorsqu'ils existent.

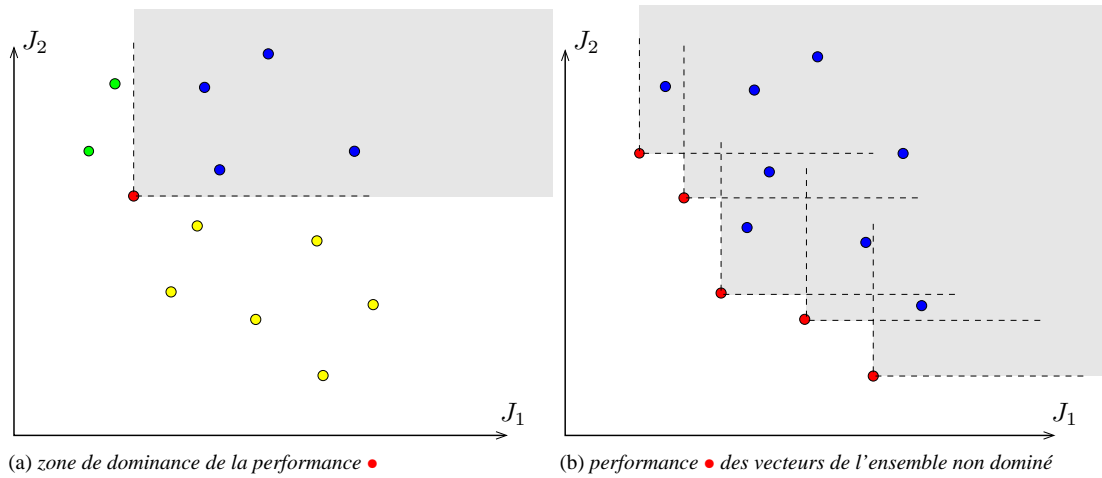


FIG. 5.1 – Dominance : on considère un ensemble de vecteurs de conception  $x$  pour lesquels on évalue les fonctions objectifs  $J_1$  et  $J_2$ . Chaque vecteur de performance  $(J_1(x), J_2(x))^T$  est représenté par un point dans l'« espace objectif »  $(\mathbb{R}^2)$ . À la Figure (a) on se place du point de vue de la performance ● : elle domine toutes les performances ●, les points ● sont plus performants par rapport au critère  $J_2$  et les points ● par rapport au critère  $J_1$ . À la Figure (b) on a identifié toutes les performances ● non dominées.

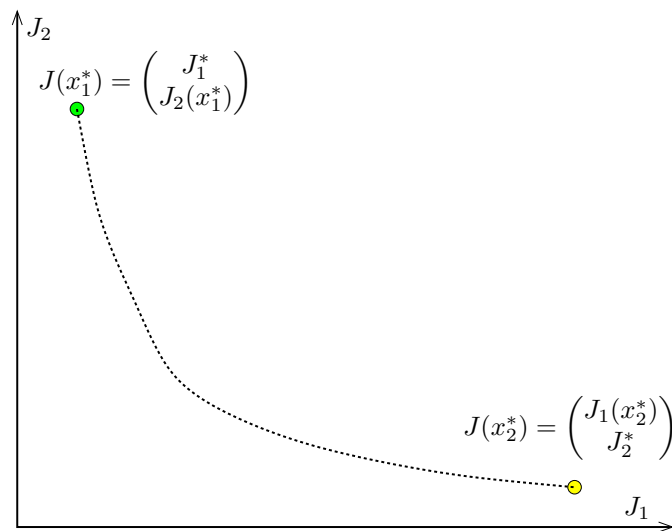


FIG. 5.2 – Exemple de front de Pareto pour deux critères : les minimums globaux  $J_1^*$  et  $J_2^*$  sont atteints en  $x_1^*$  et  $x_2^*$  respectivement (supposés uniques). Les performances  $J(x_1^*)$  et  $J(x_2^*)$  ne peuvent pas être dominées et forment donc les extrémités du front.

## 5.2 Méthodologies d'optimisation multicritère

Avec la notion d'optimalité de Pareto on a défini un critère mathématique de compromis. Sans critère supplémentaire, on est incapable de distinguer la qualité des points du front puisqu'ils sont tous mutuellement plus performants pour au moins un critère. La décision du choix d'un concept parmi l'ensemble non dominé reste en suspens. Si on prend en compte le problème compromis+décision, [97] propose une classification (d'après Fonseca et Fleming) des techniques de résolution d'un problème MO selon que la décision précède, interagit, ou suit la procédure d'optimisation.

Parmi les techniques où la décision précède l'optimisation, on considère classiquement des techniques d'agrégation comme une combinaison linéaire des fonctions objectifs (en particulier une combinaison convexe) ou une formulation minmax. La décision précède l'optimisation dans la mesure où le choix *a priori* des coefficients ou de la forme minmax détermine un critère dont on recherche le minimum global supposé unique. D'autres méthodes font appel explicitement à un cahier des charges en terme de valeurs cibles. On cherche alors à minimiser la fonction qui mesure l'écart à la cible dans l'espace des objectifs (par exemple *goal programming* [23] ou *goal attainment* [24]). En ce qui concerne les méthodes interactives, que l'on n'abordera pas ici, on réfère le lecteur à [15]. Enfin, la dernière classe de méthodes (et la plus populaire) consiste à rechercher en premier lieu une estimation du front de Pareto et de laisser le problème de décision à une procédure ultérieure.

### 5.2.1 Optimisation d'une combinaison convexe des objectifs

Soit le critère

$$J_c(x) \stackrel{\text{def.}}{=} \sum_{i=1}^N w_i J_i(x) \quad \text{tel que} \quad \sum_{i=1}^N w_i = 1 \quad \text{et} \quad \forall i \ w_i \geq 0 \quad (5.4)$$

où la valeur des coefficients reflète l'importance relative que l'on attribue à chaque critère  $J_i$ . La résolution du problème multiobjectif consiste alors à minimiser (5.4) à l'aide d'une méthode d'optimisation classique.

Pour une combinaison donnée, le point optimal de  $J_c$  est nécessairement Pareto-optimal. Cependant, l'ensemble des points Pareto-optimaux obtenus par minimisation de toutes les combinaisons convexes ne décrit pas l'ensemble de Pareto dans le cas général. En particulier, on montre que si le front n'est pas convexe alors certains points du front ne sont pas atteignables. Il est donc difficile de choisir le jeu de coefficients  $w_i$  qui assure que la performance obtenue à convergence d'un algorithme d'optimisation satisfasse le cahier des charges préalablement établi. Toutefois, on verra que cette formulation a un sens particulier dans le contexte d'un problème d'optimisation multipoint.

### 5.2.2 Minimisation du pire des cas

Une autre stratégie d'optimisation multicritère consiste à minimiser le maximum des performances individuelles :

$$J_\infty(x) = \max_{i=1, \dots, N} J_i(x) \quad (5.5)$$

On assure ainsi que toutes les performances sont bornées par la performance obtenue à convergence de l'algorithme. L'inconvénient majeur de cette méthode vient du fait que ce critère n'est pas différentiable et qu'il ne se prête donc pas aux méthodes de descente.

### 5.2.3 Recherche de front de Pareto

La majorité des algorithmes de recherche de front de Pareto fait appel aux Algorithmes Évolutionnaires (*Multiobjective Evolutionary Algorithms*, MOEA<sup>1</sup>). Le processus général consiste à guider la re-

<sup>1</sup>Une liste bibliographique régulièrement mise à jour concernant les algorithmes de type MOEA est tenue par Coello à l'adresse suivante : <http://www.lania.com.mx/~ccoello/EMOO/EMOObib.html>.

cherche vers le front de Pareto tout en assurant la diversité des individus et une répartition uniforme des performances sur le front. Pour cela on dispose d'un système de notation ou d'affectation d'un rang (*ranking*) en relation avec le front de Pareto. La manière la plus simple revient à affecter le rang 0 aux individus dont la performance n'est pas dominée et 1 aux autres. Des systèmes de notation plus élaborés sont souvent utilisés. Par exemple, Goldberg propose de considérer une succession de fronts en incrémentant le rang que l'on affecte aux points non dominés de la population à laquelle on a supprimé les éléments de rang inférieur. Fonseca et Fleming proposent un rang affecté à chaque élément de la population qui correspond au nombre de points dont la performance domine l'élément considéré (les points non dominés sont alors de rang 0).

Parmi les algorithmes les plus répandus on trouve les suivants (par ordre chronologique) :

- *Niched Pareto Genetic Algorithm* (NPGA) [51]
- *Nondominated Sorting in Genetic Algorithm* (NSGA) [92]
- *Multiobjective Genetic Algorithm* (MOGA) [44]
- *Strength Pareto Evolutionary Algorithm* (SPEA) [103]

**Remarque** On réfère généralement le premier MOEA au *Vector Evaluated Genetic Algorithm* (VEGA) [83]. Cette méthode ne fait toutefois pas appel aux notions de Pareto. Elle consiste dans un premier temps à diviser la population en autant de groupes qu'il y a d'objectifs. Les individus sont alors notés selon l'objectif attribué au groupe auquel ils appartiennent. Les meilleurs individus de chaque groupe sont sélectionnés puis croisés pour former une nouvelle génération, formant ainsi un brassage des objectifs.

### 5.2.4 Cas particulier de l'optimisation multipoint

Jusqu'à présent on s'est intéressé à l'optimisation d'une antenne dont le régime harmonique a été fixé à une fréquence donnée  $f^*$ . On a cherché à minimiser une fonction coût du type

$$\mathcal{J}(\mathcal{S}, f^*) = F(u(\mathcal{S}, f^*)) \quad (5.6)$$

où  $u(\mathcal{S}, f^*)$  est le champ rayonné par  $\mathcal{S}$  à la fréquence  $f^*$  et  $F$  une fonctionnelle champ rayonné (directivité, écart quadratique de la puissance, coefficient de réflexion, etc.).

Soit désormais la bande de fréquence  $I_f = [f^* - \frac{1}{2}\Delta f ; f^* + \frac{1}{2}\Delta f]$ . Si l'application  $\mathcal{J}$  est différentiable par rapport à  $f$  sur  $I_f$ , on peut naturellement chercher à minimiser

$$\min_{\mathcal{S}} \mathcal{J}(\mathcal{S}, f^*) + \alpha |\partial_f \mathcal{J}(\mathcal{S}, f^*)| \quad (5.7)$$

où  $\alpha$  est un coefficient de normalisation, c-à-d, on cherche à minimiser à la fois la fonctionnelle  $\mathcal{J}$  à une fréquence donnée ainsi que les variations de  $\mathcal{J}$  du premier ordre par rapport à ce paramètre. Toutefois, ce formalisme correspond davantage à l'optimisation robuste et reste valide tant que l'intervalle  $I_f$  est « petit » (la notion de « petit » étant liée à l'ordre de grandeur des variations du second ordre). Dans notre cas on sera amené à considérer de grands intervalles de fonctionnement. Alternativement, on adopte des stratégies dites *multipoints* où l'intervalle  $I_f$  est échantillonné en  $N$  points de fonctionnement  $f_1, f_2, \dots, f_N$  et à partir desquels on définit  $N$  critères  $\mathcal{J}_i(\mathcal{S}) \stackrel{\text{def.}}{=} \mathcal{J}_i(\mathcal{S}, f_i)$  à minimiser simultanément. En ce sens l'optimisation multipoint est un cas particulier de l'optimisation multicritère.

#### Minimisation en moyenne

Classiquement, on cherche à ce que la performance moyenne sur  $I_f$  soit minimisée, soit

$$\bar{\mathcal{J}}(\mathcal{S}) = \frac{1}{\Delta f} \int_{I_f} \mathcal{J}(\mathcal{S}, f) df \approx \frac{1}{\Delta f} \sum_i w_i \mathcal{J}(\mathcal{S}, f_i) \quad (5.8)$$

où les  $w_i$  sont les poids de quadrature affectés aux fréquences  $f_i$ . Cette approche revient donc à minimiser une combinaison convexe (5.4) des performances individuelles puisque les poids  $w_i$  sur  $I_f$  satisfont nécessairement  $\sum_i w_i = \Delta f$ .

**Remarque** On peut également se baser sur une formulation en norme  $L_p$  en augmentant progressivement  $p$  pour se rapprocher de la norme max tout en gardant la différentiabilité de la fonction coût

$$\mathcal{J}(\mathcal{S}) = \sum_i w_i \mathcal{J}(\mathcal{S}, f_i)^p \quad p \geq 2 \quad (5.9)$$

Cette méthode a été utilisée dans [10].

## 5.3 Méthode hiérarchique sur les points de fonctionnement

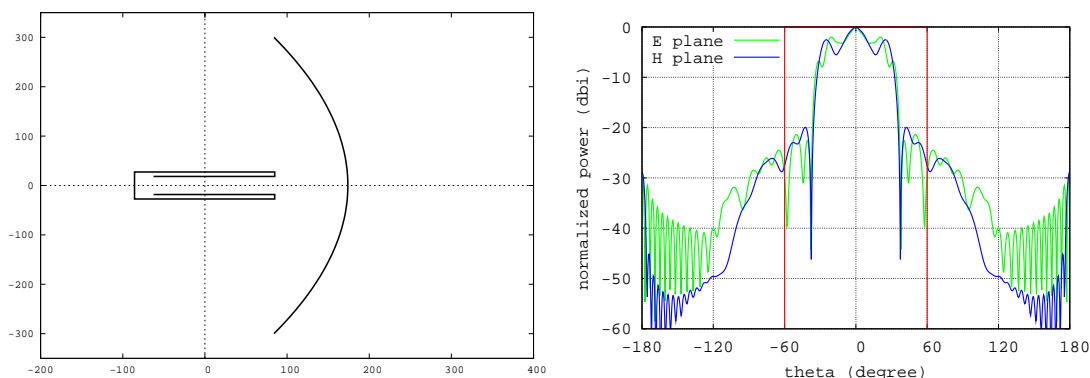
Dans cette section on propose une stratégie d'optimisation multipoint dans laquelle une hiérarchie des points de fonctionnement est utilisée. On cherche à résoudre le problème multicritère au sens d'une minimisation en moyenne. La stratégie hiérarchique consiste à raffiner progressivement le calcul de quadrature de la fonctionnelle (5.8).

### 5.3.1 Description du cas-test

On considère un problème inverse en directivité où  $\tilde{D}$  est une directivité idéale constante sur l'intervalle  $\theta \in [-60^\circ; 60^\circ]$  et nulle ailleurs (voir Figure 5.3). On cherche à minimiser l'écart quadratique entre la directivité réelle et la directivité idéale

$$J(x) = \frac{1}{2} \int_{S^2} \left( D(x, \hat{\mathbf{x}}) - \tilde{D}(\hat{\mathbf{x}}) \right)^2 dS^2 \quad (5.10)$$

sur la bande de fréquence  $I_f = [f^* - \frac{1}{2}\Delta f; f^* + \frac{1}{2}\Delta f]$  où  $f^* = 7.133333$  GHz et  $\Delta f = 1.0$  GHz.



(a) géométrie initiale : guide d'onde et réflecteur.

(b) diagramme initial (plans E et H) et diagramme cible (rouge).

FIG. 5.3 – Configuration initiale de l'antenne à optimiser : le guide d'onde est de section uniforme et le réflecteur est un paraboloïde défocalisé ; le diagramme est évalué à la fréquence  $f^* = 7.133333$  GHz.

La cible est une distribution idéale de la directivité qui en réalité n'est pas atteignable : ce problème est mal posé. En pratique, on considère qu'un diagramme est satisfaisant si la différence entre la puissance rayonnée dans la zone où l'on cherche à uniformiser le diagramme et la puissance maximale atteinte en  $\theta_{\max}$  n'excède pas 3 db (c-à-d une division de la puissance par 2). L'angle  $\theta_{3db}$  correspond à l'angle à partir duquel la puissance est inférieure à 3 db. On cherche donc à obtenir un diagramme tel que  $\theta_{3db} = 60^\circ$ . En plus de mesurer l'angle à 3 db pour les plans E et H on considère le plan de polarisation normale (ou plan N) d'après la 3<sup>e</sup> définition de Ludwig [60]. En cherchant à uniformiser le diagramme on cherche à minimiser l'amplitude des oscillations (la fonctionnelle (5.10) pénalise les variations du diagramme). On mesure donc l'amplitude maximale  $\Delta P$  en db de ces oscillations sur le plan N.

### 5.3.2 Diagramme initial sur une bande de fréquence

En anticipant sur la suite, évaluons les diagrammes de rayonnement de la forme initiale en quatre fréquences de  $I_f$ . Ces diagrammes sont représentés à la Figure 5.4. D'un point de vue qualitatif, le diagramme cible est très éloigné des diagrammes de la forme initiale : la puissance entre  $40^\circ$  et  $60^\circ$  est d'environ -20 db par rapport à la puissance maximale.

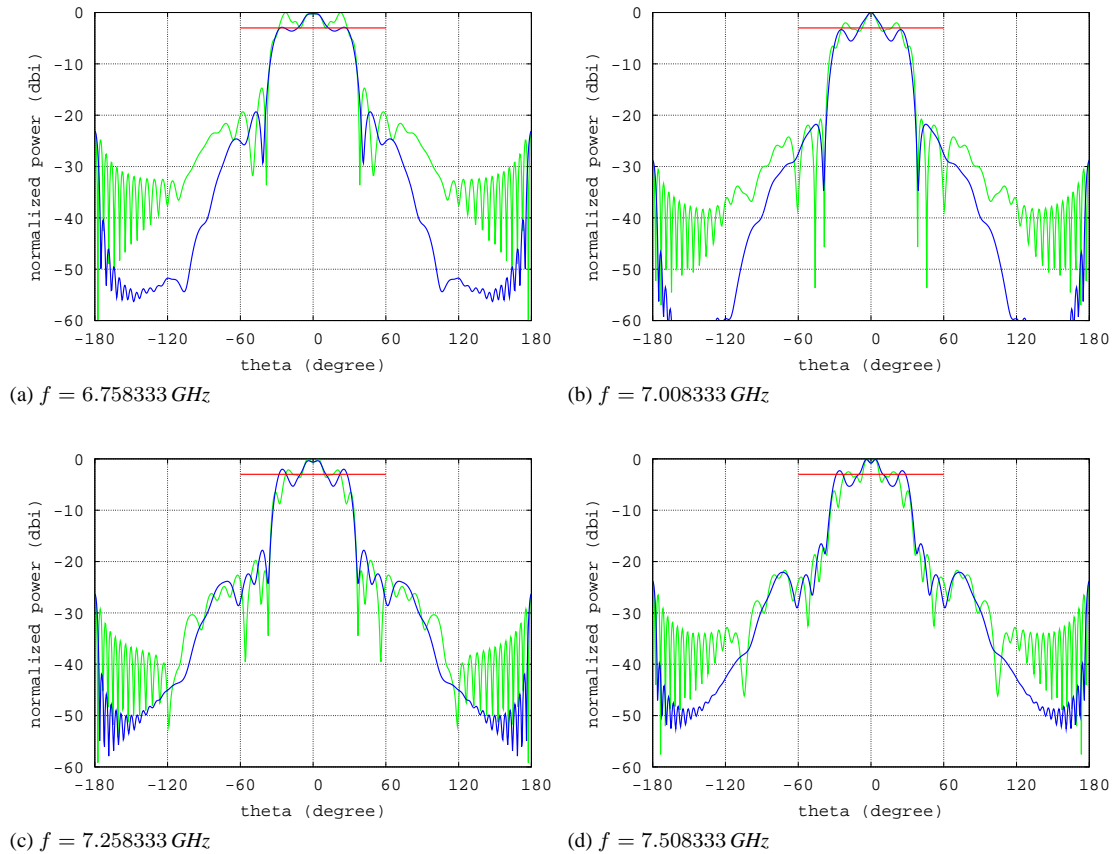


FIG. 5.4 – Diagrammes de rayonnement de l'antenne initiale évalués en quatre fréquences de fonctionnement uniformément réparties sur  $I_f$ .

### 5.3.3 Conception optimale à une fréquence

Dans un premier temps on conduit une expérimentation numérique d'optimisation à une fréquence donnée. Une première phase d'optimisation est réalisée par la méthode MPSO. L'approximation obtenue est améliorée par une phase d'optimisation de descente de type quasi-Newton BFGS. On impose comme contrainte la continuité du réflecteur au niveau de l'axe de symétrie. La fréquence de fonctionnement est fixée à  $f^*$ .

#### Phase d'optimisation robuste – MPSO

Les paramètres sont donnés dans le Tableau 5.1. La convergence de l'algorithme est illustrée à la Figure 5.5 et la configuration de l'antenne à convergence est donnée à la Figure 5.6

À première vue, le diagramme de rayonnement des plans E et H semble satisfaisant : la différence maximale de puissance dans l'intervalle  $[-60^\circ; 60^\circ]$  est de l'ordre de 3 db. Sauf pour le plan E, l'angle  $\theta_{3db}$  est proche de  $60^\circ$  (voir Tableau 5.2). La directivité est de 7.93 dbi.

Fréquence	$f = 7.133333$ GHz
Paramétrisation	Bernstein
Hiérarchie	$n = 3 \rightarrow 6 \rightarrow 9 \rightarrow 12 \rightarrow 15$
Nombre d'itérations	$i_{\max} = 20 \rightarrow 30 \rightarrow 40 \rightarrow 50 \rightarrow 60$
Nombre de particules	64

TAB. 5.1 – Paramètres de la méthode MPSO.

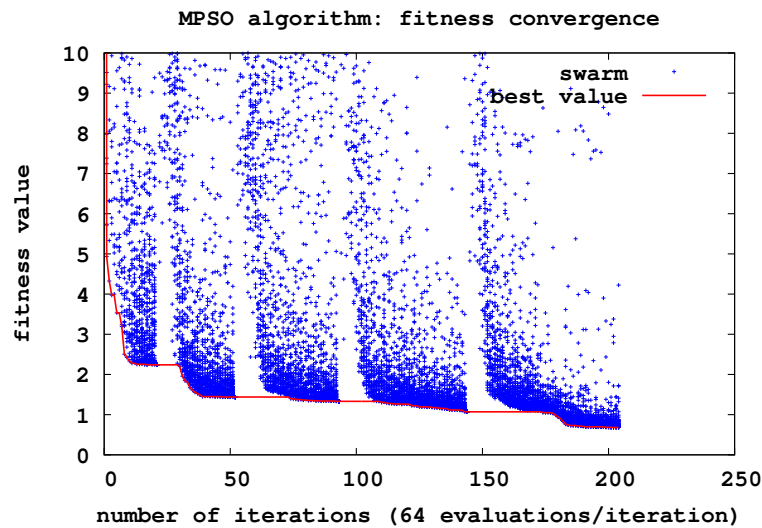
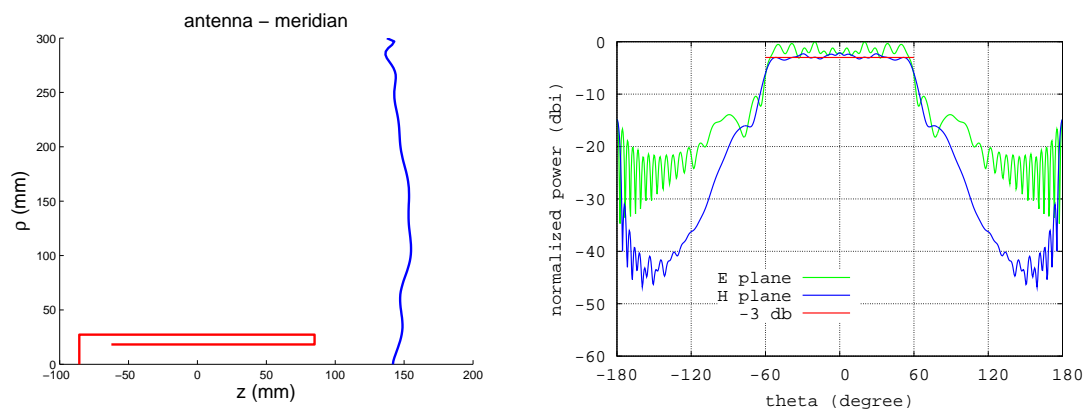


FIG. 5.5 – Convergence de l'algorithme MPSO.



(a) méridienne

(b) diagramme

FIG. 5.6 – Forme et diagramme à convergence de l'algorithme MPSO.

plan	$\theta_{\max}$ (°)	$\theta_{3db}$ (°)	$\Delta P$ (db)
E	20.0	32.0	3.12
H	0.0	58.5	1.37
N	20.5	57.5	1.51

TAB. 5.2 – Caractéristiques du rayonnement à l'issue de la phase d'optimisation robuste.



### Phase d'optimisation locale – QN-BFGS

On cherche à améliorer la performance obtenue sur la paramétrisation fine à l'aide d'un algorithme de quasi-Newton BFGS. La convergence en terme de réduction de la fonction coût et de la norme du gradient est illustrée à la Figure 5.7. En 50 itérations la fonctionnelle est diminuée de 14% et la norme du gradient de 90%. La forme convergée et son diagramme de rayonnement sont illustrés à la Figure 5.8. D'après les caractéristiques du diagramme obtenu à convergence de la procédure d'optimisation locale (voir Tableau 5.3) l'antenne est satisfaisante. La directivité est de 7.32 dBi.

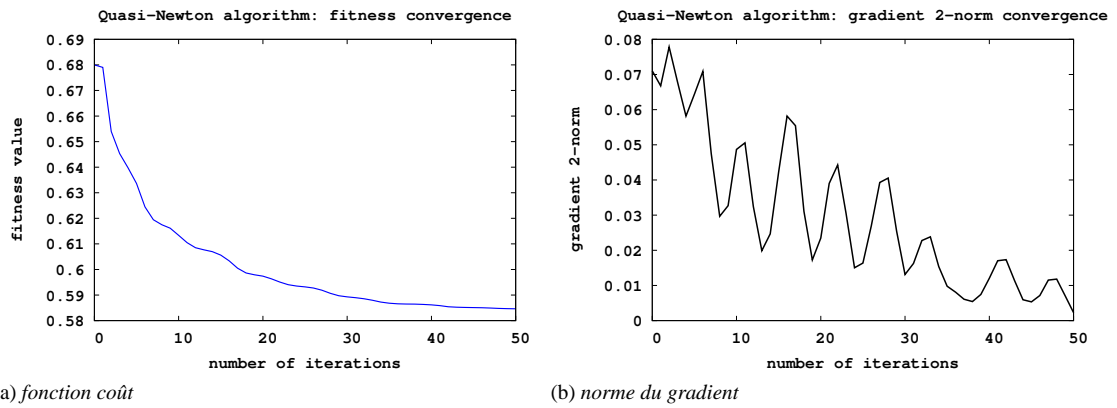


FIG. 5.7 – Convergence de l'algorithme QN/BFGS.

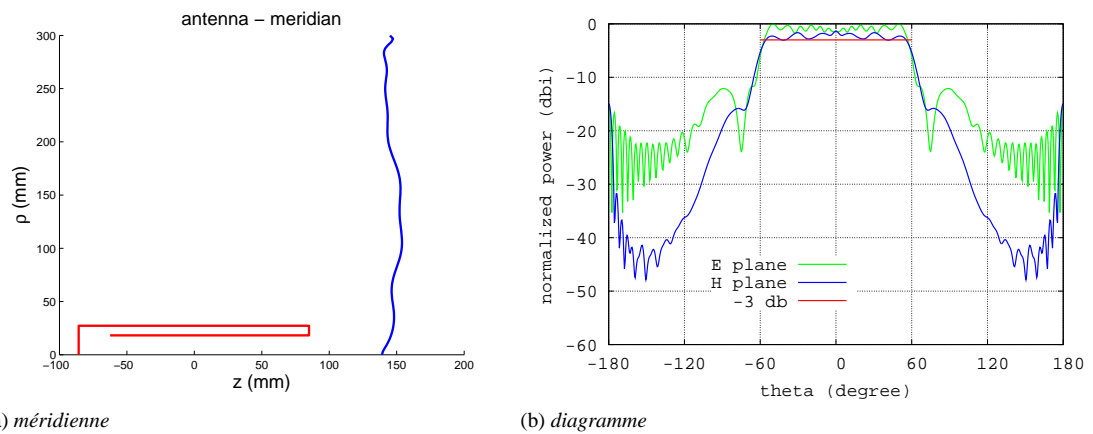


FIG. 5.8 – Forme optimisée à la fréquence  $f^*$  et diagramme de rayonnement.

plan	$\theta_{\max}$ (°)	$\theta_{3db}$ (°)	$\Delta P$ (db)
E	50.0	56.5	1.75
H	0.0	58.5	1.68
N	50.0	58.0	0.91

TAB. 5.3 – Caractéristiques du rayonnement à l'issue de la phase d'optimisation locale.

### Qualité de la forme convergée sur la bande de fréquence

On vient de réaliser une optimisation numérique de la forme du réflecteur à une fréquence donnée. Cela nous a permis d'obtenir une antenne qui satisfait la contrainte de qualité que l'on s'est fixée. On évalue la performance de cette configuration sur la bande de fréquence  $I_f$  : pour cela on calcule les valeurs de la fonction coût sur tout l'intervalle, puis on observe les diagrammes de rayonnement de la forme optimisée en certains points.

On a déjà constaté à la Figure 5.4 que l'antenne initiale était mauvaise en quatre points de  $I_f$ . L'évaluation de la fonction coût sur toute la bande de fréquence confirme la mauvaise performance de cette configuration (cf. Figure 5.9a). La Figure 5.9b montre que l'optimisation de l'antenne à une fréquence améliore significativement la performance sur tout l'intervalle : la performance la plus mauvaise est 25 fois plus petite que la plus mauvaise des performances de l'antenne initiale. En outre, les performances les plus mauvaises sont obtenues pour les fréquences les plus éloignées de la fréquence d'optimisation.

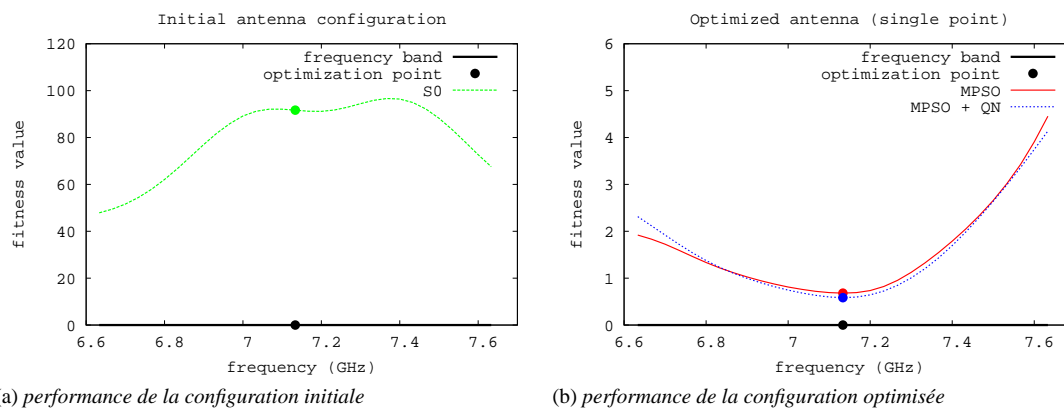


FIG. 5.9 – Valeur de la fonction coût au voisinage de la fréquence d'optimisation.

Les diagrammes de la Figure 5.10 confirment ce résultat : globalement les diagrammes sont plus proches de la cible que les diagrammes initiaux. Ceux obtenus pour des fréquences proches de la fréquence d'optimisation semblent satisfaisants (Figures 5.10b et 5.10c), contrairement à ceux les plus éloignés (Figures 5.10a et 5.10d). On trouve dans le Tableau 5.4 un résumé des caractéristiques du rayonnement. La procédure d'optimisation pour la seule fréquence  $f^*$  peut fournir un point de départ pour l'optimisation sur un voisinage de  $f^*$ .

fréquence (GHz)	$\Delta P$ (db)	plan N	directivité (dbi)
6.758333	5.45		8.88
7.008333	2.01		7.36
7.258333	2.51		7.54
7.508333	5.08		10.05

TAB. 5.4 – Caractéristiques du rayonnement sur l'intervalle  $I_f$  à l'issue de la phase d'optimisation à une fréquence.

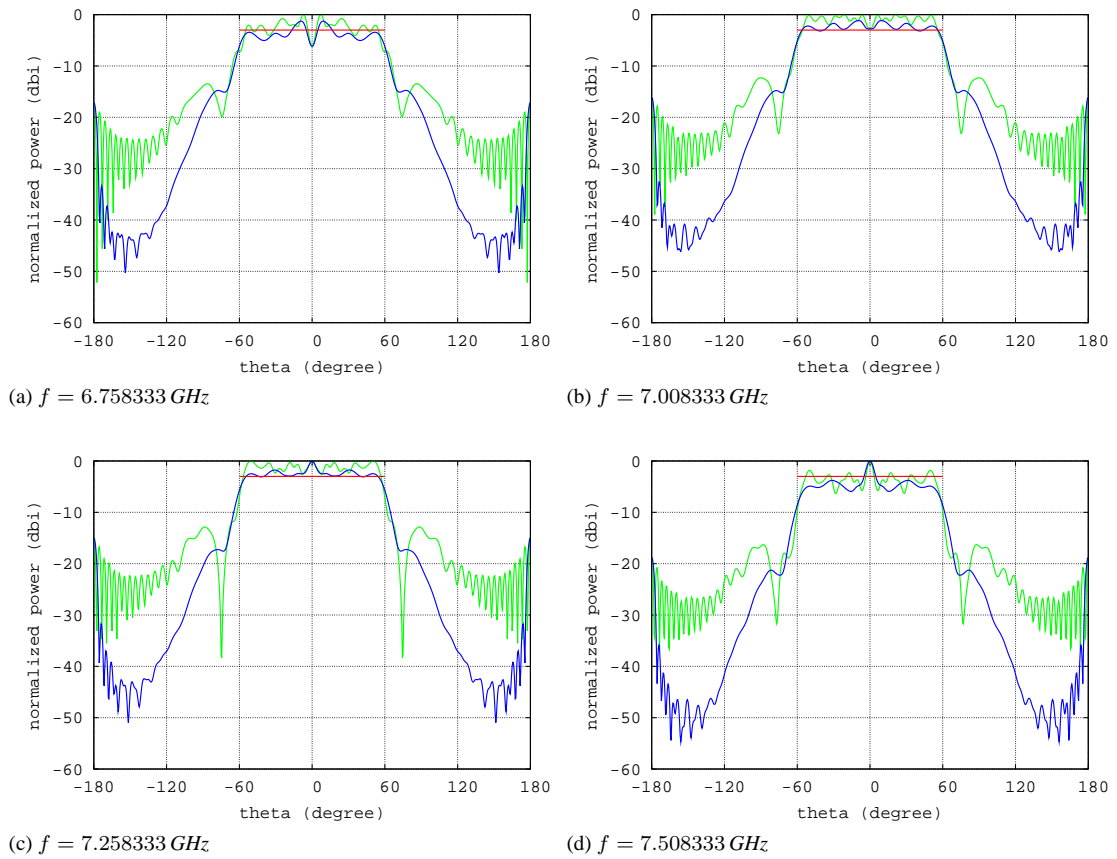


FIG. 5.10 – Diagrammes de rayonnement de l'antenne optimisée à la fréquence  $f^*$  évalués en quatre fréquences uniformément réparties sur  $I_f$ .

### 5.3.4 Conception optimale sur une bande de fréquence

Dans les expériences qui suivent on considère la formulation en moyenne (5.8) avec une quadrature de Gauss-Legendre à  $N = 9$  points. L'algorithme MPSO est appliqué à partir de la forme optimisée à la fréquence  $f^*$  (cf. Figure 5.8). La convergence et la forme obtenue sont illustrées à la Figure 5.11. On évalue *a posteriori* les performances individuelles de la forme convergée sur un échantillon de  $I_f$  beaucoup plus fin que celui de la quadrature. Elles sont comparées sur la Figure 5.12 avec les performances individuelles de la forme optimisée à la fréquence  $f^*$ .

Il n'est pas surprenant de constater que la performance est dégradée dans un voisinage de  $f^*$  alors qu'elle est améliorée ailleurs. Plus précisément, la meilleure performance est dégradée d'un facteur 1.57 (à la fréquence  $f = f^*$ ), alors que le pire des cas est amélioré d'un facteur 3.70 (à la fréquence  $f = 7.633333$  GHz).

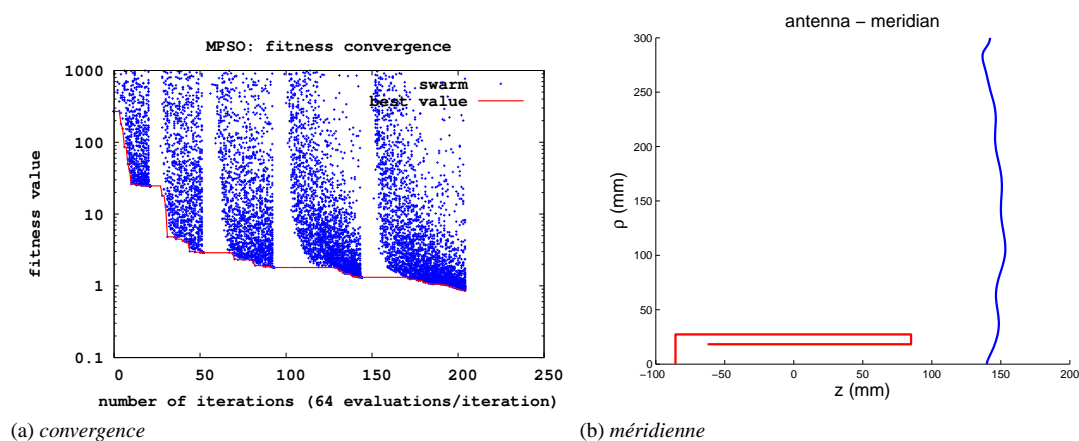


FIG. 5.11 – Convergence de l'algorithme MPSO : évolution de la fonction multicritère et forme convergée.

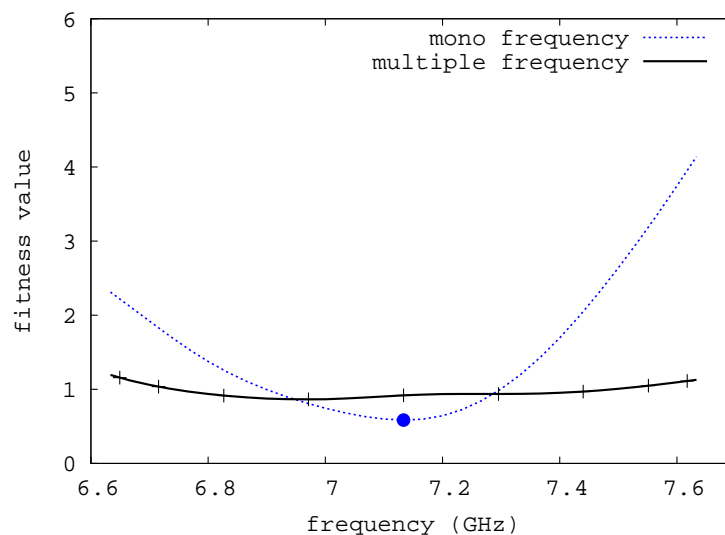


FIG. 5.12 – Performance monofréquence sur  $I_f$  après optimisation multipoint par la méthode MPSO de la forme optimisée à la fréquence  $f^*$  (●). Les + correspondent aux fréquences d'optimisation (points de Gauss).

### 5.3.5 Méthode hiérarchique sur les fréquences

Dans la section précédente, une phase d'optimisation classique à une fréquence a été réalisée afin de pourvoir un bon point de départ pour une optimisation multifréquence. On se demande s'il ne serait pas plus efficace, en terme de robustesse et de temps de calcul, d'étendre cette stratégie à plusieurs niveaux de quadrature : en s'inspirant de l'algorithme d'enrichissement progressif on propose de raffiner le calcul de la fonction objectif multifréquence (5.8) à l'aide d'une hiérarchie de points de fonctionnement.

Dans un premier temps, sur le niveau grossier de la paramétrisation, on compare les résultats d'optimisation pour plusieurs finesses de quadrature. Puis on réalise une optimisation hiérarchique en fréquence sur ce même niveau de paramétrisation. Enfin on propose un algorithme hiérarchique à la fois sur les points de fonctionnement et sur la paramétrisation.

#### Optimisations indépendantes pour plusieurs niveaux de quadrature

Soit  $w = \max_{i=1}^N w_i$  la taille (ou poids) caractéristique d'une quadrature quelconque à  $N$  points pour l'approximation de l'intégrale d'une fonction de la variable réelle sur un intervalle. À chaque quadrature correspond la fonction coût multicritère

$$\bar{\mathcal{J}}_w(\mathcal{S}) \stackrel{\text{def.}}{=} \sum_{i=1}^N w_i \mathcal{J}(\mathcal{S}, f_i) \quad (5.11)$$

dont la précision d'approximation de  $\bar{\mathcal{J}}$  (5.8) augmente avec le nombre de points. Lorsque  $N \rightarrow \infty$  on a  $w \rightarrow 0$  et donc  $\bar{\mathcal{J}}_w \rightarrow \bar{\mathcal{J}}$ . On espère alors que le minimum approché de  $\bar{\mathcal{J}}_w$  tende vers le minimum de  $\bar{\mathcal{J}}$ .

Dans cet exercice préliminaire on réalise une série d'optimisations numériques indépendantes pour les quadratures de Gauss à 1, 3, 5 et 9 points et sur le seul niveau grossier de la paramétrisation ( $n = 3$ ). On réalise 20 itérations de PSO. Par la suite, on comparera le résultat obtenu sur le niveau fin de la quadrature avec une méthode hiérarchique sur les fréquences. On représente sur la Figure 5.14 la convergence de l'algorithme et sur la Figure 5.13 la performance des formes optimisées en terme de performance individuelle sur  $I_f$ .

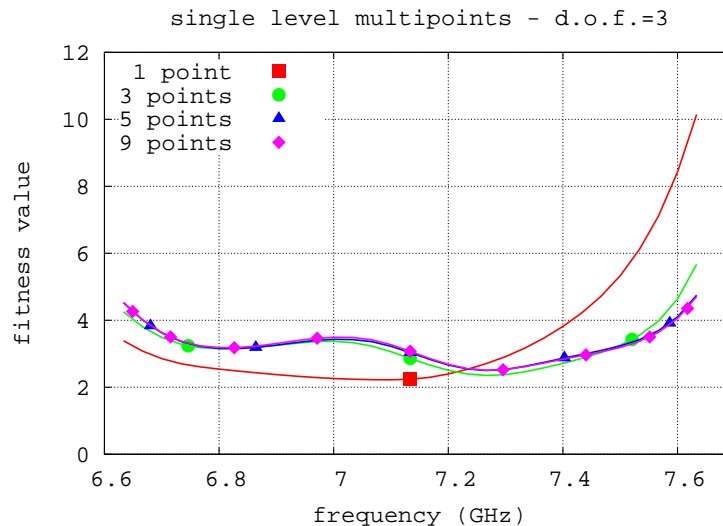


FIG. 5.13 – Performance de la forme optimisée par une méthode PSO sur le niveau grossier de la paramétrisation.

Au bout de 20 itérations, quelle que soit la quadrature, l'algorithme PSO peine à améliorer significativement le meilleur candidat. L'essentiel de la décroissance est réalisé en une dizaine d'itérations. Autrement dit, avec peu de degrés de liberté (3 d.d.l.), le problème est simple à résoudre.

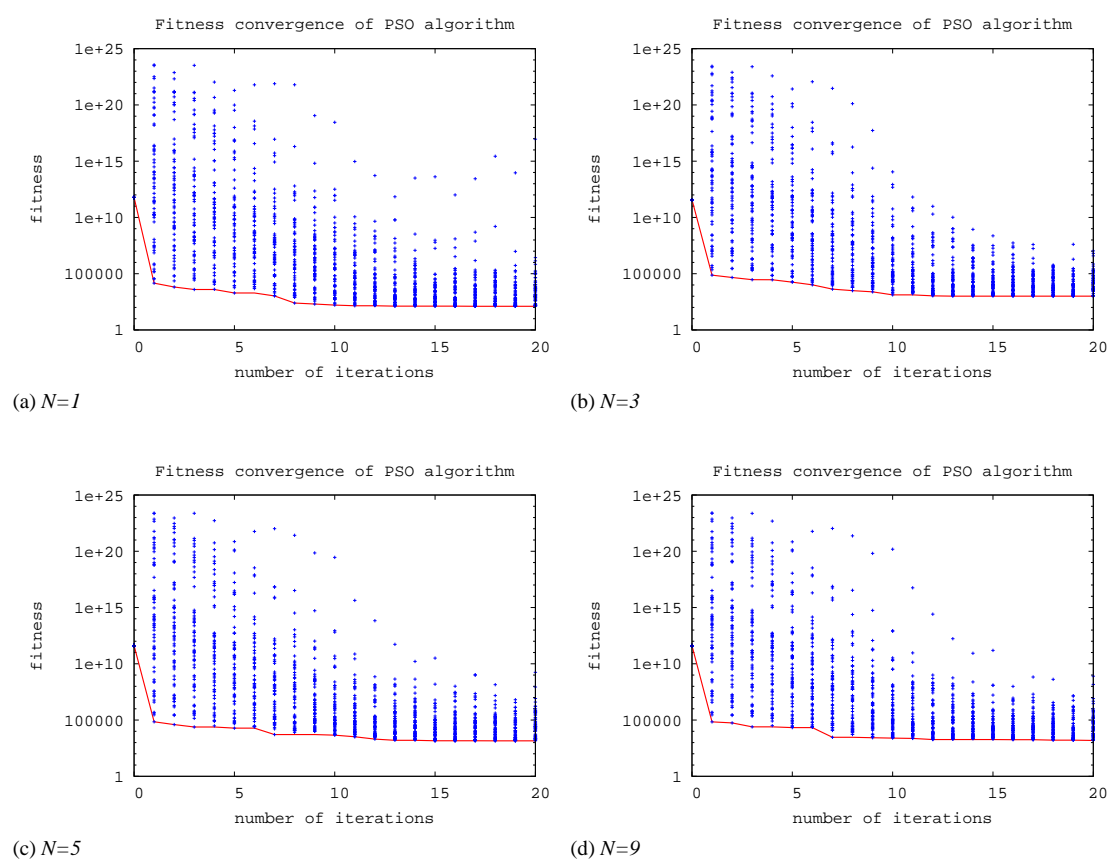


FIG. 5.14 – Convergence de l'algorithme PSO pour chaque niveau de quadrature : 20 itérations suffisent pour que l'algorithme converge sur le niveau grossier de la paramétrisation.

Les performances des formes convergées pour 3, 5 et 9 points de quadrature semblent identiques. La performance à un seul point de quadrature est plus mauvaise de 7.2 GHz à 7.6 GHz mais meilleure de 6.6 GHz à 7.2 GHz. Sachant qu'une évaluation de la fonction (5.11) est d'autant plus coûteuse qu'il y a de points, on se demande si la forme convergée à une fréquence, compte tenu de l'erreur commise sur le calcul de l'intégrale, peut tout de même fournir une bonne approximation (de coût négligeable) pour une optimisation à 3 points, puis 5 points et ainsi de suite.

### Hierarchie de points de fonctionnement sur le niveau grossier de la paramétrisation

On établit désormais une hiérarchie de quadrature de précision croissante à  $K + 1$  niveaux (cf. Figure 5.15). Soit  $N_k$  le nombre de points sur le niveau  $k$  :

$$N_k = 2^k + 1 \quad k = 0 \dots K \quad (5.12)$$

et on note  $w^k$  la taille caractéristique de la quadrature sur le niveau  $k$ . Le niveau 0 est une quadrature à un point : le point milieu  $f^*$ .

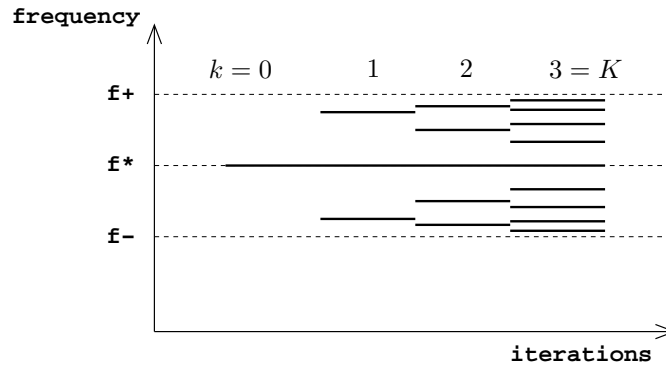


FIG. 5.15 – Illustration d'une hiérarchie de points de fonctionnement à  $K + 1 = 4$  niveaux ( $1 \rightarrow 3 \rightarrow 5 \rightarrow 9$ ) sur la bande de fréquence  $[f^- f^+]$  avec  $f^* = \frac{f^- + f^+}{2}$ . Les points de quadrature sont les points de Gauss.

On considère une paramétrisation de degré  $n$  fixé et les fonctionnelles paramétriques  $\bar{J}_{w^k}(x)$  associées. La stratégie hiérarchique consiste en une suite de phases de relaxation pour la minimisation successive des fonctionnelles paramétriques (cf. Algorithme 7).

---

#### Algorithme 7 : Optimisation multipoint avec stratégie d'enrichissement progressif

---

**Entrées :** degré fin  $K$ , concept initial  $x \in \mathbb{R}^n$

$\Delta x_0^{(0)} \leftarrow \mathbf{0}$

**pour**  $k \leftarrow 0$  à  $K$  **faire**

$\Delta x^{(k)} \leftarrow$  relaxation de  $\arg \min_{\Delta x} \bar{J}_{w^k}(x + \Delta x)$  à partir de  $\Delta x_0^{(k)}$

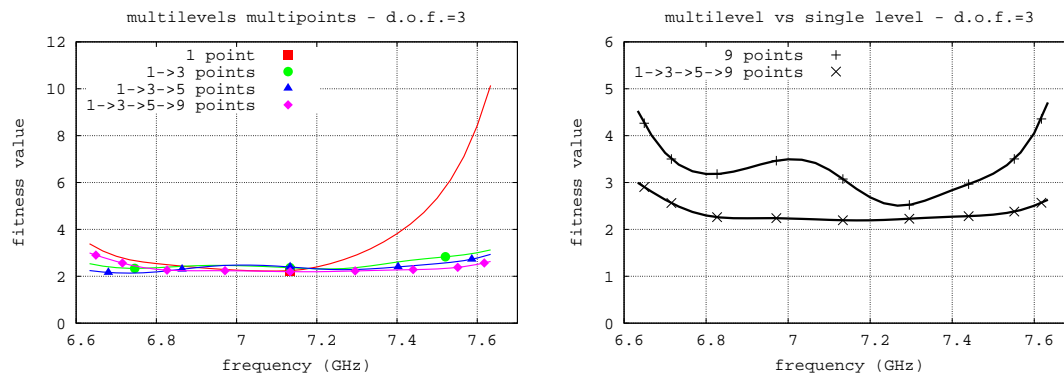
$\Delta x_0^{(k+1)} \leftarrow \Delta x^{(k)}$

**fin**

**Sorties :**  $\Delta x^{(K)}, J_{w^K}(x + \Delta x^{(K)})$

---

En guise de relaxation on réalise 10 itérations de l'algorithme PSO, ce qui revient à  $10 \times N_k \times N_{part}$  évaluations de la fonction coût où  $N_{part}$  est le nombre de particules constituant l'essaim. On représente à la Figure 5.16a la performance obtenue à chaque niveau de quadrature. À la Figure 5.16b on compare la performance obtenue avec la méthode hiérarchique et celle obtenue avec directement 9 points de quadrature. Il est clair que la méthode hiérarchique est plus efficace : toutes les performances individuelles sont améliorées.



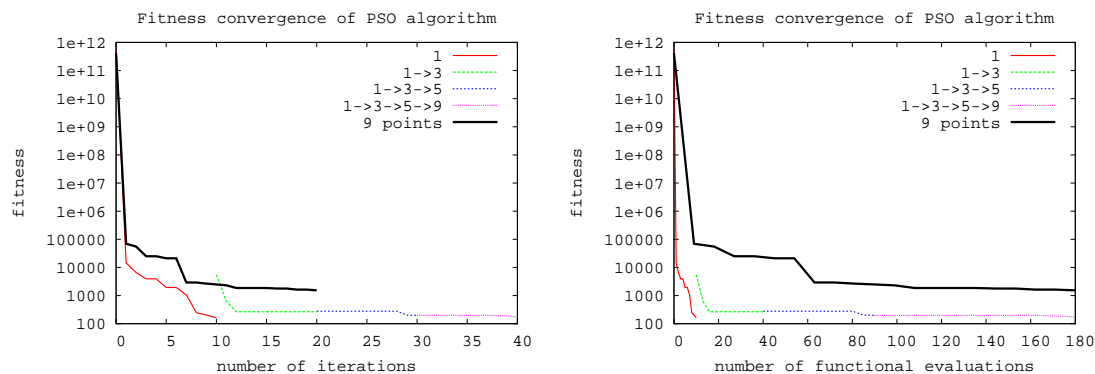
(a) performance à l'issue de la relaxation sur chaque niveau

(b) performance multiniveau en fréquence et simple niveau

FIG. 5.16 – Performance globale de la méthode d'enrichissement progressif en fréquence sur le niveau grossier de paramétrisation.

On représente à la Figure 5.17 la convergence en terme d'évolution de la fonction coût. On remarque que :

- le nombre d'itérations de relaxation est plus grand pour la méthode hiérarchique (Figure 5.17a) ; cependant les itérations sur les niveaux grossiers de quadrature sont peu coûteuses relativement au niveau fin ; à nombre égal d'évaluations de fonction coût, la méthode hiérarchique est plus performante (Figure 5.17b) ;
- en général il n'y a pas continuité des performances lorsque l'on passe d'un niveau de quadrature à un autre (sinon cela voudrait dire qu'une quadrature de degré  $k$  est aussi précise qu'une quadrature de degré  $k + 1$ ) ; en ce sens les problèmes grossiers ne sont pas consistants avec le problème fin ; toutefois, le saut de la performance entre deux niveaux diminue lorsque le nombre de points augmente (l'erreur commise dans le calcul de l'intégrale diminue).



(a) évolution de la performance en fonction du nombre d'itérations de PSO

(b) évolution de la performance en fonction du nombre d'évaluations de la fonction par nombre de particules

FIG. 5.17 – Évolution de la fonction coût des méthodes classique et d'enrichissement progressif en fréquence sur le niveau grossier de paramétrisation.

### Stratégie d'enrichissement progressif de la paramétrisation et de la quadrature

Sur le niveau grossier de paramétrisation, une méthode de raffinement du calcul de la fonction coût à l'aide d'une hiérarchie de points de fonctionnement nous a fourni une géométrie plus performante qu'avec une stratégie à un seul niveau de quadrature. En comparaison avec la performance obtenue sur le niveau fin de paramétrisation et de quadrature (cf. Figure 5.12) la performance globale est toutefois encore loin d'être satisfaisante. Il nous faut donc raffiner la déformation du réflecteur : la forme obtenue



sur le niveau grossier est choisie comme point de départ pour l'optimisation sur un niveau intermédiaire de paramétrisation.

Comme pour la paramétrisation grossière, on propose d'appliquer le même processus hiérarchique en fréquence sur ce niveau de paramétrisation dans le but de fournir une forme initiale performante pour le niveau suivant, et ainsi de suite jusqu'au niveau fin. L'ensemble de cette procédure hiérarchique d'enrichissement progressif en fréquence et en paramétrisation conduit à l'Algorithme 8 illustré à la Figure 5.18.

---

**Algorithme 8** : Stratégie d'enrichissement progressif : paramétrisation et quadrature

---

**Entrées** : degré fin de quadrature  $K$ , hiérarchie de paramétrisations  $n_1 < n_2 < \dots < n_M$

`elevation` : routine d'élévation du degré pour la paramétrisation considérée

`mfreq` : routine de relaxations successives sur une hiérarchie de quadrature (Algorithme 7)

`imax` : nombre d'itérations de relaxation

// Initialisation sur le niveau grossier

$x^{(1)} \leftarrow \mathbf{0} \in \mathbb{R}^{n_1}$

// Enrichissement de la paramétrisation

**pour**  $i = 1$  à  $M$  **faire**

    // Enrichissement de la quadrature

$\Delta x^{(i)} \leftarrow \text{mfreq}(K, x^{(i)}, \text{imax})$

    // Correction et élévation du degré

$x^{(i)} \leftarrow x^{(i)} + \Delta x^{(i)}$

**si**  $i < M$  **alors**  $x^{(i+1)} \leftarrow \text{elevation}(x^{(i)}, n_i \rightarrow n_{i+1})$

**fin**

**Sorties** :  $x^{(M)}$ ,  $\bar{J}_{wK}(x^{(M)})$

---

### Application numérique

Soit les degrés de paramétrisation  $3 \rightarrow 6 \rightarrow 9 \rightarrow 12 \rightarrow 15$ . À la Figure 5.19a-e, pour chaque niveau de paramétrisation, on représente la performance obtenue à chaque étape de la hiérarchie de quadrature. Sur la Figure 5.19f on représente la performance obtenue à l'issue de la hiérarchie de quadrature pour chaque degré de paramétrisation. Si la phase d'optimisation à une fréquence a tendance à dégrader la performance aux bornes de l'intervalle  $I_f$ , elle est cependant uniformément améliorée sur le niveau fin ( $K$  points de quadrature) pour chaque niveau de paramétrisation.

En comparaison avec la stratégie simple niveau de quadrature, pour une quantité équivalente d'évaluations de la fonction coût, la méthode multiniveau est plus performante (cf. Figure 5.20).

### Diagrammes et formes convergées

Enfin, on compare les formes (cf. Figure 5.21) et les diagrammes (cf. Figures 5.22 et 5.23) des antennes obtenues à convergence des méthodes hiérarchiques et simple niveau pour 4 fréquences données. Les caractéristiques du rayonnement (voir Tableau 5.5 et 5.6) confirment la différence de performance des deux stratégies : d'après la différence maximale de puissance sur le plan de polarisation normale et dans la zone optimisée, le rayonnement de la forme obtenue par la méthode hiérarchique est plus petite.

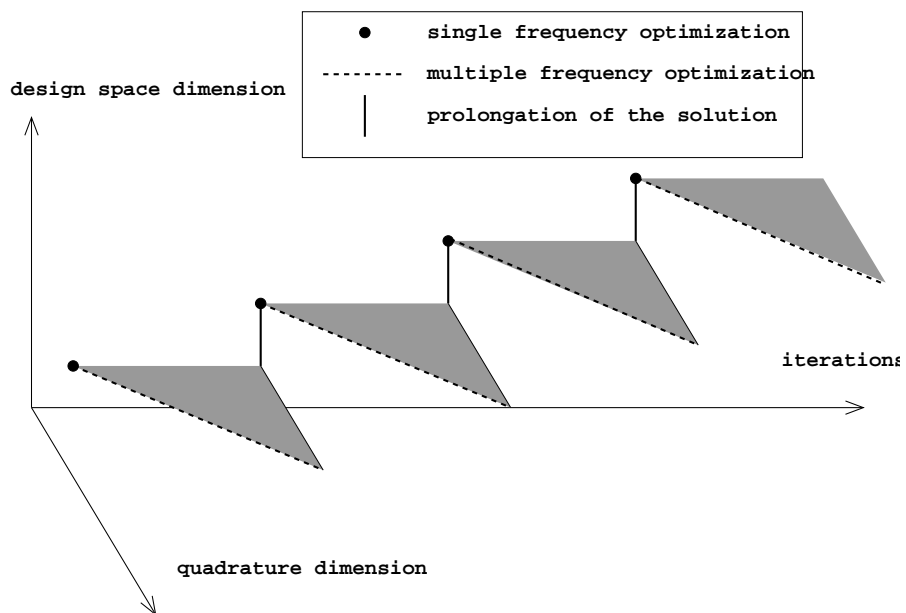
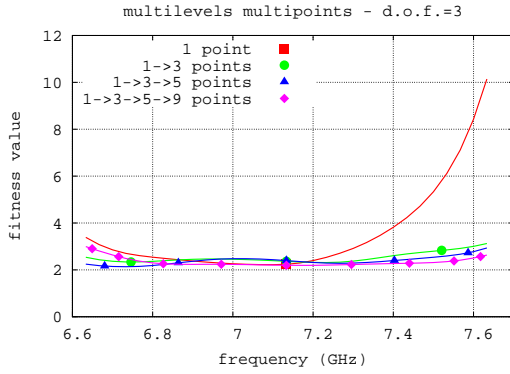
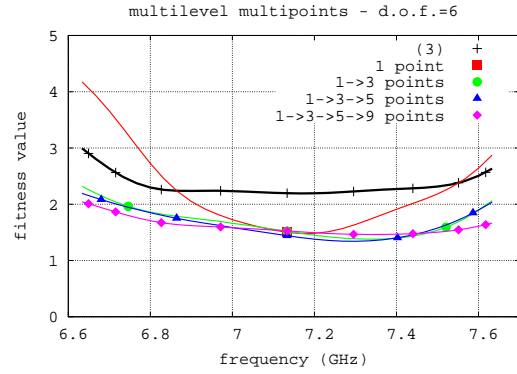


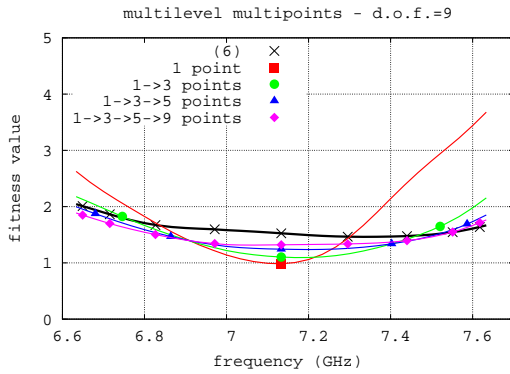
FIG. 5.18 – Schéma de l'algorithme d'enrichissement progressif en fréquence et en paramétrisation.



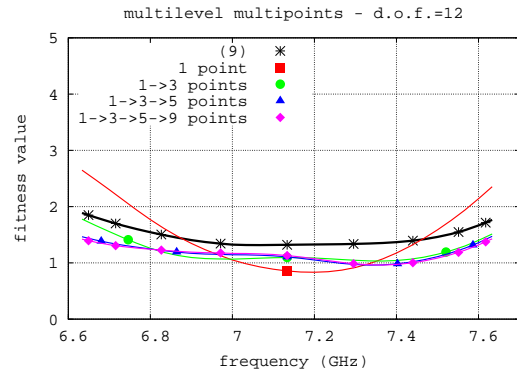
(a) 3 ddl



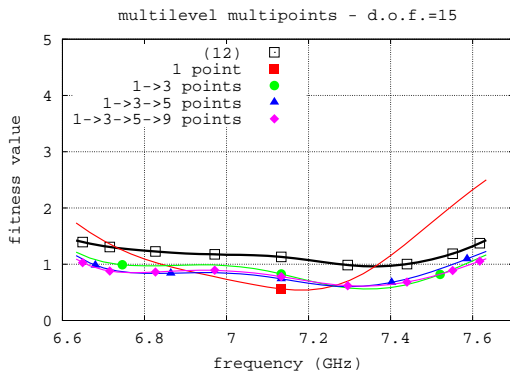
(b) 6 ddl



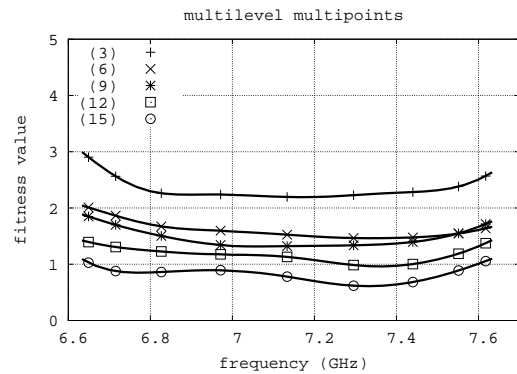
(c) 9 ddl



(d) 12 ddl



(e) 15 ddl



(f) performance finale à chaque niveau de paramétrisation

FIG. 5.19 – Performances individuelles à chaque étape de la stratégie hiérarchique en fréquence et en paramétrisation.

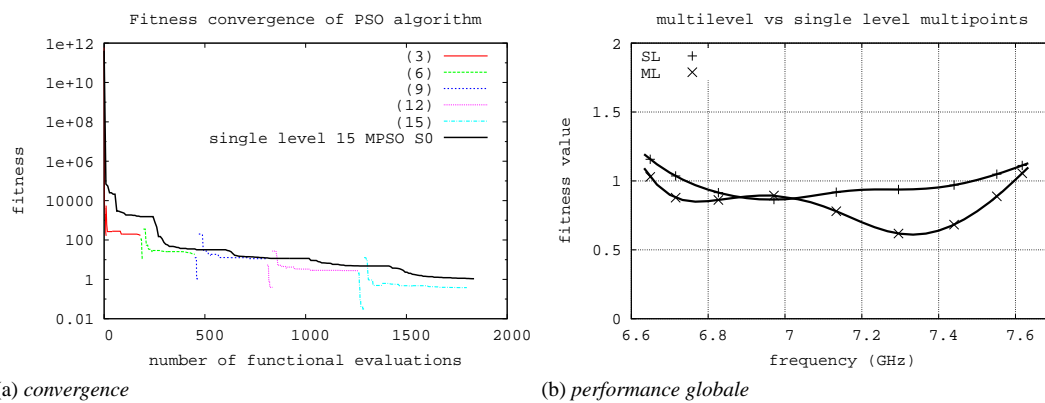


FIG. 5.20 – Comparaison de la performance finale de la méthode hiérarchique avec la méthode sur le seul niveau fin de quadrature.

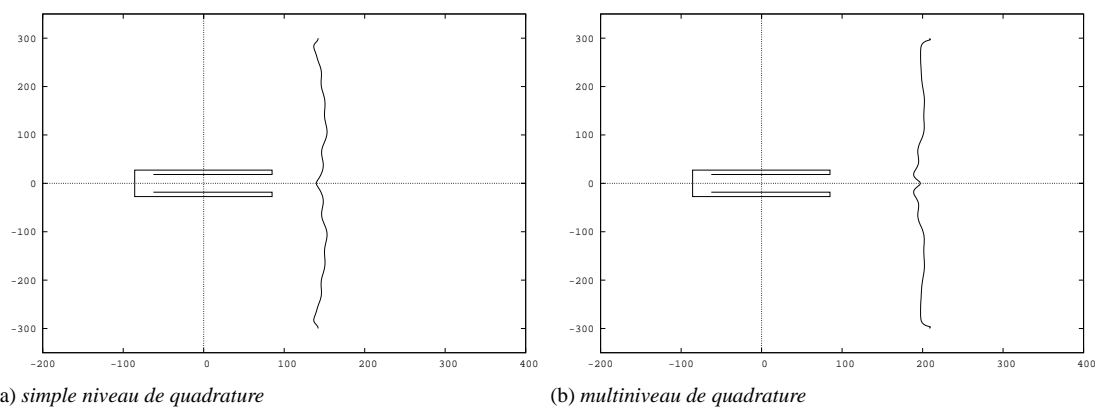


FIG. 5.21 – Formes convergées des deux stratégies.

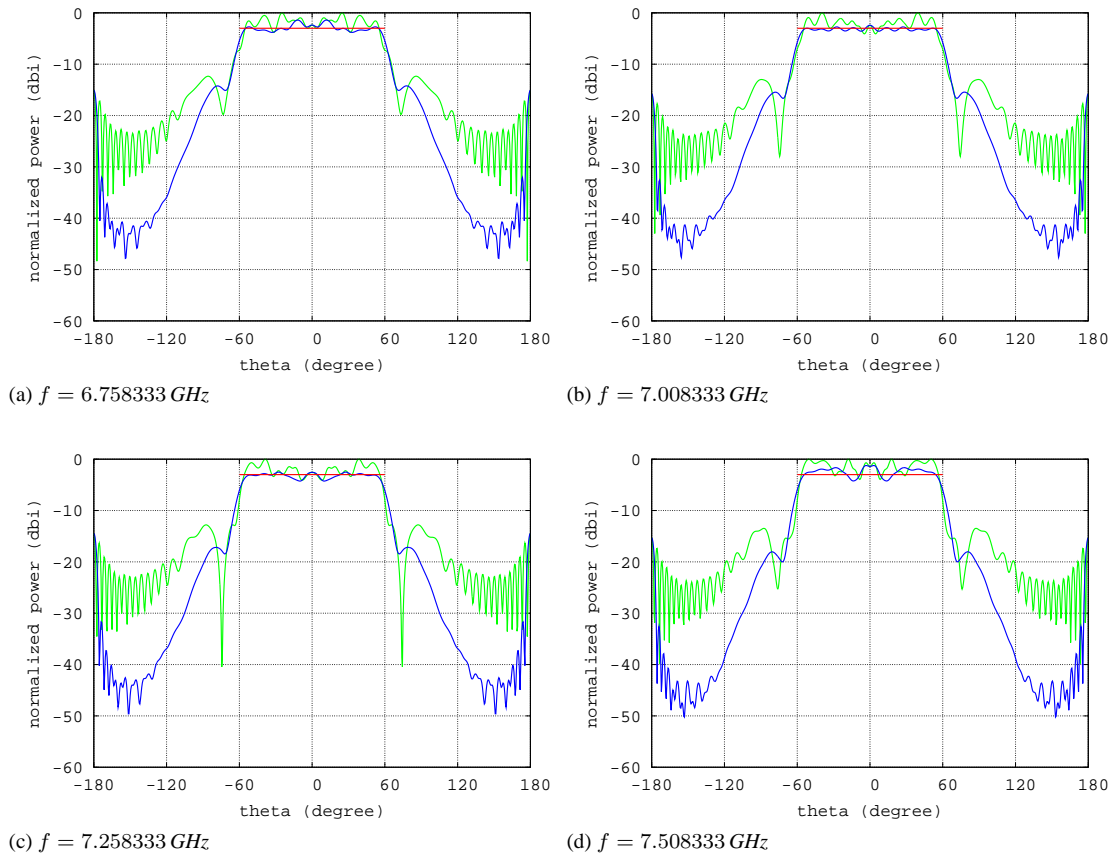


FIG. 5.22 – Diagrammes à convergence de l'algorithme MPSO à un niveau de quadrature évalués en quatre fréquences uniformément réparties sur  $I_f$ .

fréquence (GHz)	$\Delta P$ (db)	plan N	directivité (dbi)
6.758333	2.48		8.19
7.008333	2.25		8.22
7.258333	2.81		8.11
7.508333	2.60		7.62

TAB. 5.5 – Caractéristiques du rayonnement sur l'intervalle  $I_f$  à l'issue de la phase d'optimisation multifréquence à un niveau de quadrature.

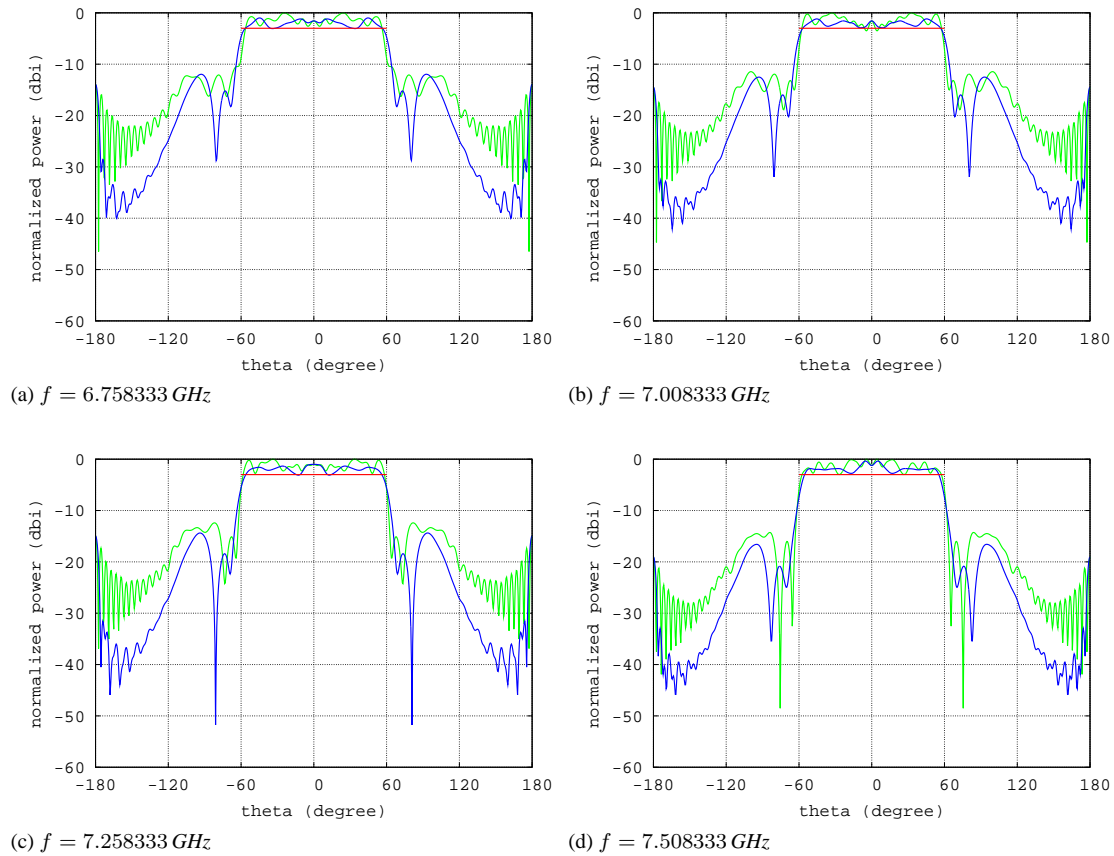


FIG. 5.23 – Diagrammes à convergence de l’algorithme MPSO avec hiérarchie du calcul de quadrature évalués en quatre fréquences uniformément réparties sur  $I_f$ .

fréquence (GHz)	$\Delta P$ (db)	plan N	directivité (dbi)
6.758333	1.21		7.44
7.008333	1.95		7.19
7.258333	1.65		7.21
7.508333	2.20		7.70

TAB. 5.6 – Caractéristiques du rayonnement sur l’intervalle  $I_f$  à l’issue de la phase d’optimisation multifréquence avec hiérarchie de quadrature.

## Conclusion

Dans ce chapitre on a considéré un problème de conception optimale d'une antenne sur une bande de fréquence. Étant donné un critère classique à une fréquence, on a adopté une formulation d'optimisation en moyenne de ce critère. La moyenne est évaluée à l'aide d'une formule de quadrature.

Toujours dans le souci d'exploiter des structures hiérarchiques d'approximation, on a développé une méthode d'optimisation multipoint à plusieurs niveaux de quadrature pour l'approximation du critère continu (la moyenne). Un « niveau » est défini par le nombre de points de quadrature, constituant un « maillage » de la bande de fréquence. Le problème d'optimisation revient à minimiser la quadrature la plus fine. Sachant que le calcul de la quadrature est d'autant plus coûteux qu'il y a de points de fonctionnement, on s'est demandé si la minimisation de quadratures grossières (donc peu coûteuses) permettait de pourvoir un concept plus performant agissant comme préconditionneur pour le problème fin. Combinée avec une stratégie d'enrichissement progressif (MPSO), nos expériences numériques montrent que la stratégie hiérarchique de raffinement en fréquence est plus performante que la stratégie MPSO seule à coût égal en terme de nombre de simulations.

Les niveaux fins, pour la quadrature comme pour la paramétrisation, ont été fixés *a priori*. En pratique, le nombre de points de fonctionnement et la richesse de l'espace de recherche nécessaires pour atteindre une précision uniforme donnée ne sont pas connus. Le processus hiérarchique associé à un critère de convergence, par exemple basé sur le gain relatif obtenu à chaque changement de niveau, permettrait de s'arrêter automatiquement sur le niveau le plus fin utile.

On a vu que la stratégie d'enrichissement progressif de la quadrature était similaire à une méthode de résolution d'un problème non-linéaire où les espaces d'approximation grossiers ne sont pas consistants avec le niveau fin. Il serait donc judicieux d'envisager une stratégie multiniveau plus élaborée de type *FAS* comme celle suggérée dans [70].

# STRATÉGIE DE MOINDRE SENSIBILITÉ POUR L'OPTIMISATION BICRITÈRE

---

<b>6.1</b>	<b>Approche d'optimisation bicritère par stratégie de jeux</b>	<b>161</b>
6.1.1	Généralités	161
6.1.2	Introduction aux jeux de Nash	161
6.1.3	Résultat de partage optimal de territoire	162
6.1.4	Méthodologie pour l'optimisation d'une antenne	163
<b>6.2</b>	<b>Optimisation d'une antenne compacte à une fréquence</b>	<b>163</b>
6.2.1	Description du cas-test	164
6.2.2	Optimisation du coefficient de réflexion	165
6.2.3	Optimisation du diagramme	166
6.2.4	Étude numérique de la sensibilité des fonctionnelles au second ordre	167
6.2.5	Optimisation dans le sous-espace de moindre sensibilité	170
<b>6.3</b>	<b>Application pour l'optimisation d'une antenne sur une bande de fréquence</b>	<b>171</b>
6.3.1	Coefficient de réflexion comme critère principal	171
6.3.2	Diagramme de rayonnement comme critère principal	175
6.3.3	Synthèse	179
	<b>Conclusion</b>	<b>180</b>

---

Dans les chapitres précédents on s'est intéressé à des problèmes d'optimisation à un seul critère physique, l'optimisation du diagramme de rayonnement, parfois en régime multifréquence. On cherche désormais à résoudre un problème couplé où les deux objectifs doivent être traités simultanément. Plutôt que l'approche multicritère exposé au chapitre précédent on adopte une méthodologie basée sur la théorie des jeux où l'espace de recherche alloué à un critère est de moindre influence pour le second critère.





## 6.1 Approche d'optimisation bicritère par stratégie de jeux

### 6.1.1 Généralités

La tendance générale de l'optimisation multicritère, basée sur les concepts de Pareto, a été brièvement exposée au chapitre précédent. On a vu que la méthodologie reposait sur la recherche de vecteurs de conception appartenant à l'espace admissible  $\mathcal{F}$  qui réalisent un compromis défini soit par agrégation des objectifs, soit par une relation d'ordre entre les vecteurs de performance (dominance).

Une autre approche, récente et peu utilisée à notre connaissance, fait appel à la théorie des jeux [99]. Il s'agit de mettre en compétition des *joueurs* pour la minimisation simultanée de chacun des critères. L'espace de décision (ou espace de recherche) alloué à un joueur est un sous-ensemble de l'espace admissible. L'ensemble des espaces de décision forment une partition de  $\mathcal{F}$  (les espaces de décision sont disjoints deux à deux et leur union est un recouvrement de  $\mathcal{F}$ ). Le vecteur de conception global résulte alors de la contribution des décisions de chaque joueur. Autrement dit, un joueur adapte sa stratégie dans l'espace qui lui est propre en fonction des stratégies des autres joueurs. On espère ainsi atteindre un équilibre pour lequel aucun des joueurs ne souhaite modifier sa stratégie. Une des difficultés d'interprétation de cet équilibre, si il existe, vient du fait qu'à notre connaissance il n'y a pas de résultat qui établisse un lien avec l'optimalité de Pareto.

La manière dont les joueurs échangent leur décision pour assembler le vecteur de conception donne naissance à de nombreuses catégories de jeux. On n'aborde ici que les jeux non coopératifs et en particulier les jeux de Nash non coopératifs [69] qui sont symétriques par rapport aux joueurs. Dans le cadre de l'optimisation multicritère on trouve également l'application de jeux de Stackelberg, qui sont asymétriques de type meneur-suiveur. Comme exemples d'application de jeux dynamiques à des problèmes de conception optimale multicritère/multidisciplinaire de systèmes gouvernés par des EDP on a [58, 100, 95] (aérodynamique) ou encore [48, 47] (structure, biologie).

### 6.1.2 Introduction aux jeux de Nash

Soit deux critères  $J_A$  et  $J_B$  définis sur  $\mathcal{F}$ . Pour la simplicité du discours considérons le cas sans contraintes  $\mathcal{F} = \mathbb{R}^n$ . On pose  $\mathcal{F} = \mathcal{F}_A \oplus \mathcal{F}_B$  (c-à-d  $\mathcal{F}_A$  et  $\mathcal{F}_B$  sont deux sous-espaces en somme directe).  $\mathcal{F}_A$  (resp.  $\mathcal{F}_B$ ) est l'espace de décision du joueur  $A$  (resp.  $B$ ) pour minimiser le critère  $J_A$  (resp.  $J_B$ ). Formellement le jeu de Nash consiste à résoudre le problème (6.1) :

$$\begin{cases} \min_{x \in \mathcal{F}_A} J_A(x + y) \\ \min_{y \in \mathcal{F}_B} J_B(x + y) \end{cases} \quad (6.1)$$

Supposons que  $\dim(\mathcal{F}_A) = n - m$  et  $\dim(\mathcal{F}_B) = m$  tel que  $0 < m < n$ , c-à-d tel qu'aucun sous-espace ne soit vide (on a bien  $\dim(\mathcal{F}_A) + \dim(\mathcal{F}_B) = \dim(\mathcal{F})$ ). On munit  $\mathcal{F}_A$  et  $\mathcal{F}_B$  des bases  $S_A \in \mathbb{R}^{n \times (n-m)}$  et  $S_B \in \mathbb{R}^{n \times m}$  respectivement. Les sous-espaces étant supplémentaires, la matrice  $S = (S_A \ S_B) \in \mathbb{R}^{n \times n}$  est donc une base de  $\mathbb{R}^n$ . On appelle  $S$  la matrice de *partage* du territoire  $\mathcal{F}$  ( $S$  pour *split*). Tout vecteur de conception  $x$  s'écrit donc

$$x(u, v) = x_A(u) + x_B(v) = S_A u + S_B v = S \begin{pmatrix} u \\ v \end{pmatrix} \quad (6.2)$$

pour  $u \in \mathbb{R}^{(n-m)}$  et  $v \in \mathbb{R}^m$ . En fonction de ces nouvelles variables, les critères s'écrivent

$$J_A(u, v) \stackrel{\text{def.}}{=} J_A(x(u, v)) \quad \text{et} \quad J_B(u, v) \stackrel{\text{def.}}{=} J_B(x(u, v)) \quad (6.3)$$

Un équilibre de Nash est un couple  $(u^*, v^*)$  tel que ni  $A$  ni  $B$  n'a intérêt à changer de stratégie étant donné la décision de l'autre joueur :

$$\begin{cases} u^* = \arg \min_u J_A(u, v^*) \\ v^* = \arg \min_v J_B(u^*, v) \end{cases} \quad (6.4)$$

En optimisation différentiable, une condition nécessaire est caractérisé par

$$\begin{aligned} \nabla_u J_A(u^*, v^*) &= 0 \\ \nabla_v J_B(u^*, v^*) &= 0 \end{aligned} \quad (6.5)$$

où l'opérateur  $\nabla_u$  (resp.  $\nabla_v$ ) désigne le gradient partiel dans le sous-espace auquel  $u$  (resp.  $v$ ) appartient.

Tel que le problème est posé, le choix des sous-espaces  $\mathcal{F}_A$  et  $\mathcal{F}_B$ , c-à-d le choix de la matrice de partage  $S$ , est arbitraire. Pourtant il détermine le déroulement du jeu. On montre qu'il existe une manière optimale de choisir ce partage [31, 1].

### 6.1.3 Résultat de partage optimal de territoire

On se base sur la méthodologie proposée dans [31] pour élaborer une stratégie d'optimisation bicritère. On montre qu'il faut mesurer dans un certain sens l'importance de l'antagonisme et déterminer dans quel(s) sous-espace(s) le conflit est le plus critique.

Dans un premier temps on suppose que les deux critères ont été hiérarchisés par le concepteur en tant que critère principal et critère secondaire. On pose arbitrairement  $J_A$  comme critère principal. Supposons qu'un processus d'optimisation pour la minimisation de ce critère a été réalisé. On aboutit alors au point stationnaire  $x_A^*$  ( $\nabla J_A(x_A^*) = \mathbf{0}$ ) pour lequel on note  $J_A^* = J(x_A^*)$ . Le développement de Taylor d'ordre 2 du critère  $J_A$  en  $x_A^*$  et dans la direction  $\Delta x \in \mathbb{R}^n$  tel que  $\|\Delta x\|_2 = 1$  est

$$\forall h > 0 \quad J_A(x^* + h\Delta x) = J_A^* + \frac{1}{2}h^2 \Delta x^T H(x_A^*) \Delta x + O(h^3) \quad (6.6)$$

À partir de ce point on cherche alors à résoudre le problème bicritère au moyen du jeu de Nash (6.7) :

$$\begin{cases} \min_{u \in \mathbb{R}^{n-m}} \tilde{J}_A(u, v) \\ \min_{v \in \mathbb{R}^m} \tilde{J}_B(u, v) \end{cases} \quad (6.7)$$

où

$$\tilde{J}_A(u, v) \stackrel{\text{def.}}{=} J_A \left( x_A^* + S \begin{pmatrix} u \\ v \end{pmatrix} \right) \quad \text{et} \quad \tilde{J}_B(u, v) \stackrel{\text{def.}}{=} J_B \left( x_A^* + S \begin{pmatrix} u \\ v \end{pmatrix} \right) . \quad (6.8)$$

Si  $x_A^*$  est un minimum local, alors  $J_A$  ne peut être que dégradé dans un voisinage de  $x_A^*$ . On utilise le théorème de Courant-Fisher 6.1 pour trouver le sous-espace parmi tous les sous-espaces de dimension  $m$  qui minimise les perturbations de  $J_A$  au second ordre. Pour cela on introduit le quotient de Rayleigh :

**Définition 6.1.** Soit  $H$  une matrice symétrique de dimension  $n$  et  $x$  un vecteur non nul de  $\mathbb{R}^n$ . Le quotient de Rayleigh  $\rho(x, H)$  est défini par

$$\rho(x, H) = \frac{x^T H x}{x^T x} \quad (6.9)$$

On peut restreindre le quotient de Rayleigh aux vecteurs  $x$  normalisés tels que  $\|x\|_2 = 1$  puisque  $\rho(\alpha x, H) = \rho(x, H)$  pour tout  $\alpha \neq 0$ .

On a le théorème suivant dont on trouvera une preuve dans [28] par exemple :

**Théorème 6.1. (Courant-Fisher minimax)** Soit  $H$  une matrice réelle symétrique. On pose  $H = \Omega \Lambda \Omega^T$  et on suppose que les valeurs propres sont ordonnées par ordre croissant  $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ , associées aux vecteurs propres  $\omega_1, \omega_2, \dots, \omega_n$ , qui sont orthogonaux en vertu de la symétrie de  $H$ . Alors on a

$$\min_{\substack{\mathcal{F}_B \subset \mathbb{R}^n \\ \dim(\mathcal{F}_B) = m}} \max_{\substack{\Delta x \in \mathcal{F}_B \\ \|\Delta x\| = 1}} \rho(\Delta x, H) = \lambda_m = \max_{\substack{\mathcal{F}_A \subset \mathbb{R}^n \\ \dim(\mathcal{F}_A) = n-m+1}} \min_{\substack{\Delta y \in \mathcal{F}_A \\ \|\Delta y\| = 1}} \rho(\Delta y, H) \quad (6.10)$$

et l'optimum est atteint pour

$$\mathcal{F}_B = \text{vect} \{ \omega_1, \dots, \omega_m \} \quad \mathcal{F}_A = \text{vect} \{ \omega_m, \dots, \omega_n \} \quad (6.11)$$

Un vecteur maximisant  $\Delta x$  et minimisant  $\Delta y$  est  $\omega_m$ .

Sachant que le développement de Taylor (6.6) s'écrit aussi à l'aide du quotient de Rayleigh

$$\forall h > 0 \quad J_A(x^* + h\Delta x) = J_A^* + \frac{1}{2}h^2\rho(\Delta x, H(x_A^*)) + O(h^3) \quad (6.12)$$

les territoires affectés aux joueurs A et B peuvent être déterminés à l'aide du théorème 6.1 : trouver le sous-espace  $\mathcal{F}_B$  de dimension  $m$  qui minimise le maximum potentiel de perturbation locale de  $J_A$  au second ordre revient à minimiser le maximum du coefficient de Rayleigh parmi tous les sous-espaces de dimension  $m$ . En effet, en utilisant la partie gauche de (6.10), un tel sous-espace est donné en terme des vecteurs propres  $\omega_k$  de  $H$ . Ainsi, pour définir la matrice de partage, on utilise la décomposition spectrale de la matrice hessienne de  $J_A$  au point  $x_A^*$

$$H(x_A^*) = \Omega_H \Lambda_H \Omega_H^T \quad (6.13)$$

où les valeurs propres sont ordonnées par ordre décroissant. Alors le partage optimal est donné par la matrice

$$S = \Omega_H. \quad (6.14)$$

Cet espace est alors alloué au joueur B pour la minimisation de  $J_B$ . En pratique, il reste à déterminer la dimension  $m$  de l'espace. Ceci peut être décidé *a posteriori* d'après la décomposition spectrale du hessien et sur la base de l'ordre de grandeur relatif des valeurs propres.

#### 6.1.4 Méthodologie pour l'optimisation d'une antenne

Dans la méthodologie précédente on a supposé qu'un des critères était prépondérant et qu'une optimisation numérique de ce critère avait été conduite. Dans le cas de l'optimisation bicritère d'une antenne, on ne sait pas *a priori* quel critère est prépondérant. On tente de le déterminer sur la base de l'étude spectrale du hessien : pour cela on réalise deux optimisations numériques indépendantes pour la minimisation de  $J_A$  et  $J_B$ . En chacun des points optimaux convergés, notés  $x_A^*$  et  $x_B^*$  respectivement, on calcule une approximation du hessien,  $H_A^* = H_A(x_A^*)$  et  $H_B^* = H_B(x_B^*)$ . On détermine alors les sous-espaces de moindre influence mutuelle d'après la décomposition spectrale des matrices hessiennes. En ce sens on mesure l'antagonisme des deux critères en deux points particuliers du front de Pareto, à savoir les points extrêmes du front (cf. Figure 5.2). On en déduit le critère prépondérant sur la base des valeurs propres.

## 6.2 Optimisation d'une antenne compacte à une fréquence

Dans cette section on traite un cas-test d'optimisation bicritère d'une antenne compacte. On cherche à la fois à satisfaire un critère en directivité et à minimiser le coefficient de réflexion dans la source guidée. Les objectifs (cahier des charges) sont donnés dans le Tableau 6.1. Dans un premier temps on considère un régime de fonctionnement fixé à la fréquence  $f=25$  GHz.

- |     |   |
|-----|---|
| (1) | maximisation de la directivité en $\theta = 0^\circ$  |
| (2) | lobe secondaire inférieur à -20 db par rapport à la puissance maximale (en $\theta = 0^\circ$ ) |
| (3) | ROS inférieur à 1.05  |

TAB. 6.1 – Cahier des charges de l'optimisation d'une antenne compacte.

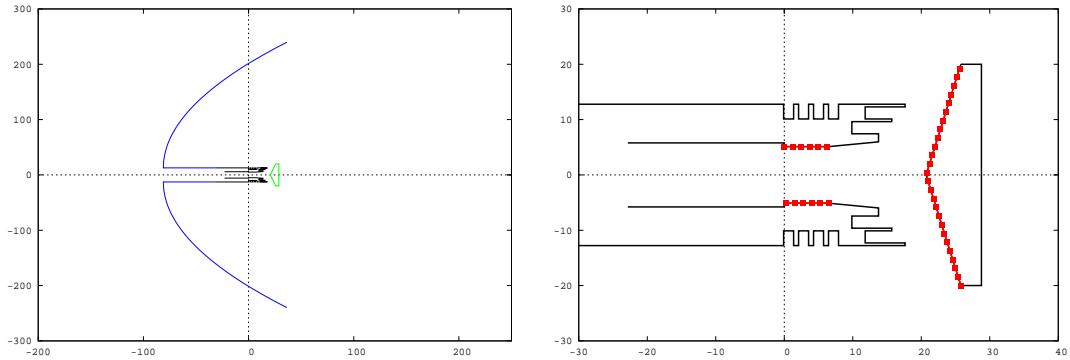
On réalise dans un premier temps une optimisation numérique de chacun des critères dans l'espace de recherche tout entier. Puis on étudie la décomposition spectrale du hessien pour déterminer le sous-espace de « moindre influence » pour les deux critères. On pourra déterminer *a posteriori* la dimension  $m$  de  $\mathcal{F}_B$  en fonction des valeurs propres.

### 6.2.1 Description du cas-test

Soit l'antenne de la Figure 6.1. Elle est composée d'un guide d'onde et de deux réflecteurs : un sous-réflecteur et un réflecteur principal. Le contrôle est formé d'une portion en amont du guide d'onde et de la portion du sous-réflecteur qui fait face à l'ouverture du guide. Le réflecteur principal est fixe.

La paramétrisation des portions à optimiser est la suivante :

- portion du guide : B-splines d'ordre 1, d.d.l. = 7
- sous-réflecteur : B-splines d'ordre 4, d.d.l. = 7
- contraintes :  $x_0 = 0$  et  $x_6 = 0$  (pour chaque portion)



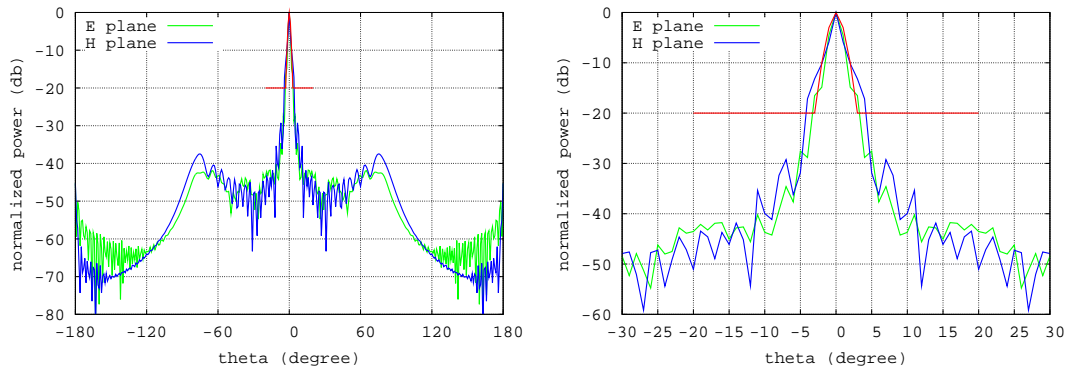
(a) géométrie : guide d'onde (noir) sous-réflecteur (vert) et réflecteur (bleu) (b) contrôle : portion amont du guide d'onde et sous-réflecteur

FIG. 6.1 – Illustration des composants de l'antenne et du contrôle. Un sous-réflecteur est situé juste devant l'ouverture du guide.

Les caractéristiques de rayonnement de la géométrie initiale sont données dans le Tableau 6.2.

- |     |   |
|-----|---|
| (1) | directivité en $\theta = 0^\circ$ : 38.90 dbi |
| (2) | diagramme : voir Figure 6.2                   |
| (3) | ROS : 1.448                                   |

TAB. 6.2 – Caractéristiques de l'antenne initiale.



(a) diagramme de rayonnement : plans E et H, cible (rouge) (b) diagramme : zoom sur  $-30^\circ \leq \theta \leq 30^\circ$

FIG. 6.2 – Rayonnement de l'antenne initiale. Le diagramme est très directif : l'angle à -3 db est de  $1^\circ$ .

Pour atteindre les objectifs 1 et 2 énoncés dans le Tableau 6.1, on choisit de minimiser la fonctionnelle

$$J_A(x) = \frac{1}{2} \int_{S^2} \left( D(u(x)) - \tilde{D} \right)^2 dS^2 \quad (6.15)$$

où  $\tilde{D}$  est un profil uniforme sur l'intervalle  $[-0.5^\circ \ 0.5^\circ]$ . En ce sens on cherche à maximiser la puissance sur  $[-0.5^\circ \ 0.5^\circ]$  tout en la minimisant ailleurs. Pour atteindre l'objectif 3, on cherche simplement à minimiser le coefficient de réflexion

$$J_B(x) = \frac{1}{2} |\Gamma(u(x))|^2. \quad (6.16)$$

### 6.2.2 Optimisation du coefficient de réflexion

À l'issue de la phase d'optimisation numérique du critère  $J_B$ , on obtient l'antenne illustrée sur la Figure 6.3 et dont les caractéristiques du rayonnement sont données dans le Tableau 6.3. La convergence est illustrée sur la Figure 6.4.

(1)	directivité en $\theta = 0^\circ$ : 38.86 dbi
(2)	diagramme : voir Figure 6.3
(3)	ROS : 1.000

TAB. 6.3 – Caractéristiques de l'antenne optimisée pour  $J_B$ .

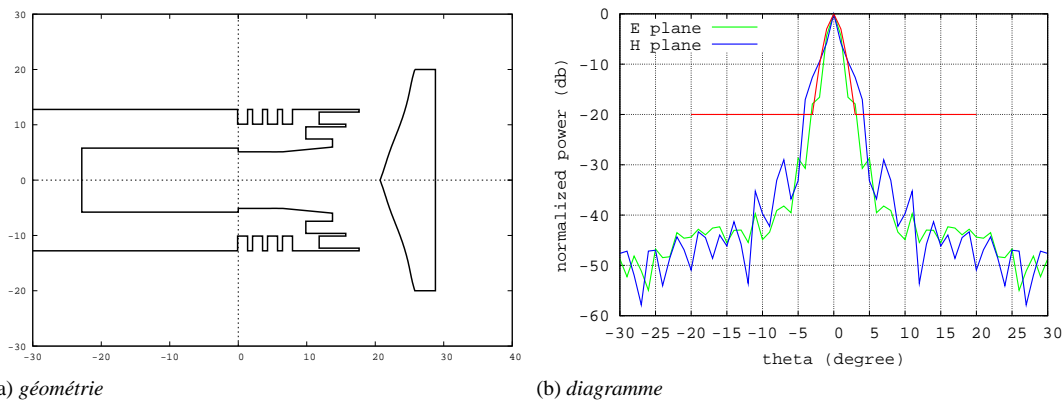


FIG. 6.3 – Configuration de l'antenne après optimisation du coefficient de réflexion seul.

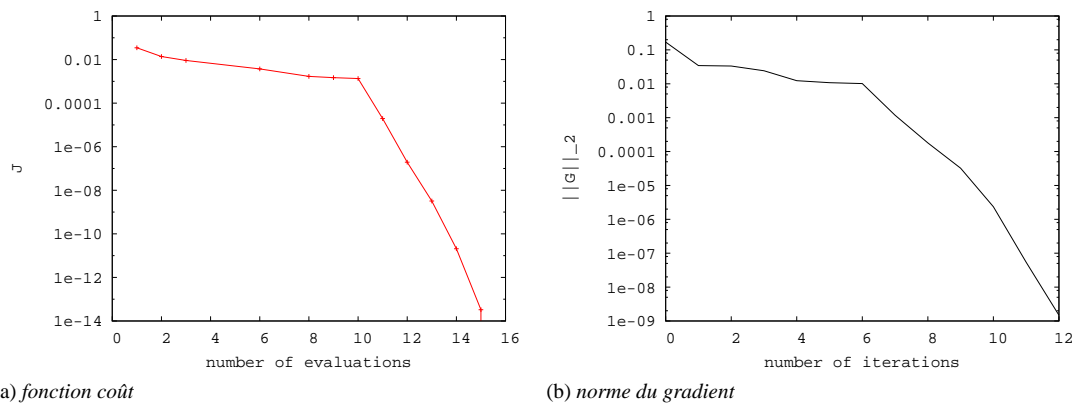


FIG. 6.4 – Convergence de l'algorithme QN-BFGS.

### 6.2.3 Optimisation du diagramme

À l'issue de la phase d'optimisation numérique du critère  $J_A$ , on obtient l'antenne illustrée sur la Figure 6.5 et dont les caractéristiques du rayonnement sont données dans le Tableau 6.4. La convergence est illustrée sur la Figure 6.6.

- |     |   |
|-----|---|
| (1) | directivité en $\theta = 0^\circ$ : 40.44 dbi |
| (2) | diagramme : voir Figure 6.5                   |
| (3) | ROS : 56.610                                  |

TAB. 6.4 – Caractéristiques de l'antenne optimisée pour  $J_A$ .

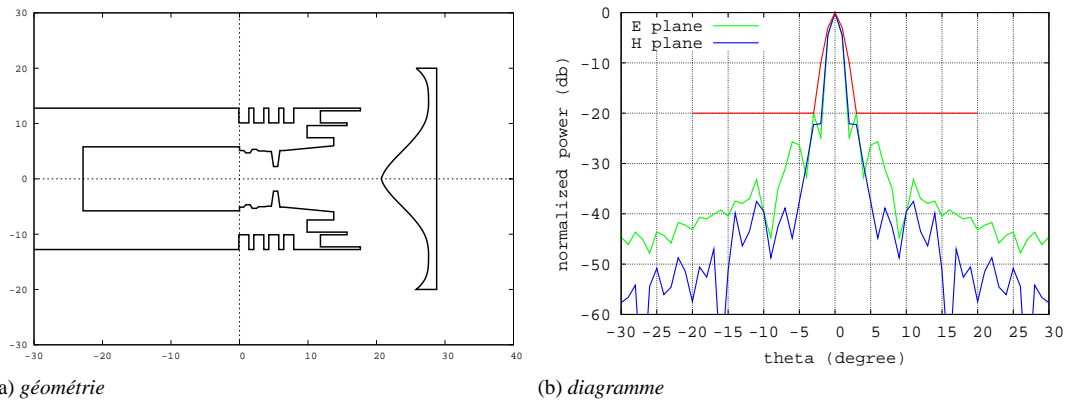


FIG. 6.5 – Configuration de l'antenne après optimisation du diagramme seul.

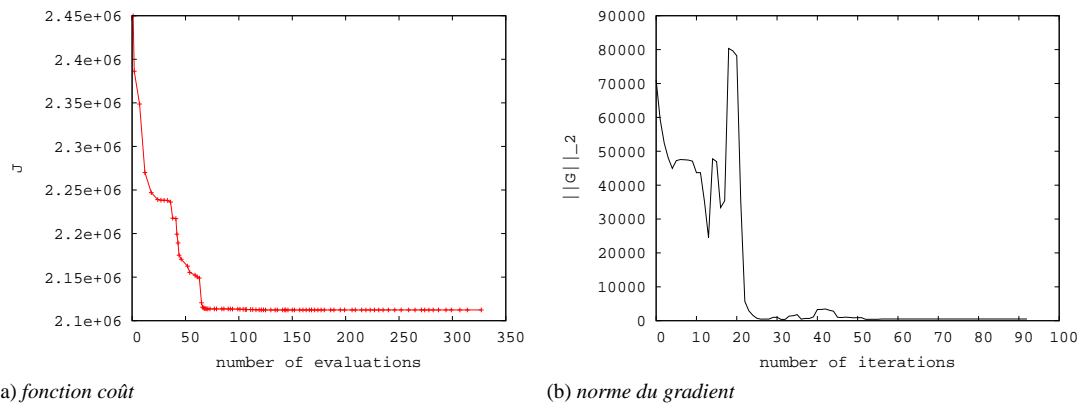


FIG. 6.6 – Convergence de l'algorithme QN-BFGS.

### 6.2.4 Étude numérique de la sensibilité des fonctionnelles au second ordre

On étudie les spectres des fonctionnelles normalisées

$$J_A(x) = \frac{J_A(x)}{J_A^*} \quad \text{et} \quad J_B(x) = \frac{J_B(x)}{J_B^*} \quad (6.17)$$

où  $J_A^*$  et  $J_B^*$  sont les valeurs de fonctionnelles obtenues à la convergence des algorithmes de descente pour les objectifs du diagramme et du coefficient de réflexion respectivement.

Il est clair, d'après les Figures 6.7, 6.8 et 6.9, que l'espace de dimension  $m = 5$  de moindre influence sur la fonctionnelle  $J_A$  en  $x_A^*$  est l'espace de déformation de la portion amont du guide d'onde. En outre, l'ordre de grandeur des plus petites valeurs propres relativement aux plus grandes (en valeur absolue) sont telles que les déformations du guide d'onde n'ont aucune influence significative sur la fonctionnelle  $J_A$  (c-à-d sur le diagramme).

Parallèlement, d'après les Figures 6.10, 6.11 et 6.12, l'espace de dimension  $m = 5$  de moindre influence sur la fonctionnelle  $J_B$  en  $x_B^*$  est donné par une base de fonctions propres composée de déformations du guide et du sous-réfecteur (les déformations du guide semblent moins influant, cf. les trois premiers vecteurs propres de la Figure 6.11). D'après l'ordre de grandeur des plus petites valeurs propres, si les déformations ont le moins d'influence, elles ne sont pas pour autant négligeables. On remarque par ailleurs que parmi les modes propres qui ont le plus d'influence sur  $J_B$  (cf. Figure 6.12) figurent principalement des modes de déformations du sous-réfecteur.

En guise de synthèse on remarque que d'une part, à partir de  $x_A^*$ , le sous-espace de déformation du guide n'a pas d'influence au second ordre sur le diagramme optimisé. On peut donc chercher à optimiser le coefficient de réflexion ( $J_B$ ) sans se soucier de dégrader la performance obtenue pour le diagramme ( $J_A$ ). D'autre part, l'approche symétrique n'est pas vraie : à partir de  $x_B^*$  la déformation dans le sous-espace de moindre influence de  $J_B$  est potentiellement néfaste à la performance obtenue. En ce sens, on confirme la hiérarchie des fonctions coût,  $J_A$  étant le critère principal devant  $J_B$ .

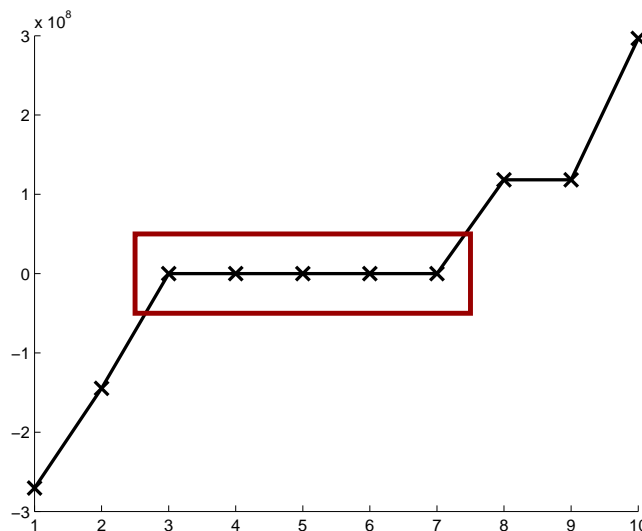


FIG. 6.7 – Valeurs propres du hessien projeté de la fonction objectif  $J_A$  évalué au point  $x_A^*$ . Le point stationnaire n'est pas un minimum local : il existe deux valeurs propres négatives. À la précision numérique du gradient près,  $x_A^*$  est donc un point selle ; l'algorithme échoue cependant à diminuer la fonction coût dans les directions propres correspondantes en raison d'un « crash » de la simulation provenant d'un problème de maillage ; autrement dit, ces directions peuvent être considérées comme des directions non admissibles. L'espace propre de dimension 5 correspondant à l'espace de moindre sensibilité par rapport à  $J_A$  correspond aux modes dont les valeurs propres sont les plus petites : elles sont encadrées en rouge.



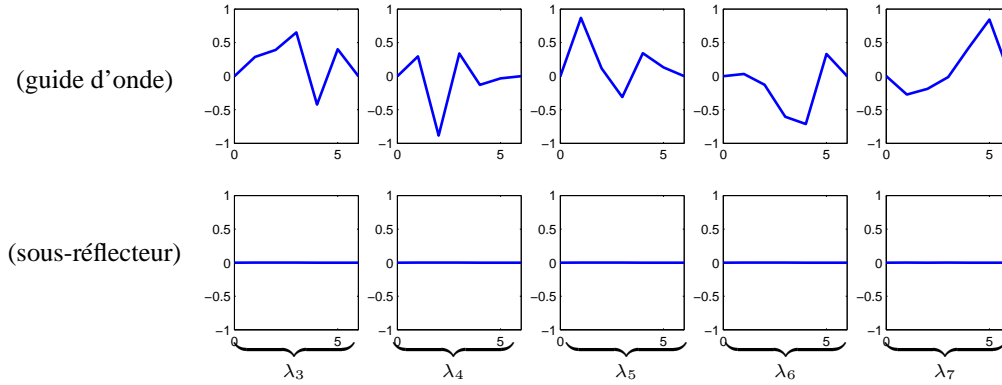


FIG. 6.8 – Vecteurs propres du hessien projeté de la fonction objectif  $J_A$  correspondant aux valeurs propres de plus petite valeur absolue. Chaque colonne représente un mode propre divisé en deux parties : la ligne supérieure correspond aux déformations du guide d'onde, la ligne inférieure aux déformations du sous-rélecteur. Ils sont triés de gauche à droite par ordre croissant des valeurs propres associées.

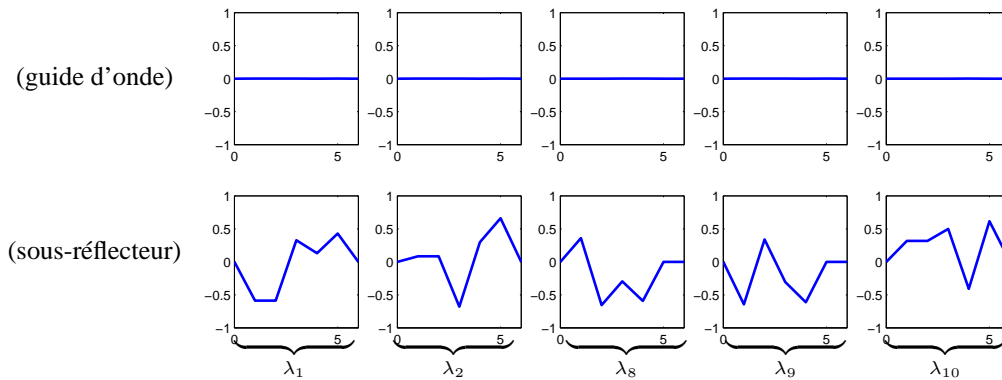


FIG. 6.9 – Vecteurs propres du hessien projeté de la fonction objectif  $J_A$  correspondant aux valeurs propres de plus grande valeur absolue. Chaque colonne représente un mode propre divisé en deux parties : la ligne supérieure correspond aux déformations du guide d'onde, la ligne inférieure aux déformations du sous-rélecteur. Ils sont triés de gauche à droite par ordre croissant des valeurs propres associées.

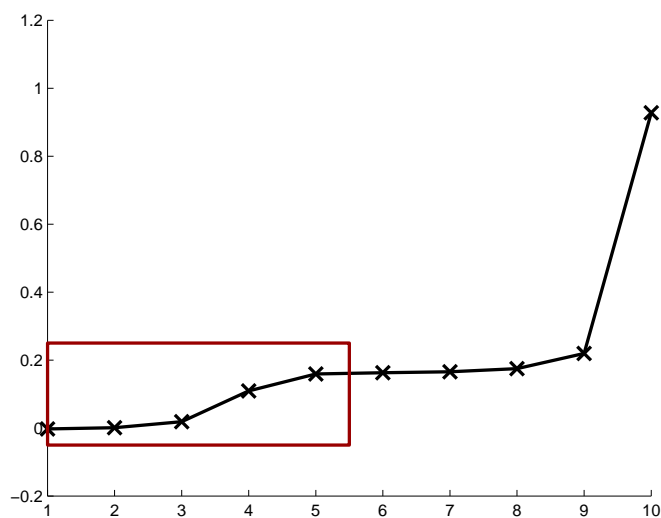


FIG. 6.10 – Valeurs propres du hessien projeté de la fonction objectif  $J_B$  évalué au point  $x_B^*$ . Le point stationnaire est un minimum local : toutes les valeurs propres sont positives. L'espace propre de dimension 5 correspondant à l'espace de moindre sensibilité par rapport à  $J_B$  correspond aux modes dont les valeurs propres sont les plus petites : elles sont encadrées en rouge.

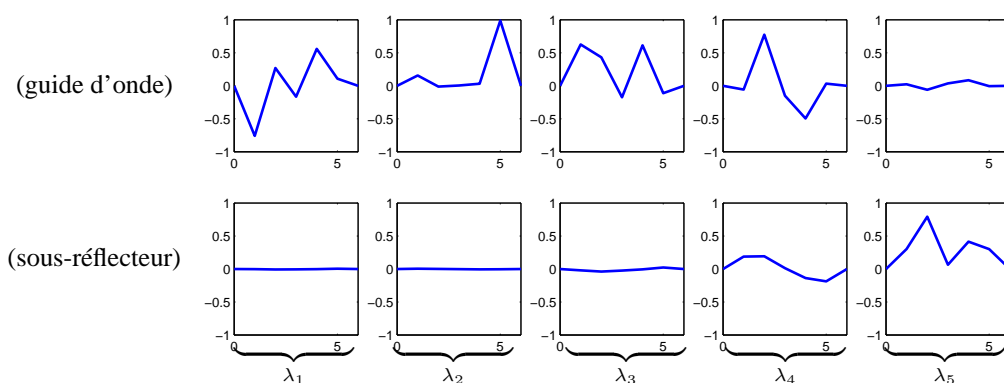


FIG. 6.11 – Vecteurs propres du hessien projeté de la fonction objectif  $J_B$  correspondant aux valeurs propres de plus petite valeur absolue.

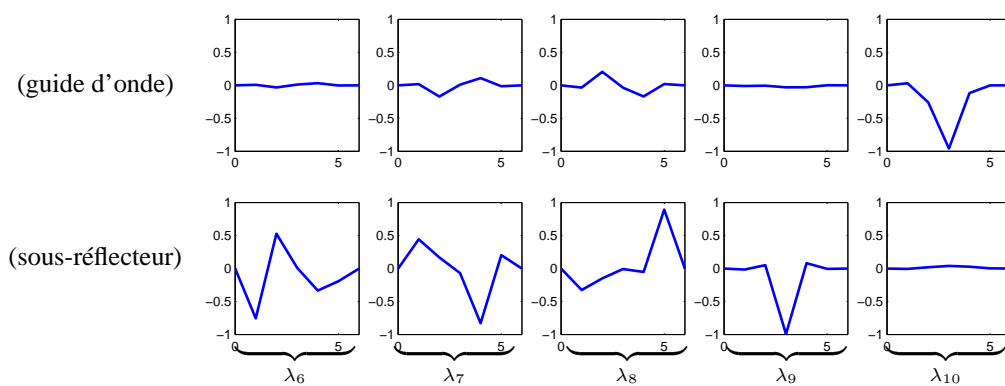


FIG. 6.12 – Vecteurs propres du hessien projeté de la fonction objectif  $J_B$  correspondant aux valeurs propres de plus grande valeur absolue.

### 6.2.5 Optimisation dans le sous-espace de moindre sensibilité

On cherche désormais à optimiser le coefficient de réflexion à partir de  $x_A^*$  et dans le sous-espace de déformation du guide d'onde. À l'issue de la phase d'optimisation numérique du critère  $J_B$ , on obtient l'antenne illustrée sur la Figure 6.13 et dont les caractéristiques du rayonnement sont données dans le Tableau 6.5. La convergence est illustrée sur la Figure 6.14.

Effectivement, la directivité a très peu baissé (-0.06 db) et la forme du diagramme reste la même (le lobe secondaire est bien inférieur à  $-20$  db), alors que le ROS a la valeur optimale de 1.0.

- |     |   |
|-----|---|
| (1) | directivité en $\theta = 0^\circ$ : 40.38 dbi |
| (2) | diagramme : voir Figure 6.13                  |
| (3) | ROS : 1.000                                   |

TAB. 6.5 – Caractéristiques de l'antenne optimisée pour  $J_A$  puis  $J_B$ .

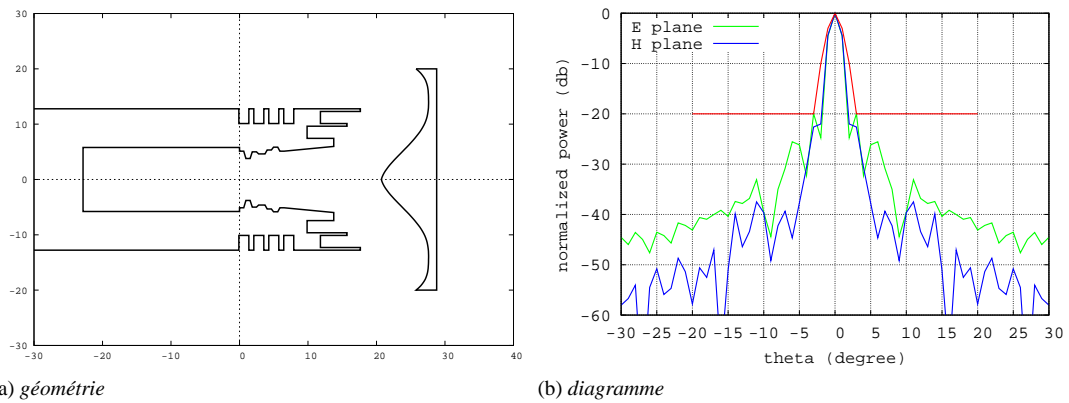


FIG. 6.13 – Configuration de l'antenne après optimisation du diagramme puis du coefficient de réflexion dans le sous-espace de moindre sensibilité de  $J_A$ .

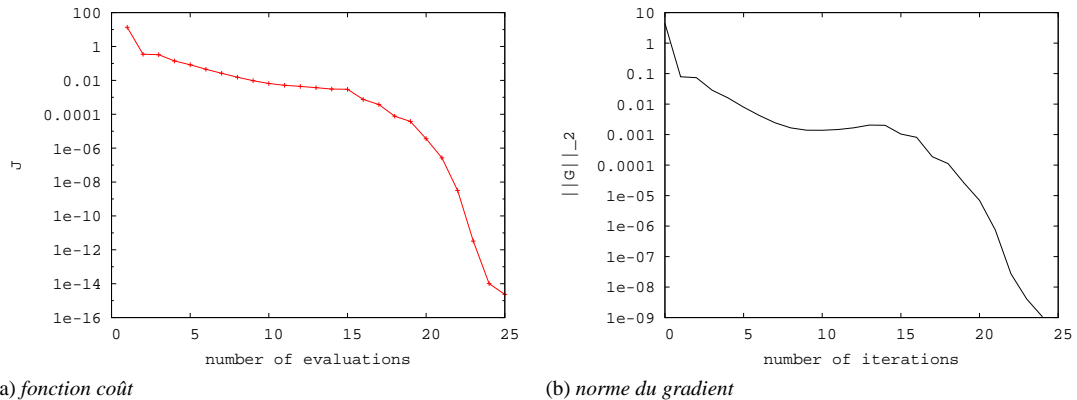


FIG. 6.14 – Convergence de l'algorithme QN-BFGS.

## 6.3 Application pour l'optimisation d'une antenne sur une bande de fréquence

Dans une configuration plus réaliste, on cherche à optimiser cette même antenne (Figure 6.1) sur la bande de fréquence  $I_f = [20\ 21]$  GHz. On cherche à minimiser en moyenne les critères (6.15) et (6.16), au sens multicritère. Pour cela on adopte la même méthodologie qu'à la section précédente en choisissant comme critère principal alternativement celui qui porte sur le coefficient de réflexion d'une part ( $J_B$ ) et celui qui porte sur le diagramme de rayonnement d'autre part ( $J_A$ ).

### 6.3.1 Coefficient de réflexion comme critère principal

Le critère  $J_B$  est choisi comme critère principal, que l'on minimise classiquement dans un premier temps. Puis on étudie le spectre du hessien numérique de la forme convergée afin d'en déduire le partage optimal. Enfin, on réalise un jeu de Nash dans le sous-espace de moindre sensibilité.

#### Optimisation du coefficient de réflexion

L'optimisation numérique de  $J_B$  en moyenne sur  $I_f$  est illustrée sur la Figure 6.15. On représente la forme convergée définie par  $x_B^*$ , la convergence de la fonction de coût, et les performances obtenues sur  $I_f$ .

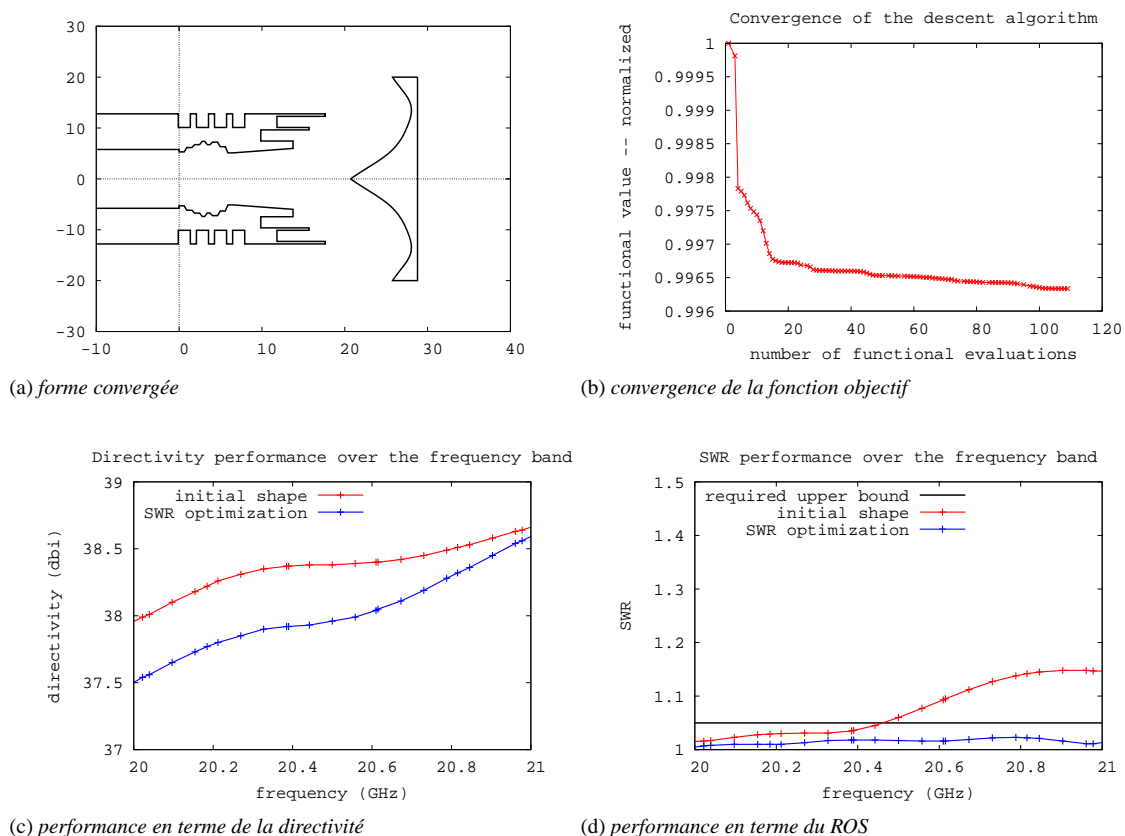


FIG. 6.15 – Optimisation numérique du coefficient de réflexion ( $J_B$ ) sur la bande de fréquence  $I_f$ . La convergence numérique est obtenue au bout de 110 évaluations de fonctionnelle. Sur les figures (c) et (d) on représente les performances de la forme initiale (rouge) et de la forme optimisée (bleu) sur  $I_f$ . On constate que le ROS a été uniformément diminué, alors que la directivité a été dégradée.

### Étude numérique de la sensibilité

Une étude numérique du hessien de  $J_B$  en  $x_B^*$  est réalisée. Sur la Figure 6.16 on représente les valeurs propres, et sur les Figures 6.17 et 6.18 les vecteurs propres. Les modes sont effectivement couplés : les vecteurs propres sont constitués de déformation à la fois du guide et du sous-réfecteur. L'espace engendré par les vecteurs propres associés aux 5 valeurs propres les plus petites (cf. Figure 6.17) constitue le territoire du joueur  $A$  pour améliorer la directivité tout en dégradant le moins possible  $J_B$  (déjà optimisé).

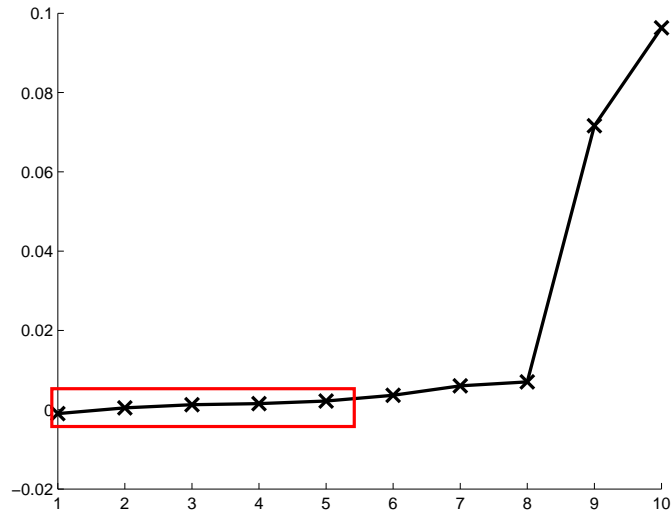


FIG. 6.16 – Valeurs propres du hessien projeté de la fonction objectif  $J_B$  évalué au point  $x_B^*$ . Le point stationnaire est un minimum local : toutes les valeurs propres sont positives. L'espace propre de dimension 5 correspondant à l'espace de moindre sensibilité par rapport à  $J_B$  correspond aux modes dont les valeurs propres sont les plus petites : elles sont encadrées en rouge.

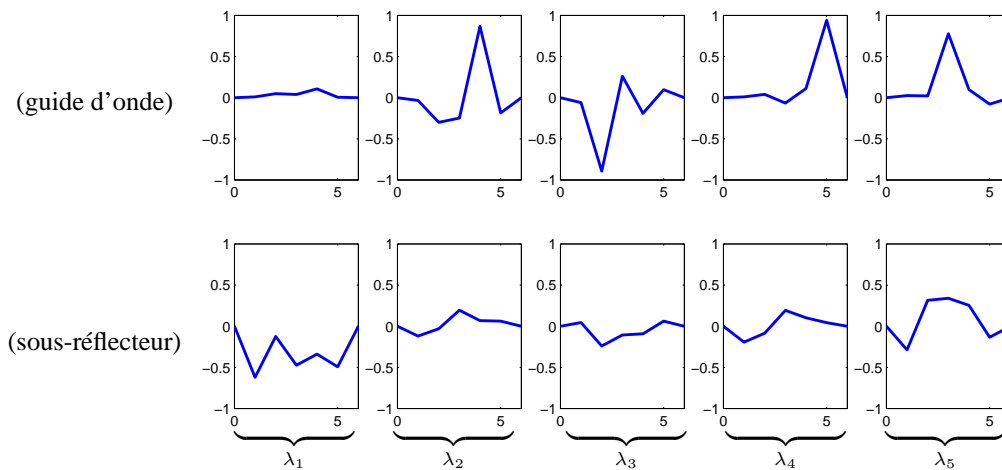


FIG. 6.17 – Vecteurs propres du hessien projeté de la fonction objectif  $J_B$  correspondant aux valeurs propres de plus petite valeur absolue.

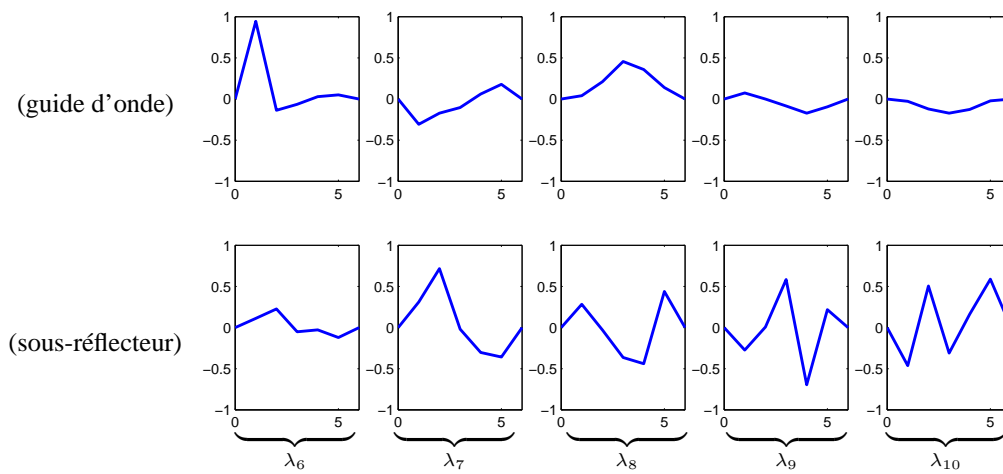


FIG. 6.18 – Vecteurs propres du hessien projeté de la fonction objectif  $J_B$  correspondant aux valeurs propres de plus grande valeur absolue.

### Jeu de Nash

Enfin, on réalise un jeu de Nash avec le partage optimal précédemment décrit. La convergence du jeu et les résultats obtenus à l'équilibre sont représentés sur la Figure 6.19. Globalement, les résultats sont peu satisfaisants, notamment à cause du caractère local de l'optimalité du partage qui semble perdue au cours de la réalisation du jeu.

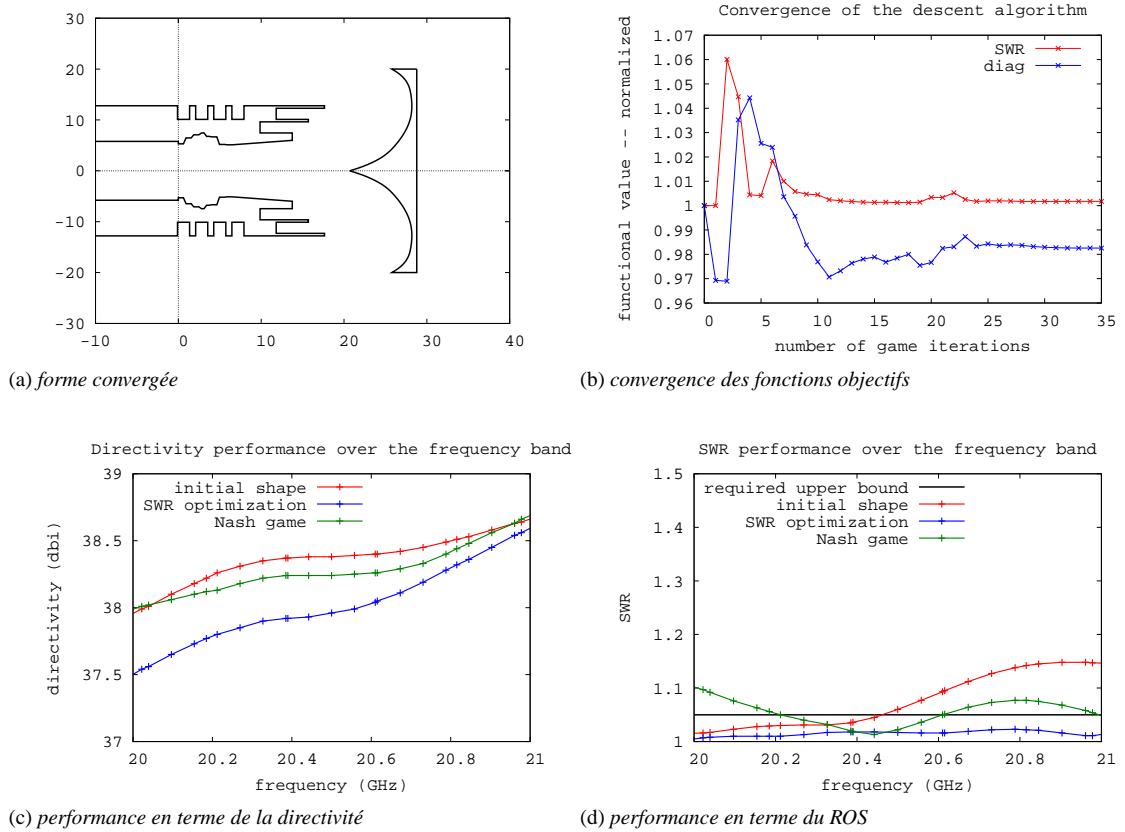


FIG. 6.19 – Résultats de la réalisation d'un jeu de Nash avec partage optimal du territoire tel que  $J_B$  soit le critère principal. Un équilibre est obtenu en 35 cycles environ. Cependant, l'optimalité du partage, qui est un résultat local, ne semble plus pertinente : en effet, dès les premières itérations du jeu, les deux critères sont dégradés. La performance de l'équilibre est représentée en vert. Si la performance en terme de directivité est améliorée par rapport à la forme précédemment optimisée pour  $J_B$ , elle est toutefois moins bonne que celle de la forme initiale. Comme prévu, le ROS est dégradé ; il reste meilleur par rapport à la forme initiale pour  $f > 20.3$  GHz et moins bon ailleurs.

### 6.3.2 Diagramme de rayonnement comme critère principal

Le critère  $J_A$  est désormais choisi comme critère principal. De la même manière, on étudie le spectre du hessien numérique de la forme convergée (par rapport à la minimisation classique de ce critère) afin d'en déduire le partage optimal. Comme pour le problème à une fréquence, on observe un sous-espace de sensibilité locale du second ordre nulle. On réalise alors une optimisation classique du coefficient de réflexion dans ce sous-espace, puis un jeu de Nash dans un sous-espace enrichi (le sous-espace de sensibilité nulle ne suffisant pas, comme on le verra).

#### Optimisation du diagramme

L'optimisation numérique de  $J_A$  en moyenne sur  $I_f$  est illustrée sur la Figure 6.20. On représente la forme convergée définie par  $x_A^*$ , la convergence de la fonction de coût, et les performances obtenues sur  $I_f$ .

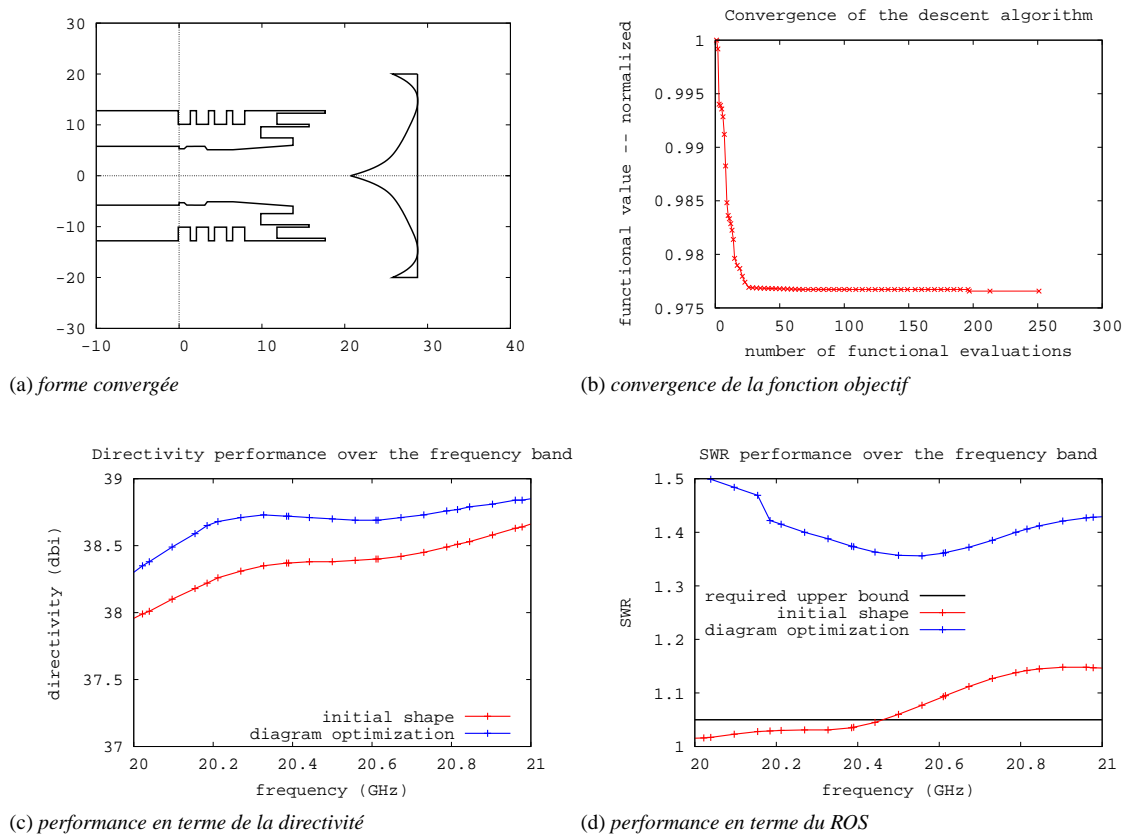


FIG. 6.20 – Optimisation numérique du diagramme de rayonnement ( $J_A$ ) sur la bande de fréquence  $I_f$ . La convergence numérique est obtenue au bout de 250 évaluations de fonctionnelle. Sur les figures (c) et (d) on représente les performances de la forme initiale (rouge) et de la forme optimisée (bleu) sur  $I_f$ . On constate que la directivité a été uniformément augmentée, alors que le ROS a été très fortement dégradé.

#### Étude numérique de la sensibilité

Une étude numérique du hessien de  $J_A$  en  $x_A^*$  est réalisée. Sur la Figure 6.21 on représente les valeurs propres, et sur les Figures 6.22 et 6.23 les vecteurs propres. On observe sur la Figure 6.22 qu'il existe un sous-espace de sensibilité nulle qui correspond à des déformations du guide d'onde uniquement. Ce



sous-espace est utilisé dans une méthodologie d'optimisation successive à la manière de la section 6.2.5. Puis on réalise un jeu de Nash en ajoutant à ce sous-espace une dimension supplémentaire (considérant alors des déformations du sous-réfecteur également).

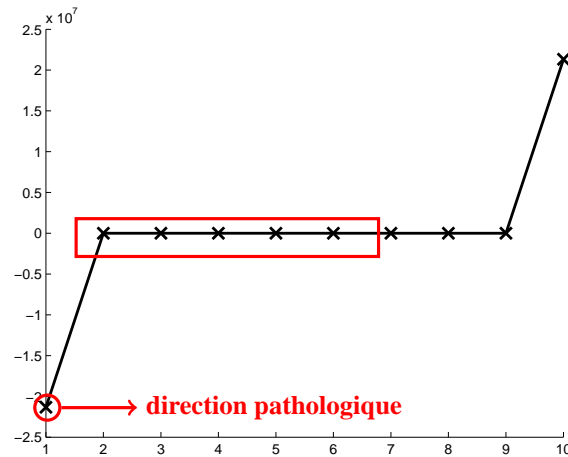


FIG. 6.21 – Valeurs propres du hessien projeté de la fonction objectif  $J_A$  évalué au point  $x_A^*$ . Le point stationnaire n'est pas un minimum local : il existe une valeur propre négative. À la précision numérique du gradient près,  $x_A^*$  est donc un point selle ; l'algorithme échoue cependant à diminuer la fonction coût le long de cette direction en raison d'un « crash » de la simulation provenant d'un problème de maillage ; autrement dit, cette direction peut être considérée comme une direction non admissible. L'espace propre de dimension 5 correspondant à l'espace de moindre sensibilité par rapport à  $J_A$  correspond aux modes dont les valeurs propres sont les plus petites : elles sont encadrées en rouge.

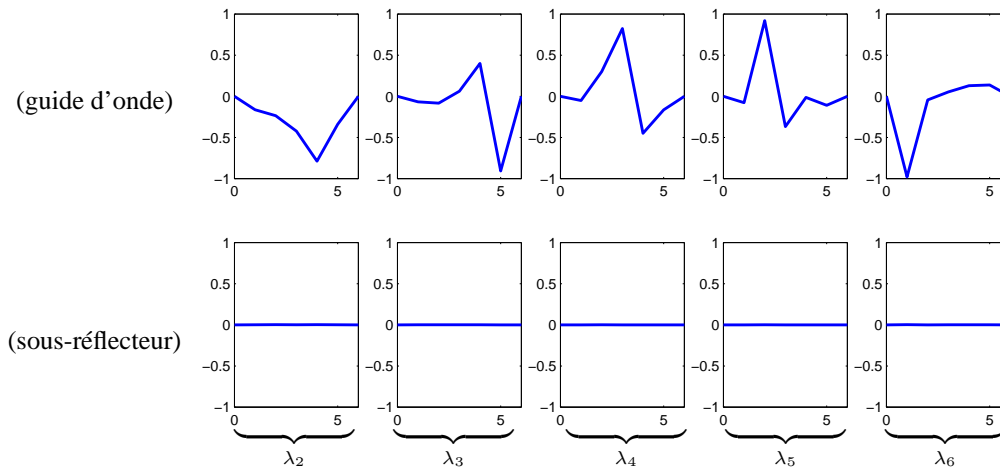


FIG. 6.22 – Vecteurs propres du hessien projeté de la fonction objectif  $J_A$  correspondant aux valeurs propres de plus petite valeur absolue : ils ne correspondent qu'à des déformations du guide uniquement.

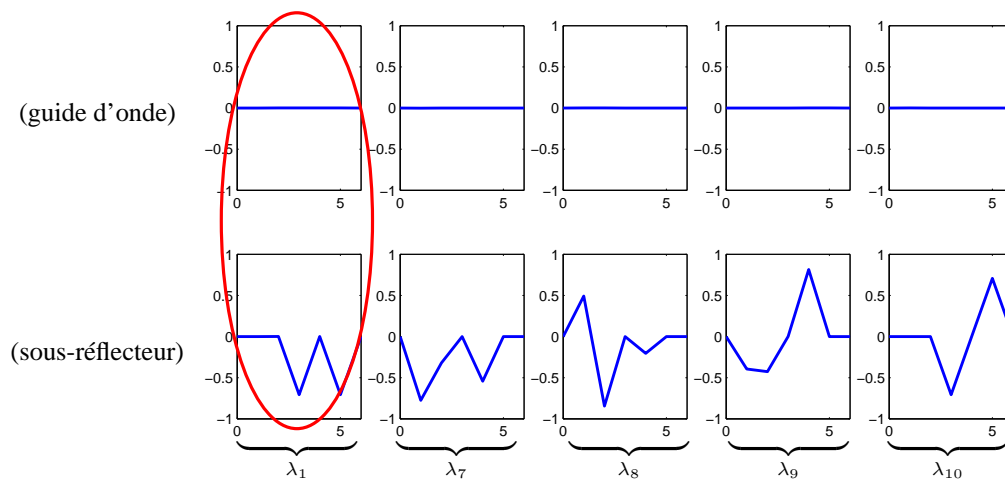


FIG. 6.23 – Vecteurs propres du hessien projeté de la fonction objectif  $J_A$  correspondant aux valeurs propres de plus grande valeur absolue : ils ne correspondent qu'à des déformations du réflecteur uniquement.

### Optimisation dans le sous-espace de moindre sensibilité

Dans un premier temps on adopte la stratégie d'optimisation successive dont les résultats sont illustrés sur la Figure 6.24. Si la directivité est effectivement très peu dégradée et que le ROS est significativement amélioré par rapport à la performance obtenue après optimisation de  $J_A$ , cela n'est pas suffisant (en ce qui concerne le ROS, la forme initiale est globalement toujours plus performante).

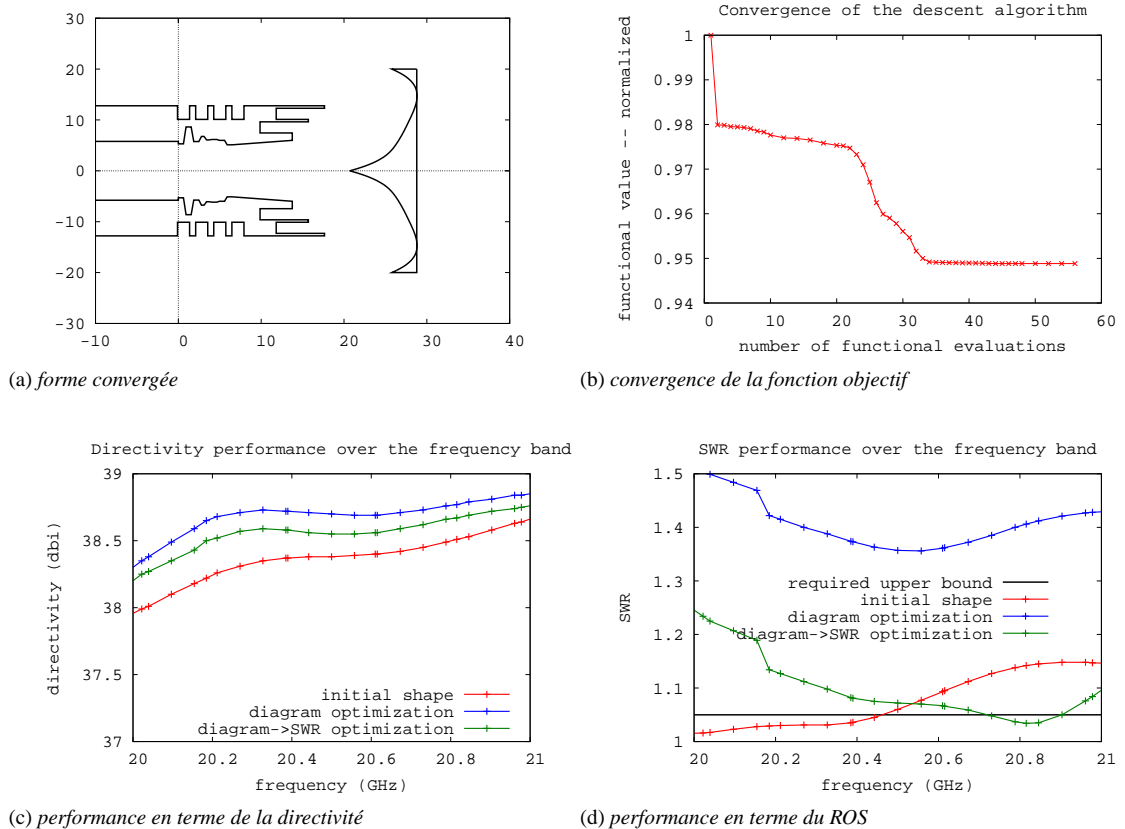


FIG. 6.24 – Optimisation numérique du ROS dans le sous-espace de moindre sensibilité (de dimension 5). Le résultat de sensibilité nulle (locale du second ordre) est vérifiée : la directivité est très peu dégradée. Le ROS est bien amélioré mais toujours moins performant que la forme initiale : le sous-espace n'est pas assez riche pour atteindre un résultat acceptable pour le concepteur d'antenne.

### Jeu de Nash

Enfin, on réalise un jeu de Nash avec le partage optimal précédemment décrit de telle sorte que le joueur  $B$  ait un territoire de dimension 6 (un mode de déformation du sous-réfecteur est donc considéré pour minimiser le ROS). La convergence du jeu et les résultats obtenus à l'équilibre sont représentés sur la Figure 6.25. Les résultats sont très satisfaisants : la directivité est très légèrement dégradée, restant supérieure à la performance de la forme initiale, tandis qu'en terme du ROS, la forme obtenue à l'équilibre est également plus performante que la forme initiale.

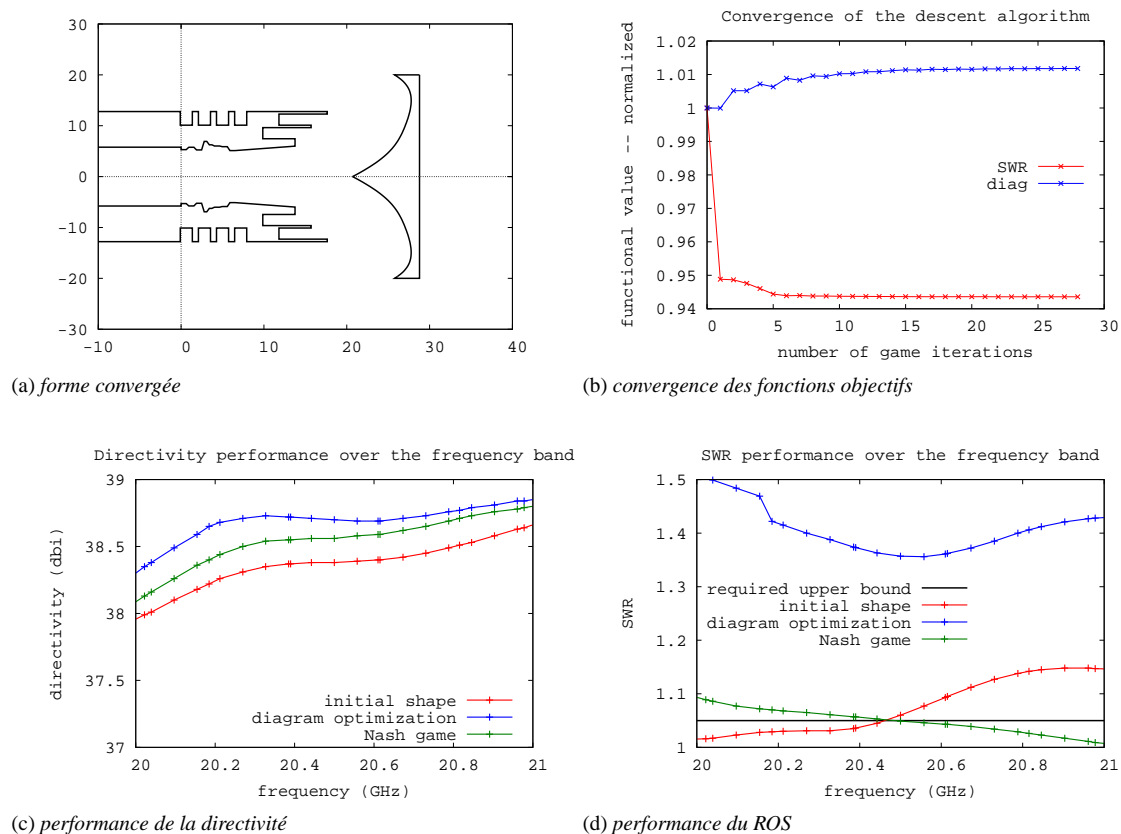


FIG. 6.25 – Résultats de la réalisation d'un jeu de Nash avec partage optimal du territoire tel que  $J_A$  soit le critère principal. Un équilibre local est obtenu en moins de 30 cycles. La performance de l'équilibre est représentée en vert. La performance en terme de directivité est très légèrement dégradée par rapport à la forme précédemment optimisée pour  $J_A$ . Le ROS est bien amélioré : le pire des cas est inférieur à 1.10 alors que celui de la forme initiale est de 1.15.

### 6.3.3 Synthèse

D'après les expériences numériques réalisées, on observe d'une part que l'étude numérique de la sensibilité des critères apporte des informations utiles pour le traitement de la compétition au moyen d'un jeu de Nash. D'autre part si le critère principal est bien choisi, alors, sur la base de cette analyse de sensibilité, le partage optimal du territoire permet effectivement d'obtenir un équilibre local dont les performances sont satisfaisantes.

## Conclusion

Dans ce chapitre on a cherché à motiver une méthodologie d'optimisation bicritère basé sur une mesure de l'antagonisme. Cette méthodologie conduit naturellement à une approche du traitement de l'optimisation multicritère par des jeux dynamiques.

Dans une approche classique basée sur les concepts de Pareto, on cherche à améliorer tous les critères à la fois « tant que cela est possible », c-à-d jusqu'à ce qu'un conflit entre deux critères apparaisse, auquel cas le concept trouvé est dit *Pareto-optimal*. Cependant, lorsque l'objectif est exprimé au moyen d'un cahier des charges, ce concept peut ne pas être satisfaisant. En outre, chercher tous les concepts Pareto-optimaux est un processus coûteux.

Dans cette nouvelle approche, on suppose que l'on part d'un point de Pareto particulier : un optimum global d'un des critères. En tant que point de Pareto, il n'est plus possible d'améliorer les deux critères à la fois. On cherche alors dans un premier temps à déterminer l'importance de l'antagonisme en ce point. Puis, pour traiter ce conflit, on propose de faire appel au formalisme des jeux non coopératifs où chaque joueur cherche à améliorer un critère dans l'espoir qu'un équilibre satisfasse le cahier des charges, sur la base d'un résultat de partage optimal de territoire. Il se trouve que dans notre cas, l'étude de l'antagonisme a mis en évidence un sous-espace de « non sensibilité », rendant le problème plus simple à résoudre, sans recourir à la réalisation du jeu.

La question reste de savoir si avec cette approche on est capable de rester sur le front de Pareto. Dans des cas plus complexes, il serait également peut-être plus pertinent de partir d'un autre point de Pareto que l'une des extrémités du front comme le suggère [32].

# CONCLUSION ET PERSPECTIVES

Cette thèse s'est articulée autour du thème central de la conception optimale d'antennes à réflecteur. Comme bon nombre de systèmes complexes en ingénierie, la géométrie est un paramètre dont la mise au point est déterminante pour le bon fonctionnement du dispositif considéré. Sur la base de la simulation de la propagation d'ondes électromagnétiques en espace libre et en régime harmonique à l'aide du logiciel SRSR développé par France Télécom R&D, on a cherché à développer et expérimenter des méthodes efficaces pour l'optimisation numérique de la forme des réflecteurs dans des configurations à symétrie de révolution.

Dans un premier temps, après avoir établi des fonctions objectifs caractéristiques pour l'optimisation d'une antenne, on a exposé nos choix sur la représentation du contrôle. Si l'optimisation de forme est intrinsèquement un problème d'optimisation en dimension infinie, l'approche paramétrique nous a semblé pertinente pour deux raisons principales : d'une part elle permet une représentation exacte du contrôle, indépendante des représentations discrètes liées à la simulation (typiquement un maillage), d'autre part elle possède une propriété régularisante (le nombre de degrés de liberté de la paramétrisation étant largement inférieur au nombre de degrés de liberté du maillage). Ce dernier point est essentiel lorsque l'on s'attache à assurer la faisabilité des concepts recherchés. En particulier on a adopté des stratégies de paramétrisation de la déformation par la technique FFD, ce qui permet un traitement assez général de la géométrie en partant de concepts existants sans se soucier d'approximations préliminaires.

La physique de la diffraction électromagnétique conduit à des problèmes difficiles d'optimisation. En particulier, on s'est rapidement aperçu que les fonctions objectifs étaient fortement multimodales. Il nous est donc apparu impératif de considérer les deux points de vue classiquement distingués au niveau du traitement numérique de l'optimisation : le point de vue global pour la robustesse et le point de vue local pour la précision. L'originalité de ce travail a consisté à aborder ces deux problèmes à l'aide de stratégies multiniveaux. Autrement dit, on a examiné les contributions potentielles de représentations hiérarchiques de la géométrie dans des contextes méthodologiques variés d'optimisation numérique. Dressons un bilan des stratégies élaborées et des résultats expérimentaux obtenus.

- Dans le contexte de l'optimisation globale, afin d'explorer l'espace de recherche, il est d'usage de faire appel aux algorithmes de population. Par ailleurs, on a vu que l'algorithme PSO était bien adapté aux stratégies multiniveaux dans la mesure où les mémoires, éléments centraux de la dynamique d'exploration de l'espace de recherche, permettent le transfert d'information entre les niveaux de paramétrisation. Une première hiérarchie de paramétrisations emboîtées a alors été considérée. Les expériences numériques réalisées semblent montrer que la robustesse de l'algorithme PSO est accrue grâce à un schéma d'enrichissement progressif. En outre, les relaxations sur les niveaux grossiers semblent avoir un effet régularisant sur les formes convergées.
- Dans le contexte de l'optimisation locale, l'élaboration d'algorithmes hiérarchiques a été motivée en deux étapes : premièrement, sur la base d'un problème modèle d'optimisation de forme, des algorithmes bigrilles idéaux directement inspirés des algorithmes multigrilles ont été décrits. Ils mettent en jeu des méthodes de descente qui jouent le rôle de relaxation. Les différentes stratégies proposées diffèrent par la nature des opérateurs de transfert entre les niveaux de paramétrisation. Le cadre du problème modèle nous a permis de conduire une étude théorique du taux de convergence et d'en déduire ainsi les stratégies efficaces. Deuxièmement, une étude numérique du hessien d'un problème de conception optimale d'antenne a été conduite afin de justifier la pertinence du problème modèle. Des expériences numériques d'optimisation d'une antenne montrent que les

stratégies susceptibles d'améliorer le taux de convergence sont effectivement en accord avec les résultats du problème modèle. Le résultat principal repose sur la nature géométrique de la paramétrisation grossière : contrairement aux multigrilles classiques, elle est composée ici de modes de haute fréquence.

- Enfin, une dernière approche hiérarchique a été abordée dans le cadre de l'optimisation en moyenne d'un critère dépendant continûment d'un paramètre. Ici, le calcul de la fonction objectif est approché par quadrature dont la précision et le coût sont liés au nombre de points. Une distribution de points est ainsi regardée comme une « grille » d'approximation. Une stratégie d'enrichissement progressif a été formulée et utilisée pour l'optimisation d'une antenne sur une bande de fréquence. D'après les résultats expérimentaux, cette stratégie est plus performante pour un coût de calcul donné. En outre, cette approche est prometteuse pour déterminer progressivement la dimension utile du niveau fin.

Avant tout, indépendamment des stratégies hiérarchiques, il apparaît que le choix de la paramétrisation est primordial. Pour une classe de formes donnée, plusieurs bases de paramétrisation sont possibles. Il s'agit donc d'un choix laissé au concepteur qui ne modifie pas nécessairement la nature du problème. Or, d'un point de vue purement géométrique, à nombre égal de degrés de liberté, le conditionnement est fortement lié au choix de la paramétrisation. La raideur induite peut donc être un facteur artificiellement pénalisant. D'après l'étude du problème modèle géométrique, on préconise l'utilisation de fonctions à support local, en particulier les courbes et surfaces B-splines : en effet, le conditionnement semble borné par une valeur indépendante du nombre de degrés de liberté ; par ailleurs, on conserve une certaine régularité de la forme tout en gardant une structure hiérarchique.

Cela étant dit, un certain nombre de critiques peuvent être formulées. Tout d'abord, il est naturel de se poser la question de l'extension des applications. Si on a considéré ici des géométries axisymétriques, on souhaite optimiser par la suite des géométries complexes 3D. Également, toujours en ce qui concerne les antennes, une plus large classe de problèmes pourra être traitée en considérant des matériaux quelconques et non plus des obstacles métalliques uniquement. On pense par exemple à l'optimisation de lentilles multicouches dont les propriétés diélectriques varient d'un milieu à l'autre. Dans ces deux configurations, la représentation des formes demande un plus grand nombre de variables et donc plus de degrés de liberté pour l'optimisation comme pour la simulation. Face à ce problème, si l'algorithme MPSO s'avère efficace, il reste néanmoins très coûteux car le nombre de particules nécessaires pour l'exploration de l'espace de recherche est lié au nombre de degrés de liberté du niveau fin. À cette critique on peut répondre que la robustesse de l'algorithme MPSO permet l'utilisation d'un plus petit nombre de particules sur les niveaux les plus grossiers, limitant ainsi le coût de la relaxation. Par ailleurs, l'algorithme PSO est bien adapté aux architectures de calcul parallèle. Pour réduire encore les coûts liés à la simulation de configurations réalistes et complexes, on peut envisager l'utilisation d'un préconditionnement par métamodèle (réseau de neurones artificiel, kriging, etc.) comme cela est proposé dans [88, 59].

Considérons le point de vue algorithmique désormais. On a montré dans des contextes indépendants que les stratégies hiérarchiques permettaient d'élargir les performances d'algorithmes classiques, soit pour la robustesse, soit pour la raideur. Il semble cependant indispensable de traiter ces deux aspects à l'aide d'un unique processus. Si on a vu que l'algorithme MPSO était robuste, rien n'a été conçu *a priori* pour la raideur numérique. En effet, une particule représente un vecteur de conception dont les éléments sont les composantes dans la base canonique d'une représentation paramétrique. Initialement, les particules sont distribuées dans des hypercubes. On se pose alors la question de la diversité d'une telle distribution de l'essaim. Comme évoqué précédemment, la raideur est liée en partie au conditionnement de la base de paramétrisation. La distance euclidienne entre les particules dans cette base ne représente donc pas nécessairement une « bonne » métrique : la diversité de l'essaim est amoindrie dans le sens où des particules « éloignées » selon cette métrique risquent d'être attirées rapidement vers le même minimum local. Intuitivement, on réalise qu'il faudrait prendre en compte une métrique qui exploite le conditionnement du problème. Pour établir cette métrique, une piste serait d'utiliser un préconditionneur basé sur l'information des évaluations de la fonctionnelle par toutes les particules. Par exemple, ce préconditionneur peut être construit à l'aide d'une régression quadratique, d'un métamodèle, ou encore d'une Analyse en Composantes Principales (ACP) comme l'algorithme CMA-ES [94] le produit dans le

cadre d'un algorithme génétique. En ce sens, tout l'essaim participe à la dynamique de recherche sur la base des meilleures performances (mémoires) et de la raideur (préconditionneur).

Dans le même ordre d'idée, si les résultats des algorithmes multiniveaux dans un contexte local (unimodal) semblent prometteurs, il arrive que la présence de minima locaux court-circuite la relaxation sur les niveaux grossiers et dirige le candidat vers un mauvais bassin d'attraction. Dans le chapitre 4 on a proposé une stratégie de type FMG pour rassembler les techniques multiniveaux qui nous ont permis d'élargir les algorithmes de base. Par ailleurs, on a déjà évoqué l'utilisation de métamodèles pour la réduction du coût de calcul. Ce même métamodèle peut également fournir une approximation du hessien pour la définition des sous-espaces de correction (cette approximation se situe comme intermédiaire entre le hessien exact, précis mais coûteux, et le hessien du problème géométrique, qui repose uniquement sur une hypothèse). Enfin, les problèmes traités n'ont été soumis qu'à des contraintes linéaires d'égalité. Dans ce cas l'espace admissible est trivial et s'étend sur chaque niveau de manière exacte. Pour traiter les non-linéarités des contraintes, il faudrait examiner des techniques plus élaborées afin de rendre les problèmes grossiers consistants par rapport aux contraintes (on pense à généraliser l'algorithme FAS par exemple).

Pour finir, on attire l'attention sur le potentiel des jeux dynamiques pour la résolution de problèmes multicritères. En particulier, à la vue des résultats de l'exercice préliminaire traité au chapitre 6, il semble qu'il faille bien distinguer dans ce domaine l'optimisation collaborative de l'optimisation compétitive. La stratégie de partage optimale de territoire a pour but de traiter « au mieux » un conflit. Or il n'y a conflit que lorsqu'un concept est Pareto-optimal. Au contraire, s'il n'y a pas de conflit, autant essayer avant tout de minimiser l'ensemble des critères. Si la Pareto-optimalité apparaît souvent comme un aboutissement du traitement d'un problème multicritère, confronté à un décideur (ou à un cahier des charges), il se peut que cette propriété ne soit pas pleinement satisfaisante. Le traitement du conflit par les jeux semble être alors une alternative à la procédure coûteuse de la recherche d'un front de Pareto. En outre, des questions théoriques en relation avec les notions de Pareto sont toujours ouvertes.





---

<b>A.1</b>	<b>Rayon spectral des algorithmes bigrille idéaux</b>	<b>187</b>
A.1.1	Transformation de similarité des matrices d'amplification	187
A.1.2	Rayon spectral du cycle idéal classique	191
A.1.3	Rayon spectral du cycle idéal en optimisation	193
A.1.4	Rayon spectral du cycle idéal avec réorganisation des valeurs propres	195
A.1.5	Rayon spectral du cycle idéal algébrique	199
<b>A.2</b>	<b>Développement logiciel : plate-forme d'optimisation</b>	<b>200</b>
A.2.1	Structure logique des composants de la plate-forme	200
A.2.2	Librairies externes	200
A.2.3	Extension de la plate-forme	201

---



## A.1 Rayon spectral des algorithmes bigrille idéaux

### A.1.1 Transformation de similarité des matrices d'amplification

#### Définitions et propriétés

On rappelle les propriétés de similarité de deux matrices carrées.

**Définition A.1.** Soit  $A$  et  $B$  deux matrices carrées de même dimension.  $A$  et  $B$  sont similaires si il existe une matrice inversible  $S$  telle que  $A = S^{-1}BS$ . On note  $A \sim B$ .

**Théorème A.1.** Si  $A \sim B$  alors  $A$  et  $B$  ont les mêmes valeurs propres.

**Théorème A.2.** Soit  $A$  et  $B$  deux matrices carrées. Alors  $AB \sim BA$ .

#### Transformations

Au chapitre 4 on a obtenu les matrices d'amplification  $G$ , d'un algorithme bigrille idéal pour le problème de Poisson et d'optimisation de forme. On a vu que  $G$  est de la forme

$$G = G_\tau \left[ I - P(RA_h P)^{-1} RA_h \right] G_\tau \quad (\text{A.1})$$

où

$$\alpha R = P^T \quad (\text{A.2})$$

pour  $\alpha > 0$  et

$$G_\tau = I - \frac{\tau}{\gamma} A_h. \quad (\text{A.3})$$

En outre on a les égalités  $A_h = \beta_h A = \beta_h \Omega \Lambda \Omega$  et  $A_{2h} = \beta_{2h} A' = \beta_{2h} \Omega' \Lambda' \Omega'$  avec  $\beta = \frac{\beta_h}{\beta_{2h}}$  indépendants de  $h$ .

Pour étudier le rayon spectral de  $G$  de chacun des problèmes on dresse une relation de similarité en utilisant les théorèmes de la section précédente :

$$\begin{aligned} G &\sim \Omega G \Omega \\ &\Omega G_\tau \Omega \left[ I - P(RA_h P)^{-1} RA_h \right] \Omega \Omega G_\tau \Omega \\ &\left( I - \frac{\tau}{\gamma} \Omega A \Omega \right) \left[ I - \beta \Omega P(A')^{-1} R \Omega \Omega A \Omega \right] \left( I - \frac{\tau}{\gamma} \Omega A \Omega \right) \\ &\left( I - \frac{\tau}{\gamma} \Lambda \right) \left[ I - \alpha \beta \Omega R^T \Omega' \Lambda'^{-1} \Omega' R \Omega \right] \left( I - \frac{\tau}{\gamma} \Lambda \right) \\ &\left( I - \frac{\tau}{\gamma} \Lambda \right) \left[ \Lambda^{-1} - \alpha \beta \Omega R^T \Omega' \Lambda'^{-1} \Omega' R \Omega \right] \Lambda \left( I - \frac{\tau}{\gamma} \Lambda \right) \\ &\sim \left[ \Lambda^{-1} - \alpha \beta \Omega R^T \Omega' \Lambda'^{-1} \Omega' R \Omega \right] \Lambda \left( I - \frac{\tau}{\gamma} \Lambda \right)^2 \end{aligned}$$

Soit  $D$  la matrice diagonale telle que  $D_{kk} = d_k = \sqrt{\lambda_k} (1 - \frac{\tau}{\gamma} \lambda_k)$ , c-à-d  $D = \Lambda^{\frac{1}{2}} (I - \frac{\tau}{\gamma} \Lambda) = (I - \frac{\tau}{\gamma} \Lambda) \Lambda^{\frac{1}{2}}$  et soit  $\sigma = \Omega' R \Omega$ . On peut réécrire

$$G \sim \left[ \Lambda^{-1} - \alpha \beta \sigma^T \Lambda'^{-1} \sigma \right] D^2 = \Sigma D^2$$

#### Structure de la matrice Sigma

On simplifie l'écriture de  $\Sigma$ . On rappelle que l'on considère deux grilles  $\mathcal{T}_h$  et  $\mathcal{T}_{2h}$  telles que  $h = \frac{1}{2^p} = \frac{1}{N+1}$  et  $2h = \frac{1}{2^{p-1}} = \frac{1}{N'+1}$ , c-à-d  $N = 2N' + 1$ . De plus, on a la propriété suivante : soit  $\lambda_k$ ,

$k = 1, \dots, N$  les valeurs propres sur  $\mathcal{T}_h$  et  $\lambda'_j, j = 1, \dots, N'$  les valeurs propres sur  $\mathcal{T}_{2h}$ . Alors, pour tout  $j = 1, \dots, N'$  on a  $\lambda'_j = \lambda_{2j}$ .

Premièrement on évalue  $\sigma$ . On rappelle de (3.24) que  $S_{ik} = \sqrt{\frac{2}{N+1}} \sin(ik\pi h) = \sqrt{2h} s_{ik}$ .

$$\sigma = S'RS = 2\sqrt{2}hs'R_s \quad (\text{A.4})$$

avec  $s'_{ik} = \sin(ik\pi 2h) = s_{2i,k} = s_{i,2k}$ . Donc un élément de  $s'R_s$  s'écrit

$$(s'R_s)_{jk} = \frac{1}{2\alpha} \sum_{i=1}^{N'} s_{2j,i} [s_{2i-1,k} + 2s_{2i,k} + s_{2i+1,k}] = \frac{1}{2\alpha} \tilde{\sigma}_{jk}. \quad (\text{A.5})$$

Donc

$$\sigma = \frac{h\sqrt{2}}{\alpha} \tilde{\sigma}. \quad (\text{A.6})$$

Après quelques transformations trigonométriques (voir [29] par exemple) on a

$$\tilde{\sigma}_{jk} = \begin{cases} \frac{1}{h} \cos^2\left(\frac{\theta_k}{2}\right) & 1 \leq j = k \leq N' \\ -\frac{1}{h} \cos^2\left(\frac{\theta_k}{2}\right) & 1 \leq j = N+1-k \leq N' \\ 0 & \text{sinon} \end{cases} \quad (\text{A.7})$$

Maintenant que nous avons une forme simplifiée de  $\sigma$  on peut réécrire  $\Sigma$

$$\Sigma = \Lambda^{-1} - \frac{\beta 2h^2}{\alpha} \tilde{\sigma}^T \Lambda'^{-1} \tilde{\sigma} = \frac{\beta 2h^2}{\alpha} C. \quad (\text{A.8})$$

où

$$C_{jk} = \sum_{i=1}^{N'} \frac{\tilde{\sigma}_{i,j} \tilde{\sigma}_{i,k}}{\lambda_{2i}}, \quad j, k = 1 \dots N. \quad (\text{A.9})$$

On distingue les différents cas possibles (voir Figure A.1) :

**(A)**  $j = k$

**(A1)**  $1 \leq j = k \leq N'$

**(A2)**  $N' + 1 = j = k = N' + 1$

**(A3)**  $N' + 2 \leq j = k \leq N$

**(B)**  $j = N + 1 - k$

**(C)**  $j \neq k$  and  $j \neq N + 1 - k$

On déduit de (A.7) les valeurs de  $C_{jk}$  :

$$C_{jk} = \begin{cases} \frac{\cos^4\left(\frac{\theta_k}{2}\right)}{\lambda_{2k} h^2} & \mathbf{A1} \\ 0 & \mathbf{A2} \\ \frac{\cos^4\left(\frac{\theta_k}{2}\right)}{\lambda_{2(N+1-k)} h^2} & \mathbf{A3} \\ -\frac{\cos^2\left(\frac{\theta_j}{2}\right) \cos^2\left(\frac{\theta_k}{2}\right)}{\lambda_{2j} h^2} & \mathbf{B} \\ 0 & \mathbf{C} \end{cases} \quad (\text{A.10})$$

Finalement, on donne une forme générale de  $\Sigma$  :

$$\Sigma_{jk} = \begin{cases} \frac{1}{\lambda_k} - \frac{2\beta \cos^4\left(\frac{\theta_k}{2}\right)}{\alpha \lambda_{2k}} & \mathbf{A1} \\ \frac{1}{\lambda_k} & \mathbf{A2} \\ \frac{1}{\lambda_k} - \frac{2\beta \cos^4\left(\frac{\theta_k}{2}\right)}{\alpha \lambda_{2(N+1-k)}} & \mathbf{A3} \\ \frac{2\beta \cos^2\left(\frac{\theta_j}{2}\right) \cos^2\left(\frac{\theta_k}{2}\right)}{\alpha \lambda_{2j}} & \mathbf{B} \\ 0 & \mathbf{C} \end{cases} \quad (\text{A.11})$$

Il apparaît que  $\Sigma$  est non nulle sur la diagonale et l'« anti-diagonale » seulement.

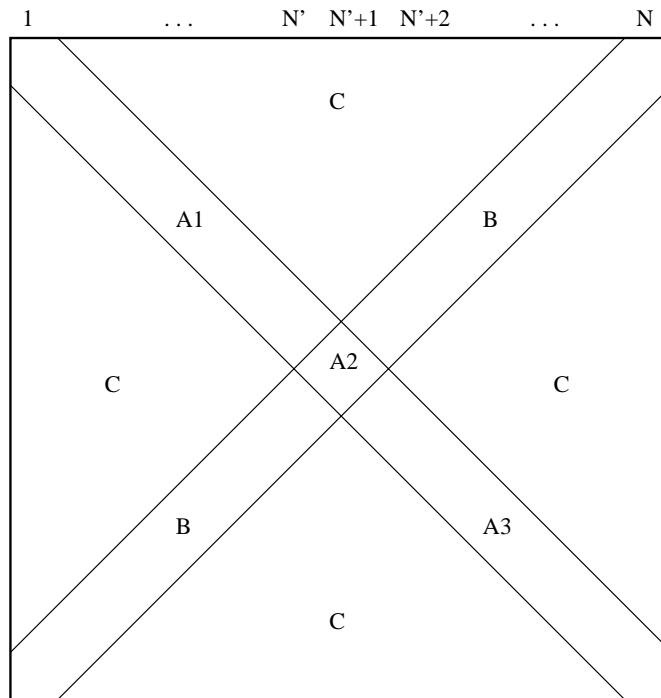


FIG. A.1 – Structure de la matrice  $\Sigma$ .

### Application au problème de Poisson

- $\alpha = 2$
- $\beta = 4$
- $\lambda_k = 2 - 2 \cos(\theta_k) = 4 \sin^2\left(\frac{\theta_k}{2}\right)$

**A1** :  $1 \leq k \leq N'$

$$\begin{aligned}
 \Sigma_{kk} &= \frac{1}{4 \sin^2(\frac{\theta_k}{2})} - 4 \frac{\cos^4(\frac{\theta_k}{2})}{4 \sin^2(\theta_k)} \\
 &= \frac{1}{4 \sin^2(\frac{\theta_k}{2})} - \frac{\cos^4(\frac{\theta_k}{2})}{4 \sin^2(\frac{\theta_k}{2}) \cos^2(\frac{\theta_k}{2})} \\
 &= \frac{1 \cos^2(\frac{\theta_k}{2})(1 - \cos^2(\frac{\theta_k}{2}))}{4 \sin^2(\frac{\theta_k}{2}) \cos^2(\frac{\theta_k}{2})} \\
 &= \frac{1}{4}
 \end{aligned}$$

**A2** :  $k = N' + 1 = \frac{1}{2h}$

$$\begin{aligned}
 \Sigma_{kk} &= \frac{1}{4 \sin^2(\frac{\theta_k}{2})} \\
 &= \frac{1}{4 \sin^2(\frac{\pi}{4})} \\
 &= \frac{1}{2}
 \end{aligned}$$

**A3** :  $N' + 2 \leq k \leq N$

$$\begin{aligned}
 \Sigma_{kk} &= \frac{1}{4 \sin^2(\frac{\theta_k}{2})} - 4 \frac{\cos^4(\frac{\theta_k}{2})}{4 \sin^2((N + 1 - k)\pi h)} \\
 &= \frac{1}{4 \sin^2(\frac{\theta_k}{2})} - 4 \frac{\cos^4(\frac{\theta_k}{2})}{4 \sin^2(\theta_k)} \\
 &= \frac{1}{4}
 \end{aligned}$$

**B** :  $1 \leq j = N + 1 - k = \frac{1}{h} - k \leq N'$

$$\begin{aligned}
 \Sigma_{jk} &= 4 \frac{\cos^2(\frac{\theta_j}{2}) \cos^2(\frac{\theta_k}{2})}{4 \sin^2(\theta_j)} \\
 &= \frac{\cos^2(\frac{\theta_j}{2}) \cos^2(\frac{\theta_k}{2})}{4 \sin^2(\frac{\theta_j}{2}) \cos^2(\frac{\theta_j}{2})} \\
 &= \frac{1 \cos^2(\frac{\theta_k}{2})}{4 \sin^2(\frac{\pi}{2} - \frac{\theta_k}{2})} \\
 &= \frac{1}{4}
 \end{aligned}$$

$$\Sigma_{kj} = \Sigma_{jk}$$

### Application au problème d'optimisation

- $\alpha = 1$
- $\beta = \frac{1}{2}$
- $\lambda_k = 4 + 2 \cos(\theta_k)$

**A1** :  $1 \leq k \leq N'$

$$\Sigma_{kk} = \frac{1}{\lambda_k} - \frac{\cos^4\left(\frac{\theta_k}{2}\right)}{\lambda_{2k}}$$

**A2** :  $k = N' + 1 = \frac{1}{2h}$

$$\begin{aligned}\Sigma_{kk} &= \frac{1}{4 + 2 \cos(\theta_k)} \\ &= \frac{1}{4 + 2 \cos\left(\frac{\pi}{2}\right)} \\ &= \frac{1}{4}\end{aligned}$$

**A3** :  $N' + 2 \leq k \leq N$

$$\begin{aligned}\Sigma_{kk} &= \frac{1}{\lambda_k} - \frac{\cos^4\left(\frac{\theta_k}{2}\right)}{4 + 2 \cos(2(N+1-k)\pi h)} \\ &= \frac{1}{\lambda_k} - \frac{\cos^4\left(\frac{\theta_k}{2}\right)}{\lambda_{2k}}\end{aligned}$$

**B** :  $1 \leq k = N + 1 - j \leq N'$

$$\Sigma_{jk} = \frac{\sin^2\left(\frac{\theta_k}{2}\right) \cos^2\left(\frac{\theta_k}{2}\right)}{\lambda_{2k}}$$

$$\Sigma_{kj} = \Sigma_{jk}$$

### A.1.2 Rayon spectral du cycle idéal classique

On rappelle que  $A_h = \frac{1}{h^2}A = \frac{1}{h^2}\Omega\Pi\Omega$  et  $A_{2h} = \frac{1}{4h^2}A' = \frac{1}{4h^2}\Omega'\Pi'\Omega'$ . Soit  $R$  l'opérateur de restriction « moyenne »  $R = \frac{1}{2}P^T$  de telle sorte que  $RA_hP = A_{2h}$ . D'après A.1.1 et A.1.1 on a la propriété de similarité suivante

$$G \sim \left[ \Pi^{-1} - 8\sigma^T\Pi'^{-1}\sigma \right] D^2 = \Sigma D^2$$

avec  $D = \Pi^{\frac{1}{2}}(I - \frac{\tau}{2}\Pi)$ ,  $\sigma = \Omega'R\Omega$ , et

$$\Sigma = \begin{pmatrix} \frac{1}{4} & & & & & & & \frac{1}{4} \\ & \frac{1}{4} & & & & & & \\ & & \ddots & & & & & \\ & & & \frac{1}{4} & & & & \\ & & & & \frac{1}{2} & & & \\ & & & & & \frac{1}{4} & & \\ & & & & & & \ddots & \\ & & & & & & & \frac{1}{4} \\ \frac{1}{4} & & & & & & & \frac{1}{4} \end{pmatrix}. \quad (\text{A.12})$$





### A.1.3 Rayon spectral du cycle idéal en optimisation

On rappelle que  $H_h = \frac{h}{6}B = \frac{h}{6}\Omega\Lambda\Omega$  et  $H_{2h} = \frac{2h}{6}B' = \frac{2h}{6}\Omega'\Lambda'\Omega'$ . Pour un souci de consistance avec les notations précédentes, on écrit les transferts en terme de l'opérateur de restriction. D'après la section A.1.1 on a la relation de similarité

$$G \sim \left[ \Lambda^{-1} - \frac{1}{2}\sigma^T\Lambda'^{-1}\sigma \right] D^2 = \Sigma D^2$$

avec  $D = \Lambda^{\frac{1}{2}}(I - \frac{\tau}{4}\Lambda)$ ,  $\sigma = \Omega'R\Omega$ , et

$$\Sigma_{jk} = \begin{cases} \frac{1}{\lambda_k} - \frac{\cos^4\left(\frac{\theta_k}{2}\right)}{\lambda_{2k}} & j = k \neq N' + 1 \\ \frac{1}{4} & j = k = N' + 1 \\ \frac{\cos^2\left(\frac{\theta_k}{2}\right)\sin^2\left(\frac{\theta_k}{2}\right)}{\lambda_{2k}} & j + k = N + 1 \\ 0 & \text{sinon} \end{cases} \quad (\text{A.15})$$

$\Sigma$ , et par conséquent  $\Sigma D^2$ , a la structure décrite à la Figure A.1. Soit  $\tau = \frac{4}{5}$  le paramètre de relaxation optimal obtenu à la section 3.2.3. Les valeurs de la matrice diagonale  $D^2$  sont

$$\begin{aligned} d_k^2 &= \lambda_k \left(1 - \frac{\tau}{4}\lambda_k\right)^2 \\ &= \lambda_k \left(1 - \frac{1}{5}(4 + 2\cos\theta_k)\right)^2 \\ &= \frac{1}{25}\lambda_k (1 - 2\cos\theta_k)^2. \end{aligned}$$

Ainsi, les entrées non nulles de  $\Sigma D^2$  sont

i)  $j = k \neq N' + 1$

$$\begin{aligned} (\Sigma D^2)_{kk} &= \frac{1}{25} \left( \frac{1}{\lambda_k} - \frac{\cos^4\left(\frac{\theta_k}{2}\right)}{\lambda_{2k}} \right) \lambda_k (1 - 2\cos\theta_k)^2 \\ &= \frac{1}{25} \frac{\lambda_{2k} - \lambda_k \cos^4\left(\frac{\theta_k}{2}\right)}{\lambda_{2k}} (1 - 2\cos\theta_k)^2 \end{aligned}$$

ii)  $j = k = N' + 1$

$$\begin{aligned} (\Sigma D^2)_{N'+1, N'+1} &= \frac{1}{25} \frac{1}{4} \lambda_{N'+1} (1 - 2\cos\frac{\pi}{2})^2 \\ &= \frac{1}{25} \frac{1}{4} 4 \\ &= \frac{1}{25} \end{aligned}$$

iii)  $j + k = N + 1$

$$(\Sigma D^2)_{jk} = \frac{1}{25} \frac{\cos^2\left(\frac{\theta_k}{2}\right)\sin^2\left(\frac{\theta_k}{2}\right)}{\lambda_{2k}} \lambda_k (1 - 2\cos\theta_k)^2$$

On cherche les couples de valeurs et vecteurs propres de  $\Sigma D^2$ . Un couple est évident : pour  $k = N' + 1$ , il est facile de vérifier que  $\mathbf{w}_k = \mathbf{e}_k$  est un vecteur propre associé à la valeur propre  $\alpha_k = \frac{1}{25}$ . L'expression des  $N - 1 = 2N'$  autres couples n'est pas aussi simple que pour l'équation de Poisson. Pour les obtenir on adopte la stratégie suivante : dans un premier temps on effectue une hypothèse sur les entrées non nulles des vecteurs propres  $\mathbf{w}_k$  ; on en déduit un système linéaire à résoudre ; ces systèmes

sont résolus en utilisant le logiciel de calcul symbolique Maple<sup>®</sup> ; enfin on vérifie l'indépendance linéaire des vecteurs propres.

Pour tout  $k \leq N'$  supposons que  $\mathbf{w}_k$  est tel que  $\mathbf{w}_k = x_k \mathbf{e}_k + x_{N+1-k} \mathbf{e}_{N+1-k}$  :

$$\mathbf{w}_k = \begin{pmatrix} x_k \\ \\ \\ x_{N+1-k} \end{pmatrix}.$$

On suppose sans perte de généralité que  $x_{N+1-k} = 1$  puisque les vecteurs propres sont définis à une constante multiplicative près (de manière équivalente on peut supposer que les vecteurs propres sont normalisés,  $x_k^2 + x_{N+1-k}^2 = 1$ , et on choisit  $x_k \geq 0$ ). Pour tout  $\mathbf{w}_k$  le système linéaire suivant doit être résolu :

$$(\Sigma D^2) \mathbf{w}_k = \alpha_k \mathbf{w}_k. \quad (\text{A.16})$$

En injectant l'hypothèse sur  $\mathbf{w}_k$  dans (A.16) on a

$$\begin{cases} x_k (\Sigma D^2)_{k,k} + (\Sigma D^2)_{k,N+1-k} & = \alpha_k x_k \\ x_k (\Sigma D^2)_{N+1-k,k} + (\Sigma D^2)_{N+1-k,N+1-k} & = \alpha_k \end{cases}$$

où les inconnues sont  $x_k$  et  $\alpha_k$ . On définit ainsi  $N'$  systèmes linéaires pour  $k = 1, \dots, N'$ . Chaque système peut être réécrit comme une équation du second ordre en  $x_k$  et une équation linéaire en  $\alpha_k$ . Il s'ensuit, si elles existent, deux paires de solutions pour chaque système : les  $2N'$  couples restant.

D'après Maple on obtient les résultats suivant :

i) les  $N'$  vecteurs linéairement indépendant

$$\begin{aligned} \mathbf{w}_k &= \cos^2\left(\frac{\theta_k}{2}\right) \left(1 - 8 \cos^2\left(\frac{\theta_k}{2}\right) + 16 \cos^4\left(\frac{\theta_k}{2}\right)\right) \mathbf{e}_k \\ &\quad - \sin^2\left(\frac{\theta_k}{2}\right) \left(1 - 8 \sin^2\left(\frac{\theta_k}{2}\right) + 16 \sin^4\left(\frac{\theta_k}{2}\right)\right) \mathbf{e}_{N+1-k} \end{aligned}$$

appartiennent au noyau de  $G$ . Ce sont des vecteurs propres de valeur propre  $\alpha_k = 0$ .

ii) Les  $N'$  vecteurs linéairement indépendants

$$\begin{aligned} \mathbf{w}_{N'+1+k} &= \sin^2\left(\frac{\theta_k}{2}\right) \left(1 + 2 \sin^2\left(\frac{\theta_k}{2}\right)\right) \mathbf{e}_k \\ &\quad + \cos^2\left(\frac{\theta_k}{2}\right) \left(1 + 2 \cos^2\left(\frac{\theta_k}{2}\right)\right) \mathbf{e}_{N+1-k} \end{aligned}$$

sont des vecteurs propres associés aux valeurs propres

$$\alpha_{N'+1+k} = \frac{1}{25} \frac{27 - 144 \cos^2\left(\frac{\theta_k}{2}\right) + 304 \cos^4\left(\frac{\theta_k}{2}\right) - 320 \cos^6\left(\frac{\theta_k}{2}\right) + 160 \cos^8\left(\frac{\theta_k}{2}\right)}{3 - 8 \cos^2\left(\frac{\theta_k}{2}\right) + 8 \cos^4\left(\frac{\theta_k}{2}\right)}.$$

Le rayon spectral du cycle idéal s'écrit donc

$$\begin{aligned} \rho(\Sigma D^2) &= \max_{k=1..N} |\alpha_k| \\ &= \max_{k=N'+1..N} |\alpha_k| \end{aligned}$$

On peut montrer que  $\alpha_k > \frac{1}{25}$  pour  $k = 1 \dots N'$  et que  $\alpha$  est une fonction monotone décroissante sur  $\theta \in ]0, \frac{\pi}{2}[$ . Alors le maximum est atteint en  $\theta = \theta_1$ , c-à-d  $\rho(\Sigma D^2) = \alpha_{N'+2}$  qui est dépendant du maillage. Lorsque on fait tendre la taille du maillage vers 0 ( $h \rightarrow 0$ ) on a  $\alpha_{N'+2} \rightarrow \frac{9}{25} = \left(\frac{3}{5}\right)^2$ . C'est exactement le comportement de 2 itérations de Jacobi (voir la section 3.3.6). En d'autres termes la correction de grille grossière est inutile.

### A.1.4 Rayon spectral du cycle idéal avec réorganisation des valeurs propres

On considère toujours la paramétrisation P1 sur chacun des niveau et  $E_{N'}^N = P$ . On conduit une analyse spectrale de la matrice d'amplification (3.88) avec  $Q = \Omega \mathbb{P} \Omega^T P$ . Comme pour le calcul du rayon de la méthode classique on adopte la méthodologie suivante :

- L'expression du problème grossier  $H' = Q^T H_h Q$  est simplifié.
- On en déduit une forme plus simple de  $I - Q H'^{-1} Q^T H_h$  et des transformations de similarité sont appliquées à  $G$  ; Il s'ensuit que  $G$  est similaire à une matrice de la forme  $\Sigma D^2$  où  $D$  est diagonale.
- Le calcul des entrées de  $\Sigma$  exhibe une structure illustrée à la Figure A.1.
- On propose une hypothèse sur la structure des vecteurs propres de  $\Sigma D^2$  et le système linéaire résultant est résolu avec l'aide de Maple. On en déduit une formule analytique des vecteurs propres et par conséquent le rayon spectral.

On rappelle que  $H_h = \beta_h B = \beta_h \Omega \Lambda \Omega^T$  et  $G_\tau = I - \frac{\tau}{4} B$ .

**a** La matrice du sous-problème est

$$\begin{aligned} Q^T H_h Q &= \beta_h P^T \Omega \mathbb{P} \Omega^T \Omega \Lambda \Omega^T \Omega \mathbb{P} \Omega^T P \\ &= \beta_h P^T \Omega \mathbb{P} \Lambda \mathbb{P} \Omega^T P \\ &= \beta_h P^T C P \end{aligned}$$

où on a noté  $C = \Omega \mathbb{P} \Lambda \mathbb{P} \Omega^T$ . Le produit  $\Omega \mathbb{P}$  permute les colonnes tandis que le produit  $\mathbb{P} \Omega$  permute les lignes. Les vecteurs propres  $S_k$  ont les propriétés de symétrie suivantes : si  $k$  est impair, alors  $S_k$  est pair par rapport à la valeur du milieu et inversement si  $k$  est pair, alors  $S_k$  est impair par rapport à la valeur du milieu (voir la section 3.2.3). Formellement cela s'écrit

$$S_{(N+1-j)k} = \begin{cases} S_{jk} & k \text{ impair} \\ -S_{jk} & k \text{ pair} \end{cases}$$

Alors on a

$$\Omega \mathbb{P} = \begin{pmatrix} S_N & \cdots & S_1 \end{pmatrix} \quad \text{et} \quad \mathbb{P} \Omega = \begin{pmatrix} S_1 & -S_2 & S_3 & \cdots & -S_{N-1} & S_N \end{pmatrix}.$$

En outre comme  $(\mathbb{P} \Omega)^T = \Omega \mathbb{P}$ , alors

$$C = \Omega \mathbb{P} \Lambda \mathbb{P} \Omega^T = \begin{pmatrix} S_1^T \\ -S_2^T \\ \vdots \\ -S_{N-1}^T \\ S_N^T \end{pmatrix} \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_N \end{pmatrix} \begin{pmatrix} S_1 & -S_2 & \cdots & -S_{N-1} & S_N \end{pmatrix}.$$

$$\begin{aligned} C_{jk} &= \sum_{i=1}^N \lambda_i (-1)^{j+k} s_{ij} s_{ik} \\ &= 2h (-1)^{j+k} \sum_{i=1}^N (4 + 2 \cos \theta_i) \sin(i\theta_j) \sin(i\theta_k) \end{aligned}$$

L'orthogonalité des vecteurs propres donne

$$\begin{aligned} \sum_{i=1}^N 4 \sin(i\theta_j) \sin(i\theta_k) &= \frac{4}{2h} \delta_{jk} \\ \sum_{i=1}^N 2 \cos(\theta_i) \sin(i\theta_j) \sin(i\theta_k) &= \frac{1}{2h} (\delta_{j,k-1} + \delta_{j,k+1}) \end{aligned}$$

Donc

$$C = \begin{pmatrix} 4 & -1 & & & \\ -1 & 4 & -1 & & \\ & & \ddots & & \\ & & & -1 & 4 & -1 \\ & & & & -1 & 4 \end{pmatrix}.$$

Finalement, de simples multiplications avec l'opérateur d'interpolation linéaire  $P$  donne

$$Q^T H_h Q = \beta_h P^T C P = 4\beta_h I \quad (\text{A.17})$$

**b** D'après (A.17) on a

$$\begin{aligned} I - Q(Q^T H_h Q)^{-1} Q^T H_h &= I - \frac{1}{4\beta_h} Q Q^T \beta_h \Omega \Lambda \Omega \\ &= I - \frac{1}{4} \Omega \mathbb{P} \Omega P P^T \Omega \mathbb{P} \Lambda \Omega \end{aligned}$$

On applique des transformations de similarité sur  $G$  en utilisant le fait que  $\Omega$  est symétrique et orthogonal :

$$\begin{aligned} G &= G_\tau \left[ I - Q(Q^T H_h Q)^{-1} Q^T H_h \right] G_\tau \\ &= G_\tau \Omega \Omega \left[ I - \frac{1}{4} \Omega \mathbb{P} \Omega P P^T \Omega \mathbb{P} \Lambda \Omega \right] \Omega \Omega G_\tau \\ &= G_\tau \Omega \left[ I - \frac{1}{4} \mathbb{P} \Omega P P^T \Omega \mathbb{P} \Lambda \right] \Lambda^{-1} \Lambda \Omega G_\tau \\ &\sim \Omega G_\tau \Omega \left[ \Lambda^{-1} - \frac{1}{4} \mathbb{P} \Omega P P^T \Omega \mathbb{P} \right] \Lambda \Omega G_\tau \Omega \\ &\sim \left[ \Lambda^{-1} - \frac{1}{4} \mathbb{P} \Omega P P^T \Omega \mathbb{P} \right] \Lambda \left( I - \frac{\tau}{4} \Lambda \right)^2 = \Sigma D^2 \end{aligned}$$

où on a noté

$$\Sigma = \left[ \Lambda^{-1} - \frac{1}{4} \mathbb{P} \Omega P P^T \Omega \mathbb{P} \right] \quad \text{et} \quad D^2 = \Lambda \left( I - \frac{\tau}{4} \Lambda \right)^2. \quad (\text{A.18})$$

De plus, puisque  $\mathbb{P}$  est orthogonal, on a  $G \sim (\mathbb{P} \Sigma \mathbb{P})(\mathbb{P} D^2 \mathbb{P})$  avec

$$\mathbb{P} \Sigma \mathbb{P} = \left[ \mathbb{P} \Lambda^{-1} \mathbb{P} - \frac{1}{4} \mathbb{P} \Omega P P^T \Omega \right] \quad \text{et} \quad \mathbb{P} D^2 \mathbb{P} = \mathbb{P} \Lambda \mathbb{P} \left( I - \frac{\tau}{4} \mathbb{P} \Lambda \mathbb{P} \right)^2. \quad (\text{A.19})$$

**c** Pour évaluer  $\mathbb{P} \Sigma \mathbb{P}$  on doit calculer  $\Omega P P^T \Omega$ . Pour  $i = 1 \dots N'$  et  $k = 1 \dots N$  on a

$$(P^T \Omega)_{ik} = \frac{1}{2} (S_{2i-1,k} + 2S_{2i,k} + S_{2i+1,k})$$

Donc, puisque  $\Omega P P^T \Omega = (P^T \Omega)^T P^T \Omega$ , alors

$$\begin{aligned} (\Omega P P^T \Omega)_{jk} &= \frac{1}{4} \sum_{i=1}^{N'} (S_{2i-1,j} + 2S_{2i,j} + S_{2i+1,j}) (S_{2i-1,k} + 2S_{2i,k} + S_{2i+1,k}) \\ &= \frac{2h}{4} \sum_{i=1}^{N'} s_{2i-1,j} (s_{2i-1,k} + s_{2i,k} + s_{2i+1,k}) \\ &\quad + 2s_{2i,j} (s_{2i-1,k} + s_{2i,k} + s_{2i+1,k}) \\ &\quad + s_{2i+1,j} (s_{2i-1,k} + s_{2i,k} + s_{2i+1,k}) \end{aligned}$$

où on a posé  $S_{jk} = \sqrt{2h}s_{jk} = \sqrt{2h} \sin(j\theta_k)$ . En utilisant la formule

$$\sin(a - b) + \sin(a + b) = 2 \sin(a) \cos(b) \quad (\text{A.20})$$

on obtient

$$\begin{aligned} s_{2i-1,j} + s_{2i+1,j} &= \sin(2i\theta_j - \theta_j) + \sin(2i\theta_j + \theta_j) \\ &= 2 \sin(2i\theta_j) \cos(\theta_j) \\ &= 2s_{ij} \cos(\theta_j) \end{aligned}$$

donc

$$(\Omega P P^T \Omega)_{jk} = h(1 + \cos \theta_j) \sum_{i=1}^{N'} s_{2i,j} (s_{2i-1,k} + s_{2i,k} + s_{2i+1,k}).$$

D'après la définition de  $\tilde{\sigma}_{jk}$  en (A.5) et puisque  $\Omega P P^T \Omega$  est symétrique, on a

$$(\Omega P P^T \Omega)_{jk} = 2h \cos^2 \left( \frac{\theta_j}{2} \right) \begin{cases} \tilde{\sigma}_{jk} & 1 \leq j \leq N', \quad \forall k \\ \tilde{\sigma}_{kj} & 1 \leq k \leq N', \quad \forall j \\ \frac{1}{h} \cos^2 \left( \frac{\theta_k}{2} \right) & N' + 2 \leq j = k \leq N \\ 0 & \text{sinon} \end{cases}$$

D'après (A.7) qui donne les valeurs de  $\tilde{\sigma}_{jk}$  pour  $j = 1 \dots N'$  et  $k = 1 \dots N$  alors

$$(\Omega P P^T \Omega)_{jk} = \begin{cases} 2 \cos^4 \left( \frac{\theta_k}{2} \right) & 1 \leq j = k \leq N, \quad j, k \neq N' + 1 \\ -2 \cos^2 \left( \frac{\theta_j}{2} \right) \cos^2 \left( \frac{\theta_k}{2} \right) & j + k = N + 1, \quad j, k \neq N' + 1 \\ 0 & \text{sinon} \end{cases}$$

et

$$(\mathbb{P}\Sigma\mathbb{P})_{jk} = \frac{\delta_{jk}}{\lambda_{N+1-k}} - \frac{1}{4} (\Omega P P^T \Omega)_{jk} \quad (\text{A.21})$$

Avec

$$\lambda_{N+1-k} = 4 + 2 \cos(\theta_{N+1-k}) = 4 - 2 \cos \theta_k$$

on obtient finalement

i)  $j = k$  et  $j, k \neq N' + 1$

$$(\mathbb{P}\Sigma\mathbb{P})_{jk} = \frac{1 - \cos^4 \left( \frac{\theta_k}{2} \right) (2 - \cos \theta_k)}{4 - 2 \cos \theta_k}$$

ii)  $j = k = N' + 1$

$$(\mathbb{P}\Sigma\mathbb{P})_{jk} = \frac{1}{\lambda_{N'+1}} = \frac{1}{4}$$

iii)  $j + k = N + 1$  et  $j, k \neq N' + 1$

$$(\mathbb{P}\Sigma\mathbb{P})_{jk} = \frac{1}{2} \cos^2 \left( \frac{\theta_k}{2} \right) \sin^2 \left( \frac{\theta_k}{2} \right)$$

iv) sinon

$$(\mathbb{P}\Sigma\mathbb{P})_{jk} = 0$$

Il s'ensuit que la structure des éléments non nuls de  $(\mathbb{P}\Sigma\mathbb{P})$  est la même que celle décrite à la Figure A.1.

**d** On calcule désormais les valeurs et vecteurs propres de  $(\mathbb{P}\Sigma\mathbb{P})(\mathbb{P}D^2\mathbb{P})$ . D'après (A.18) la matrice diagonale  $\mathbb{P}D^2\mathbb{P}$  est telle que

$$\mathbb{P}D^2\mathbb{P} = \mathbb{P}\Lambda\mathbb{P}\left(I - \frac{\tau}{4}\mathbb{P}\Lambda\mathbb{P}\right)^2$$

dont les entrées  $d_k^2$  sont

$$d_k^2 = \lambda_{N+1-k} \left( I - \frac{\tau}{4}\lambda_{N+1-k} \right)^2$$

Avec  $\tau = \frac{4}{5}$  en particulier on a

$$d_k^2 = \frac{1}{25}\lambda_{N+1-k}(1 + 2\cos\theta_k)^2$$

Du fait de la structure de  $(\mathbb{P}\Sigma\mathbb{P})(\mathbb{P}D^2\mathbb{P})$ , il est évident que le vecteur  $\mathbf{w}_{N'+1} = \mathbf{e}_{N'+1}$  est un vecteur propre associé à la valeur propre  $\alpha_{N'+1} = \frac{1}{4}d_{N'+1}^2 = \frac{1}{25}$ . Pour les autres couples on adopte la même stratégie qu'à la section 3.4.5.

Soit  $k \leq N'$  et considérons l'hypothèse suivante sur la structure des vecteurs propres  $\mathbf{w}_k$  de  $\mathbb{P}\Sigma\mathbb{P}\mathbb{P}D^2\mathbb{P}$  : supposons que  $\mathbf{w}_k$  est tel que  $\mathbf{w}_k = x_k\mathbf{e}_k + x_{N+1-k}\mathbf{e}_{N+1-k}$ . Par ailleurs, sans perte de généralité, soit  $x_{N+1-k} = 1$ , soit  $\mathbf{w}_k$  est normalisé (c-à-d  $x_k^2 + x_{N+1-k}^2$ ) et  $x_k \geq 0$ . Pour chaque  $\mathbf{w}_k$  on doit résoudre le système linéaire

$$(\mathbb{P}\Sigma\mathbb{P}\mathbb{P}D^2\mathbb{P})\mathbf{w}_k = \alpha_k\mathbf{w}_k.$$

En injectant l'hypothèse on obtient

$$\begin{cases} x_k(\mathbb{P}\Sigma\mathbb{P}\mathbb{P}D^2\mathbb{P})_{k,k} + (\mathbb{P}\Sigma\mathbb{P}\mathbb{P}D^2\mathbb{P})_{k,N+1-k} & = \alpha_k x_k \\ x_k(\mathbb{P}\Sigma\mathbb{P}\mathbb{P}D^2\mathbb{P})_{N+1-k,k} + (\mathbb{P}\Sigma\mathbb{P}\mathbb{P}D^2\mathbb{P})_{N+1-k,N+1-k} & = \alpha_k \end{cases}$$

où les inconnues sont  $x_k$  et  $\alpha_k$ . On a donc défini  $N'$  systèmes linéaires. Chaque système peut être séparé en une équation du second degré en  $x_k$  et une équation linéaire en  $\alpha_k$ . On a ainsi deux couples  $(x_k, \alpha_k)$  de solutions par système : les  $2N'$  couples restant.

D'après Maple on a le résultat suivant :

i) Les  $N'$  vecteurs non nuls et linéairement indépendants

$$\begin{aligned} \mathbf{w}_k &= \left( (4\sin^2(\frac{\theta_k}{2}) - 3)^2 \sin^2(\frac{\theta_k}{2}) - 1 \right) \mathbf{e}_k \\ &\quad + (4\sin^2(\frac{\theta_k}{2}) - 3)^2 \sin^2(\frac{\theta_k}{2}) \mathbf{e}_{N+1-k} \end{aligned}$$

appartiennent au noyau de  $G$ . Ce sont des vecteurs propres associés à la valeur propre  $\alpha_k = 0$ .

ii) Les  $N'$  vecteurs non nuls et linéairement indépendants

$$\begin{aligned} \mathbf{w}_{N'+1+k} &= \sin^2(\frac{\theta_k}{2})(2\sin^2(\frac{\theta_k}{2}) - 3)\mathbf{e}_k \\ &\quad (2\sin^2(\frac{\theta_k}{2}) + 1)(\sin^2(\frac{\theta_k}{2}) - 1)\mathbf{e}_{N+1-k} \end{aligned}$$

sont des vecteurs propres associés aux valeurs propres

$$\alpha_{N'+1+k} = \frac{1}{25} \left( 1 - 8\cos^2(\frac{\theta_k}{2}) \sin^2(\frac{\theta_k}{2}) \cos^2(\theta_k) \right)$$

Le rayon spectral du cycle idéal est donc

$$\begin{aligned} \rho(\mathbb{P}\Sigma\mathbb{P}\mathbb{P}D^2\mathbb{P}) &= \max_{k=1..N} |\alpha_k| \\ &= \max_{k=N'+1..N} \alpha_k \end{aligned}$$

Si on étudie la fonction  $f : \theta \mapsto 1 - 8\cos^2(\theta) \sin^2(\theta) \cos^2(2\theta)$  et sa dérivée sur l'intervalle  $]0, \frac{\pi}{4}[$  on peut montrer que  $f$  est positive et bornée supérieurement par 1. Alors, pour tout  $k > N' + 1$ ,  $\alpha_k$  est plus petit que  $\frac{1}{25}$ . Puisque  $\alpha_{N'+1} = \frac{1}{25}$  alors le rayon spectral est donné par

$$\rho(\mathbb{P}\Sigma\mathbb{P}\mathbb{P}D^2\mathbb{P}) = \alpha_{N'+1} = \frac{1}{25}. \quad (\text{A.22})$$

### A.1.5 Rayon spectral du cycle idéal algébrique

On rappelle qu'on a  $Q = \Omega_2$ ,  $G_\tau = I - \frac{\tau}{4}B$  et  $H_h = \beta_h B$  où  $\beta_h = \frac{h}{6}$ . On examine la matrice GCA  $Q^T H_h Q$  du problème grossier :

$$\begin{aligned} Q^T H_h Q &= \beta_h \Omega_2^T [\Omega_1 \quad \Omega_2] \begin{pmatrix} \Lambda_1 & 0 \\ 0 & \Lambda_2 \end{pmatrix} \begin{bmatrix} \Omega_1^T \\ \Omega_2^T \end{bmatrix} \Omega_2 \\ &= \beta_h [0 \quad I] \begin{pmatrix} \Lambda_1 & 0 \\ 0 & \Lambda_2 \end{pmatrix} \begin{bmatrix} 0 \\ I \end{bmatrix} \\ &= \beta_h \Lambda_2. \end{aligned}$$

Il est donc facile d'inverser le problème grossier, ce qui donne

$$\begin{aligned} \Sigma &= I - Q (Q^T H_h Q)^{-1} Q^T H_h \\ &= I - \beta_h^{-1} \Omega_2 \Lambda_2^{-1} \Omega_2^T \beta_h B \\ &= I - \Omega_2 \Lambda_2^{-1} \Omega_2^T [\Omega_1 \quad \Omega_2] \begin{pmatrix} \Lambda_1 & 0 \\ 0 & \Lambda_2 \end{pmatrix} \begin{bmatrix} \Omega_1^T \\ \Omega_2^T \end{bmatrix} \\ &= I - \Omega_2 \Lambda_2^{-1} \Lambda_2 \Omega_2^T \\ &= I - \Omega_2 \Omega_2^T \end{aligned}$$

En appliquant des transformations de similarité sur  $G$  on obtient

$$\begin{aligned} G &= G_\tau \Sigma G_\tau \\ &= G_\tau \Omega \Omega^T \Sigma \Omega \Omega^T G_\tau \\ &\sim (\Omega^T \Sigma \Omega) (\Omega^T G_\tau \Omega)^2 \end{aligned}$$

où

$$\begin{aligned} \Omega^T \Sigma \Omega &= I - \Omega^T \Omega_2 \Omega_2^T \Omega \\ &= I - \begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} \end{aligned}$$

et

$$\Omega^T G_\tau \Omega = I - \frac{\tau}{4} \Lambda$$

Donc  $G$  est similaire à

$$G \sim \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} \left( I - \frac{\tau}{4} \Lambda \right)^2 = \begin{pmatrix} \left( I - \frac{\tau}{4} \Lambda_1 \right)^2 & 0 \\ 0 & 0 \end{pmatrix}$$

de telle sorte que le rayon spectral est donné par

$$\rho(G) = \left( 1 - \frac{\tau}{4} \min \Lambda_1 \right)^2 = \left( 1 - \frac{\tau}{4} \lambda_{N'+1} \right)^2.$$

En particulier, avec  $\tau = \frac{4}{5}$  et sachant que  $\lambda_{N'+1} = 4$  on a

$$\rho(G) = \left( \frac{1}{5} \right)^2 = \frac{1}{25}.$$

Le rayon spectral est identique à la méthode de réorganisation des paires de vecteurs et valeurs propres.



## A.2 Développement logiciel : plate-forme d'optimisation

les travaux de cette thèse ont abouti au développement d'une plate-forme d'optimisation de forme pour la conception optimale d'antenne à réflecteur. Cette plate-forme intègre :

- des outils de manipulation géométrique de type CAO, éventuellement en formulation *Free-Form* ;
- des outils d'optimisation numérique robuste (PSO) et locale (méthodes de descente : plus grande descente, gradient Conjugué, Quasi-Newton BFGS) ;
- une interface avec le logiciel de simulation numérique des équations de Maxwell en régime harmonique pour les Structures Rayonnantes à Symétrie de Révolution (SRSR) développé à France Télécom R&D La Turbie ;
- des algorithmes avancés utilisant des techniques hiérarchiques inspirées des Multigrilles et visant à améliorer la robustesse et/ou la vitesse de convergence (tels qu'ils ont été exposés dans la présente thèse) ;
- des techniques de parallélisation pour accélérer le code d'optimisation en réalisant des simulations simultanées dans le cadre d'une optimisation robuste, multipoint, ou encore pour le calcul du gradient par différences finies ;
- des outils pratiques de compilation du code, notamment avec GNU/Make.

### A.2.1 Structure logique des composants de la plate-forme

Le processus d'optimisation numérique de la forme d'une antenne respecte le schéma classique *géométrie* → *maillage* → *analyse* → *optimisation (évaluation/sélection)* → *déformation* → *géométrie* → ... . Chacune de ces tâches correspond à un *compartiment logiciel*. La plate-forme est l'assemblage de ces compartiments dont une illustration est donnée sur la Figure A.2.

Concrètement, ces compartiments correspondent à des routines de calcul dont le domaine d'application est plus ou moins étendu :

- Le compartiment *déformation* comprend des outils très généraux de CAO [39] (Bézier, B-splines, etc.) ainsi que des outils de FFD [86].
- Le compartiment *géométrie* assure l'interface entre les outils de CAO et le format de description des géométries de SRSR.
- Les compartiments *maillage* et *analyse* sont assurés par des routines internes de SRSR.
- Le compartiment *optimisation* comprend des algorithmes classiques d'optimisation numérique paramétrique (méthodes de descente, PSO, [45, 54, 77]), réutilisable dans d'autres contextes et voué à être enrichi.
- Le cœur de la plate-forme assemble ces outils dans un processus d'optimisation qui met en jeu des algorithmes élaborés, hiérarchiques, et dont le but est d'améliorer la robustesse et la vitesse de convergence.

La plate-forme comprend également deux dossiers de dépôt : un dépôt de géométries au format SRSR et un dépôt de fichiers *cibles* pour les problèmes inverses.

### A.2.2 Bibliothèques externes

L'ensemble de la plate-forme est codé en FORTRAN 77, excepté les routines C de lecture des fichiers de géométrie SRSR. Les *flottants* sont en double précision. Elle dépend des bibliothèques externes BLAS<sup>1</sup> et LAPACK<sup>2</sup>. Par ailleurs, il existe deux versions du code : une version séquentielle et une version parallèle basée sur les instructions MPI. Les routines parallèles sont utilisées pour les évaluations simultanées de la fonctionnelle dans le contexte de l'algorithme PSO et/ou dans le contexte d'une optimisation multifr-

<sup>1</sup><http://www.netlib.org/blas>

<sup>2</sup><http://www.netlib.org/lapack>

quence ou encore pour le calcul d'un gradient par différences finies. Le code nécessite alors une librairie qui intègre les routines de communication MPI comme par exemple Open MPI<sup>3</sup> ou MPICH2<sup>4</sup>.

### A.2.3 Extension de la plate-forme

Chaque composant est prévu pour être enrichi ou remplacé, et ce de manière indépendante (moyennant parfois quelques routines d'interfaces) :

- les techniques de paramétrisation et de déformation bidimensionnelle s'étendent très simplement aux problèmes tridimensionnels. En particulier, la technique FFD s'applique pour tout objet inclus dans une *boîte* 3D. Si cet objet est défini par un maillage, chacun des points du maillage est déformé selon le même principe, c'est-à-dire grâce au déplacements de points de contrôle qui agissent sur des produits tensoriels de fonctions de base (Bernstein, B-spline, etc.). Les outils de déformation inclus dans les librairies sont déjà prêts pour leur utilisation dans des configurations 3D ;
- la librairie des algorithmes classiques d'optimisation peut être facilement enrichie par d'autres algorithmes (recuit simulé, algorithmes génétiques, simplexe, etc.) ; ces algorithmes sont ainsi intégrée de manière transparente dans les processus hiérarchiques ; également, au sein d'une même hiérarchie, plusieurs algorithmes peuvent être utilisés ; cela est prévu dans le but d'expérimenter des techniques multiniveaux d'hybridation ;
- enfin, il est prévu que le calcul du gradient soit remplacé ultérieurement par un code adjoint.

---

<sup>3</sup><http://www.open-mpi.org>

<sup>4</sup><http://www.mcs.anl.gov/research/projects/mpich2>

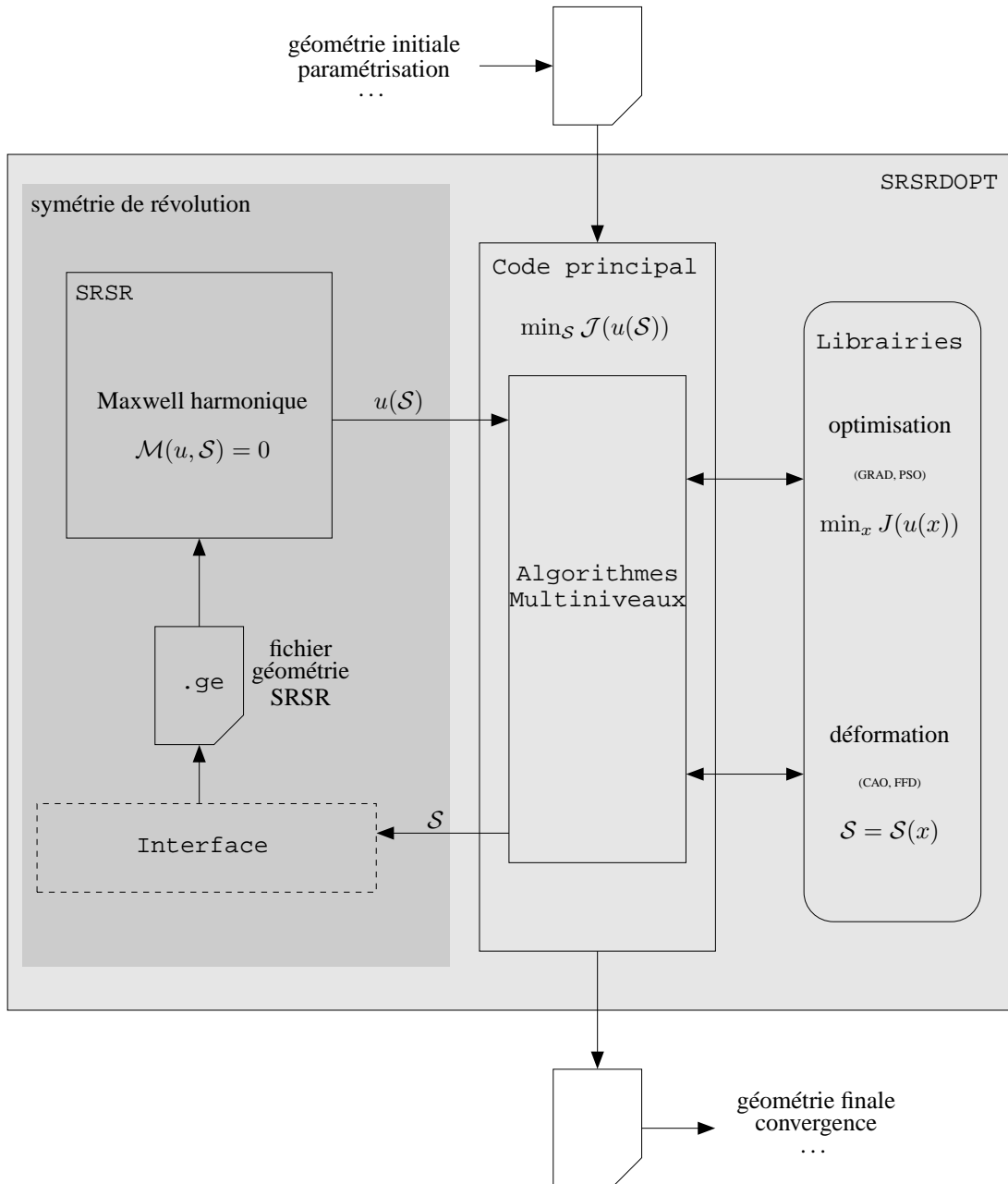


FIG. A.2 – Illustration des unités logiques de la plate-forme d'optimisation de structures rayonnantes à symétrie de révolution.

# BIBLIOGRAPHIE

- [1] Badr Abou el Majd. *Algorithmes hiérarchiques et stratégies de jeux pour l'optimisation multidisciplinaire, Application à l'optimisation de la voilure d'un avion d'affaires*. PhD thesis, Université de Nice-Sophia Antipolis, 2007. (Cité pages 44 et 162.)
- [2] Grégoire Allaire, Frédéric De Gournay, François Jouve, and Anca-Maria Toader. Structural optimization using topological and shape sensitivity via a level set method. *Control and Cybernetics*, 34 :59–80, 2005. (Cité page 30.)
- [3] Michele Andreoli, Janka Ales, and Jean-Antoine Désidéri. Free-form-deformation parameterization for multilevel 3D shape optimization in aerodynamics. Research Report RR-5019, INRIA, 2003. (Cité page 132.)
- [4] Thomas S. Angell and Andreas Kirsch. *Optimization Methods in Electromagnetic Radiation*. Springer Monographs in Mathematics. Springer-Verlag, 2004. (Cité page 21.)
- [5] Eyal Arian and Shlomo Ta'asan. Smoothers for optimization problems. In N. Duane Melson, T.A. Manteuffel, S.F. McCormick, and C.C. Douglas, editors, *Seventh Copper Mountain Conference on Multigrid Methods*, volume CP3339, pages 15–30, Hampton, VA, US, 1995. NASA, NASA Conference Publication. (Cité page 59.)
- [6] John S. Asvestas. The physical optics method in electromagnetic scattering. *Journal of Mathematical Physics*, 21(2) :290–299, 1980. (Cité page 14.)
- [7] Thomas Bäck. *Evolutionary algorithms in theory and practice : evolution strategies, evolutionary programming, genetic algorithms*. Oxford University Press, Oxford, UK, 1996. (Cité page 44.)
- [8] Raymond Bills. *Problème extérieur pour les équations de Maxwell Application aux antennes de révolution*. PhD thesis, Université de Nice, 1982. (Cité pages 8, 12, 13, 15 et 20.)
- [9] Raymond Bills, R. Carré, and Jean-Louis Desvilles. *SRSR-D v4 User Guide*. France Télécom R&D La Turbie, 2006. (Cité page 15.)
- [10] Louis Blanchard. *Conception d'antenne avec optimisation des lobes réseau : Application au partitionnement en sous-réseaux d'une antenne radar*. PhD thesis, ENSMP, November 2007. (Cité pages 30 et 141.)
- [11] J.F. Bonnans, J.Ch. Gilbert, C. Lemaréchal, and C. Sagastizábal. *Numerical Optimization – Theoretical and Practical Aspects*. Universitext. Springer Verlag, Berlin, second edition, 2006. (Cité page 28.)
- [12] Alfio Borzi and Volker H. Schulz. Multigrid methods for PDE optimization. 2009. (Cité page 59.)
- [13] Alain Bossavit. *Électromagnétisme en vue de la modélisation*. Springer-Verlag, deuxième édition, 2004. (Cité pages 7 et 11.)
- [14] Achi Brandt. Multi-level adaptive solutions to boundary-value problems. *Mathematics of Computation*, *American Mathematical Society*, 31(138) :333–390, 1977. (Cité page 59.)

- [15] Jürgen Branke, Kalyanmoy Deb, Kaisa Miettinen, and Roman Slowinski, editors. *Multiobjective Optimization. Interactive and Evolutionary Approaches*. Springer. Lecture Notes in Computer Science Vol. 5252, Berlin, Germany, 2008. (Cité page 139.)
- [16] William L. Briggs, Van Emden Henson, and Steve F. McCormick. *A Multigrid Tutorial (2nd ed.)*. Society for Industrial and Applied Mathematics, second edition, 2000. (Cité page 59.)
- [17] Jean Cea. Conception optimale ou identification de formes, calcul rapide de la dérivée directionnelle de la fonction coût. *RAIRO, Modélisation Mathématiques et Analyse Numérique*, 20(3) :371–402, 1986. (Cité page 30.)
- [18] Jean Cea, Stéphane Garreau, Philippe Guillaume, and Mohamed Masmoudi. The shape and topological optimizations connection. *Computer Methods in Applied Mechanics and Engineering*, 188(4) :713 – 726, 2000. IVth World Congress on Computational Mechanics. (II). Optimum. (Cité page 30.)
- [19] Benoît Chaigne. Shape reconstruction in electromagnetics. Master’s thesis, Chalmers University of Technology, 2006. (Cité page 105.)
- [20] Benoît Chaigne and Jean-Antoine Désidéri. Méthodes hiérarchiques pour la conception optimale de forme d’antenne à réflecteur. Research Report 6625, INRIA, Sophia-Antipolis, France, September 2008. (Cité page 79.)
- [21] Praveen Chandrashekarappa and Régis Duvigneau. Metamodel-assisted particle swarm optimization and application to aerodynamic shape optimization. Research Report RR-6397, INRIA, 2007. (Cité page 112.)
- [22] Praveen Chandrashekarappa and Régis Duvigneau. Radial Basis Functions and Kriging Metamodels for Aerodynamic Optimization. Research Report RR-6151, INRIA, 2007. (Cité page 112.)
- [23] A. Charnes and W.W. Cooper. *Management models and industrial applications of linear programming*. J. Wiley, New York, 1961. (Cité page 139.)
- [24] Y.-L. Chen and C.-C. Liu. Multiobjective VAR planning using the goal-attainment method. *Generation, Transmission and Distribution, IEEE Proceedings*, 141(3) :227–232, May 1994. (Cité page 139.)
- [25] M. Clerc and J. Kennedy. The particle swarm - explosion, stability, and convergence in a multidimensional complex space. *Evolutionary Computation, IEEE Transactions on*, 6(1) :58–73, 2002. (Cité page 45.)
- [26] Paul F. Combes. *Micro-ondes*, volume 2. Dunod, 1997. (Cité pages 7, 12 et 19.)
- [27] Carl De Boor. *A practical guide to splines*. Springer-Verlag, New York, 1978. (Cité page 31.)
- [28] James W. Demmel. *Applied numerical linear algebra*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 1997. (Cité pages 63, 94 et 162.)
- [29] Jean-Antoine Désidéri. *Modèles discrets et schémas itératifs. Application aux algorithmes multigrilles et multidomaines*. Hermès, 1998. (Cité pages 59, 63, 67, 71, 75 et 188.)
- [30] Jean-Antoine Désidéri. Two-level ideal algorithm for parametric shape optimization. pages 65–85. Institute of Numerical Mathematics, Russian Academy of Sciences, 2006. Proc. of two International Conferences : Moscow, Institute of Numerical Mathematics, Russian Academy of Sciences, Sept. 16-17, 2006 and Houston, 2006. (Cité pages 70 et 79.)
- [31] Jean-Antoine Désidéri. Split of territories in concurrent optimization. Research Report 6108, INRIA, Sophia-Antipolis, France, February 2007. (Cité page 162.)

- [32] Jean-Antoine Désidéri. Multiple-Gradient Descent Algorithm (MGDA). Technical report, INRIA, 2009. (Cité page 180.)
- [33] Thomas Dreyer, Bernd Maar, and Volker H. Schulz. Multigrid optimization in applications. *Journal of Computational and Applied Mathematics*, 120(1-2) :67–84, 2000. (Cité page 59.)
- [34] Pierre Dubois. *Optimisation de structures rayonnantes métalliques 3D par déformation de surfaces iso-niveaux en régime harmonique*. PhD thesis, École des Mines de Paris, Décembre 2005. (Cité pages 27, 30 et 44.)
- [35] Régis Duvigneau. Aerodynamic Shape Optimization with Uncertain Operating Conditions using Metamodels. Research Report RR-6143, INRIA, 2007. (Cité page 112.)
- [36] Régis Duvigneau. Robust design in aerodynamics using metamodels. 2007. EUCCO 2007, European Congress on Computational Optimization, Montpellier, France, April 2007. (Cité page 112.)
- [37] Régis Duvigneau, Benoît Chaigne, and Jean-Antoine Désidéri. Multi-level parameterization for shape optimization in aerodynamics and electromagnetics using a particle swarm optimization algorithm. Research Report 6003, INRIA, Sophia-Antipolis, France, October 2006. (Cité page 112.)
- [38] Régis Duvigneau and Praveen Chandrashekarappa. Meta-modeling for robust design and multi-level optimization. 2007. 42e Congrès d’Aérodynamique Appliquée AAAF, Sophia-Antipolis, France, March 2007. (Cité page 112.)
- [39] Gerald Farin. *Curves and Surfaces for CAGD : a practical guide*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, fifth edition edition, 2002. (Cité pages 31, 89 et 200.)
- [40] R.P. Fedorenko. A relaxation method for solving elliptic difference equations. *Russian Journal of Computational and Applied Mathematics*, 1 :1092–1096, 1961. (Cité page 59.)
- [41] R.P. Fedorenko. The speed of convergence of one iterative process. *USSR Computational Mathematics and Mathematical Physics*, 4(3) :227–235, 1964. (Cité page 59.)
- [42] Ralph Ferrayé, Jean-Yves Dauvignac, and Christian Pichot. An inverse scattering method based on contour deformations by means of a level set method using frequency hopping technique. *Antennas and Propagation, IEEE Transactions on*, 51(5) :1100–1113, May 2003. (Cité page 30.)
- [43] Ralph Ferrayé, Pierre Dubois, Ioannis Aliferis, Jean-Yves Dauvignac, Christian Pichot, Claude Dedebean, and Jean-Paul Zolésio. Shape-gradient optimization applied to the reconstruction of 2-d and 3-d metallic objects. In *Proceedings of the IEEE Antennas and Propagation Society International Symposium (AP-S 2009)*, Charleston, SC, USA, June 1–5 2009. (Cité page 30.)
- [44] Carlos M. Fonseca and Peter J. Fleming. Multiobjective Optimization and Multiple Constraint Handling with Evolutionary Algorithms—Part I : A Unified Formulation. *IEEE Transactions on Systems, Man, and Cybernetics, Part A : Systems and Humans*, 28(1) :26–37, 1998. (Cité page 140.)
- [45] Philipp E. Gill, Walter Murray, and Margaret H. Wright. *Practical Optimization*. Academic Press, London ; New York, 1981. (Cité pages 28, 48 et 200.)
- [46] David E. Goldberg. *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1989. (Cité page 44.)
- [47] Abderrahmane Habbal. A topology nash game for tumoral antiangiogenesis. *Structural and Multidisciplinary Optimization*, 30(5) :404–412, November 2005. (Cité page 161.)
- [48] Abderrahmane Habbal, Joakim Petersson, and Mikael Thellner. Multidisciplinary topology optimization solved as a nash game. *International Journal for Numerical Methods in Engineering*, 61(7) :949–963, August 2004. Copyright © 2004 John Wiley & Sons, Ltd. (Cité page 161.)

- [49] Wolfgang Hackbusch. *Multigrid Methods and Applications*. Springer-Verlag, New York, U.S., 1985. (Cité page 59.)
- [50] Laurent Hascoët. Tapenade : a tool for automatic differentiation of programs. In *Proceedings of 4<sup>th</sup> European Congress on Computational Methods, ECCOMAS'2004, Jyvaskyla, Finland, 2004*. (Cité page 44.)
- [51] Jeffrey Horn, Nicholas Nafpliotis, and David E. Goldberg. A Niche Pareto Genetic Algorithm for Multiobjective Optimization. In *Proceedings of the First IEEE Conference on Evolutionary Computation, IEEE World Congress on Computational Intelligence*, volume 1, pages 82–87, Piscataway, New Jersey, June 1994. IEEE Service Center. (Cité page 140.)
- [52] John David Jackson. *Classical Electrodynamics*. John Wiley & Sons, 1975. (Cité page 7.)
- [53] Ioannis C. Karpolis and Kyriakos C. Giannakoglou. A multilevel approach to single- and multiobjective aerodynamic optimization. *Computer Methods in Applied Mechanics and Engineering*, 197(33-40) :2963–2975, June 2008. (Cité page 59.)
- [54] C.T. Kelley. *Iterative Methods for Optimization*. SIAM, Raleigh, North Carolina, 1999. (Cité pages 48 et 200.)
- [55] J. Kennedy and R. C. Eberhart. Particle swarm optimization. *IEEE International Conference on Neural Networks*, 4 :1942–1948, 1995. (Cité page 45.)
- [56] Robert Michael Lewis and Stephen G. Nash. A multigrid approach to the optimization of systems governed by differential equations. In *8th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization*, Long Beach, CA, US, 2000. AIAA. (Cité page 59.)
- [57] Robert Michael Lewis and Stephen G. Nash. Model problems for the multigrid optimization of systems governed by differential equations. *SIAM J. Sci. Comput.*, 26(6) :1811–1837, 2005. (Cité pages 59 et 63.)
- [58] Hélène Lièvre, Jean-Antoine Désidéri, and Abderrahmane Habbal. Optimisation numérique de profils d'aile par algorithmes génétiques et jeux de nash. Research Report RR-4275, INRIA Sophia-Antipolis (SINUS), 2001. (Cité page 161.)
- [59] Antonio Luchetta, Stefano Manetti, Luca Pellegrini, Giuseppe Pelosi, and Stefano Selleri. Design of waveguide microwave filters by means of artificial neural networks. *International Journal of RF and Microwave Computer-Aided Engineering*, 16(6) :554–560, 2006. Copyright © 2006 Wiley Periodicals, Inc. (Cité page 182.)
- [60] A. C. Ludwig. The definition of cross polarization. *IEEE Transactions on Antennas and Propagation*, vol. AP-21 :116–119, January 1973. (Cité page 141.)
- [61] Nathalie Marco and François Beux. Multilevel optimization : Application to one-shot shape optimum design. Research Report 2068, INRIA, Sophia-Antipolis, France, 1993. (Cité page 59.)
- [62] Nathalie Marco and Alain Dervieux. Multilevel parametrization for aerodynamical optimization of 3d shapes. *Finite Elements in Analysis and Design*, 26(4) :259–277, 1997. (Cité page 59.)
- [63] Massimiliano Martinelli and François Beux. Multi-level gradient-based methods and parametrization in aerodynamic shape design. *Revue Européenne de Mécanique Numérique*, 17(1-2) :169–197, 2008. (Cité page 59.)
- [64] Mohamed Masmoudi. A synthetic presentation of shape and topological optimization. In *Conference on Inverse Problems, Control and Shape Optimization*. PICO, 1998. (Cité page 30.)

- [65] Mohamed Masmoudi, Julien Pommier, and Bessem Samet. Application of the topological asymptotic expansion to inverse scattering problems. In Yuri Kuznetsov, Pekka Neittaanmäki, and Olivier Pironneau, editors, *Numerical Methods for Scientific Computing, Variational Problems and Applications*. CIMNE Publication, 2003. (Cité page 30.)
- [66] Mohamed Masmoudi, Julien Pommier, and Bessem Samet. The topological asymptotic expansion for the maxwell equations and some applications. *Inverse Problems*, 21(2) :547–564, April 2005. (Cité page 30.)
- [67] Kaisa Miettinen. *Nonlinear Multiobjective Optimization*, volume 12 of *International Series in Operations Research and Management Science*. Kluwer Academic Publishers, Dordrecht, 1999. (Cité page 137.)
- [68] Frédérique Millo. *Conception optimale de structures rayonnantes*. PhD thesis, Université de Nice, 1991. (Cité pages 30 et 44.)
- [69] John Nash. Non-cooperative games. *The Annals of Mathematics*, 54(2) :286–295, 1951. (Cité page 161.)
- [70] Stephen G. Nash. A multigrid approach to discretized optimization problems. *Optimization Methods and Software*, 14 :99–116, 2000. (Cité pages 59 et 158.)
- [71] Stephen G. Nash and Jorge Nocedal. A numerical study of the limited memory BFGS method and the truncated-newton method for large scale optimization. *SIAM Journal on Optimization*, 1(3) :358–372, August 1991. (Cité page 132.)
- [72] J. A. Nelder and R. Mead. A simplex method for function minimization. *The Computer Journal*, 7(4) :308–313, January 1965. (Cité page 44.)
- [73] Jorge Nocedal. Updating quasi-newton matrices with limited storage. *Mathematics of Computation*, 35(151) :773–782, 1980. (Cité page 132.)
- [74] Jorge Nocedal and Stephen J. Wright. *Numerical Optimization*. Springer, 1999. (Cité page 28.)
- [75] Sophocles J. Orfanidis. Electromagnetic waves and antennas. 2008. (Cité pages 7, 11, 12 et 19.)
- [76] Stanley Osher and James A. Sethian. Fronts propagating with curvature dependent speed : Algorithms based on hamilton-jacobi formulations. *Journal of Computational Physics*, 79 :12–49, 1988. (Cité page 30.)
- [77] K. E. Parsopoulos and M. N. Vrahatis. UPSO : A Unified Particle Swarm Optimization scheme. In *Computational Methods in Sciences and Engineering*, volume 1 of *Lecture Series on Computer and Computational Sciences*, pages 868–873, Zeist, The Netherlands, 2004. VSP International Science Publishers. (Cité pages 46 et 200.)
- [78] K. E. Parsopoulos and M. N. Vrahatis. Parameter selection and adaptation in unified particle swarm optimization. *Mathematical and Computer Modelling*, 46(1-2) :198–213, 2007. (Cité page 46.)
- [79] Christian Pichot, Jean-Yves Dauvignac, Ioannis Aliferis, Emmanuel Le Brusq, Ralph Ferrayé, and Vincent Chatelee. Recent nonlinear inversion methods and measurement system for microwave imaging. In *Imaging Systems and Techniques, 2004. (IST). 2004 IEEE International Workshop on*, pages 95–99, May 2004. (Cité page 30.)
- [80] Les Piegl and Wayne Tiller. *The NURBS book*. Springer-Verlag, London, UK, 1995. (Cité page 31.)
- [81] J. W. Ruge and K. Stuben. *Algebraic Multigrid*, chapter 4, pages 73–130. SIAM, Philadelphia, Pennsylvania, 1987. (Cité page 59.)



- [82] Bessem Samet, Samuel Amstutz, and Mohamed Masmoudi. The topological asymptotic for the helmholtz equation. *SIAM Journal on Control and Optimization*, 42(5) :1523–1544, 2003. (Cité page 30.)
- [83] J. David Schaffer. *Multiple Objective Optimization with Vector Evaluated Genetic Algorithms*. PhD thesis, Vanderbilt University, 1984. (Cité page 140.)
- [84] Volker H. Schulz. Solving discretized optimization problems by partially reduced sqp methods. *Computing and Visualization in Science*, 1(2) :83–86, 1998. (Cité page 59.)
- [85] Thomas W. Sederberg, David L. Cardon, G. Thomas Finnigan, Nicholas S. North, Jianmin Zheng, and Tom Lyche. T-spline simplification and local refinement. *ACM Trans. Graph.*, 23(3) :276–283, 2004. (Cité page 31.)
- [86] Thomas W. Sederberg and Scott R. Parry. Free-Form Deformation of solid geometric models. *SIGGRAPH Computer Graphics*, 20(4) :151–160, 1986. (Cité pages 35 et 200.)
- [87] Thomas W. Sederberg, Jianmin Zheng, Almaz Bakenov, and Ahmad Nasri. T-splines and t-nurccs. In *SIGGRAPH '03 : ACM SIGGRAPH 2003 Papers*, pages 477–484, New York, NY, USA, 2003. ACM. (Cité page 31.)
- [88] Stefano Selleri, Stefano Manetti, and Giuseppe Pelosi. Neural network applications in microwave device design. *International Journal of RF and Microwave Computer-Aided Engineering*, 12(1) :90–97, 2002. Copyright © 2002 John Wiley & Sons, Inc. (Cité page 182.)
- [89] J. Sokolowski and A. Zochowski. On the topological derivative in shape optimization. *SIAM J. Control Optim.*, 37(4) :1251–1272, 1999. (Cité page 30.)
- [90] Jan Sokolowski and Jean-Paul Zolesio. *Introduction to Shape Optimization*. Springer-Verlag, 1992. (Cité pages 30, 37 et 44.)
- [91] W. Spendley, G.R. Hext, and F.R. Himsworth. Sequential application of simplex designs in optimization and evolutionary operation. *Technometrics*, 4 :441–461, 1962. (Cité page 44.)
- [92] N. Srinivas and Kalyanmoy Deb. Multiobjective Optimization Using Nondominated Sorting in Genetic Algorithms. *Evolutionary Computation*, 2(3) :221–248, Fall 1994. (Cité page 140.)
- [93] Julius Adams Stratton. *Electromagnetic Theory*. McGraw-Hill, New York, 1st ed. edition, 1941. (Cité page 7.)
- [94] T. Suttorp, N. Hansen, and C. Igel. Efficient covariance matrix update for variable metric evolution strategies. *Machine Learning*, 72(2) :167–197, 2009. (Cité page 182.)
- [95] Zhili Tang, Jean-Antoine Désidéri, and Jacques Périaux. Multicriterion aerodynamic shape design optimization and inverse problems using control theory and nash games. *Journal of Optimization Theory and Applications*, 135(3) :599–622, December 2007. (Cité page 161.)
- [96] Virginia Joanne Torczon. *Multidirectional Search : a direct search algorithm for parallel machines*. PhD thesis, Rice University, Houston, TX, USA, 1989. Chairman-Dennis, Jr., John E. (Cité page 44.)
- [97] David A. Van Veldhuizen and Gary B. Lamont. Multiobjective evolutionary algorithms : Analyzing the state-of-the-art. *Evolutionary Computation*, 8(2) :125–147, 2000. (Cité page 139.)
- [98] M. Vázquez, A. Dervieux, and B. Koobus. Multilevel optimization of a supersonic aircraft. *Finite Elements in Analysis and Design*, 40(15) :2101–2124, 2004. (Cité page 59.)
- [99] John Von Neumann and Oskar Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, 1944. (Cité page 161.)

- [100] Jiang Feng Wang and Jacques Periaux. Multi-Point Optimization using GAs and Nash/Stackelberg Games for High Lift Multi-airfoil Design in Aerodynamics. In *Proceedings of the Congress on Evolutionary Computation 2001 (CEC'2001)*, volume 1, pages 552–559, Piscataway, New Jersey, May 2001. IEEE Service Center. (Cité page 161.)
- [101] Peter Wesseling. *An Introduction to Multigrid Methods*. John Wiley & Sons, 1991. (Cité pages 59, 71 et 75.)
- [102] Ji Chao Zhao, Jean-Antoine Désidéri, and Badr Abou el Majd. Two-level correction algorithms for model problems. Research Report 6246, INRIA, Sophia-Antipolis, France, July 2007. (Cité page 79.)
- [103] Eckart Zitzler and Lothar Thiele. Multiobjective Evolutionary Algorithms : A Comparative Case Study and the Strength Pareto Approach. *IEEE Transactions on Evolutionary Computation*, 3(4) :257–271, November 1999. (Cité page 140.)

**Titre** Méthodes hiérarchiques pour l'optimisation géométrique de structures rayonnantes

**Résumé** Une antenne à réflecteur est un dispositif encore largement utilisé pour la communication satellite. La durée de vie d'un tel dispositif est étroitement liée à la fatigue due à la consommation d'énergie pour émettre le signal. Un des enjeux de la conception optimale d'une antenne revient donc à produire des systèmes dont le rendement est le meilleur possible par rapport à une tâche donnée. La particularité d'une antenne à réflecteur se traduit par la présence de surfaces rayonnantes dont la géométrie constitue le paramètre principal pour assumer cette tâche. Sur la base de la simulation de la propagation d'une onde électromagnétique en espace libre et en régime harmonique, on est capable de développer des méthodes d'optimisation numérique de la forme de surfaces rayonnantes. On cherche à minimiser un critère qui traduit en terme mathématique la tâche à effectuer d'un point de vue énergétique. Cependant, les méthodes utilisées sont souvent soumises à des difficultés liées au fait que ces problèmes sont mal posés et numériquement raides. Le contrôle étant géométrique, on a examiné dans cette thèse les contributions potentielles de représentations hiérarchiques afin d'étendre les performances d'algorithmes classiques d'optimisation. Ces extensions empruntent leurs fondements aux méthodes multigrilles pour la résolution d'EDP. Un exemple théorique d'optimisation de forme permet d'asseoir les stratégies appliquées à l'optimisation d'antennes. Puis des expériences numériques d'optimisation montrent que les algorithmes de bases sont améliorés en terme de robustesse comme en terme de vitesse de convergence.

**Mots-clés** électromagnétisme, antenne à réflecteur, optimisation de forme paramétrique, méthodes de descente, Optimisation par Essaim de Particules, méthodes multiniveaux, optimisation multipoint

**Title** Hierarchical methods for optimal shape design of radiating structures

**Abstract** A reflector antenna is a device that is widely used for satellite communication. The life length of such a device depends highly on the fatigue due to the energy consumption for the signal emission. Thus, one of the goals of the optimal design of an antenna is to improve the productivity of systems designed for a given task, for a fixed input power. A reflector antenna is characterized by radiating surfaces whose geometry is the main parameter that can be controlled to fulfill the task. Based on the time-harmonic wave propagation simulation in free space, numerical procedures for the optimal design of the shape of radiating structures are examined. Namely, we aim at minimizing a criterion that represents the task in mathematical terms. In this framework, classical optimization methods are often submitted to challenging difficulties related to the fact that the problems are ill-posed because multimodal and numerically stiff. Since the control is the geometry of the reflectors, we have investigated in this thesis the potential enhancements of basic algorithms using hierarchical parametric representations. The theoretical foundations of the proposed algorithms rely on the Multigrid methods for solving PDE. A theoretical example for shape optimization is considered in order to derive different multilevel strategies. These strategies are then applied to real-case problems for the optimal design of reflector antennas. Numerical experiences show that basic algorithms are effectively enhanced in terms of robustness and convergence rate.

**Keywords** electromagnetics, reflector antenna, parametric shape optimization, descent methods, Particle Swarm Optimization, multilevel methods, multipoint optimization