

Simulation électromagnétique utilisant une méthode modale de décomposition en ondelettes

Ana Maria Armeanu

▶ **To cite this version:**

Ana Maria Armeanu. Simulation électromagnétique utilisant une méthode modale de décomposition en ondelettes. Micro et nanotechnologies/Microélectronique. Université de Grenoble, 2011. Français. NNT : . tel-00690418

HAL Id: tel-00690418

https://theses.hal.science/tel-00690418

Submitted on 23 Apr 2012

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 GRENOBLE

ÉCOLE DOCTORALE
MICRO ET NANO ELECTRONIQUE

Mémoire de thèse

présenté par

Ana Maria Armeanu

pour obtenir le grade de

DOCTEUR D'UNIVERSITÉ
Spécialité : MICRO ET NANO ELECTRONIQUE

Simulation électromagnétique utilisant une méthode modale de decomposition en ondelettes

Rapporteurs :	Mme. Christine Letrou
	M. Brahim Guizal
Examinateurs :	M. Patrick Schiavone - Directeur de thèse
	M. Kofi Edee - Co- directeur de thèse
	M. Gérard Granet
	M. Vincent Farys
	M. Olivier Toublan

REMERCIEMENTS

Je tiens tout d'abord à exprimer ma profonde reconnaissance à mon directeur de thèse Patrick Schiavone, qui a consacré à l'encadrement de ma thèse un temps considérable et qui m'a accordé une grande disponibilité - malgré son emploi du temps chargé et son éloignement géographique. Je le remercie pour la bienveillance de père de famille qu'il sait manifester à ses doctorants, pour les longues et fructueuses discussions par téléphone, ou de vive voix.

Je remercie également Kofi Edee [1], qui en tant que co-directeur de thèse, s'est chargé de la proximité géographique. J'ai ainsi pu profiter de sa grande compétence scientifique pour faire mes premiers pas dans le domaine des ondelettes. Il a su éveiller en moi la curiosité et l'envie d'approfondir et d'étendre toujours plus mes connaissances. Je lui en suis très reconnaissante.

J'adresse toute ma gratitude à Gérard Granet pour sa rigueur scientifique et son sens critique qui ont été très profitables à mes travaux de recherche. Je le remercie aussi d'avoir manifesté un grand intérêt pour l'ensemble de ma thèse et d'avoir pris la peine de la relire en détail.

Mes remerciements ne seraient pas complets s'ils ne s'adressaient pas aussi à tous les membres du laboratoire LTM que j'ai côtoyés pendant ces trois années de thèse. Une pensée particulière pour Mame Kouna Top qui a un coeur d'or, Ayse Akbalik qui m'a "supportée" comme collègue de bureau au LTM, Nassima Zegaoui, toujours gaie et active, Tzvety, ardente défenseuse des droits de la femme, Martin Kogelschatz, grand spécialiste de la bible. Je salue également Stéphanie Berger, Sylvaine Cetra et Jumana Boussey, leur sympathie et leur professionnalisme m'ont été très précieux.

J'ai eu la chance de collaborer pendant ma thèse avec le Lasmea, un laboratoire de recherche très vivant et accueillant, où règne une ambiance conviviale et dynamique. Parmi les artisans de ce bel équilibre je dois d'abord citer Antoine Moreau, Rémi Pollès, Julien Arlandis, Jessica Benedicto. J'ai vraiment apprécié nos discussions sur une vaste palette des sujets (en partant de sujets scientifiques et en arrivant à des sujets comme Picasso dans l'art contemporain). Je salue aussi les autres collègues de l'équipe Matelec et Gravir.

Je remercie mes parents qui ont eu confiance en moi et ont encouragé l'aventureuse décision de faire une thèse en France. Je remercie Titi, mon frère qui a partagé les instants de doutes et de joies liés à mon travail de thèse. Un grand salut amical à la famille Venet-Vrinceanu. Leur amitié et leur soutien m'ont allégé le mal du pays. Je remercie également tous mes amis que je ne peux pas mentionner ici.

1. dit aussi " Le Dieu des Ondelettes ".

Table des matières

Notations et Abréviations **5**

Introduction générale **7**

1 Présentation de la méthode numérique **11**

1.1 Introduction 12

1.2 Méthodes de résolution numérique des équations de Maxwell 13

1.2.1 Méthode intégrale 13

1.2.2 Méthode de Rayleigh 14

1.2.3 Méthode différentielle 14

1.2.4 Méthodes modales 15

1.3 Méthode Modale : principe général 18

1.4 Les équations de Maxwell 20

1.4.1 Les équations de propagation 21

1.4.2 Milieux homogènes 25

1.4.3 Milieux inhomogènes 25

1.5 Conditions aux limites et calcul des efficacités 26

1.6 Bilan 29

2 Méthode Modale avec des fonctions Splines 31

2.1 Introduction 32

2.2 Description de la méthode 33

2.2.1 Méthode des résidus pondérés 33

2.2.2 Application aux équations de Maxwell 35

2.2.3 Description des fonctions B_n et T_q 37

2.2.4 Périodicité et pseudo périodicité 42

2.3 Résultats numériques 44

2.3.1 Influence des fonctions test 45

2.3.2 Influence des fonctions de base 51

2.3.3 Influence de la grille de discrétisation 52

2.3.4 Comparaison avec d'autres méthodes 55

2.3.4.1 Réseau diélectrique 55

2.3.4.2 Réseaux métalliques 57

2.3.4.3 Réseaux avec des matériaux à indice négatif 59

2.3.4.4 Réseaux résine sur silicium 60

2.3.5 Résolution spatiale adaptative 61

2.4 Bilan 66

3 Méthode Modale avec décomposition en ondelettes 69

3.1 Introduction 70

3.2 Eléments de théorie sur les ondelettes 71

3.2.1 Transformée en ondelettes continue 71

3.2.2 Transformée en ondelettes discrète 73

3.2.3 Analyse multirésolution 73

3.2.4 Quelques propriétés intéressantes des ondelettes 75

3.2.5 Ondelettes orthogonales 78

3.2.6 Ondelettes biorthogonales 81

3.2.7 Transformée en ondelettes rapide 84

3.3 Applications aux équations de Maxwell 87

3.3.1 La forme des matrices au premier niveau de résolution 89

3.3.2 La forme des matrices au deuxième niveau de résolution 91

3.3.3 Décimation des détails : le zoom 92

3.4 Résultats 94

3.4.1 L'influence de la longueur des filtres 94

3.4.2 Ajouter un niveau de détail 97

3.4.3 Comparaison avec la littérature 103

3.5 Bilan 104

Conclusion et perspectives **109**

Bibliographie **112**

Annexes **121**

A Formalisme des matrices S **123**

B Calcul des efficacités **127**

C Publications et conferences **131**

Notations et Abréviations

E	Champ électrique
H	Champ magnétique
D	Déplacement électrique
B	Excitation magnétique
H	Champ magnétique
TE	Polarisation transverse électrique **E**=(0,0,E_z)
TM	Polarisation transverse magnétique **H**=(0,0,H_z)
d	période du réseau
f	facteur de remplssage
ε	permitivité électrique relative
ε_0	permitivité électrique du vide
ν	indice de refraction
θ	angle d'incidence
μ	perméabilité magnétique
μ_0	perméabilité magnétique du vide
ω	pulsation
h	hauteur du sillon
λ	longueur d'onde
$\mathfrak{Re}(x)$	Partie réelle du compexe x
$\mathfrak{Im}(x)$	Partie imaginaire du compexe x
$x*$	Complexe conjugué de x
∂_x	Dérivée par raport à x

Introduction générale

Cette thèse s'est déroulée dans le contexte d'une collaboration entre le Laboratoire des Technologies de la Microélectronique, Grenoble (LTM) le Laboratoire des Sciences et Matériaux pour l'Electronique et d'Automatique, Clermont - Ferrand (LASMEA) et l'industriel STMicroelectronics Crolles 2. Ce manuscrit rend compte de cette collaboration où les mathématiques appliquées en électromagnétisme ont été mises au service de la simulation numérique en vue d'applications en lithographie et en métrologie pour l'industrie de la microélectronique.

Cadre général

De nos jours, on utilise quotidiennement des produits de la microélectronique. On les retrouve dans les ordinateurs, les télécommunications, l'industrie automobile, l'armement, l'aérospatial et même dans des produits électroménagers. Les applications de la microélectronique sont diverses et ont tendance à se diversifier encore plus [2]. Dans ces conditions, la recherche et le développement en microélectronique est plus indispensable que jamais. Des investissements colossaux sont faits afin de développer des procédés plus innovants et plus rentables du point de vue économique. Chaque année le marché est renouvelé, avec l'apparition d'appareils de plus en plus petits, de plus un plus performants, consommant peu d'énergie et à coût de fabrication réduit. Pour ces raisons, le pourcentage de chiffre d'affaire directement réinvesti dans le domaine R&D est un des plus élevés de toute l'industrie. On peut mentionner ici l'exemple de STMicroelectronics, un des acteurs principaux de la recherche en microélectronique grenobloise. ST a consacré environ 28% de son chiffre d'affaires pour la R&D.

Les progrès de l'industrie microélectronique et en particulier la miniaturisation des com-

2. Par exemple, en 2000, l'électronique représentait 20% du coût total d'une voiture, aujourd'hui, en 2010, la part s'élève à 40% [1] et dans les prochains années ce pourcentage va encore croître grâce aux systèmes guidés.

posants ont rendu nécessaire l'amélioration des méthodes de mesure des structures fabriquées. Des appareils comme l'AFM, le microscope électronique à balayage, l'ellipsométre sont utilisés de nos jours quotidiennement dans la mesure des structures. Une innovation majeure est venue avec la scattérométrie qui présente l'avantage considérable d'être une technique non destructive. Les dimensions des objets sont obtenues à partir du traitement du problème inverse de la diffraction. Dans le contexte de la microélectronique, on compare les spectres mesurés avec un instrument optique à des spectres calculés pour un grand nombre de structures et stockés dans une bibliothèque. Comme les objets fabriqués sont extrêmement complexes, les mesures portent pour l'instant uniquement sur des motifs tests canoniques qui témoignent du processus de fabrication. Ces motifs test sont en général des réseaux de lignes périodiques. Ce genre de mesures a commencé à être utilisé au début des années 2000 et l'utilisation des ces techniques a débuté, au sein de l'Alliance Crolles 2 en 2001.

Les besoins en calcul numérique se font d'autant plus ressentir que le coût de fabrication d'un objet structuré à l'échelle de la centaine de nanomètres reste très élevé. Ainsi, pour concevoir un composant, mieux vaut avoir recours au calcul numérique plutôt qu'à la fabrication de multiples échantillons. Pour autant, le calcul numérique ne constitue pas une solution de facilité. Il s'agit en effet de calculer la diffraction par des structures géométriques complexes composées de matériaux présentant de forts contrastes d'indice, donc loin du régime perturbatif, et faisant coexister des éléments de grande et de petites tailles par rapport à la longueur d'onde. L'algorithme de calcul numérique doit être capable de décrire toutes ces structures avec précision et rapidité.

L'analyse des réseaux a dû considérablement évoluer dans les dernières années lorsque la longueur d'onde λ est devenue de l'ordre de grandeur de la période du réseau ou de sa hauteur. En effet, dans ce cas les théories de l'optique physique ne permettent pas de décrire la diffraction par des réseaux car elle ne tient pas compte de la nature vectorielle du problème. Les équations fondamentales de l'étude des réseaux sont les équations de Maxwell. Elles permettent de décrire la diffraction de la lumière illuminant le réseau. Étant donné que, en général, dans les milieux discontinus ces équations n'ont pas de solution analytique, des méthodes numériques sont requises afin d'approcher le champ électromagnétique et la quantité d'énergie portée par chaque ordre de diffraction.

Objectifs de la thèse

Dans les dernières décennies, un grand nombre de méthodes numériques de résolution des équations de Maxwell ont été mises en œuvre : méthodes intégrales [2, 3], développement sur les ondes planes [4, 5], méthodes différentielles [6, 7, 8]. Parmi ces dernières, des méthodes plus spécifiques comme les méthodes modales qui se sont avérées être bien adaptées pour l'analyse des réseaux. Elles ont été développées en parallèle par plusieurs chercheurs [8, 9, 10, 11]. Dans le contexte industriel de la microélectronique, la RCWA [9, 12, 13] est la méthode d'analyse la plus utilisée grâce à sa facilité de mise en œuvre. Malheureusement, elle repose sur l'emploi des bases de Fourier dont les défauts sont bien connus. L'analyse des structures à fort contraste d'indice ou à faible facteur de forme soulève des difficultés ou des instabilités à cause notamment du phénomène de Gibbs. Le calcul rapide et fiable pour des structures tridimensionnelles reste toujours un défi.

Pour remédier aux insuffisances de l'analyse de Fourier, une famille de fonctions appelées ondelettes a été inventée [14, 15, 16, 17, 18] aux début des année 80. La transformée en ondelettes permet d'exprimer sans perte d'information un signal dans une base dans laquelle on obtient la contribution locale de chaque fréquence au signal. Cette analyse a révolutionné les performances numériques dans le domaine du traitement de signal. De plus, elle a été appliquée avec succès dans de nombreux autres domaines dans lesquels l'analyse de Fourier était habituellement utilisée. Les ondelettes ont ouvert une voie alternative, souvent mieux adaptée et elles constituent des outils importants de l'analyse moderne. Le but de cette thèse est d'exploiter et d'appliquer ces bases d' ondelettes, dans le contexte des méthodes modales, à la résolution du problème de la diffraction par des réseaux. Ces bases, grâce à l'analyse multirésolution peuvent décrire les variations brusques du champ électromagnétique sur plusieurs niveaux de détails. Elles amènent un aspect novateur dans la mise en œuvre numérique de la méthode modale. Elles arrivent à éviter des phénomènes non désirés, comme celui de Gibbs et à traiter des cas pour lesquels d'autres méthodes classiques sont en échec.

Organisation du mémoire

Ce mémoire se divise en 3 chapitres. Le premier, *Outils numériques*, de nature théorique, passe brièvement en revue les méthodes numériques les plus souvent utilisées pour résoudre le problème de la diffraction par des réseaux. Les différentes méthodes possèdent chacune

des avantages et des inconvénients, autant dans leur mise en œuvre, que dans leur description des phénomènes électromagnétiques. Nous détaillons plus particulièrement la méthode modale puisque c'est celle que nous avons retenue dans le cadre de notre travail.

Nous nous proposons par la suite, dans le deuxième chapitre, *Méthode Modale avec des fonctions Splines*, d'adapter la théorie sur la méthode modale pour utiliser les fonctions B-spline comme fonctions de base. Ces fonctions représentent le premier pas vers l'analyse multirésolution en utilisant les ondelettes splines ou plus généralement les ondelettes à support compact. Une compréhension approfondie du comportement des bases à support compact sur un seul niveau de détail est nécessaire avant de développer une analyse sur plusieurs échelles de détail. Nous présentons de façon détaillée l'applications des ces fonctions à la résolution des équations de Maxwell pour les deux cas de polarisation. On fait aussi un passage en revue du choix des fonctions B-splines qu'on utilisera par la suite (il s'agit des splines de degré 0, 1, 2 et 3). Le calcul analytique des produits scalaires est détaillé en tenant compte des propriétés de symétrie, de support compact des fonctions B-splines et de leur définition explicite. On détaille également les modifications à faire pour rendre la base des splines périodique et écrire la condition de pseudo-périodicité. Ce chapitre contient également des résultats numériques qui mettent en lumière l'influence des paramètres optogéométriques sur la vitesse de convergence des valeurs propres, des efficacités de diffraction et du champ à la surface du réseau. Nous comparons nos résultats avec ceux issus d'autres implémentations de la méthode modale. De plus, nous introduisons une résolution spatiale adaptative et nous présentons les améliorations apportées par cette technique.

Le troisième chapitre *Méthode Modale avec des fonctions Ondelettes* constitue la partie la plus originale du manuscrit. L'analyse multirésolution est utilisée pour résoudre le problème de la diffraction par un réseau dans le cadre des méthodes modales. Nous commençons par développer la théorie permettant la résolution de notre problème. Ensuite, nous décrivons, avec les bases d'ondelettes, comment les équations de Maxwell deviennent un système aux valeurs propres. Enfin, nous détaillons la façon mathématique de construire "le zoom" sur les zones de variations brusques du champ électromagnétique en utilisant plusieurs niveaux de détail. Pour finir, des résultats numériques illustrent ce chapitre.

Chapitre 1

Présentation de la méthode numérique

Sommaire

1.1 Introduction 12
1.2 Méthodes de résolution numérique des équations de Maxwell 13
1.2.1 Méthode intégrale 13
1.2.2 Méthode de Rayleigh 14
1.2.3 Méthode différentielle 14
1.2.4 Méthodes modales 15
1.3 Méthode Modale : principe général 18
1.4 Les équations de Maxwell 20
1.4.1 Les équations de propagation 21
1.4.2 Milieux homogènes 25
1.4.3 Milieux inhomogènes 25
1.5 Conditions aux limites et calcul des efficacités 26
1.6 Bilan 29

L'objectif de ce premier chapitre est consacré à une brève présentation des méthodes numériques de résolution des problèmes de diffraction par des réseaux les plus couramment employées. Puisque l'approche que nous avons développée fait partie de la famille des méthodes modales, on donne l'aperçu du calcul général de résolution des équations de Maxwell et des équations fondamentales de la théorie des réseaux, en utilisant ce type de méthode.

1.1 Introduction

La réduction des tailles de semi-conducteurs fait ressentir de plus en plus le besoin de techniques de mesure scatterométriques précises. Les masques de lithographie doivent prendre en compte les effets de bords dues au phénomène de diffractions de la lumière. D'un point de vu pratique, il est plus avantageux de concevoir des routines de calcul numérique qui décrivent le phénomène plutôt que fabriquer à l'aveugle de coûteux échantillons. Cependant le calcul numérique s'avère difficile à mettre en œuvre quand on doit décrire des structures géométriques complexes en se plaçant dans le régime résonant. Pour obtenir un modèle de simulation à la fois rapide et rigoureux, des hypothèses comme la périodicité des masques où le découpage en tranches des formes trop complexes sont faites afin d'alléger les calculs numériques. Le model théorique nous donne des informations avant le processus de fabrication. Dans la figure 1.1 on a représenté un motif typique qui peut par exemple réprésenter des lignes de résine obtenues par lithographie optique. Une onde électromagnétique incidente qui rencontre une telle structure,

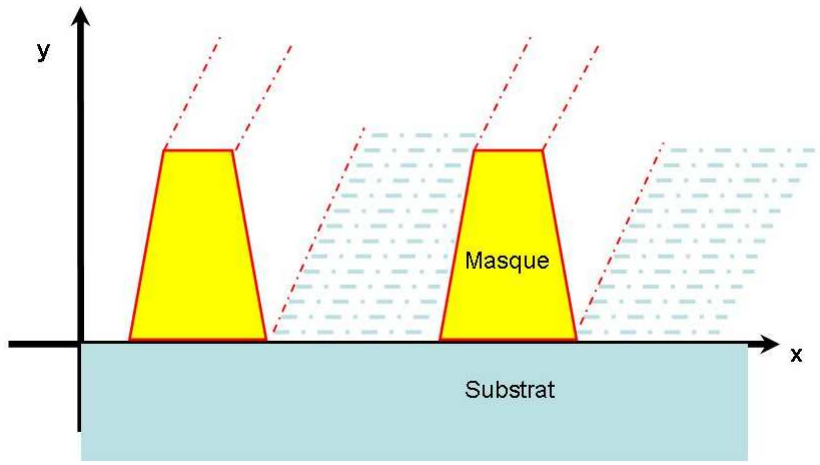


FIGURE 1.1 – *Réseau de diffraction quelconque*

voit son amplitude et sa phase modifiées. Cette modification se traduit par une répartition de l'intensité de l'onde incidente suivant des directions de diffraction. Pour simuler la répartition de ces intensités et décrire le champ électromagnétique la résolution des équations de Maxwell à l'aide de méthodes numériques rigoureuses est requise.

1.2 Méthodes de résolution numérique des équations de Maxwell

De nombreuses méthodes existent, mais de gros efforts restent à faire afin d'obtenir des codes pour des structures tridimensionnelles qui soient suffisamment robustes et rapides. Dans ce qui suit, on fait un bref passage en revue des méthodes les plus couramment utilisées.

1.2.1 Méthode intégrale

C'est une des plus anciennes méthodes rigoureuses élaborées [2, 3]. L'idée de base consiste à résoudre le problème de la diffraction par une simple intégration. A partir du moment où le champ et ses dérivées sont connus sur la surface de l'élément diffractant, le théorème de Green permet alors de connaître le champ dans tout l'espace [19] :

$$F(x,y) = -i\omega\mu_0 \int G(x-x^{'}, y-f(x^{'}))j_s(x')dl^{'}$$

où $G(x-x^{'}, y-f(x^{'}))$ représente la fonction de Green, $(x^{'}, f(x^{'}))$ est un point d'observation, j_s le courant de surface et $dl^{'}$ un élément de la frontière de la surface. Le principe physique sous-jacent est le suivant : un champ incident crée en un point x de la surface du réseau une densité surfacique de courant $j_s(x)$. La densité de courant créée en un point $x^{'}$ par ce champ incident sera alors perçue par le point x comme une source secondaire. La densité totale de courant en un point dépend donc à la fois du champ incident, mais aussi de la densité des points dans son environnement. Or le champ diffracté correspond au champ généré par l'ensemble de ces courants de surface. Le problème revient alors à déterminer une densité de courant élémentaire et à l'intégrer sur la surface totale du réseau.

Une des difficultés de cette méthode est que lorsque les sources secondaires doivent être rapprochées les une des autres, le couplage entre deux sources voisines peut croître indéfiniment créant ainsi une singularité numérique. On est obligé de localiser ces singularités analytiquement et de les extraire de l'intégration numérique (pour plus de détails le lecteur pourra consulter [20]). Un autre moyen de contourner ces singularités numériques est la méthode des sources fictives dans laquelle les sources ne sont plus disposées sur la surface du réseau mais dans un très proche voisinage.

Un des grands avantages de la méthode intégrale est que l'on recherche les sources à la surface même du réseau sans jamais le traverser. Il y a un prix à payer : des codes de calcul plus

complexes donc plus coûteux en temps.

1.2.2 Méthode de Rayleigh

Lord Rayleigh a établi la théorie électromagnétique permettant de calculer les efficacités de diffraction d'un réseau [4, 5] au début du siècle dernier. Cette théorie [21] a permis d'expliquer les anomalies décrites par Wood [22] en 1902, caractérisées par une variation notable des coefficients de réflexion et transmission lors d'une faible modification des paramètres caractéristiques de l'onde incidente. Le développement de Rayleigh fait apparaître une somme infinie d'ondes planes. Le champ électromagnétique est exprimé par une série d'ondes planes montantes et descendantes ; ces ondes pouvant être soit propagatives, soit évanescentes. Cette relation est valable dans les couches homogènes. La méthode de Rayleigh fait l'hypothèse que le développement du champ en ondes planes reste encore valable dans la région du réseau. Les conditions aux limites au niveau du réseau peuvent être exprimées par une relation simple liant les amplitudes des ordres diffractés, transmis et réfléchis. Le problème est ainsi ramené à la résolution d'un système d'équation linéaire. Une des limitations des cette approche est que l'hypothèse n'est acceptable que pour des réseaux de faible profondeur [1].

Des améliorations ont été proposées pour des cas particuliers, comme la méthode de Rayleigh-Fourier pour les réseaux sinusoïdaux où il devient possible de traiter des structures cinq fois plus profondes que la limite précédente. La formulation variationnelle de l'hypothèse de Rayleigh permet aussi une adaptation aux cas des réseaux triangulaires.

1.2.3 Méthode différentielle

La méthode différentielle classique [6, 7, 8] a été développée au début des années 60. Les équations de Maxwell peuvent s'exprimer, d'un point de vue mathématique, comme des équations aux dérivées partielles à coefficients non constants. En utilisant la périodicité de la structure on décompose le champ électromagnétique ainsi que la permittivité en série de base (le plus couramment utilisées étant les bases de Fourier) :

$$F(x,y) = \sum_m F_m(y) B_m(x) \qquad \varepsilon(x,y) = \sum_m \varepsilon_m(y) B_m(x)$$

1. de l'ordre de $\lambda/10$

Une fois projetée sur cette base, le système d'équations aux dérivées partielles devient un système d'équations différentielles ordinaires, qui peut être résolu par des algorithmes d'intégration numériquen classiques.

1.2.4 Méthodes modales

Les méthodes modales, aussi nommées méthodes spectrales [23], sont utilisées dans de nombreux problèmes physiques. Elles ont été appliquées avec succès aux problèmes de réseaux. C'est aussi la méthode sur laquelle nous allons travailler dans ce mémoire. Le calcul du champ électromagnétique est applicable à des structures de forme quelconque à condition que la structure soit coupée en tranches invariantes selon l'axe $0y$ (voir figure 1.2). Une fois le décou-

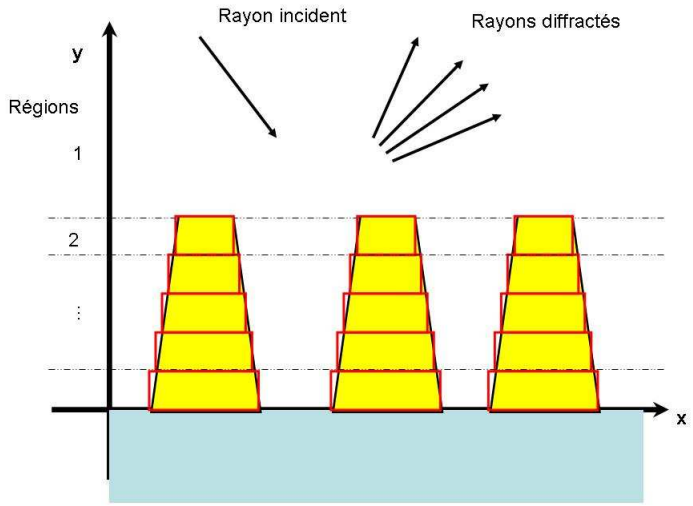


FIGURE 1.2 – *Principe de découpage d'une structure périodique de forme quelconque en rectangles y-invariants. Le contour de l'objet est représenté par les contours noirs. Le profile de l'objet est approximé par des rectangles rouges.*

page fait, on approxime les contours de l'objet dans chaque tranche par des rectangles. Cette approche est nécessaire pour assurer la condition d'invariance de la permittivité dans chaque tranche. Cela nous permet de faire l'intégration des équations de Maxwell dans une seule direction de l'espace. Le fait que pour ce type de profil, la permittivité ne dépend plus de la direction verticale dans la région du réseau ramène le problème à un système d'équations aux valeurs propres. Les méthodes modales, du fait de l'intégration analytique des équations de Maxwell

dans une direction de l'espace, possèdent une signification physique forte. Elles reposent sur le calcul des échanges d'énergie aux interfaces entre tous les modes, propagatifs et évanescents de la structure. Dans de nombreuses situations, peu de modes (souvent les quelques premiers modes propagatifs ou faiblement évanescents) participent réellement au transport de l'énergie. En identifiant ces modes, le calcul numérique permet une compréhension fine de phénomènes physiques observés expérimentalement ou prédits numériquement.

Il y a de nombreuses variantes de méthodes modales. Parmi eux on rappelle **La Méthode Modale Exacte** développée par Botten et al. [8, 24, 25]. Elle nous a servi de référence pour le calcul des erreurs pour certaines valeurs propres spécifiques des couches inhomogènes. Cette méthode a le grand avantage de fournir une solution "exacte". L'idée de base est de partir de l'équation de Helmoltz [9] et d'obtenir les valeurs propres comme solution d'une équation transcendante. Le champ global est obtenu comme combinaison linéaire de chacun des ces modes. Des problèmes se posent dans les cas de milieux avec des permittivités complexes où il y a beaucoup de modes évanescents qui doivent être pris en compte. Lorsque l'on cherche une valeur propre dans le plan complexe, on a besoin de méthodes mathématiques sophistiquées [26] pour pouvoir dépasser les difficultés d'ordre numérique. Même avec ces méthodes, la détermination des solutions propres est difficile et il y a toujours un risque de ne pas prendre en compte toutes les valeurs et vecteurs propres évanescents.

La méthode la plus utilisée pour résoudre le problème des réseaux est sans doute **la Méthode Modale de Fourier**. Introduite pour la première fois par Knop [9] en 1978 et popularisée ensuite par Mohoram [12, 13] sous le nom de Rigorous Coupled Wave Analysis (RCWA), on la retrouve aujourd'hui dans la littérature dans nombreux articles de Li [27, 28, 29, 30] ou de Granet [31] sous l'appellation Modal Method by Fourier Expansion (MMFE). L'idée de base est de développer toutes les composantes des champs en séries de Fourier et de résoudre le système aux valeurs propres associé aux opérateurs différentiels obtenus. Il faut cependant noter que cette méthode très efficace pour les réseaux diélectriques, peut présenter des divergences ou des convergences lentes dans le cas des réseaux métalliques. Des améliorations fondées sur des considérations empiriques ont été faites dans le cas de la polarisation TM par Granet et al. [32] ou en se basant sur des critères physiques par Lalanne et al. [33]. Finalement, Li [34] a donné des bases mathématiques solides à ces approches en élaborant la théorie connue sous le nom de règles de factorisation. La vitesse de convergence à été encore améliorée par l'introduction de systèmes des coordonnées qui resserrent les points autour des discontinuités [35, 36]. Le cas de l'incidence conique [37, 38] et dernièrement même les motifs apériodiques [39] ont pu être étudiés. Un des inconvénients de la méthode de Fourier est qu'elle est mise en difficulté lorsqu'il s'agit de réseaux avec des motifs très localisés et/ou des forts contrastes d'indices.

Dans le cadre des méthodes modales, on rappelle aussi **La Méthode C** introduite par Chandezon et al. [10, 40] utilisée pour traiter des profil plus généraux. Grâce à un changement de coordonnées approprié, il est possible de "transformer" le profil réel en un plan. Les équations de Maxwell sont résolues séparément de part et d'autre du "nouveau profil" puis raccordées sur l'interface. L'idée de base paraît simple, mais elle entraîne une complication lors de la résolution des équations de Maxwell : on se retrouve confronté à un système d'équations aux dérivées partielles à coefficients non constants, même dans le cas du milieu homogène. Depuis son apparition, en 1980, jusqu'à nos jours, des nombreux améliorations ont été faites et la méthode C s'est avérée utile dans plusieurs applications [41, 42]. Li [43] et Plumey et al. [44] ont étendu la méthode C à la diffraction conique. Granet et al. [45], Li et al. [46] et Preist et al. [47] ont permis de traiter des couches à faces non-parallèles. Puis la méthode C a été étendue aux réseaux bi-périodiques [48, 49] et aux cristaux anisotropes [50]. Li [43] améliore encore la stabilité numérique de la méthode C grâce à l'algorithme de propagation des matrices S.

Une autre variante de la méthode modale est son application avec des **Différences Finies** [51]. Son implémentation est simple et son application aux réseaux s'est avérée fructueuse. Les dérivées partielles de l'équation de propagation sont remplacées par des différences finies. Les expressions des différences finies des champs sont calculées sur des grilles entrelacées aux points x_i :

$$\frac{\partial}{\partial x} = \frac{x_{i+\Delta x} - x_i}{\Delta x}$$

où Δx est le pas de discrétisation. Cette méthode peut être rattachée à la famille des méthodes modales puisqu'après avoir remplacé les dérivées partielles par des différences finies, l'équation de propagation est résolue en cherchant les modes propres. Pour améliorer la vitesse de convergence, une grille non uniforme a été utilisée [11].

Notre méthode se place dans le cadre de la famille des méthodes modales. La différence par rapport aux autres est la base de développement. Notre choix est d'utiliser les bases d'ondelettes [2] [14] à support [3] compact. Les ondelettes, grâce à l'analyse multirésolution, ont déjà prouvé leur potentiel dans de nombreux domaines où l'analyse de Fourier était traditionnellement employée. Elles ont ouvert une voie alternative, souvent mieux adaptée dans des domaines comme le traitement de signal, la compression d'images, etc. L'idée d'extrapoler ces outils pour

2. Ce nom a été inventé par le géophysicien Jean Morlet au début des années 1980 quand les ondelettes ont été employées pour la première fois afin de résoudre des problèmes de traitement des signaux pour la prospection pétrolière. Les ondelettes sont devenues vite très célèbres en devenant des outils fondamentaux pour l'analyse harmonique moderne. L'appellation "ondelettes a été ensuite traduite en anglais par le terme "wavelet".

3. le support d'une fonction f, noté le plus couramment $supp(f)$, est l'ensemble des points de son domaine de définition pour lesquelles la valeur de la fonction est non nulle $supp(f) = \{x; f(x) \neq 0\}$

le traitement des réseaux date d'il y a cinq ans, et des premières études ont été faites au sein du laboratoire LTM. La stratégie multirésolution s'applique aux réseaux lamellaires en utilisant plusieurs échelles de détails pour les zones de discontinuités ou de variations brusques. Ceci peut donner une approximation précise des zones où le champ électromagnétique varie rapidement. Dans le même temps on a aussi une bonne approximation, en utilisant un nombre minimal de données, dans les zones uniformes. Cela nous permet de caractériser d'une manière stable et efficace le champ électromagnétique d'une grande variété de structures.

1.3 Méthode Modale : principe général

Avec une approche modale, on peut résoudre le problème de la diffraction par une structure multicouche périodique en calculant les modes et leurs échanges d'énergie aux interfaces. Étant donné que le cas général d'un réseau périodique selon deux dimensions apporte beaucoup de difficultés d'ordre calculatoire (temps de calcul, capacité mémoire), il est naturel de ne considérer dans un premier temps que le problème de la diffraction par le réseau périodique selon une seule dimension, communément appelé réseau 1D. Il est nécessaire de caractériser la structure 1D, d'analyser les avantages et les inconvénients engendrés par la méthode utilisée, d'optimiser le temps de calcul et de réduire les besoins en mémoire avant de faire l'extension à 2D.

Le schéma général de calcul du champ électromagnétique se divise en plusieurs étapes (voir la figure 1.3) : le calcul des modes propres de la structure dans chaque tranche ; le calcul des couplages entre les modes des diverses tranches ; la résolution finale compte tenu des conditions d'excitation ; le calcul des efficacités de diffraction et du champ électromagnétique. Le calcul est rigoureux dans le sens où, lorsque le nombre de termes utilisés dans le développement du champ augmente, le calcul tend vers la solution exacte.

La brique élémentaire du problème à résoudre est celui d'un réseau de diffraction mono périodique lamellaire (présenté dans la figure 1.4) éclairé par une onde plane monochromatique. La structure est périodique selon l'axe $0x$ et infiniment longue selon l'axe $0z$. Le réseau est éclairé par une onde électromagnétique plane, monochromatique de vecteur d'onde $\mathbf{k}$ situé dans le plan $x0y$ qui fait un angle θ avec l'axe vertical Oy.

Ceci donne naissance à des ondes diffractées et des ondes transmises dans les directions données par la formule des réseaux :

$$sin\theta_m = sin\theta + m\frac{\lambda}{d}$$

1.3 Méthode Modale : principe général

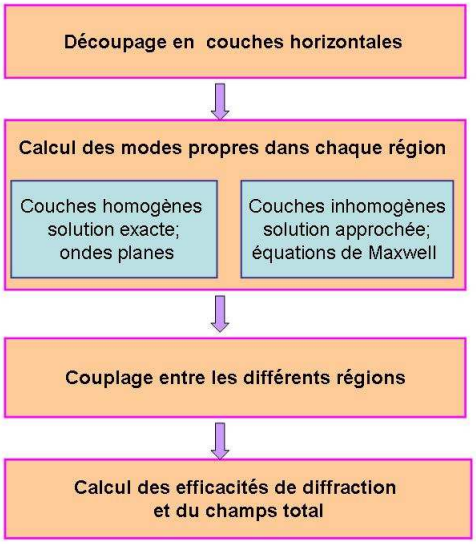


FIGURE 1.3 – *Les grandes étapes de résolution à l'aide d'une méthode modale du problème de diffraction par un réseau*

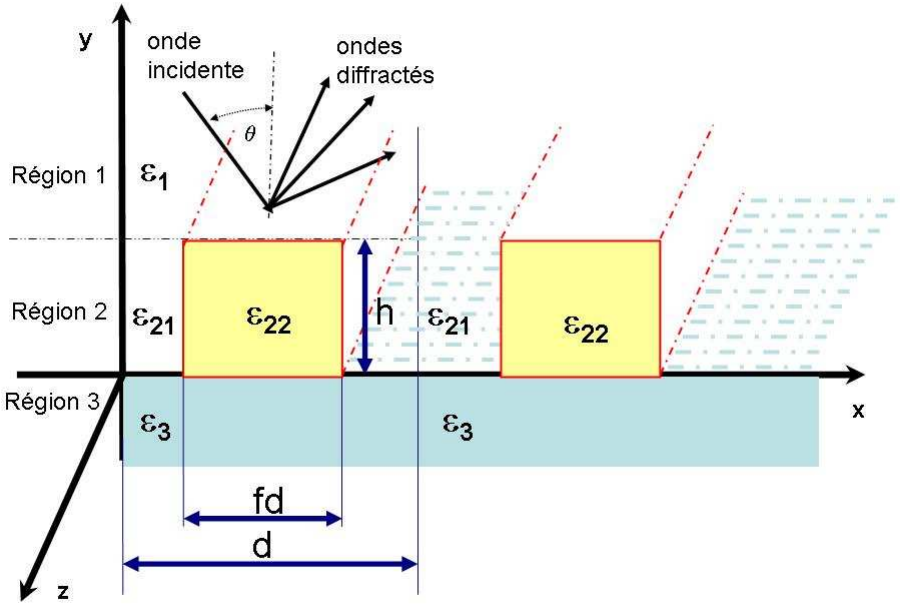


FIGURE 1.4 – *La géométrie du problème de diffraction*

où m est un entier représentant le nombre ordinal d'un ordre diffracté, λ la longueur d'onde, d la période du réseau et θ_m repère la direction de l'onde diffractée dans l'ordre m. Le nombre d'ordres diffractés dépend des paramètres du réseau. En tenant compte que le sinus est une fonction bornée $\left(\mid sin\theta_m \mid \leq 1 \Leftrightarrow \mid sin\theta + m\frac{\lambda}{d} \mid \leq 1 \right)$, les valeurs de m sont tous les entiers compris dans l'intervalle $-(1+sin\theta)\frac{d}{\lambda} \leq m \leq (1-sin\theta)\frac{d}{\lambda}$. Le nombre des ordres difractés est d'autant plus grand que le rapport $\frac{\lambda}{d}$ est petit. Dans le cas ou $\lambda \cong d$, on peut avoir que un ou deux ordres propagatifs : c'est le cas de la diffraction résonante. Le "problème du réseau" consiste à déterminer le champ diffracté et la répartition de l'énergie entre les différents ordres de diffraction.

La géométrie est séparée en trois régions : le milieu incident, le réseau lamellaire, et le substrat. Le milieu incident est formé du demi-plan $y \geq h$ (h étant la hauteur du sillon) et caractérisé par la permittivité ε_1. Le réseau lamellaire est formé de la portion du plan $0 \leq y \leq h$ où la permittivité est modulée. On peut la décrire comme une fonction périodique suivant l'axe $0x$, donnée par la formule :

$$\varepsilon_2(x) = \begin{cases} \varepsilon_{21} & \text{si} \quad 0 \leq x \leq fd \\ \\ \varepsilon_{22} & \text{si} \quad fd \leq x \leq d \end{cases} \tag{1.1}$$

f est un paramètre qui varie entre 0 et 1, appelé facteur de remplissage qui traduit le pourcentage de la période du réseau occupé par le matériau de permittivité ε_{22}.

Le substrat est formé du demi-plan $y \leq 0$ et il est caractérisé par la permittivité ε_3. Ce découpage en trois régions nous assure la condition d'invariance de la permittivité dans chaque région dans la direction $0y$.

1.4 Les équations de Maxwell

Sous leur forme différentielle les équations de Maxwell peuvent s'écrire [52, 53] :

$$\begin{cases} \mathbf{rot}\ \mathbf{H} = \mathbf{J} + \dfrac{\partial \mathbf{D}}{\partial t} \\ \\ \mathbf{rot}\ \mathbf{E} = -\dfrac{\partial \mathbf{B}}{\partial t} \end{cases} \tag{1.2}$$

$$\begin{cases} div\ \mathbf{D} = \rho \\ \\ div\ \mathbf{B} = 0 \end{cases} \tag{1.3}$$

où **E** et **H** sont respectivement le champ électrique et le champ magnétique ; **D** est le déplacement électrique ; **B** l'excitation magnétique ; ρ la densité de charge électrique et **J** la densité de courant.

Sous cette forme, ces équations sont indépendantes du milieu. Pour les résoudre, on doit connaître les relations constitutives. Elles contiennent l'information sur la nature du milieu dans lequel baignent les champs. Dans notre cas les matériaux sont supposés linéaires, isotropes [4], diélectriquement parfaits [5] et non magnétiques [6]. Les relations de milieu s'écrivent :

$$\mathbf{D} = \varepsilon_0 \varepsilon \mathbf{E} \quad \text{et} \quad \mathbf{B} = \mu_0 \mathbf{H}$$

avec ε_0 et μ_0 la permittivité et la perméabilité du vide. Les matériaux ne contiennent ni charges électrostatique localisées ($\rho = 0$) ni courants ($\mathbf{J} = 0$). En général, ces grandeurs dépendent de la géométrie et du temps. Les interfaces séparant deux milieux différents sont invariantes dans le temps, c-à-d leur forme géométrique n'évolue pas au cours du temps. On se place en régime harmonique et on adopte la notation complexe. La dépendance temporelle est exprimée par le facteur $exp(i\omega t)$, où ω représente la pulsation. Le système 1.2 s'écrit en régime harmonique sous la forme :

$$\begin{cases} \mathbf{rot}\ \mathbf{H} = i\omega\varepsilon_0\varepsilon\mathbf{E} \\ \\ \mathbf{rot}\ \mathbf{E} = -i\omega\mu_0\mathbf{H} \end{cases} \tag{1.4}$$

1.4.1 Les équations de propagation

En projetant la relation (1.4) dans un système des coordonnées cartésiennes d'axes $(0, \mathbf{e_x}, \mathbf{e_y}, \mathbf{e_z})$ on obtient un système de six équations à six inconnues, les inconnues étant les six composantes

4. la permittivité électrique relative ε est une grandeur scalaire
5. dépourvu de charge et de courant électrique
6. la perméabilité magnétique $\mu = \mu_0$

scalaires du champ électromagnétique, chacune dépendant des trois variables spatiales x, y, z :

$$\begin{cases} \partial_y E_z - \partial_z E_y = -i\omega\mu_0 H_x \\ \partial_z E_x - \partial_x E_z = -i\omega\mu_0 H_y \\ \partial_x E_y - \partial_y E_x = -i\omega\mu_0 H_z, \\ \partial_y H_z - \partial_z H_y = -i\omega\varepsilon_0\varepsilon E_x \\ \partial_z H_x - \partial_x H_z = -i\omega\varepsilon_0\varepsilon E_y \\ \partial_x H_y - \partial_y H_x = -i\omega\varepsilon_0\varepsilon E_z, \end{cases} \tag{1.5}$$

Dans le cas de notre problème unidimensionnel, la géométrie ainsi que l'excitation (l'onde plane incidente) présentent une invariance suivant la direction $0z$. On démontre aisément que le champ diffracté par le réseau présente la même invariance. Les solutions du système 1.5 sont donc indépendantes de la variable z. L'opérateur différentiel ∂_z est identiquement nul ($\partial_z = 0$). En tenant compte de cette remarque le système 1.5 s'écrit :

$$\begin{cases} \partial_y E_z = -i\omega\mu_0 H_x \\ \partial_x E_z = i\omega\mu_0 H_y \\ \partial_x E_y - \partial_y E_x = -i\omega\mu_0 H_z, \\ \partial_y H_z = -i\omega\varepsilon_0\varepsilon E_x \\ \partial_x H_z = i\omega\varepsilon_0\varepsilon E_y \\ \partial_x H_y - \partial_y H_x = -i\omega\varepsilon_0\varepsilon E_z, \end{cases} \tag{1.6}$$

On remarque que ce système de six équations à six inconnues peut se scinder en deux sous-systèmes indépendants de trois équations à trois inconnues :

$$\begin{cases} \partial_y E_z = -i\omega\mu_0 H_x \\ \partial_x E_z = i\omega\mu_0 H_y \\ \partial_x H_y - \partial_y H_x = i\omega\varepsilon_0\varepsilon E_z, \end{cases} \tag{1.7}$$

et

$$\begin{cases} \partial_y H_z = i\omega\varepsilon_0\varepsilon E_x \\ \partial_x H_z = -i\omega\varepsilon_0\varepsilon E_y \\ \partial_x E_y - \partial_y E_x = -i\omega\mu_0 H_z, \end{cases} \tag{1.8}$$

Dans le système 1.7 le champ électromagnétique a pour composantes H_x, H_y et E_z. La seule composante non nulle du champ électrique étant celle transversale au plan d'incidence (voir fig. 1.5). Cette polarisation est qualifiée de transverse électrique (TE). De même, le système

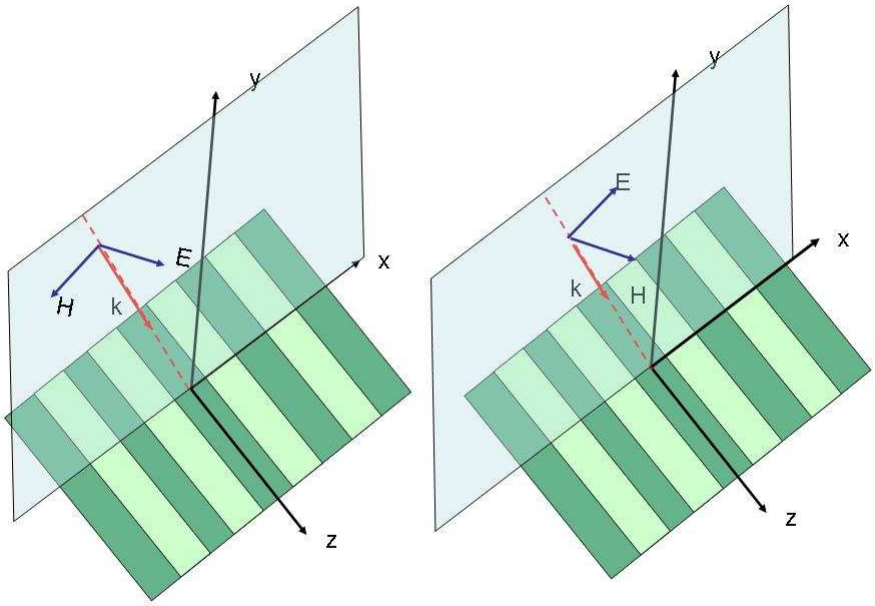


FIGURE 1.5 – *Orientation des champs en polarisation TE et en polarisation TM. L'onde plane, définie par son vecteur d'onde* $\mathbf{k}$*, le champ électrique* $\mathbf{E}$ *et le champ magnétique* $\mathbf{H}$ *forment un trièdre direct.*

1.8 dont la seule composante non nulle du champ magnétique est la composante H_z décrit la polarisation transverse magnétique (TM). Un calcul supplémentaire permet de démontrer que ces composantes transversales du champ électromagnétique, vérifient, dans les deux cas de

polarisation, les équations de propagation suivantes :

$$\begin{aligned} &\mathrm{TE} : \left(\partial_x \partial_x + \partial_y \partial_y + k^2 \varepsilon(x)\right) E_z(x,y) = 0 \\ &\mathrm{TM} : \left(\partial_x \frac{1}{\varepsilon(x)} \partial_x + \frac{1}{\varepsilon(x)} \partial_y \partial_y + k^2\right) H_z(x,y) = 0. \end{aligned} \tag{1.9}$$

Ces deux équations peuvent s'écrire sous la forme symbolique suivante :

$$[\mathcal{L}_x + \partial_y^2] F(x,y) = 0 \tag{1.10}$$

où l'opérateur $\mathcal{L}_x$ agit sur des fonctions dépendant uniquement de la variable spatiale x :

$$\mathcal{L}_x = \begin{cases} \left(\partial_x \partial_x + k^2 \varepsilon(x)\right) & \textbf{TE} \\ \\ \left(\dfrac{1}{\varepsilon(x)}\right)^{-1} \left(\partial_x \dfrac{1}{\varepsilon(x)} \partial_x + k^2\right) & \textbf{TM}, \end{cases} \tag{1.11}$$

et $F(x,y)$ désigne la composante $E_z(x,y)$ ou $H_z(x,y)$ suivant le cas de polarisation. Puisque la permittivité du milieu ne dépend que de x,la technique de séparation des variables peut être employée pour résoudre l'équation 1.10. En posant :

$$F(x,y) = X(x)Y(y)$$

l'équation 1.10 devient :

$$[\mathcal{L}_x + \partial_y^2] XY = 0. \tag{1.12}$$

En divisant par XY (la solution nulle étant exclue), on obtient :

$$\frac{1}{X} \mathcal{L}_x X + \frac{1}{Y} \partial_y^2 Y = 0 \tag{1.13}$$

Les deux termes du premier membre agissant sur des fonctions dépendant des deux variables indépendantes x et y, il existe une constante que nous noterons $k^2 r^2$ telle que :

$$\frac{1}{X} \mathcal{L}_x X = k^2 r^2 \tag{1.14}$$

$$\frac{1}{Y} \partial_y^2 Y = -k^2 r^2. \tag{1.15}$$

De l'équation 1.15 on déduit la dépendance exponentielle du champ électromagnétique suivant la direction de $\mathbf{0y}$:

$$Y(y) = e^{-ikry}.$$

Reste à calculer la dépendance en $X(x)$, solution de l'équation aux valeurs propres 1.14 vérifiant les conditions de pseudopériodicité :

$$X(x+d) = e^{-ik\alpha_0 d} X(x). \tag{1.16}$$

1.4.2 Milieux homogènes

Dans le cas d'un milieu homogène les solutions de l'équation 1.14 sont analytiques. Lorsque la fonction, $\varepsilon(x) = \nu^2(x) = cste$ [7], (dans notre cas le milieu incident et le substrat), le champ $F(x, y)$ vérifie l'équation de Helmoltz dans les deux cas de polarisation :

$$\left(\partial_x\partial_x + \partial_y\partial_y + k^2\nu^2\right) F(x, y) = 0$$

Le champ est alors obtenu comme une combinaison linéaire d'ondes planes appelée le developement de Rayleigh [53] :

$$F(x, y) = \sum_m a_m^{\pm 1} e^{-ik\alpha_m x} e^{-ik\beta_m y} \tag{1.17}$$

où α_m et β_m respectent :

$$\alpha_m^2 + \beta_m^2 = \nu^2.$$

Lorsque $|\alpha_m| \leq \nu$, $\beta_m = \pm\sqrt{\nu^2 - \alpha_m^2}$. Il existe $\theta_m \in (R)$ tel que $\alpha_m = \nu sin(\theta_m)$ auquel cas $\beta_m = \nu cos(\theta_m)$. Ces valeurs de α_m et β_m sont associées à des ondes se propageant suivant les directions θ_m. Pour $\beta_m \geq 0$ ces ondes sont progressives (se propagent vers les valeurs de y croissantes) dans le cas contraire elles sont régressives. Lorsque $|\alpha_m| \geq \nu$, β_m est imaginaire pur, $\beta_m = \pm i\sqrt{\alpha_m^2 - \nu^2}$. Ces solutions sont associées à des ondes évanescentes (c'est-à-dire exponentiellement amorties en y) et antiévanescentes (exponentiellement amorties en $-y$).

Les a_m^{-1} et a_m^{+1} représentent les amplitudes des ondes se propageant vers les y négatifs et respectivement vers les y positifs. On considère que le réseau est éclairé par une seule onde incidente. Tous les a_m^{-1} sont nuls sauf a_0^{-1} que l'on prend par convention égale à 1. Les a_m^{+1} pour les valeurs de m pour lesquels β_m est réel s'identifient aux amplitudes des ondes planes réfléchies, les autres étant les amplitudes des ondes évanescentes.

Dans le substrat, les a_m^{-3}, pour lesquelles β_m du substrat est réel correspondent aux ordres transmis. Les a_m^{+3} correspondraient aux ondes incidentes qui illumineraient la structure par le dessous. La structure étant supposée éclairée uniquement par le dessus, ces ondes n'existent pas. On prendra $a_m^{+3} = 0$, $\forall m$.

1.4.3 Milieux inhomogènes

Lorsque ε n'est plus une constante nous sommes ramenés à la résolution d'une équation aux valeurs propres 1.14 qui peut s'écrire sous la forme symbolique suivante valable pour les deux

7. ν désigne l'indice de réfraction

types de polarisation :

$$\mathcal{L}F = r^2 F \tag{1.18}$$

Les r^2 sont les valeurs propres de l'opérateur $\mathcal{L}$. Désignons par F_m les solutions propres de l'opérateur $\mathcal{L}$ vérifiant les conditions de pseudopériodicité. Les propriétés de complétude de cette base (formée par les vecteurs propres) permettent de développer toute fonction pseudopériodique, en particulier F :

$$F(x,y) = \sum_m a_m e^{\pm ikr_m y} F_m(x) \tag{1.19}$$

$$F_m(x+d) = F_m(x) e^{-iksin\theta d} \tag{1.20}$$

où les r_m sont les racines des valeurs propres r^2 de partie réelle positive et de partie imaginaire négative. Le calcul numérique consistera donc à évaluer numériquement les fonctions propres $F_m(x)$.

Dans la méthode modale de Fourier ces fonctions propres sont développées sur la base de Fourier :

$$F_m(x) = \sum_n F_{nm} e^{-ik\alpha_n x} \tag{1.21}$$

avec $\alpha_n = \alpha_0 + n\frac{\lambda}{d}$.

En utilisant les bases de Fourier, des problèmes de convergence lente à cause du phénomène de Gibbs nous ont incité à étudier les convergences en utilisant d'autres type de base. Les modes propres sont développés comme suit :

$$F_m(x) = \sum_n F_{nm} B_n(x) \tag{1.22}$$

avec $B_n(x)$ la fonction de base.

1.5 Conditions aux limites et calcul des efficacités

Afin de déterminer le champ dans chaque milieu, il faut déterminer les vecteurs des amplitudes A des rayons réfléchis et transmis de chaque région (voir figure 1.6) :

$$A_l^{\pm} = \begin{pmatrix} \vdots \\ a_{lm}^{\pm} \\ \vdots \end{pmatrix} \tag{1.23}$$

l repère la région, $l = \{1, 2, 3\}$. Une fois les amplitudes déterminées on peut calculer les efficacités de diffraction et le champ total de la structure. Les composantes tangentielles doivent vérifier les conditions de continuité lors de la traversée d'une surface de discontinuité physique. Cette discontinuité peut être le fait d'une rupture de permittivité ou perméabilité. Dans le cas du réseau lamellaire, le calcul du champ total de la structure nous a conduit à subdiviser la structure en trois parties suivant les demi-plans $y = 0$ et $y = h$. Nous sommes dans un cas de discontinuité de permittivité, chaque milieu est décrit par une fonction de permittivité différente. La transcription de la continuité des composantes tangentielles du champ à l'interface

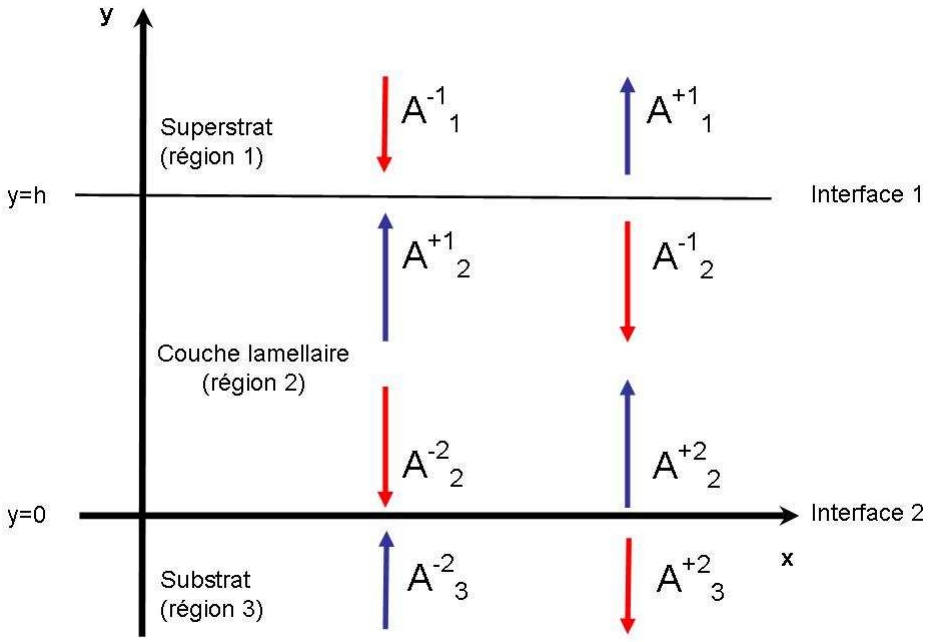


FIGURE 1.6 – ***A*** *désigne une amplitude du champ. Les signes* $+$ *et* $-$ *correspondent aux direction de propagation (vers les* y *positifs ou vers les* y *négatifs). Le chiffre de l'exposant désigne l'interface sur laquelle interagit l'amplitude et le chiffre en indice est pour la région dans laquelle elle se trouve.)*

$y = 0$ conduit au système d'équation suivant :

$$\begin{cases} F_1(x, y = 0) = F_2(x, y = 0) \\ \\ \dfrac{1}{\chi_k^1}\dfrac{\partial}{\partial y}F_1(x, y = 0) = \dfrac{1}{\chi_k^2}\dfrac{\partial}{\partial y}F_2(x, y = 0) \end{cases} \tag{1.24}$$

De même, la continuité du champ à l'interface $y = h$ s'écrit :

$$\begin{cases} F_2(x, y = h) = F_3(x, y = h) \\ \\ \dfrac{1}{\chi_k^2}\dfrac{\partial}{\partial y}F_2(x, y = h) = \dfrac{1}{\chi_k^3}\dfrac{\partial}{\partial y}F_3(x, y = h) \end{cases} \tag{1.25}$$

En polarisation TE

$$\begin{cases} F(x, y) = E_z(x, y) \\ \\ \chi_k = \dfrac{1}{\omega\mu_0\mu} \end{cases} \tag{1.26}$$

En polarisation TM

$$\begin{cases} F(x, y) = H_z(x, y) \\ \\ \chi_k = \dfrac{1}{\omega\varepsilon_0\varepsilon} \end{cases} \tag{1.27}$$

les chiffres de l'indice pour l'expression de $F(x, y)$ faisant référence à la région dans laquelle le champ se trouve.

Grâce aux conditions de continuité, nous pouvons déterminer toutes les amplitudes des ondes réflechies et transmises. De plus, au sujet de l'étude du problème de diffraction, il est également important de chercher à définir les amplitudes des ondes entrantes et sortantes.

Les amplitudes peuvent être reliées entre elles en utilisant plusieurs formalismes comme par exemple les formalisme des matrices de dispersion **S**, des matrices de transmission **T**, des matrices d'impedance **R**, voire le formalisme Dirichlet to Newmann. Notre choix est le formalisme des matrices **S**, connu pour sa stabilité [54]. Il est détaillé dans l'annexe A.

Une fois calculées, les amplitudes des ondes réflechies et transmises on peut déduire le champ dans tout l'espace et on peut aussi détérminer les efficacités, qui représentent la fraction d'énergie réflechie et transmise dans les différentes directions par rapport à l'énergie incidente. L'efficacité en réflexion selon l'ordre de réflexion q est donnée par :

$$R_q = \frac{a_q^+ a_q^{+*} r_q \sum_n F_{nq} F_{nq}^*}{a_0^- a_0^{-*} r_0 \sum_n F_{n0} F_{n0}^*} \tag{1.28}$$

La formule 1.28 est valable dans les deux cas de polarisation. En transmission, pour le cas de polarisation TE l'efficacité transmise selon l'ordre q de transmission s'écrit comme :

$$T_q = \frac{a_q^- a_q^{-*} r_q \sum_n F_{nq} F_{nq}^*}{a_0^- a_0^{-*} r_0 \sum_n F_{n0} F_{n0}^*} \tag{1.29}$$

En polarisation TM, l'efficacité de transmission est donnée par :

$$T_q = \frac{\varepsilon_1}{\varepsilon_3} \frac{a_q^- a_q^{-*} r_q \sum_n F_{nq} F_{nq}^*}{a_0^- a_0^{-*} r_0 \sum_n F_{n0} F_{n0}^*} \tag{1.30}$$

On détaille en Annexe B la façon de les obtenir.

1.6 Bilan

Dans ce chapitre nous avons rappelé les différentes notions fondamentales d'électromagnétisme nécessaires à la présentation du travail effectué durant la thèse ainsi que les principales méthodes de résolution utilisées dans la littérature. Il est à noter que l'intégration analytique dans une direction de l'espace autorise la modélisation d'objets très étendus dans cette direction. C'est un avantage important des méthodes modales en général par rapport aux méthodes reposant sur une discrétisation tridimensionnelle.

Pour caractériser les structures 1D, il existe une gamme large des méthodes et on peut obtenir des résultats fiables pour la plupart des structures étudiées. Les travaux de cette thèse ont été motivés par la possibilité d'introduction des bases hiérarchiques dans le but gagner en précision et rapidité. Ces travaux représentent les préliminaires pour l'extension au cas 2D, la où la plupart des méthodes classiques échouent à donner des convergences avec des temps de calcul raisonnables.

Chapitre 2

Méthode Modale avec des fonctions Splines

Sommaire

2.1 Introduction 32

2.2 Description de la méthode 33

2.2.1 Méthode des résidus pondérés 33

2.2.2 Application aux équations de Maxwell 35

2.2.3 Description des fonctions B_n et T_q 37

2.2.4 Périodicité et pseudo périodicité 42

2.3 Résultats numériques 44

2.3.1 Influence des fonctions test 45

2.3.2 Influence des fonctions de base 51

2.3.3 Influence de la grille de discrétisation 52

2.3.4 Comparaison avec d'autres méthodes 55

2.3.5 Résolution spatiale adaptative 61

2.4 Bilan 66

Ce chapitre est consacré à l'aperçu du calcul spécifique de résolution des équations de Maxwell en utilisant la méthode modale de décomposition avec des fonctions splines. Puis, les performances de la méthode sont étudiées en utilisant des cas (théoriques) couramment utilisés dans la littérature ainsi que des cas pratiques utilisées pour la lithographie.

2.1 Introduction

Ethymologiquement, le mot *spline* signifie *latte*. Cette définition a été inspirée du domaine de l'industrie où on utilise des lattes souples en bois pour tracer des courbes assez harmonieuses de formes variées qui passent par un certain nombre de points, ne pouvant pas être tracées avec un compas.

Les fonctions splines ont été introduites pour la première fois par Schoenberg [55] en 1946. Dès leur apparition, les splines se sont avérées des instruments adaptés pour les calculs numériques et leurs propriétés. Elles ont éveillé l'intérêt de plusieurs chercheurs [56, 57, 58]. Entre temps, dans l'industrie automobile d'importants progrès ont été faits par Bézier [59] en utilisant des polynômes de Bernstein et par Casteljau [60]. Ces techniques ont eu l'avantage de permettre la description des formes lisses par des polynômes lisses. Les fonctions B-splines sont apparues comme une progression naturelle, étant définies par morceaux comme des polynômes de Bernstein [61,62]. L'étude générale de ces fonctions et de leurs propriétés peuvent être trouvées dans les livres de Boor [63] et Schumaker [64].

Aujourd'hui, les fonctions B-splines sont universellement reconnues comme des outils très puissants dans la théorie de l'approximation et elles sont devenues des outils de base dans beaucoup de domaines d'application comme les logiciels de dessin de surfaces [1] [65], la compression des données [2] [66], l'analyse des images [3] [67], le traitement de signal [68] et la liste peut encore continuer. Grâce à leur définition analytique comme morceaux de polynômes et à leurs propriétés comme la continuité, le support compact, les fonctions splines fournissent des algorithmes stables, efficaces et rapides à calculer.

Nous cherchons à appliquer les fonctions splines au problème de l'électromagnétisme et bénéficier de leurs propriétés. Des premiers travaux ont démarré en 2005 au sein du laboratoire LTM [69]. A la lumière de ces résultats encourageants, on a suivi cette voie prometteuse. Elle nous a permis d'obtenir plusieurs résultats intéressants en utilisant des changement de coordonnées [70, 71, 72] qui adaptent spatialement la grille de discrétisation et ressèrent les points autour des discontinuités. Récemment une approche basée sur les développements sur des bases B-spline en utilisant une technique des nœuds confondus aux discontinuités a été implémentée

1. Tous les logiciels de dessin, dès plus sophistiqués comme par exemple CAGD (utilisé dans la constructions automobile ou aéronautique), à des logiciels très simples comme Microsoft Draw, utilisent pour représenter des géométries complexes des algorithmes à la base des fonctions splines

2. Pour la visualisation des objets en 3D en temps réel il est devenu impératif d'élaborer des représentations compactes des ces donnés

3. Dans les services de télémédecine, les jeux vidéo ou la production de films d'animation 3D

par Bouchon et al [73].

2.2 Description de la méthode

Les équations des systèmes 1.7 et 1.8 peuvent être écrites sous la forme concise suivante :

$$\mathcal{L}f = g \tag{2.1}$$

où $\mathcal{L}$ est un opérateur linéaire et f, g des fonctions inconnues. Une technique de résolution numérique de ce type d'équations, est la méthode des résidus pondérés appelé communément la méthode des moments [74, 75], méthode des projections [76] ou encore méthode de Petrov-Galerkin [76].

2.2.1 Méthode des résidus pondérés

La méthode des résidus pondérés est une méthode qui permet de résoudre numériquement un système d'équations dans le cas où celui-ci ne présenterait pas de solution analytique. Cette méthode comporte deux étapes. Dans un premier temps on définit une base B_n sur laquelle sont développées les fonctions f et g :

$$f = \sum_n f_n B_n \quad g = \sum_n g_n B_n. \tag{2.2}$$

f_n et g_n sont des coefficient inconnus. Dans la pratique, seulement un nombre fini N d'éléments de base est pris en compte. L'équation 2.2 devient :

$$f = \sum_{n=1}^{N} f_n B_n \quad g = \sum_{n=1}^{N} g_n B_n. \tag{2.3}$$

La seconde étape consiste en la détermination des coefficients f_n et g_n. Pour ce faire, on définit une fonction résidu R comme suit :

$$R_N = \sum_n^N \mathcal{L} f_n B_n - \sum_n^N g_n B_n, \tag{2.4}$$

On cherche à minimiser le résidu R_N. Cette minimisation est équivalente à chercher les fonctions f_n et g_n qui donnent la meilleure solution numérique de l'équation 2.1. Dans ce but on

utilise une technique dite de pondération. Ceci consiste à choisir un ensemble de fonctions de pondération [4] T_q et à imposer que :

$$< Tq, R_N >= 0 \quad q \in \{1, 2, \ldots N\} \tag{2.5}$$

où $<,>$ désigne le produit scalaire :

$$< u, v >= \int u^* v dx \tag{2.6}$$

u et v étant deux fonctions quelconques avec $u*$ le complexe conjugué de u.

On observe que la méthode de pondération permet de transformer l'équation 2.1 en une équation matricielle qui peut être écrite comme :

$$\sum_n^N < T_q, \mathcal{L}B_n > f_n - \sum_n^N < T_q, B_n > g_n = 0. \tag{2.7}$$

où f_n et g_n sont les vecteurs colonne à déterminer. La question qui se pose est comment choisir les fonctions T_q. Plusieurs choix existent mais les techniques les plus utilisées sont :

a) *Technique de collocation par points,* appelé aussi "point matching" [77,74]. Dans cette technique, on détermine les valeurs des coefficients f_n en imposant que le résidu soit nul en certains points $(x = x_q)$. Dans ce cas, les fonctions T_q sont des fonctions Dirac. L'avantage principal de cette méthode est que l'intégration représentée par le produit scalaire 2.7 devient triviale :

$$\mathcal{L} \sum_n f_n B_n(x_q) - \sum_n g(x_q) = 0 \quad \forall q \in \{1, 2, \ldots\} \tag{2.8}$$

En résolvant l'équation 2.8 on obtient les coefficients f_n qui sont aussi les solutions de l'équation 2.1. Les limitations de cette technique résident dans le fait qu'on ne regarde la valeur du résidu qu'en des points précis. Au voisinage des nœuds x_q la reconstruction de f à partir de la base B_n peut faire apparaître d'éventuelles oscillations du résidu qui peuvent engendrer des instabilités numériques.

b) *Technique de collocation par sous domaines.* [74] Dans cette technique on impose au résidu d'avoir une moyenne nulle sur un certain nombre des sous domaines (intervalles).

Dans ce cas, les fonctions T_q sont des fonctions à support compact. La plupart du temps on utilise des fonctions portes définies comme :

$$T_q(x) = \begin{cases} 1 & x_q \leq x \leq x_{q+1} \\ \\ 0 & \text{ailleurs.} \end{cases} \tag{2.9}$$

4. couramment appelées fonctions test ou bien fonctions de projections

La méthode de collocation par sous domaines est intéressante car elle fait intervenir une mesure moins locale des performances de la fonction test à satisfaire l'équation 2.7. Cette technique n'est pas exemptée de danger puisqu'elle se base sur un moyennage du résidu. Par exemple dans le cas de la fonction porte, on pourrait imaginer un résidu qui passerait par des valeurs très grandes, positives et négatives mais dont la moyenne demeurerait nulle. On aurait alors une mauvaise approximation de la solution exacte.

c) *Méthode des moindres carrés.* Pour éviter les problèmes de mauvaise approximation de la fonction exacte liée au fait qu'on moyenne des valeurs positives et négatives, on peut imaginer un critère encore plus sévère : minimiser le carré du résidu sur tout le domaine de solution. La minimisation du résidu revient à trouver la plus petite distance entre g et $\mathcal{L}B_n$ au sens de la norme L^2. Pour ce faire on choisit $T_q = \mathcal{L}B_n$.

d) *Méthode de Galerkin*[5] [78]. Dans cette technique, les pondérations sont choisies parmi les fonctions (B_n) qui ont servi à construire la solution approchée, $T_q \in \{B_n, n \in \mathbb{N}^*\}$. La méthode de Galerkin est la méthode des résidus la plus utilisée.

Dans la suite, on appelle génériquement B_n les fonctions de base et T_q les fonctions test avec $n, q \in \{1, 2, \ldots, N\}$. Ces fonctions nous permettent de trouver les coefficients inconnus, f_n. Ces coefficients ne donnent accès qu'à une approximation de la solution. Cette approximation est entachée d'une erreur qui sera plus ou moins grande. Il est difficile d'établir une hiérarchie a priori de ce qui est la "meilleure approximation" puisque dans la pratique on ne connaît pas la solution exacte. Il y a cependant des critères numériques a posteriori comme la stabilité, la vitesse de convergence, la robustesse qui nous donne des indication sur la qualité de l'approximation.

2.2.2 Application aux équations de Maxwell

La première étape consiste à développer les fonctions inconnues du problème (les composantes du champ) sur des fonctions de base :

$$f = \sum_{n=1}^{N} f_n B_n$$

5. Boris G. Galerkin est mathématicien russe réputé pour ses contributions à l'étude de poutres et plaques élastiques. Son nom est lié à la méthode de résolution approchée des structures élastiques qui est une de bases de la méthode des éléments finis. La méthode de Galerkin est applicable uniquement quand l'opérateur $\mathcal{L}$ est autoadjoint.

où f désigne une composante quelconque du champ électromagnétique (E_z, H_x, H_y pour la polarisation TE et H_z, E_x, E_y pour la polarisation TM). La deuxième étape consiste à déterminer les coefficients du champ f_n (E_{z_n}, H_{x_n}, H_{y_n} pour TE et H_{z_n}, E_{x_n}, E_{y_n} pour TM) en projetant le système obtenu sur des fonctions test T_q. Les équations de Maxwell (1.7 et 1.8) deviennent :

$$\mathbf{TE}: \begin{cases} kr \sum_{n=1}^{N} \langle T_q, B_n \rangle E_{z_n} = \omega\mu_0 \sum_{n=1}^{N} \langle T_q, B_n \rangle H_{x_n} \\ \sum_{n=1}^{N} \left\langle T_q, \dfrac{dB_n}{dx} \right\rangle E_{z_n} = i\omega\mu_0 \sum_{n=1}^{N} \langle T_q, B_n \rangle H_{y_n} \\ \sum_{n=1}^{N} \left\langle T_q, \dfrac{dB_n}{dx} \right\rangle H_{y_n} + ikr \sum_{n=1}^{N} \langle T_q, B_n \rangle H_{x_n} = i\omega\varepsilon_0 \sum_{n=1}^{N} \langle T_q, \varepsilon B_n \rangle E_{z_n} \end{cases} \tag{2.10}$$

et

$$\mathbf{TM}: \begin{cases} \dfrac{-kr}{\varepsilon_0} \sum_{n=1}^{N} \left\langle T_q, \dfrac{1}{\varepsilon} B_n \right\rangle H_{z_n} = \omega \sum_{n=1}^{N} \langle T_q, B_n \rangle E_{x_n} \\ \sum_{n=1}^{N} \left\langle T_q, \dfrac{dB_n}{dx} \right\rangle H_{z_n} = -i\omega\varepsilon_0 \sum_{n=1}^{N} \langle T_q, \varepsilon B_n \rangle E_{y_n} \\ \sum_{n=1}^{N} \left\langle T_q, \dfrac{dB_n}{dx} \right\rangle E_{y_n} + ikr \sum_{n=1}^{N} \langle T_q, B_n \rangle E_{x_n} = -i\omega\mu_0 \sum_{n=1}^{N} \langle T_q, B_n \rangle H_{z_n} \end{cases} \tag{2.11}$$

$\forall q \in \{1, N\}$. Considérons le cas de la polarisation TE (équation 2.10). En substituant les vecteurs $[H_{x_n}]$ (de la première équation) et $[H_{y_n}]$ (deuxième équation) dans la troisième équation de 2.10 on obtient une équation de propagation vérifiée par le vecteur $[E_{z_n}]$. En procédant de même pour la polarisation TM, on obtient l'équation de propagation vérifiée par $[H_{z_n}]$:

$$\left[\mathbf{G}^{-1}\left(\frac{1}{k^2}\mathbf{D}\mathbf{G}^{-1}\mathbf{D} + \mathbf{G}^{\varepsilon}\right)\right][E_{z_n}] = r^2[E_{z_n}] \tag{2.12a}$$

$$\left[\left(\mathbf{G}^{\frac{1}{\varepsilon}}\right)^{-1}\left(\frac{1}{k^2}\mathbf{D}\left(\mathbf{G}^{\varepsilon}\right)^{-1}\mathbf{D} + \mathbf{G}\right)\right][H_{z_n}] = r^2[H_{z_n}] \tag{2.12b}$$

Les matrices $\mathbf{G}$, $\mathbf{D}$, $\mathbf{G}^{\varepsilon}$ et $\mathbf{G}^{\frac{1}{\varepsilon}}$ contiennent les produits scalaires entre les fonctions de base, les fonctions test et leurs derivées :

$$\mathbf{G} = \begin{pmatrix} \langle T_1, B_1 \rangle & \cdots & \langle T_1, B_N \rangle \\ \vdots & \ddots & \vdots \\ \langle T_N, B_1 \rangle & \cdots & \langle T_N, B_N \rangle \end{pmatrix}, \tag{2.13}$$

$$\mathbf{D} = \begin{pmatrix} \left\langle T_1, \frac{d}{dx} B_1 \right\rangle & \cdots & \left\langle T_1, \frac{d}{dx} B_N \right\rangle \\ \vdots & \ddots & \vdots \\ \left\langle T_N, \frac{d}{dx} B_1 \right\rangle & \cdots & \left\langle T_N, \frac{d}{dx} B_N \right\rangle \end{pmatrix}, \tag{2.14}$$

$$\mathbf{G}^{\varepsilon} = \begin{pmatrix} \langle T_1, \varepsilon B_1 \rangle & \cdots & \langle T_1, \varepsilon B_N \rangle \\ \vdots & \ddots & \vdots \\ \langle T_N, \varepsilon B_1 \rangle & \cdots & \langle T_N, \varepsilon B_N \rangle \end{pmatrix}, \tag{2.15}$$

$$\mathbf{G}^{\frac{1}{\varepsilon}} = \begin{pmatrix} \left\langle T_1, \frac{1}{\varepsilon} B_1 \right\rangle & \cdots & \left\langle T_1, \frac{1}{\varepsilon} B_N \right\rangle \\ \vdots & \ddots & \vdots \\ \left\langle T_N, \frac{1}{\varepsilon} B_1 \right\rangle & \cdots & \left\langle T_N, \frac{1}{\varepsilon} B_N \right\rangle \end{pmatrix}. \tag{2.16}$$

2.2.3 Description des fonctions B_n et T_q

Le choix des fonctions de base et de test est le point clé dans l'application de la méthode des moments. Le but est d'obtenir des résultats stables, des matrices bien conditionnées même pour des nombres de troncature élevés. Un couple des fonctions base/test bien adapté va donner des résultats d'une précision acceptable avec un faible effort de calcul. Cela peut ne pas être valable pour toutes les structures. On peut par exemple avoir des comportements différents quand on traite des réseaux diélectriques ou des réseaux métalliques. Pour certaines situations précises, comme par exemple quand on a à simuler le comportement d'un petit motif sur une grande période ou une structure avec un grand contraste entre les indices optiques le choix le plus approprié peut encore différer.

Notre choix des fonctions de base et de test s'est porté sur les fonctions splines. L'idée d'utiliser ce type des fonctions n'est pas fortuite. Ces fonctions ont déjà fait leurs preuves dans la résolution des problèmes d'interpolation. Elles sont facile à manipuler et constituent le premier pas vers l'analyse multi résolution et les ondelettes splines. Dans des nombreux domaines les ondelettes et l'analyse multi résolution ont ouvert une voie alternative aux problèmes où l'analyse de Fourier était classiquement utilisée et ont résolu des problèmes comme celui de phénomène de Gibbs [6] qui apparaît aussi dans le cadre du problème de diffraction par des ré-

6. Le phénomène de Gibbs, nommé d'après le physicien J. Williard Gibbs, apparaît lorsqu'on utilise des séries de Fourier pour approcher numériquement des fonctions périodiques avec des fortes discontinuités. Ce phénomène crée des oscillations de la fonction approchée autour des discontinuités. Il est très indésirable et dans des domaines

seaux.

Au cours des années, les bases de splines les plus étudiées ont été celles des splines avec des fonctions des puissances tronquées (définies pour la première fois par Schoenberg [55]), les splines cardinales (définies par des produits de convolution) et les B-splines (définies de manière récursive [61, 62]). Les premières, malgré le fait qu'elles soient bien adaptées pour l'analyse, se sont avérées fournir des matrices mal conditionnées [61]. Les deuxièmes sont difficiles à manipuler et, dans leur utilisation pour des problèmes d'interpolation il apparaissent des phénomènes indésirables comme celui de Runge [7]. Boor [61] et Cox [62], en continuant les travaux de Schoenberg, ont montré que les B-splines fournissent des méthodes stables avec des matrices bien conditionnées.

Les splines qu'on utilise dans notre approche sont les fonctions B-splines qui sont définies par morceaux sous forme de polynômes. Sur tous les morceaux, on utilise des polynômes de même degré qui déterminent le degré de la B-spline. La lettre B souligne que les B-splines forment une base, donc n'importe quelle fonction peut être représentée comme combinaison linéaire des B-splines. Une fonction B-spline possède des propriétés intéressantes comme la symétrie, le support compact, elle est définie positive et les morceaux des polynômes qui la définissent sont raccordés entre eux.

Considérons une période du réseau lamellaire (voir figure 1.4) $I = [0, d]$ qui est divisé en N sous intervalles $[x_n, x_{n+1}]$ à l'aide d'une suite croissante de $N + 1$ points :

$$0 = x_0 < x_1 < \cdots < x_n < x_{n+1} < \cdots < x_{N+1} = d$$

Les B-splines sont définies sur un support compact minimal qui respecte les propriétés de symétrie et de raccordement entre les morceaux des polynômes. Ce support varie en fonction du degré de la spline. On fait la convention qu'une fonction B-spline de degré p qui commence à être non nulle à partir du point x_n est notée B_n^p. Son support, qui contient $p+2$ points consécutifs sera $[x_n, x_{n+p+1}]$. Les B-splines ont été définies d'une manière récursive par [62, 61] comme :

$$B_n^0(x) = \begin{cases} 1 & x_n \leq x \leq x_{n+1} \\ \\ 0 & \text{ailleurs.} \end{cases} \tag{2.17}$$

comme le traitement de signal, l'imaginerie par résonance magnétique, il a été évité en utilisant l'analyse multi résolution à la place de l'analyse de Fourier.

7. C'est un phénomène découvert par Carl David Tolmé Runge qui se produit dans certain problèmes d'interpolation des fonctions. Lorsqu'on augmente le nombre des points d'interpolation il arrive que la fonction interpolante oscille et s'écarte de plus en plus de la fonction à interpoler. Ce phénomène est similaire au phénomène de Gibbs

$$B_n^p(x) = \frac{x - x_n}{x_{n+p} - x_n} B_n^{p-1}(x) + \frac{x_{n+p+1} - x}{x_{n+p+1} - x_{n+1}} B_{n+1}^{p-1}(x) \tag{2.18}$$

Dans nos travaux, on a utilisé des B-splines de degré au plus égal à trois. Les splines de degré zéro, définies par 2.17, ont un support de longueur $x_{n+1} - x_n$. Elles ont l'allure des fonctions portes (voir figure 2.1).

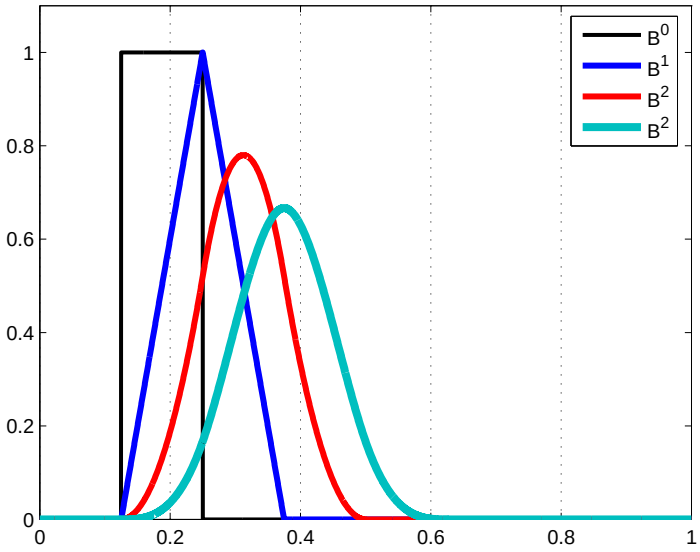


FIGURE 2.1 – *Fonctions B-spline de degré 0,1,2 et 3.*

En utilisant la formule de recurence 2.18, on obtient les splines de degré 1 qui sont définies comme :

$$B_n^1(x) = \begin{cases} \dfrac{x - x_n}{x_{n+1} - x_n} & \text{si} \quad x_n \leq x \leq x_{n+1} \\ \dfrac{x_{n+2} - x}{x_{n+2} - x_{n+1}} & \text{si} \quad x_{n+1} \leq x \leq x_{n+2} \\ 0 & \text{ailleurs} \end{cases} \tag{2.19}$$

Elles ont un support de $[x_n, x_{n+2}]$ et une forme triangulaire (voir figure2.1).

Les splines quadratiques de support $[x_n, x_{n+3}]$ sont construites à partir de polynômes de

dégré 2 :

$$B_n^2(x) = \begin{cases} (x - x_n)^2 & \text{si} \quad x_n \leq x \leq x_{n+1} \\ \left(x - \dfrac{x_{n+2} - 3x_{n+1}}{2}\right)^2 + 3\left(\dfrac{x_{n+2} - x_{n+1}}{2}\right)^2 & \text{si} \quad x_{n+1} \leq x \leq x_{n+2} \\ (x - x_{n+3})^2 & \text{si} \quad x_{n+2} \leq x \leq x_{n+3} \\ 0 & \text{ailleurs} \end{cases} \tag{2.20}$$

Les splines cubiques ont un support de longueur $x_{n+4} - x_n$. Elles sont définies comme :

$$B_n^3(x) = \begin{cases} \dfrac{(x - x_n)^3}{6(x_{n+1} - x_n)^3} & \text{si} \quad x_n \leq x \leq x_{n+1} \\ \dfrac{(x_{n+2} - x)^3}{2(x_{n+2} - x_{n+1})^3} - \dfrac{(x_{n+2} - x)^2}{(x_{n+2} - x_{n+1})^2} + \dfrac{2}{3} & \text{si} \quad x_{n+1} \leq x \leq x_{n+2} \\ \dfrac{(x - x_{n+2})^3}{2(x_{n+3} - x_{n+2})^3} - \dfrac{(x - x_{n+2})^2}{(x_{n+3} - x_{n+2})^2} + \dfrac{2}{3} & \text{si} \quad x_{n+2} \leq x \leq x_{n+3} \\ \dfrac{(x_{n+4} - x)^3}{6(x_{n+4} - x_{n+3})^3} & \text{si} \quad x_{n+3} \leq x \leq x_{n+4} \\ 0 & \text{ailleurs} \end{cases} \tag{2.21}$$

Une fonction B-spline B_n^p est de classe [8] C^{p-1}. La fonction B_n^p est non nulle sur l'intervalle $[x_n, x_{n+p+1}]$. Sur l'intervalle $[x_n, x_{n+1}]$ il y a exactement $p+1$ B-splines non nulles, la première étant B_{n-p+1}^p qui finit dans le nœud x_{n+1} et la dernière B_n^p qui commence dans le nœud x_n. A titre d'exemple, on considère la superposition des fonctions B-spline de degré 1, B^1 illustré dans la figure 2.2. On remarque que dans chaque intervalle $[x_n, x_{n+1}]$ il y a exactement deux triangles non nuls. Grâce à cette propriété on a :

$$B_n^p(x) B_m^p(x) = 0 \text{ pour tout } |m - n| \geq p. \tag{2.22}$$

Dans le développement d'une fonction arbitraire f, qu'on écrit comme combinaison linéaire

8. La fonction et ses dérivées jusqu'à l'ordre p sont continues. Les fonctions B_n^1 sont de classe C^0 puisque toutes les dérivées sont discontinues. B_n^0 est de classe C^{-1}. Par convention on dit qu'une fonction appartient à la classe C^{-1} si elle est discontinue.

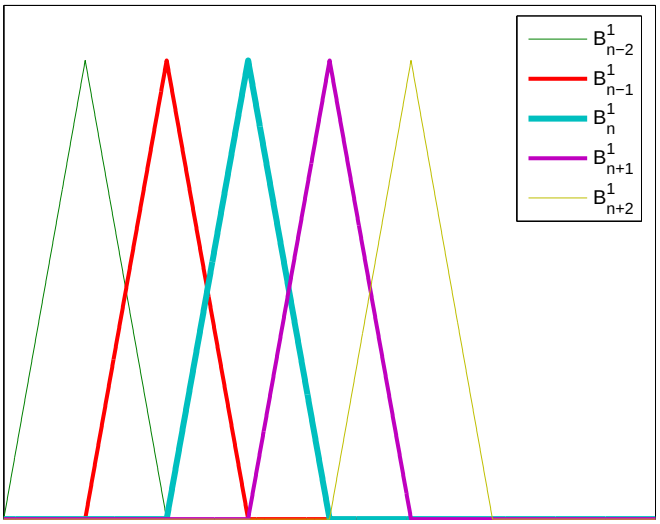


FIGURE 2.2 – *Superposition des fonctions B-spline de degré 1. Chaque triangle se superpose avec son voisin de gauche et son voisin de droite, donc pour un n fixé, seuls les produits $B^1_n(x)B^1_{n-1}(x)$, $B^1_n(x)B^1_n(x)$ et $B^1_n(x)B^1_{n+1}(x)$ sont non nuls.*

des fonctions de base B^p (voir l'équation 2.3), en utilisant par exemple la méthode de Galerkin (les fonctions test sont choisies parmi les fonctions de base) on a :

$$< B^p_q|f >=< B^p_n|\sum_{n=1}^{N} f_n B^p_n >= \sum_{n=q-p}^{n=q+p} f_n < B^p_q|B^p_n > \quad q \in \{1, 2, \ldots, N\}. \tag{2.23}$$

Même si les sommations varient de $n = 1$ jusqu'à N, le nombre des termes à sommer est bien moindre. Grâce à la propriété 2.22 beaucoup de produits scalaires sont nuls. On effectue donc un nombre minimal d'opérations de sommations.

Notons que la fonction B^p_n n'est rien d'autre qu'une translation d'un pas de discrétisation de la fonction B^p_{n-1} (voir l'exemple présenté dans la figure 2.2). Grâce à cela les produits scalaires entre $B^p_n(x)B^p_n(x)$ ont la même valeur $\forall n$. Par ailleurs, les produits scalaires entre $B^p_n(x)B^p_{n-m}(x)$ et $B^p_n(x)B^p_{n+m}(x)$ ont aussi la même valeur puisque les fonctions splines sont symétriques.

2.2.4 Périodicité et pseudo périodicité

Afin d'appliquer le développement sur des fonctions B-spline à notre problème de diffraction par un réseau périodique, on modifie la base de telle manière qu'elle soit pseudopériodique puisque le champ est pseudopériodique. Cette opération se fait en deux étapes, premièrement on rend la base périodique et après on multiplie la base par le facteur de pseudo périodicité. La périodisation se fait facilement grâce au fait que les splines ne sont rien d'autre que des translations selon l'axe $0x$ de B_1^p. Au voisinage du point $x = 0$ (la limite inférieure de l'intervalle étudié) on a les fonctions B_0^p, $B_{-1}^p, \ldots, B_{-p}^p$ qui finissent dans les points x_1,$x_2, \ldots, x_p$. A la limite supérieure $x = d$ les splines B_{N-p}^p, $B_{N-p+1}^p, \ldots, B_N^p$ dépassent de l'intervalle $[0, d]$. Pour prendre en compte la périodicité il suffit de relier les bouts des fonctions qui coupent les deux extrémités de l'intervalle $[0, d]$. Dans la figure 2.3 on donne un exemple en utilisant les splines quadratiques. La spline B_{-2} a pour support $[x_{-2}, x_1]$, donc elle rentre en étant non nulle sur l'intervalle $[x_0, x_1]$. La spline B_{N-2} a pour support $[x_{N-2}, x_{N+1}]$, donc il y a une partie de son support qui dépasse l'intervalle $[0, d]$. Grâce aux propriétés énoncées ci-dessous, le bout de la spline B_{N-2} qui dépasse l'intervalle de périodicité est exactement égal à la position de la spline B_{-2} à une translation de x_N près. Il suffit donc de relier la spline B_{N-2} avec la spline B_{-2}. En faisant de même pour les splines B_{-1} et B_{N-1} on assure la périodicité.

$$B_{-1} = B_{-1} \cup B_{N-1}, \; B_{N-2} = B_{-2} \cup B_{N-2}$$

Après une simple renumérotation, ces splines vont être indexées par des indices de 1 à N.

$$B_{-1},\ B_0,\ B_1,\ \ldots,\ B_{N-3},\ B_{N-2} \longrightarrow B_1,\ \ldots,\ B_{N-3},\ B_{N-2},\ B_{N-1},\ B_N$$

Dans les matrices des produits scalaires, grâce à la propriété 2.22 de nombreux produits scalaires sont nuls. La périodisation de la base se traduit par l'addition des produits scalaires non nuls dans les coins supérieur droit et inférieur gauche. Puisque toutes les matrices ont le même nombre d'éléments non nuls, on donnera la forme de la matrice $\mathbf{G}$ en exemple :

$$\mathbf{G} = \begin{pmatrix} g_{11} & g_{12} & g_{13} & 0 & \cdots & 0 & g_{13} & g_{12} \\ g_{12} & g_{11} & g_{12} & g_{13} & 0 & \cdots & 0 & g_{13} \\ g_{13} & g_{12} & g_{11} & g_{12} & g_{13} & 0 & \cdots & 0 \\ & \ddots & \ddots & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \ddots & \ddots & \\ 0 & \cdots & 0 & g_{13} & g_{12} & g_{11} & g_{12} & g_{13} \\ g_{13} & 0 & \cdots & 0 & g_{13} & g_{12} & g_{11} & g_{12} \\ g_{12} & g_{13} & 0 & \cdots & 0 & g_{13} & g_{12} & g_{11} \end{pmatrix} \tag{2.24}$$

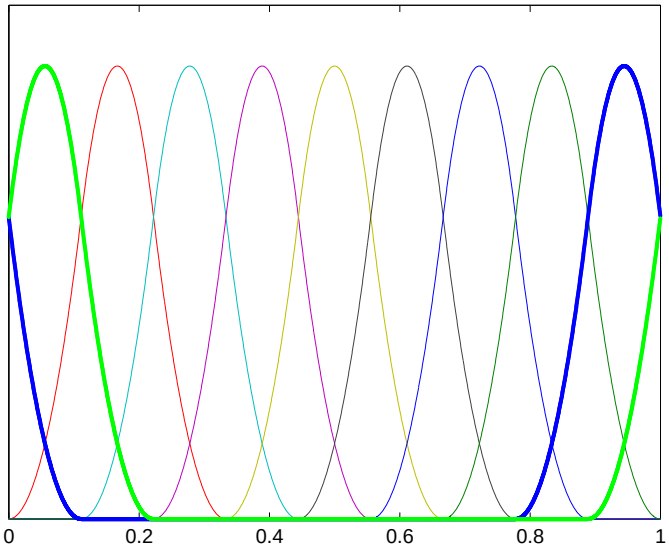


FIGURE 2.3 – *Exemple de périodisation pour les fonctions B-spline de degré 2. Les splines dessinées en gras coupent l'intervalle étudié. Pour assurer la périodicité elles sont reliées :* $\mathbf{B_{-2}}$ *avec* $\mathbf{B_{N-2}}$ *et* $\mathbf{B_{-1}}$ *avec* $\mathbf{B_{N-1}}$.

Remarquons que la symétrie des fonctions induit une symétrie dans les matrices des produits scalaires $\mathbf{G}$ et $\mathbf{D}$. Cette symétrie n'est pas conservée pour les produits scalaires qui prennent en compte le saut de la permittivité dans la couche inhomogène, les matrices $\mathbf{G}^{\varepsilon}$ et $\mathbf{G}^{\frac{1}{\varepsilon}}$ des équations 2.15 et 2.16.

Dans la suite, pour alléger la notation, on renonce au terme p qui définit la valeur du degré. Nous la rappellerons au besoin. Les conditions de périodicité simples ne sont valides que dans le cas particulier de l'incidence normale. Pour une incidence quelconque le champ électromagnétique est pseudopériodique. Pour prendre en compte la pseudo périodicité du champ, on doit simplement multiplier les fonctions de base et de test par le facteur de pseudo périodicité $exp(-ik\alpha_0 x)$:

$$\widetilde{B}_n(x) = e^{-ik\alpha_0 x} B_n(x) \qquad \text{avec } n \in \{1, 2, \ldots, N\}$$

$$\widetilde{T}_q(x) = e^{-ik\alpha_0 x} T_q(x) \qquad \text{avec } q \in \{1, 2, \ldots, N\}$$

où le symbole $\sim$ indique la prise en compte de la pseudopériodicité.

$$< \widetilde{T}_q | \widetilde{B}_n >= \int \widetilde{T}_q^* \widetilde{B}_n dx = \int e^{ik\alpha_0 x} T_q^* e^{-ik\alpha_0 x} B_n dx =< T_q | B_n > . \tag{2.25}$$

Le court calcul 2.26 montre que les matrices $\mathbf{G}$ qui contient les produits scalaires $< T_q|B_n >$ restent inchangées par la pseudopériodicité.

Concernant les produits scalaires contenant les dérivées $\frac{d}{dx}$, le changement à faire est :

$$\begin{aligned} &< \widetilde{T}_q|\frac{d}{dx}\widetilde{B}_q >= \int e^{ik\alpha_0 x}T_q^* \left[(-ik\alpha_0)e^{-ik\alpha_0 x}B_n + e^{-ik\alpha_0 x}\frac{d}{dx}B_n\right] dx = \\ &=< T_q|\frac{d}{dx}B_n > -ik\alpha_0 < T_q|B_n > . \end{aligned} \tag{2.26}$$

En terme de matrices, ce changement conduit à exprimer $\widetilde{\mathbf{D}}$ de la manière suivante :

$$\widetilde{\mathbf{D}} = \mathbf{D} - ik\alpha_0\mathbf{G}$$

2.3 Résultats numériques

Le problème à étudier est celui de la diffraction d'une onde plane par le réseau présenté à la figure 1.4. Nous allons étudier la convergence de certaines valeurs propres et de certains ordres de diffraction en fonction du nombre de fonctions de base introduites dans le calcul. Soit N le paramètre qui correspond à la taille des matrices. L'étude est faite en choisissant une palette variée de cas pour illustrer l'influence des différents paramètres numériques et physiques ainsi que faire une comparaison avec d'autres méthodes modales. Afin de comparer l'influence de ces paramètres sur la méthode utilisée, on a étudié la vitesse de convergence des valeurs propres, des efficacités ainsi que l'allure du champ proche à la surface du réseau. Pour l'étude des valeurs propres et des efficacités la convergence est mesurée, pour un nombre de troncature donné, à l'aide d'une fonction d'erreur définie comme suit :

$$err_{val}(N) = |\frac{val^{(N)} - val}{val}|. \tag{2.27}$$

où val est la valeur de référence. Dans un problème numérique il est impossible de trouver une solution "exacte". En conséquence nous estimons que la Méthode Modale décrite par Botten et al. [8, 24, 25] est celle qui présente le moins d'artefacts numériques puisqu'elle consiste à trouver les zéros d'une équation transcendante. C'est pour cette raison qu'on a choisi cette méthode comme méthode de référence pour le calcul des valeurs propres. En ce qui concerne les efficacités, trouver la solution "exacte" s'avère encore plus délicat puisque le calcul des efficacités est une conjonction de plusieurs résultats intermédiaires (valeurs propres, vecteurs propres, chaînage des matrices S) tous entachés d'une plus ou moins grande erreur numérique. Dans ce

cas, pour trouver des références on a procédé de deux manières : pour le réseau métallique d'indices $\nu = 0.22 - 6.71i$ abondamment traité dans la littérature nous avons utilisé comme valeurs de référence les valeurs considérés "exactes" par [79]. Pour les autres exemples, qui n'étaient pas traité dans la littérature nous avons utilisé comme référence, non plus une méthode unique, mais le résultat donné par une combinaison des méthodes les plus utilisées dans la littérature : MMFE [12], MMFE avec adaptation spatiale [35] et FDMM [11]. Plus précisément, à partir de l'analyse de la convergence de ces méthodes, nous avons choisi comme référence la valeur qui possède le plus grand nombre de chiffres significatifs en commun. Cependant aucune validation expérimentale ne permet de mesurer une efficacité à plus de 2 voire 3 chiffres significatifs. Aller plus loin n'a qu'un intérêt "académique" lorsqu'on veut comparer et valider les méthodes numériques entre elles. C'est pourquoi nous avons choisi de nous limiter aux 5 premiers ciffres après la virgule.

2.3.1 Influence des fonctions test

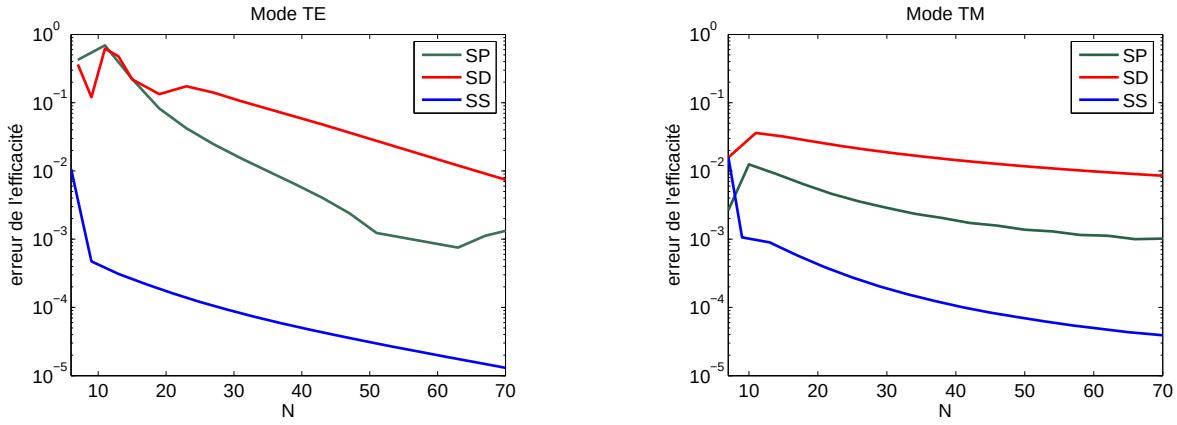
Nous essayons de voir l'impact des fonctions test sur la convergence des résultats. Nous choisissons pour ce faire comme fonction de base des splines quadratiques ; les résultats obtenus pouvant être généralisés aux cas des splines de degré quelconque. Dans ce cas-ci trois types de fonctions test ont été utilisés :

- une distribution de Dirac, dans ce cas nous parlerons de méthode de collocation par point ou *point-matching method* ou dans notre terminologie *spline-Dirac* (SD)
- des fonctions portes, nous conduisant à la méthode de collocation par sous domaines, dans la terminologie locale nous désignerons par *spline-porte* (SP) ce type de méthode des moments.
- des fonctions B-splines quadratiques : méthode de Galerkin ou *spline-spline* (SS).

Pour les illustrations nous prenons un cas type de structure utilisée en lithographie en UV profond à 193 nm. Il s'agit d'un réseau diélectrique sur un substrat de silicium. La configuration simulée est représentative d'une étape de métrologie dite scatterométie, réalisée à l'aide d'un ellipsomètre. Ce type d'instrument optique permet de mesurer le changement de polarisation induit par la réflexion sur un objet (ici le réseau). La résolution du problème inverse de diffraction permet de remonter aux caractéristiques géométriques de l'objet diffractant. C'est une technique de nano métrologie dimensionnelle de plus en plus utilisée dans l'industrie microélectronique. La résolution du problème inverse se ramène ici à une optimisation paramétrique qui nécessite la solution rigoureuse des équations de Maxwell pour calculer le changement de polarisation.

On se place dans l'air, $\varepsilon_1 = \varepsilon_{21} = 1$. On regarde l'efficacité de diffraction pour les deux cas de polarisation avec $\lambda = 193$ nm. La période du réseau est $d = 130$ nm et sa hauteur $h = 120$ nm. Le masque de résine a un indice de réfraction de $\nu_{22} = 1,6 - 0,02i$ et un facteur de remplissage $f = 0.5$. A cette longueur d'onde l'indice de réfraction du silicium est $\nu_3 = 0.885 - 2.37i$. L'ellipsomètre est incliné à $\theta = 72.32^\circ$ par rapport à la normale.

Pour ce réseau, il existe deux ordres diffractés : l'ordre spéculaire 0 et l'ordre -1 ; on a établi numériquement que l'efficacité R_0 diffractée vaut : 0.73971 en polarisation TE et 0.23761 en polarisation TM. Nous avons représenté sur les figures 2.4(a) et 2.4(b) les courbes de convergence $R_0(N)$, pour les trois méthodes de calcul cité plus haut. De façon générale, pour les deux cas de polarisation, et comme nous devrions nous y attendre, la méthode de Galerkin offre la meilleure vitesse de convergence. En polarisation TE, pour un ordre de troncature $N = 41$, la méthode de collocation par points ou point matching n'offre qu'une faible précision inférieure à 10^{-1}, tandis que la méthode de collocation par sous domaines fournit une précision de 10^{-2}. A cet ordre de troncature ($N = 41$), la meilleure précision est offerte par la méthode de Galerkin soit 10^{-4}. Dans le cas de la polarisation TM, on regarde les performances des méthodes toujours à la troncature $N = 41$. La méthode de collocation par points SD fournit une précision sur le premier chiffre significatif soit 10^{-1}, 10^{-2} pour la méthode de collocation par sous domaines SP et enfin une meilleure précision à savoir 10^{-3} est obtenue par la méthode de Galerkin SS. Rappelons que les valeurs qu'on à considéré "exactes" sont à la précision de la 5^{me} décimale, on peut donc considèrer que SS est arrivée à la valeur exacte dans les deux cas de polarisation.



(a) Polarisation TE *(b) Polarisation TM*

FIGURE 2.4 – *Erreur de l'efficacité de diffraction pour un réseau : $\theta = 72.32^\circ$, $\lambda = 193$ nm, $d = 130$ nm, $h = 120$ nm, $f = 0.5$, $\varepsilon_1 = \varepsilon_{21} = 1$, $\varepsilon_{22} = (1,6 - 0,02i)^2$, $\varepsilon_3 = (0.885 - 2.37i)^2$.*

Les figures 2.5 et 2.6 montrent respectivement le module du champ proche *i.e.* les composantes tangentielles du champ électromagnétique normalisées par rapport à leur valeur maximale, sur le réseau ($y = h$) en polarisation TE respectivement en polarisation TM. Sur chaque figure, sont représentés, pour les deux cas de polarisation, les résultats des calculs obtenus par la méthode de collocation par point SD fig.2.5(a), la méthode de collocation par sous domaines SP fig. 2.5(b) et la méthode de Galerkin SS fig. 2.5(c) pour deux valeurs de l'ordre de troncature $N = 41$ et $N = 81$.

Les meilleurs résultats sont obtenus par la méthode de Galerkin. Les oscillations parasites qu'on observe, dues à la représentation du champ par des B-splines sont très apparentes dans le cas de la SD, sont moins prononcées dans le cas de la méthode de collocation par sous domaines SP et encore moins dans le cas de la méthode SS. Pour expliquer cette différence de comportement, nous nous référons à la figure 2.7 sur laquelle les différentes fonctions test sont superposées à la fonction de base (ici les splines quadratiques).

Rappelons que la projection sur les fonctions test consiste à calculer une valeur moyenne pondérée sur un support qui est l'intersection des supports des fonctions de base et des fonctions test. Dans le cas de la méthode de collocation, aucune moyenne n'est effectuée pour le calcul du champ, les fonctions test étant des distributions de Dirac. Dans le cas de la méthode de collocation par sous domanines SP, le calcul du champ consiste en une moyenne sur l'intervalle $[x_k, x_{k+1}]$ ce qui réduit les oscillations. Un calcul effectué par la méthode SS élargit le support d'intégration au support de la spline interpolatrice $[x_k, x_{k+3}]$ réduisant considérablement les oscillations.

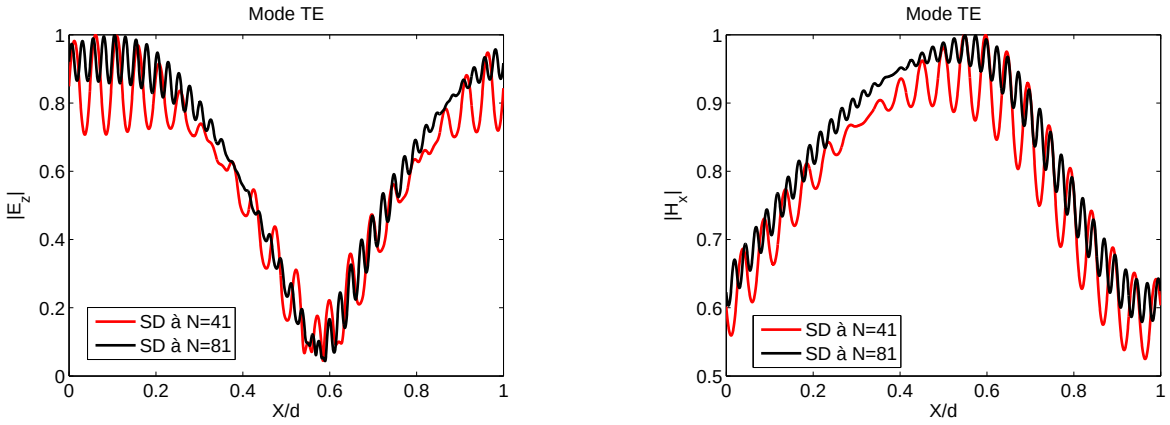
L'étude de la convergence du champ proche en fonction de la troncature est présentée sur les figures 2.8. Le critère de convergence retenu est défini comme :

$$S_N = \sum_{x_k} |F_{N+1}(x_k, h) - F_N(x_k, h)|^2, \tag{2.28}$$

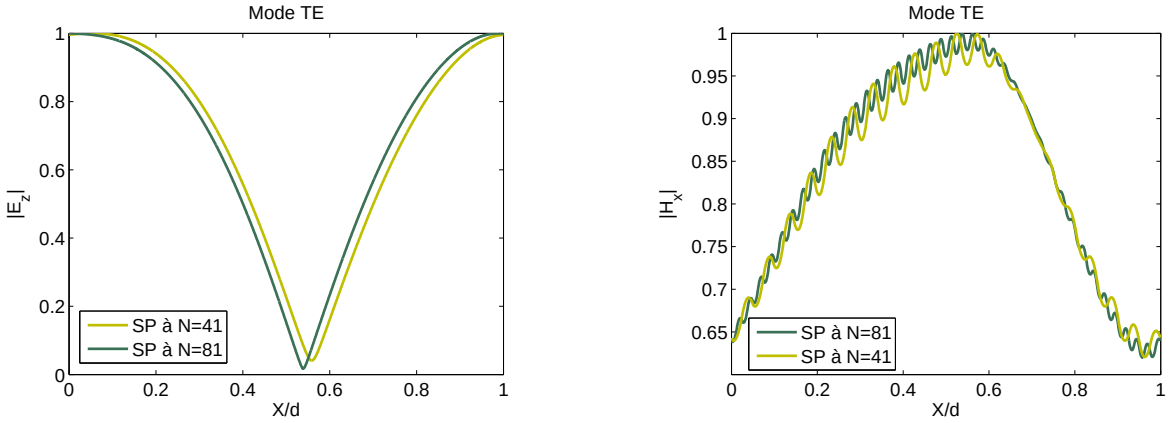
où N est l'ordre de troncature et F désigne une composante tangentielle quelconque du champ électromagnétique pour les deux types de polarisation.

Sur ces figures, on peut observer que les pentes des courbes de convergence en erreur quadratique obtenues avec les SS sont beaucoup plus grande par rapport aux SP ou SD. L'ordre des pentes de convergence du champ varie comme celle des efficacités. Les résultats sont conformes à ceux montrés sur les figures 2.4(a) et 2.4(b). Même si la convergence du champ proche dans le cas de polarisation TM, n'est pas illustré ici, on mentionne que l'ordre des performances est aussi le même. Cependant l'étude de la convergence du champ proche telle que

(a) SD première et deuxième composantes du champ



(b) SP première et deuxième composantes du champ



(c) SS première et deuxième composantes du champ

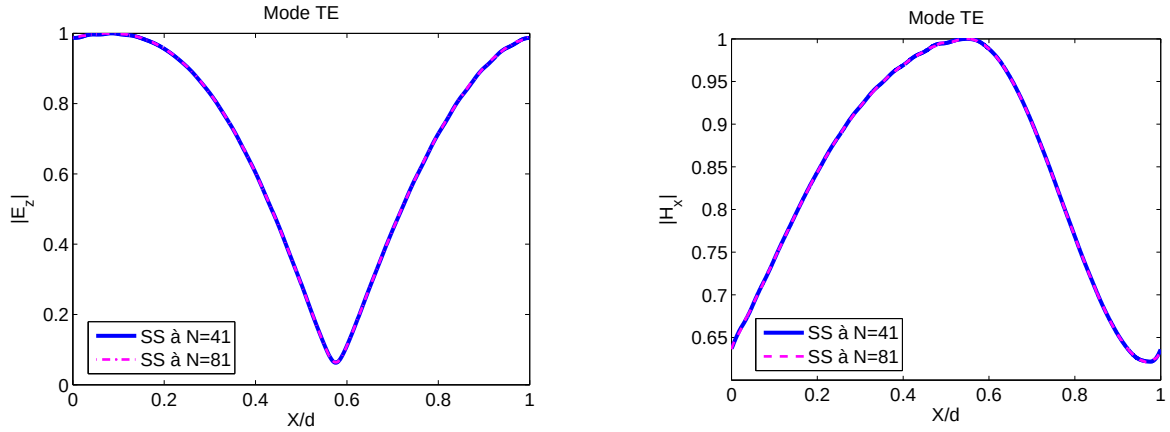
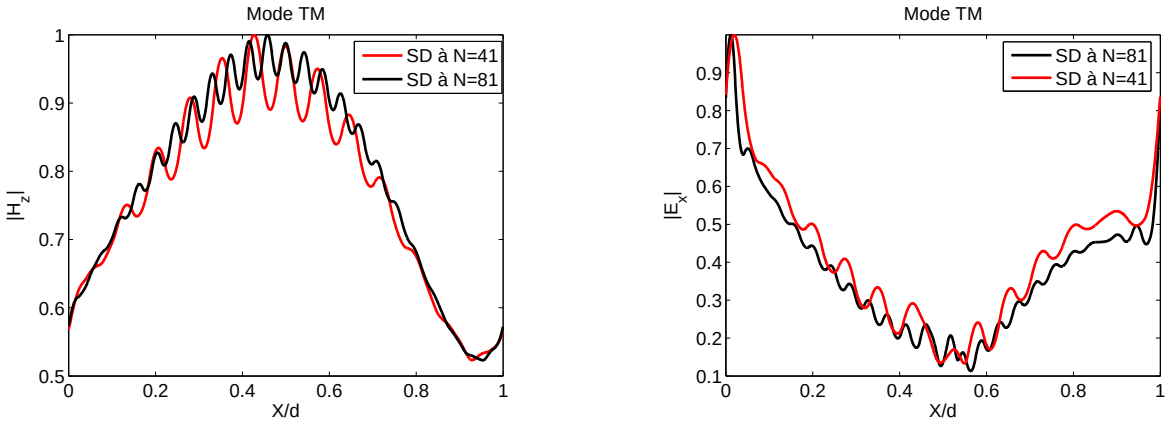
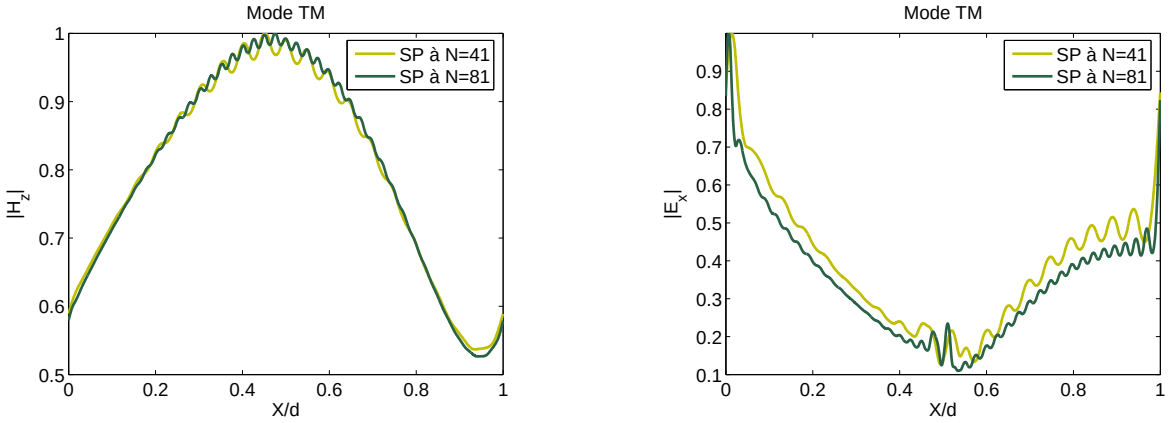


FIGURE 2.5 – *Polarisation TE : les champs E_z et H_x pour un réseau avec les mêmes paramètres que ceux de la figure 2.4*

(a) SD première et deuxième composantes du champ



(b) SP première et deuxième composantes du champ



(c) SS première et deuxième composantes du champ

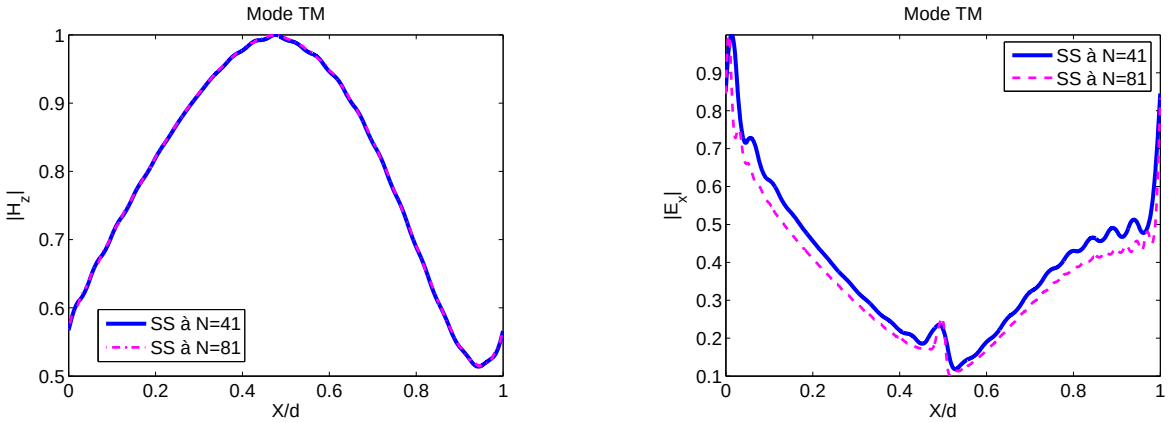


FIGURE 2.6 – *Polarisation TM : les champs H_z et E_x pour un réseau avec les mêmes paramètres que ceux de la figure 2.4*

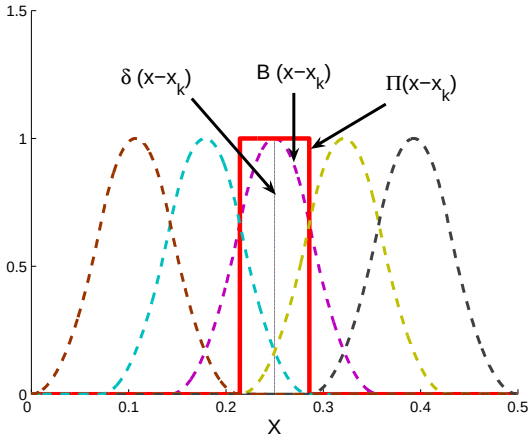
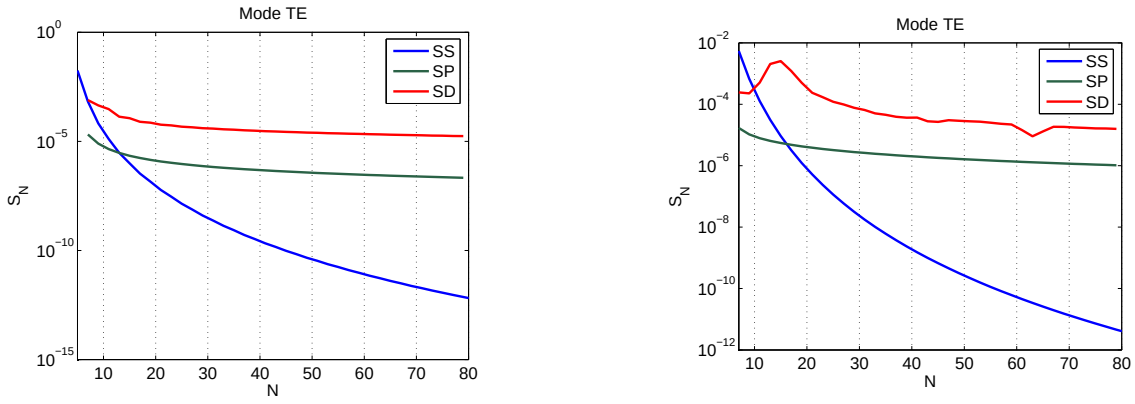


FIGURE 2.7 – *Description des fonctions de base et de test*

présentée par l'équation 2.28 est équivalente à l'étude de la convergence d'une suite de Cauchy alors que la fonction d'erreur de l'équation 2.27 introduit une convergence uniforme. De ce fait il n'est pas judicieux de comparer la vitesse de convergence du champ proche avec celle des efficacités même si les deux grandeurs (champ proches et efficacités) convergent en fonction de N.



(a) composante E_z du champ

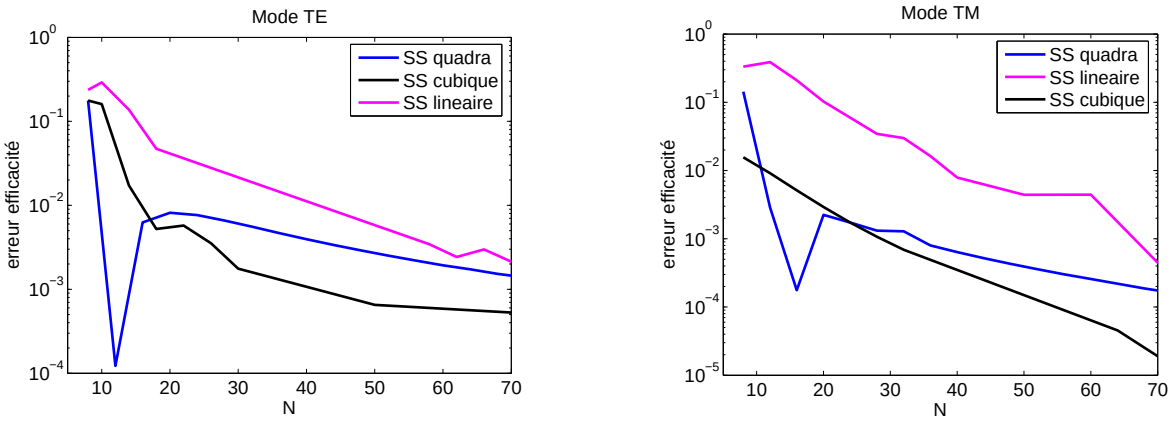
(b) composante H_x du champ

FIGURE 2.8 – *Convergence en erreur quadratique pour les champ E_z et H_x pour un réseau avec les mêmes paramètres que ceux de la figure 2.4*

2.3.2 Influence des fonctions de base

Pour illustrer l'influence des fonctions de base sur la vitesse de convergence, nous avons testé sur le même réseaux des splines de degré de 1 à 3, les fontions test restant inchangées. Nous avons représenté sur les figures 2.9 (a) et 2.9 (b) des courbes comparatives de convergence de l'efficacité diffractée dans l'ordre spéculaire par un réseau diélectrique d'indice de réfraction $\nu_{22} = 2.35$ éclairé sous une incidence $\theta = 15°$. Les autres paramètres numériques sont : $f = 0.5$, $h = d = \lambda = 1$. Pour ce cas, les valeurs de référence sont : 0.10872 en TE et 0.06890 en TM. Ces calculs sont effectués par la méthode Galerkin avec des splines de degré 1 (courbe en magenta), de degré 2 (courbe en bleu) et de degré 3 (courbe en noir). Dans le cas d'une polarisation TM, une précision de 10^{-4} est atteinte avec un ordre de troncature $N = 55$ pour les splines de degré 3 alors que une approximation du champ par des spline de degré 2 ne fournie qu'une précision de 10^{-3} et 10^{-2} dans le cas des splines de degré 1. Dans le cas TE, pour le même ordre de troncature $N = 55$, les splines de degré 1 et 2 fournissent indifféremment une précision de 10^{-2} alors que l'approximation au degré 3 permet d'augmenter la précision d'un chiffre significatif.

Une interprétation possible des performances respectives des trois méthodes est la meilleure prise en compte des sauts de la permittivité par les fonctions splines de degré plus élevé. En effet, augmenter le degré des splines élargit leur support et la discontinuité est intercepté par plusieurs fonctions spline.



(a) Polarisation TE

(b) Polarisation TM

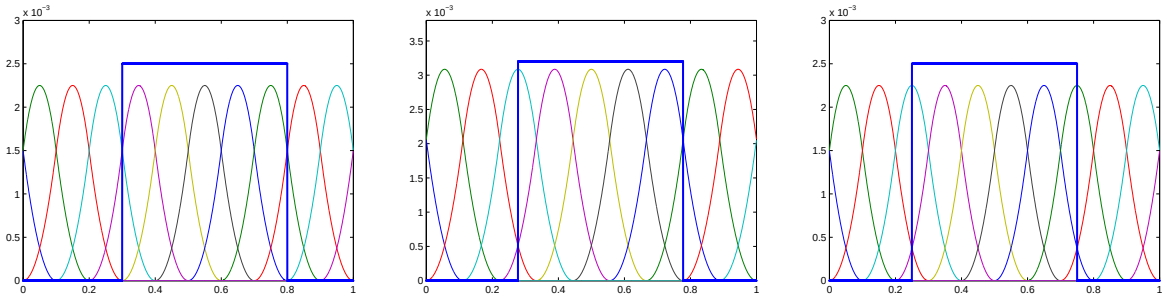
FIGURE 2.9 – *Erreur de l'efficacité de diffraction en utilisant Galerkin avec des splines de degré* 1 *(courbe en magenta), de degré* 2 *(courbe en bleu) et de degré* 3 *(la courbe en noire). Les paramètres du réseau sont :* $\theta = 15^{\circ}$, $\lambda/d = h/d = 1$, $f = 0.5$, $\varepsilon_1 = \varepsilon_{21} = 1$, $\varepsilon_{22} = \varepsilon_3 = 2.35^2$.

2.3.3 Influence de la grille de discrétisation

Dans le cas des bases nodales, les discontinuités de la fonction de permittivité sont "vues" par un certain nombre des fonctions de base. Ce nombre est déterminé par le degré de la fonction de base utilisée. Par exemple, dans le cas des splines de degré 1, chaque fonction chevauche sa voisine de droite et de gauche. Le nombre maximal des fonctions qui se chevauchent dans un point fixé sur l'axe $0x$ est égal à deux et le nombre minimal est 1. Suivant l'endroit où le saut de la permittivité se situe par rapport à ces chevauchements des fonctions de base, la vitesse de convergence peut varier. Plus précisément, on obtient des résultats d'une précision plus élevée lorsque le saut de la fonction de permittivité est "vu" par le maximum des fonctions de base. Des variations similaires ont aussi été remarquées en utilisant la méthode MMFE [12], mais la différence entre les vitesses de convergence sont moins importantes.

Afin d'illustrer ce phénomène, on a choisi d'utiliser des fonctions spline quadratiques comme fonctions de base et de test. On définit trois grilles de discrétisation : la première, qu'on appelera $DS1$ est illustrée à la figure 2.3.3(a). Elle correspnd au cas où les discontinuités de la permittivité (dessinée en bleu sur les figures 2.3.3) chevauchent un nombre minimum de fonctions de base ; dans le cas des splines quadratiques il y en a seulement deux. La deuxième grille de discrétisation, qu'on nommera $DS2$, est illustrée dans la figure 2.3.3(b). Elle correspond au cas où une discontinuité de la permittivité est moyennée sur deux fonctions de base et l'autre discontinuité sur trois ; et enfin avec la troisième grille de discrétisation (notée $DS3$) les deux sauts de la

permittivité sont à l'intersection de trois splines 2.3.3(c). Le but de ce paragraphe est de montrer

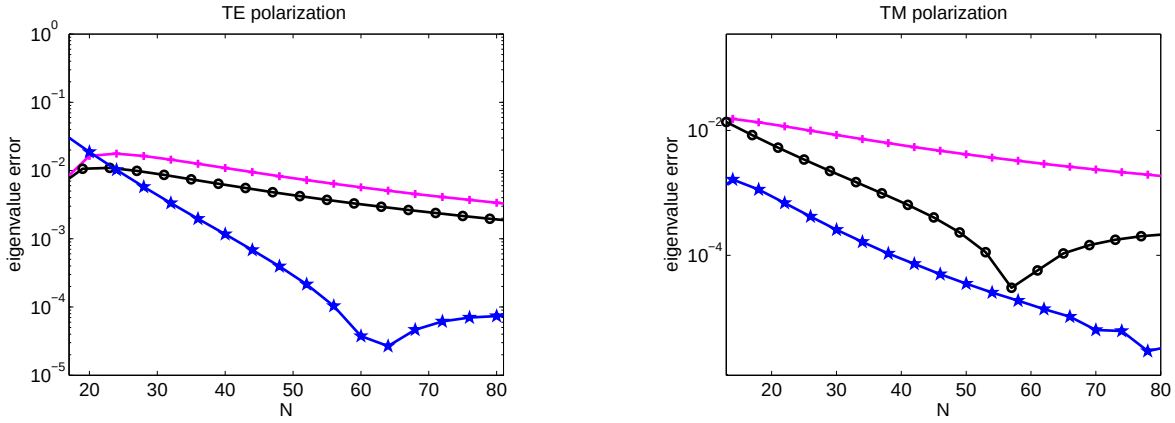


(a) DS1 *(b) DS2* *(c) DS3*

FIGURE 2.10 – *Différents schéma de discrétisation. Les fonctions de base sont les splines quadratiques. La permittivité de la couche inhomogène est dessinée en bleu. DS1-les deux discontinuités de la permittivité sont à l'intersection de deux splines ; DS2-une des discontinuités de la permittivité est à l'intersection de deux splines et l'autre est à l'intersection des 3 splines ; DS3 les deux discontinuités de la permittivité sont à l'intersection de trois splines*

l'influence de ces grilles de discrétisation en utilisant des pas équidistants. L'exemple considéré utilise un facteur de remplissage d'une moitié de la période. Dans ce cas, on peut étudier la convergence de chaque grille de discrétisation sans avoir recours à des techniques d'adaptation spatiale où à des grilles non uniformes. Il suffit d'imposer que la première discontinuité soit située à l'intersection des trois splines et on obtient automatiquement le schéma DS3 dans le cas des troncatures paires et le schéma DS2 dans le cas des troncatures impaires. Le schéma DS1 est obtenu avec des troncatures paires quand on impose que le premier saut de discontinuité de la permittivité soit à l'intersection des deux splines. Pour des structures d'un facteur de remplissage quelconque on peut respecter ces schémas de discrétisation en utilisant une fonction qui resserre les points de telle manière que la meilleure discrétisation soit respectée. Cette fonction a déjà été détaillée dans [70].

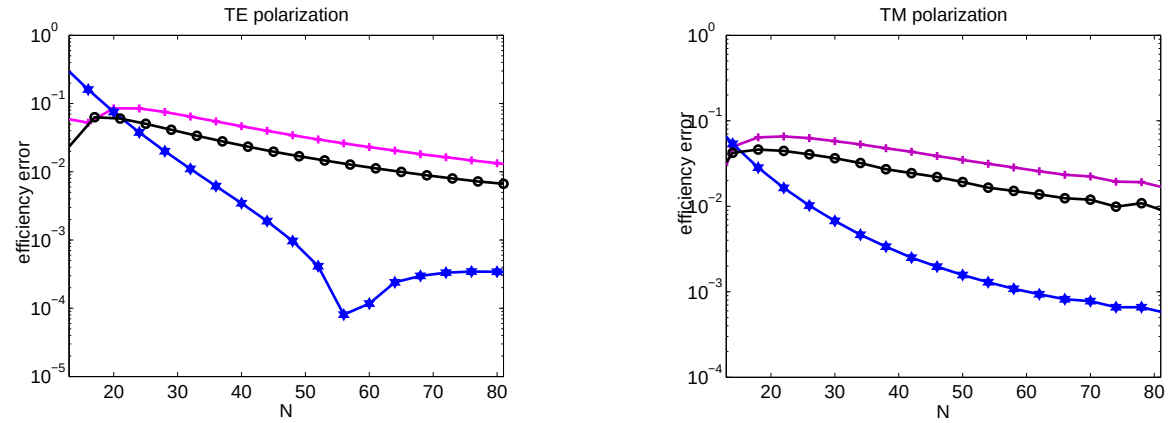
Pour illustrer ce type de comportement on a choisi un réseau dont les paramètres sont : $d/\lambda = 1$, $h/\lambda = 1$, $f = 0.5$, $\varepsilon_1 = \varepsilon_{21} = 1$, $\varepsilon_{22} = \varepsilon_3 = (0.22 - 6.71i)^2$ et $\theta = 30^\circ$. Dans les figures 2.11(a) et 2.11(b) on peut voir le comportement des erreurs des valeurs propres à l'échelle logarithmique en fonction de la taille des matrices. Le courbes tracées avec des "+" sont les résultats obtenus avec la grille de discrétisation DS1 ; les "o" sont pour la grille DS2 et les "*" sont pour la grille DS3. On peut observer qu'une importante amélioration est obtenue en utilisant la grille de discrétisation DS3. L'erreur décroît rapidement, elle est significativement inférieure, même pour des petit ordres de troncature.



(a) Polarisation TE *(b) Polarisation TM*

FIGURE 2.11 – *Les erreurs d'une valeur propre de la couche inhomogène. Les paramètres du réseau sont les mêmes que ceux de la figure 2.15.*

Pour cet exemple, on a aussi comparé les performances en calculant les efficacités. On prend en considération l'ordre -1 de diffraction. Sur cet ordre les valeurs de référence sont 0.73428 en polarisation TE et 0.84848 en polarisation TM. Les graphiques 2.12(a) et 2.12(b) montrent l'évolution de l'erreur sur l'efficacité en échelle logarithmique en fonction du nombre de troncature. On peut observer une fois de plus que les résultats fournis en utilisant la grille DS3 sont les meilleurs. Les résultats ci-dessous montrent clairement que la vitesse de convergence



(a) Polarisation TE *(b) Polarisation TM*

FIGURE 2.12 – *Les erreurs de l'efficacité reflechie sur l'ordre -1. Les paramètres du réseau sont les mêmes que ceux de la figure 2.15.*

est influencée par la grille de discrétisation, plus précisément, par la position des discontinuités

de la permittivité par rapport à cette grille. Globalement, on observe que dans ces figures les meilleurs résultats sont obtenus en utilisant la grille DS3, suivies de la grille DS2 et puis DS1. L'ordre est respecté dans toutes les figures.

L'amélioration est due au fait que les sauts de la fonction de permittivité coupent un maximum de fonctions splines non nulles. En utilisant le schéma DS3 on moyenne donc la fonction de permittivité sur le maximum de sous intervalles, c'est à dire trois. En utilisant la grille DS2, l'une des discontinuités est moyennée par le maximum des splines (dans ce cas trois) et l'autre seulement par deux. Les résultats utilisant la DS2 sont moins bonnes par rapport à DS3 mais supérieures aux performances de DS1 qui moyenne chaque discontinuité sur le minimum de sous intervalles de discrétisation. On peut en déduire qu'une façon d'accélérer la vitesse de convergence (au moins pour les B-splines de degré 1 et 2) est de prendre en compte le saut de la permittivité de telle manière que la discontinuité soit moyennée sur le maximum de sous intervalles.

2.3.4 Comparaison avec d'autres méthodes

Nous comparons dans ce paragraphe nos résultats avec ceux de la littérature en l'occurrence avec la Méthode Modale de Fourier [12] et la Méthode Modale des différences finies dans le domaine fréquentiel [11]. Nous nous sommes intéressés à quatre types des réseaux :

- un réseau diélectrique à fort contraste d'indices étudié par Granet dans la référence [35].
- un réseau métallique d'indices $\nu = 0.22 - 6.71i$ étudié pour la première fois par Li [80] et repris dans la littérature par plusieurs auteurs [11, 35, 72, 73, 79].
- un réseau métallique d'indice de réfraction purement imaginaire $\nu = -6.71i$.
- un réseau utilisé dans la pratique au laboratoire LTM.

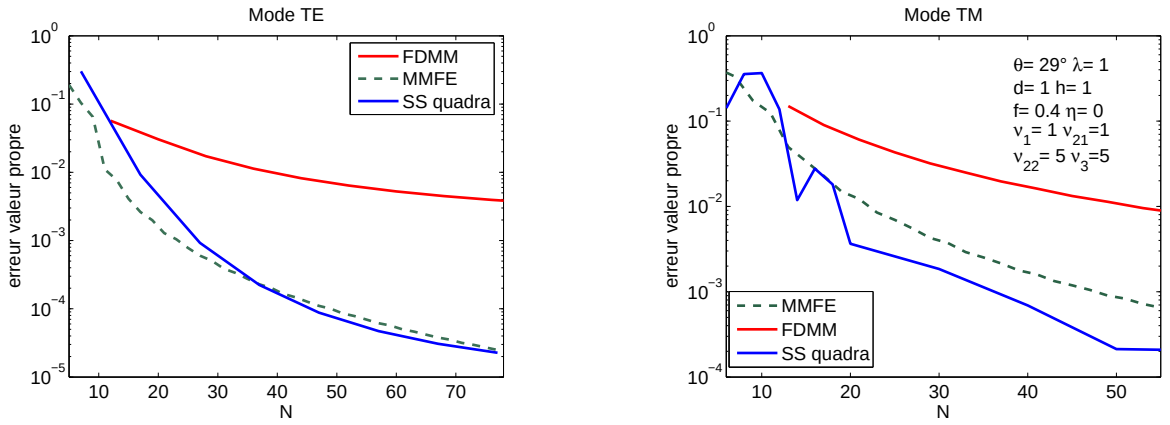
Pour ce faire, nous allons comparer dans un premier temps certaines valeurs propres des matrices (2.12a) et (2.12b) avec les valeurs de référence. Dans un second temps, les valeurs des efficacités réfléchies pour certains ordres seront confrontées aux résultats des références citées plus haut.

2.3.4.1 Réseau diélectrique

Le premier exemple pris en compte traite un réseau diélectrique avec un très fort contraste d'indice. Un contraste d'indice important permet de tester les performances d'une méthode, sachant qu'on veut obtenir une précision donnée avec une matrice de taille minimale. Les pa-

ramètres du réseau sont : $d/\lambda = 1$, $\varepsilon_{21} = 1$, $\varepsilon_3 = \varepsilon_{22} = 25$. Ce réseau est éclairé par une onde plane en polarisation TE ou TM sous un angle d'incidence $\theta = 29°$. Pour cet exemple, un calcul par la méthode modale [8] de la deuxième valeur propre réelle (les valeurs propres étant classées par l'ordre croissant) donne $r = 3.35101975722312$ pour un cas de polarisation TE et $r = 2.81329903403930$ pour un cas de polarisation TM.

Nous avons comparé sur les figures 2.13(a) et 2.13(b) la fonction d'erreur décrite par l'équation 2.27 pour $val = r$, en fonction de N, dans le cas TE resp TM obtenue avec notre approche (courbe en bleu) avec celles obtenues par des méthodes considérées comme des références dans le domaine : la méthode modale de Fourier (MMFE) [12, 35] (courbe en pointillé) et une méthode modale basée sur un calcul de différences finies dans le domaine fréquentiel (FDMM) [11] (courbe en rouge). A l'aide de notre approche nous obtenons une bonne adéquation avec les valeurs exactes. Dans les deux cas de polarisation, une précision de 10^{-3} est atteinte pour un ordre de troncature avoisinant $N = 50$. En comparant les trois courbes d'erreur, on observe que l'approche Galerkin par des splines quadratiques se situe entre la MMFE et la FDMM en termes de convergence dans le cas de la polarisation TE. Par contre en polarisation TM, les résultats obtenus par l'approche spline quadratique supplantent les deux autres.



(a) Polarisation TE *(b) Polarisation TM*

FIGURE 2.13 – *Erreur sur une valeur propre de référence dans la couche inhomogène d'un réseau diélectrique avec les paramètres : $\theta = 29°$, $d/\lambda = 1$, $\varepsilon_{21} = 1$, $\varepsilon_{22} = 25$. La courbe en pointillé est obtenue avec la MMFE [35], la courbe bleu est obtenue en utilisant la Méthode de Galerkin avec des B-splines quadratiques et la courbe en rouge correspond FDMM [11].*

2.3.4.2 Réseaux métalliques

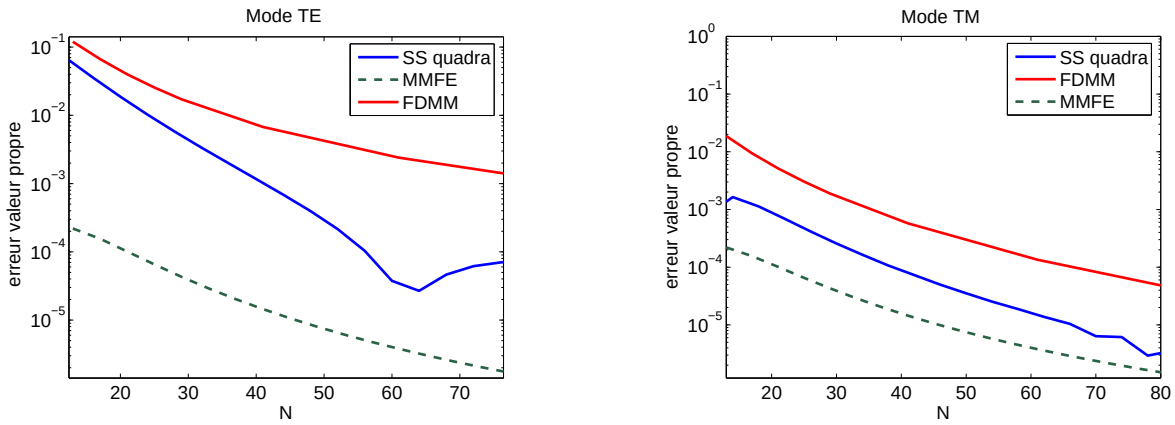
Le déuxième exemple numérique est celui du réseau traité dans [11, 35, 72, 73, 79]. Dans la littérature, cet exemple a servi à tester la robustesse de la méthode de Fourier, la FDMM, etc. Les paramètres du réseau sont : $d/\lambda = 1$, $h/\lambda = 1$, $f = 0.5$, $\varepsilon_1 = \varepsilon_{21} = 1$ et $\varepsilon_{22} = \varepsilon_3 = (0.22 - 6.71i)^2$. Le réseau est éclairé par une onde plane sous en angle d'incidence $\theta = 30^\circ$. Ce cas est présenté dans les références [80, 11, 35, 72, 73]. Comme dans le cas diélectrique, nous comparons dans la couche inhomogène, les erreurs sur la deuxième valeur propre par rapport à la méthode modale de Botten [8]. La méthode de Galerkin avec des B-splines quadratiques est comparée à la FDMM et à la MMFE. La valeur de référence obtenue par la méthode modale [8] est :

$$r = 0.40565997728692 - 0.00570953767335i \quad \text{pour le cas TE}$$

$$r = 1.05070585861225 - 0.00180066465604i \quad \text{pour le cas TM}$$

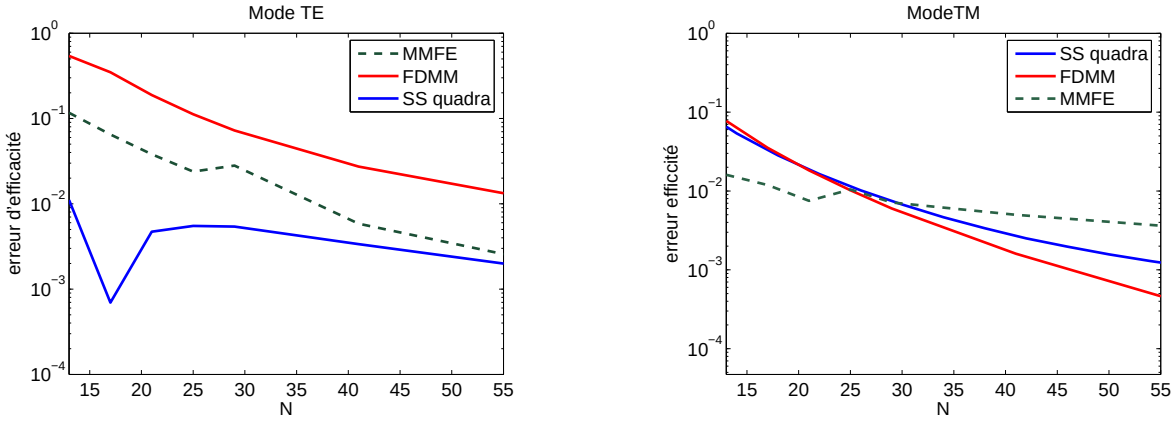
Les figures 2.14(a) et 2.14(b) relatent le comportement de la fonction d'erreur $err_r(N)$ pour les deux cas de polarisation. Pour ces deux cas les résultats sont concordants et une précision de 10^{-3} est atteinte pour une valeur N d'environ 50. Pour se situer par rapport à la MMFE et la FDMM [11], nous avons représenté sur les mêmes graphiques les erreurs de ces deux méthodes : la courbe en pointillé étant pour MMFE et la courbe en rouge pour la FDMM. Pour les deux cas de polarisation les résultats obtenus par la méthode de Galerkin par des splines quadratiques se situent entre ceux de de la MMFE et la FDMM.

Pour cet exemple, uniquement les ordres zero et -1 sont diffractés. Pour comparer la précision des résultats numériques sur les efficacités, on prend en considération l'ordre de diffraction -1. Sur cet ordre les valeurs prises comme référence en [79] sont 0.73428 en polarisation TE et 0.84848 en polarisation TM. Les graphiques 2.15(a) et 2.15(b) montrent l'évolution de l'erreur sur les efficacités en l'échelle logarithmique en fonction du nombre de troncature. Comme on peut le remarquer sur ces graphiques, notre méthode fournit des bons résultats même pour des troncatures assez faibles. En terme de performances, en polarisation TE on obtient des résultats meilleurs que les méthodes de référence. En polarisation TM, l'approche Galerkin avec des splines quadratiques a la même allure que celle fournie par la méthode FDMM. Les deux résultats supplantent les performances de la la MMFE pour les troncatures au delà de 25.



(a) Polarisation TE *(b) Polarisation TM*

FIGURE 2.14 – *Erreur sur une valeur propre de référence dans la couche inhomogène d'un réseau métallique avec les paramètres :* $\theta = 30^{\circ}$, $d = 1$, $\lambda = 1$, $h = 1$, $\varepsilon_{21} = 1$, $\varepsilon_3 = \varepsilon_{22} = (0.22 - i6.71)^2$. *La courbe en pointillé est obtenue MMFE [35], la courbe rouge est obtenue en utilisant les différences finies [11], la courbe bleu est obtenue en utilisant la Méthode de Galerkin avec des B-splines quadratiques.*



(a) Polarisation TE *(b) Polarisation TM*

FIGURE 2.15 – *Erreur de l'ordre -1 de diffraction pour un réseau métallique avec les paramètres :* $\theta = 30^{\circ}$, $d = 1$, $\lambda = 1$, $h = 1$, $\varepsilon_{21} = 1$, $\varepsilon_3 = \varepsilon_{22} = (0.22 - i6.71)^2$.

2.3.4.3 Réseaux avec des matériaux à indice négatif

Des travaux récents ont montré l'intérêt d'utiliser, dans le domaine du spectre visible ou même infrarouge, des réseaux avec des matériaux dont les indices de réfraction sont négatifs. On a testé notre méthode pour un des ces cas. C'est l'exemple d'un réseau (déjà étudié dans la littérature en [11]) avec un indice de réfraction purement imaginaire et négatif. Les paramètres de simulation sont $\theta = 30^{\circ}$, $d = 1$, $\lambda = 1$, $h = 1$, $\nu_{21} = 1$, $\nu_3 = \nu_{22} = -i6.71$. Les résultats sont illustrés dans la figure 2.16. On compare la Méthode modale de Fourier avec notre méthode en utilisant comme base de développement les B-splines quadratiques. Comme on peut l'observer

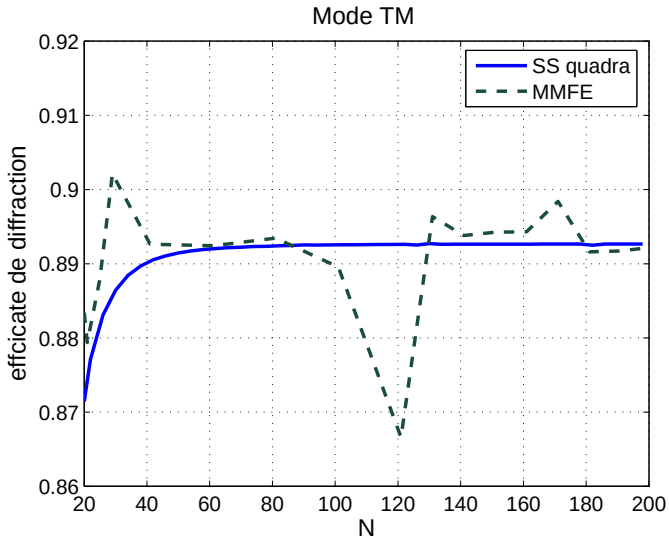


FIGURE 2.16 – *Convergence de l'ordre -1 de diffraction pour le cas de polarisation TM pour une structure métallique :* $\theta = 30^{\circ}$, $d = 1$, $\lambda = 1$, $h = 1$, $\varepsilon_{21} = 1$, $\varepsilon_3 = \varepsilon_{22} = (-i6.71)^2$.

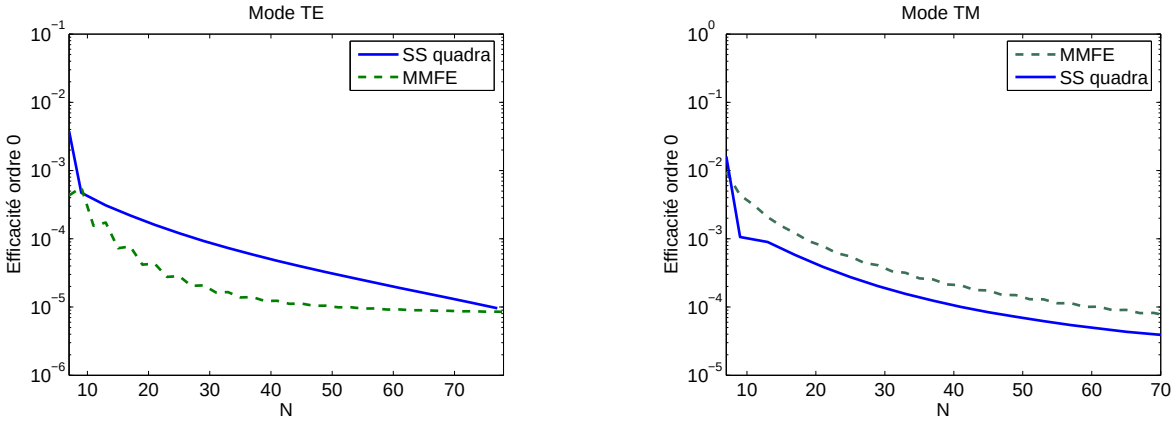
dans la figure 2.16 en utilisant les splines on obtient la convergence tandis que ce n'est pas le cas avec la méthode modale de Fourier. On peut expliquer que la MMFE ne converge pas dans ce cas par le fait que les règles de factorisation de Li [27] ne sont pas démontrées quand la permittivité ε passe par zéro. La convergence est également obtenue avec les autres bases subsectionnales (Dirac, portes, triangles, etc.). Il semble donc que l'utilisation des bases à support compact soit le seul choix possible dans ce type de cas, au moins si l'alternative est la MMFE. D'ailleurs, contrairement à tous les cas précédents, nous n'avons pu utiliser cette dernière comme méthode de référence pour disposer d'une valeur exacte.

Le fait que les bases subsectionales, se prêtent bien à l'analyse de ce type de cas, est un résultat très positif pour notre méthode dans la mesure où ce type de matériaux focalise l'attention

grâce à leurs nombreuses potentialités sur le plan scientifique et technologique (matériaux main gauche, cape d'invisibilité, etc.).

2.3.4.4 Réseaux résine sur silicium

Le dernier exemple comparatif est celui déjà traité dans 2.4. Dans les figures 2.17 on peut remarquer que la convergence de notre méthode et celle de MMFE sont comparables. Cet exemple



(a) Polarisation TE *(b) Polarisation TM*

FIGURE 2.17 – *Erreur de l'efficacité de diffraction pour un réseau :* $\theta = 72.32^{\circ}$, $\lambda = 193$, $d = 130$, $h = 120$, $f = 0.5$, $\varepsilon_1 = \varepsilon_{21} = 1$, $\varepsilon_{22} = (1,6 - 0,02i)^2$, $\varepsilon_3 = (0.885 - 2.37i)^2$.

illustre le fait que les fonctions splines se prêtent bien pour le traitement des réseaux aussi bien pour de cas déjà étudiés dans la littérature qui testent la robustesse de la méthode mais aussi pour des cas d'intérêt applicatif puisque le réseau étudié ici est un des cas typiques de réseaux à traiter dans notre laboratoire.

Pour conclure, les résultats obtenus en utilisant la méthode modale avec développement en fonctions splines sont fiables. Les courbes convergent vers les valeurs de référence. D'autres simulations non reportés dans ce paragraphe, montrent que d'une manière générale les résultats obtenus avec la méthode de Galerkin avec des B-splines quadratiques sont plus performantes que celles obtenues avec la FDMM de [11]. Cela est dû au fait que les différences finies peuvent être vues, en terme de méthode des moments, comme des fonctions portes, c'est à dire des fonctions B-spline de degré 0. Cependant, la manière dont Lalanne et al [11] ont implémenté la FDMM présente certaines différences par rapport à notre méthode : la FDMM utilise des grilles intercalées pour le calcul des différentes composantes du champ (E et H ne sont pas sur

la même grille).

Concernant les performances de nos résultats comparés ceux obtenus par la méthode MMFE, on obtient des résultats comparables, parfois plus performants, parfois moins performants. Il n'est pas possible d'affirmer qu'une des deux méthodes est meilleure, numériquement parlant, que l'autre. Rappelons que la principale différence entre les deux méthodes réside dans le fait que l'une utilise des bases à support compact et l'autre des bases définies dans tout le domaine.

2.3.5 Résolution spatiale adaptative

Il a été montré [81] que l'interpolation d'une fonction f par des splines peut conduire à des oscillations dans le cas où la fonction présente de fortes courbures. Dans le milieu inhomogène et en polarisation TM, le champ H est continu mais sa dérivée par rapport à la direction de périodicité est discontinue. On peut donc s'attendre à ce que sa décomposition sur une base de splines présente également des oscillations. Prenons comme exemple une structure utilisée en scatterométrie. Il s'agit de lignes de cuivres assez isolées déposées sur un substrat d'oxyde de silicium. Nous avons représenté sur la figure fig. 2.18 le module du champ magnétique $H_z(x, 0)$

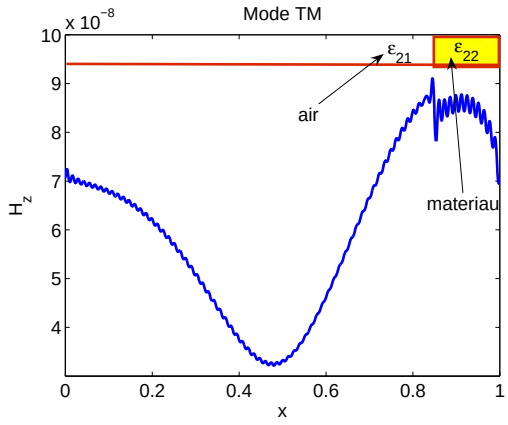


FIGURE 2.18 – *L'allure du champ H_z.*

avec les paramètres : $\varepsilon_{22} = (1,133 - i2,558)^2$, $\varepsilon_3 = 1,462^2$, $d = 350$ nm, $\lambda = 500$ nm, $h = 100$ nm, $f = d/7$, $N = 41$, $\varepsilon_1 = \varepsilon_{21} = 1$ et $\theta = 40^\circ$.

Comme on peut l'observer dans cette figure, le champ présente une forte variation en $x =$

$fd = 300$nm. Des oscillations parasites apparaissent sur le champ évalué à partir des splines :

$$H_z(x, y = 0) = \sum_{n=1}^{N} \sum_{m=1}^{N} a_m^{\pm} F_{nm} B_n(x) \tag{2.29}$$

Ces oscillations peuvent être gommées en introduisant une résolution spatiale adaptive. Dans le contexte de notre problème, cette méthode consiste à resserrer les points de discrétisation autour des variations rapides du champ. Cette approche à été déjà utilisé avec succès pour résoudre le problème de diffraction par des réseaux à l'aide d'une méthode modale par plusieurs chercheurs : [35], [11], [36], [70], [72]. Il existe plusieurs approches pour mettre en œuvre cette technique, tous ont en commun d'introduire une fonction $x(u)$ qui resserre les points autour des endroits de forte variation du champs électromagnétique. Ces variations se produisent autour des discontinuités de la fonction de permittivité $\varepsilon(x)$.

Le raffinement numérique introduit par cette nouvelle représentation de l'espace a pour conséquence une vitesse de convergence plus rapide. Du point du vue pratique, à précision identique, les valeurs propres et les efficacités sont obtenues avec des matrices de dimension plus faible avec la résolution spatiale adaptative.

Dans [70], nous avons déjà utilisé les splines de degré 1 et la résolution spatiale adaptive. Le point du vue adopté consistait à dire que $x = x(u)$, $y = v$, $z = w$ définissent un nouveau système des coordonnées (u, v, w) déduit du système cartésien (x, y, z) et à écrire les équations de Maxwell sous forme covariante.

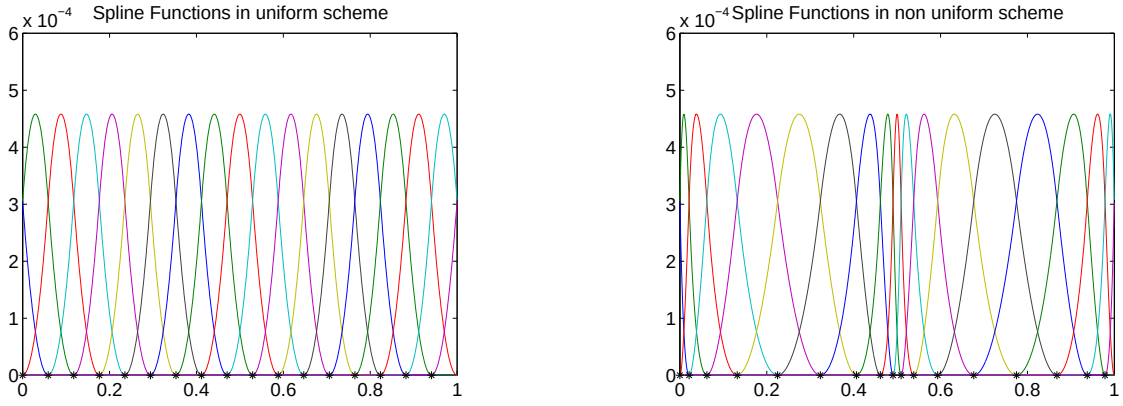
Pour un cas simple comme celui ci, on peut également considérer $x(u)$ comme un changement de variable qui permet de construire un ensemble des points x_k non équidistants à partir des points u_k équidistants. Ceci permet de rester en coordonnées cartesiennes. Il sufit de réécrire les composantes du champ électromagnétique en utilisant la dérivation des fonctions composées.

Considérons comme $x(u)$ la fonction déjà utilisée en [35] et [72]. Elle est définie comme suit :

$$x\,(u) = \begin{cases} u - \eta \frac{f}{2\pi} \sin\left(\frac{2\pi u}{fd}\right) & \text{pour} \quad 0 \leq u < fd \\ \\ u + \eta \frac{1-f}{2\pi} \sin\left(\frac{2\pi(d-u)}{(1-f)d}\right) & \text{pour} \quad fd \leq u \leq d \end{cases} \tag{2.30}$$

où η est un paramètre compris entre 0 et 1 qui permet de compresser les points autour des discontinuités. On peut observer que dans le cas où $\eta = 0$, la grille reste enchangée puisque

dans ce cas $x(u) = u$, donc le repère des coordonnées (u, v, w) coïncide avec (x, y, z). Le resserrage des points est maximal pour η plus proche de 1. La figure 2.19 montre l'impact de la grille avant 2.19(a) et après 2.19(b) la compression des points pour une valeur de $\eta = 0.75$.



(a) *(b)*

FIGURE 2.19 – *Illustration d'une grille uniforme à gauche et d'une grille non-uniforme à droite. Les points en noir sont les points de discretisation.*

Les modifications à apporter aux opérateurs pour tenir compte de la nouvelle répartition des points de discrétisation sont minimes. Faire le changement de variable d'une composante quelconque, F du champ électromagnétique revient à écrire :

$$\frac{\partial}{\partial x}F = \frac{\partial F}{\partial u}\frac{\partial u}{\partial x}.$$

Puisqu'on connaît $x(u)$ et pas $u(x)$, on calcule $\dfrac{\partial u}{\partial x}$ en inversant $\dfrac{\partial u}{\partial u}$. En tenant compte des ces remarques, les opérateurs des équations (2.12a) et (2.12b) deviennent :

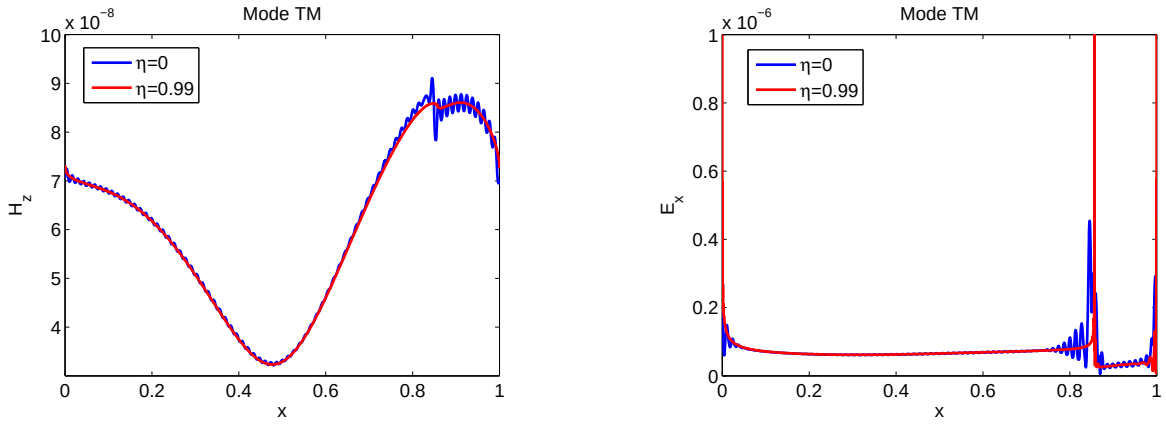
$$\mathrm{TE} : \left[\mathbf{G}^{-1}\frac{1}{k^2}\left(\dot{\mathbf{X}}^{-1}\mathbf{D}\mathbf{G}^{-1}\dot{\mathbf{X}}^{-1}\mathbf{D} + \mathbf{G}^{\varepsilon}\right)\right] \tag{2.31a}$$

$$\mathrm{TM} : \left[\left(\mathbf{G}^{\frac{1}{\varepsilon}}\right)^{-1}\frac{1}{k^2}\left(\dot{\mathbf{X}}^{-1}\mathbf{D}\left(\mathbf{G}^{\varepsilon}\right)^{-1}\dot{\mathbf{X}}^{-1}\mathbf{D} + \mathbf{G}\right)\right] \tag{2.31b}$$

où $\dot{\mathbf{X}}$ est la forme matricielle de $\dfrac{d}{du}x(u)$ aux points x_k.

Pour illustrer l'impact de la résolution spatiale adaptative, on reprend le réseau étudié précédemment (voir figure 2.18) et on compare les allures de chaque composante du champ sans utiliser l'adaptation spatiale ($\eta = 0$) avec les résultats obtenus en utilisant un paramètre de

contraction très influent ($\eta = 0.99$). Les figures 2.20(a) et 2.20(b) montrent la comparaison des modules des composantes tangentielles du champ H_z et E_x, calculés pour $\eta = 0$ (la courbe en bleu) et pour $\eta = 0.99$ (la courbe en rouge). On peut observer que les oscillations parasites sont

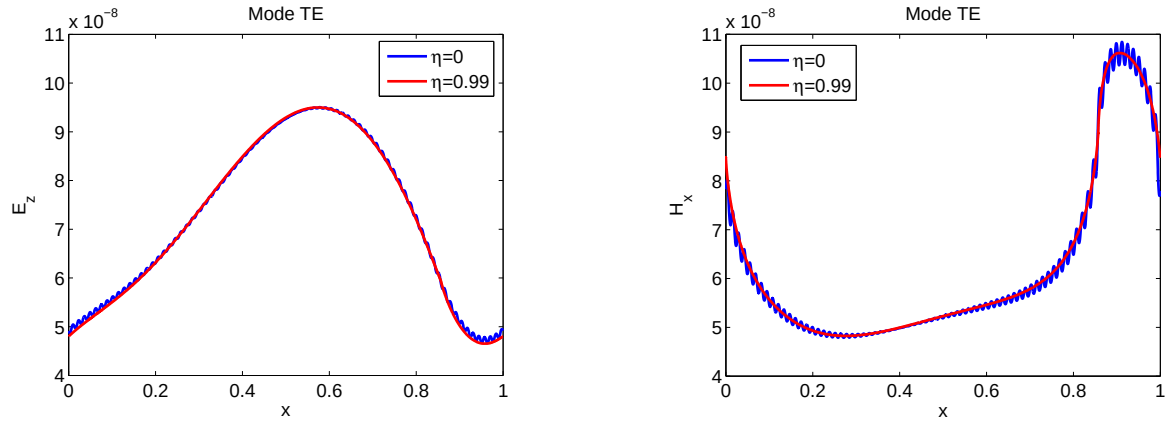


(a) première composante du champ *(b) deuxième composante du champ*

FIGURE 2.20 – *Polarisation TM. Masques en cuivre,* $\varepsilon_{22} = (1,133 - i2,558)^2$*, sur un substrat d'oxyde de silicium* $\varepsilon_3 = 1,462^2$*. Les autres paramètres sont :* $\theta = 40^\circ$*,* $d = 350$ *nm,* $\lambda = 500$ *nm,* $h = 100$ *nm,* $f = 1/7$*,* $\varepsilon_1 = \varepsilon_{21} = 1$*,* $N = 41$*.*

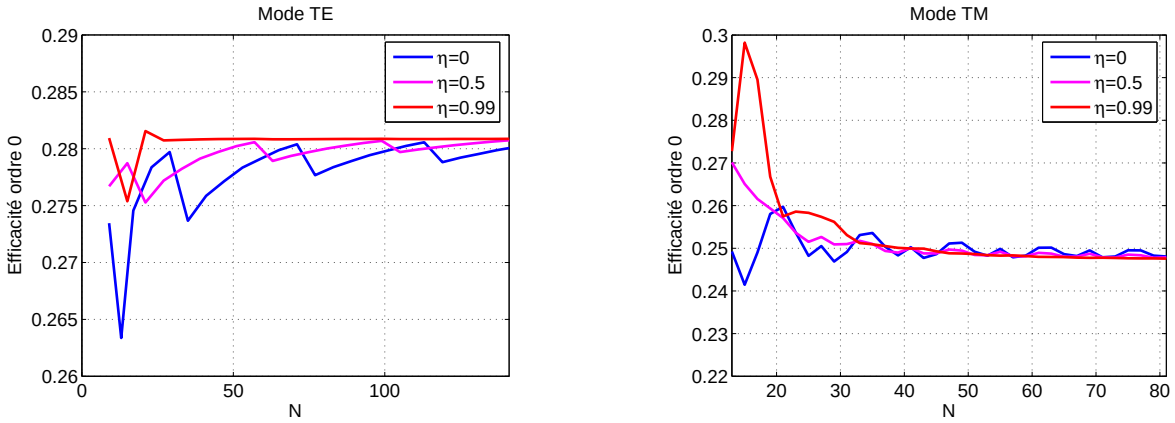
totalement gommées par la formulation paramétrique mais de plus la discontinuité de E_x est bien représenté et ce sans aucune oscillation. Ce résultat est tout aussi valable dans le cas TE où des oscillations du champ apparaissent également malgré une variation moins rapide du champ. Cf fig.2.21.

Un autre avantage de l'utilisation de la résolution spatiale adaptative est qu'on améliore la vitesse de convergence des efficacités de diffraction. Les figures 2.22(a) et 2.22(b) représentent les courbes de convergence de l'efficacité de l'ordre spéculaire en fonction de N pour différentes valeurs du paramètre η, dans les deux cas de polarisation. La courbe en bleu n'utilise pas la résolution spatiale adaptative ($\eta = 0$) ; la courbe en magenta utilise un paramètre de contraction intermédiaire ($\eta = 0.5$) et la courbe en rouge utilise le paramètre η très proche de 1. On remarque clairement l'impact sur la vitesse de convergence des résultats. Lorsque le paramètre de résolution spatiale devient de plus en plus influent les efficacités convergent plus vite. Dans la figure 2.22 (a) on peut clairement remarquer que les oscillations dues à l'effet de discrétisation décrit dans le paragraphe 2.3.3, présentes sur la courbe bleu sont atténuées sur la courbe en magenta qui utilise l'adaptation spatiale avec un paramètre $\eta = 0.5$ et totalement gommées dans le cas où le paramètre η est très influent ($\eta = 0.99$).



(a) première composante du champ *(b) deuxième composante du champ*

FIGURE 2.21 – *Polarisation TE. Masques en cuivre,* $\varepsilon_{22} = (1,133 - i2,558)^2$*, sur un substrat d'oxyde de silicium* $\varepsilon_3 = 1,462^2$*. Les autres paramètres sont :* $\theta = 40^\circ$*,* $d = 350$ *nm,* $\lambda = 500$ *nm,* $h = 100$ *nm,* $f = 1/7$*,* $\varepsilon_1 = \varepsilon_{21} = 1$*,* $N = 41$*.*



(a) Polarisation TE *(b) Polarisation TM*

FIGURE 2.22 – *Convergences en utilisant la méthode de Galerkin avec des B-splines quadratiques pour différentes valeurs l'adaptation spatiale et un réseau avec les mêmes paramètres que pour la figure 2.21.*

2.4 Bilan

Ce chapitre nous a permis de présenter une approche originale de résolution des équations de Maxwell en utilisant les fonctions B-splines. Pour chacun de ces cas test on a calculé les courbes de convergence des valeurs propres et des efficacités de diffraction. Le choix des fonctions de base, de test, la grille de discrétisation ainsi que des techniques d'adaptation spatiale ont une influence sur les performances des méthodes.

Nous avons testé notre méthode en faisant varier l'ensemble des paramètres opto-géométriques du problème de diffraction. Bien que tous les résultats ne soient pas reportés ici, quelques points méritent d'être mentionnées. De manière générale, la méthode est robuste aux changements des paramètres physiques. Comme pour les autres méthodes modales on obtient des convergences plus lentes pour des facteurs de remplissage petit et/ou des contrastes d'indices élevés. Les convergences ont été accélérées grâce à l'introduction de la résolution spatiale adaptative. D'autres méthodes modales utilisent également l'adaptation de la résolution spatiale avec la même fonction $x(u)$. Le grand avantage de notre formulation, par rapport aux autres est que le paramètre optimal [9] de contraction des points autour des discontinuités, η, est toujours voisin de 1, ce qui n'est pas le cas d'autres formulations. Du point du vue pratique cela constitue un avantage considérable.

Des convergences encore plus rapides pourraient être obtenues en augmentant le degré des B-spline. En contre partie, on obtiendrait des matrices moins creuses. Si on utilisait la méthode de Galerkin avec des B-splines de degré p, le nombre minimal, en terme de taille des matrices est $2p + 1$. En effet, comme chaque spline se superpose avec p voisins à droite et p à gauche, on obtiendrait des matrices avec $2p + 1$ diagonales. Même si on est obligé de partir de plus loin en troncature (au minimum au $2p + 1$), on gagne en précision initiale des résultats. Il faut donc trouver le compromis entre le degré de la spline utilisée et la robustesse des résultats. Dans des problèmes voisins, comme ceux d'interpolation, le meilleur compromis a été la spline cubique, on s'est aussi arrêté au degré 3.

La première motivation pour l'utilisation des fonctions B-spline dans la méthode modale n'était pas de trouver une méthode encore plus efficace pour caractériser les réseaux. Pour cela, les méthodes classiques donne des résultats fiables dans la plupart des cas. Ces travaux ont été motivés par la possibilité d'introduction des bases hiérarchiques et l'analyse multirésolution en utilisant des ondelettes splines. Ce type d'analyse a été appliqué avec succès dans des do-

9. Ceci est valable au moins dans tous les cas que nous avons testé mais ça n'est pas une démonstration.

maines variés. Elle est bien adaptée pour décrire les discontinuités brusques avec des indices très contrastés ou localisés.

L'analyse multirésolution peut être vue comme une forme de résolution spatiale adaptative particulière. L'idée de base est de faire un zoom sur les zones à variation brusque du champ en utilisant une analyse à plusieurs échelles de détail. Notre travail, a consisté à l'appliquer aux équations de Maxwell. Ceci est développé dans le chapitre suivant 3 où nous exposons en détail la théorie et son application dans le contexte de la méthode modale et de la diffraction par des réseaux.

Chapitre 3

Méthode Modale avec décomposition en ondelettes

Sommaire

3.1 Introduction 70
3.2 Eléments de théorie sur les ondelettes 71
3.2.1 Transformée en ondelettes continue 71
3.2.2 Transformée en ondelettes discrète 73
3.2.3 Analyse multirésolution 73
3.2.4 Quelques propriétés intéressantes des ondelettes 75
3.2.5 Ondelettes orthogonales 78
3.2.6 Ondelettes biorthogonales 81
3.2.7 Transformée en ondelettes rapide 84
3.3 Applications aux équations de Maxwell 87
3.3.1 La forme des matrices au premier niveau de résolution 89
3.3.2 La forme des matrices au deuxième niveau de résolution 91
3.3.3 Décimation des détails : le zoom 92
3.4 Résultats 94
3.4.1 L'influence de la longueur des filtres 94
3.4.2 Ajouter un niveau de détail 97
3.4.3 Comparaison avec la littérature 103
3.5 Bilan 104

3.1 Introduction

Les ondelettes sont issues de l'intuition de l'ingénieur en géophysique, Jean Morlet, dans les années 1980. Ce géophysicien les a employées pour la première fois afin de résoudre des problèmes de traitement des signaux pour la prospection pétrolière. [1] A partir des simulations sur un petit ordinateur de bureau, Morlet créa une méthode empirique pour décomposer un signal en ondelettes et le reconstruire à l'aide de ces ondelettes. Sous l'impulsion du physicien A. Grossman [14], Morlet a justifié mathématiquement ces résultats empiriques. Ils ont montré que, lorsque l'on représente un signal avec des ondelettes, l'énergie [2] du signal ne change pas. Les mathématiciens S. Mallat [15] et Y. Meyer [18], ont établi une structure d'ordre en introduisant un environnement qu'ils appellent analyse multirésolution (à plusieurs échelles de détails) à l'aide d'une fonction appelée fonction d'échelle. Ils ont montré qu'on peut décomposer un signal en ondelettes à des résolutions différentes en utilisant une succession des filtres. Mallat créa une méthode systématique de construction d'ondelettes orthogonales [16], connu comme l'algorithme de Mallat, qui fournit une analyse complète, concise et rapide. En 1988, I. Daubechies [17] élargit la théorie en créant les ondelettes non analytiques, fabriquées à partir d'itérations, tout en gardant les propriétés d'orthogonalité, régularité et support compact. Puis la contrainte de régularité a pu être relaxée quand des ondelettes biorthogonales [82] sont apparues. L'approche est moins facile à mettre en œuvre puisqu'il faut jongler avec deux analyses multirésolution en même temps et donc avec quatre types d'ondelettes. Néanmoins les ondelettes biorthogonales ont de nombreux avantages. Elles sont très utiles pour les applications de codage [83] ou traitement d'image [84].

Initialement conçues pour le traitement du signal, les ondelettes et les techniques multirésolution ont été appliquées avec succès dans différents domaines, notamment en traitement d'images pour des problèmes tels que l'estimation du mouvement, la reconnaissance de forme, la recherche dans des bases de données et la transmission d'informations. Elles sont également des outils efficaces pour le débruitage et la compression de données. La propriété essentielle, exploitée dans ces applications, est la possibilité d'approximer les images à plusieurs échelles.

1. La méthode standard de détection du pétrole sous terre consiste à envoyer des vibrations dans le sol et à analyser les échos. De cette manière, on tente de déterminer la profondeur des diverses couches, leur épaisseur et leur constitution. Les fréquences des ces échos sont toutes mélangées entre elles et pour extraire l'information de cet enchevêtrement d'échos à l'aide de l'analyse de Fourier, Morlet avait besoin d'ordinateurs de plus en plus puissants. Il a essayé de donner une solution intuitive au problème des ces interférences en liant chaque fréquence à l'épaisseur d'une couche sous terre : les hautes fréquences correspondaient grossièrement aux couches minces et les basses fréquences aux couches épaisses.

2. L'énergie désigne ici la valeur moyenne du carré de l'amplitude.

On peut partir d'une vue "grossière" qu'on peut affiner au cours de traitements successifs. Aujourd'hui, de nombreux logiciels [3] rendent accessible la mise en œuvre d'une transformation en ondelettes en utilisant diverses familles d'ondelettes.

Parmi les nombreuses familles d'ondelettes notre choix s'est porté sur la famille des ondelettes splines. Ces ondelettes sont faciles à manipuler parce qu'elles ont des formules explicites dans les deux domaines (temporel et fréquentiel). De plus, elles possèdent des propriétés très utiles comme la régularité, la symétrie où l'antisymétrie ; elles peuvent être conçues pour avoir un support compact. L'ondelette père [4] est une fonction B-spline qui est (pour un degré fixé) la fonction la plus régulière qu'on peut définir sur un support compact minimal. Cette propriété importante nous fournit une localisation temps-fréquence optimale. Unser [88] montre que, pour un degré fixé, cette famille de fonctions a les meilleurs propriétés d'approximation parmi toutes les ondelettes connues. Le premier exemple d'une ondelette spline a été donné par Haar [89], l'ondelette de Haar est une spline de degré 0. Les exemples les plus connus d'ondelettes splines sont les ondelettes orthogonales de Battle-Lemarié [90, 91], les ondelettes B-spline [92, 93], les ondelettes splines biorthogonales [94].

Un autre type d'ondelettes utilisé dans le cadre de notre travail sont les ondelettes symlets. Puisque ces ondelettes sont orthogonales, elles sont plus faciles à manipuler que les ondelettes B-spline. En plus les symlets sont les ondelettes les plus symétriques possibles parmi toutes les familles d'ondelettes orthogonales.

3.2 Eléments de théorie sur les ondelettes

3.2.1 Transformée en ondelettes continue

La transformée en ondelettes constitue une alternative à l'analyse temps fréquence faite en employant la transformée de Fourier. La transformée en ondelettes continue consiste à créer à partir d'une fonction mère, couramment appelée ondelette analysante, ψ qui ressemble à une petite onde, une famille d'ondelettes $\psi(\alpha x + k)$ où α et k sont des nombres réels. α sert à dilater

3. Parmi les logiciels de calcul numérique ou analytique, on mentionne la Wavelet Toolbox, une boîte à outils MATLAB [85] et Wavelet Explorer, une bibliothèque de Mathematica [86] ainsi que WaveLab qui est une bibliothèque complète de fonctions MATLAB élaborée à l'université de Stanford [87].

4. Une ondelette père, appelée couramment fonction d'échelle, est une fonction génératrice de l'analyse à plusieurs niveaux de détails. A partir de cette fonction on construit les fonctions génératrices de la résolution supérieure.

la fonction ψ et k à la translater. Quand on analyse une fonction $f(x)$ avec ces ondelettes, on la transforme en une fonction de deux variables [95] :

$$C(\alpha,k) = |\alpha|^{1/2} \int_{-\infty}^{\infty} f(x)\psi(\alpha(x-k))dx =< f, |\alpha|^{1/2}\psi^*(\alpha(x-k)) >, \alpha \neq 0 \qquad (3.1)$$

Si l'on regarde la formule (3.1), on observe qu'il y a une similitude entre la transformation en ondelettes et la transformation de Fourier. Avec la transformée de Fourier on transforme une fonction à l'aide de sinus et cosinus. En utilisant la transformée en ondelettes à la place des sinus et cosinus, on utilise la famille de fonctions

$$\{|\alpha|^{1/2}\psi(\alpha(x-k)), (\alpha,k) \in \mathbb{R}^* \text{ x } \mathbb{R}\}.$$

Cette famille est formée à partir d'extensions/contractions (selon que $|\alpha| < 1$ ou $|\alpha| > 1$) et translations de $\psi(x)$. Notons que le facteur $|\alpha|^{1/2}$ joue seulement le rôle d'un terme de normalisation assurant que chaque ondelette de la famille soit de même amplitude. L'ondelette $|\alpha|^{1/2}\psi(\alpha(x-k))$ est localisée spatialement autour du point k, qui est appelé paramètre de localisation. Une valeur élevée de $|\alpha|$ correspond à des échelles fines alors qu'une faible valeur correspond à des échelles grossières. Le paramètre $1/\alpha$ est appelé facteur d'échelle.

Une propriété importante de la transformée en ondelettes est que la fonction peut être reconstruite à partir de $C(\alpha,k)$, en la multipliant par l'ondelette analysante et en calculant l'intégrale du produit. On obtient la relation suivante qui définit la transformée inverse :

$$f(x) = \frac{1}{K_\psi} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} C(\alpha,k)|\alpha|^{1/2}\psi(\alpha(x-k))d\alpha dk \qquad (3.2)$$

avec

$$K_\psi = \int_{-\infty}^{\infty} \frac{|\Psi(f)|^2}{|f|} df,$$

$\Psi(f)$ étant la transformée de Fourier de l'ondelette analysante ψ. La formule d'inversion (3.2) nécessite que la constante K_ψ soit finie. Cette dernière condition (dite d'admissibilité) implique que :

$$\Psi(0) = \int_{-\infty}^{\infty} \psi(x)dx = 0 \qquad (3.3)$$

ce qui signifie que l'ondelette analysante doit être de moyenne nulle. Grâce à cette propriété les régions où la fonction $f(x)$ ne varie pas donnent des coefficients nuls. Ainsi les ondelettes n'encodent que les variations de la fonction.

L'approche de la transformée inverse en ondelettes est une fois de plus similaire à la transformée inverse de Fourier. Elle permet de reconstruire la fonction $f(x)$ à partir de sa transformée en ondelettes en additionnant des ondelettes dilatées, contractées et translatées de la même

manière qu'on reconstruit une fonction en additionnant des sinus et cosinus avec l'analyse de Fourier. La particularité de la transformée en ondelettes est que l'on peut analyser une fonction avec des ondelettes dilatées et contractées à la fois. Les ondelettes s'adaptent aux différentes variations de la fonction initiale : les variations brusques seront encodées par des ondelettes contractées tandis que les variations lentes seront encodées par des ondelettes dilatées.

3.2.2 Transformée en ondelettes discrète

L'équation 3.2 permet d'exprimer $f(x)$ en fonction des ondelettes. Cette représentation intégrale fait intervenir les ondelettes correspondant à toutes les localisations spatiales et à tous les facteurs d'échelle. Cela la rend difficile à mettre en œuvre. De plus il y a des représentations redondantes [5]. Une manière plus économique et plus commode de reconstruction serait de disposer d'un développement de la fonction sur un ensemble dénombrable d'ondelettes. Pour cela on est conduit à utiliser une transformation en ondelettes discrète. Cela est possible si l'on ne fait varier α et k que sur des valeurs discrètes. Habituellement, l'on choisit une grille dyadique : $\alpha = 2^{-j}, j \in \mathbb{Z}$ et $k \in \mathbb{Z}$.

3.2.3 Analyse multirésolution

Définition. Une analyse multirésolution (AMR) de $L^2(\mathbb{R})$ est une suite de sous-espaces vectoriels fermés V_j de $L^2(\mathbb{R})$ vérifiant les conditions suivantes :

1. $(V_j)_{j\in\mathbb{Z}}$ constitue une suite d'espaces emboîtés :

$$\forall j \in \mathbb{Z}, V_j \subset V_{j+1} \tag{3.4}$$

2. $\lim_{j->-\infty} V_j = \{0\}$, $\lim_{j->\infty} V_j$ est dense dans $L^2(\mathbb{R})$.
3. Une fonction $f(x) \in V_j$ si et seulement si sa dilatée $f(2x) \in V_{j+1}$.
4. Il existe $\varphi(x) \in L^2(\mathbb{R})$, appelée fonction d'échelle ou ondelette père telle que $\{\varphi(x-k), k \in \mathbb{Z}\}$ forme une base orthonormale de V_0.

5. Conformément au théorème d'échantillonnage, pour pouvoir convertir un signal continu en une forme discrète, la fréquence d'échantillonnage du signal doit être supérieure ou égale au double de la fréquence maximale contenue dans le signal.
Si les fréquences présentes dans le signal respectent le théorème d'échantillonnage, on obtient la transformée vraie. Dans le cas contraire, on obtient un phénomène appelé "recouvrement du spectre". Celui ci est gênant parce qu'il provoque des distorsions dans les fréquences plus proches de la fréquence maximale du signal.

Dans ce contexte, appliquer l'AMR à notre problème de diffraction consiste à calculer les projections orthogonales de chaque composante du champ électromagnétique sur les espaces $V_j, j \in \mathbb{Z}$.

Conséquences. La relation d'inclusion 1 nous indique que V_j est un sous-espace de V_{j+1}, voir figure 3.1. Ainsi, la projection d'une composante quelconque f sur V_j constitue nécessai-

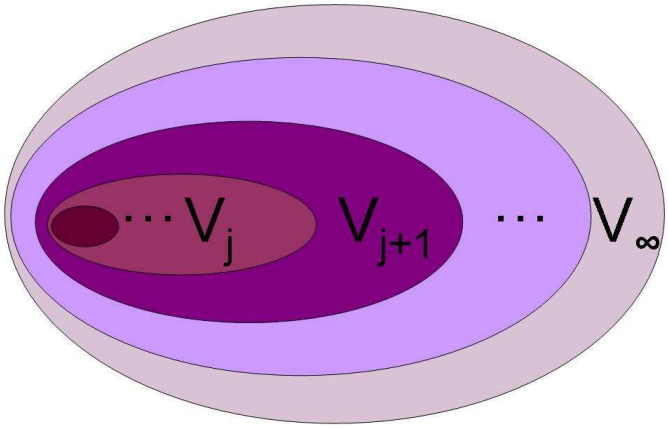


FIGURE 3.1 – *Suite des espaces emboités V_j.*

rement une moins bonne approximation que sa projection sur V_{j+1}. Par ailleurs, la relation 1 peut être interprétée comme une propriété de causalité qui montre que l'approximation sur V_{j+1} contient toute l'information nécessaire pour calculer l'approximation sur V_j.

Les conditions 2 signifient que, lorsque j tend vers $-\infty$, on perd toutes les informations sur la fonction et quand j tend vers ∞, l'approximation est asymptotiquement égale à la fonction d'origine.

D'après la condition 3, si on augmente progressivement l'échelle d'un facteur deux, sans se préoccuper des échelles intermédiaires, on double le nombre d'ondelettes qui échantillonnent la période du réseau (l'intervalle $[0, d]$). Puisqu'on passe de V_j à V_{j+1} en multipliant par 2 le facteur d'échelle, j est appelé niveau de résolution. Par récursivité, on déduit facilement de cette propriété que $\forall j \in \mathbb{Z}$, une composante du champ, $f(x)$, appartient à V_0 si et seulement si $f(2^j x)$ appartient à V_j.

La condition 4 implique que, pour tout $j \in \mathbb{Z}$, l'espace V_j est invariant par la translation d'un multiple de 2^j :

$$\forall k \in \mathbb{Z},\ f(x) \in V_j \Rightarrow\ f(x - k2^j) \in V_j.$$

Les deux derniers résultats se traduisent par le fait que $\{2^{-j/2}\varphi(2^j x - k),\ k \in \mathbb{Z}\}$ constitue une base orthonormale de V_j. Pour un j et k fixés, on notera par la suite $2^{-j/2}\varphi(2^j x - k)$ par

$$\varphi_{j,k}(x) = 2^{-j/2}\varphi(2^j x - k). \tag{3.5}$$

Interprétation intuitive Pour donner un aperçu sur les principes de l'analyse multirésolution, prenons l'exemple de l'approximation du nombre $153/13 = 11,769230\ldots$. Selon la précision désirée, on peut arrondir $153/13$ par 10 ; 11 ; $11,7$; $11,76$, etc. De la même manière la fonction d'échelle comprimée ou dilatée donne l'image de la fonction à une résolution donnée. Les ondelettes encodent la différence d'information entre deux résolutions successives. Dans l'exemple précèdent, en utilisant le système décimal la différence entre deux résolutions consécutives est de facteur 10. Entre la première résolution où l'on a comme information 10 et la deuxième ou l'on a 11, c'est la différence 1 qui est encodée par les ondelettes. Entre 11 et 11,7 des ondelettes plus petites encodent le détail 0,7 ; d'autres ondelettes vont encoder le détail 0.06 et ainsi de suite. La finesse des détails qu'on rajoute progressivement améliore l'approximation. Dans l'autre sens, quand on étire la fonction d'échelle, on finit par ne plus rien voir, comme si on essayait approximer $153/13$ avec des multiples de 100. Les nombres décimaux peuvent approximer n'importe quel nombre sans redondance et avec une précision arbitraire. La multirésolution fait de même avec n'importe quelle fonction.

3.2.4 Quelques propriétés intéressantes des ondelettes

Une propriété importante des ondelettes est le nombre de moments nuls. De façon générale le moment d'ordre m d'une fonction est l'intégrale de cette fonction multipliée par sa variable élevée à la puissance m :

$$\mathfrak{M}_m = \int_{-\infty}^{+\infty} x^m \psi(x) dx \quad m = 0, 1, 2\ldots \tag{3.6}$$

Le nombre de moments nuls (lié aux oscillations) précise l'interaction entre une ondelette et un monôme de degré quelconque. Si le moment d'ordre 1 d'une ondelette est nul, celle-ci est orthogonale aux fonctions linéaires. L'intérêt de travailler avec des ondelettes possédant un grand nombre de moments nuls est de concentrer un maximum d'information dans très peu de coefficients. Ainsi une variation locale abrupte d'une fonction fournira de très grands coefficients

d'ondelettes qui émergent de l'arrière-plan des coefficients nuls ou petits. L'exemple de la transformées en ondelette rapide d'une fonction porte décrivant dans notre cas pratique la fonction permittivité électrique ε d'un réseau lamellaire est représenté sur la figure 3.2. Dans ce cas de

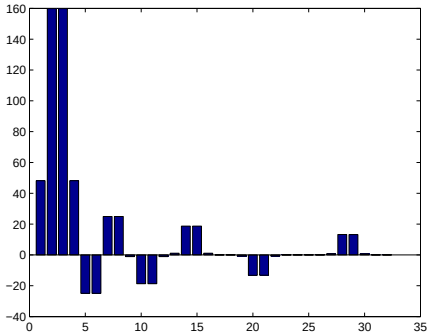


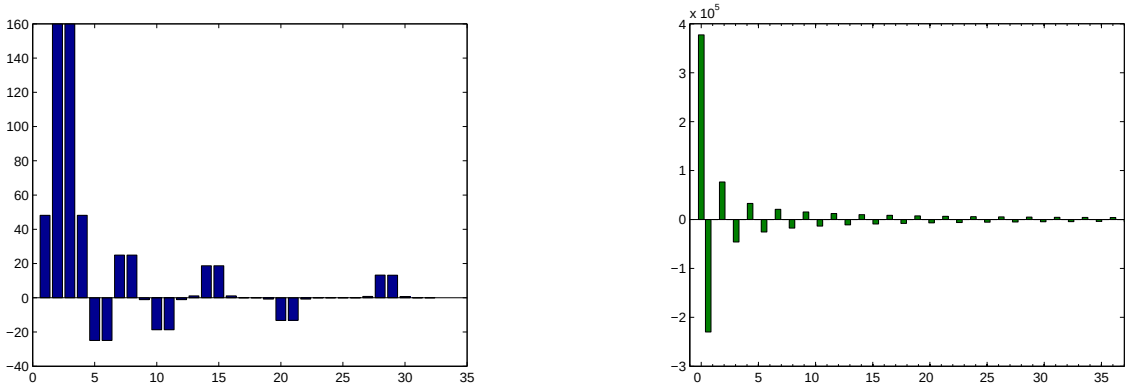
FIGURE 3.2 – *Coefficients de la transformée en ondelettes splines de la fonction de permitivité* ε.

figure le nombre de coefficients d'ondelette de la fonction ε décroît très rapidement en fonction de la résolution mais d'une résolution à l'autre, l'information portée par des coefficients de moins en moins significatifs. Les quatre premiers coefficients correspondent aux fonctions père $\varphi_{1,1}$,..,$\varphi_{1,4}$; les quatre suivant correspondent aux fonctions mère $\psi_{1,1}$,..,$\psi_{1,4}$. Les coefficients entre 9 et 16 correspondent aux ondelettes $\varphi_{2,1}$,..,$\varphi_{2,8}$, les coefficients de 17 à 32 correspondent aux ondelettes $\varphi_{3,1}$,..,$\varphi_{3,16}$.

Afin de souligner les différences entre la transformée en ondelettes et la transformée de Fourier on a représenté sur les figures 3.3(a) et 3.3(b) les deux transformées pour la même fonction porte.

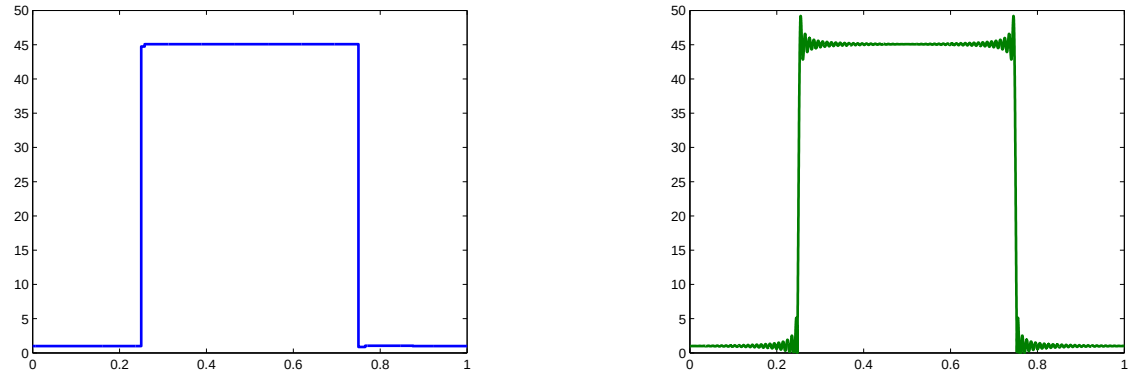
Les figures 3.4(a) et 3.4(b) montrent la reconstitution de la fonction d'origine (la fonction de permittivité $\varepsilon(x)$) en utilisant les coefficients de la transformée en ondelettes respectivement les coefficients de la transformée de Fourier. Comme on peut le voir sur la figure 3.4(b) la reconstitution à l'aide des coefficients de Fourier fait apparaître le phénomène de Gibbs, tandis qu'avec la reconstitution à l'aide des coefficients d'ondelettes, voire figure 3.4(a), ce phénomène est absent.

Une remarque simple mais essentielle pour la suite : la fonction d'échelle φ appartient à l'espace V_0. Puisque V_0 est contenu dans V_1, φ peut être aussi exprimée comme une combinaison



(a) *(b)*

FIGURE 3.3 – *Coefficients des deux tranformées : (a)-la transformée en ondelettes splines de degré 0 ; (b)-la transformée de Fourier.*



(a) *(b)*

FIGURE 3.4 – *Reconstitution de la fonction $\varepsilon(x)$ en utilisant : (a)-les coefficients de la transformée en ondelettes splines de degré 0 ; (b)-les coefficients de la transformée de Fourier.*

linéaire des fonctions $(\varphi(2x-n))_{n\in\mathbb{Z}}$ qui engendrent l'espace V_1 :

$$\varphi(x) = \sum_{n\in\mathbb{Z}} h_n \varphi(2x-n) \tag{3.7}$$

où les coefficients h_n sont communément appelés coefficients de filtres. L'équation d'échelle (3.7) exprime à elle seule l'emboîtement des espaces V_j puisque par changement de variable on obtient :

$$\varphi_{j,k} = \frac{1}{\sqrt{2}} \sum_{n\in\mathbb{Z}} h_n \varphi_{j+1,2k+n}. \tag{3.8}$$

En général les sommes (3.7) et (3.8) ne sont pas finies. Pour qu'elles soient finies on doit choisir des ondelettes à support compact. On peut interpréter l'équation (3.7) comme un produit de convolution dans l'espace direct et la réécrire comme un produit simple dans l'espace de Fourier comme suit :

$$\widehat{\varphi}(\xi) = m_0(\xi/2)\widehat{\varphi}(\xi/2), \tag{3.9}$$

où la fonction $\widehat{\varphi}$ est la transformée de Fourier de φ :

$$\widehat{\varphi}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \varphi(x) e^{ix\xi} dx \tag{3.10}$$

et la fonction m_0 est une fonction périodique définie comme suit :

$$m_0(\xi) = 2^{-\frac{1}{2}} \sum_{n=0}^{L_f-1} h_n e^{ix\xi}, \tag{3.11}$$

avec L_f la longueur des filtres h_n. Cette longueur est liée au nombre des moments nuls m donnés par l'équation 3.6. Le choix des filtres déterminent complètement la fonction φ et donc l'analyse multirésolution, grâce à la définition récursive de la relation d'échelle (3.8).

Une conséquence importante de la multirésolution est qu'il est possible d'effectuer la transformée en ondelette discrète d'une fonction sans nécessairement disposer d'expression analytique des ondelettes père ou mère. Le produits scalaire de la fonction par l'ondelette se faisant grâce à un produit de convolution du signal avec le banc de filtres $(h_n)_{n\in\mathbb{Z}}$.

3.2.5 Ondelettes orthogonales

Il existe une technique systématique pour construire des bases d'ondelettes dans le cas où les fonctions $(\varphi_{j,k})_{k\in\mathbb{Z}}$ constituent une base orthonormée. Dans ce cas particulier, on dispose du projecteur orthogonal :

$$P_j f = \sum_{k\in\mathbb{Z}} \langle f, \varphi_{j,k} \rangle \, \varphi_{j,k} \tag{3.12}$$

L'ondelette analysante, ψ est alors définie par la relation :

$$\psi(x) = \sum_{n\in\mathbb{Z}} g_n \varphi(2x-n) \quad \text{avec } g_n = (-1)^n h_{L_f-1-n} \tag{3.13}$$

les g_n étant les coefficients de filtres associées aux ondelettes analysantes. Les filtres $(h_n)_{n\in\mathbb{Z}}$ et $(g_n)_{n\in\mathbb{Z}}$ sont connus sous le nom des filtres miroir en quadrature (quadratic mirror filter-QMF). L'équation (3.13) peut être exprimée comme :

$$\widehat{\psi}(\xi) = m_1(\xi/2)\widehat{\varphi}(\xi/2) \tag{3.14}$$

avec :

$$m_1(\xi) = 2^{-\frac{1}{2}} \sum_{n=0}^{L-1} g_n e^{ix\xi} \tag{3.15}$$

On appellera par la suite W_j, le complémentaire orthogonal de V_j dans V_{j+1}. On peut montrer que les fonctions $(\psi_{j,k})_{k\in\mathbb{Z}}$ ainsi construites constituent une base dans W_j et que l'on a donc :

$$Q_j f = (P_{j+1} - P_j) f = \sum_{k\in\mathbb{Z}} \langle f, \psi_{j,k}\rangle \, \psi_{j,k}. \tag{3.16}$$

En pratique, pour appliquer l'AMR à notre problème de diffraction, on décompose les composantes du champ électromagnétique sur la base orthonormée de $L^2(\mathbb{R})$:

$$f = P_1 f + \sum_{j\geq 1} Q_j f = \sum_{k\in\mathbb{Z}} \langle f, \varphi_{1,k}\rangle \, \varphi_{1,k} + \sum_{j\geq 1}\sum_{k\in\mathbb{Z}} \langle f, \psi_{j,k}\rangle \, \psi_{j,k} \tag{3.17}$$

où f représente une composante tangentielle du champ électromagnétique.

La différence entre deux approximations consécutives, $(proj_{V_{j+1}} f(x) - proj_{V_j} f(x))$, représente l'information de " détails ", qui existait à l'échelle 2^{j+1} et qui est perdue à l'échelle 2^j. Cette information est contenue dans le sous-espace W_j orthogonal à V_j tel que :

$$V_{j+1} = V_j \oplus W_j \tag{3.18}$$

où le symbole $\oplus$ désigne la somme directe de sous-espaces vectoriels. Cette relation peut s'appliquer d'une manière récurrente afin d'obtenir une relation entre des niveaux de résolution non successifs.

La projection orthogonale $proj_{W_j} f(x)$ d'une composante du champ sur ce sous-espace donne l'information de " détails " au niveau de résolution j. Compte tenu de la relation (3.18), la projection sur V_{j+1} peut être décomposée en la somme des projections sur V_j et W_j :

$$proj_{V_{j+1}} f(x) = proj_{V_j} f(x) + proj_{W_j} f(x)$$

Le problème central qui subsiste est la construction de fonctions d'échelles φ telles que $(\varphi_{j,k})_{k\in\mathbb{Z}}$ constitue une base orthonormée, et ayant des propriétés numériques intéressantes : support compact, régularité, ordre élevé d'approximation pour les espaces V_j correspondants. Ce problème a été résolu par I. Daubechies en 1988 dans un célèbre article fondateur de la théorie moderne des bases d'ondelettes [17]. L'approche suivie consiste à définir directement la fonction φ comme une solution de l'équation d'échelle (3.7) et à mettre au point les coefficients h_n dans cette équation de façon à assurer les propriétés requises sur la solution φ. On peut ainsi vérifier que l'orthonormalité des translatées de φ se traduit par les équations algébriques :

$$\sum_n h_n h_{n+2k} = 2 \; si \; k = 0, 0 \; sinon \tag{3.19}$$

et que les propriétés d'approximation d'ordre N pour les espaces V_j se traduisent par les équations linéaires :

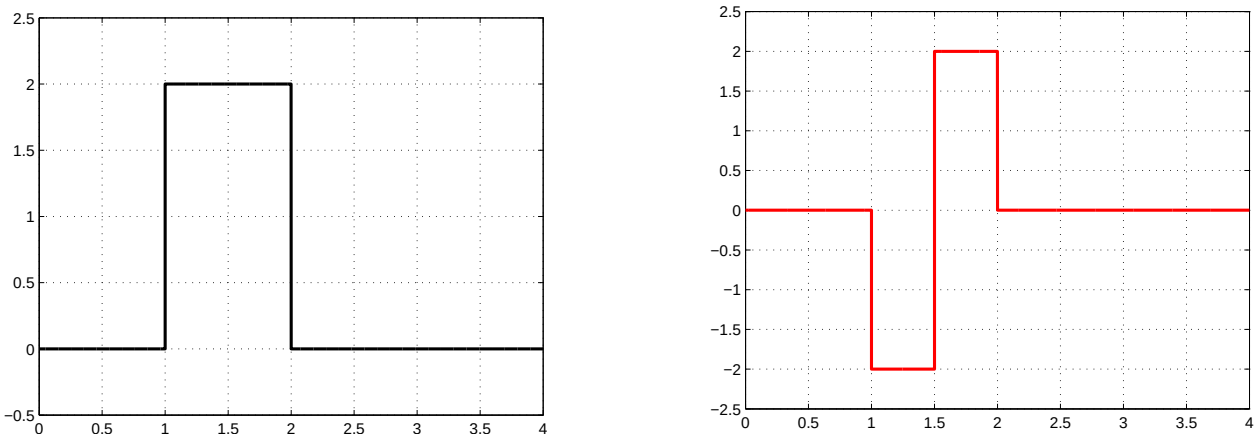
$$\sum_n h_n = 2 \;\; et \;\; \sum_n (-1)^n n^m h_n = 0, m = 0, ..., N-1 \tag{3.20}$$

Partant de ces équations, I. Daubechies construit pour tout N une suite de coefficients h_n non nuls pour les indices $\{0, 1, \ldots, 2N-1\}$, telle que la solution φ de l'équation d'échelle, à support compact dans $[0, 2N-1]$, engendre une base orthonormée $(\varphi_{j,k})_{k\in\mathbb{Z}}$ d'espaces d'approximation d'ordre N. Elle montre en outre que la régularité de la fonction φ tend vers $+\infty$ avec N : comme pour les splines, la régularité est accessible au prix d'un support plus large. Le cas $N = 0$ correspond simplement au système de Haar. On voit que, dans ce cas,

$$\varphi(x) = \varphi(2x) + \varphi(2x-1).$$

En appliquant l'équation d'échelle (3.7) ont obtient que les filtres associés à l'ondelette de Haar sont $h_0 = h_1 = 1$, $h_n = 0$ sinon. La longueur des filtres h_n est $L = 2$. Par consequence, les filtres qui définissent l'ondelette analysante, obtenus grâce à l'équation 3.13 sont : $g_0 = 1$, $g_1 = -1$, $g_n = 0$ sinon. Les figures 3.5 (a) et 3.5 (b) illustrent les ondelettes père et mère pour le cas d'une spline de dégré 0. Ces ondelettes sont connues sous le nom d'ondelettes de Haar.

Pour les splines de degré plus élevé, $N > 1$, puisque les fonctions φ sont définies positives et qu'elles se chevauchent entre elles, $(\varphi_{j,k})_{k\in\mathbb{Z}}$ ne constitue plus une base orthogonale. Pour $N = 1$ par exemple, les intégrales $\int \varphi_k(x)\varphi_{k+1}(x)dx$ et $\int \varphi_k(x)\varphi_{k-1}(x)dx$ sont non nulles. Une conséquence malheureuse de cette remarque est la difficulté de la manipulation pratique de la projection orthogonale sur l'espace V_j. La valeur dans en point de $P_j f$ est influencée par les valeurs voisines de f. Une manière élégante de résoudre cet inconvénient est d'utiliser comme espace de projection un espace $\widetilde{V}_j$ qui soit orthogonal à l'espace V_j. Cette approche nous amène à utiliser des ondelettes biorthogonales.



(a) *(b)*

FIGURE 3.5 – *Splines de degré 0 : (a)-la fonction d'échelle φ ; (b) l'ondelette analysante ψ.*

3.2.6 Ondelettes biorthogonales

Les ondelettes biorthogonales proposées en 1992 par I. Daubechies, J.C. Fauveau [94], emploient deux bases d'ondelettes (non orthonormales) de $L^2(\mathbb{R})$, $\{2^{-j/2}\psi(2^j x-k),(j,k)\in\mathbb{Z}^2\}$ et $\{2^{-j/2}\widetilde{\psi}(2^j x-k),(j,k)\in\mathbb{Z}^2\}$ qui vérifient la propriété de biorthogonalité :

$$\left\langle \frac{1}{2^{j/2}}\psi\Big(2^j x-k\Big),\frac{1}{2^{j/2}}\widetilde{\psi}\Big(2^{j'}x-k'\Big)\right\rangle = 1 \text{ si } \; j=j' \text{ et } \; k=k', 0 \text{ sinon} \tag{3.21}$$

Par analogie avec l'équation (3.5), on peut écrire $\psi_{j,k}(x) = 2^{-j/2}\psi(2^j x-k)$. La relation de dualité décrite par l'équation (3.21) peut donc être décrite dans une forme simplifiée comme :

$$\left\langle \psi_{j,k},\widetilde{\psi}_{j,l}\right\rangle = 1 \; si \; k=l, \; 0 \; sinon \tag{3.22}$$

Une fonction $f(x)$ de $L^2(\mathbb{R})$ se décompose de manière duale sur les éléments de l'une ou l'autre base :

$$f(x)=\sum_{j=-\infty}\sum_{k=-\infty}\langle f(\tau),\psi_{j,k}(\tau)\rangle\,\widetilde{\psi}_{j,k}(x)=\sum_{j=-\infty}\sum_{k=-\infty}\left\langle f(\tau),\widetilde{\psi}_{j,k}(\tau)\right\rangle\psi_{j,k}(x)$$

On introduit une fonction d'échelle duale $\widetilde{\varphi}$ orthogonale à φ :

$$\langle\varphi_{j,k},\widetilde{\varphi}_{j,l}\rangle = 1 \; si \; k=l, \; 0 \; sinon \tag{3.23}$$

Cette fonction est, elle aussi, solution d'une équation d'échelle du type :

$$\widetilde{\varphi}(x)=\sum_{n\in\mathbb{Z}}\widetilde{h}_n\widetilde{\varphi}(2x-n) \tag{3.24}$$

Une telle fonction permet de définir des projecteurs sur V_j :

$$P_j f = \sum_{k\in\mathbb{Z}} \langle f, \widetilde{\varphi}_{j,k}\rangle \, \varphi_{j,k} \tag{3.25}$$

La relation entre les ondelettes de l'espace direct et celles de l'espace dual est faite à travers les coefficients des filtres comme suit :

$$\psi(x) = \sum_{n\in\mathbb{Z}} g_n \varphi(2x-n) \text{ et } \widetilde{\psi}(x) = \sum_{n\in\mathbb{Z}} \widetilde{g}_n \widetilde{\varphi}(2x-n) \tag{3.26}$$

avec $g_n = (-1)^n \widetilde{h}_{\widetilde{L_f}-1-n}$ et $\widetilde{g}_n = (-1)^n h_{L_f-1-n}$. Grâce aux relations d'orthogonalité on obtient que :

$$\left\langle \varphi_{j,k} \, , \widetilde{\psi}_{j,l} \right\rangle = \langle \widetilde{\varphi}_{j,k} \, , \psi_{j,l} \rangle = 0 \tag{3.27}$$

Ces ondelettes sont précisément l'outil nécessaire pour caractériser le complément entre deux niveaux d'approximation dont on montre qu'il est donné par le projecteur Q_j :

$$Q_j f = (P_{j+1} - Pj) f = \sum_{k\in\mathbb{Z}} \left\langle f, \widetilde{\psi}_{j,k} \right\rangle \psi_{j,k} \tag{3.28}$$

On aboutit ainsi la décomposition d'une fonction f (qui, dans notre cas, sera une composante quelconque du champ électromagnétique) dans une base biorthogonale de $L^2(\mathbb{R})$:

$$f = P_1 f + \sum_{j\geq 1} Q_j f = \sum_{k\in\mathbb{Z}} \langle f, \widetilde{\varphi}_{1,k}\rangle \, \varphi_{1,k} + \sum_{j\geq 1} \sum_{k\in\mathbb{Z}} \left\langle f, \widetilde{\psi}_{j,k} \right\rangle \psi_{j,k} \tag{3.29}$$

Notons que le cas des ondelettes orthonormales correspond au cas particulier où les fonctions d'analyse $(\widetilde{\varphi}, \widetilde{\psi})$ coïncident avec les fonctions de synthèse (φ, ψ).

Une propriété importante des bases d'ondelettes biorthogonales est la stabilité : on peut trouver deux constantes $C_1, C_2 > 0$ telles que pour tout $f \in L^2(\mathbb{R})$:

$$C_1 \|f\|_2^2 \leq \sum_{k\in\mathbb{Z}} |\, \langle f, \widetilde{\varphi}_{1,k}\rangle \,|^2 + \sum_{j\geq 1} \sum_{k\in\mathbb{Z}} |\, \left\langle f, \widetilde{\psi}_{j,k} \right\rangle \,|^2 \leq C_2 \|f\|_2^2 \tag{3.30}$$

Cette propriété est bien entendu vérifiée dans le cadre orthonormal avec $C_1 = C_2 = 1$.

Le problème est donc ici la construction des fonctions d'échelles duales φ et $\widetilde{\varphi}$. La démarche est de nouveau de partir des équations d'échelles et de mettre au point les coefficients h_n et $\widetilde{h}_n$ de façon à assurer les propriétés requises sur les solutions φ et $\widetilde{\varphi}$. On peut ainsi vérifier que les relations de dualité se traduisent par les équations :

$$\sum_n \widetilde{h}_n \widetilde{h}_{n+2k} = 2 \;\; si \;\; k = 0, \; 0 \; sinon \tag{3.31}$$

Cette approche offre beaucoup plus de flexibilité que la précédente : on peut par exemple fixer les coefficients h_n donc la fonction φ, et chercher les coefficients $\widetilde{h}_n$ comme solutions de ces équations linéaires. Notons que ces solutions ne sont en général pas uniques et que l'on peut ainsi les sélectionner de sorte que $\widetilde{\varphi}$ ait de bonnes propriétés numériques (support compact, régularité). Cette approche permet en particulier d'obtenir des fonctions duales pour les splines de degré arbitraire.

Si on prend l'exemple d'une fonction spline φ, telle qu'elle a été définie par Boor [57] on peut vérifier qu'à un facteur d'échelle près les filtres associées à la splines de degré N sont donnés par :

$$h_n = 2^{-N} \frac{(N+1)!}{n!(N+1-n)!} \tag{3.32}$$

pour $n = 0, ..., N+1$ et $h_n = 0$ sinon.

Les coefficients de filtres pour une spline quadratique sont : $h_0 = 1/4, \quad h_1 = 3/4, \quad h_2 = 3/4$ et $h_3 = 1/4$, donc d'après l'équation d'échelle 3.7, $\varphi(x) \in V_0$ s'exprime comme combinaison des fonctions d'échelle de niveau $\varphi(2x) \in V_1$:

$$\varphi(x) = \frac{1}{4}\varphi(2x-3) + \frac{3}{4}\varphi(2x-2) + \frac{3}{4}\varphi(2x-1) + \frac{1}{4}\varphi(2x).$$

La figure 3.6 illustre les fonctions d'échelle du niveau de résolution 1 (en traits continus) multipliées par les coefficients de filtre adéquats afin de reconstruire la spline quadratique du niveau 0 (en traits pointillées).

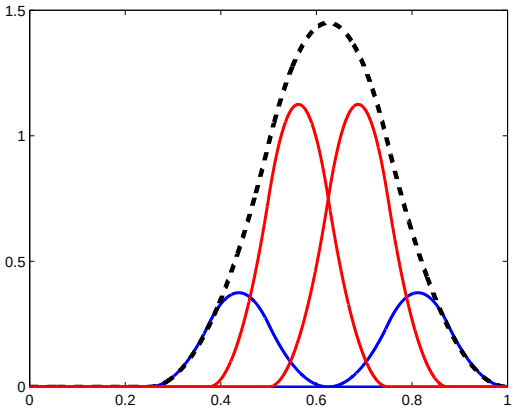
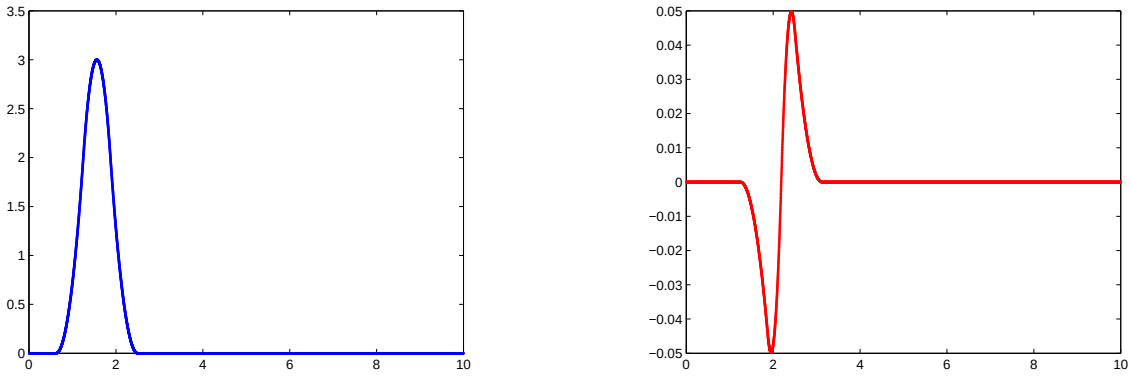


FIGURE 3.6 – *Reconstitution d'une ondelette spline à l'aide de filtres et d'ondelettes de l'échelle suppérieure.*

La relation (3.13) fournit les coefficient de filtres g_n qui aident à la construction de l'ondelette analysante. En fonction du nombre de moments nuls, ces fonctions peuvent différer. Pour le cas d'une spline quadratique, les coefficients de filtre associés aux ondelettes analysantes avec 2 moments nuls sont : $g_{-2} = 1/4,\ \ g_{-1} = -3/4,\ \ g_0 = 3/4$ et $g_1 = -1/4$. L'ondelette analysante s'écrit aussi comme combinaison linéaire des fonctions $\varphi(2x) \in V_1$:

$$\psi(x) = \frac{1}{4}\varphi(2x+2) - \frac{3}{4}\varphi(2x+1) + \frac{3}{4}\varphi(2x) - \frac{1}{4}\varphi(2x-1).$$

La figure 3.7 présente ainsi des fonctions φ et $\widetilde{\varphi}$ dans le cas où φ est la fonction associée aux B-splines quadratiques.



(a) *(b)*

FIGURE 3.7 – *Splines de degré 2 : (a)-la fonction d'échelle φ ; (b) l'ondelette analysante ψ.*

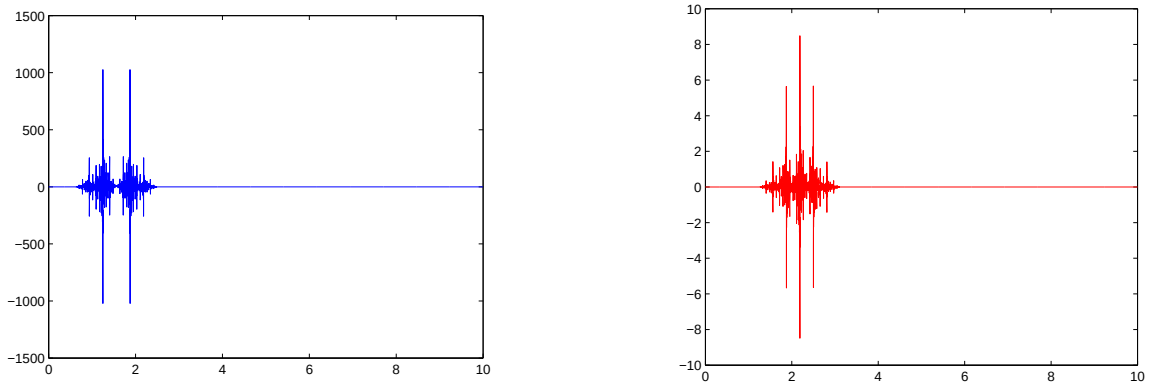
Les coefficients de filtres de l'équation d'échelle duale $\widetilde{h}$ et $\widetilde{g}$ s'écrivent à l'aide des filtres de l'espace direct h et g :

$$\widetilde{h}_n = -(-1)^n g_{L_g-1-n} \; et \; \widetilde{g}_n = (-1)^n h_{L_h-1-n}. \tag{3.33}$$

Nous illustrons sur la figure 3.8 les fonctions $\widetilde{\varphi}$ et $\widetilde{\psi}$ pour le cas où φ est une spline quadratique et l'ondelette analysante ψ a deux moments nuls.

3.2.7 Transformée en ondelettes rapide

Cette transformée est appliquée dans à notre problème pour décomposer la fonction de permittivité $\varepsilon(x)$. Les coefficients de la décomposition de $\varepsilon(x)$ sur une base d'ondelettes se



(a) *(b)*

FIGURE 3.8 – *Splines de degré 2 : (a)-la fonction d'échelle duale $\widetilde{\varphi}$; (b)-l'ondelette analysante duale $\widetilde{\psi}$.*

calculent par un algorithme rapide qui cascade des convolutions discrètes des filtres h et g. La transformée en ondelettes rapide (FWT) décompose successivement chaque approximation $P_{j+1}f$ en une approximation plus grossière P_jf et en coefficients d'ondelettes correspondant à Q_jf. Le point de départ est de définir les suites $c_J = (c_{j,k})$ et $d_J = (d_{j,k})$ comme des produits scalaire de ε et des fonctions père φ et respectivement des fonctions mère ψ :

$$c_{j,k} = \langle \varepsilon, \widetilde{\varphi}_{j,k} \rangle \;\; et \;\; d_{j,k} = \langle \varepsilon, \widetilde{\psi}_{j,k} \rangle. \tag{3.34}$$

En substituant 3.34 dans la représentation de ε en base d'ondelettes de l'équation (3.29) on obtient :

$$f = c_1\varphi + \sum_{j\geq 1}\sum_{k\in\mathbb{Z}} d_{j,k}\psi_{j,k} \tag{3.35}$$

On cherche des relations entre les suites $c_{j,k}$ et $d_{j,k}$. En utilisant la relation d'échelle 3.7, on peut écrire $c_{j,k}$ comme suit :

$$c_{j,k} = \langle \varepsilon, \widetilde{\varphi}_{j,k} \rangle = \langle \varepsilon, \frac{1}{\sqrt{2}} \sum_{n\in\mathbb{Z}} \widetilde{h}_n \widetilde{\varphi}_{j+1,2k+n} \rangle = \frac{1}{\sqrt{2}} \sum_{n\in\mathbb{Z}} \widetilde{h}_n c_{j+1,2k+n} = \frac{1}{\sqrt{2}} \sum_{n\in\mathbb{Z}} \widetilde{h}_{n-2k} c_{j+1,n}. \tag{3.36}$$

Un calcul similaire s'effectue pour $d_{j,k} = \langle \varepsilon, \widetilde{\psi}_{j,k} \rangle$ et on obtient les formules de décomposition :

$$c_{j,k} = \frac{1}{\sqrt{2}} \sum_{n\in\mathbb{Z}} \widetilde{h}_{n-2k} c_{j+1,n} \quad et \quad d_{j,k} = \frac{1}{\sqrt{2}} \sum_{n\in\mathbb{Z}} \widetilde{g}_{n-2k} c_{j+1,n}. \tag{3.37}$$

Il est intéressant d'interpréter cet algorithme de décomposition en termes de filtrage. Les formules (3.37) peuvent en effet s'interpréter comme des convolutions de la suite c_{j+1} avec les

suites $\widetilde{h}_{-n}/\sqrt{2}$ et $\widetilde{g}_{-n}/\sqrt{2}$ suivies d'une décimation consistant à ne retenir que les échantillons d'indices pairs. Un tel dispositif, représenté sur la figure 3.9 avec une normalisation évitant les facteurs $1/\sqrt{2}$, est bien connu dans la théorie du signal sous le nom de filtrage en sous-bande, les suites c_j et d_j représentant respectivement les bandes obtenues par filtrage passe-bas et passe-haut de c_{j+1}. c_{j+1}. La représentation en ondelettes est constituée des coefficients d'ondelettes

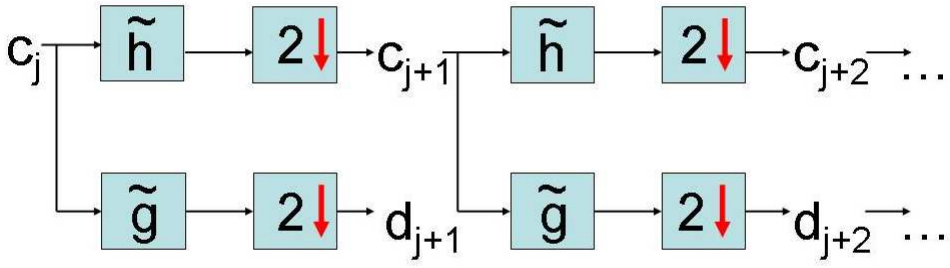


FIGURE 3.9 – Transformée en ondelettes rapide par cascades de filtrages par $\widetilde{h}$ et $\widetilde{g}$ suivis de sous-échantillonnage d'un facteur 2.

$\{c_j\}$ calculés aux échelles $2^J < 2^j \leq 2^L$ complétée par les coefficients $\{d_j\}_{J<j\leq L}$. Elle se calcule de manière récursive à partir de (3.9) pour $J < j \leq L$. Le coefficient initial c_J joue un rôle particulier.

La fonction d'origine $\varepsilon(x)$ est échantillonnée. On considère une distance d'échantillonnage quelconque N^{-1}, où N une puissance de 2 ($2^L = N$). Puis on calcule $c_J = \langle \varepsilon, \widetilde{\varphi}_{J,k} \rangle$. A un facteur de normalisation près celui-ci représente la moyenne de la fonction $\varepsilon(x)$ pondérée au voisinage des points d'échantillonnage $N^{-1}n$. Le calcul du produit scalaire du premier terme ($\langle \varepsilon, \widetilde{\varphi}_{J,k} \rangle$) se fait d'après la formule

$$\langle \varepsilon, \widetilde{\varphi}_{J,k} \rangle = \frac{1}{2^L} \sum_{n=1}^{N} \varepsilon(x_n) \widetilde{\varphi}_{J,k}(x_n) \tag{3.38}$$

Dans la pratique, pour que le calcul de la transformée en ondelette rapide soit précis il faut que L soit très grand par rapport à J. Une mauvaise approximation initiale du c_J peut engendrer des erreurs d'approximations des coefficients $\{c_j\}$, $J < j \leq L$.

3.3 Applications aux équations de Maxwell

On cherche à résoudre les équations de Maxwell en utilisant la méthode des moments avec le développement sur différents types de bases d'ondelettes. Puisque le cas orthogonal n'est qu'un cas particulier du cas biorthogonal (quand les constantes C_1 et C_2 de l'équation (3.30) sont égales) on donne l'aperçu de la résolution des équations de Maxwell dans le cas biorthogonal. Pour ce faire, on doit jongler avec deux analyses :

- une analyse multirésolution dans l'espace direct définit par la suite des espaces emboités $(V_j)_{j\in\mathbb{Z}}$, avec les fonctions génératrices père φ et mère ψ.
- une analyse de l'espace dual définit par la suite des espaces emboîtés $(\widetilde{V}_j)_{j\in\mathbb{Z}}$, avec les fonctions génératrices père $\widetilde{\varphi}$ et mère $\widetilde{\psi}$.

Reprenons les équations de Maxwell :

$$\mathbf{TE}: \begin{cases} kr \sum_{n=1}^{N} \langle T_q, B_n \rangle E_{z_n} = \omega\mu_0 \sum_{n=1}^{N} \langle T_q, B_n \rangle H_{x_n} \\ \sum_{n=1}^{N} \left\langle T_q, \dfrac{dB_n}{dx} \right\rangle E_{z_n} = i\omega\mu_0 \sum_{n=1}^{N} \langle T_q, B_n \rangle H_{y_n} \\ \sum_{n=1}^{N} \left\langle T_q, \dfrac{dB_n}{dx} \right\rangle H_{y_n} + ikr \sum_{n=1}^{N} \langle T_q, B_n \rangle H_{x_n} = i\omega\varepsilon_0 \sum_{n=1}^{N} \langle T_q, \varepsilon B_n \rangle E_{z_n} \end{cases} \tag{3.39}$$

et

$$\mathbf{TM}: \begin{cases} \dfrac{-kr}{\varepsilon_0} \sum_{n=1}^{N} \left\langle T_q, \dfrac{1}{\varepsilon} B_n \right\rangle H_{z_n} = \omega \sum_{n=1}^{N} \langle T_q, B_n \rangle E_{x_n} \\ \sum_{n=1}^{N} \left\langle T_q, \dfrac{dB_n}{dx} \right\rangle H_{z_n} = -i\omega\varepsilon_0 \sum_{n=1}^{N} \langle T_q, \varepsilon B_n \rangle E_{y_n} \\ \sum_{n=1}^{N} \left\langle T_q, \dfrac{dB_n}{dx} \right\rangle E_{y_n} + ikr \sum_{n=1}^{N} \langle T_q, B_n \rangle E_{x_n} = -i\omega\mu_0 \sum_{n=1}^{N} \langle T_q, B_n \rangle H_{z_n} \end{cases} \tag{3.40}$$

où les fonctions de base B_n sont définies dans l'espace V_j (sont donc les fonctions $\varphi_{1,k}$ et $\psi_{i,k}$ aux différents niveaux de résolution) tandis que les fonctions T_q sont définies dans l'espace dual $\widetilde{V}_j$ (les fonctions $\widetilde{\varphi}_{1,k}$ et $\widetilde{\psi}_{i,k}$ aux différents niveaux de résolution). Chaque composante du champ électromagnétique est décrite comme une somme finie de fonctions de base de l'espace direct V_j :

$$f(x) = \sum_{k\in\mathbb{Z}} F_k \varphi_{1,k}(x) + \sum_{j\geq 1}\sum_{l\in\mathbb{Z}} F_l \psi_{j,l}(x) \tag{3.41}$$

où f désigne une composante du champ ((E_z, H_x, H_y) en polarisation TE et (H_z, E_x, E_y) en polarisation TM) ; F_k et F_l sont les coefficients à déterminer. La somme sur k exprime la combinaison linéaire des fonctions d'échelle (qui encodent l'approximation "grossière"), la somme sur l la combinaison linéaire d'ondelettes analysantes (qui encodent les détails). La somme sur j exprime la profondeur des détails. En termes géométriques, la fonction d'échelle $\varphi_{1,k}$ lisse le champ f tandis que les ondelettes analysantes $\psi_{j,l}$ révèlent les différences. Pratiquement, puisque les ondelettes sont décrites par des filtres (voir les équations 3.7 et 3.13), on effectue un produit de convolution de f par deux filtres numériques. Pour transformer f en ondelettes, on convolue à chaque résolution f échantillonné et la suite de nombres $h_1, h_2, h_3, h_4, ...$, le filtre passe-bas associé à la fonction d'échelle $\varphi_{1,k}$, puis par la suite $g_1, g_2, g_3, g_4, ...$ le filtre passe-haut associé à l'ondelette analysante $\psi_{j,l}$.

On résout les équations de Maxwell par la méthode des moments, de la même manière que dans le chapitre précédant et on arrive aux équations de propagation qui sont décrites sous forme matricielle par l'équation 2.12. Puisque les fonctions test T_q sont orthogonales aux fonctions de base B_n, la matrice $\mathbf{G}$ est la matrice identité. La matrice $\mathbf{D}$ contient les produits scalaires entre les fonctions de test et les dérivées des fonctions de base. Les éléments des matrices $\mathbf{G}^{\varepsilon}$ et $\mathbf{G}^{\frac{1}{\varepsilon}}$ différent des éléments de la matrice identité dans les endroits où la fonction de permittivité casse l'orthogonalité des produits scalaires. En effet, les fonctions de base qui coupent les discontinuités ($\varepsilon(x)$ est non constant) contiendront les coefficient qui caractérisent les discontinuités :

$$\mathbf{G}^{\varepsilon}_{q,n} = \int T_q(x)\varepsilon(x)B_n(x)dx \quad \mathbf{G}^{\frac{1}{\varepsilon}}_{q,n} = \int T_q(x)\frac{1}{\varepsilon}(x)B_n(x)dx \tag{3.42}$$

Les matrices ne contiennent des coefficients non nuls qu'aux endroits où les fonctions de base et de test se superposent à un saut de permittvité. Ailleurs ε est une constante donc les conditions d'orthogonalité sont satisfaites et les produits scalaires sont nuls. Tenant compte de ces dernières remarques, l'équation de propagation devient :

$$\left[\left(\frac{1}{k^2}\mathbf{D}\mathbf{D} + \mathbf{G}^{\varepsilon}\right)\right][E_{z_n}] = r^2[E_{z_n}] \tag{3.43a}$$

$$\left[\left(\mathbf{G}^{\frac{1}{\varepsilon}}\right)^{-1}\left(\frac{1}{k^2}\mathbf{D}\left(\mathbf{G}^{\varepsilon}\right)^{-1}\mathbf{D} + \mathbf{I}\right)\right][H_{z_n}] = r^2[H_{z_n}] \tag{3.43b}$$

3.3.1 La forme des matrices au premier niveau de résolution

En utilisant un seul niveau de résolution on a $j = 1$. Pour alléger la notation on renonce à l'indice j. Chaque composante du champ s'écrit comme somme des ondelettes mère et père :

$$f(x) = \sum_{k\in\mathbb{Z}} F_k \varphi_k(x) + \sum_{l\in\mathbb{Z}} F_l \psi_l(x) \tag{3.44}$$

Puis les équations de Maxwell sont projetées sur les ondelettes mère et père de l'espace dual. Les matrices des produits scalaires $\mathbf{D}$, $\mathbf{G}^{\varepsilon}$ et $\mathbf{G}^{\frac{1}{\varepsilon}}$ auront la forme :

$$\mathbf{M} = \begin{pmatrix} \mathbf{M11} & \mathbf{M12} \\ \mathbf{M21} & \mathbf{M22} \end{pmatrix} \tag{3.45}$$

où $\mathbf{M11}$ et $\mathbf{M22}$ sont matrices carrées qui contiennent les produits scalaires homogènes, c'est-à-dire entre les fonctions de même type (les pères de $\widetilde{V}_j$ avec les pères de V_j et respectivement les mères de $\widetilde{V}_j$ avec les mères de V_j). Les matrices $\mathbf{M12}$ et $\mathbf{M21}$ contiennent les produits scalaires "hétérogènes", entre les fonctions de type différent (les pères de $\widetilde{V}_j$ avec les mères de V_j et respectivement les mères de $\widetilde{V}_j$ avec les pères de V_j).

Pour le cas de $\mathbf{G}^{\varepsilon}$ les éléments des sous matrices $\mathbf{G}11^{\varepsilon}$, $\mathbf{G}12^{\varepsilon}$, $\mathbf{G}21^{\varepsilon}$ et $\mathbf{G}22^{\varepsilon}$ sont définis comme :

$$G11^{\varepsilon}_{nm} = \int \varepsilon(x)\widetilde{\varphi}_n(x)\varphi_m(x)dx \tag{3.46a}$$

$$G12^{\varepsilon}_{nm} = \int \varepsilon(x)\widetilde{\varphi}_n(x)\psi_m(x)dx \tag{3.46b}$$

$$G21^{\varepsilon}_{nm} = \int \varepsilon(x)\widetilde{\psi}_{j,n}(x)\varphi_{j,m}(x)dx \tag{3.47a}$$

$$G22^{\varepsilon}_{nm} = \int \varepsilon(x)\widetilde{\psi}_{j,n}(x)\psi_{j,m}(x)dx \tag{3.47b}$$

Pour le cas de $\mathbf{G}^{\frac{1}{\varepsilon}}$:

$$G11^{\frac{1}{\varepsilon}}_{nm} = \int \frac{1}{\varepsilon(x)}\widetilde{\varphi}_n(x)\varphi_m(x)dx \tag{3.48a}$$

$$G12^{\frac{1}{\varepsilon}}_{nm} = \int \frac{1}{\varepsilon(x)}\widetilde{\psi}_n(x)\psi_m(x)dx \tag{3.48b}$$

$$G21^{\frac{1}{\varepsilon}}_{nm} = \int \frac{1}{\varepsilon(x)}\widetilde{\psi}_n(x)\varphi_m(x)dx \tag{3.49a}$$

$$G22^{\frac{1}{\varepsilon}}_{nm} = \int \frac{1}{\varepsilon(x)}\widetilde{\psi}_n(x)\psi_m(x)dx \tag{3.49b}$$

Les matrices $\mathbf{G}^{\varepsilon}$ et $\mathbf{G}^{\frac{1}{\varepsilon}}$ sont calculées grâce à la transformée en ondelettes rapide.

Pour le cas de la matrice $\mathbf{D}$ les éléments des sous matrices $\mathbf{D11}$, $\mathbf{D12}$, $\mathbf{D21}$ et $\mathbf{D22}$ sont définis comme :

$$D11_{nm} = \int \widetilde{\varphi}_n(x)\frac{\partial}{\partial x}\varphi_m(x)dx, \tag{3.50a}$$

$$D12_{nm} = \int \widetilde{\varphi}_n(x)\frac{\partial}{\partial x}\psi_m(x)dx, \tag{3.50b}$$

$$D21_{nm} = \int \widetilde{\psi}_n(x)\frac{\partial}{\partial x}\varphi_m(x)dx, \tag{3.51a}$$

$$D22_{nm} = \int \widetilde{\psi}_n(x)\frac{\partial}{\partial x}\psi_m(x)dx, \tag{3.51b}$$

Ces produits scalaires sont exactement détérminés grâce aux coefficients de filtres. Numériquement, les filtres déterminent exactement les valeurs des produits scalaires. Cette détermination se fait à l'aide d'une équation à valeurs propres. Les dérivées des ondelettes (mère ou père) vérifient aussi les équations d'échelle (3.7) et (3.13). Les éléments de la matrice $\mathbf{D}$ peuvent être déduits de :

$$\int \widetilde{\varphi}_n(x)\frac{\partial}{\partial x}\varphi_m(x)dx = \frac{1}{2}\sum_{p\in Z}\sum_{q\in Z}\widetilde{h}_{p-2n}h_{q-2m}\int \widetilde{\varphi}_p(x)\frac{\partial}{\partial x}\varphi_q(x)dx \tag{3.52}$$

$$\int \widetilde{\varphi}_n(x)\frac{\partial}{\partial x}\psi_m(x)dx = \frac{1}{2}\sum_{p\in Z}\sum_{q\in Z}\widetilde{h}_{p-2n}g_{q-2m}\int \widetilde{\varphi}_p(x)\frac{\partial}{\partial x}\psi_q(x)dx \tag{3.53}$$

$$\int \widetilde{\psi}_n(x)\frac{\partial}{\partial x}\varphi_m(x)dx = \frac{1}{2}\sum_{p\in Z}\sum_{q\in Z}\widetilde{g}_{p-2n}h_{q-2m}\int \widetilde{\psi}_p(x)\frac{\partial}{\partial x}\varphi_q(x)dx \tag{3.54}$$

$$\int \widetilde{\psi}_n(x)\frac{\partial}{\partial x}\psi_m(x) = \frac{1}{2}\sum_{p\in Z}\sum_{q\in Z}\widetilde{g}_{p-2n}g_{q-2m}\int \widetilde{\psi}_p(x)\frac{\partial}{\partial x}\psi_q(x)dx \tag{3.55}$$

Ces équations peuvent être écrites sous la forme :

$$\mathbf{M}X = \alpha X \tag{3.56}$$

où $\mathbf{M}$ est la matrice qui contient les produits des filtres, X est le vecteur propre contenant les produits scalaires $< \widetilde{\varphi},\ \varphi^{'} >$, $< \widetilde{\varphi},\ \psi^{'} >$ $< \widetilde{\psi},\ \varphi^{'} >$, $< \widetilde{\psi},\ \psi^{'} >$ et $\alpha = 2$ la valeur propre associée à ces vecteurs propres. Par exemple, dans l'équation 3.52 le produit scalaire

$$\int \widetilde{\varphi}_n(x)\frac{\partial}{\partial x}\varphi_m(x)dx$$

est vecteur propre d'une matrice $\mathrm{M}\widetilde{h}h$ dont l'élément de rang nm est :

$$\sum_{p\in Z}\sum_{q\in Z}\widetilde{h}_{p-2n}h_{q-2m}.$$

La valeur $\alpha = 2$ est une conséquence de l'équation d'échelle 3.7. Quand on dérive cette équation on obtient :

$$\varphi'(x) = 2\sum_{n}\varphi'(2x-n) \tag{3.57}$$

Dans l'équation 3.56 pour la valeur $\alpha = 1$ on obtient les produits scalaires $<\widetilde{\varphi},\ \varphi>, <\widetilde{\varphi},\ \psi>$ $<\widetilde{\psi},\ \varphi>, <\widetilde{\psi},\ \psi>$. De manière plus générale, pour $\alpha = 2^p$ on obtient les produits scalaires entre les fonctions test et les dérivées d'ordre p des fonctions de base puisque

$$\varphi(p)(x) = 2^p\sum_{n}\varphi(p)(2x-n) \tag{3.58}$$

Etant donné que les produits scalaires sont obtenus comme vecteurs propres des équations ci-dessus, une normalisation s'impose pour éliminer le coefficient de phase. La normalisation de l'opérateur différentiel d'ordre quelconque p respecte les relations :

$$\int \varphi(x)\frac{\partial^p}{\partial x^p}x^p dx = p!,\ \ x^p = \sum_{q} q^p\varphi(x-q) \tag{3.59a}$$

$$\int \psi(x)\frac{\partial^p}{\partial x^p}x^p dx = p!,\ \ x^p = \sum_{q} q^p\psi(x-q) \tag{3.59b}$$

3.3.2 La forme des matrices au deuxième niveau de résolution

Quand on rajoute un niveau de résolution supplémentaire, chaque élément du champ électromagnétique s'écrit comme combinaison linéaire des mères et des pères du niveau de résolution 1 et des mères du niveau de résolution 2, l'équation 3.44 devient :

$$f(x) = \sum_{k\in\mathbb{Z}} F_k\varphi_{j=1,k}(x) + \sum_{l\in\mathbb{Z}} F_l\psi_{j=1,l}(x) + \sum_{p\in\mathbb{Z}} F_p\psi_{j=2,p}(x) \tag{3.60}$$

En ajoutant un niveau de détail supplémentaire, les matrices des produits scalaires deviennent :

$$\boldsymbol{D} = \begin{bmatrix} \mathbf{D}11 & \mathbf{D}12 & \mathbf{D}13 \\ \mathbf{D}21 & \mathbf{D}22 & \mathbf{D}23 \\ \mathbf{D}31 & \mathbf{D}32 & \mathbf{D}33 \end{bmatrix} \tag{3.61}$$

$$\boldsymbol{G}^{\varepsilon} = \begin{bmatrix} \boldsymbol{G}11^{\varepsilon} & \boldsymbol{G}12^{\varepsilon} & \boldsymbol{G}13^{\varepsilon} \\ \boldsymbol{G}21^{\varepsilon} & \boldsymbol{G}22^{\varepsilon} & \boldsymbol{G}23^{\varepsilon} \\ \boldsymbol{G}31^{\varepsilon} & \boldsymbol{G}32^{\varepsilon} & \boldsymbol{G}33^{\varepsilon} \end{bmatrix} \tag{3.62}$$

$$\boldsymbol{G}^{\frac{1}{\varepsilon}} = \begin{bmatrix} \boldsymbol{G}11^{\frac{1}{\varepsilon}} & \boldsymbol{G}12^{\frac{1}{\varepsilon}} & \boldsymbol{G}13^{\frac{1}{\varepsilon}} \\ \boldsymbol{G}21^{\frac{1}{\varepsilon}} & \boldsymbol{G}22^{\frac{1}{\varepsilon}} & \boldsymbol{G}23^{\frac{1}{\varepsilon}} \\ \boldsymbol{G}31^{\frac{1}{\varepsilon}} & \boldsymbol{G}32^{\frac{1}{\varepsilon}} & \boldsymbol{G}33^{\frac{1}{\varepsilon}} \end{bmatrix} \tag{3.63}$$

Les sous matrices $\mathbf{D13}$, $\mathbf{D23}$, $\mathbf{D31}$, $\mathbf{D32}$ et $\mathbf{D33}$ encodent la différence entre le niveau $j = 1$ et $j = 2$. Ces matrices sont calculées facilement grâce aux filtres. Les produits scalaires de la résolution $j + 1$ s'expriment en fonction des filtres et les produits scalaires du niveau de résolution j. Ces relations sont une conséquence des équations d'échelle 3.7 et 3.13. Par exemple, pour le cas la matrice $\mathbf{D13}$, les produits scalaires à calculer sont ceux entre $\widetilde{\varphi}_{j=1,p}$ et $\psi'_{j=2,q}$. Grâce à l'équation d'échelle, on peut déduire que :

$$\int \widetilde{\varphi}_{j=1,p}(x)\frac{\partial}{\partial x}\psi_{j=2,q}(x)dx = \sum_{n\in Z}\widetilde{h}_n \int \widetilde{\varphi}_{j=2,2p+n}(x)\frac{\partial}{\partial x}\psi_{j=2,q}(x)dx. \tag{3.64}$$

On n'a donc à calculer que les produits scalaires entre les fonctions de base et de test de la même résolution. Une fois calculés, les autres produits scalaires ne sont rien d'autre que des combinaisons linéaires de ces produits scalaires multipliés par les filtres qui leur correspondent.

3.3.3 Décimation des détails : le zoom

A chaque niveau de résolution les ondelettes rétrécissent leur support de moitié par rapport au niveau inférieur et l'amplitude d'une ondelette devient plus grande. La figure 3.10 montre l'allure des ondelettes analysantes ψ pour les ondelettes symlets à trois niveaux de résolution différents. D'un niveau de résolution à un autre le nombre d'ondelettes double. Par exemple si on est à la troncature N et au niveau de résolution 0, on n'a que N fonctions d'échelle qui génèrent des matrices carrées de taille N, (les matrices $\mathbf{M}11$). Quand on rajoute un niveau de résolution (les ondelettes analysantes) la taille des matrices devient $2N$(on ajoute les matrices $\mathbf{M}12$, $\mathbf{M}21$ et $\mathbf{M}22$). Au deuxième niveau de résolution on rajoute les ondelettes analysantes $\psi_{j=2,k}$ qui sont deux fois plus nombreuses que celles du niveau 1. On aura donc des matrices de taille $4N$, au troisième niveau on rajoute $\psi_{j=3,k}$ qui sont deux fois plus nombreuses que celles du niveau 2, on aura donc des matrices de taille $8N$, etc.

Avec la multirésolution on cherche à approximer grossièrement les zones uniformes et plus précisément les zones où le champ varie rapidement (i.e. on s'intéresse à obtenir une approximation très précise autour des points où la fonction $\varepsilon(x)$ est discontinue). On n'a pas besoin

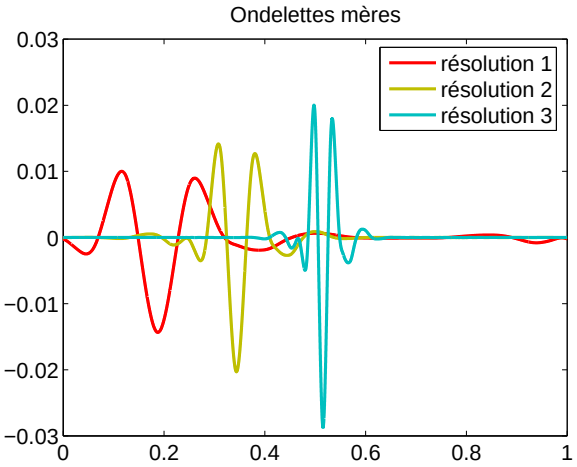


FIGURE 3.10 – *Ondelettes mères aux résolution 1 ; résolution 2 et résolution 3.*

des détails dans les zones avec un ε constant et on peut les décimer en utilisant une fonction de seuillage. Pour une résolution j fixée, cette fonction ne va retenir que les ondelettes analysantes qui vont couper la discontinuité.

$$s(k) = 1 \Leftrightarrow \left\{ \int \varepsilon(x)\psi_{j,k}(x) \neq 0 \right\} \quad s(k) = 0 \quad sinon \tag{3.65}$$

Puisque les ondelettes analysantes sont de moyenne nulle, les intégrales de l'équation (3.65) sont nulles dans les zones où la fonction de permittivité est constante. Grâce à cette propriété on ne perd pas d'information utile quand on fait la décimation des lignes et des colonnes dans les matrices de produits scalaires qui correspondent aux zones homogènes.

La figure 3.11 montre, à titre d'exemple, la décimation des lignes et colonnes pour le cas de la matrice $\boldsymbol{G}^{\varepsilon}$. Cette figure représente la forme d'une matrice des produits scalaires avec trois niveaux de résolution. Au niveau de résolution 0, on retient toutes les produits scalaires (ce sont les produits scalaires entre les ondelettes père et la fonction de permittivité). Les niveaux 1 et 2 font intervenir dans les calculs les ondelettes mère. Grâce au fait que ces ondelettes sont de moyenne nulle ainsi qu'aux propriétés d'orthogonalité, les intervalles où la fonction $\varepsilon(x)$ est une constante vont générer des produits scalaires nuls. On peut faire une décimation des fonctions de base (et de test) qui ne traversent pas les discontinuités de la fonction $\varepsilon(x)$ en utilisant la fonction de seuillage $s(k)$ décrite par l'équation 3.65. On fait supprimer des lignes et de colonnes de la matrice $\boldsymbol{G}^{\varepsilon}$ sans perdre l'information nécessaire pour la suite. Sur la figure 3.11 les "régions" délimitées par les rectangles verts seront retenus dans les calculs. Ce sont les k pour lesquels $s(k) \neq 0$. Les coefficients se retrouvant dans les "régions" jaunes du dessin

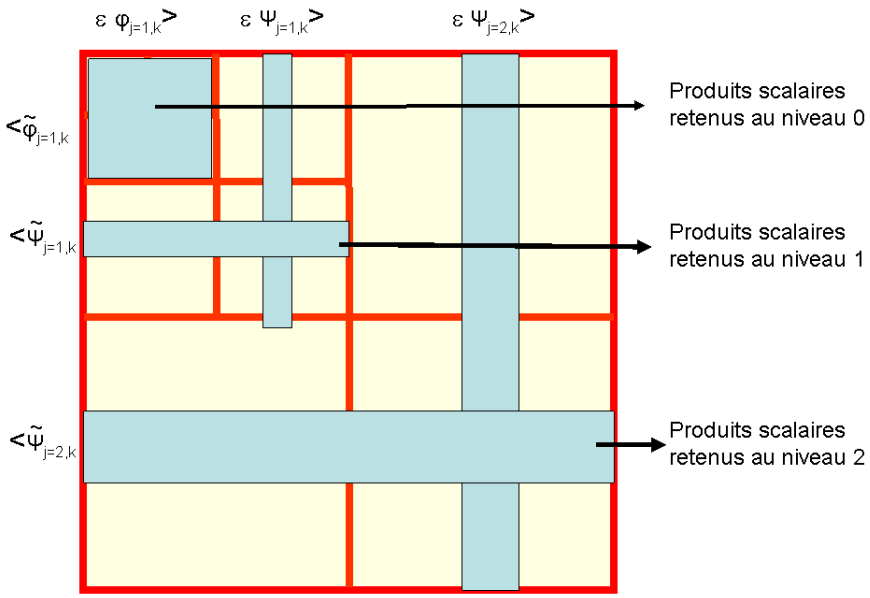


FIGURE 3.11 – *Décimation des lignes et colonnes dans les matrices des produits scalaires.*

seront décimés (dans ces régions $s(k) = 0$ i.e. les produits scalaires seront nuls).

3.4 Résultats

Ce paragraphe traite plusieurs cas de réseaux qui mettent en difficulté ou même en échec la méthode modale avec les bases usuelles et en particulier la MMFE. Pour cette raison, dans la plupart des exemples numériques on ne peut pas obtenir une valeur de référence.

3.4.1 L'influence de la longueur des filtres

Le paramètre de régularité est donné par la longueur des filtres (L_f). Pour illuster l'influence de ce paramètre sur la vitesse de convergence on a choisi les ondelettes splines biorthogonales quadratiques définies en [94] et connues aussi sous le nom de CDF[6] en [87]. Ces ondelettes

6. CDF vient des acronymes des chercheurs qui les ont construits : Cohen, Daubechies et Fauveau. Ces ondelettes ont été beaucoup utilisées dans la pratique : pour la compression d'images avec JPEG2000 ou pour la

possèdent des propriétés intéressantes. Le lecteur pourra trouver plus de détails sur leurs propriétés dans l'article de Unser [96] *"Ten good reasons for using spline wavelets"*. Les ondelettes splines sont symétriques, ont le support le plus petit pour un nombre de moments nuls donné. Les filtres correspondant aux ondelettes père φ de l'espace direct (donnés par la formule 3.32) sont $h = [1\ 3\ 3\ 1]$ à un facteur de normalisation près. La normalisation choisie pour ces travaux est celle standard [7], utilisée dans Wavelab850 [8]. Dans ce cas les valeurs de h sont :

$$h = [0.17677669529664\ \ 0.53033008588991\ \ 0.53033008588991\ \ 0.17677669529664]$$

Les filtres g_n correspondant aux ondelettes analysantes,ψ, sont également symétriques et en fonction du nombre des moments nuls de ψ, la longueur de leur support varie. Pour un nombre de moments nuls fixé, m, on aura des filtres de longueur $2m$. La figure 3.12 montre des ondelettes analysantes à 2, 6 et 10 moments nuls. La longueur des filtres g, appelée par la suite L_g, va être égale à 4, 12 et 20 respectivement. On peut remarquer qu'en ajoutant des moments nuls à l'ondelette analysante, son support devient plus grand. Il est intéressant de travailler avec

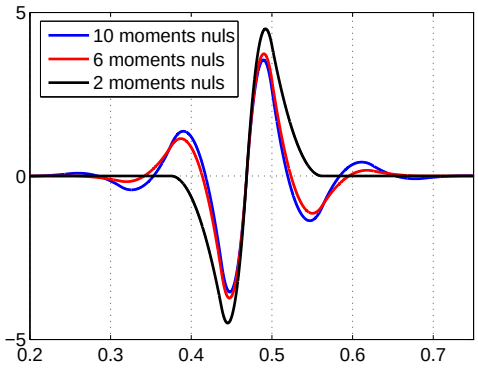


FIGURE 3.12 – *Ondelette mère d'une spline quadratique à divers moments nuls.*

des ondelettes possédant un grand nombre de moments nuls. Comme la formule 3.66, le montre, les filtres de l'espace dual $\widetilde{V}$ associés aux ondelettes père et mère sont corrélés aux filtres de l'espace direct g et h.

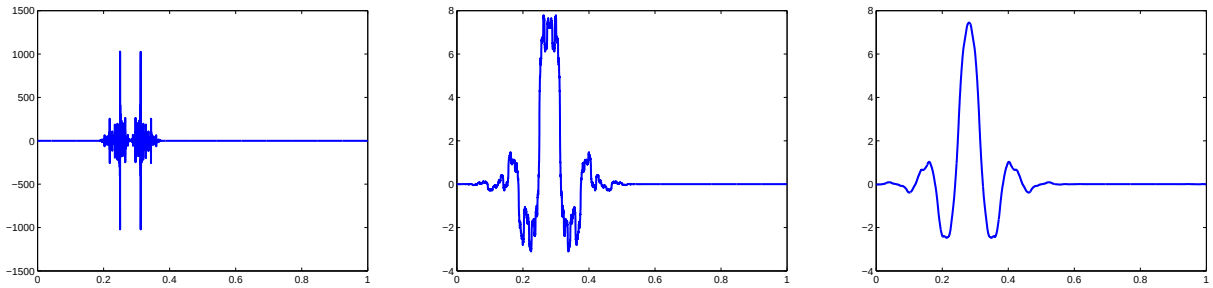
$$\widetilde{h}_n = -(-1)^n g_{L_g-1-n}\ et\ \widetilde{g}_n = (-1)^n h_{L_h-1-n}. \tag{3.66}$$

compression avec des pertes.

7. Cette normalisation est basée sur le fait que $\int \varphi_{j,k}(x)\varphi_{j,k}(x)dx = 1$.

8. WaveLab est une bibliothèque complète de fonctions MATLAB élaborée à l'université de Stanford [87]. Elle constitue une base de référence technique pour l'utilisation des ondelettes dans différents domaines afin de construire des applications spécifiques.

Ajouter des moment nuls à l'ondelette analysante de l'espace direct a une conséquence sur la régularité de la fonction d'échelle de l'espace dual. La figure 3.13 montre les fonctions d'échelle $\widetilde{\varphi}$ associées aux filtres $\widetilde{h}$. Les coefficients des ces filtres sont listés dans le tableau 3.1. Le prix à

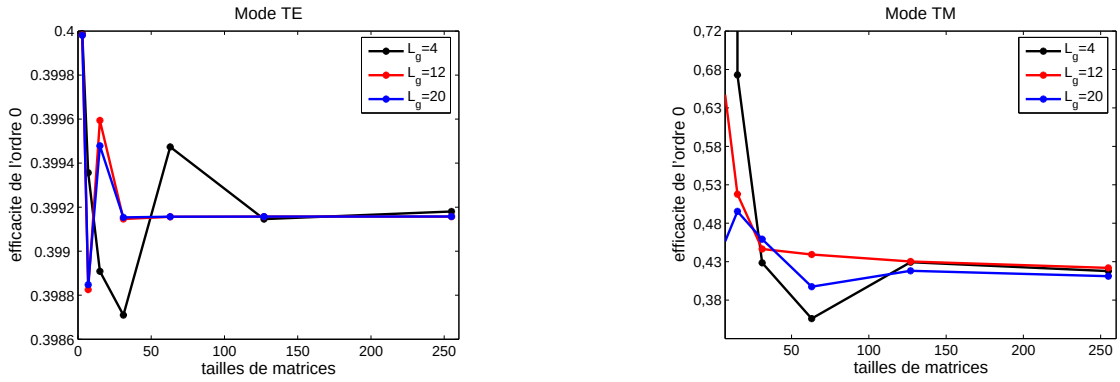


(a) $L_g = 4$ *(b)* $L_g = 10$ *(c)* $L_g = 20$

FIGURE 3.13 – *Les fonctions* $\widetilde{\varphi}$ *correspondant aux filtres de longueur* L_g *du tableau 3.1.*

payer pour des $\widetilde{\varphi}$ plus régulières est un support plus étendu.

Afin d'illustrer l'impact du paramètre de régularité sur la vitesse de convergence on a choisi l'exemple d'un réseau avec les paramètres : $\lambda/d = 0.1$, $h/d = 0.1$, $f = 0.90$ et $\theta = 10°$. Le milieu incident est l'air ($\varepsilon_1 = \varepsilon_{21} = 1$) et les permittivités des matériaux du réseaux sont $\varepsilon_{22} = \varepsilon_3 = 25$. Les figures 3.14(a) et 3.14(b) montrent que la convergence de l'efficacité de l'ordre spéculaire pour ce réseau.



(a) -Polarisation TE *(a) -Polarisation TM*

FIGURE 3.14 – *Convergence de l'efficacité pour un réseau des paramètres :* $\theta = 10°$, $\lambda/d = h/d = 0.1$, $f = 0.9$, $\varepsilon_1 = \varepsilon_{21} = 1$, $\varepsilon_{22} = \varepsilon_3 = 5$.

Comme on peut l'observer dans ces figures, les résultats deviennent plus stables au fur et à

$\widetilde{h}$ pour $L_g = 4$	$\widetilde{h}$ pour $L_g = 12$	$\widetilde{h}$ pour $L_g = 20$
-0.35355339059327	-0.01381067932005	-0.00067974437278
1.06066017177982	0.04143203796015	0.00203923311835
1.06066017177982	0.05248058141619	0.00506031921961
-0.35355339059327	-0.26792717880897	-0.02061891264111
	-0.07181553246426	-0.01411278793018
	0.96674755240348	0.09913478249423
	0.96674755240348	0.01230013626942
	-0.07181553246426	-0.32019196836078
	-0.26792717880897	0.00205002271157
	0.05248058141619	0.94212570067821
	0.04143203796015	0.94212570067821
	-0.01381067932005	0.00205002271157
		-0.32019196836078
		0.01230013626942
		0.09913478249423
		-0.01411278793018
		-0.02061891264111
		0.00506031921961
		0.00203923311835
		-0.00067974437278

TABLE 3.1 – *Les filtres des ondelettes CDF.*

mesure que le nombre des moments nuls d'ondelette analysante ψ croît, donc au fur et à mesure que les fonctions génératrices $\widetilde{\varphi}$ deviennent plus régulières.

3.4.2 Ajouter un niveau de détail

Les ondelettes utilisées dans ce paragraphe sont les symlets. On a fait ce choix puisque ces ondelettes, étant orthogonales, sont plus facile à manipuler par rapport aux ondelettes B-splines. Deuxièment, on a observé que parmi les autres familles d'ondelettes orthogonales, les symlets sont celles qui fournissent les meilleurs résultats. Grâce aux propriétés d'orthogonalité, l'espace de développement V_j et l'espace de projection $\widetilde{V}_j$ coïncident. Les fonctions de base et les fonctions test coïncident aussi ($\varphi = \widetilde{\varphi}$ et $\psi = \widetilde{\psi}$). Comme dans le cas des ondelettes

spline, les symlets sont des ondelettes avec un support minimal pour un nombre de moments nuls donné. Elles gardent l'allure la plus symétrique [9] possible. Le nombre de moments nuls peut être entre 4 et 10, avec 4 on obtient les ondelettes symlets les moins régulières, mais sur le support le plus petit. Pour 10 moments nuls on obtient les symlets les plus régulières. Le nombre des moments nul est égal à la moitié du nombre des unités de longueur compris dans le support de la fonction. Pour l'exemple pris en considération, dans cette section on a choisi les symlets à dix moments nuls (ceci donne des symlets sur un support des 20 unités de longueur). Dans ce cas les coefficients de filtres sont :

0.001089170447	0.000135245020	-0.012220642630	-0.002072363923	0.064950924579
0.016418869426	-0.225558972234	-0.100240215031	0.667071338154	1.088251530500
0.542813011213	-0.050256540092	-0.045240772218	0.070703567550	0.008152816799
-0.028786231926	-0.001137535314	0.006495728375	0.000080661204	-0.000649589896

L'exemple pris en compte est celui d'un réseau métallique, déjà étudié dans le chapitre précedent 2 et par plusieurs chercheurs en [11, 35, 72, 73, 79]. On cherche à approximer l'efficacité de l'ordre -1 en utilisant les ondelettes symlets.

Dans un premier temps on utilise un seul niveau de détail composé par les ondelettes père (qui donnent l'approximation) et les ondelettes analysantes (qui donnent la première génération des détails). La figure 3.15 illustre les ondelettes père et mères pour le niveau de détails.

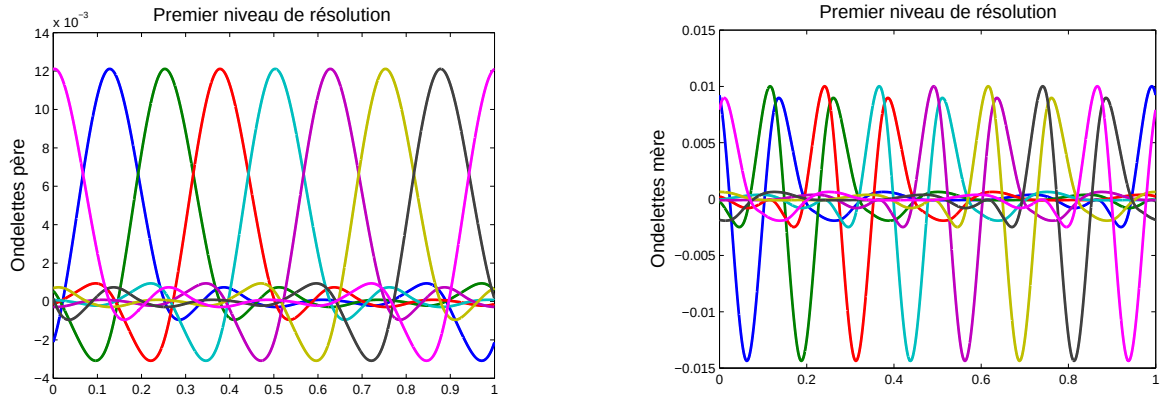


FIGURE 3.15 – *Symlets : (a)-la fonction d'échelle φ ; (b) l'ondelette analysante ψ.*

Le tableau 3.2 est composé de 3 colonnes : la première (J) désigne le paramètre de résolu-

9. Il n'existe pas d'ondelettes à support compact dans une AMR orthogonale qui soient symétriques. L'exception est l'ondelette de Haar qui est antisymétrique.

J	N	R_{-1} à L=14
1	3	0.1165
2	7	0.5605
3	15	0.8761
4	31	0.8535
5	63	0.8489
6	127	0.8486

TABLE 3.2 – *1 niveau de résolution ; polarisation TM : Convergence de l' l'efficacité de l'ordre -1 de diffraction pour différents ordres de troncature. le cas de pour une structure métallique :* $\theta = 30^{\circ}$, $d = 1$, $\lambda = 1$, $h = 1$, $\varepsilon_{21} = 1$, $\varepsilon_3 = \varepsilon_{22} = (0.22 - i6.71)^2$.

tion, la deuxième (N) désigne la taille des matrices des produits scalaires ($\boldsymbol{G}$, $\boldsymbol{D}$, $\boldsymbol{G}^{\varepsilon}$ et $\boldsymbol{G}^{\frac{1}{\varepsilon}}$) et la troisième désigne la convergence de l'efficacité (R_{-1}) l'ordre -1. Notons que le paramètre J influence la taille des matrices. Pour un J fixé, on a 2^J ondelettes père et 2^J ondelettes analysantes. Pour cet exemple, au niveau de résolution 1, on prend en compte dans nos calculs les 2^J ondelettes père et les premières $2^J - 1$ ondelettes analysantes ce qui nous amène à travailler avec une matrice de taille $N = 2^{J+1} - 1$. Remarquons que les valeurs de R_{-1} tendent vers la valeur de référence (qui est 0.84848). La convergence est lente et on a l'inconvénient de travailler avec des matrices qui ont des tailles correspondant à une puissance de 2. Il est difficile à construire des graphiques d'erreur qui comparent la convergence des ondelettes avec celle des méthodes classiques qui utilisent des tailles de matrices correspondant aux nombres impairs. Par exemple, la MMFE utilise des tailles de matrices $2M + 1$, $M \in \mathbb{N}$ étant le nombre de troncature. Cet inconvénient peut être levé en utilisant plusieurs générations de détails. Par exemple, si on utilise deux niveaux de détails, on rajoute dans la base les ondelettes analysantes du niveau 2, $\psi_{j=2,k}$, (voir l'équation 3.44). Même si ces ondelettes sont deux fois plus nombreuses que leurs "ancêtres" $\psi_{j=1,k}$, on n'utilise que celles qui traversent les discontinuités. L'intégrale de l'équation 3.65 est nulle sur les intervalles où la fonction de permittivité est constante. On ne retient donc que les positions k des ondelettes pour lesquelles l'intégrale est non nulle. La figure 3.16 illustre les ondelettes analysantes qui restent dans la décomposition pour assurer le zoom sur les détails de résolution $j = 2$.

Le tableau 3.3 montre la convergence de l'ordre -1 en fonction des tailles des matrices en utilisant deux niveaux de résolution. Le calcul des efficacités est fait pour deux valeurs différentes de L : $L = 13$ et $L = 16$. Le paramètre L représente le nombre des points qui sont utilisés dans la définition d'une ondelette. Comme on peut le voir dans ce tableau les valeurs de R_{-1} convergent plus vite lorsque L croît. Ces différences sont dues au fait que les ondelettes bougent

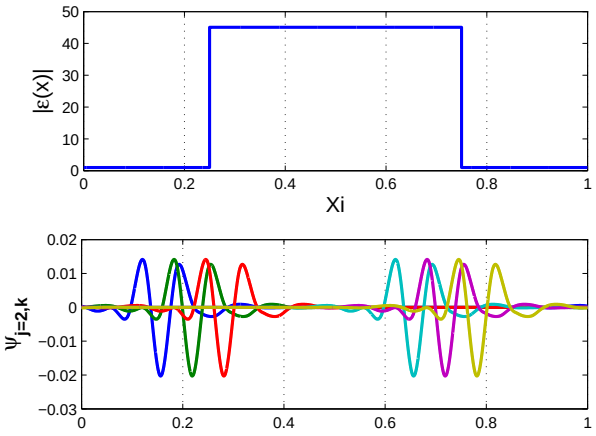


FIGURE 3.16 – *En haut le module de la fonction de permittivité $\varepsilon(x)$. En bas, les ondelettes analysantes qui assurent le zoom au niveau de résolution 2.*

légèrement en fonction de L (voir la figure 3.17). Les différences entre les R_{-1} calculées avec $L = 13$ et $L = 16$ sont plus prononcées lorsqu'on a des petites troncatures.

Les tableaux 3.4 montrent la convergence de l'ordre -1 en utilisant 3, 4 et 5 niveaux de résolution. On peut toujours observer des différences entre le calcul de l'efficacité à $L = 13$ et $L = 16$ à cause du phénomène illustré dans la figure 3.17. Ces différences sont plus estompées quand on utilise plusieurs étages de résolution. Si on compare par exemple les valeurs obtenues avec trois étages de résolution, on peut voir que, pour des ordres de troncature assez grands ($N \geq 33$), la différence se joue à la troisième décimale tandis qu'en utilisant 5 niveaux de ré-

J	N	R_{-1}	
		L=13	L=16
1	7	0.5606	0.5603
2	13	0.8758	0.8762
3	31	0.8534	0.8536
4	51	0.8490	0.8489
5	83	0.8488	0.8485

TABLE 3.3 – *2 niveaux de résolution ; polarisation TM : Convergence de l' l'efficacité de l'ordre -1 de diffraction pour différents ordres de troncature. le cas de pour une structure métallique : $\theta = 30^{\circ}$, $d = 1$, $\lambda = 1$, $h = 1$, $\varepsilon_{21} = 1$, $\varepsilon_3 = \varepsilon_{22} = (0.22 - i6.71)^2$.*

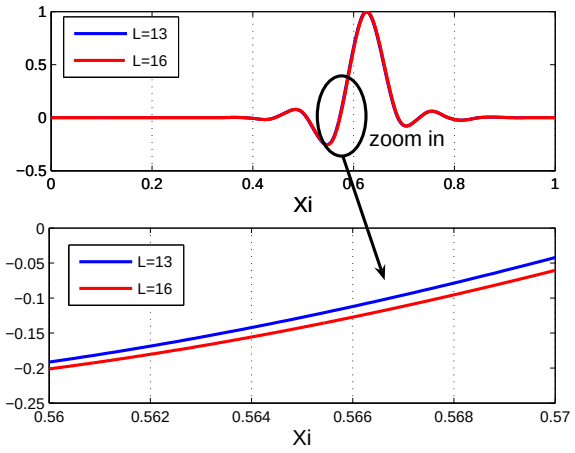


FIGURE 3.17 – *Différence entre une ondelette définie avec $L = 13$(en bleu) et une ondelette avec $L = 16$ (en rouge)*

solution la différence se joue à la quatrième décimale. Si on veut plus de précision il suffit juste d'utiliser encore plus de niveaux de détail. Le prix à payer est des troncatures plus élevées puisqu'au fur et à mesure qu'on descend d'un niveau de résolution les produits scalaires contiennent des détails de plus en plus fins.

Du point du vue pratique, on cherche à obtenir des convergences jusqu'à la troisième décimale. Si on s'intéresse à la vitesse de convergence en fonction de tailles de matrices pour les trois premières décimales exactes, les meilleures convergences sont obtenues en utilisant trois niveaux de résolutions (avec $L = 16$). En effet, en utilisant un seul niveau de résolution on ob-

J	N	R_{-1}	
		L=13	L=16
1	13	0.6671	0.6744
2	19	0.8292	0.8177
3	33	0.8492	0.8485
4	49	0.8497	0.8486

3 niveaux de résolution.

J	N	R_{-1}	
		L=13	L=16
1	13	0.7245	0.7247
2	21	0.8494	0.8476
3	37	0.8496	0.8482
4	53	0.8496	0.8485

4 niveaux de résolution.

J	N	R_{-1}	
		L=13	L=16
1	27	0.6472	0.6536
2	43	0.8344	0.8376
3	63	0.8484	0.8485
4	87	0.8481	0.8485

5 niveaux de résolution.

TABLE 3.4 – *Polarisation TM : Convergence de l' l'efficacité de l'ordre -1 de diffraction pour différents ordres de troncature. le cas de pour une structure métallique : $\theta = 30^{\circ}$, $d = 1$, $\lambda = 1$, $h = 1$, $\varepsilon_{21} = 1$, $\varepsilon_3 = \varepsilon_{22} = (0.22 - i6.71)^2$.*

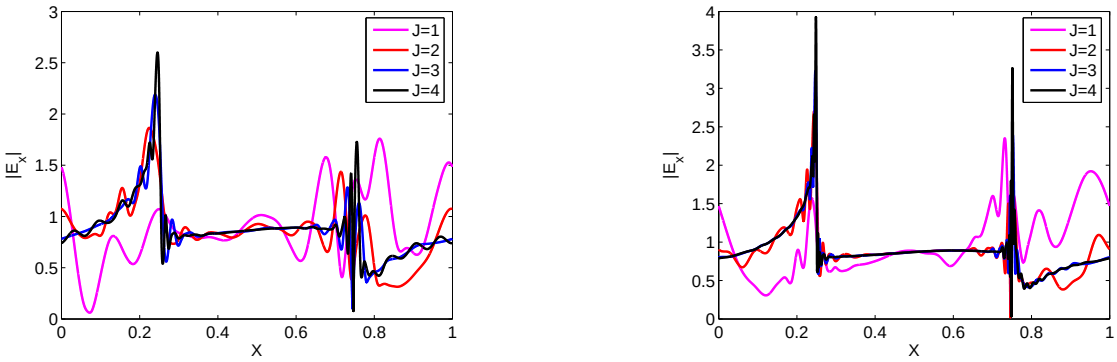
tient cette convergence à $N = 127$. Quand on rajoute un niveau de détail, cette convergence est obtenue à $N = 51$, puis avec trois niveaux de détail on diminue encore la troncature à $N = 33$. En utilisant 4 ou 5 niveaux de résolution, on peut observer que la convergence de la troisième décimale est obtenue à $N = 37$ et $N = 63$. Ce phénomène de ralentissement de la vitesse de convergence en fonction de la taille des matrices est dû au fait qu'en utilisant beaucoup de niveaux de résolution on rajoute des lignes et colonnes supplémentaires aux matrices $\mathbf{D}$, $\mathbf{G}^{\varepsilon}$ et $\mathbf{G}^{\frac{1}{\varepsilon}}$ avec des produits scalaire contenant des détails de plus en plus fins.

Si on cherche une convergence rapide jusqu'à une précision donnée, une des limitations de la méthode est que le nombre des niveaux de résolution "optimal" diffère en fonction du cas traité. On ne peut pas savoir à l'avance à quel niveau de résolution on doit s'arrêter pour obtenir la convergence la plus rapide.

Les figures 3.18 illustrent l'allure de la deuxième composante tangentielle (E_x) en utilisant plusieurs niveaux de détails. La première remarque est que, si on regarde une figure particulière (3.18(a) par exemple) on observe que le champ et décrit de plus en plus précisement lorsqu'on augmente J. Ce comportement est tout à fait normal, puisque J est relié à la taille des matrices et quand on augmente la taille des matrices le champ est mieux décrit. Analysons la figure 3.18(a). Pour $J = 1$ on utilise des matrices de taille $N = 7$. La taille des ces matrices est donnée par : 2 fonctions père ($\varphi_{1,k=1}$ et $\varphi_{1,k=2}$), 2 fonctions mère de la résolution 1 ($\psi_{1,k=1}$ et $\psi_{1,k=2}$) et 3 fonctions mère de la résolution 2 ($\psi_{2,k}$, $k = 1, 2, 3$). L'information grossière étant portée par deux fonctions père on obtient un très mauvais résultat sur l'allure de la composante E_x (la courbe en magenta). La courbe en rouge correspond à $J = 2$. Le champ est décrit par des matrices de taille $N = 13$ on peut voir une amélioration par rapport à la première. Cette amélioration vient du fait qu'on utilise plus de fonctions père (2^2 plus précisément) qui encodent l'information grossière. La courbe en bleu correspond à $J = 3$ ($N = 31$). La courbe en noir correspond à $J = 4$ (ce qui fait $N = 51$: 2^4 ondelettes père, 2^4 ondelettes analysantes au niveau 1 et 19 ondelettes analysantes du niveau 2). On peut observer que cette courbe offre la meilleure approximation du champ. Si on augmente J on aura une approximation encore plus précise, mais la taille des matrices devient $N = 81$.

Deuxième remarque : pour des J égaux, on peut observer que les oscillations du champ sont atténuées au fur et à mesure qu'on rajoute un niveau de résolution supplémentaire. Si on regarde la courbe en noir (correspondant à $J = 4$) dans les deux graphiques 3.18(a) et 3.18(b) on observe que les oscillations du champ sont de plus en plus atténuées, même si les tailles des matrices sont approximativement égales : pour deux niveaux de résolution on a $N = 51$ et pour 4 niveaux $N = 53$. Les atténuations des oscillations sont dues à l'utilisation des ondelettes

analysantes sur plusieurs niveaux de détail. Ces ondelettes décrivent les variations rapides du champ donc elles jouent un rôle décisif quand on veut gommer les oscillations.



(a) 2 étages de résolution. (b) 4 étages de résolution.

FIGURE 3.18 – *Le module du champ E_x décrit par plusieurs niveaux de résolution.*

Pour conclure, ajouter des niveaux de détail nous permet d'avoir une flexibilité supplémentaire par rapport aux méthodes classiques. Introduire la multirésolution aide à minimiser le phénomène de Gibbs. En plus, on obtient un raffinement numérique important à des ordres de troncature faibles. Cela est valable pour la convergence de l'efficacité et également pour la description des composantes du champ.

3.4.3 Comparaison avec la littérature

L'utilisation d'ondelettes est particulièrement intéressante quand on a à traiter des cas où le facteur de remplissage est très petit (où très grand) par rapport à la période. Cette situation peut représenter par exemple une ligne quasi isolée. Les difficultés s'ajoutent si ce type de géométrie se présente conjointement avec des matériaux de permittivité très différentes. Dans ces conditions, les méthodes couramment utilisées dans la simulation de la diffraction par des réseaux, comme la MMFE échouent car les matrices deviennent mal conditionnées. Nous illustrons ce phénomène en nous appuyant sur l'exemple de trois réseaux de lignes dont le produit $f\varepsilon_{22} = 2.5$ est maintenu constant. Les autres paramètres sont : $\lambda/d = 0.1$, $h/d = 0.1$, $\theta = 10°$, $\varepsilon_1 = \varepsilon_{21} = 1$ et $\varepsilon_3 = \varepsilon_{22}$. Maintenir constante la valeur de $f\varepsilon_{22}$, nous permet de conserver la valeur moyenne de la permittivité du réseau. En faisant diminuer que f, sans augmenter $f\varepsilon_{22}$ donc le contraste, on pourrait atteindre une configuration où les méthodes numériques n'arrivent plus à " distinguer " le réseau. Les trois couples de valeurs sont les suivants : $f = 0.3$ et $\varepsilon_{22} = 25/3$;

$f = 0.2$ et $\varepsilon_{22} = 25/2$; $f = 0.1$ et $\varepsilon_{22} = 25$. La comparaison de la convergence des efficacités de diffraction entre la MMFE et les ondeltettes est illustrée par les figures 3.19(a), 3.19(b) et 3.19(c) pour les deux cas de polarisation.

Les courbes en pointillée sont les résultats fournis par la méthode MMFE, tandis que les courbes en bleu sont obtenus en utilisant les ondelettes splines de degré deux avec un seul niveau de détail (les ondelettes père et les ondelettes mère du premier niveau de résolution). Il apparaît clairement que la méthode MMFE se trouve mise en difficulté au fur et à mesure que le facteur de remplissage devient plus petit par rapport à la période.

On peut aussi remarquer que l'approche utilisant les ondelettes fournit des courbes lisses, sans oscillations. Ce comportement est dû au fait que l'utilisation des ondelettes mère révéle les détails dans les endroits de forte discontinuité tandis que l'utilisation des ondelettes père décrit le champ ailleurs.

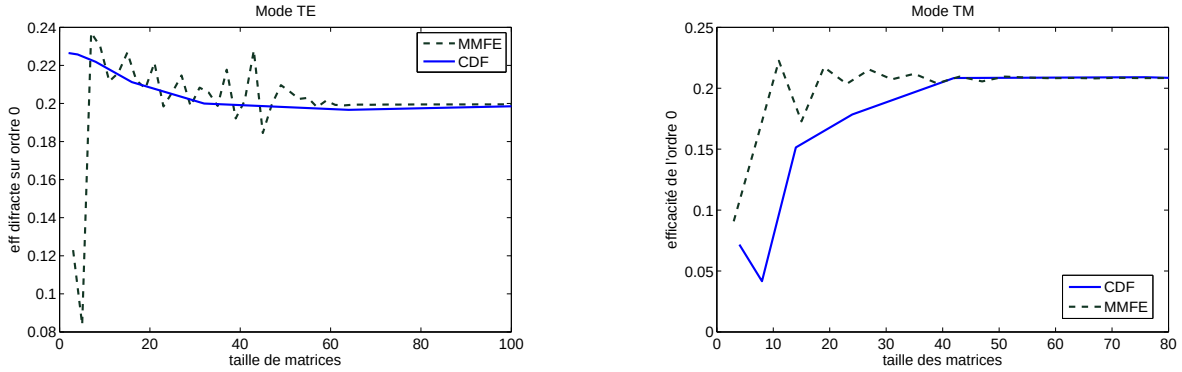
Ce résultat est très encourageant pour l'utilisation des ondelettes. On peut ainsi obtenir des résultats qui caractérisent les erreurs engendrées par un défaut de fabrication ou bien on peut obtenir des résultats de mesure pour des motifs très isolés par rapport à la période. Dans ce genre de cas, les méthodes classiques se trouvent en difficulté.

Des résultats plus précis à des troncatures plus faibles pourraient être obtenus en utilisant plusieurs générations d'ondelettes mère. En effet, dès que le motif devient plus localisé par rapport à la période, les méthodes classiques s'avèrent inefficaces. Dans le même temps, pour les motifs localisés, les ondelettes sont encore plus efficaces. Pour faire un zoom sur le motif isolé on rajoute un très petit nombre d'ondelettes mère dans les régions du saut de la permittivité (il s'agit des ondelettes qui traversent les saut de la fonction $\varepsilon(x)$). Si on a besoin de résultats encore plus précis, on rajoute une génération d'ondelettes mère. A la génération suivante les ondelettes mère vont doubler leur nombre, mais grâce à la décimation, celui-ci reste assez petit si l'on compare avec le nombre total d'ondelettes.

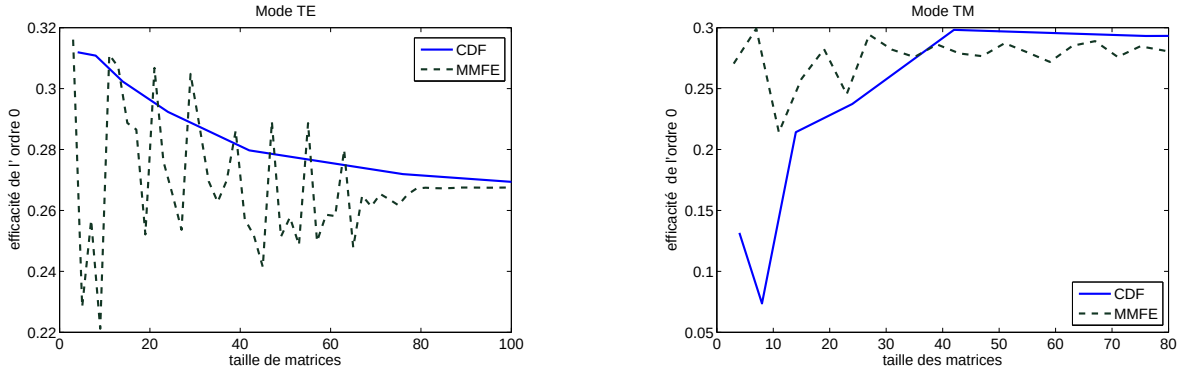
3.5 Bilan

Les ondelettes apparaissent comme une découverte importante en mathématiques appliquées. Dans de nombreux domaines d'applications où l'analyse de Fourier était traditionnellement employée, elles ont ouvert une voie alternative, souvent mieux adaptée à la nature des phénomènes étudiés. Leurs avantages sont liés au fait qu'on peut maîtriser leur position et leur

(a) $f = 0.3$, $\varepsilon_{22} = 25/3$.



(b) $f = 0.2$, $\varepsilon_{22} = 25/2$.



(c) $f = 0.1$, $\varepsilon_{22} = 25$.

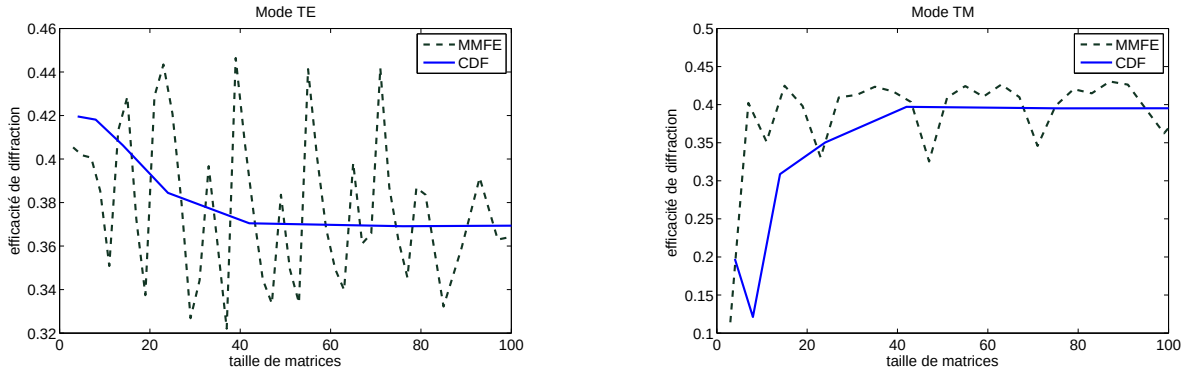


FIGURE 3.19 – *Convergence de l'efficacité pour différents réseaux qui respectent le rapport* $f * \varepsilon_{22} = cst$. *Les autres paramètres sont :* $\theta = 10^{\circ}$, $\lambda/d = h/d = 0.1$, $\varepsilon_1 = \varepsilon_{21} = 1$, $\varepsilon_3 = \varepsilon_{22}$.

étendue spatiale dans le domaine d'étude alors que les bases de Fourier ont pour support le domaine tout entier. Comme on a pu le constater dans ce chapitre, les ondelettes se prêtent bien à l'analyse des réseaux. Les ondelettes fournissent de bons résultats dans les cas où on doit traiter de fortes discontinuités du champ. Ces cas mettent souvent en difficulté les méthodes classiques.

Les programmes de simulation ont été conçus de façon à permettre le changement des fonctions test et des fonctions de base sans modifier le noyau du code. Il suffit de modifier les valeurs des coefficients des filtres pour travailler avec des ondelettes différentes. On a pu ainsi aisément étudier la convergence de notre méthode en fonction de divers type d'ondelettes.

Une première étude a consisté à vérifier l'impact de la longueur de filtres sur la vitesse de convergence. Pour cela, on a choisi comme couple de fonctions de base et test les ondelettes splines de degré 2. On a observé que les filtres les plus longs fournissent les meilleures convergences. Ceci est dû au fait que les filtres plus longs génèrent des ondelettes plus régulières. Un autre résultat intéressant, non illustré dans ce rapport, nous a permis de constater que la convergence s'améliore en augmentant le degré des ondelettes splines. On a également observé des phénomènes de convergence en dent de scie lorsque l'on utilise les ondelettes splines de degré 1. Ceci est explicable par le fait que ces fonctions ont des dérivées discontinues et pour les réseaux lamellaires, en fonction de l'ordre de troncature, les discontinuités de la fonction de permittivité peuvent coïncider ou non avec les discontinuités des dérivés des splines de degré 1. Il faudrait donc concevoir des codes de simulations avec des contraintes supplémentaires pour éliminer ce problème. Cependant, sur des cas très simples, on a observé qu'en augmentant le degré de l'ondelette spline, les convergences s'améliorent. Les splines de degré deux semblent un bon compromis entre la simplicité du code et son efficacité.

Une autre observation intéressante est qu'il existe une corrélation entre la propriété de la symétrie des fonctions de base et la vitesse de convergence. Plus précisément, les meilleures convergences sont obtenues quand les fonctions sont symétriques. De même dans le cas des ondelettes orthogonales, qui ne sont pas symétriques, les meilleurs résultats sont obtenus par les ondelettes symlets qui " ressemblent " le plus à des ondelettes symétriques.

D'autres investigations, comme le changement de divers paramètres opto-géométriques ainsi que la comparaison avec les méthodes classiques utilisées dans la résolution des équations de Maxwell, nous ont permis d'avoir une image globale sur la méthode modale de décomposition de ondelettes. En général, d'un point de vue pratique, on peut considérer comme satisfaisante la vitesse de convergence pour tout type de réseau. La comparaison de notre méthode avec la

MMFF ou avec la FDMM, a montré que l'utilisation des ondelettes peut être une solution intéressante quand il s'agit de cas métalliques, avec de grands contrastes d'indices. Pour ce type de réseaux, la décomposition en ondelettes s'avère être plus performante que les méthodes classiques. Il serait pour autant faux de croire que les ondelettes sont la meilleure solution à tous les problèmes. Lorsque les variations d'indices sont petites et le facteur de remplissage de l'ordre d'une demi période, numériquement parlant, la MMFE reste la plus efficace puisqu'elle fournit des convergences rapides avec très peu d'harmoniques.

Conclusion et perspectives

Les progrès de la technologie permettent la miniaturisation de plus en plus poussée des circuits intégrés. En même temps apparaissent des besoins nouveaux en métrologie pour contrôler les processus de fabrication. En pratique, on grave des motifs test qui sont souvent des réseaux de lignes périodiques dont on estime les dimensions par des mesures optiques d'intensité ou ellipsométriques. Les dimensions sont déduites de la comparaison entre les spectres mesurés des spectres calculés. Le besoin de modèles théoriques de simulation de la réponse spectrale des divers objets fabriqués se fait de plus en plus ressentir surtout quand on se place dans le domaine résonant. Parmi les méthodes numériques de calcul électromagnétique de diffraction à travers des réseaux, la méthode la plus couramment utilisée est la Méthode Modale de Fourier. Les problèmes intrinsèques à cette méthode, comme le phénomène de Gibbs, s'avèrent d'autant plus indésirables qu'on diminue la taille des transistors.

Notre objectif dans le cadre de cette thèse a été d'explorer une nouvelle voie en utilisant des bases de développement différentes qui ne présentent pas les inconvénients des bases de Fourier. Nous nous sommes inspirés des travaux effectués dans de nombreux domaines comme le traitement de signal, le traitement d'images, la compression des données où les ondelettes et l'analyse multi-résolution constituent maintenant une solution alternative à l'analyse de Fourier. Nous avons utilisé les ondelettes et l'analyse à plusieurs niveaux de détails pour la résolution, par une méthode modale, du problème de la diffraction par un réseau.

Ces outils de calcul de l'analyse moderne ont été appliqués dans de nombreux domaines comme le traitement de signal, traitement d'images, compression des données, etc. Les ondelettes et l'analyse multirésolution constituent maintenant une solution alternative, souvent mieux adaptée pour des problèmes où l'analyse de Fourier était employée.

Pour mieux comprendre l'impact de l'analyse à plusieurs niveaux de détails, nous avons tout d'abord fait une analyse à un seul niveau en utilisant des bases à support compact et plus précisément des fonctions B-spline de différents degrés (0,1, 2 et 3). Cette étude nous a tout

d'abord montré que ces fonctions sont bien adaptées pour le problème de la diffraction par des réseaux. On a montré que la méthode de Galerkin fournit les meilleurs résultats en terme de convergence des valeurs propres, d'efficacité ou de description des composantes du champ électromagnétique. On a aussi observé qu'on obtient des meilleurs résultats en utilisant la méthode de Galerkin et en augmentant le degré de spline. La position de la grille de discrétisation par rapport aux sauts de permittivité est un autre facteur qui influence les vitesses de convergence. Plus précisément, on obtient de meilleurs résultats si les sauts de discontinuité interceptent un nombre maximal de fonctions de base. Le même comportement a été observé pour les splines de degré 1, 2 et 3, mais les résultats illustrés dans ce manuscrit se limitent aux splines de degré 2. Un des inconvénients de l'utilisation des fonctions splines est que le champ est mal décrit dans les zones où le champ varie brusquement. Cet inconvénient - déjà observé dans des problèmes d'interpolation - a été corrigé grâce à l'utilisation d'une résolution spatiale adaptative qui resserre les points de discrétisation autour des discontinuités. On gomme les oscillations du champ grâce à cette technique tout en améliorant la vitesse de convergence.

La deuxième partie du travail s'est concentrée sur l'utilisation d'ondelettes. Parmi les nombreuses familles d'ondelettes biorthogonales on a orienté notre choix vers les ondelettes splines, connues sous le nom de CDF (d'après les initiales de leur inventeurs : Cohen, Daubechies et Feauveau). Ces ondelettes ont déjà fait leurs preuves dans le domaine de visualisation des données en temps réel, la compression des données, etc. Ces ondelettes ont des propriétés intéressantes comme la formulation explicite, la régularité, la symétrie ou l'antisymétrie. L'analyse multirésolution se construit à partir des fonctions B-spline. Les résultats fournis dans le chapitre 3 utilisent les ondelettes biorthogonales à un seul niveau de détail. On utilise donc des ondelettes mère et des ondelettes père (qui sont des fonctions B-spline) dans notre problème de diffraction. On a observé des résultats encourageants, de bonnes convergences, en particulier dans des cas où l'analyse de Fourier est mise en échec. De plus, grâce à l'utilisation des ondelettes mère on arrive à gommer les oscillations dues à l'utilisation des fonctions splines. Nous avons introduit plusieurs échelles de détail, en nous limitant aux bases d'ondelettes orthogonales. Les résultats rapportés ici utilisent les ondelettes symlet, les ondelettes les plus symétriques possibles. On a observé qu'en introduisant plusieurs niveaux de détails, on arrive à améliorer les performances en terme de convergence et à mieux décrire les discontinuités ou les variations brusques du champ.

Perspectives

Il est vraisemblable que l'on peut encore améliorer les performances de la méthode modale associée à des splines. En effet, mathématiquement parlant, les vecteurs propres sont des solutions de l'équation de propgation qui satisfont la condition de pseudo-périodicité et des conditions de continuité au bord des domaines définis par la fonction epsilon. Ces conditions ne sont pas exprimées explicitement mais sont portées par le schéma numérique. Ainsi, nous avons vu que la position de la discontinuité de epsilon par rapport aux splines influait beaucoup sur la vitesse de convergence. Nous pensons que l'introduction de splines décrites avec des nœuds confondus positionnés sur les discontinuités, permettrait d'améliorer la convergence. En effet, mettre des nœuds confondus construit d'autres splines dont la dérivée n'est pas continue. Or, en polarisation TM, les fonctions que nous recherchons ne sont pas à dérivée continue. Les premiers essais que nous avons fait dans cette voie sont prometteurs.

A la lumière des résultats obtenus à l'aide de l'analyse multirésolution on peut dire que celle-ci est bien adaptée au problème de diffraction par des réseaux. La méthode a beaucoup d'avantages (elle est robuste au changement des fonctions de base, on arrive a obtenir des résultats dans des situations où les méthodes classiques sont en échec). Cependant il existe un problème numérique gênant lié à la transformée rapide en ondelettes. Celle ci est un processus récursif dont la précision globale dépend fortement du calcul d'une valeur initiale dont on obtient une approximation à partir d'échantillons. Si on veut garder la robustesse des codes vis à vis des fonctions de base et de test, on est obligé d'utiliser cette transformation puisqu'on ne connaît pas toujours la forme analytique de ces dernières.

Une façon de contourner ce problème est de se restreindre à des ondelettes définies analytiquement et de faire un maximum de calculs analytiques. Pour cette raison, l'utilisation des ondelettes B-splines devient intéressante. Puisque les ondelettes splines ne sont pas orthogonales, on doit jongler avec deux analyses multirésolution afin de pouvoir arriver à résoudre notre problème. Grâce aux coefficients de filtres on peut obtenir d'une manière précise toutes les valeurs des produits scalaires. La définition analytique aide aussi quand il s'agit de faire la transformation en ondelettes de la fonction de permittivité.

Une autre direction d'amélioration serait d'optimiser la routine de décimation d'ondelettes. Pour l'instant on fait la décimation sur un critère numérique ; on ne garde dans les calculs que les ondelettes mère qui traversent la discontinuité. On pourrait imaginer un critère plutôt physique qui nous permettrais de dire si une ondelette doit rester dans la base ou pas. Par exemple un calcul complet de champ avec un faible nombre de fonctions de base donne déjà

des informations sur celui ci. A leur tour, dans une itération suivante ces informations peuvent éclairer sur les ondelettes à conserver ou pas.

En perspective plus lointaine, une fois les améliorations en 1D mises au point, un programme de simulation des structures en trois dimensions à l'aide d'ondelettes peut s'avérer une alternative efficace aux problèmes de caractérisation des motifs 3D isolées ou très denses.

Bibliographie

[1] C. Saunier. L'évolution du secteur des semiconducteurs et ses liens avec les micro et nanotechnologies. $http://www.senat.fr/rap/r02-138/r02-138_mono.html\sharp toc0$, consulté en juin 2010.

[2] R. Petit and M. Cadilhac. 259. *C. R. Acad. Sci. Paris Vol **17**(6) pp : 2077*, 1964.

[3] A. Wirgin. Theoretical and experimental investigation of a new type of blazed grating. *Rev. Opt. Vol.**9** pp : 449*, 1964.

[4] L. Rayleigh. On the dynamical theory of grating. *Procedings of the Royal Society A Vol.**79** pp : 399-416*, 1907.

[5] L. Rayleigh. On the dynamical theory of grating. *Reprint in Selected Papers on diffraction Gratings Editor D. Maystre SPIE Milestone Series Vol.**MS83** pp : 399-416*, 1993.

[6] R. Petit. Diffraction d'une onde plane par un réseau métallique. *Rev. Opt. Vol **8** pp : 353-370*, 1966.

[7] M. Nevière, R. Petit, and M. Cadilhac. About the theory of optical grating coupler waveguide systems. *J.Opt.Soc. Am. A Vol.**8** pp : 113-117*, 1973.

[8] L.C. Botten, M. Cadilhac, G.H. Derrick, D. Maystre, R.C. McPherdran, M. Nevière, R. Petit, and P. Vincent. *Electromagnetic Theory of Gratings*. R. Petit Edition, 1980.

[9] K. Knop. Rigorous diffraction theory for transmission phase gratings with deep rectangular grooves. *J. Opt. Soc. Am. A Vol.**66** pp : 1206-1210*, 1978.

[10] J. Chandezon, D. Maystre, and G. Raoult. A new theoretical method for diffraction gratings and its numerical application. *J. Opt. (Paris) Vol.***11** *pp : 235-241*, 1980.

[11] P. Lalanne and J.P Hugonin. Numerical performance of finite-difference modal method for the electromagnetic analysis of one-dimensional grating. *J. Opt.Soc.Am./ Vol **17**(6) pp : 1033-1042*, 2000.

[12] M. G. Moharam and T. K. Gaylord. Diffraction analysis of dielectric surface-relief gratings. *J. Opt. Soc. Am. A Vol.* **72** *pp : 1385-1392*, 1982.

[13] M. G. Moharam and T. K. Gaylord. Rigourous coupled-wave analysis of metallic surface-relief gratings. *J. Opt. Soc. Am. A Vol.**3** No.11 pp : 1780-1787*, 1986.

[14] A. Grossmann and J. Morlet. Decomposition of hardy functions into square integrable wavelets of constant shape. *SIAM J. Math. Anal. Vol.**15** pp : 723-736*, 1984.

[15] S. Mallat. A theory for multiresolution signal decomposition : the wavelet representation. *IEE Transf. Patt. Anal. Intell. PAMI Vol.**11**(7) pp : 674-693*, 1989.

[16] S. Mallat. Multiresolution approximation and wavelet. *Trans. Amer. Math. Soc. Vol.**315** pp : 69-88*, 1989.

[17] I. Daubechies. *Orthonormal bases of compactly supported wavelets*. Com. Pure and Appl. Math. XLI pp : 909-996, 1988.

[18] Y. Meyer. *Ondelettes. Ondelettes et opérateurs.* Vol. 1 Hermann Paris, 1990.

[19] D. Maystre. Rigourous vector theories of diffraction gratings. *Progress in Optics(Elsevier North-Holland Amsterdam) Vol.**XXXI** pp : 2-67*, 1984.

[20] R. Petit. *Electromagnetic theory of gratings Topics in current physics*. Springer Verlag Vol.**22**, 1980.

[21] L. Rayleigh. Note on the remarkable case of diffraction spectra described by prof. wood. *Philos.Mag. Vol. **14** page : 60*, 1907.

[22] R.W. Wood. On a remarkable case of uneven distribution of light in a diffraction grating spectrum. *Philos.Mag. Vol. **4** pp : 396-402*, 1902.

[23] D. Gottlieb and S.A. Orszag. *Numerical analysis of spectral methods : Theory and Applications*. Society of Industrial and Applied Mathematics Philadelphia, 1989.

[24] L.C. Botten, M.C. Craig, R.C. McPherdran, L. R. Adams, and J.R. Andrewartha. The dielectric lamellar diffraction grating. *Opt. Acta Vol. **28** pp : 413-428*, 1981.

[25] L.C. Botten, M.C. Craig, R.C. McPherdran, L. R. Adams, and J.R. Andrewartha. The finitely conducting lamellar diffraction grating. *Opt. Acta Vol. **28** pp : 1087-1102*, 1981.

[26] L.C. Botten, M.C. Craig, and R.C. McPherdran. Highly conducting lamellar gratings. *Opt. Acta Vol. **28** pp : 1103-1106*, 1981.

[27] L. Li. Use of fourier series in the analysis of discontinuous periodic structures. *J. Opt. Soc. Am. A (Optics Image Science and Vision) Vol. **13**(9) pp : 1870-1876*, 1996.

[28] L. Li. New formulation of the fourier modal method for crossed surfaces-relief gratings. *J. Opt. Soc. Am. A (Optics Image Science and Vision) Vol. **14**(10) pp : 2758-2767*, 1997.

[29] L. Li. Reformulation of the fourier modal method for surfaces-relief gratings made with anisotropic materials. *J. of Modern Optics Vol. **45**(7)*, 1998.

[30] L. Li. Fourier modal method for crossed anisotropic gratings with arbitrary permittivity and permeability tensors. *J. of Optics A : Pure and Applied Optics Vol.**5**(4) pp : 345-355*, 2003.

[31] G. Granet. Analysis of diffraction by crossed gratings using a non-orthogonal coordinate system. *Pure Appl. Opt. Vol. **4** pp : 777-793*, 1995.

[32] G. Granet and B. Guizal. Efficient implementation of the coupled-wave method for metallic lamellar gratings in tm polarization. *J.Opt.Soc. Am. A Vol. **13** pp : 1019-1023*, 1996.

[33] Ph. Lalanne and G. M. Morris. Highly improved convergence of the coupled-wave method for tm polarization. *J.Opt.Soc. Am. A Vol.**13** pp : 779-784*, 1996.

[34] L. Li. Formulation and comparison of two recursive matrix algorithms for modeling layered diffraction gratings. *J. Opt. Soc. Am. A Vol. **13**(5) pp : 1024-1035*, 1996.

[35] G. Granet. Reformulation of the lamellar grating problem through the concept of adaptive spatial resolution. *J.Opt.Soc.Am. Vol.**16** pp : 2510-2516*, 1999.

[36] B. Guizal, H. Yala, and D. Felbacq. Reformulation of the eigenvalue problem in the fourier modal method with spatial adaptive resolution. *Opt. Lett. Vol.**34**(18) pp : 2790-2792*, 2009.

[37] P. Popov and L. Mashev. Conical diffraction mounting generalization of a rigorous differential method. *J. of Optics Vol.**17** pp : 175-180*, 1986.

[38] L. Li. A modal analysis of lamellar diffraction gratings in conical mountings. *J. Mod. Opt. Vol. **40** pp : 553-573*, 1993.

[39] J. P. Bérenger. A perfectly matched layer for the absorption of electromagnetic waves. *J. Comput. Phys. Vol.***114***(2) pp : 185-200*, 1994.

[40] J. Chandezon, M.T. Dupus, G. Cornet, and D. Maystre. Multicoated grating : a differential formalism applicable in the entire optical region. *J. Opt. Soc. Am. Vol.***72** *pp : 839-846*, 1982.

[41] G. Granet and J. Chandezon. The method of curvilinear coordinates applied to the problem of scattering from surface relief gratings defined by parametric equations : application to scattering from cycloidal grating. *Pure Appl. Opt. Vol.***6** *pp : 727-740*, 1997.

[42] L. Li, J. Chandezon, G. Granet, and J.-P. Plumey. Rigorous and efficient grating-analysis method made easy for optical engineers. *App. Opt. Vol.***38** *pp : 304-313*, 1999.

[43] L. Li. Multilayer coated diffraction gratings : differential method of chandezon et al. revisited. *J. Opt. Soc. Am. A Vol. **11**(6) pp : 2816-2828*, 1994.

[44] J.P. Plumey, G. Granet, and J. Chandezon. Differential covariant formalism for solving maxwell's equations in curvilinear coordonate : oblique scattering from lossy periodic surfaces. *IEEE Trans. Antennas Propag.. Vol. **43***, 1995.

[45] G. Granet, J.P. Plumey, and J. Chandezon. Scattering by a periodically corrugated dielectric layer with non-identical faces. *Pure Appl. Opt. Vol. **4** pp : 1-5*, 1995.

[46] L. Li, G. Granet, J.P. Plumey, and J. Chandezon. Some topics in extenting the c method to multilayer-coated gratings of differents profiles. *Pure Appl. opt. Vol. **5** pp : 141-156*, 1996.

[47] T.W. Preist, N.P.K. Cotter, and J.R. Sambles. Periodic multilayer gratings of arbitrary shape. *J. Opt. Soc. Am. A Vol.**5** pp : 1740-1748*, 1995.

[48] G. Granet. Analysis of diffraction by surface relief crossed gratings with use of chandezon method : application to multilayer crossed gratings. *J.Opt.Soc. Am. A Vol. **15** pp : 1121-1131*, 1998.

[49] J.B. Harris, T.W. Preist, J.R. Sambles, J.N. Thorpe, and R.A. Watts. Optical response of bigrating. *J.Opt.Soc. Am. A Vol. **13** pp : 2041-2049*, 1996.

[50] J.B. Harris, T.W. Preist, and J.R. Sambles. Differential formalism for multilayer diffraction gratings made with uniaxial materials. *J.Opt.Soc. Am. A Vol. **12** pp : 1965-1973*, 1995.

[51] G. E. Forsythe and W. R. Wasow. *Finite-Difference Methods for partial Differential Equations*. John Wiley and Sons Inc. New York, 1960.

[52] J. D. Jackson. *Classical Electrodynamics*. John Wiley and Sons : New York, 1975.

[53] R. Petit. *Ondes électromagnétiques en radioélectricité et en optique*. Masson, 1993.

[54] N. Chateau and J.-P. Hugonin. Algorithm for the rigorous coupled wave analysis of grating diffraction. *J. Opt. Soc. Am. A Vol.**11**(4) pp : 1321-1331*, 1994.

[55] I. J. Schoenberg. Contributions to the problem of approximation of equidistant data by analytic functions. *Quart. Appl. Math. Vol.**4** pp : 45-99 112-141*, 1946.

[56] H. B. Curry and I. J. Schoenberg. On polya frequency functions iv : The fundamental spline approximations. *J. Anal. Math. Vol.**17** pp : 71-107*, 1966.

[57] C. Boor. On uniform approximation by splines. *J. Approximation Theory Vol.**1** pp : 219-235*, 1968.

[58] M. A. Sabin. Parametric splines in tension. *Tecnical Note VTO/MS/160 British Aircraft corporation*, 1970.

[59] P. Bézier. Example of an existing system in motor industry : The unisurf system. *Proc. Royal Soc. of London Vol.**A321** pp : 207-218*, 1971.

[60] P. Casteljau. *Shape Mathematics and CAD*. Kogan Page London, 1986.

[61] C. Boor. On calculating with b-splines. *J. Approximation Theory Vol.**6** pp : 50-62*, 1972.

[62] M.G. Cox. Numerical evaluation of b-splines. *J. of the Institute of Math. and its Appl. Vol. **10**2 pp : 134-149*, 1972.

[63] C. Boor. *A practical guide to splines*. Applied mathematical science Vol.**27** Springer-Verlag New York, 1978.

[64] L. Schumaker. *Splines functions : Basic Theory*. John Wiley and Sons New York, 1981.

[65] G. E. Farin. *Curves and Surfaces for Computer-Aided Geometric Design :A practical code*. Academic Press Inc Orlando Fl USA, 1996.

[66] P. Alliez and C. Gotsman. Recent advances in compression of 3d meshes. *Advances in Multiresolution for Geometric Modelling Mathematics and Visualization Part I pp 3-26*, 2005.

[67] K. Mamou, T. Zaharia, and F. Prêteux. A skinning approach for dynamic 3d mesh compression. *Computer Animation and Virtual Wordls Vol.***17** *pp : 337-346*, 2006.

[68] M. Unser, A. Aldroubi, and M. Eden. B-spline signal processing : Part i - theory. *IEEE Transaction on Signal Processing Vol.**41**(2) pp : 821-832*, 1966.

[69] K. Edee, P. Schiavone, and G. Granet. Analysis of defect in extreme uv lithography mask using a modal method based on nodal b-spline expansion. *Japanese Journal of Applied Physics. Vol. **44**(9A) pp : 6458-6462*, 2005.

[70] G. Granet, L.B. Andriamanampisoa, K. Raniriharinosy, A.M.Armeanu, and K.Edee. Modal analysis of lamellar gratings using the moment method with subsectional basis and adaptive spatial resolution. *J.Opt.Soc.Am. Vol.**27**(6)*, 2010.

[71] A. Armeanu, K. Edee, G. Granet, and P. Schiavone. The lamellar diffraction grating problem : a spectral method based on spline expansion. *Proceedings of ICMI 2 conference Vol.**19**(2) pp : 37-46*, 2009.

[72] A.M. Armeanu, K. Edee, G. Granet, and P. Schiavone. Modal method based on spline expansion for the electromagnetic analysis of the lamellar grating. *P.I.E.R. Vol.**106** pp : 243-261*, 2010.

[73] P. Bouchon, F. Pardo, R. Haïdar, and J. L. Pelouard. Fast modal method for subwavelength gratings based on b-spline formulation. *J. Opt. Soc. Am. A. /Vol **27**(4) pp : 696-702*, 2010.

[74] R. Harrington. *Field computation by Moment Methods*. New York : The Macmillan, 1968.

[75] R. Harrington. Origin and development of the method of moments for field computation. *IEEE. Ant. Prop. Soc. Magasine pp : 31-36*, 1990.

[76] R. Harrington. Matrix methods field problems. *Proc. IEEE. Vol.**55** pp : 136-149*, 1967.

[77] K. K. Mei and J. Van Bladel. Scattering by perfectly conducting rectangular cilinders. *IEEE Trans. Ant. Prop Vol.**AP-II** pp : 185-192*, 1963.

[78] L. V. Kantorovich and V. I. Krylov. *Approximate methods of higher analysis*. John Wiley and Sons Inc. New York, 1959.

[79] D. Song and Y. Y. Lu. Chebyshev collocation dirichlet-to-newmann map method for diffraction gratings. *J. Opt. Soc. A. Vol.**26**(9) pp : 1980-1988*, 2009.

[80] L. Li and C. Haggans. Convergence of the coupled-wave method for metallic lamelar diffraction gratings. *J. Opt. Soc. Am. A Vol. **10** pp : 1184-1189*, 1993.

[81] A. Quarteroni, R. Sacco, and F. Saleri. *Méthodes Numériques ; Algorithmes analyse et application*. Spriger Verlag Italia, 2004.

[82] I. Daubechies. *Ten lectures on wavelets*. SIAM Philadelphia USA, 1992.

[83] M.Vetterli and J. Kovacevic. *Wavelets and subband coding*. Prentice Hall Englewood Cliffs NJ, 1995.

[84] M. Unser, A. Aldroubi, and M. Eden. Fast b-spline transforms for continuous image representation and interpolation. *IEEE Trans. Pattern Anal. Machine Intell. Vol.**13**(3) pp : 277-285*, 1991.

[85] Mathworks. $http : //www.mathworks.com$, consulté en aout 2010.

[86] Mathematica. $http : //store.wolfram.com$, consulté en aout 2010.

[87] Stanford University. $http : //www - stat.stanford.edu/\ wavelab$, consulté en aout 2010.

[88] M. Unser. Approximation power of biorthogonal wavelets expansion. *IEEE Transaction on Signal Processing Vol.**44**(3) pp : 519-527*, 1996.

[89] A. Haar. Zur theorie des orthogonalen funktionensysteme. *Math. Ann Vol.**69** pp : 331-371*, 1910.

[90] G. Battle. A block spin construction of ondelettes. part i : Lemarié functions. *Commun Math. Phys. Vol.**110** pp : 601-615*, 1987.

[91] P. G. Lemarié. Ondelettes à localisation exponentielle. *J. Math. pures et appl. **67**(3) pp : 227-236*, 1988.

[92] C.K. Chui and J.Z. Wang. On compactly supported splines wavelets and a duality principle. *Trans. Amer. Math. Soc. Vol. **330**2 pp : 903-915*, 1992.

[93] M. Unser, A. Aldroubi, and M. Eden. On the asymptotic convergence of b-spline wavelets to gabor functions. *IEEE Trans. Information Theory. Vol.**38**(2) pp : 864-872*, 1992.

[94] A. Cohen, I. Daubechies, and J.C. Feauveau. Biorthogonal bases of complactly supported wavelets. *Comm. ACM Vol.**45** pp : 485-560*, 1992.

[95] R. Carmona, W.L. Hwang (W.L.), and Torrésani. *Pratical Time-Frequency Analysis : Wavelet and Gabor Transforms with an Implementation in S. Wavelet analysis and its applications series*. Academic Press, New York, 1998.

[96] M. Unser. Ten good reasons for using spline wavelets. *Proc. SPIE Wavelets Application in Signal and Image Processing Vol.**3169** pp : 422-431*, 1997.

Annexes

Annexe A

Formalisme des matrices S

Ce formalisme nous permet de déterminer les amplitudes des rayons réfléchis et transmis et de calculer les efficacités de diffraction. La matrice de dispersion, que nous appelons par la

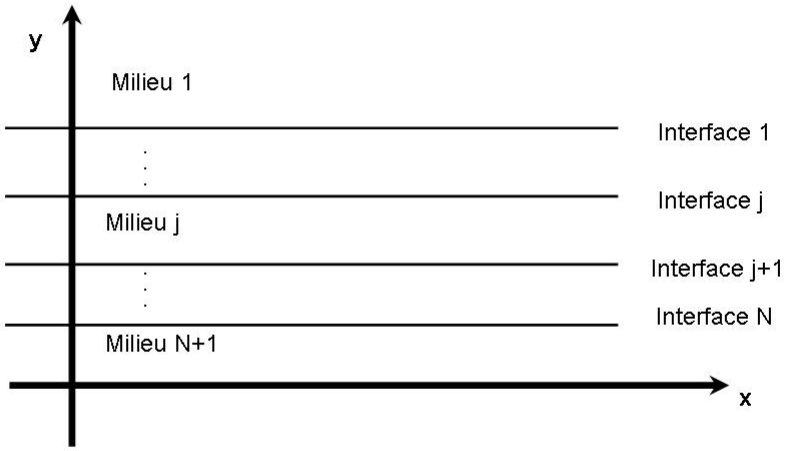


FIGURE A.1 – structure multi-couche

suite matrice **S**, relie les amplitudes entrantes dans l'interface j aux amplitudes sortantes :

$$\begin{bmatrix} A_1^- \\ A_{N+1}^+ \end{bmatrix} = \mathbf{S} \begin{bmatrix} A_{N+1}^- \\ A_1^+ \end{bmatrix} \tag{A.1}$$

A_1^- et A_1^+ sont les vecteurs colonne contenant les amplitudes entrantes et sortantes pour le milieu incident tandis que A_{N+1}^- et A_{N+1}^+ désignent les amplitudes entrantes et sortantes pour le dernier millieu. La matrice **S** est définie comme un bloc de 4 matrices :

$$\mathbf{S} = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \tag{A.2}$$

La matrice **S** est calculé à l'aide d'un algorithme récursif. On definit la matrice S^j comme :

$$\begin{bmatrix} A_j^- \\ A_{j+1}^+ \end{bmatrix} = S^j \begin{bmatrix} A_{j+1}^- \\ A_j^+ \end{bmatrix} \tag{A.3}$$

A_j^+ et A_{j+1}^+ sont les vecteurs colonne contenant les amplitudes entrantes pour les milieux j et $j+1$ tandis que A_j^- et A_{j+1}^- désignent les amplitudes sortantes. S^j est la matrice de transfert entre la couche j et la couche $j+1$. Ceci est illustré dans la figure A.2. Dans notre cas, on

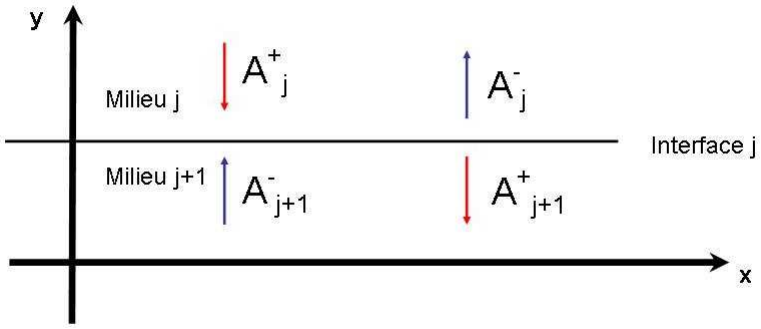


FIGURE A.2 – Les flèches rouges représentent les amplitudes entrantes et les flèches bleues les amplitudes sortantes correspondant à chaque milieu.

n'a que 3 milieux, donc le cascadage des matrices **S**, revient au cascadage des matrices S^1 et S^2. Ceci est illustré dans la figure A.3. Les coefficients de milieu lamellaire sont expriment un déphasage,

$$\begin{pmatrix} A_2^{+1} \\ A_2^{-2} \end{pmatrix} = \begin{pmatrix} 0 & \phi \\ \phi & 0 \end{pmatrix} \begin{pmatrix} A_2^{-1} \\ A_2^{+2} \end{pmatrix}$$

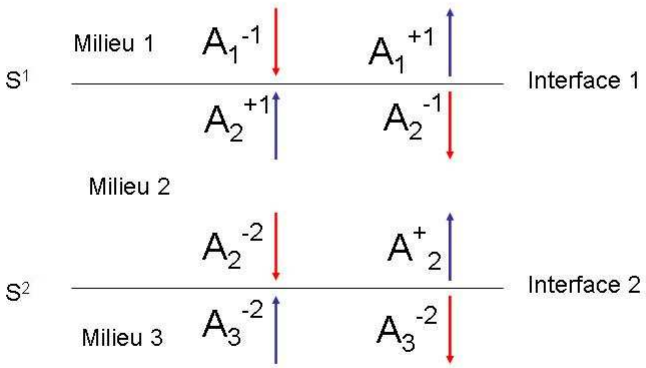


FIGURE A.3 – Les amplitudes entrantes et sortantes pour notre problème de diffraction.

où $\phi = exp(-ikrh)$, ou r et h désignent les valeurs propres et la hauteur du milieu 2. La matrice de passage entre $j = 1$ et $j = 3$, est donnée par :

$$\begin{pmatrix} A_1^{+1} \\ A_3^{-2} \end{pmatrix} = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \begin{pmatrix} A_1^{-1} \\ A_3^{+2} \end{pmatrix}$$

tel que :

$$\begin{cases} S_{11} = S_{11}^1 + S_{12}^1 \phi \Delta_2 S_{11}^2 \phi + S_{21}^1 \\ S_{12} = S_{12}^1 \phi \Delta_2 S_{12}^2 \\ S_{21} = S_{21}^2 \phi \Delta_1 S_{21}^1 \\ S_{22} = S_{22}^2 + S_{21}^2 \phi \Delta_1 S_{22}^1 \phi + S_{12}^2 \end{cases}$$

avec $\begin{cases} \Delta_1 = (I - S_{22}^1 \phi S_{11}^2)^{-1} \\ \Delta_2 = (I - S_{11}^2 \phi S_{22}^1)^{-1} \end{cases}$

Annexe B

Calcul des efficacités

Les vecteurs d'onde des ondes réfléchies et transmises se répartissent dans un plan perpendiculaire à la surface du réseau et contenant le vecteur d'onde de l'onde incidente. La valeur des efficacités de chaque onde s'obtient en faisant le rapport de la puissance réfléchie ou transmise de cette onde avec la puissance incidente.

Considérons un parallélépipède rectangle (voir figure B.1) qui contient une période de la structure du réseau. On note par $\mathbf{S}$ la surface du parallélépipède dans le plan $x0z$. La partie réelle du flux du vecteur de Poynting [1] [52] à travers une surface $\mathbf{S}$ est donnée par :

$$\mathbf{\Upsilon} = \mathfrak{Re} \int_S \mathbf{n}. \left(\frac{1}{2} \mathbf{E} \wedge \mathbf{H}^* \right) ds, \tag{B.1}$$

où $\mathbf{n}$ désigne la normale sortante à la surface et $\mathbf{H}^*$ le conjugué de $\mathbf{H}$.

Le réseau étant infiniment long en z la variation du flux dans cette direction est nulle. La variation du flux suivant l'axe $0x$ se compense puisque le réseau est périodique, donc seulement la composante en y du flux intervient dans le calcul des efficacités. Le flux suivant y à travers la surface S est donné par :

$$\mathbf{\Upsilon}_y = \frac{1}{2} \int_S \mathbf{n}_y \mathfrak{Re} \left(E_z H_x^* - E_x H_z^* \right) \mathbf{e}_y ds. \tag{B.2}$$

Le vecteur $\mathbf{e}_y$ est le vecteur unitaire suivant la direction $0y$, donc $\mathbf{n}_y = \pm \mathbf{e}_y$ en fonction de la face considérée. Pour calculer une puissance, on se place loin de la surface $y \longrightarrow \infty$; l'integrale

1. vecteur dont la direction indique, dans un milieu isotrope la direction de propagation d'une onde électromagnétique et dont l'intensité vaut la densité de puissance véhiculée par cette onde.

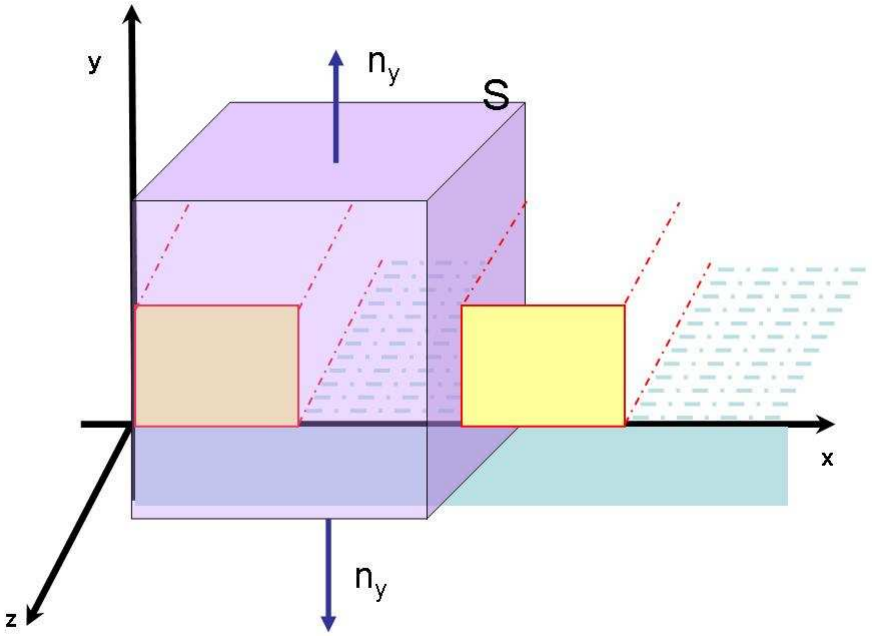


FIGURE B.1 – *Parallélépipède utilisé pour le calcul des efficacités*

sur la surface devient une intégrale dépendant que de x, d'où :

$$\Upsilon_y = \pm\frac{1}{2}\int \mathfrak{Re}\left(E_z H_x^* - E_x H_z^*\right) dx. \tag{B.3}$$

En polarisation TE, E_x est nulle, d'où :

$$\Upsilon_y = \pm\frac{1}{2}\int \mathfrak{Re}\left(E_z H_x^*\right) = \pm\frac{1}{2\omega\mu}\int \mathfrak{Re}\left(E_z \partial_y E_z^*\right) dx \tag{B.4}$$

avec

$$E_z(x,y) = \sum_m (a_m^- e^{-ikr_m y} + a_m^+ e^{ikr_m y}) \sum_n F_{nm} B_n(x) \tag{B.5}$$

$$\frac{\partial E_z^*(x,y)}{\partial y} = ik \sum_q r_q^* (a_q^{-*} e^{ikr_q^* y} - a_q^{+*} e^{-ikr_q^* y}) \sum_p F_{pq}^* B_p^*(x) \tag{B.6}$$

Le flux Υ_y est un flux linéique à cause de l'évanescence selon $0z$. En remplaçant les équations B.5 et B.6 dans B.4 on obtient :

$$\Upsilon_y = \mathfrak{Re}\frac{1}{2}\sum_{m,n,p,q}\int r_m^*(a_m^- e^{-ikr_m y} + a_m^+ e^{ikr_m y})(a_q^{-*} e^{ikr_q^* y} - a_q^{+*} e^{-ikr_q^* y}) F_{nm} F_{pq}^* B_n(x) B_p^*(x) dx \tag{B.7}$$

Pour un m fixé la somme $\sum_n F_{nm}B_n(x)$ est une onde plane, de même la somme $\sum_p F_{pq}B_p(x)$ est une onde plane pour un p fixe, donc

$$\sum_{n,p} \int r_m^* F_{nm} F_{pq}^* B_n(x) B_p^*(x) dx = r_m^* \delta_{m,q} \tag{B.8}$$

En tenant compte des ces remarques, le flux écrit :

$$\Upsilon_y = \frac{1}{2}\mathfrak{Re}(\sum_{m,q} r_m^*(a_m^- e^{-ikr_m y} + a_m^+ e^{ikr_m y})(a_q^{-*} e^{ikr_q^* y} - a_q^{+*} e^{-ikr_q^* y})) \tag{B.9}$$

Les produits de l'équation B.9 font intervenir des 4 termes avec des produits entre deux exponentielles. Les premiers deux sont :

$$e^{\pm ik(r_q + r_q^*)y} = e^{\pm 2k\mathfrak{Re}(r_q)y}$$

$$e^{\pm ik(r_q - r_q^*)y} = e^{\mp 2k\mathfrak{Im}(r_q)y}$$

Les deux autres termes sont conjugués l'un de l'autre, leur différence est imaginaire et la partie réelle de celle-ci est nulle.

$$\mathfrak{Re}\left(r_q^* e^{\pm 2k\mathfrak{Im}(r_q)y}\right) = r_q \text{ si } r_q \text{ réelle, et 0 si } r_q \text{ imaginaire}$$

On distigue :

Les ondes montantes, données par l'expression $\mathfrak{Re}\left(r_q e^{-2k\mathfrak{Im}(r_q)y}\right) \longrightarrow 0$ quand $y \longrightarrow \infty$

Les ondes descendantes, données par l'expression $\mathfrak{Re}\left(r_q e^{2k\mathfrak{Im}(r_q)y}\right) \longrightarrow 0$ quand $y \longrightarrow \infty$

Finalement pour les ondes propagatives (r_q réelles) le flux d'énergie de l'onde montante s'écrit :

$$\sum_q a_q^+ a_q^{+*} r_q \sum_n F_{nq} F_{nq}^* \tag{B.10}$$

Le flux de l'énergie descendante s'écrit :

$$\sum_q a_q^- a_q^{-*} r_q \sum_n F_{nq} F_{nq}^* \tag{B.11}$$

Une efficacité est définie comme le rapport du flux d'énergie transporté par une onde diffractée (respectivement transmise) au flux d'énergie apporté par l'onde incidente. L'efficacité réfléchie et transmise sur l'ordre q de diffraction s'écrit :

$$R_q = \frac{\Upsilon_{reflechi}^q}{\Upsilon_{incident}} = \frac{a_q^+ a_q^{+*} r_q \sum_n F_{nq} F_{nq}^*}{a_0^- a_0^{-*} r_0 \sum_n F_{n0} F_{n0}^*} \tag{B.12}$$

$$T_q = \frac{\Upsilon^q_{transmis}}{\Upsilon_{incident}} = \frac{a_q^- a_q^{-*} r_q \sum_n F_{nq} F^*_{nq}}{a_0^- a_0^{-*} r_0 \sum_n F_{n0} F^*_{n0}} \tag{B.13}$$

En polarisation TM, H_x est nulle d'où, en utilisant l'équation B.3 le vecteur de pointing selon $0y$ devient :

$$\mathbf{\Upsilon}_y = -\frac{1}{2} \int \mathfrak{Re} \left(E_x H_z^* \right) dx.$$

Le calcul est identique à la polarisation TE à la différence près que μ doit être remplacé par $\varepsilon = \varepsilon_0 \nu^2$. Les efficacités étant définies comme des rapports de flux il faut distinguer le cas où les flux sont dans le même milieu (auquel cas ε disperse comme en TE) et le cas où les flux sont dans des milieux différents. L'expression de l'efficacité réfléchie reste donc inchangée. Pour les efficacités transmises il faut multiplier l'expression B.13 par $\frac{\varepsilon_1}{\varepsilon_3}$, d'où :

$$T_q = \frac{\varepsilon_1}{\varepsilon_3} \frac{a_q^- a_q^{-*} r_q \sum_n F_{nq} F^*_{nq}}{a_0^- a_0^{-*} r_0 \sum_n F_{n0} F^*_{n0}} \tag{B.14}$$

Annexe C

Publications et conferences

Publications

A.M. Armeanu, K. Edee, G. Granet et P. Schiavone, "Modal method based on spline expansion for the electromagnetic analysis of the lamellar grating", P.I.E.R. Vol.106 pages : 243-261, 2010.

G. Granet, L. Andriamanapisoa, K. Raniriharinosy, **A.M. Armeanu** et K. Edee, "Modal analysis of lamellar gratings using the moment method with subsectionnal basis and adaptive spatial resolution", J. Opt. Soc. Am. A. Vol. 27(6), 2010.

A.M. Armeanu, K. Edee, L. Andriamanapisoa, G. Granet and P. Schiavone, "Modal method with spline expansion : the importance of the discontinuities positions in the discretisation scheme of a lamellar grating",(en cours de redaction).

Conferences

A.M.Armeanu, L.B. Andriamanampisoa, K.Edee, G.Granet, P. Schiavone, "Modal analysis of conical diffraction by binary gratings using subsectional basis expansion", Presentation orale, EOS, Paris, Octobre 2010

G.Granet, **A.M.Armeanu**, L.B. Andriamanampisoa, K.Edee, "Modal analysis of lamellar gratings : a new formulation based on the moment method with subsectional basis and adaptive spatial resolution", Presentation orale, PIERS, Boston, Juillet 2010

A.M.Armeanu, L.B. Andriamanampisoa, K.Edee, G.Granet, P.Schiavone, "Electromagnetic analysis of lamellar diffraction gratings using a spectral method based on spline expansion with a non uniform sampling scheme", Conférence invité, MSMW10, Kharkov, Ukraine Juin 2010

G. Granet, **A.M. Armeanu** et K. Edee, "The lamellar diffraction grating problem : a spectral method based on subsectional basis and adaptive spatial resolution", Présentation orale, D02010, Koli, Finlande, Fevrier 2010

A.M. Armeanu, R.Pollès, A. Moreau, K. Edee et G. Granet, "Nouvelle formulation de la méthode modale adaptée aux modes exotiques d'une structure lamellaire", Présentation poster, GDR-Ondes,Paris, Nov 2009

A.M. Armeanu, K. Edee, G. Granet et P. Schiavone, "The lamellar diffraction grating problem : a spectral method based on spline expansion", Présentation orale ICMI, Bacau, Roumanie, 8-10 Sept 2009

A.M. Armeanu, K. Edee, et P. Schiavone, "Convergence efficiency of the wavelets spline modal method and comparison with fourier modal method", Présentation orale JNTE, Toulouse 19-21 Nov 2008