

## THÈSE EN COTUTELLE

*Université de Rennes 1 et Université de Tunis El Manar*

Préparée à

**l'École Nationale d'Ingénieurs de Tunis**

et

**l'Institut de Recherche Mathématique de Rennes**

Présentée en vue de l'obtention du

**DIPLÔME DE DOCTORAT**

*Mention : Mathématiques et applications*

par

**Rania RAIS**

Couplage entre éléments finis et représentation intégrale pour les problèmes de diffraction acoustique et électromagnétique : analyse de convergence des méthodes de Krylov et méthodes multipôles rapides

**Thèse soutenue à Rennes  
le 14 Février 2014**

devant le jury composé de :

**Saloua Aouadi**

Professeur à la faculté des sciences de Tunis  
/ Rapporteur

**Victorita Dolean**

Maître de conférences, HDR à l'université de Nice  
/ Rapporteur

**Moez Khenissi**

Maître de conférences, HDR à la faculté des sciences  
de Monastir / Examineur

**Karim Ramdani**

Directeur de Recherche à l'Inria Nancy  
/ Examineur

**Nabil Gmati**

Professeur à l'école nationale d'ingénieurs de Tunis  
/ Directeur de thèse

**Florian Méhats**

Professeur à l'université de Rennes 1  
/ Directeur de thèse

**Eric Darrigrand**

Maître de conférences à l'université de Rennes 1  
/ Co-directeur de thèse



## Remerciements

Je désire exprimer ma profonde gratitude à mes directeurs de thèse Eric Darrigrand, Nabil Gmati et Florian Méhats pour avoir accepté de diriger cette thèse. J'ai beaucoup appris grâce à leur rigueur et leur précieux conseils. Je tiens à leur exprimer ma profonde reconnaissance pour m'avoir initiée à la recherche, accepté de codiriger ce travail. Leur disponibilité, leur encouragement, leur confiance et leurs conseils m'ont été d'un grand soutien durant mon stage de Master et les années de ma thèse. J'aimerais remercier en particulier Eric pour son soutien sur le plan humain et pour la bonne humeur dans laquelle ce travail a été accompli. Je le remercie également pour ses exigences lors de la relecture de mon manuscrit.

Je remercie vivement Saloua Aouadi et Victorita Dolean d'avoir accepté de rapporter mon travail. Je remercie également Moez Khenissi, Karim Ramdani de me faire l'honneur d'être membres de mon jury.

Je remercie vivement les membres du laboratoire LAMSIN de l'Ecole Nationale d'Ingénieurs de Tunis et les membres de l'IRMAR à l'Université de Rennes 1 qui m'ont apporté une aide scientifique et morale. Mes séjours à l'IRMAR ont été très enrichissants grâce à la multidisciplinarité de ses équipes et à la convivialité de ses membres. J'adresse un grand merci à Daniel Martin qui m'a beaucoup apporté par son expérience pendant mon stage de mastère et au début de ma thèse avant son départ à la retraite. Je remercie Yvon Lafranche pour son intérêt envers mon travail et l'excellente collaboration qui a pu s'établir avec lui pour le développement de la librairie MÉLINA++.

Je désire en outre remercier tous les responsables administratifs qui favorisent les bonnes conditions aux doctorants. Je tiens à exprimer tout particulièrement mes remerciements à Marie-Aude Verger, Carole Wosiak, Claude Boschet, Chantal Halet et Raoudha Jelassi.

Je voudrais remercier tous ceux qui ont partagé avec moi ces années de ma vie en tant que thésard. J'exprime ma profonde sympathie à tous les doctorants du LAMSIN et de l'IRMAR que j'ai croisés pendant ces années. En particulier, Zeineb Chebbi et Meryam Cherichi. Je tiens à remercier Julie Sauzeau et Hélène Hivert qui ont partagé avec moi le même bureau durant les derniers mois de ma thèse. Sans oublier tous les amis que j'ai rencontrés pendant mon séjour en France. J'ai eu l'occasion de faire des connaissances et partager de bons moments avec des personnes de différentes cultures très riches et diversifiées que ce soit au niveau de la religion, de la musique, des arts, de la cuisine... J'adresse mon remerciement à mes très chers amis avec qui j'ai passé des moments superbes et inoubliables : Dany Obeid (Liban), Corinna Burkard (Allemagne), Andreea Vulcan, Cristiana Dumitrescu et Cristina Tabarca (Roumanie), Wardia Fedaoui, Rahma Khelil et Zaineb Hassainia (Algérie), Safa Schrigui (Tunisie), Asma Abounan (Maroc), Samira Sulaiman et Khaled Ewhida (Libye). J'ai trouvé une ambiance familiale et j'ai partagé avec eux des différents moments conviviaux. Avec quelques un, j'étais confrontée

à des débats très intéressants qui m'ont permis d'échanger des idées et développer mon sens critique et mon esprit d'analyse.

Je tiens très sincèrement à remercier toute personne que j'ai rencontrée durant mon parcours et qui a contribué d'une façon ou d'une autre, de près ou de loin, à aller au bout de cette aventure. Je remercie également l'Agence Universitaire de la Francophonie (AUF) qui a financé mes travaux de thèse.

Enfin, je tiens à exprimer ma reconnaissance à ma famille qui m'a soutenue durant ce travail doctoral. Je remercie mes parents pour leur confiance, leurs encouragements et pour avoir supporté mes angoisses pendant les moments difficiles au cours de mes travaux de thèse. Un grand merci à mes sœurs Noura, Amira, Sandra et mes frères Othmane et Mohamed. Je remercie tous les membres de ma famille en particulier mon grand père qui était heureux de ma réussite ... sans oublier mes tantes, mes oncles et mes cousins, notamment ma tante Ahlem, ma première enseignante en maternelle.

*Je dédie cette thèse à mes parents  
à mes soeurs et à mes frères  
à Malek et Ali*



# Table des matières

|          |   |           |
|----------|---|-----------|
| <b>0</b> | <b>Introduction générale</b>  | <b>1</b>  |
| <b>I</b> | <b>Résolution du problème de la propagation d'onde acoustique en domaine non borné</b>                                  | <b>5</b>  |
| <b>1</b> | <b>Comparaison de CEFRI à des méthodes locales</b>  | <b>7</b>  |
| 1.1      | Position du problème de Helmholtz extérieur . . . . .   | 7         |
| 1.2      | Méthode <b>CEFRI</b> pour le problème de Helmholtz . . . . .  | 7         |
| 1.3      | Conditions aux limites transparentes locales . . . . .  | 12        |
| 1.3.1    | Condition de rayonnement à distance finie . . . . .   | 12        |
| 1.3.2    | Condition transparente de type <b>BGT</b> . . . . .   | 13        |
| 1.4      | Tests numériques . . . . .  | 14        |
| 1.4.1    | Etude de l'erreur en fonction de la taille du domaine . . . . .   | 14        |
| 1.4.2    | Etude de l'erreur en fonction du pas de discrétisation . . . . .  | 17        |
| 1.4.3    | Etude de l'erreur de <b>CEFRI</b> en fonction de la discrétisation à un nombre de couches éléments finis fixé . . . . . | 19        |
| 1.4.4    | Comparaison du temps de calcul . . . . .  | 20        |
| 1.4.5    | Conclusions . . . . .   | 22        |
| <b>2</b> | <b>Mise en œuvre de la méthode multipôles rapide</b>  | <b>25</b> |
| 2.1      | Historique de la FMM . . . . .  | 25        |
| 2.2      | Exemple d'application de la FMM en configuration 1D . . . . .   | 25        |
| 2.3      | Application de la FMM 1 niveau à CEFRI en 3D . . . . .  | 28        |
| 2.3.1    | Construction du pavage et notion de voisinage . . . . .   | 29        |
| 2.3.2    | Développement multipolaire . . . . .  | 30        |
| 2.3.3    | Algorithme et calcul de la complexité de la FMM un niveau . . . . .   | 34        |
| 2.4      | Application de la FMM multiniveaux à <b>CEFRI</b> en 3D . . . . .   | 36        |
| 2.4.1    | Présentation générale . . . . .   | 36        |

|  |   |           |
|--|---|-----------|
| 2.4.2  | Algorithme : FMM multiniveaux . . . . .                       | 37        |
| 2.4.3  | Interpolation et anterpolation . . . . .                      | 39        |
| 2.5  | Tests numériques : cas FMM un niveau . . . . .                | 42        |
| 2.6  | FMM stable pour toutes fréquences . . . . .                   | 45        |
| 2.7  | Présentation de la librairie FastMMLib . . . . .              | 47        |
| <b>II Résolution du problème de la propagation d'onde électromagnétique en domaine non borné</b> |   | <b>51</b> |
| <b>3</b>   | <b>Méthode de Schwarz – cas classique</b>                     | <b>53</b> |
| 3.1  | Éléments constitutifs du modèle . . . . .                     | 53        |
| 3.1.1  | Présentation des équations de Maxwell . . . . .               | 53        |
| 3.1.2  | Conditions de transmission entre deux milieux . . . . .       | 55        |
| 3.1.3  | Conditions aux limites d'un conducteur parfait . . . . .      | 56        |
| 3.2  | Problème de diffraction électromagnétique . . . . .           | 56        |
| 3.2.1  | Problème de Maxwell classique en domaine non borné . . . . .  | 56        |
| 3.2.2  | Problème en domaine borné . . . . .                           | 57        |
| 3.2.3  | Formulation variationnelle, existence et unicité . . . . .    | 58        |
| 3.3  | Méthode de Schwarz avec recouvrement total . . . . .          | 60        |
| 3.4  | Diffraction par une sphère conductrice . . . . .              | 61        |
| 3.4.1  | Outils mathématiques . . . . .                                | 62        |
| 3.4.2  | Investigation analytique en 3D . . . . .                      | 63        |
| <b>4</b>   | <b>Méthode de Schwarz – cas régularisé</b>                    | <b>71</b> |
| 4.1  | Equation de Maxwell régularisée . . . . .                     | 71        |
| 4.2  | Problème de Maxwell régularisé en domaine non borné . . . . . | 72        |
| 4.3  | Formulation variationnelle, existence et unicité . . . . .    | 74        |
| 4.4  | Méthode de Schwarz avec recouvrement total . . . . .          | 75        |
| 4.5  | Investigation analytique en 3D : . . . . .                    | 77        |
| 4.6  | Analyse de la méthode GMRES . . . . .                         | 80        |
| 4.6.1  | Formulation du problème sur le bord artificiel . . . . .      | 80        |
| 4.6.2  | Etude de la convergence du GMRES . . . . .                    | 81        |
| <b>5</b>   | <b>Mise en œuvre numérique</b>                                | <b>87</b> |
| 5.1  | Approximation par éléments finis de Lagrange . . . . .        | 87        |
| 5.2  | Tests de validation du code CEFRI Maxwell . . . . .           | 88        |
| 5.2.1  | Problème de Maxwell avec une condition de Neumann . . . . .   | 89        |

|       |   |            |
|-------|---|------------|
| 5.2.2 | Problème de Maxwell avec une condition de Dirichlet . . . . . | 92         |
| 5.3   | Résolution du problème de Maxwell initial . . . . .           | 101        |
|       | <b>Conclusions et perspectives</b>                            | <b>109</b> |
|       | <b>Annexe A</b>   | <b>111</b> |
|       | <b>Annexe B</b>   | <b>117</b> |
|       | <b>Annexe C</b>   | <b>119</b> |



# Chapitre 0

## Introduction générale

Les problèmes de propagation d'ondes acoustiques et électromagnétiques sont présents dans de multiples domaines. Citons pour la propagation des ondes acoustiques la réduction des nuisances sonores des automobiles, avions, hélicoptères, l'optimisation de l'acoustique des salles de concert. Pour la propagation des ondes électromagnétiques, on peut citer le dimensionnement des antennes de télécommunications, la mesure des effets de la téléphonie mobile, la détection et la caractérisation de menaces (missiles, avions de combat, ...), la protection des équipements électroniques embarqués. Le travail présenté dans cette thèse concerne la résolution du problème de diffraction d'une onde incidente par un obstacle. Cela consiste à déterminer le champ diffracté par l'obstacle étudié. Nous nous intéressons tout d'abord au problème de la diffraction d'une onde acoustique harmonique en temps. Ce premier problème est régi par les équations de Helmholtz. Nous nous intéresserons ensuite au problème de diffraction d'une onde électromagnétique harmonique en temps, régi par les équations de Maxwell. Pour résoudre ces problèmes extérieurs, il est nécessaire de se ramener à un problème équivalent posé dans un domaine borné. Parmi les méthodes les plus simples à mettre en œuvre, nous pouvons citer les méthodes basées sur des conditions aux limites transparentes locales. Ces méthodes consistent à imposer sur la frontière fictive du domaine de calcul, des conditions artificielles approchées, telles que les conditions de Sommerfeld à distance finie [1], ou les conditions aux limites transparentes de type Bayliss-Gunzburger-Turkel de premier et de second ordre ([2] [3] [4] [5]). Une discrétisation par éléments finis standard du système résultant de ces méthodes donne un système linéaire dont la matrice est creuse. L'inconvénient de ces méthodes est que pour obtenir des résultats satisfaisants, on est amené à imposer le bord du domaine de calcul suffisamment loin de la frontière de l'obstacle. Cela peut être un facteur limitant de ces méthodes du point de vue du coût du calcul et de la mémoire. Nous citons par ailleurs la méthode PML ([6]), qui peut également être une bonne alternative à la résolution des problèmes extérieurs. Elle consiste à remplacer la frontière artificielle par une couche absorbante de telle sorte à annuler toute réflexion. Cette méthode bien que très efficace pour les guides d'ondes, reste plus difficile à généraliser pour des obstacles quelconques. Une autre alternative consiste à restreindre les calculs au bord de l'obstacle : la méthode des équations intégrales de frontière ([7][8][9]). La matrice résultant de la discrétisation est dense et mal conditionnée. De plus les noyaux de Green (potentiel de simple couche et de double couche) intervenant

dans la matrice sont singuliers, ce qui oblige à avoir recours à des stratégies complexes de traitement des singularités. Malgré le développement de préconditionneurs très performants ces dernières années, plusieurs problèmes restent difficiles à aborder à l'aide de cette méthode. Notons à titre d'exemple la diffraction par des cavités lorsque l'ouverture de la cavité est de l'ordre de la longueur d'onde. Pour les problèmes de diffraction d'une onde électromagnétique par un obstacle diélectrique (coefficients variables à l'intérieur de l'obstacle) la construction de bons préconditionneurs reste très difficile. L'objectif de cette thèse est de nous intéresser plus particulièrement à une méthode utilisant une condition artificielle exacte définie par la représentation intégrale de la solution à l'extérieur du domaine de calcul. Cette méthode est appelée méthode de couplage entre éléments finis et représentation intégrale et sera désignée par l'acronyme **CEFRI** ([10] [11] [12] [13]). L'un des avantages de cette méthode est le fait d'éviter le traitement des singularités présentes dans la méthode des éléments finis de frontière. Par ailleurs, la frontière fictive qui a permis de construire le domaine de calcul peut être choisie assez proche de celle de l'obstacle. Malgré la bonne connaissance des problèmes mathématiques posés par la méthode **CEFRI**, plusieurs questions restent posées concernant leur mise en œuvre numérique. La méthode de couplage entre éléments finis et représentation intégrale s'interprète comme une méthode de décomposition de domaines pour la résolution d'un problème en domaine non borné.

Plusieurs aspects et possibilités restent à explorer : une première expérimentation analytique et numérique de la technique de Schwarz pour la résolution de l'équation de Helmholtz dans la thèse de F. Jelassi ([14], [15]) a orienté le choix d'un préconditionneur à une méthode de type Krylov. L'exploration, de type analytique, qui est faite dans la thèse de F. Jelassi illustre bien leur fiabilité et leur apport. Il reste, bien entendu, à étendre ces résultats aux équations de Maxwell ce qui est fait dans le cadre de mes travaux de thèse. Nous montrons l'analogie de certains algorithmes utilisés pour la résolution du système linéaire obtenu par la méthode **CEFRI** et la méthode de Schwarz, pour la résolution des équations de Maxwell en domaine extérieur. Cette étude est examinée dans un premier temps, puis pour le problème de Maxwell régularisé dans un deuxième temps. Cette relecture de la méthode **CEFRI** nous a permis d'orienter le choix d'un préconditionneur pour les méthodes de Krylov. Nous avons justifié la convergence de la méthode de GMRES ([16]) pour le problème de Maxwell régularisé. En fait, une analyse approfondie du comportement de celui-ci est faite prouvant sa convergence superlinéaire est examinée pour les problèmes continus et se base sur des résultats de la théorie spectrale. Sur le plan numérique, le projet de thèse est basé sur la participation au développement de la librairie **MÉLINA++** [17] (en langage C++) initialement développée par Daniel Martin à l'université de Rennes 1 et désormais par l'équipe Analyse Numérique de l'IRMAR. Nous nous sommes intéressés dans un premier temps au cas de la diffraction par une onde acoustique. Une mise en œuvre numérique est établie par les librairies **MÉLINA** [18] et **MÉLINA++**. Dans un second volet, nous nous sommes intéressés à la résolution numérique du problème de diffraction électromagnétique par un conducteur parfait. Pour la mise en œuvre numérique, l'objectif était de greffer au sein de la librairie éléments finis **MÉLINA++** les outils nécessaires à la résolution du problème. Ce travail a contribué de manière très significative au développement de la librairie. Il consiste à intégrer l'ensemble des intégrands relatifs à son problème mathématique, des intégrands volumiques et des intégrands surfaciques relatifs à des opérateurs intégraux. L'utilisation de la li-

brairie MÉLINA++ pour la résolution du problème de Maxwell a mené à l'utilisation de la stratégie de la régularisation des équations de Maxwell ([13]) et la pénalisation de la condition essentielle imposée sur le bord de l'obstacle. Ces spécificités ont nécessité une analyse fine du comportement numérique afin d'obtenir la validation de chacun des aspects. Au cours de ces travaux, j'ai également mis en évidence quelques bugs initiaux de la librairie et permis ainsi la correction et la validation de quelques objets essentiels de la librairie dont je n'étais pas l'auteur.

Lors de la résolution d'un système faisant intervenir des opérateurs intégraux, il est indispensable d'avoir recours à des méthodes de réduction des coûts relatifs aux intégrands issus des opérateurs intégraux. En fait, la discrétisation par la méthode éléments finis induit la présence d'une matrice pleine issue de la méthode de couplage entre éléments finis et représentation intégrale, du fait de l'apparition d'un opérateur intégral couplant les deux frontières (le bord de l'obstacle et le bord artificiel) du domaine. Les produits matrice-vecteur peuvent devenir un facteur limitant en temps de calculs et en coût de mémoire. En effet, la partie éléments-finis induisant des matrices creuses, les calculs relatifs à la représentation intégrale peuvent rapidement, selon les cas, représenter un pourcentage gigantesque du coût de résolution. La méthode multipôles rapide (Fast Multipole Method ou FMM) permet de repousser ces limitations. Pour le cas des équations de Helmholtz nous pouvons nous référer aux papiers [19] [20] [21] [22]. En effet, le calcul limitant un algorithme itératif est le produit matrice-vecteur, qui pour une matrice pleine de taille  $N \times N$ , est proportionnel à  $N^2$ . Dans des applications physiques, on est amené à considérer des situations où le nombre d'onde est très grand. Et comme  $N$  croît avec le nombre d'onde de manière quadratique pour la discrétisation d'une surface, le produit matrice-vecteur devient très coûteux. La méthode multipôles rapide réduit les coûts en temps de calcul et en place mémoire du produit matrice-vecteur. La FMM existe sous deux formes : la méthode 1 niveau permet de calculer le produit matrice-vecteur en  $O(N^{3/2})$  opérations. La méthode multiniveaux permet d'obtenir une complexité de l'ordre de  $O(N \ln N)$ . Nous avons mis en œuvre la méthode multipôles rapide 1 niveau dans le cadre de la résolution de problème de Helmholtz par la méthode CEFRI. En fait, nous avons implémenté la méthode FMM dans la librairie MÉLINA. Vu la limitation de l'allocation dynamique de MÉLINA qui est liée aux spécificités du fortran 77, nous étions restreints à des cas tests de taille modeste. Cet obstacle peut être surmonté en utilisant MÉLINA++. La stratégie adoptée dans notre développement de la méthode FMM se base sur un développement du noyau de Green selon la série de Gegenbauer. Cette méthode s'avère limitée face au problème dit de basses fréquences. Ce problème apparaît aussi à haute fréquence lorsque la dimension des détails géométriques est petite devant la longueur d'onde du problème physique traité. Les travaux de E. Darve et P. Havé ([23] [24]) ont donné lieu à une nouvelle version de la FMM valable pour un large spectre de fréquences incluant les basses fréquences. Pour cela, nous contribuons au développement de la librairie FastMMLib [25]. L'objectif de cette librairie est d'offrir à la communauté scientifique une version générique des méthodes multipôles rapides. Plus précisément, nous participons au développement de la version multiniveaux de la FMM.

### Plan de la thèse

Ce mémoire de thèse comporte deux parties : la première concerne l'étude du problème de Helmholtz en domaine extérieur et la seconde partie concerne le problème de Max-

well en domaine non borné. Le but de la première partie est de se familiariser avec les différentes méthodes auxquelles nous nous intéressons pour la résolution d'un problème de diffraction en domaine non borné (méthode **CEFRI**, méthode FMM). L'objectif majeur consiste à appliquer ces méthodes dans le cadre des équations de Maxwell. Le premier chapitre présente une comparaison de la méthode **CEFRI** avec quelques méthodes à condition aux limites transparentes locales. A travers une série d'expériences numériques avec la librairie **MÉLINA**, nous montrons l'efficacité de la méthode **CEFRI** par rapport à des méthodes basées sur des conditions de rayonnement à distance finie ou des conditions aux limites transparentes de type **BGT** de premier et de second ordre. Le chapitre 2 est dédié à présenter la méthode multipôles rapide. La plupart des développements liés à la FMM ne sont pas nouveaux. Le but est d'appliquer ces notions et de montrer l'apport de la FMM dans le cadre de la résolution du problème de Helmholtz dans un domaine tridimensionnel non borné résolu par la méthode **CEFRI**.

Dans un second volet, nous nous sommes intéressés au problème de diffraction électromagnétique par un conducteur parfait. Nous présentons au troisième chapitre le modèle physique auquel nous nous intéressons. Nous abordons l'algorithme de Schwarz avec recouvrement total pour le problème de Maxwell classique extérieur en 3D. Nous avons montré l'analogie de certains algorithmes utilisés pour la résolution du système linéaire obtenu par la méthode **CEFRI** et la méthode de Schwarz, pour la résolution des équations de Maxwell en domaine extérieur. L'analogie, que nous avons faite, avec la méthode de Schwarz permet une exploration analytique de la méthode afin d'illustrer les taux de convergence. Pour la mise en œuvre numérique de ce problème dans la librairie éléments finis **MÉLINA++**, nous nous intéressons à un problème équivalent dit problème de Maxwell régularisé. Au chapitre 4, nous étudions le comportement de la méthode de Schwarz pour ce problème. Nous montrons que la méthode **CEFRI** est interprétée comme une méthode de Schwarz qui suggère un préconditionneur pour un solveur de type Krylov. Une analyse théorique est aussi faite de la convergence de la méthode de GMRES pour le problème de Maxwell régularisé. Le dernier chapitre présente des simulations numériques qui ont permis la validation de la mise en œuvre de la méthode **CEFRI** pour le problème de Maxwell dans **MÉLINA++** et du choix de la stratégie itérative.

À la fin de ce manuscrit de thèse, nous avons regroupé dans l'annexe A des résultats mathématiques utiles pour l'étude des phénomènes de propagation d'ondes acoustiques et électromagnétiques. Nous présentons à travers l'annexe B une brève introduction concernant la méthode de Schwarz. L'annexe C donne une idée sur les différents outils intégrés dans **MÉLINA++** ainsi que la structure du code de la mise en œuvre de la méthode **CEFRI**.

# Première partie

## Résolution du problème de la propagation d'onde acoustique en domaine non borné

Cette partie est dédiée à la présentation des différentes méthodes auxquelles nous nous intéressons afin de résoudre un problème de diffraction en domaine non borné. Nous avons choisi de commencer avec le problème de Helmholtz vu sa simplicité par rapport au problème de Maxwell. Plus précisément, nous nous intéressons à la méthode de couplage entre éléments finis et représentation intégrale **CEFRI**. Nous avons mené une étude comparative avec quelques méthodes de résolution des problèmes extérieurs. Du point de vue de la résolution du système linéaire, nous avons tiré quelques caractéristiques de la méthode **CEFRI**. D'autre part, nous avons évalué l'utilité de faire appel aux méthodes multipôles rapides FMM pour limiter le coût en temps de calculs et en place mémoire. Pour cela, nous présentons deux variantes de méthodes FMM. La première est destinée plutôt aux problèmes de hautes fréquences. La deuxième méthode est plus générale et peut être employée pour un large spectre de fréquences.



# Chapitre 1

## Comparaison de la méthode CEFRI à des méthodes à une condition aux limites transparente locale

### 1.1 Position du problème de Helmholtz extérieur

On s'intéresse au calcul du potentiel de perturbation  $u$  d'une onde incidente plane induite par la présence d'un obstacle  $\Omega_i$  de frontière  $\Gamma$ , solution du problème extérieur :

$$\left\{ \begin{array}{l} \Delta u + k^2 u = 0 \text{ dans } \Omega^e = \mathbb{R}^3 \setminus \overline{\Omega}_i, \\ \partial_{\mathbf{n}} u = f \text{ sur } \Gamma, \\ \lim_{R \rightarrow \infty} \int_{\|x\|=R} |\partial_{\mathbf{n}} u - iku|^2 d\gamma = 0. \end{array} \right. \quad (1.1)$$

où  $k > 0$  est le nombre d'onde et  $\mathbf{n}$  est la normale sortante du domaine extérieur. La donnée  $f$  de la première condition est l'opposée de la dérivée normale de l'onde incidente sur l'obstacle et la seconde condition est appelée condition de rayonnement de Sommerfeld; elle définit le comportement de la solution à l'infini. Le problème est posé sur le complémentaire du corps  $\Omega_i$  dans  $\mathbb{R}^3$  et pour tout  $f$  dans  $L^2(\Gamma)$ , ce problème admet une unique solution dans  $H_{loc}^1(\Omega^e)$ .

On s'intéresse à une classe de méthodes qui mène la résolution du problème de diffraction extérieur à un problème posé dans un domaine borné entre l'obstacle et la frontière fictive simulant l'infini. Plus précisément, on se propose dans cette partie de présenter une étude comparative entre quelques méthodes transparentes locales et la méthode de couplage entre éléments finis et représentation intégrale qu'on désignera dans la suite par **CEFRI** ([10] [11] [12] [13]).

### 1.2 Méthode CEFRI pour le problème de Helmholtz

La méthode **CEFRI** consiste à introduire une frontière  $\Sigma$  essentiellement arbitraire et qui entoure l'obstacle sur lequel nous construisons une condition aux limites au moyen de

la représentation intégrale sur la frontière  $\Gamma$  : pour tout  $x$  appartenant à  $\Omega^e$

$$u(x) = \int_{\Gamma} \partial_{\mathbf{n}_y} u(y) G(x, y) - u(y) \partial_{\mathbf{n}_y} G(x, y) d\gamma(y) \quad (1.2)$$

où  $G$  est la fonction de Green associée à l'équation de Helmholtz en dimension 3 est définie par

$$G(x, y) = \frac{e^{ik\|x-y\|}}{4\pi\|x-y\|}.$$

$\mathbf{n}_y$  représente le vecteur normal unitaire à la surface  $\Gamma$  dirigé vers l'extérieur du domaine  $\Omega$ . Nous nous ramenons alors à un problème posé dans le domaine borné  $\Omega$ , délimité par les frontières  $\Gamma$  et  $\Sigma$  :

$$\left\{ \begin{array}{l} \Delta u + k^2 u = 0 \text{ dans } \Omega, \\ \partial_{\mathbf{n}_y} u = f \text{ sur } \Gamma, \\ (\partial_{\mathbf{n}_x} - ik)u(x) = (\partial_{\mathbf{n}_x} - ik) \int_{\Gamma} f(y) G(x, y) - u(y) \partial_{\mathbf{n}_y} G(x, y) d\gamma(y) \text{ sur } \Sigma. \end{array} \right. \quad (1.3)$$

$\mathbf{n}_x$  représente le vecteur normal unitaire à la surface  $\Sigma$  dirigé vers l'extérieur du domaine  $\Omega$ . Les deux problèmes (1.1) et (1.3) sont équivalents ([26]) au sens où la solution de (1.3) est la restriction de la solution de (1.1) à  $\Omega$  et la solution de (1.3) peut être prolongée d'une façon unique dans  $\Omega^e$  à l'aide de (1.2). Bien que l'équivalence entre le problème extérieur (1.1) et (1.3) est établie, nous exposerons dans la suite une démonstration légèrement différente et qui donne une interprétation en termes de méthode de décomposition de domaines. Pour cela on va procéder en deux étapes :

- Montrer l'équivalence entre le problème (1.1) et le problème (1.4-1.6,1.7-1.10) défini dans la suite.
- Montrer l'équivalence entre les problèmes (1.3) et (1.4-1.6,1.7-1.10).

En fait, le problème (1.1) peut s'interpréter comme issu d'une technique de décomposition de domaine. Ce résultat est énoncé dans la proposition suivante :

**Proposition 1.2.1.** *Si  $(u_1, u_2)$  est solution du problème (1.4-1.6,1.7-1.10)*

$$\left\{ \begin{array}{l} \Delta u_1 + k^2 u_1 = 0 \text{ dans } \Omega, \end{array} \right. \quad (1.4)$$

$$\left\{ \begin{array}{l} \partial_{\mathbf{n}_y} u_1 = f \text{ sur } \Gamma, \end{array} \right. \quad (1.5)$$

$$\left\{ \begin{array}{l} (\partial_{\mathbf{n}_x} - ik)u_1(x) = (\partial_{\mathbf{n}_x} - ik)u_2(x) \text{ sur } \Sigma. \end{array} \right. \quad (1.6)$$

$$\left\{ \begin{array}{l} \Delta u_2 + k^2 u_2 = 0 \text{ dans } \Omega_i \cup \Omega^e, \end{array} \right. \quad (1.7)$$

$$\left\{ \begin{array}{l} [\partial_{\mathbf{n}_y} u_2] = \partial_{\mathbf{n}_y} u_1 \text{ sur } \Gamma, \end{array} \right. \quad (1.8)$$

$$\left\{ \begin{array}{l} [u_2] = u_1 \text{ sur } \Gamma, \end{array} \right. \quad (1.9)$$

$$\left\{ \begin{array}{l} \text{Condition de rayonnement.} \end{array} \right. \quad (1.10)$$

alors

$$u = \left\{ \begin{array}{ll} u_1 & \text{dans } \Omega \\ u_2 & \text{dans } \Omega^e \end{array} \right.$$

est solution de (1.1) et  $u_1 = u_2$  dans  $\Omega \cap \Omega^e$ . Et réciproquement, si  $u$  est une solution de (1.1) alors

$$u_1 = u|_{\Omega} \quad \text{et} \quad u_2 = \begin{cases} 0 & \text{dans } \Omega_i \\ u & \text{dans } \Omega^e \end{cases} \quad (1.11)$$

sont respectivement solutions de (1.4-1.6) et (1.7-1.10).



Preuve :

Pour montrer l'implication directe, on considère  $u_1$  solution de (1.4-1.6) et  $u_2$  solution de (1.7-1.10) et nous allons montrer que :

$$u = \begin{cases} u_1 & \text{dans } \Omega \\ u_2 & \text{dans } \Omega^e \end{cases}$$

est solution de (1.1) et que  $u_1 = u_2$  dans  $\Omega \cap \Omega^e$ .

Il est clair que  $\partial_{\mathbf{n}_y} u = f$  sur  $\Gamma$  et que  $u$  vérifie la condition de rayonnement. De plus  $u$  vérifie l'équation de Helmholtz dans  $\Omega^e$ . Il reste à montrer que  $u_1$  et  $u_2$  coïncident sur  $\Omega$ . Pour cela considérons la fonction suivante :

$$\psi = \begin{cases} u_2 - u_1 & \text{dans } \Omega \\ u_2 & \text{dans } \Omega_i \end{cases}$$

On a donc

$$\begin{cases} \Delta \psi + k^2 \psi = 0 & \text{dans } \bar{\Omega}_i \cup \Omega, \\ (\partial_{\mathbf{n}_x} - ik)\psi = 0 & \text{sur } \Sigma. \end{cases} \quad (1.12)$$

Ce problème admet pour unique solution  $\psi = 0$  dans  $\bar{\Omega}_i \cup \Omega$  ([26]). Par suite,  $u_1 = u_2$  dans  $\Omega$  et  $u_2 = 0$  dans  $\Omega_i$ .

Etant donné une solution  $u$  de (1.1), on vérifie aisément que  $u_1$  (respectivement  $u_2$ ) donné par (1.11) vérifie les équations (1.4) et (1.5) (respectivement (1.7) et (1.10)). De plus

$$\begin{aligned} [u_2]_{\Gamma} &= u|_{\Gamma} \\ &= u_1|_{\Gamma}, \end{aligned}$$

et

$$\begin{aligned} [\partial_{\mathbf{n}_y} u_2]_{\Gamma} &= \partial_{\mathbf{n}_y} u|_{\Gamma} \\ &= \partial_{\mathbf{n}_y} u_1|_{\Gamma}. \end{aligned}$$

Par ailleurs, on a d'une part,  $u = u_1$  dans  $\Omega$  donc en particulier :

$$u = u_1 \quad \text{sur } \Sigma \quad \text{et} \quad \partial_{\mathbf{n}_x} u = \partial_{\mathbf{n}_x} u_1 \quad \text{sur } \Sigma.$$

D'autre part, en vertu de la formule de représentation intégrale (voir Annexe 1) nous aurons

$$u_2(x) = \int_{\Gamma} [\partial_{\mathbf{n}_y} u_2(y)] G(x, y) - [u_2(y)] \partial_{\mathbf{n}_y} G(x, y) d\gamma(y) \quad \forall x \in \Omega_i \cup \Omega^e.$$

Ceci implique que

$$u_2(x) = \int_{\Gamma} \partial_{\mathbf{n}_y} u(y) G(x, y) - u(y) \partial_{\mathbf{n}_y} G(x, y) d\gamma(y) \quad \forall x \in \Omega_i \cup \Omega^e.$$

En particulier on a sur  $x \in \Sigma$ .

$$(\partial_{\mathbf{n}_x} - ik)u_2(x) = (\partial_{\mathbf{n}_x} - ik) \int_{\Gamma} \partial_{\mathbf{n}_y} u(y) G(x, y) - u(y) \partial_{\mathbf{n}_y} G(x, y) d\gamma(y) \quad \text{sur } \Sigma.$$

Cela donne

$$\begin{aligned} (\partial_{\mathbf{n}_x} - ik)u_2(x) &= (\partial_{\mathbf{n}_x} - ik)u(x) \quad \text{sur } \Sigma \\ &= (\partial_{\mathbf{n}_x} - ik)u_1(x) \quad \text{sur } \Sigma, \end{aligned}$$

ce qui permet de montrer le second sens de la proposition.

**Proposition 1.2.2.** *Le problème en domaine borné (1.3) est équivalent au problème (1.4-1.6, 1.7-1.10).*

Preuve :

Soit  $u$  solution de (1.3) alors en posant

$$v(x) = \int_{\Gamma} \partial_{\mathbf{n}_y} u(y) G(x, y) - u(y) \partial_{\mathbf{n}_y} G(x, y) d\gamma(y), \quad \forall x \in \Omega_i \cup \Omega^e,$$

on a

$$(\partial_{\mathbf{n}_x} - ik)u(x) = (\partial_{\mathbf{n}_x} - ik)v(x) \quad \text{sur } \Sigma.$$

**Proposition 1.2.3.** *Considérons un domaine  $O$ . Soit  $u$  la fonction définie sur  $O \cup O'$  par :*

$$u(x) = \int_{\partial O} [\partial_{\mathbf{n}_y} u(y)] G(x, y) - [u(y)] \partial_{\mathbf{n}_y} G(x, y) d\gamma(y),$$

où  $O'$  est le domaine complémentaire de  $O$  et  $\mathbf{n}_y$  la normale extérieure à  $O'$ .

Alors  $(\Delta + k^2)u = 0$  séparément dans  $O$  et  $O'$  et  $u$  vérifie la condition de rayonnement.

Compte tenu de la proposition précédente  $(u, v)$  est solution du problème (1.4-1.6, 1.7-1.10). Inversement, si  $u_1$  et  $u_2$  sont des solutions respectives de (1.4-1.6) et (1.7-1.10) on a d'après la formule de représentation intégrale dans  $\Omega_i \cup \Omega^e$  (voir annexe A)

$$u_2(x) = \int_{\Gamma} \partial_{\mathbf{n}_y} u_1(y) G(x, y) - u_1(y) \partial_{\mathbf{n}_y} G(x, y) d\gamma(y) \quad \forall x \in \Omega_i \cup \Omega^e.$$

Or puisque

$$(\partial_{\mathbf{n}_x} - ik)u_1(x) = (\partial_{\mathbf{n}_x} - ik)u_2(x) \quad \text{sur } \Sigma,$$

on a alors

$$(\partial_{\mathbf{n}_x} - ik)u_1(x) = (\partial_{\mathbf{n}_x} - ik) \int_{\Gamma} f(u) G(x, y) - u_1(y) \partial_{\mathbf{n}_y} G(x, y) d\gamma(y) \text{ sur } \Sigma.$$

Ainsi,  $u_1$  est solution de (1.3) et ceci achève la démonstration. On peut donc conclure, d'après les deux propositions 1.2.1 et 1.2.2, que le problème non borné (1.1) est équivalent au problème tronqué (1.3).

La formulation variationnelle du problème (1.3) consiste à :

$$\left\{ \begin{array}{l} \text{Trouver } u \in H^1(\Omega) \text{ tel que pour tout } v \in H^1(\Omega) \\ \int_{\Omega} \nabla u(x) \nabla v(x) dx - k^2 \int_{\Omega} u(x) v(x) dx - ik \int_{\Sigma} u(x) v(x) d\sigma(x) \\ + \int_{\Sigma} D_{\Gamma} u(x) v(x) d\sigma(x) = \int_{\Gamma} f(x) v(x) d\gamma(x) + \int_{\Sigma} S_{\Gamma} f(x) v(x) d\sigma(x) \end{array} \right. \quad (1.13)$$

où les opérateurs  $D_{\Gamma}$  et  $S_{\Gamma}$  sont définis par :  $\forall v \in L^2(\Gamma) \forall x \in \Sigma$

$$D_{\Gamma} v(x) = \int_{\Gamma} v(y) \partial_{\mathbf{n}_y} K(x, y) d\gamma(y) \quad \text{et} \quad S_{\Gamma} v(x) = \int_{\Gamma} v(y) K(x, y) d\gamma(y)$$

avec le noyau défini par

$$K(x, y) = (\partial_{\mathbf{n}_x} - ik) G(x, y).$$

La discrétisation de (1.13) par une approximation d'éléments finis standard (de Lagrange), dont la base est notée  $\{w_{\alpha}\}_{\alpha}$  conduit au système linéaire dont les inconnues sont les valeurs nodales  $u_{\alpha}$  de l'approximation  $\sum_{\alpha} u_{\alpha} w_{\alpha}$  de la solution  $u$  :

$$(A + C) \mathbf{U} = \mathbf{F}^{\Gamma} - \mathbf{F}^{\Sigma} \quad (1.14)$$

avec

$$\begin{aligned} A_{\beta\alpha} &= \int_{\Omega} \nabla w_{\alpha}(x) \nabla w_{\beta}(x) - k^2 w_{\alpha}(x) w_{\beta}(x) dx - ik \int_{\Sigma} w_{\alpha}(x) w_{\beta}(x) d\sigma(x), \\ C_{\beta\alpha} &= \int_{\Sigma} D_{\Gamma}(w_{\alpha})(x) w_{\beta}(x) d\sigma(x), \\ \mathbf{F}_{\alpha}^{\Gamma} &= \int_{\Gamma} f(x) w_{\alpha}(x) d\gamma(x), \quad \mathbf{F}_{\beta}^{\Sigma} = \int_{\Sigma} S_{\Gamma}(f)(x) w_{\beta}(x) d\sigma(x). \end{aligned}$$

La matrice  $A$  est inversible, creuse, symétrique et à coefficients complexes, par contre la matrice  $C$  comporte un bloc plein dû au couplage des deux frontières du domaine de calcul. La matrice  $A + C$  est complexe et non hermitienne; de plus  $A + C$  est mal conditionnée et si nous choisissons  $A$  comme préconditionneur on obtient le système

$$(I + A^{-1}C) \mathbf{U} = A^{-1}(\mathbf{F}^{\Gamma} - \mathbf{F}^{\Sigma}). \quad (1.15)$$

Si on choisit comme solveur l'algorithme GMRES, au cours des itérations, la multiplication de  $A^{-1}$  par un vecteur donné est effectuée simplement par la résolution d'un système

linéaire de matrice  $A$  par la méthode LU en résolvant deux systèmes linéaires triangulaires par la méthode de descente et la méthode de remontée.

On pose  $H(x, y) = \partial_{\mathbf{n}_y} K(x, y)$  pour tout  $(x, y) \in \Sigma \times \Gamma$ , et on approche la fonction  $H$  par interpolation sur les fonctions de base éléments finis définies sur les nœuds de  $\Sigma \times \Gamma$  :

$$H(x, y) \approx \sum_{\sigma} \sum_{\gamma} H(x_{\sigma}, y_{\gamma}) w_{\sigma}(x) w_{\gamma}(y),$$

où l'indice  $\sigma$  (resp.  $\gamma$ ) est générique des nœuds de  $\Sigma$  (resp.  $\Gamma$ ). La matrice  $C$  peut s'écrire alors sous forme du produit des matrices

$$C = P_{\Sigma}^t M_{\Sigma} \mathcal{H}_{\Sigma\Gamma} M_{\Gamma} P_{\Gamma}$$

où  $P_{\Gamma}$  (resp.  $P_{\Sigma}$ ) est la (matrice de) projection sur  $\Gamma$  (resp. sur  $\Sigma$ ),  $M_{\Gamma}$  (resp.  $M_{\Sigma}$ ) la matrice de masse sur  $\Gamma$  (resp. sur  $\Sigma$ ) et  $\mathcal{H}_{\Sigma\Gamma}$  la matrice noyau de coefficients  $H(x_{\sigma}, y_{\gamma})$ . Avec des notations analogues, le terme  $\mathbf{F}^{\Sigma}$  peut s'écrire comme produit matrice-vecteur

$$\mathbf{F}_{\sigma}^{\Sigma} = \sum_{\gamma} \mathcal{R}_{\sigma\gamma} f_{\gamma} \quad \text{où} \quad \mathcal{R} = P_{\Sigma}^t M_{\Sigma} \mathcal{K}_{\Sigma\Gamma} M_{\Gamma}$$

et  $f_{\gamma}$  est la  $\gamma$ -ème composante de l'interpolée de la donnée  $f$  sur  $\Gamma$  :

$$f(x) \approx \sum_{\gamma} f_{\gamma} w_{\gamma}(x)$$

et  $\mathcal{K}_{\Sigma\Gamma}$  la matrice noyau de coefficients  $K(x_{\sigma}, y_{\gamma})$ .

## 1.3 Méthodes basées sur des conditions aux limites transparentes locales

### 1.3.1 Condition de rayonnement à distance finie

L'idée de cette méthode consiste à introduire une frontière fictive  $\Sigma$  entourant l'obstacle sur laquelle on impose une condition aux limites provenant de la condition de radiation de type Sommerfeld ([1]). Cette condition aux limites sur  $\Sigma$  est aussi désignée condition de rayonnement à distance finie (**CRDF**). Ceci, nous amène à étudier le problème suivant :

$$\begin{cases} \Delta u + k^2 u = 0 \text{ dans } \Omega, \\ \partial_{\mathbf{n}_y} u = f \text{ sur } \Gamma, \\ (\partial_{\mathbf{n}_x} - ik)u(x) = 0 \text{ sur } \Sigma. \end{cases} \quad (1.16)$$

Avec cette méthode, il est nécessaire d'éloigner suffisamment la frontière  $\Sigma$  pour avoir une bonne approximation de la condition de rayonnement de Sommerfeld. La formulation variationnelle consiste à

$$\begin{cases} \text{Trouver } u \in H^1(\Omega) \text{ tel que pour tout } v \in H^1(\Omega) \\ \int_{\Omega} \nabla u(x) \nabla \bar{v}(x) dx - k^2 \int_{\Omega} u(x) \bar{v}(x) dx - ik \int_{\Sigma} u(x) \bar{v}(x) d\sigma(x) \\ = \int_{\Gamma} f(x) \bar{v}(x) d\gamma(x) \end{cases}$$

Après une discrétisation par éléments finis, on aboutit au système linéaire suivant :

$$AU = \mathbf{F}^\Gamma.$$

où  $A$  et  $\mathbf{F}^\Gamma$  sont les mêmes éléments du système (1.14) défini dans la section précédente.

### 1.3.2 Condition transparente de type BGT

Les conditions aux limites transparentes locales désignées **BGT** de premier et de second ordre sont les conditions introduites par Bayliss, Gunzburger et Turkel ([2]) pour des frontières fictives circulaires et sphériques. Ces méthodes ont été ensuite étendues dans ([3],[4], [5]) pour des frontières artificielles de formes générales convexes et pour un ordre élevé et donc plus précises. Le problème extérieur (1.1) se ramène à résoudre le problème suivant :

$$\begin{cases} \Delta u + k^2 u = 0 \text{ dans } \Omega, \\ \partial_{\mathbf{n}_y} u = f \text{ sur } \Gamma, \\ \partial_{\mathbf{n}_x} u = -\mathcal{M}_l u \text{ sur } \Sigma ; l = 1, 2. \end{cases}$$

En 2D :

- Condition de premier ordre **BGT1** :  $\partial_{\mathbf{n}_x} u = -\mathcal{M}_1 u := iku - \frac{\kappa}{2}u$
- Condition de deuxième ordre **BGT2** :

$$\partial_{\mathbf{n}_x} u = \mathcal{M}_2 u := -\mathcal{M}_1 u + \frac{\kappa^2}{8(\kappa - ik)}u + \partial_s \left( \frac{1}{2(\kappa - ik)} \partial_s u \right)$$

où  $\kappa$  est la courbure de la surface et  $s$  l'abscisse curviligne.

La formulation variationnelle consiste à

$$\begin{cases} \text{Trouver } u \in H^1(\Omega) \text{ tel que pour tout } v \in H^1(\Omega) \\ \int_{\Omega} \nabla u(x) \nabla \bar{v}(x) dx - k^2 \int_{\Omega} u(x) \bar{v}(x) dx + \int_{\Sigma} \mathcal{M}_l u(x) \bar{v}(x) d\sigma(x) \\ = \int_{\Gamma} f(x) \bar{v}(x) d\gamma(x) \end{cases}$$

Une discrétisation par éléments finis mène à la résolution de ce système linéaire :

$$BU = \mathbf{F}^\Gamma.$$

où  $B$  est définie comme suit

$$B_{\beta\alpha} = \int_{\Omega} \nabla w_\alpha(x) \nabla w_\beta(x) - k^2 w_\alpha(x) w_\beta(x) dx + \int_{\Sigma} \mathcal{M}_l w_\alpha(x) w_\beta(x) d\sigma(x)$$

et  $\mathbf{F}^\Gamma$  est défini dans la section précédente. On remarque que la matrice  $B$  n'est autre que la somme de la matrice  $A$  définie précédemment et une matrice creuse. Par exemple pour le cas de la condition **BGT1** en 2D, on a

$$B = A + \frac{\kappa}{2}M$$

où  $M$  est la matrice de masse sur  $\Sigma$ .

Les méthodes à condition aux limites transparentes locales conduisent à des systèmes linéaires dont la matrice est à coefficients complexes, creuse, symétrique et non hermitienne. On peut utiliser des méthodes directes de résolution telles que la méthode d'élimination de Gauss ou la factorisation LU. On peut choisir aussi un solveur itératif adapté au problème traité. Les méthodes directes sont robustes en comparaison avec les méthodes itératives mais elles peuvent devenir coûteuses en termes de mémoire et de temps du calcul dès qu'il s'agit de problèmes en 3D.

## 1.4 Tests numériques

On présente des résultats numériques qui ont été obtenus à l'aide de la bibliothèque MÉLINA dans sa version Fortran 77 [18] développée au sein de l'IRMAR par D. Martin et au sein de l'ENSTA. MÉLINA a été utilisée dans d'autres applications que les problèmes de diffraction acoustique (voir section Pressbook de [18]).

On considère le cas d'une onde incidente plane et un nombre d'onde  $k$ . On vise à résoudre le problème de Helmholtz à l'extérieur du disque unité. Dans ce cas, on dispose d'une solution exacte. Les simulations numériques sont faites dans un premier temps par une interpolation P1 puis en P2. Dans la suite, on désigne par  $h$  le pas de discrétisation du maillage et  $R$  le rayon du bord artificiel entourant l'obstacle.

### 1.4.1 Etude de l'erreur en fonction de la taille du domaine, à discrétisation fixée

On s'intéresse au cas où le pas de discrétisation  $h$  est fixé et on fait varier la distance délimitée entre la frontière de l'obstacle et la frontière artificielle de rayon  $R$ . On trace l'évolution de l'erreur quadratique relative entre la solution calculée par **CEFRI** et la solution exacte, l'erreur quadratique relative entre la solution calculée par **CRDF** et la solution exacte et l'erreur quadratique relative entre la solution calculée par **BGT2** et la solution exacte qu'on désigne respectivement  $err_c$ ,  $err_r$  et  $err\_bgt2$ .

Au regard de la figure 1.1, on remarque que l'erreur relative entre la solution numérique et la solution exacte est plus faible pour la méthode **CEFRI**. De plus, l'erreur  $err_c$  donne une bonne précision pour  $R = 1.06$ . Puis, lorsqu'on augmente davantage la distance entre l'obstacle et la frontière fictive, l'erreur se dégrade de nouveau vraisemblablement à cause du conditionnement du système qui devient mauvais lorsque la taille de la matrice éléments finis grandit. Pour l'erreur  $err_r$  et l'erreur  $err\_bgt2$ , lorsque  $R$  augmente l'approximation s'améliore puis à partir de  $R = 1.6$  les erreurs d'arrondi deviennent trop importantes et masquent l'amélioration quand  $R$  augmente. On remarque que la valeur optimale de l'erreur  $err_r$  et l'erreur  $err\_bgt2$  n'est pas satisfaisante par rapport à l'erreur  $err_c$  mais l'erreur  $err\_bgt2$  reste toujours meilleure que l'erreur  $err_r$ . Dans la figure 1.2, on refait la même étude de précision en augmentant le nombre de points par longueur d'onde : on a fait le choix d'une discrétisation d'environ 20 points par longueur d'onde. On constate les mêmes observations. On remarque que

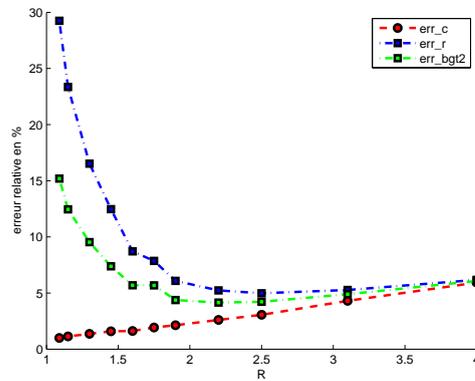
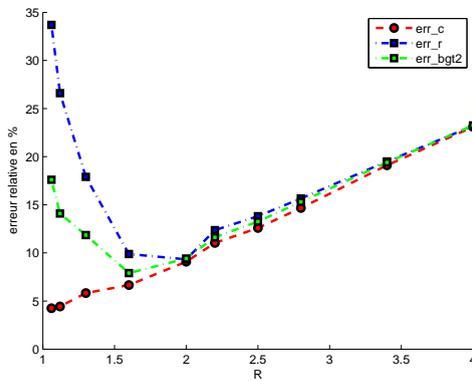


FIGURE 1.1 – Cas interpolation P1,  $k = 10$  ;  $\lambda = 10h$  ( $h = 0.06$ ).  
 FIGURE 1.2 – Cas interpolation P1,  $k = 10$  ;  $\lambda = 20h$  ( $h = 0.03$ ).

pour la méthode **CRDF**, l'erreur  $\text{err}_r$  atteint une valeur optimale égale à environ 5% pour un  $R = 2.5$  qui reste toujours supérieure à celle de l'erreur  $\text{err}_c$  qui est égale à 1% pour  $R = 1.09$ . La distance entre l'obstacle et le bord artificiel joue un rôle important sur la précision de la méthode **CRDF** et de la méthode **BGT2**. Mais en comparant les deux figures 1.1 et 1.2, on observe que le comportement de la précision de ces deux méthodes dépend aussi de la finesse du maillage. En fait, dans le cas test où on a mis dix points par longueur d'onde l'erreur  $\text{err}_r$  optimale est environ 9% pour  $R = 1.6$ , tandis que pour un choix de vingt points par longueur d'onde, l'erreur  $\text{err}_r$  pour  $R = 2.5$  est environ 5%. Au regard des deux figures, on voit bien que l'erreur  $\text{err}_r$  aussi bien que l'erreur  $\text{err}_bgt2$  décroît puis commence à croître tandis que ce phénomène apparaît plus vite dans le test 1 (figure 1.1) que dans le test 2 (figure 1.2) où on a mis plus de points par longueur d'onde. Dans les figures 1.3 et 1.4, on refait la même étude en choi-

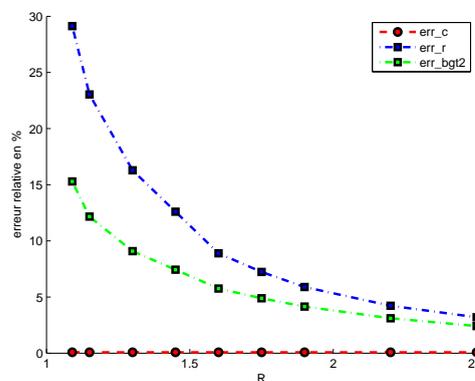
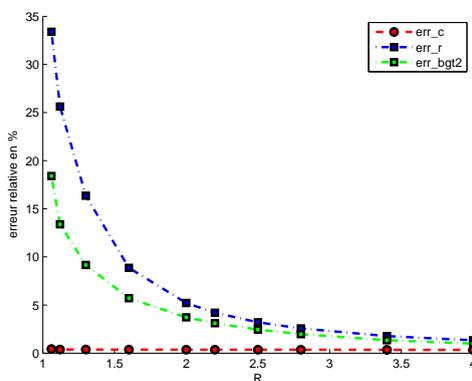


FIGURE 1.3 – Cas interpolation P2,  $k = 10$  ;  $\lambda = 10h$  ( $h = 0.06$ ).  
 FIGURE 1.4 – Cas interpolation P2,  $k = 10$  ;  $\lambda = 20h$  ( $h = 0.03$ ).

issant  $k = 10$  et une interpolation P2 : le pas de discrétisation est  $h = 0.06$  à gauche et  $h = 0.03$  à droite. On remarque toujours que la précision de la méthode **CEFRI** est beaucoup plus satisfaisante que celle des méthodes **CRDF** et **BGT2**. Mais en comparant

avec les résultats en P1, on observe qu'avec une interpolation P2 il n'y a qu'une légère différence entre les résultats donnés avec une discrétisation par dix points par longueur d'onde et ceux obtenus par vingt points par longueur d'onde. Dans 1.5 et 1.6, on se place

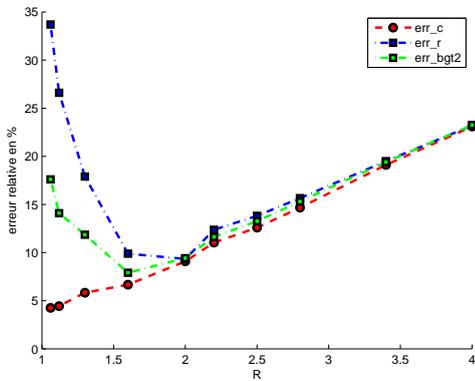


FIGURE 1.5 – Cas interpolation P1,  $k = 10$ ,  $\lambda = 10h$  ( $h = 0.06$ ).

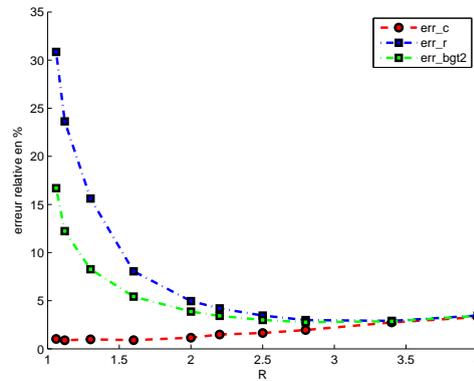


FIGURE 1.6 – Cas interpolation P2,  $k = 20$ ,  $\lambda = 5h$  ( $h = 0.06$ ).

dans une configuration où on a le même pas de discrétisation du maillage ( $h=0.06$ ). Pour une interpolation P1 (figure 1.5), on a mis dix points par longueur d'onde. Pour le cas d'une interpolation en P2 (figure 1.6) on a 5 points par longueur d'onde. Le but de ces choix est de se situer dans un cas similaire pour les deux interpolations. On refait la même étude. On observe que les résultats en P2 sont meilleurs que ceux en P1.

Dans la suite, on se propose de regarder chaque méthode indépendamment : **CEFRI**, **CRDF** ou **BGT2**. Sur les figures (1.7), (1.8) et (1.9), on représente l'erreur relative commise en fonction de  $R$  pour différents degrés d'interpolation de la méthode éléments finis P1 et P2. On voit que les résultats en P2 sont plus précis et plus stables que ceux obtenus en P1.

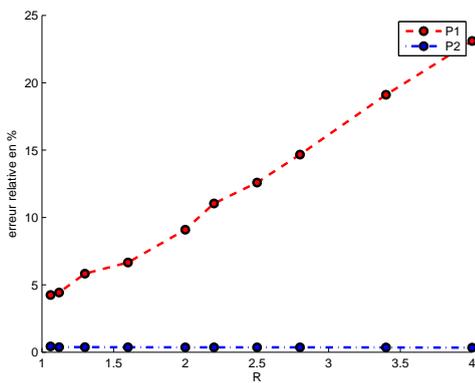


FIGURE 1.7 – Etude de la précision de la méthode **CEFRI** en P1 et P2,  $h = 0.06$ ,  $k = 10$  ( $\lambda = 10h$ ).

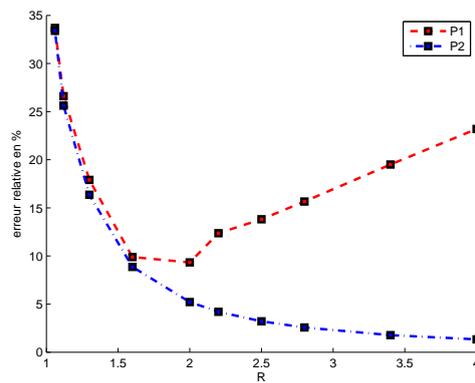


FIGURE 1.8 – Etude de la précision de la méthode **CRDF** en P1 et P2,  $h = 0.06$ ,  $k = 10$  ( $\lambda = 10h$ ).

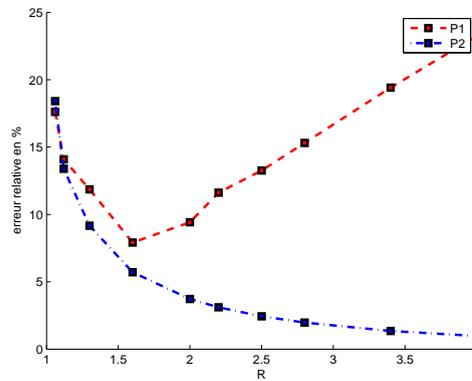


FIGURE 1.9 – Etude de la précision de la méthode **BGT2** en P1 et P2,  $h = 0.06$ ,  $k = 10$  ( $\lambda = 10h$ ).

### 1.4.2 Etude de l'erreur en fonction du nombre d'onde à un domaine fixé

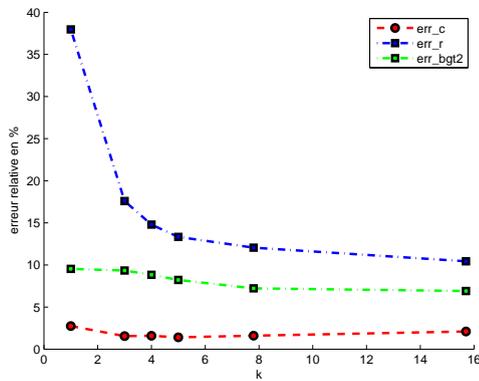
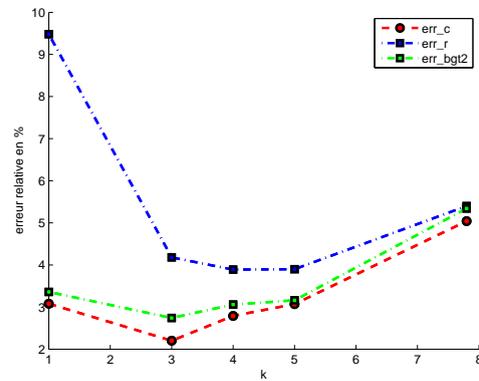
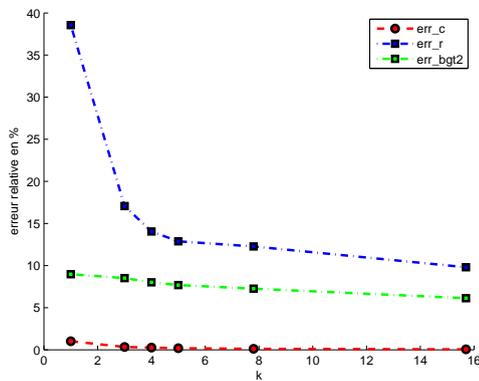
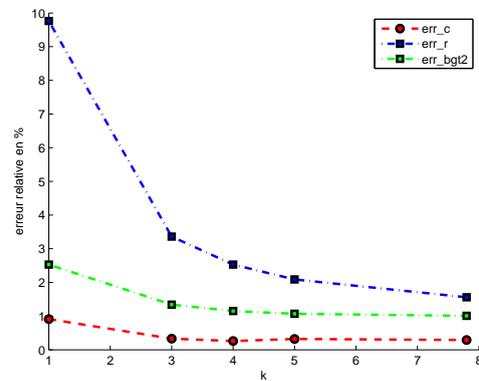
On se propose d'étudier le cas où on fait varier conjointement le nombre d'onde et le pas de discrétisation  $h$  et on fixe la distance délimitée par la frontière de l'obstacle de rayon 1 et la frontière artificielle de rayon  $R$ . Dans un premier temps, on choisit  $R = 1.5$  puis, on regardera ce qui se passe pour le cas  $R = 4$ .

|                   |       |       |       |       |       |       |
|-------------------|-------|-------|-------|-------|-------|-------|
| $h$               | 0.3   | 0.1   | 0.08  | 0.06  | 0.04  | 0.02  |
| $k$               | 1     | 3     | 4     | 5     | 7.8   | 15.7  |
| nombre d'éléments | 109   | 918   | 1479  | 2569  | 6050  | 24213 |
| err_c en %        | 2.75  | 1.55  | 1.59  | 1.41  | 1.6   | 2.1   |
| err_r en %        | 37.97 | 17.59 | 14.79 | 13.33 | 12.05 | 10.43 |
| err_bgt2 en %     | 9.54  | 9.34  | 8.83  | 8.22  | 7.22  | 6.91  |

TABLE 1.1 – Cas interpolation P1 :  $R = 1.5$ ,  $\lambda=20h$ .

|                   |      |       |       |       |       |
|-------------------|------|-------|-------|-------|-------|
| $h$               | 0.3  | 0.1   | 0.08  | 0.06  | 0.04  |
| $k$               | 1    | 3     | 4     | 5     | 7.8   |
| nombre d'éléments | 1192 | 11132 | 18074 | 31236 | 71993 |
| err_c en %        | 3.08 | 2.29  | 2.79  | 3.07  | 5.04  |
| err_r en %        | 9.48 | 4.18  | 3.89  | 3.94  | 5.41  |
| err_bgt2 en %     | 3.36 | 2.74  | 3.06  | 3.16  | 5.34  |

TABLE 1.2 – Cas interpolation P1,  $R = 4$ ,  $\lambda=20h$ .

FIGURE 1.10 – Cas interpolation P1,  $R = 1.5$ ,  $\lambda = 20h$ .FIGURE 1.11 – Cas interpolation P1,  $R = 4$ ,  $\lambda = 20h$ .FIGURE 1.12 – Cas interpolation P2,  $R = 1.5$ ,  $\lambda = 20h$ .FIGURE 1.13 – Cas interpolation P2,  $R = 4$ ,  $\lambda = 20h$ .

Dans cette partie, on précise au lecteur qu'on va étudier la précision des méthodes **CEFRI**, **CRDF** et **BGT2** pour différents problèmes physiques. En fait, pour chaque choix du nombre d'onde  $k$  on a mis environ 20 points par longueur d'onde (Voir tableau 2.1 et 2.2). Au regard des figures 1.10, 1.11, 1.12 et 1.13, on remarque que l'erreur  $\text{err}_c$  est plus faible que l'erreur  $\text{err}_r$  et l'erreur  $\text{err}_{\text{bgt2}}$ . Mais on remarque que l'erreur  $\text{err}_r$  dans 1.10 diminue en raffinant le maillage du domaine de calcul pour atteindre une valeur optimale égale à environ 10%. On observe que cette valeur est atteinte dans la figure 1.11 pour  $R$  choisi suffisamment grand (cas  $R = 4$ ) et pour un choix du nombre d'onde  $k = 1$ . Pour ce choix de  $R$ , l'erreur  $\text{err}_r$  a une valeur minimale d'environ 4%. On remarque aussi que l'erreur se dégrade pour  $h = 0.04$  ce qui correspond à un maillage de 71993 triangles d'ordre 1. On peut expliquer cette instabilité numérique par une accumulation des erreurs d'arrondi des calculs éléments finis P1 vu la taille du domaine de calcul. Et ce phénomène n'apparait pas pour des éléments finis d'ordre 2 (voir figures 1.12 et 1.13).

A travers cette série de tests numériques, on a constaté que la méthode **CEFRI** donne un résultat plus satisfaisant que les méthodes **CRDF** et **BGT2**. En comparant les résultats pour le choix d'une frontière fictive de rayon  $R = 1.5$  au cas où  $R = 4$ , on a vérifié que les

méthodes basées sur des conditions aux limites transparentes locales se comportent mieux pour une frontière artificielle assez loin de l'obstacle. Ce comparatif ne mérite pas d'être fait pour la méthode **CEFRI**. Concernant la méthode **CEFRI**, on se propose de mener une étude plus pertinente dans la section suivante : analyse à un nombre de couches éléments finis fixé.

### 1.4.3 Etude de l'erreur de CEFRI en fonction de la discrétisation à un nombre de couches éléments finis fixé

Dans cette partie, on se propose d'étudier le comportement de la méthode **CEFRI** en fonction de la finesse du maillage. Ceci correspond à définir la distance entre le bord de l'obstacle  $\Gamma$  et la frontière artificielle  $\Sigma$  en fonction du choix du pas de la discrétisation  $h$ , de telle sorte qu'on a un nombre de couches éléments finis fixé. Dans ce cas, le domaine de calcul devient de plus en plus petit quand  $h$  tend vers 0. On désigne par CEF le groupe nominal : couche éléments finis. On présente des résultats numériques pour un choix d'une seule couche éléments finis pour une interpolation P1 puis une interpolation P2. La même étude est effectuée pour deux couches éléments finis. Dans les figures 1.15 et 1.17,

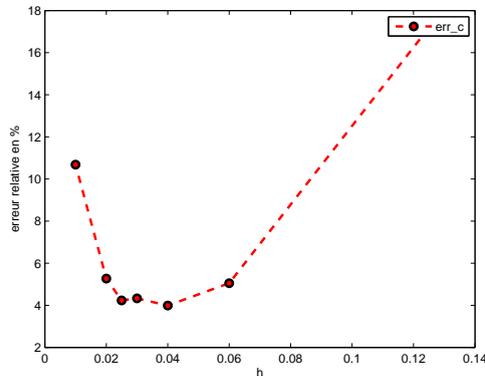


FIGURE 1.14 – Cas interpolation P1, 1 CEF,  $k = 10$ .

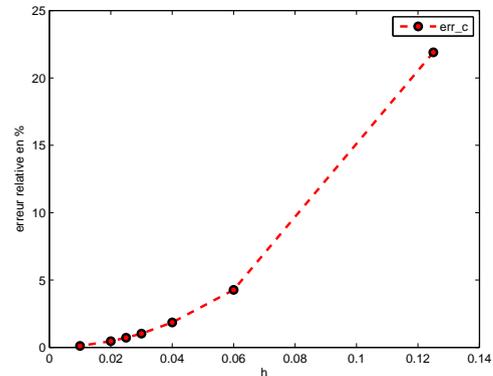
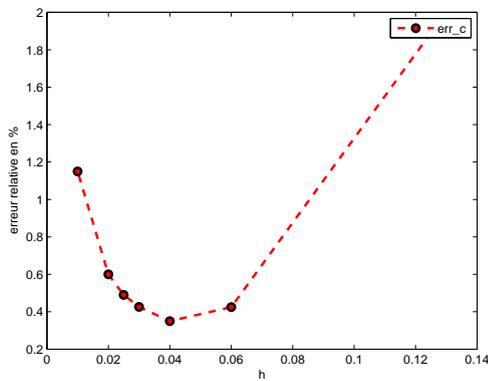
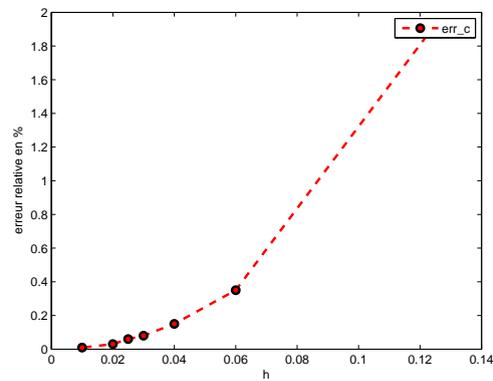
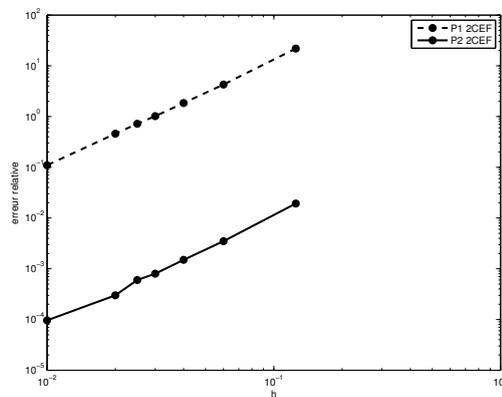


FIGURE 1.15 – Cas interpolation P1, 2 CEF,  $k = 10$ .

on constate, pour le cas d'un choix de deux couches éléments finis une bonne stabilité du comportement de l'erreur  $err_c$  en P1 aussi bien que en P2. L'erreur  $err_c$  décroît en choisissant un pas de discrétisation  $h$  de plus en plus fin. Par contre avec une seule couche éléments finis, (voir figures 1.14 et 1.16) on observe que l'erreur décroît puis on remarque une instabilité numérique : l'erreur  $err_c$  augmente en raffinant de plus en plus le maillage. La courbe 1.18 représente une autre manière de présenter le comportement de la précision  $err_c$  en utilisant l'échelle logarithmique. En traçant les résultats par exemple du cas d'une interpolation P1 et P2 avec deux couches éléments finis dans le domaine de calcul, on obtient une droite qui traduit encore une fois la stabilité numérique expliquée précédemment.

FIGURE 1.16 – Cas interpolation P2, 1 CEF,  $k = 10$ .FIGURE 1.17 – Cas interpolation P2, 2 CEF,  $k = 10$ .FIGURE 1.18 – Cas interpolation P1, P2 ; 2 CEF ;  $k=10$  en échelle logarithmique

#### 1.4.4 Comparaison du temps de calcul

On se propose dans cette partie de mener une étude comparative du temps de calculs pour les méthodes auxquelles on s'intéresse. L'analyse est faite pour deux choix d'interpolation éléments finis P1 et P2.

A partir de la figure 1.19, la méthode **CEFRI** atteint une précision optimale pour un rayon du bord artificiel  $R = 1.06$  qui correspond à un temps de calcul égal à 0.11. On présente dans la figure 1.19 l'évolution du temps de calcul mis pour résoudre le problème avec la méthode **CRDF** et **BGT2** en fonction de  $R$ . On trace la droite  $t_{\text{cefri}}$  qui représente le temps écoulé pour avoir la meilleure précision de la méthode **CEFRI**. On observe que les courbes du temps de calcul de **CRDF** et **BGT2** désignées respectivement  $t_{\text{crdf}}$  et  $t_{\text{bgt2}}$  croisent la courbe de temps  $t_{\text{cefri}}$  présentée pour le cas  $R = 1.6$ . Dans la figure 1.19, un choix de rayon du bord artificiel égal à 1.6 correspond à une erreur  $\text{err}_r$  de l'ordre de 10% et l'erreur  $\text{err}_{\text{bgt2}}$  est de l'ordre de 8% qui sont toutes les deux supérieures à l'erreur  $\text{err}_c$  qui est de l'ordre de 6%.

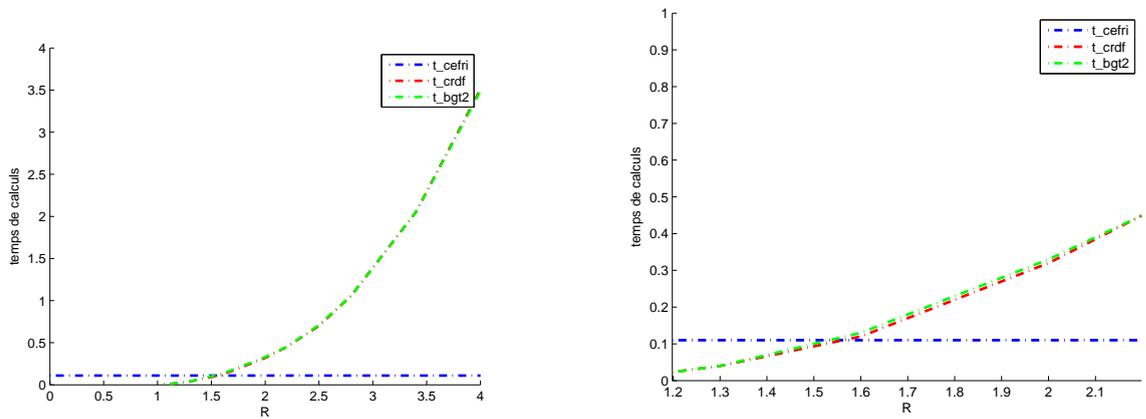


FIGURE 1.19 – Cas  $k = 10$ ,  $\lambda = 10h$  ( $h = 0.06$ ), temps de calcul en P1 (la figure à droite est un zoom de celle qui est à gauche).

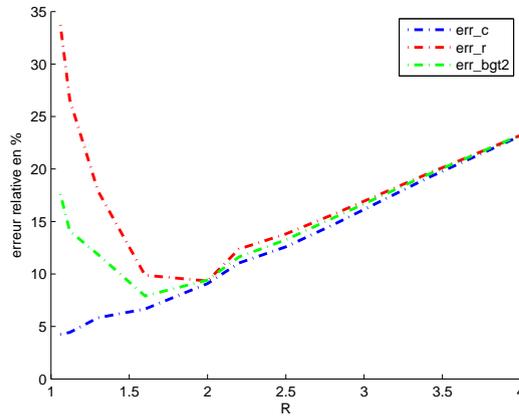


FIGURE 1.20 – Cas interpolation P1,  $k = 10$ ;  $\lambda = 10h$  ( $h = 0.06$ ).

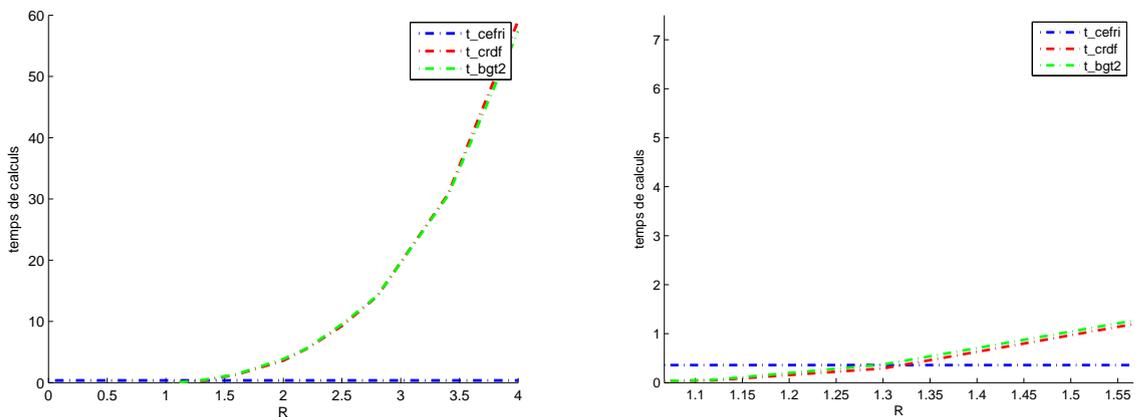
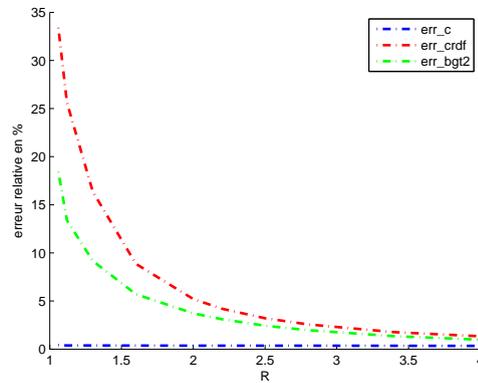


FIGURE 1.21 – Cas  $k = 10$ ,  $\lambda = 10h$  ( $h = 0.06$ ), temps de calcul en P2 (la figure à droite est un zoom de celle qui est à gauche).

FIGURE 1.22 – Cas interpolation P2,  $k = 20$ ,  $\lambda = 5h$  ( $h = 0.06$ ).

A travers la figure 1.21, on mène une étude comparative entre la précision des méthodes **CEFRI**, **CRDF** et **BGT2** pour le cas d’une interpolation P2. On remarque que l’erreur  $err_r$  et l’erreur  $err\_bgt2$  sont plus au moins de l’ordre de 1% pour le cas d’un  $R = 4$ , que le temps mis pour la résolution est à peu près égal à 60 secondes (voir figure 1.21). La méthode **CEFRI** atteint cette précision mais avec un domaine de calcul beaucoup plus petit qui correspond à  $R = 1.06$  ce qui favorise un temps de calcul très petit (égal à 0.36 seconde) par rapport à celui donné par les méthodes transparentes.

| erreur        | t_cefri | t_crdf | t_bgt2 |
|---------------|---------|--------|--------|
| $\simeq 1\%$  | 0.36    | 59.11  | 57.39  |
| $\simeq 5\%$  | -       | 3.61   | 1.38   |
| $\simeq 10\%$ | -       | 1.31   | 0.37   |

TABLE 1.3 – Etude du temps de calcul à une précision fixée, cas interpolation P2 :  $h = 0.06$ ,  $k = 10$ .

D’autre part, nous étudions dans le tableau 1.3 le temps mis pour chacune de ces méthodes à une précision donnée par éléments finis en P2. Nous rappelons que la méthode **CEFRI** permet d’avoir une précision de 1% en variant le domaine de calcul (on peut consulter la figure (1.9)). Nous remarquons que les méthodes **CRDF** et **BGT2** atteignent aussi cette précision mais en mettant une durée à peu près égale à 170 fois le temps mis par la méthode **CEFRI**. Nous constatons aussi que les méthodes à condition aux limites transparentes locales peuvent être satisfaisantes pour la résolution du problème extérieur. Ceci est vrai si on n’exige pas d’avoir une très bonne précision.

### 1.4.5 Conclusions

A travers ces tests numériques, nous constatons que la méthode **CEFRI** est robuste du point de vue précision et complexité en temps de calcul. Ceci, est vrai même pour un choix de deux couches éléments finis. Nous n’avons pas besoin d’élargir le domaine de

calcul comme c'est le cas pour les méthodes **CRDF** et **BGT2**. Le seul inconvénient par rapport aux méthodes à condition aux limites transparentes locales telles que les méthodes de **CRDF** et **BGT** est la mise en œuvre numérique. En fait, les méthodes **CRDF** et **BGT2** sont faciles à implémenter dans une librairie éléments finis. Par contre, la méthode **CEFRI** mérite plus de développement numérique surtout pour l'implémentation du terme de couplage qui fait interagir la surface de l'obstacle et le bord artificiel (voir l'équation 1.14). Ces techniques existent déjà dans la librairie **MÉLINA** dans le cadre de la résolution du problème de Helmholtz extérieur. La mise en œuvre de la méthode **CEFRI** dans une librairie en C++ (**MÉLINA++**) pour la résolution des problèmes de propagation d'onde acoustique et électromagnétique est l'un des objectifs de mes travaux de thèse.



## Chapitre 2

# Mise en œuvre de la méthode multipôles rapide pour la méthode CEFRI

### 2.1 Historique de la FMM

Le calcul limitant un algorithme itératif est le coût du produit matrice-vecteur, qui pour une matrice pleine de taille  $N$ , est proportionnel à  $N^2$ . Dans des applications physiques, on est amené à considérer des situations où le nombre d'onde est très grand. Et comme  $N$  croît avec le nombre d'onde de manière quadratique pour une surface, le produit matrice-vecteur devient très coûteux. Pour surmonter ce problème, on fait appel à la méthode multipôles rapide (FMM) qui réduit les coûts en temps de calcul et en place mémoire du produit matrice-vecteur. Historiquement, les méthodes multipôles rapides ont été introduites pour le problème à N-corps de l'astronomie ([27]). Elles ont été ensuite étendues au problème des équations de Laplace résolues par des méthodes intégrales (V. Rokhlin et L. Greengard dans [28] et [29]). Le cas des équations de Helmholtz a été abordé par V. Rokhlin dans le cas 2D dans l'article [19] puis en 3D dans [20]. Ces algorithmes conduisent à une complexité en  $O(N^{\frac{3}{2}})$  pour la méthode à un niveau et  $N \ln N$  pour la méthode multiniveaux. Le développement FMM le plus répandu est basé sur la série de Gegenbauer ([30]).

### 2.2 Exemple d'application de la FMM en configuration 1D

La méthode multipôles rapide notée FMM permet de réaliser de manière économique des produits matrice-vecteur. Nous présentons dans ce qui suit le principe et les notions fondamentales de la méthode FMM à partir d'un cas simplifié en dimension 1. Pour cela, nous considérons une matrice  $A$  définie par la formule suivante :

$$A_{ij} = \frac{1}{x_i - y_j}.$$

On se propose d'accélérer le produit matrice-vecteur  $AV$  où  $V$  est un vecteur donné de  $\mathbb{R}^N$ . Il s'agit d'accélérer le calcul de l'expression suivante :

$$\forall i = \{1, \dots, N\}, \quad \sum_{j=1}^N \frac{1}{x_i - y_j} \mathbf{V}_j.$$

Les points  $x_i$  sont nommés "points observateurs" ou "points cibles" et les  $y_j$  sont appelés "points sources". L'idée clef de la méthode FMM est le découplage des points  $x_i$  et  $y_j$ . Pour cela, nous utilisons un regroupement adéquat de ces points comme suit : soit le segment  $[0,2]$  tel que les points  $x_i$  et  $y_j$  sont répartis de manière approximativement uniforme respectivement sur  $[0,1[$  et  $]1,2]$ . On regroupe ces points dans des boîtes : les  $x_i$  sont dans la boîte  $B_1 = [0, \frac{1}{2}[$  et la boîte  $B_2 = [\frac{1}{2}, 1[$  et les points  $y_j$  sont dans la boîte  $B_3 = [1, \frac{3}{2}[$  et la boîte  $B_4 = [\frac{3}{2}, 2[$ . Notons  $C_k$  le centre de la boîte  $B_k$  avec  $k \in \{1, \dots, 4\}$ .

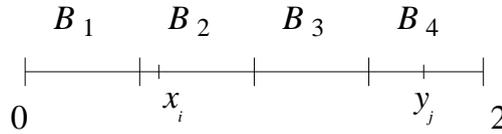


FIGURE 2.1 – Constuction des boîtes 1D

Si  $x_i \in B_2$  et  $y_j \in B_4$  (voir Fig. 2.1) alors

$$|C_2 - x_i| < |C_2 - y_j|. \quad (2.1)$$

En écrivant  $A_{ij}$  sous la forme suivante :

$$\begin{aligned} \frac{1}{x_i - y_j} &= \frac{1}{x_i - C_2 + C_2 - y_j} \\ &= \frac{1}{(C_2 - y_j) \left(1 - \frac{C_2 - x_i}{C_2 - y_j}\right)} \end{aligned}$$

et en utilisant (2.1) on a

$$A_{ij} = \frac{1}{C_2 - y_j} \sum_{n=0}^{\infty} \left( \frac{C_2 - x_i}{C_2 - y_j} \right)^n.$$

Si on se donne une tolérance  $\varepsilon > 0$ , il existe un nombre  $L_\varepsilon$  tel que

$$\left| \frac{1}{x_i - y_j} - \sum_{l=0}^{L_\varepsilon} \frac{(C_2 - x_i)^l}{(C_2 - y_j)^{l+1}} \right| \leq \varepsilon.$$

Soit  $\alpha \in \mathbb{R}$  tel que

$$\left| \frac{C_2 - x_i}{C_2 - y_j} \right| < \alpha < 1$$

alors

$$\left| \sum_{l=L_\varepsilon+1}^{\infty} \frac{(C_2 - x_i)^l}{(C_2 - y_j)^{l+1}} \right| \leq C \left( \frac{1}{1 - \alpha} \right) \alpha^{L_\varepsilon+1}.$$

Nous avons choisi  $L_\varepsilon$  tel que  $\tilde{c} \alpha^{L_\varepsilon+1} \leq \varepsilon$  avec  $\tilde{c} = C \frac{1}{1 - \alpha}$ . Ceci implique

$$\begin{aligned} \ln \tilde{c} + (L_\varepsilon + 1) \ln \alpha &\leq \ln \varepsilon \\ L_\varepsilon &\geq \frac{\ln \varepsilon^{-1} + \ln \tilde{c}}{\ln \alpha^{-1}} - 1 \\ L_\varepsilon &\simeq \ln \varepsilon^{-1}. \end{aligned}$$

La matrice  $A$  peut être décomposée sous la forme suivante :

$$A_{ij} = A_{ij}^{proche} + A_{ij}^{loin}$$

avec

$$A_{ij}^{proche} = \begin{cases} A_{ij} & \text{si } (x_i \in B_2 \text{ et } y_j \in B_3) \\ 0 & \text{sinon} \end{cases}$$

et

$$A_{ij}^{loin} = \begin{cases} A_{ij} & \text{si } (x_i \in B_2 \text{ et } y_j \in B_4) \text{ et } (x_i \in B_1 \text{ et } y_j \in B_3 \cup B_4) \\ 0 & \text{sinon.} \end{cases}$$

La matrice  $A^{proche}$  reste inchangée car la condition de convergence (2.1) n'est pas vérifiée. Par contre pour la matrice  $A^{loin}$ , nous utilisons l'approximation que nous avons écrite précédemment. Ainsi, on a pour tout  $i \in \{1, \dots, N\}$  tel que  $x_i \in B_2$  :

$$\begin{aligned} (AV)_i &= \sum_{j=1}^N A_{ij} \mathbf{V}_j \\ &= \sum_{j/y_j \in B_3} \frac{1}{x_i - y_j} \mathbf{V}_j + \sum_{j/y_j \in B_4} \frac{1}{C_2 - y_j} \sum_{n=0}^{L_\varepsilon} \left( \frac{C_2 - x_i}{C_2 - y_j} \right)^n \mathbf{V}_j \\ &= \sum_{j/y_j \in B_3} \frac{1}{x_i - y_j} \mathbf{V}_j + \sum_{n=0}^{L_\varepsilon} (C_2 - x_i)^n \sum_{j/y_j \in B_4} \frac{\mathbf{V}_j}{(C_2 - y_j)^{n+1}}. \end{aligned}$$

On peut conclure que l'expression  $(AV)_i$  est décomposée en une somme d'interactions proches ( $x_i$  et  $y_j$  appartiennent à des boîtes voisines) et d'interactions lointaines ( $x_i$  et  $y_j$  appartiennent à des boîtes non voisines), où les interactions lointaines sont approchées par la FMM.

### Etude de la complexité

On désigne par  $B_i$  la boîte contenant le point observateur  $x_i$  et  $\mathcal{V}(B_i)$  la réunion des boîtes voisines avec  $B_i$ . En général, " $B$  est une boîte voisine de  $B'$ " veut dire que les deux boîtes  $B$  et  $B'$  se touchent. On note  $M$  le nombre de boîtes regroupant les points  $x_i$  et  $y_j$ .

– Interactions proches :

$$\sum_{j/y_j \in \mathcal{V}(B_i)} \frac{\mathbf{V}_j}{x_i - y_j}.$$

La complexité du calcul des interactions proches est de  $O\left(\frac{N^2}{M}\right)$ .

– Interactions lointaines :

$$\sum_{n=1}^{L_\varepsilon} (C_i - x_i)^n \sum_{j/y_j \notin \mathcal{V}(B_i)} \frac{\mathbf{V}_j}{(C_i - y_j)^{n+1}}.$$

Ce calcul se fait en 2 étapes :

1. Pour chaque boîte  $B_i$  et pour chaque  $n \in \{1, \dots, L_\varepsilon\}$ , nous calculons l'expression

$$S_i^n = \sum_{j/y_j \notin \mathcal{V}(B_i)} \frac{\mathbf{V}_j}{(C_i - y_j)^{n+1}}.$$

Cette étape nécessite  $O(M.L_\varepsilon.N)$  opérations.

2. Pour chaque boîte  $B_i$ , pour chaque point  $x_i$  de  $B_i$ , nous calculons

$$\sum_{n=0}^{L_\varepsilon} (C_i - x_i)^n S_i^n.$$

Cette somme nécessite  $O(N.L_\varepsilon)$  opérations.

La complexité du calcul des interactions lointaines est de  $O(M.N.\ln \varepsilon^{-1})$ . On en déduit que la complexité totale est

$$\frac{N^2}{M} + M.N.\ln \varepsilon^{-1}.$$

Le choix du nombre de boîtes se fait de façon à minimiser ce coût algorithmique. Ainsi, le choix optimal est  $M \sim \sqrt{N}$  et la complexité optimale est  $(N^{3/2}.\ln \varepsilon^{-1})$ .

**Remarque 2.2.1.** *Les boîtes utilisées sont de même taille. On parle dans ce cas d'une méthode **FMM à un niveau**. Pour l'exemple présenté précédemment, nous avons remarqué que la condition de convergence (2.1) n'est pas vérifiée pour  $x_i \in B_2$  et  $y_j \in B_3$ . Mais nous pouvons subdiviser à nouveau chacune de ces deux boîtes en deux boîtes identiques. On les note  $B_{21}$ ,  $B_{22}$ ,  $B_{31}$ ,  $B_{32}$ . Dans ce cas, si  $x_i \in B_{22}$  et si  $y_j \in B_{32}$  la condition de convergence est vérifiée alors on peut suivre le même raisonnement que précédemment. Dans ce cas nous avons fait un regroupement des points en fonction de l'éloignement par rapport au point observateur. On parle de méthode **FMM multiniveaux**.*

## 2.3 Application de la FMM 1 niveau à CEFRI en 3D

Maintenant que la structure générique de l'algorithme multipôles rapide a été présentée, nous allons étudier son application au problème de Helmholtz en domaine non borné résolu par la méthode **CEFRI**. Plus précisément, nous étudions le produit de la matrice  $\mathcal{H}$ , vue dans le chapitre 1, par un vecteur donné.

### 2.3.1 Construction du pavage et notion de voisinage

#### Construction géométrique des boîtes :

Soit un ensemble de points  $(x_\sigma)$  repartis de manière assez uniforme sur un domaine  $\Sigma$  et des points  $(y_\gamma)$  repartis sur  $\Gamma$ . On construit un pavage de la région de l'espace qui contient les  $(x_\sigma)$  et les  $(y_\gamma)$ . En 3D, nous choisissons un cube qui contient toutes les particules  $(x_\sigma)$  et  $(y_\gamma)$ . Nous construisons le pavage de manière récursive : à chaque étape, nous subdivisons chaque cube en huit cubes identiques. Ainsi de suite jusqu'à obtenir une taille de boîtes convenables et un nombre de particules par boîte convenable. Par abus d'écriture, on notera  $B_\sigma$  la boîte contenant un point  $x_\sigma$  donné et  $C_\sigma$  le centre de  $B_\sigma$ . Pour une boîte  $B$  du pavage, on note  $\mathcal{V}(B)$  l'ensemble de boîtes "proches" de  $B$ .  $B' \notin \mathcal{V}(B)$  signifie qu'il existe  $\alpha < 1$  tel que

$$\forall x \in B \quad \forall y \in B' \quad |y - C_{B'}| + |x - C_B| < \alpha |C_B - C_{B'}|$$

avec  $C_B$  est le centre de la boîte  $B$  et  $C_{B'}$  le centre de la boîte  $B'$ . La matrice  $\mathcal{H}$  ( voir chapitre 1) peut être décomposée sous la forme suivante :

$$\mathcal{H} = \mathcal{H}^{proche} + \mathcal{H}^{loin}$$

avec

$$\mathcal{H}_{lm}^{proche} = \begin{cases} \mathcal{H}_{lm} & \text{si } x_\sigma \text{ et } y_\gamma \text{ appartiennent à des boîtes proches} \\ 0 & \text{sinon.} \end{cases}$$

et

$$\mathcal{H}_{\sigma\gamma}^{loin} = \begin{cases} 0 & \text{si } x_\sigma \text{ et } y_\gamma \text{ appartiennent à des boîtes proches} \\ \mathcal{H}_{\sigma\gamma} & \text{sinon} \end{cases}$$

La matrice  $\mathcal{H}^{proche}$  est calculée d'une manière classique. La partie  $\mathcal{H}^{loin}$  est considérée au travers de la méthode FMM. Le produit de la matrice  $\mathcal{H}$  par un vecteur donné  $\mathbf{v}$  est

$$(\mathcal{H}\mathbf{v})_\sigma = (\mathcal{H}\mathbf{v})_\sigma^{proche} + (\mathcal{H}\mathbf{v})_\sigma^{loin} \quad \text{tels que}$$

$$\begin{aligned} (\mathcal{H}\mathbf{v})_\sigma^{proche} &= \sum_{B' \in \mathcal{V}(B_\sigma)} \sum_{\gamma/y_\gamma \in B'} H(x_\sigma, y_\gamma) \mathbf{v}_\gamma \\ &= \sum_{B' \in \mathcal{V}(B_\sigma)} \sum_{\gamma/y_\gamma \in B'} \frac{\partial K(x_\sigma, y_\gamma)}{\partial n_{y_\gamma}} \mathbf{v}_\gamma \\ &= \sum_{B' \in \mathcal{V}(B_\sigma)} \sum_{\gamma/y_\gamma \in B'} \frac{\partial}{\partial n_{y_\gamma}} \left( \left( \frac{\partial}{\partial \mathbf{n}_{x_\sigma}} - ik \right) G(x_\sigma, y_\gamma) \right) \mathbf{v}_\gamma. \end{aligned}$$

Le calcul de  $(\mathcal{H}\mathbf{v})_\sigma^{loin}$  est basé sur un développement judicieux du noyau de Green. En fait, l'idée de base de la méthode multipôles rapide est de tenter de séparer les variables  $x_\sigma$  et  $y_\gamma$ . Il nous faut réécrire le noyau de Green différemment. Pour cela, nous présentons quelques outils mathématiques que nous allons utiliser par la suite.

### 2.3.2 Développement multipolaire

La fonction de Green de l'équation de Helmholtz pour un domaine tridimensionnel infini est, à un coefficient multiplicatif près, la fonction de Hankel sphérique de première espèce et d'ordre zéro :

$$G(k, |x_\sigma - y_\gamma|) = \frac{ik}{4\pi} h_0(k|x_\sigma - y_\gamma|).$$

**Théorème 2.3.1.** : (Théorème d'addition de Gegenbauer)

$$h_0(k|r_0 + r|) = \sum_{q=0}^{\infty} (-1)^q (2q + 1) h_q(k|r_0|) j_q(k|r|) P_q(\cos(r, r_0))$$

où  $r$  et  $r_0$  sont des vecteurs tel que  $|r| < |r_0|$  avec

$$\begin{aligned} r &= r_\sigma - r_\gamma \quad \text{avec} \quad r_\sigma = x_\sigma - C_\sigma \quad \text{et} \quad r_\gamma = y_\gamma - C_\gamma \\ r_0 &= C_\sigma - C_\gamma \end{aligned}$$

$\cos(r, r_0)$  est le cosinus de l'angle entre les vecteurs  $r$  et  $r_0$ ,  $h_q$  est la fonction de Hankel de première espèce,  $j_q$  est la fonction de Bessel et  $P_q$  est un polynôme de Legendre.

La convergence de cette série est assurée si et seulement si  $r$  et  $r_0$  vérifient la relation  $|r| < |r_0|$ . Cette contrainte oblige de distinguer les couples de boîtes voisines des couples de boîtes non voisines.

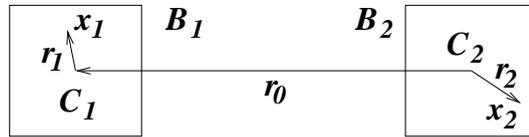


FIGURE 2.2 – Cas de deux boîtes non voisines ( $\sigma = 1$  et  $\gamma = 2$ ).

**Proposition 2.3.1.** : (Formule de Funk-Hecke)

$$j_q(k|r|) P_q(\cos(r, r_0)) = \frac{(i)^q}{4\pi} \int_{\mathcal{S}} P_q(\cos(s, r_0)) e^{ik(s, x_\sigma - C_\sigma)} e^{-ik(s, y_\gamma - C_\gamma)} ds$$

où  $\mathcal{S}$  est la sphère unité de  $\mathbb{R}^3$  et  $(\dots)$  désigne le produit scalaire de  $\mathbb{R}^3$ .

On tronque la série de Gegenbauer et en utilisant la formule de Funk-Hecke, on obtient :

$$h_0(k|r_0 + r|) \simeq \sum_{q=0}^Q \frac{(-i)^q}{4\pi} (2q + 1) h_q(k|r_0|) \int_{\mathcal{S}} P_q(\cos(s, r_0)) e^{ik(s, r)} ds$$

où  $Q$  est le paramètre de troncature de la série de Gegenbauer

On note :

$$\mathcal{T}_{r_0}^Q(s) = \sum_{q=0}^Q (-i)^q (2q + 1) P_q(\cos(s, r_0)) h_q(k|r_0|). \quad (2.2)$$

$\mathcal{T}_{r_0}^Q(s)$  est appelée fonction de transfert.

Ensuite, on fait commuter la sommation avec l'intégrale donc on peut écrire

$$\begin{aligned} h_0(k|r_0 + r|) &\simeq \frac{1}{4\pi} \int_S \mathcal{T}_{r_0}^Q(s) e^{ik(s,r)} ds \\ &= \frac{1}{4\pi} \int_S \mathcal{T}_{r_0}^Q(s) e^{ik(s,x_\sigma - C_\sigma)} e^{-ik(s,y_\gamma - C_\gamma)} ds. \end{aligned}$$

Dans la suite, nous allons exploiter ces résultats dans le but de déterminer le calcul des interactions lointaines  $(\mathcal{H}\mathbf{v})_\sigma^{loin}$ .

$$\begin{aligned} H(x_\sigma, y_\gamma) &= \frac{\partial}{\partial \mathbf{n}_{y_\gamma}} \left[ \left( \frac{\partial}{\partial \mathbf{n}_{x_\sigma}} - ik \right) G(x_\sigma, y_\gamma) \right] \\ &= \frac{ik}{4\pi} \frac{\partial}{\partial \mathbf{n}_{y_\gamma}} \left[ \left( \frac{\partial}{\partial \mathbf{n}_{x_\sigma}} - ik \right) h_0(k|x_\sigma - y_\gamma|) \right]. \end{aligned}$$

Nous avons besoin alors de déterminer le calcul de ces quantités :

$$\frac{\partial h_0(k|x_\sigma - y_\gamma|)}{\partial \mathbf{n}_{y_\gamma}}, \quad \frac{\partial h_0(k|x_\sigma - y_\gamma|)}{\partial \mathbf{n}_{x_\sigma}} \quad \text{et} \quad \frac{\partial^2 h_0(k|x_\sigma - y_\gamma|)}{\partial \mathbf{n}_{y_\gamma} \partial \mathbf{n}_{x_\sigma}}.$$

$$\begin{aligned} \frac{\partial h_0(k|x_\sigma - y_\gamma|)}{\partial \mathbf{n}_{x_\sigma}} &\simeq \frac{ik}{4\pi} \int_S \mathcal{T}_{r_0}^Q(s) (s, \mathbf{n}_{x_\sigma}) e^{ik(s,x_\sigma - C_\sigma)} e^{-ik(s,y_\gamma - C_\gamma)} ds, \\ \frac{\partial h_0(k|x_\sigma - y_\gamma|)}{\partial \mathbf{n}_{y_\gamma}} &\simeq \frac{-ik}{4\pi} \int_S \mathcal{T}_{r_0}^Q(s) (s, \mathbf{n}_{y_\gamma}) e^{ik(s,x_\sigma - C_\sigma)} e^{-ik(s,y_\gamma - C_\gamma)} ds, \\ \frac{\partial^2 h_0(k|x_\sigma - y_\gamma|)}{\partial \mathbf{n}_{y_\gamma} \partial \mathbf{n}_{x_\sigma}} &\simeq \frac{k^2}{4\pi} \int_S \mathcal{T}_{r_0}^Q(s) (s, \mathbf{n}_{x_\sigma}) (s, \mathbf{n}_{y_\gamma}) e^{ik(s,x_\sigma - C_\sigma)} e^{-ik(s,y_\gamma - C_\gamma)} ds, \end{aligned}$$

$$H(x_\sigma, y_\gamma) \simeq \frac{-ik^3}{(4\pi)^2} \int_S \mathcal{T}_{r_0}^Q(s) (1 - (s, \mathbf{n}_{x_\sigma})) e^{ik(s,x_\sigma - C_\sigma)} e^{-ik(s,y_\gamma - C_\gamma)} (s, \mathbf{n}_{y_\gamma}) ds.$$

L'étape suivante consiste à discrétiser l'intégrale. Pour cela, nous choisissons  $S$  points  $s_p$  de quadrature sur la sphère  $S$  et un ensemble de poids  $w_p$  tels que

$$\int_S f(s) ds \simeq \sum_{p=1}^S f(s_p) w_p.$$

La discrétisation sur la sphère unité nous donne l'approximation suivante

$$H(x_\sigma, y_\gamma) \simeq \frac{-ik^3}{(4\pi)^2} \sum_{p=1}^S w_p (1 - (s_p, \mathbf{n}_{x_\sigma})) e^{ik(s_p, x_\sigma - C_\sigma)} \mathcal{T}_{r_0}^Q(s_p) e^{-ik(s_p, y_\gamma - C_\gamma)} (s_p, \mathbf{n}_{y_\gamma}).$$

On remarque bien que deux approximations sont nécessaires pour l'évaluation numérique de notre calcul. La première est la troncature de la série :

$$\sum_{q=0}^{\infty} (-1)^q (2q+1) h_q(k|r_0|) j_q(k|r|) P_q(\cos(r, r_0)).$$

Le choix du paramètre de troncature  $Q$  se fait de telle sorte qu'on a

$$\left| \frac{ik}{4\pi} \sum_{q=Q+1}^{\infty} (-1)^q (2q+1) h_q(k|r_0|) j_q(k|r|) P_q(\cos(r, r_0)) \right| < \varepsilon$$

où  $\varepsilon$  est l'erreur d'approximation désirée. Une analyse asymptotique du terme  $j_q(k|r|)$  donne

$$j_q(k|r|) \sim \frac{(k|r|)^q}{(2q+1)(2q-1)\dots 3.1} \left( 1 - \frac{(k|r|)^2}{2(2q+3)} + \dots \right)$$

donc  $q \mapsto j_q(k|r|)$  converge très vite vers 0 dès que  $q > k|r|$ . L'analyse asymptotique de  $h_q$  est donnée par le comportement de la fonction de Neumann  $y_q$  du fait que  $h_q = j_q + iy_q$ . Asymptotiquement  $y_q(k|r_0|)$  s'écrit

$$y_q(k|r_0|) \sim -\frac{(2q-1)\dots 3.1}{(k|r_0|)^{q+1}} \left( 1 - \frac{(k|r_0|)^2}{2(1-2q)} + \dots \right)$$

ce qui implique que  $q \mapsto h_q(k|r_0|)$  diverge lorsque  $q$  tend vers  $\infty$  dès que  $q > k|r_0|$ . Alors, nous ne pouvons pas augmenter indéfiniment la valeur du paramètre de troncature  $Q$ . Par suite, on a

$$(2q+1)j_q(k|r|)y_q(k|r_0|) \sim \frac{1}{k|r_0|} \left( \frac{|r|}{|r_0|} \right)^q.$$

Cette analyse asymptotique montre que ce produit se comporte comme une série géométrique de raison  $\frac{|r|}{|r_0|}$ . Nous montrons les variations de  $j_q(30)$  et  $y_q(60)$  dans la figure (2.3).

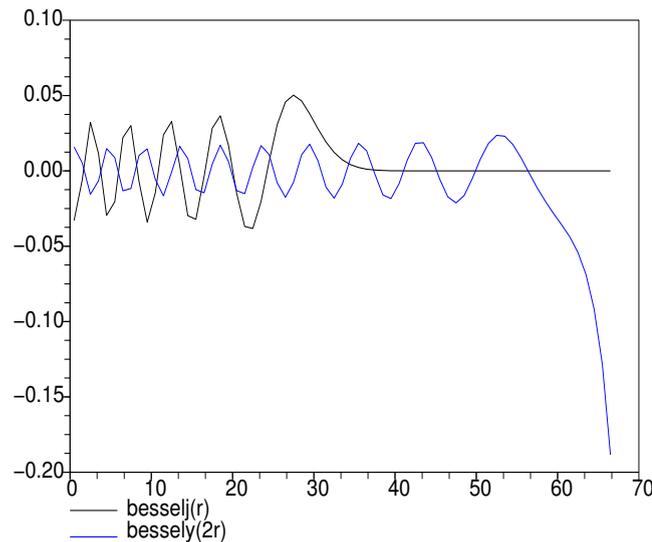


FIGURE 2.3 – Variation des fonctions de Bessel et de Hankel

Pour la fonction  $q \mapsto j_l(30)$ , on voit que tout d'abord (pour une trentaine de termes environ) il n'y a pas de convergence du tout. Puis à partir de  $q > 30 = k|r|$  apparaît une

zone où la décroissance est assez rapide. Inversement,  $q \mapsto y_q(60)$  diverge très vite lorsque  $q$  devient plus grand que  $60 = k|r_0|$ . Avec la condition  $|r| < |r_0|$ , le comportement de décroissance rapide de  $j_q(k|r|)$  intervient avant la divergence de  $y_q(k|r_0|)$ , la convergence vers 0 du produit  $j_q(k|r|) y_q(k|r_0|)$  a lieu dès que  $q$  devient plus grand que  $k|r|$ . Cette convergence aboutit au résultat suivant : Pour tout  $\varepsilon$ , il existe  $Q$  tel que

$$\left| \frac{ik}{4\pi} \sum_{q=Q+1}^{\infty} (-1)^q (2q+1) h_q^{(1)}(k|r_0|) j_q(k|r|) P_q(\cos(r, r_0)) \right| < \varepsilon.$$

Dans la littérature, on trouve de nombreuses formules empiriques. Par exemple une étude de WC Chew [31] a abouti à la formule suivante :

$$Q = k|r| + C(k|r|)^{\frac{1}{3}}.$$

Après une série de tests numériques, nous choisissons  $C = 2.5$  le plus adéquat à notre configuration.

La seconde approximation va être la discrétisation sur la sphère unité. La discrétisation de l'intégrale sur la sphère unité revient à déterminer des points de quadrature  $s_p$  et des poids  $w_p$  qui intègrent exactement les harmoniques sphériques  $Y_{lm}$ . Nous rappelons que

$$Y_{lm}(\theta, \varphi) = \sqrt{\frac{2l+1}{4\pi} \frac{(l-m)!}{(l+m)!}} P_{\sigma}^m(\cos \theta) e^{im\varphi}.$$

En fait, en tenant compte du développement suivant

$$e^{i(s,r)} = \sum_{q=0}^{\infty} (-1)^q (2q+1) j_q(k|r|) P_q(\cos(s, r)) \quad (2.3)$$

et le théorème d'addition des harmoniques sphériques [32] :

**Théorème 2.3.2.** Soient deux vecteurs de coordonnées sphériques  $(r, \theta, \varphi)$  et  $(r', \theta', \varphi')$ . Notons  $\gamma$  l'angle entre ces deux vecteurs. Alors

$$P_{\sigma}(\cos \gamma) = \frac{4\pi}{2l+1} \sum_{m=-l}^l Y_{lm}^*(\theta', \varphi') Y_{lm}(\theta, \varphi)$$

où  $\cos \gamma = \cos \theta \cos \theta' + \sin \theta \sin \theta' \cos(\varphi - \varphi')$ .

L'intégrand  $\mathcal{T}_{r_0}^Q(s) e^{ik(s,r)}$  peut alors être approché par une combinaison d'harmoniques sphériques. On choisit des points  $s_p$  de la forme  $(\theta_j, \varphi_i)$  tels que  $0 \leq i < I$  et  $0 \leq j < J$ . On suppose qu'une bijection a été établie entre  $\{s_p, p = 1, \dots, IJ - 1\}$  et  $\{(\theta_j, \varphi_i), i = 1 \dots I - 1, j = 1, \dots, J - 1\}$ . Une formule d'intégration de Gauss-Legendre à  $J$  points permet d'intégrer exactement tous les polynômes de Legendre de degré inférieur ou égal à  $2J$ . Si on note  $s_p = (\theta_j, \varphi_i)$ , on définit le poids  $w_p$  associé à  $s_p$  comme suit

$$w_p = \frac{2\pi}{I} GL_j$$

avec  $GL_j$  les poids de Gauss-Legendre. Les  $\theta_j$  sont définis sur  $[0, \pi]$  et  $\varphi_i = \frac{2\pi i}{I}$ . Si nous choisissons  $I = 2Q + 1$  et  $J = Q + 1$ , nous obtenons une formule exacte pour toutes les harmoniques sphériques de degré inférieur ou égal à  $2Q$ . On peut donc prendre

$$S = (2Q + 1) \times (Q + 1).$$

Pour les études de complexité que nous allons présenter dans la suite, nous considérons  $Q \approx k|r|$  et donc  $S \approx 2(k|r|)^2$ .

Le calcul des interactions lointaines consiste alors à :

$$\begin{aligned} (\mathcal{H}\mathbf{v})_\sigma^{loin} &= \sum_{B' \notin \mathcal{V}(B_\sigma)} \sum_{\gamma/y_\gamma \in B'} \frac{-ik^3}{(4\pi)^2} \sum_{p=1}^S w_p (1 - (s_p, \mathbf{n}_{x_\sigma})) e^{ik(s_p, x_\sigma - C_\sigma)} \mathcal{T}_{r_0}^Q(s_p) \\ &\quad e^{-ik(s_p, y_\gamma - C_\gamma)} (s_p, \mathbf{n}_{y_\gamma}) \mathbf{v}_\gamma \\ &= \frac{-ik^3}{(4\pi)^2} \sum_{p=1}^S w_p (1 - (s_p, \mathbf{n}_{x_\sigma})) e^{ik(s_p, x_\sigma - C_\sigma)} \sum_{B' \notin \mathcal{V}(B_\sigma)} \mathcal{T}_{r_0}^Q(s_p) \\ &\quad \sum_{\gamma/y_\gamma \in B'} e^{-ik(s_p, y_\gamma - C_\gamma)} (s_p, \mathbf{n}_{y_\gamma}) \mathbf{v}_\gamma. \end{aligned}$$

### 2.3.3 Algorithme et calcul de la complexité de la FMM un niveau

– Etape 0 : Partie indépendante du vecteur  $\mathbf{v}$

$$\begin{aligned} \forall r_0 \quad \forall s_p \quad \mathcal{T}_{r_0}^Q(s_p) &= \sum_{q=0}^Q \frac{(-i)^q}{4\pi} (2q + 1) P_q(\cos(s_p, r_0)) h_q(k|r_0|). \\ \forall y_\gamma \quad \forall s_p \quad f_\gamma(s_p) &= (s_p, \mathbf{n}_{y_\gamma}) e^{-ik(s_p, y_\gamma - C_\gamma)}. \\ \forall x_\sigma \quad \forall s_p \quad g_\sigma(s_p) &= (1 - (s_p, \mathbf{n}_{x_\sigma})) e^{ik(s_p, x_\sigma - C_\sigma)}. \end{aligned}$$

– Etape 1 – “Far fields” : Rassemblement de l’information dans le centre de la boîte source.

$$\forall B_\gamma \quad \forall s_p \quad F_{B_\gamma}(s_p) = \sum_{\gamma/y_\gamma \in B_\gamma} f_\gamma(s_p) \mathbf{v}_\gamma.$$

– Etape 2 – “Close fields” : Transfert de l’information des boîtes sources  $B_\gamma$  vers la boîte cible  $B_\sigma$ .

$$\forall B_\sigma \quad \forall s_p \quad G_{B_\sigma}(s_p) = \sum_{B_\gamma \notin \mathcal{V}(B_\sigma)} \mathcal{T}_{r_0}^Q(s_p) F_{B_\gamma}(s_p).$$

- Etape 3 : Passage de l'information de  $C_\sigma$  vers le point cible  $x_\sigma$ .

$$\forall B_\sigma \quad \forall \sigma/x_\sigma \in B_\sigma \quad (\mathcal{H}\mathbf{v})_\sigma^{loin} = \frac{ik^3}{(4\pi)^2} \sum_{p=1}^S w_p g_\sigma(s_p) G_{B_\sigma}(s_p)$$

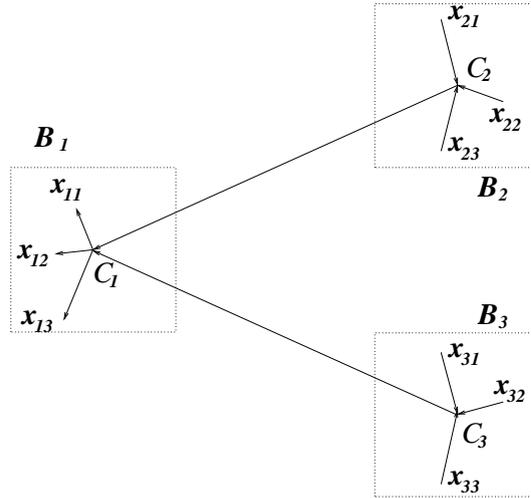


FIGURE 2.4 – Schématisation du passage de l'information : cas d'une boîte cible et deux boîtes sources  $B_1$  et  $B_2$

### Complexité des interactions lointaines

Notons  $N_\Sigma$  (resp.  $N_\Gamma$ ) le nombre total de degrés de liberté sur  $\Sigma$  (resp sur  $\Gamma$ ) et  $M_\Sigma$  (resp.  $M_\Gamma$ ) le nombre de boîtes contenant la surface  $\Sigma$  (resp. la surface  $\Gamma$ ). La complexité du calcul des interactions lointaines résulte des décomptes par étapes suivants :

- L'étape 0 est une étape indépendante du vecteur  $\mathbf{v}$  et du second membre du système qui nécessite un nombre d'opérations de l'ordre de  $SQM_\Sigma + N_\Gamma S + N_\Sigma S$  et une place mémoire de l'ordre de  $SM_\Sigma + N_\Gamma S + N_\Sigma S$ .
- L'étape 1 nécessite un coût de calcul de l'ordre de  $N_\Gamma S$  opérations et un coût de place mémoire de l'ordre de  $M_\Gamma S$ . L'étape 2 coûte de l'ordre de  $M_\Gamma M_\Sigma S$  opérations et  $O(SM_\Sigma)$  place mémoire.
- L'étape 3 nécessite un coût de calcul de l'ordre de  $N_\Sigma S$  opérations et une place mémoire en  $O(N_\Sigma)$ .

Comme le nombre de degrés de liberté sur  $\Sigma$  est du même ordre que celui sur  $\Gamma$  qu'on désigne par  $N$  dans la suite. Le nombre de boîtes  $M_\Sigma$  est du même ordre que le nombre de boîtes  $M_\Gamma$  qu'on note  $M$ . Par suite le calcul des interactions lointaines nécessite un nombre d'opérations de l'ordre de  $SQM + M^2 S + NS$  et une place de mémoire de l'ordre de  $SM + NS + MS$ . Si on note  $d$  le diamètre des boîtes comme  $Q \sim k|r|$ , on a  $Q \sim kd$  et  $S \sim (kd)^2$ . Cela donne  $S \sim \frac{N}{M}$  et le coût du calcul des interactions lointaines devient  $O\left(MN + \frac{N^2}{M}\right)$  et  $O\left(\frac{N^2}{M}\right)$  place mémoire.

### Complexité des interactions proches

En adoptant les mêmes notations que précédemment, le calcul des interactions proches nécessite  $O\left(\frac{N^2}{M}\right)$  opérations et  $O(N)$  place mémoire.

En conclusion, le produit de la matrice  $\mathcal{H}$  par un vecteur nécessite  $O\left(\frac{N^2}{M} + NM\right)$  opérations et  $O\left(N + \frac{N^2}{M}\right)$  place mémoire. Si on choisit  $M = \sqrt{N}$  dans le but de minimiser l'expression  $MN + \frac{N^2}{M}$  alors la complexité optimale est  $N^{3/2}$  pour les calculs et pour la place mémoire.

## 2.4 Application de la FMM multiniveaux à CEFRI en 3D

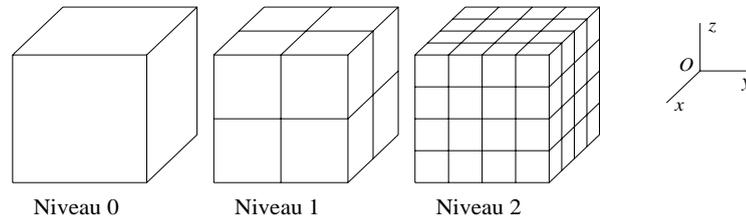
### 2.4.1 Présentation générale

La méthode FMM multiniveaux repose sur une division de chaque boîte en huit boîtes plus petites. Donc, de nouvelles boîtes non voisines sont ainsi créées. Par suite, les calculs des interactions proches sont limitées aux boîtes adjacentes au niveau le plus fin de la subdivision. La FMM multiniveaux permet d'obtenir une complexité optimale de l'ordre de  $N \ln N$  ou  $N \ln^2 N$  suivant l'implémentation, plus précisément suivant la méthode d'interpolation choisie, tandis que le stockage est toujours de l'ordre de  $N \ln N$ .

**Remarque 2.4.1.** *Avec la FMM 1 niveau, si les boîtes deviennent nombreuses alors le transfert est de plus en plus long. En utilisant la méthode multiniveaux, on réduit considérablement le nombre de transferts. Pour cela, on commence par utiliser les boîtes les plus grandes de notre pavage. Elles sont en nombre réduit. On effectue pour ce niveau tous les transferts pour les couples de boîtes qui ne sont pas voisines l'une de l'autre.*

#### – Construction du pavage et notion de voisinage

Soit un ensemble de points  $(x_\sigma)$  repartis de manière assez uniforme sur un domaine  $\Sigma$  et des points  $(y_\gamma)$  repartis sur  $\Gamma$ . On construit un pavage de la région de l'espace qui contient les  $(x_\sigma)$  et les  $(y_\gamma)$ . En 3D, on choisit un cube qui contient toutes les particules  $(x_\sigma)$  et  $(y_\gamma)$ , il s'agit du niveau 0. On construit le pavage de manière récursive : à chaque étape, on subdivise chaque cube en huit cubes ou boîtes identiques, ainsi de suite jusqu'à obtenir une taille de boîtes convenable et un nombre de particules par boîte convenable. Si le niveau  $n$  est construit alors le niveau  $(n + 1)$  se fait en subdivisant chaque cube  $B_D^n$  du niveau  $n$  en 8 boîtes de côtés de longueur  $l_n$  avec  $l_{n+1} = \frac{l_n}{2}$ . Ces boîtes sont appelées boîtes filles de  $B_D^n$ . Et  $B_D^n$  est la boîte mère de  $B_d^{n+1}$ , où  $B_d^{n+1}$  désigne une boîte donnée du niveau  $(n + 1)$  contenue dans  $B_D^n$ .



**Remarque 2.4.2.** Les boîtes voisines sont des boîtes ayant au moins un sommet en commun et se trouvant au même niveau du pavage. On conserve la notation  $\mathcal{V}$  pour désigner les boîtes voisines. Définissons de nouveaux ensembles de boîtes :

$$TL^{n+1}(d) =$$

$$\{d' \in \{1, \dots, M_{n+1}\} \text{ tels que } B_{d'}^{n+1} \text{ non voisine de } B_d^{n+1} \text{ et } B_{D'}^n \text{ voisine de } B_D^n\}$$

où  $B_D^n$  (respectivement  $B_{D'}^n$ ) désigne la boîte mère de la boîte  $B_d^{n+1}$  (respectivement  $B_{d'}^{n+1}$ ) et  $M_{n+1}$  est le nombre de boîtes au niveau  $n + 1$ .

**Remarque 2.4.3.** Si la répartition des éléments du maillage est non uniforme alors la complexité des interactions proches devient grande, ce qui fait perdre l'intérêt de la FMM multiniveaux.

#### – Choix du niveau final

On a vu que le passage d'un niveau à un autre plus fin nécessite la division de la longueur des côtés des boîtes par deux. Par suite, en 3D surfacique d'un niveau  $n$  donné à un autre, on garde un nombre de boîtes

$$M_n \approx 4^n.$$

Il en résulte que le niveau le plus fin est donné par la relation suivante

$$n_{fin} \approx \frac{\ln M}{\ln 4}$$

où  $M$  est le nombre de boîtes souhaitées dans le niveau le plus fin.

## 2.4.2 Algorithme : FMM multiniveaux

### – Etape 0 : Partie indépendante du vecteur $\mathbf{v}$

Calcul de la fonction de transfert pour tout niveau  $n$

$$\forall r_{DD'}^n \quad \forall s_p \quad \mathcal{T}_{r_{DD'}^n}^{Q^n}(s_p) = \sum_{q=0}^{Q^n} \frac{(-i)^q}{4\pi} (2q+1) P_q(\cos(s_p, r_{DD'}^n)) h_q^1(k | r_{DD'}^n |)$$

où  $r_{DD'}^n = C_D^n - C_{D'}^n$ ,  $C_D^n$  (resp.  $C_{D'}^n$ ) étant le centre de la boîte donnée  $B_D^n$  (resp.  $B_{D'}^n$ ).  $Q^n$  étant le paramètre de troncature de l'opérateur de transfert au niveau  $n$ . Cette étape doit être effectuée pour tout vecteur  $r_{DD'}^n$ , tel que  $B_{D'}^n \in TL^n(B_D^n)$ , lorsque  $n$  décrit l'ensemble des niveaux.

### Calcul des moments au niveau le plus fin

$$\forall y_\gamma \quad \forall s_p^{(nfin)} \quad f_\gamma(s_p^{(nfin)}) = (s_p^{(nfin)}, \mathbf{n}_{y_\gamma}) e^{-ik(s_p^{(nfin)}, y_\gamma - C_D^{(nfin)})}.$$

$$\forall x_\sigma \quad \forall s_p^{(nfin)} \quad g_\sigma(s_p^{(nfin)}) = (1 - (s_p^{(nfin)}, \mathbf{n}_{x_\sigma})) e^{ik(s_p^{(nfin)}, x_\sigma - C_{D'}^{(nfin)})}.$$

Pour  $n$  allant de  $nfin - 1$  à  $2$ ,  $\forall D \in \{1, \dots, M_n\}$  et  $\forall s_p^{(n)}$

$$h_\sigma(s_p^{(n)}) = e^{ik(s_p^{(n)}, C_D^n - C_d^{n+1})}.$$

Pour  $n$  allant de  $nfin - 1$  à  $2$ ,  $\forall D \in \{1, \dots, M_n\}$  et  $\forall s_p^{(n)}$

$$\tilde{h}_\sigma(s_p^{(n)}) = e^{ik(s_p^{(n)}, C_D^{n+1} - C_D^n)}.$$

– Etape 1 : Calcul de la fonction de radiation lointaine au niveau le plus fin

$$F_D^{(nfin)}(s_p^{(nfin)}) = \sum_{\gamma/x_\gamma \in B_D^{nfin}} f_D(s_p^{(nfin)}) \mathbf{v}_\sigma.$$

– Etape 2 : Agrégation des fonctions jusqu'au niveau 2

pour  $n$  allant de  $nfin - 1$  à  $2$ ,  $\forall D \in \{1, \dots, M_n\}$

$$F_D^{(n)}(s_p^{(n)}) = \sum_{d \in \text{fille}(D, n)} h_\sigma(s_p^{(n)}) \mathcal{I}nt(F_d^{(n+1)})(s_p^{(n)})$$

où  $\text{fille}(D, n) = \{d \text{ tel que } B_d^{n+1} \text{ est une boîte fille de } B_D^n\}$  et  $\mathcal{I}nt$  désigne l'opérateur d'interpolation.

– Etape 3 : Transferts de l'information des boîtes sources  $B_{D'}^n$  vers la boîte cible  $B_D^n$ .

Le transfert se fait en deux étapes :

### Transfert entre boîtes non voisines pour le niveau grossier égal à 2

$\forall D \in \{1, \dots, M_2\}$

$$G_D^{(2)}(s_p^{(2)}) = \sum_{B_{D'}^2, \notin \mathcal{V}(B_D^2)} \mathcal{T}_{r_{DD'}^2}^{Q^2}(s_p) F_{D'}^{(2)}(s_p^{(2)}).$$

Transfert entre boîtes banlieues ce qui revient à calculer

$$\forall n \in \{3, \dots, n_{fin}\}, \forall D \in \{1, \dots, M_n\}$$

$$G_D^{(n)}(s_p^{(n)}) = \sum_{D' \in TL^n(D)} \mathcal{T}_{r_{DD'}^n}^{Q^n}(s_p) F_{D'}^{(n)}(s_p^{(n)}).$$

- Étape 4 : “Désagrégation” des fonctions de radiation locales (les fonctions transférées) jusqu’au niveau le plus fin  $n_{fin}$ .

Pour  $n$  allant de 2 à  $n_{fin} - 1$ ,  $\forall D \in \{1, \dots, M_n\}$

$$G_d^{(n+1)}(s_p^{(n+1)}) = G_d^{(n+1)}(s_p^{(n+1)}) + \mathcal{Ant}(\tilde{h}_\sigma(s_p^{(n)}) G_D^{(n)})(s_p^{(n+1)})$$

$G_D^{(n)}$  est la fonction de radiation locale pour la boîte mère  $B_D^n$  de la boîte  $B_d^{n+1}$  et  $\mathcal{Ant}$  désigne l’opérateur d’antépolation.

- Étape 5 : Passage de l’information de  $C_d^{n_{fin}}$  vers le point observateur.

$$\forall d \in \{1, \dots, M_{n_{fin}}\}, \forall \sigma/x_\sigma \in B_d^{n_{fin}}$$

$$(\mathcal{H}\mathbf{v})_\sigma^{loin} = \frac{ik^3}{(4\pi)^2} \sum_{p=1}^{S_{n_{fin}}} w_p g_\sigma(s_p^{n_{fin}}) G_d^{n_{fin}}(s_p^{n_{fin}}).$$

En conclusion, la méthode multiniveaux consiste à :

- Evaluer les fonctions de radiation au niveau le plus fin.
- Interpoler les fonctions de radiation pour obtenir celles des niveaux précédents.
- Effectuer les transferts à tous les niveaux.
- Antépoler les fonctions de radiation du niveau le plus grossier vers le niveau final.
- Transmettre l’information au point observateur.

On remarque que deux nouvelles étapes par rapport à la méthode FMM 1 niveau sont l’interpolation et l’antépolation. En fait, lors de la montée et de la descente en niveau, on est ramené à déterminer les fonctions de radiation sur des discrétisations différentes. Les opérateurs d’interpolations et d’antépolations permettent de prendre en considération le changement de discrétisation. La présentation de ces notions sera l’objectif de la partie qui suit.

### 2.4.3 Interpolation et antépolation

Le but de l’interpolation est l’évaluation des fonctions de radiation  $F^{(n)}$  en  $S_n$  points de discrétisation au niveau  $n$  à partir des fonctions connues en  $S_{n+1}$  points au niveau  $n + 1$ . En tenant compte du fait qu’au niveau  $n$  les boîtes sont deux fois plus grandes que celles au niveau  $n + 1$ , on a la relation suivante :  $S_n = 4S_{n+1}$ . On remarque aussi

que plus le niveau  $nfin$  est grand plus l'appel à l'opérateur d'interpolation est fréquent. Parmi les méthodes d'interpolation on peut citer par exemple la méthode semi-naïve [33]. C'est un algorithme de complexité  $S^{\frac{3}{2}}$  où  $S$  désigne le nombre de points de discrétisation sur la sphère unité  $\mathcal{S}$ . Une deuxième méthode est une amélioration de la précédente en se ramenant à la FMM 1D ([34], [35], [36]). C'est un algorithme approché et de complexité  $S \ln S$ . Dans [22], E. Darve propose aussi un algorithme de complexité  $S$ . Pour la mise en œuvre numérique on se contente ici d'utiliser la première méthode.

### Interpolation semi-naïve

On fait appel à l'interpolation au cours de l'étape 2 de l'algorithme multiniveaux. En effet, pour  $n$  allant de  $nfin - 1$  à 2 et  $\forall D \in \{1, \dots, M_n\}$

$$\begin{aligned} F_D^{(n)}(s_p^{(n)}) &= \sum_{x_\sigma \in B_D} (s_p^{(n)}, \mathbf{n}_{x_\sigma}) e^{ik(s_p^{(n)}, C_D^{(n)} - x_\sigma)} \mathbf{v}_\sigma \\ &= \sum_{d \in \text{fille}(D, n)} \sum_{x_\sigma \in B_d^{(n+1)}} (s_p^{(n)}, \mathbf{n}_{x_\sigma}) e^{ik(s_p^{(n)}, C_D^{(n)} - C_d^{(n+1)})} e^{ik(s_p^{(n)}, C_d^{(n+1)} - x_\sigma)} \mathbf{v}_\sigma \\ &= \sum_{d \in \text{fille}(D, n)} e^{ik(s_p^{(n)}, C_D^{(n)} - C_d^{(n+1)})} \mathcal{I}nt(F_d^{(n+1)})(s_p^{(n)}). \end{aligned}$$

La fonction  $F_d^{(n+1)}$  peut être évaluée aux points  $s_p^{(n+1)}$  et non pas aux points  $s_p^{(n)}$ . C'est la raison pour laquelle nous utilisons l'opérateur d'interpolation  $\mathcal{I}nt$ . D'après (2.3), la fonction à interpoler peut être approchée comme suit :

$$\begin{aligned} F_d^{(n+1)}(s) &= \sum_{x_\sigma \in B_d} (s, \mathbf{n}_{x_\sigma}) e^{ik(s, C_d^{(n+1)} - x_\sigma)} \mathbf{v}_\sigma \\ &\approx \sum_{x_\sigma \in B_d} (s, \mathbf{n}_{x_\sigma}) v_\sigma \sum_{p=0}^{l'} (2p+1) i^p j_p(k | C_d^{(n+1)} - x_\sigma |) P_p(\cos(s, C_d^{(n+1)} - x_\sigma)). \end{aligned}$$

On a donc pour tout  $s$

$$\begin{aligned} F_d^{(n+1)}(s) &\approx \sum_{p=0}^{l'} (s, \mathbf{n}_{x_\sigma}) (2p+1) i^p \sum_{x_\sigma \in B_d} \mathbf{v}_\sigma j_p(k | C_d^{(n+1)} - x_\sigma |) P_p(\cos(s, C_d^{(n+1)} - x_\sigma)) \\ &:= \tilde{F}_d^{(n+1)}(s). \end{aligned} \quad (2.4)$$

Or,  $\tilde{F}_d^{(n+1)}(s)$  peut être décomposée sur la base des harmoniques sphériques comme suit

$$\tilde{F}_d^{(n+1)}(s) = \sum_{l=0}^{\infty} \sum_{m=-l}^l [\tilde{F}_d^{(n+1)}]_{lm} Y_{lm}(s), \quad (2.5)$$

avec  $[\tilde{F}_d^{(n+1)}]_{lm}$  le coefficient de Fourier relatif à cette base définie par

$$[\tilde{F}_d^{(n+1)}]_{lm} = \int_{\mathcal{S}} Y_{lm}^*(s) \tilde{F}_d^{(n+1)}(s) ds,$$

pour tous les  $-l \leq m \leq l$ ,  $0 \leq l \leq l'$ . L'intérêt des fonctions harmoniques sphériques est que notre fonction est très bien approchée par une combinaison linéaire finie de ces fonctions et que l'on sait construire une quadrature efficace pour ces fonctions. La fonction  $Y_{lm}^*(s)\tilde{F}_d^{(n+1)}(s)$  est de largeur de bande égale à  $2l'$  donc on peut calculer l'intégrale exactement avec  $2l' + 1$  points dans la direction de  $\varphi$  et  $l'$  points dans la direction de  $\theta$ . Soit

$$[\tilde{F}_d^{(n+1)}]_{lm} = \sum_{s_p} w_p Y_{lm}^*(s) \tilde{F}_d^{(n+1)}(s_p).$$

En tenant compte des formules (2.4) et (2.5), on a

$$\begin{aligned} \mathcal{I}nt(F_d^{(n+1)})(s_p^n) &\approx \sum_{\substack{-l \leq m \leq l \\ 0 \leq l \leq l'}} Y_{lm}(s_p^n) \sum_{s_p^{(n+1)}} w_p Y_{lm}^*(s_p^{(n+1)}) \tilde{F}_d^{(n+1)}(s_p^{(n+1)}) \\ &= \sum_{s_p^{(n+1)}} w_p \tilde{F}_d^{(n+1)}(s_p^{(n+1)}) \sum_{\substack{-l \leq m \leq l \\ 0 \leq l \leq l'}} Y_{lm}(s_p^n) Y_{lm}^*(s_p^{(n+1)}). \end{aligned} \quad (2.6)$$

Nous rappelons que les harmoniques sphériques peuvent s'exprimer sous la forme

$$Y_{lm}(\theta, \varphi) = \frac{1}{\sqrt{4\pi}} \bar{P}_{lm}(x) e^{im\varphi}$$

avec  $\bar{P}_{lm}(x)$  la fonction de Legendre normalisée définie par

$$\bar{P}_{lm}(x) = \sqrt{(2l+1) \frac{(l-m)!}{(l+m)!}} P_l(x).$$

**Proposition 2.4.1.** : (Formule de sommation de Christoffel-Darboux)

Pour tout  $x, x'$  dans  $[-1, 1]$  tel que  $x \neq x'$  on a

$$\sum_{l=|m|}^{l'} \bar{P}_{lm}(x) \bar{P}_{lm}(x') = \varepsilon_{l'+1,m} \frac{\bar{P}_{l'+1,m}(x') \bar{P}_{l',m}(x) - \bar{P}_{l'+1,m}(x) \bar{P}_{l',m}(x')}{x' - x}$$

$$\text{avec } \varepsilon_{l,m} = \sqrt{\frac{l^2 - m^2}{4l^2 - 1}}.$$

$$\text{Dans la suite on note } \mathcal{P}_{lm}(x, x') = \varepsilon_{l'+1,m} \frac{\bar{P}_{l'+1,m}(x') \bar{P}_{l',m}(x) - \bar{P}_{l'+1,m}(x) \bar{P}_{l',m}(x')}{x' - x}.$$

En utilisant la formule de sommation de Christoffel-Darboux ,

$$\begin{aligned} \mathcal{I}nt(F_d^{(n+1)})(s_p^n) &\approx \sum_{s_p^{(n+1)}} \frac{w_p}{4\pi} \tilde{F}_d^{(n+1)}(s_p^{(n+1)}) \sum_{\substack{-l \leq m \leq l \\ 0 \leq l \leq l'}} e^{im(\varphi^{(n)} - \varphi_i^{(n+1)})} \mathcal{P}_{l'm}(\cos \theta^{(n)}, \cos \theta_j^{(n+1)}). \end{aligned}$$

En décomposant la somme sur  $s_p$  en une somme sur  $i$  et une somme sur  $j$ , on a

$$\mathcal{I}nt(F_d^{(n+1)})(s_p^n) \approx \sum_{\substack{-l \leq m \leq l \\ 0 \leq l \leq l'}} e^{im\varphi^{(n)}} \sum_{j=1}^{J^{(n+1)}} \frac{w_j}{4\pi} \mathcal{P}_{lm}(\cos \theta^{(n)}, \cos \theta_j^{(n+1)}) \sum_i^{I^{(n+1)}} e^{-im\varphi_i^{(n+1)}} \tilde{F}_d^{(n+1)}(s_p^{(n+1)})$$

où  $s_p^{(n)} = (\theta^{(n)}, \varphi^{(n)})$  et  $s_p^{(n+1)} = (\theta_j^{(n+1)}, \varphi_i^{(n+1)})$ . En conclusion, l’algorithme de l’interpolation semi-naïve se fait en trois phases :

- Transformée de Fourier rapide pour calculer pour tout  $m$

$$\tilde{F}^1(m, \theta_j^{n+1}) = \sum_i^I e^{-im\varphi_i^{(n+1)}} F_d^{(n+1)}(s_p^{(n+1)}).$$

- Produit matrice-vecteur pour tous les angles  $\theta^{(n)}$  d’interpolation

$$\tilde{F}^2(m, \theta^n) = \sum_{j=1}^J \frac{w_j}{4\pi} \mathcal{P}_{lm}(\cos \theta^{(n)}, \cos \theta_j^{(n+1)}) \tilde{F}^1(m, \theta_j^{n+1}).$$

- Transformée de Fourier inverse pour tout  $\varphi^{(n)}$

$$F_d^{(n+1)}(s_p^n) \approx \sum_{\substack{-l \leq m \leq l \\ 0 \leq l' \leq l'}} e^{im\varphi^{(n)}} \tilde{F}^2(m, \theta^n).$$

Pour la même raison qu’il a fallu utiliser l’opérateur d’interpolation du niveau le plus fin vers le niveau le plus grossier, lors des calculs des fonctions de radiation locales au niveau le plus fin, on va faire appel à l’opérateur d’interpolation pour passer d’un niveau à un niveau plus fin. L’algorithme de l’interpolation est identique à celui de l’interpolation mais il est moins précis que celui de l’interpolation. En fait, les fonctions à interpoler sont les fonctions qui dépendent des fonctions de transferts, ce qui n’est pas le cas pour l’interpolation. La présence de l’opérateur de transfert induit des erreurs numériques lors de la “désaggrégation” des fonctions de radiation locales (voir étape 4 de l’algorithme FMM multiniveaux).

## 2.5 Tests numériques : cas FMM un niveau

Nous présentons dans cette partie des expériences numériques en trois dimensions de la résolution du problème (1.15) réalisées à l’aide de l’environnement de programmation MÉLINA. Nous considérons un obstacle sphérique de rayon  $R = 1$  centré à l’origine et soumis à une onde incidente plane. Nous désignons par  $\mathbf{u}_{GMRES}$  la solution donnée par la méthode GMRES.  $\mathbf{u}_{FMM}$  est la solution calculée par un solveur itératif GMRES combiné avec la méthode FMM et  $\mathbf{u}_{EXA}$  la solution exacte du problème. On note

- Temps E.F : temps de calcul des éléments finis intervenant dans le calcul de la matrice  $A$  et du second membre.
- Temps LU : temps de la factorisation LU de la matrice  $A$ .
- Temps GMRES : temps mis après une résolution par la méthode de GMRES.
- Temps GMRES+FMM : temps mis après une résolution par la méthode de GMRES combinée avec la méthode FMM.

Le tableau (2.1) montre bien que nous avons une bonne précision et des temps de calculs très satisfaisants, compte tenu de la taille des problèmes traités. Le gain du calcul de la

|  |                         | $k = 4.1$<br>1024 points | $k = 8.2$<br>4098 points |
|--|-------------------------|--------------------------|--------------------------|
| Opérations communes (s)  | Temps EF                | 0.36                     | 3.09                     |
|  | Temps LU                | 0.16                     | 2.8                      |
| Opérations<br>GMRES sans FMM (s)   | Calcul noyau de green   | 8.40                     | 136.43                   |
|  | Calcul matrice de green | 0.37                     | 19.17                    |
|  | Assem. mat. Green       | 0.02                     | 0.83                     |
|  | Résolution              | 0.22                     | 2.78                     |
| Opérations<br>GMRES+FMM (s)  | Précalculs              | $Q = 4$<br>0.59          | $Q = 8$<br>29.51         |
|  | Résolution              | 0.43                     | 3.91                     |
| $\frac{\  \mathbf{u}_{GMRES} - \mathbf{u}_{EXA} \ _2}{\  \mathbf{u}_{EXA} \ _2}$ |                         | 0.21e-1                  | 0.22e-1                  |
| $\frac{\  \mathbf{u}_{FMM} - \mathbf{u}_{EXA} \ _2}{\  \mathbf{u}_{EXA} \ _2}$   |                         | 0.19e-1                  | 0.22e-1                  |

TABLE 2.1 – Temps d'exécution et précision pour GMRES et GMRES+FMM : cas 1024 points sur  $\Gamma$  et  $\Sigma$  et 4098 points sur  $\Gamma$  et  $\Sigma$ .

FMM est dû au fait que pour un point donné de  $\Sigma$  on ne calcule plus tous les noyaux de Green pour tout point de  $\Gamma$  comme c'est fait dans MÉLINA et qui est très couteux surtout lorsqu'on a un très grand nombre de nœuds sur les deux frontières du domaine de calcul. En fait le produit FMM est décomposé en somme d'interactions proches et lointaines comme c'est présenté dans la section 2.3. Le calcul des interactions proches nécessite le calcul des noyaux de Green pour un nombre de points sur  $\Gamma$  bien défini et qui est nettement inférieur à celui de MÉLINA et ceci est fait dans la partie précalculs de la FMM (Voir tableau 2.1). Et pour le calcul des interactions lointaines c'est juste un assemblage des informations précalculées (voir l'algorithme FMM un niveau) sans avoir besoin de calculer intégralement les matrices définies par les noyaux de Green.

Dans la suite, nous étudions la précision et le temps d'exécution de chaque solveur en faisant varier le paramètre de troncature  $Q$  (voir tableau suivant).

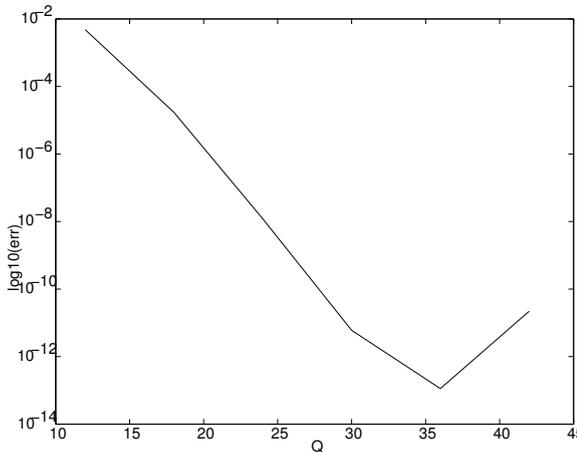
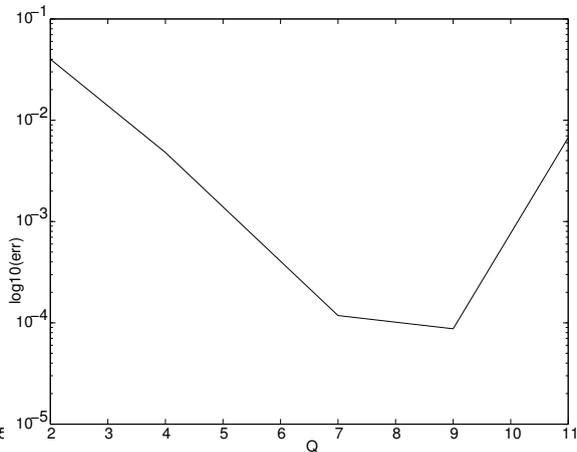
| $Q$  | 4                    | 8                    | 12                   | 16                   |
|--|----------------------|----------------------|----------------------|----------------------|
| $\frac{\  \mathbf{u}_{FMM} - \mathbf{u}_{GMRES} \ _2}{\  \mathbf{u}_{GMRES} \ _2}$ | $0.66 \cdot 10^{-2}$ | $0.97 \cdot 10^{-4}$ | $0.55 \cdot 10^{-5}$ | $0.1 \cdot 10^{-5}$  |
| $\frac{\  \mathbf{u}_{FMM} - \mathbf{u}_{EXA} \ _2}{\  \mathbf{u}_{EXA} \ _2}$     | $0.19 \cdot 10^{-1}$ | $0.21 \cdot 10^{-1}$ | $0.21 \cdot 10^{-1}$ | $0.21 \cdot 10^{-1}$ |
| Temps GMRES+FMM (s)  | 1.54                 | 2.08                 | 3.19                 | 4.63                 |

TABLE 2.2 – Précision et temps de calcul en fonction de  $Q$ , Cas 1024 points sur  $\Gamma$  et  $\Sigma$ .

On peut conclure que l'erreur sur le calcul du noyau de Green décroît lorsque  $Q$  augmente. Par contre cette erreur reste bien inférieure à l'erreur éléments finis qui est

pour ce cas  $0.21 \cdot 10^{-1}$ . Ainsi, l'erreur multipôles  $\frac{\|\mathbf{u}_{FMM} - \mathbf{u}_{GMRES}\|_2}{\|\mathbf{u}_{GMRES}\|_2}$  ne change pas la précision entre la solution calculée  $\mathbf{u}_{FMM}$  et la solution exacte qui est de même ordre que celle entre la solution exacte du problème et la solution calculée  $\mathbf{u}_{GMRES}$ . Le choix de  $Q$  est imposé par le code de calcul de telle sorte que nous n'avons pas d'instabilité numérique. Mais, on se permet de modifier  $Q$  dans ce test numérique afin de montrer l'impact de  $Q$  sur le temps de calcul sans gain de précision entre la solution calculée  $\mathbf{u}_{FMM}$  et la solution exacte. Par conséquent, il est inutile d'augmenter le paramètre de troncature  $Q$ . Il suffit de choisir  $Q = 4$ .

Les développements classiques de la FMM présentent une instabilité lorsque le maillage a des densités de maillage locales importantes devant la longueur d'onde. Ce problème est connu sous le nom d'instabilité basses fréquences. Pour spécifier l'instabilité numérique due à l'accumulation des fonctions de Hankel dans l'opérateur des transferts (2.2), regardons une configuration simple :  $\Sigma = [(0, 0, 0), (1, 0, 0)]$  et  $\Gamma = [(0, 0.1, 0.1), (1, 0.1, 0.1)]$  avec une discrétisation 80 points par domaine. Dans la figure 2.5, on a  $k = 50$  (soit  $\lambda = 10h$ ) et dans la figure 2.6 on a  $k = 0.5$  (soit  $\lambda = 1000h$ ). Nous remarquons une instabilité numérique pour les deux cas, mais elle survient plus vite dans le deuxième test, montrant l'importance de la cohérence entre la discrétisation et le nombre d'onde. En fait, pour le cas test où on a 10 points par longueur d'onde l'erreur décroît jusqu'à  $10^{-13}$ . Par contre, l'erreur est limitée à  $10^{-4}$  pour le cas d'un choix de 1000 points par longueur d'onde.

FIGURE 2.5 – Cas 80 points,  $k = 50$ .FIGURE 2.6 – Cas 80 points,  $k = 0.5$ .

La résolution numérique est faite par l'intégration des modules en fortran 90 d'un code FMM dans la librairie MÉLINA. Vu la limitation de l'allocation dynamique de MÉLINA qui est liée aux spécificités du fortran 77, nous n'avons pas pu avoir des résultats pour des cas tests où le nombre de degrés de liberté est très grand. C'est la raison pour laquelle nous avons choisi d'utiliser la librairie MÉLINA++ (version C++) par la suite. Notre but est la mise en œuvre de la méthode FMM multiniveaux dans MÉLINA++ afin de quantifier ses apports pour la résolution des problèmes de diffraction acoustiques et électromagnétiques avec la méthode CEFRI. Dans ce cadre un développement d'une librairie FastMMLib est en cours et a pour objectif d'offrir une version générique des méthodes multipôles rapides. Vu que le choix de développement du noyau de Green selon une série de Gegenbauer

s'avère limité face aux problèmes connus sous le nom de problème à basses fréquences, nous souhaitons intégrer de récentes versions de la FMM valables pour un large spectre de fréquences incluant les basses fréquences.

## 2.6 FMM stable pour toutes fréquences

Plusieurs développements ont été réalisés pour l'élaboration d'une FMM valable en régime "basses fréquences" ([37], [38], [39]). Greengard et al. ([40]) ont proposé un développement séparant et traitant différemment les termes évanescents et les termes propagatifs. Darve et al ([24]) se sont inspirés de ces travaux pour établir une FMM valide à toute fréquence. En fait, pour éviter le problème de la stabilité numérique, Greengard et al. [40] ont proposé un autre développement de la solution fondamentale de l'équation de Helmholtz en 3D par le théorème suivants :

### Théorème 2.6.1.

$$\frac{e^{ik|r|}}{|r|} = \frac{ik}{2\pi} \int_{\mathcal{S}^{z^+}} e^{ik(\sigma,r)} d\sigma + \frac{1}{2\pi} \int_{\phi=0}^{2\pi} \int_{\chi=0}^{+\infty} e^{-\chi z} e^{i\sqrt{\chi^2+k^2}(x \cos \phi + y \sin \phi)} d\chi d\phi \quad (2.7)$$

où  $\mathcal{S}^{z^+} = \{r = (x, y, z) \in \mathbb{R}^3, |r| = 1, z > 0\}$ .

Soit le vecteur  $r' = (x', y', z')$ . On désigne par  $w_k^\sigma, \sigma_k$  et  $w_k^{\chi\phi}, \chi_k, \phi_k$  les poids et les points de quadrature. La formule (2.7) peut alors s'écrire

$$\frac{e^{ik|r+r'|}}{|r+r'|} \approx \frac{ik}{2\pi} \sum_p w_p^\sigma e^{ik(\sigma_p, r')} + \frac{1}{2\pi} \sum_p w_p^{\chi\phi} e^{-\chi_k z'} e^{i\sqrt{\chi_k^2+k^2}(x' \cos \phi_p + y' \sin \phi_p)}.$$

Dans cette section, on se propose de présenter une version de la méthode FMM stable pour toute fréquence développée dans [24]. En fait dans cet article, E. Darve et P. Havé ont donné lieu à une nouvelle version de la FMM valable pour un large spectre de fréquences. Nous remarquons dans la nouvelle formule (voir 2.7), ni fonctions de Hankel ni série à l'infini mais plutôt une intégrale généralisée. Nous présentons comment la formule (2.7) est discrétisée.

- Traitement de  $\int_{\mathcal{S}^{z^+}} e^{ik(\sigma,r)} d\sigma$

On a

$$\int_{\mathcal{S}^{z^+}} e^{ik(\sigma, r+r')} d\sigma = \int_{\mathcal{S}} T_r(\sigma) e^{ik(\sigma, r')} d\sigma \quad (2.8)$$

où  $T_r(\sigma) = \mathbf{1}_{\mathcal{S}^{z^+}}(\sigma) e^{ik(\sigma, r)}$  pour  $\sigma \in \mathcal{S}^2, r \in \mathbb{R}^3$ . Nous pouvons approcher (2.8) par une décomposition sur la base des harmoniques sphériques en se basant sur les résultats suivants :

**Proposition 2.6.1.** [24] Pour tout  $\varepsilon > 0$ , nous avons :

$$| e^{ik(\sigma,r)} - \sum_{0 \leq |m| \leq l \leq L} [e^{ik(\sigma,r)}]_{lm} Y_{lm}(\sigma) | \leq \varepsilon$$

avec  $[e^{ik(\sigma,r)}]_{lm} = \int_{\mathcal{S}} e^{ik(\sigma,r)} Y_{lm}^*(\sigma) d\sigma$  et  $L \geq k|r| + 1.8 d_0^{2/3} (k|r|)^{1/3}$  tel que  $d_0 = \log_{10}(1/\varepsilon)$ .

**Proposition 2.6.2.** [24] Soit  $T_r(\sigma)$  la fonction de transfert définie par

$$T_r(\sigma) = \mathbf{1}_{\mathcal{S}z^+}(\sigma) e^{ik(\sigma,r)} \text{ pour } \sigma \in \mathcal{S}, r \in \mathbb{R}^3,$$

Soient  $\varepsilon > 0$  l'erreur considérée sur l'approximation de l'intégrand du terme  $\int_{\mathcal{S}z^+} e^{ik(\sigma,r)} d\sigma$  et  $d$  le diamètre des boîtes au niveau où  $T_r$  opère. Soit  $\tilde{T}_r^L$  la projection de la fonction de transfert  $T_r$  dans une base d'harmoniques sphériques :

$$\tilde{T}_r^L(\sigma) = \sum_{0 \leq |m| \leq l \leq L} [T_r]_{lm} Y_{lm}(\sigma)$$

avec  $L \approx kd + 1.8 d_0^{2/3} (kd)^{1/3}$ ,  $d_0 = \log_{10}(1/\varepsilon)$ .  $\tilde{T}_r^L$  est une approximation de  $T_r$  par une formule de quadrature à  $(L+1) \times (L+1)$  points.

En conclusion, le terme  $\tilde{T}_r^L e^{ik(\sigma,r')}$  est exactement intégré par  $(L+1)^2$  points.

- Traitement de  $\int_{\phi=0}^{2\pi} \int_{\chi=0}^{+\infty} e^{-\chi z} e^{i\sqrt{\chi^2+k^2}(x \cos \phi + y \sin \phi)} d\chi d\phi$

Nous nous proposons de rappeler les outils mathématiques pour la discrétisation de l'intégrale auquel nous nous sommes intéressés dans cette partie. En fait, nous présentons comment se fait la troncature de l'intégrale généralisée afin de discrétiser l'intégrale suivant  $\chi$  par la méthode de décomposition en valeurs singulière. Nous exposons les résultats concernant la discrétiser l'intégrale suivant  $\phi$ . A la fin de cette section nous écrivons les étapes de l'algorithme FMM à toutes fréquences. Pour cela, nous adoptons les notations suivantes :

$$V(x, y, z) = \frac{1}{2\pi} \int_0^\infty d\chi \int_0^{2\pi} d\phi K(\chi, z, x \cos \phi + y \sin \phi)$$

tel que

$$K(x, y, z) = e^{-\chi z} e^{i(x \cos \phi + y \sin \phi) \sqrt{\chi^2+k^2}}.$$

Nous définissons l'approximation de  $V$  à  $\varepsilon$  près par

$$\tilde{V}_\varepsilon(x, y, z) = \frac{1}{2\pi} \int_0^{\chi'_{max}(\varepsilon)} d\chi \int_0^{2\pi} d\phi K(\chi, z, x \cos \phi + y \sin \phi)$$

tel que  $\chi'_{max}(\varepsilon)$  vérifie l'inégalité suivante

$$\chi'_{max}(\varepsilon) \geq -\frac{\log(z_{min}\varepsilon)}{z_{min}}.$$

avec  $z_{min}$  désigne la borne inférieure suivant  $z$ . Soit  $R_0$  le côté d'une boîte au niveau auquel on souhaite calculer  $V$ . La discrétisation suivant la variable  $\chi$  se fait par une décomposition en valeurs singulières du noyau  $K$  [41]

**Proposition 2.6.3.** [24] Pour tout  $\varepsilon > 0$ , il existe un ensemble de fonctions réelles orthonormales  $\{u_p(\chi)\}_{0 \leq p \leq N}$ , un ensemble de fonctions complexes orthonormales  $\{v_p(\chi)\}_{0 \leq p \leq N}$  et des coefficients réels positifs  $\{c_p\}_{0 \leq p \leq N}$  tels que

$$\forall(\chi, z, \rho) \in [0, \chi_{max}] \times [z_{min}, z_{max}] \times [-\rho_{max}, \rho_{max}],$$

$$\left| K(\chi, z, \rho) - \sum_{p=0}^{N-1} c_p u_p(\chi) v_p(z, \rho) \right| \leq \varepsilon$$

avec  $\rho = x \cos \phi + y \sin \phi$ .

Une application de l'approche de quadrature gaussienne généralisé de [42], offre la possibilité de construire une quadrature en  $\chi$  de  $N/2$  points et  $N/2$  poids permettant d'intégrer exactement les  $N$  fonction  $u_p(\chi)$ . E. Darve et P. Havé ont montré à travers des expériences numériques que nous pouvons approcher d'une manière efficace  $K(\chi, z, \rho)$  avec un nombre réduit de terme dans la décomposition en valeurs singulières. La discrétisation suivant  $\phi$  se fait par la technique de décomposition de Fourier par une quadrature uniforme de taille  $L_\phi + 1$  avec  $L_\phi = L_0 + 1.8d_0^{2/3}L_0^{1/3}$ , tel que  $d_0 = \log(1/\varepsilon)$  et  $L_0 = \frac{R_0}{\sqrt{2}} \sqrt{\left(\frac{2d_0}{R_0}\right)^2 + k^2}$ .

## 2.7 Présentation de la librairie FastMMLib

Les méthodes multipôles rapides ont participé à améliorer la performance des codes de calculs des problèmes aux dérivées partielles. Citons par exemple le problème de Laplace et les problèmes des ondes acoustiques et électromagnétiques. La mise en œuvre des ces méthodes est souvent propre à une thématique précise. La librairie FastMMLib initiée par Eric Darrigrand et Yvon Lafranche, à l'IRMAR, université de Rennes 1, est destinée à intégrer plusieurs familles de méthodes rapides. Elle est développée dans un contexte générique de sorte à pouvoir traiter une gamme de problèmes la plus vaste possible. Elle devra être simple à utiliser et elle pourra être appelée par des librairies de résolution numérique comme les librairies éléments finis MÉLINA et MÉLINA++. Pour un point donné  $x_\sigma$  dans la boîte  $B_\sigma$  et en supposant que  $y_\gamma$  dans  $B_\gamma$  tel que  $B_\gamma \notin \mathcal{V}(B_\sigma)$ , la quantité  $H(x_\sigma, y_\gamma)$  admet le développement suivant :

$$H(x_\sigma, y_\gamma) = c \sum_{p=1}^P c_p g_{\gamma, B}^{(p)} \sum_{\tilde{p}=1}^{\tilde{P}} T_{B, \tilde{B}}^{(p, \tilde{p})} f_{\sigma, \tilde{B}}^{(p, \tilde{p})}.$$

Dans le cadre de l'application de la FMM à la méthode CEFRI pour l'équation de Helmholtz,  $H$  est définie comme suit

$$H(x_\sigma, y_\gamma) \simeq \frac{-ik^3}{(4\pi)^2} \sum_{p=1}^S w_p (1 - (s_p, \mathbf{n}_{x_\sigma})) e^{ik(s_p, x_\sigma - C_\sigma)} \mathcal{T}_{r_0}^Q(s_p) e^{-ik(s_p, y_\gamma - C_\gamma)} (s_p, \mathbf{n}_{y_\gamma})$$

où  $c = \frac{-ik^3}{(4\pi)^2}$ ,  $c_p = w_p$ ,  $T_{B,\hat{B}}^{(p;\hat{p})} = \mathcal{T}_{r_0}^Q(s_p)$ ,  $f_{\sigma,\hat{B}}^{(p;\hat{p})} = (1 - (s_p, \mathbf{n}_{x_\sigma}))e^{ik(s_p, x_\sigma - C_\sigma)}$   
 et  $g_{\gamma,B}^{(p)} = (s_p, \mathbf{n}_{y_\gamma})e^{-ik(s_p, y_\gamma - C_\gamma)}$ .

La librairie FastMMlib est constituée de plusieurs types de classes :

- Des classes internes dans la librairie :
  - classe Particles
  - classe Level
  - classe FMBox
  - classe TargetBox.
  
- Des classes manipulées par l'utilisateur :
  - classe FMM
  - classe Levels
  - classe SetofParticles
  - classe GreenKernel.

La classe FMM permet de définir un objet qui porte le produit matrice-vecteur FMM et qui est calculé par la méthode `matVec`. L'utilisateur doit choisir son propre noyau de Green par la classe `MyGreenKernelClass` en fonction du problème auquel il s'intéresse. Il doit spécifier ses propres particules à l'aide de la classe `MyParticleClass`. En fait, les particules sont associées aux points dans la formulation discrétisée du problème à résoudre (cas de la dynamique moléculaire ou cas de nœuds éléments finis d'un maillage). L'utilisateur fournit ses informations en définissant ses objets en suivant la syntaxe suivante :

```
vector<Particle*> MyParticles;
// chaque element du vecteur pointe vers une particule

SetOfParticles MySetsofParticles(SpaceDim, SetDim, MyParticles);
//definition des donnees de la FMM:
//dimension de l'espace,
//dimension de l'ensemble contenant les particules,
//les particules.

Levels MyLevels(MySetsofParticles);
// construction des boites

GreenKernel MyKernel;
// definition du noyau de Green

FMM MyFMM(MyLevels, MyKernel);
// definition de l'objet qui porte le produit matrice-vecteur

V = MyFMM.MatVec(U);
```

```
// calcul du produit FMM
```

Les alternatives que nous envisageons d'étudier, d'un point de vue insertion dans la librairie FastMMLib, sont les suivantes :

- En parallèle à la FMM, O. Bruno et ses co-auteurs ([43], [44], [45]) ont développé une méthode basée sur la transformée de Fourier rapide. Elle s'apparente fortement aux méthodes FMM et sur le plan de la mise en œuvre numérique, la différence majeure réside dans la considération des boîtes : perçues au travers de leur centre pour la FMM, dans la méthode de O. Bruno, elles interviennent au travers d'un ensemble de points uniformément répartis sur les faces des boîtes.
- W. Hackbusch et son équipe ont développé la méthode dite des H-matrices ([46], [47], [48]). Compte tenue de sa structure octree semblable à celle de la FMM, cette méthode inspirée des décompositions en valeurs singulières par bloc, devrait pouvoir être intégrée dans la librairie FastMMLib.
- Une dernière alternative que l'on peut citer est la méthode tensorielle présentée par B. Khoromskij ([49]) lors des Journées Singulières Augmentées 2013 à l'IRMAR.



## Deuxième partie

# Résolution du problème de la propagation d'onde électromagnétique en domaine non borné

Dans cette partie, nous nous intéressons à la résolution du problème de diffraction électromagnétique par un conducteur parfait. Dans un premier volet, nous justifions l'interprétation de la méthode **CEFRI** comme une méthode de Schwarz avec recouvrement total. Ceci nous a permis d'étudier la convergence de cet algorithme et d'orienter le choix d'un préconditionneur pour les méthodes de Krylov. Dans un second volet, la mise en œuvre de la résolution du problème nous a permis d'enrichir la librairie **MÉLINA++** en intégrant des nouveaux objets qui ouvrent les portes à la résolution numérique d'autres configurations de problèmes électromagnétiques par la méthode **CEFRI**. Compte tenu de **MÉLINA++**, nous avons choisi de travailler avec des équations de Maxwell régularisées plutôt qu'avec des équations de Maxwell classiques afin de pouvoir utiliser les éléments finis de Lagrange. De plus, la prise en compte de la condition essentielle est établie par la méthode de pénalisation. Nous avons analysé chacun de ces aspects afin de valider le couplage entre éléments finis et représentation intégrale pour le problème de Maxwell extérieur.



# Chapitre 3

## Méthode de Schwarz avec recouvrement total pour le problème de Maxwell classique extérieur en 3D

### 3.1 Eléments constitutifs du modèle

Dans ce document, nous désignons respectivement,  $\nabla$ ,  $\text{div}$  et  $\text{rot}$  le gradient, la divergence et le rotationnel. Nous notons respectivement  $\text{div}_x$  et  $\text{rot}_x$  le rotationnel et la divergence par rapport à un point  $x \in \mathbb{R}^3$ . Pour une matrice  $M$  donnée, le rotationnel de la matrice  $M$ , qu'on note  $\text{rot } M$ , est une matrice dont la  $l^{\text{ième}}$  colonne est  $\text{rot } M^l$  où  $M^l$  est la  $l^{\text{ième}}$  colonne de  $M$ . Nous définissons aussi la divergence de la matrice  $M$  comme un vecteur dont la  $l^{\text{ième}}$  composante est la divergence de la  $l^{\text{ième}}$  colonne de  $M$ . Nous notons aussi le produit scalaire par  $\cdot$  et le produit vectoriel par  $\wedge$ .

#### 3.1.1 Présentation des équations de Maxwell

Le champ électromagnétique est constitué par quatre champs vectoriels qui vérifient les équations de Maxwell données par ([50]) :

$$\text{rot } \mathcal{E} + \frac{\partial \mathcal{B}}{\partial t} = 0 \quad \text{Loi de Faraday,} \quad (3.1)$$

$$\text{rot } \mathcal{H} - \frac{\partial \mathcal{D}}{\partial t} = \mathcal{J} \quad \text{Loi de Maxwell-Ampère,} \quad (3.2)$$

$$\text{div } \mathcal{D} = \rho \quad \text{Loi de Gauss,} \quad (3.3)$$

$$\text{div } \mathcal{B} = 0 \quad \text{Loi de Gauss-magnétique,} \quad (3.4)$$

avec les notations usuelles ( pour  $\mathbf{F} = (\mathbf{F}_1, \mathbf{F}_2, \mathbf{F}_3)$ ,  $x = (x_1, x_2, x_3)$ )

$$\text{div } \mathbf{F} = \sum_{i=1}^3 \frac{\partial \mathbf{F}_i}{\partial x_i},$$
$$\text{rot } \mathbf{F} = \left( \frac{\partial \mathbf{F}_3}{\partial x_2} - \frac{\partial \mathbf{F}_2}{\partial x_3}, \frac{\partial \mathbf{F}_1}{\partial x_3} - \frac{\partial \mathbf{F}_3}{\partial x_1}, \frac{\partial \mathbf{F}_2}{\partial x_1} - \frac{\partial \mathbf{F}_1}{\partial x_2} \right).$$

et

$\mathcal{E} = \mathcal{E}(x, t)$  : l'intensité du champ électrique,

$\mathcal{D} = \mathcal{D}(x, t)$  : la densité du flux électrique,

$\mathcal{H} = \mathcal{H}(x, t)$  : l'intensité du champ magnétique,

$\mathcal{B} = \mathcal{B}(x, t)$  : la densité du flux magnétique,

$\mathcal{J} = \mathcal{J}(x, t)$  : la densité du courant électrique,

$\rho = \rho(x, t)$  : la densité de la charge électrique.

**Remarque 3.1.1.** Par dérivation par rapport au temps  $t$  de (3.3) et en appliquant l'opérateur de divergence à (3.2), on obtient la relation dite de conservation de l'électricité :

$$\frac{\partial \rho}{\partial t} + \operatorname{div} \mathcal{J} = 0. \quad (3.5)$$

On se place en régime harmonique où chaque grandeur physique varie périodiquement en temps sous la pulsation  $\omega = 2\pi f$  avec  $f$  la fréquence :

$$\mathcal{F}(x, t) = \mathcal{R}e(\mathbf{F}(x) \exp^{-i\omega t})$$

où  $\mathcal{F}$  désigne une des grandeurs physiques  $\mathcal{B}$ ,  $\mathcal{H}$ ,  $\mathcal{E}$ ,  $\mathcal{D}$ ,  $\mathcal{J}$  et  $\rho$ . On note par  $\mathcal{R}e$  la partie réelle de l'expression entre parenthèses. Les équations (3.1) et (3.2) s'écrivent alors sous la forme suivante :

$$\operatorname{rot} \mathbf{E} - i\omega \mathbf{B} = 0, \quad (3.6)$$

$$\operatorname{rot} \mathbf{H} + i\omega \mathbf{D} = \mathbf{J}. \quad (3.7)$$

Pour résoudre le problème de Maxwell (3.1-3.4), nous rajoutons des conditions dites constitutives qui sont des lois de comportement, entre  $\mathbf{E}$  et  $\mathbf{D}$ ,  $\mathbf{B}$  et  $\mathbf{H}$ . Elles dépendent des considérations physiques et décrivent les propriétés du matériau considéré. Ces relations sont généralement données sous les formes suivantes :

$$\begin{cases} \mathbf{D} = \varepsilon \mathbf{E}, \\ \mathbf{B} = \mu \mathbf{H}. \end{cases} \quad (3.8)$$

avec  $\varepsilon = \varepsilon_0 \cdot \varepsilon_r$  la permittivité du milieu où  $\varepsilon_r$  est la permittivité relative et  $\varepsilon_0$  est la permittivité du vide. De même  $\mu = \mu_0 \cdot \mu_r$  est la perméabilité magnétique du matériau où  $\mu_r$  est la perméabilité relative et  $\mu_0$  est la perméabilité du vide. Lorsque  $\varepsilon$  et  $\mu$  sont constants, on dit que le milieu est homogène. Ces paramètres sont des tenseurs pour des milieux anisotropiques et scalaires pour des milieux isotropiques. Remarquons que la dépendance linéaire des termes dans les relations (3.8) n'est pas générale. Elle n'est pas vérifiée pour une classe importante de matériaux. Dans ce qui suit, nous considérons le cas où  $\varepsilon$  et  $\mu$  sont des variables scalaires et indépendantes du temps. Aussi on ne considère que les matériaux vérifiant le système (3.8) qui caractérise les milieux linéaires sans polarisation.

Dans le cas d'un corps conducteur, d'après la loi d'Ohm le champ électrique  $\mathbf{E}$  est proportionnel à la densité du courant :

$$\mathbf{J} = \sigma \mathbf{E}$$

où  $\sigma$  représente la conductivité. C'est une quantité positive, nulle dans les isolants comme le vide ou l'air.

On obtient alors le problème de Maxwell harmonique suivant :

$$\left\{ \begin{array}{l} \mathbf{rot} \mathbf{E} - i\omega\mu\mathbf{H} = 0, \\ \mathbf{rot} \mathbf{H} + i\omega\varepsilon\mathbf{E} = 0, \\ \mathbf{div}\varepsilon\mathbf{E} = \rho, \\ \mathbf{div}\mu\mathbf{H} = 0. \end{array} \right. \quad (3.9)$$

où  $\mathbf{E}$  (resp.  $\mathbf{H}$ ) désigne le champ électrique (resp. magnétique).

**Remarque 3.1.2.** *L'expression de la permittivité dépend du milieu de propagation. En fait, dans le cas d'un domaine de permittivité  $\varepsilon$ , s'il est de plus conducteur le problème harmonique a une permittivité  $\tilde{\varepsilon}$  donnée par  $\tilde{\varepsilon} = \varepsilon + \frac{i\sigma}{\omega}$ .*

En éliminant  $\mathbf{H}$  dans (3.9), on obtient une EDP d'ordre 2 en  $\mathbf{E}$

$$\mathbf{rot}\left(\frac{1}{\mu}\mathbf{rot} \mathbf{E}\right) - \omega^2\varepsilon\mathbf{E} = 0. \quad (3.10)$$

Les composantes  $\mathbf{E}$  et  $\mathbf{H}$  étant liées par la relation

$$\mathbf{H} = -\frac{i}{\omega\mu}\mathbf{rot} \mathbf{E}.$$

**Remarque 3.1.3.** *Dans le cas où  $\mu$  et  $\varepsilon$  sont constants l'équation (3.10) prend la forme simplifiée suivante :*

$$\mathbf{rot} \mathbf{rot} \mathbf{E} - k_s^2\mathbf{E} = 0 \quad (3.11)$$

avec  $k_s^2 = \omega^2\varepsilon\mu$ .

### 3.1.2 Conditions de transmission entre deux milieux

On étudie le comportement des champs au passage d'une surface. Soit une surface régulière  $\Gamma$  séparant un milieu  $\Omega_1$  d'un milieu  $\Omega_2$ . Soit  $\mathbf{n}$  la normale à  $\Gamma$  orientée du milieu 1 vers le milieu 2. On note par  $\rho_\Gamma$  et  $\mathbf{J}_\Gamma$  respectivement la densité de charge et de courant sur la surface  $\Gamma$ . Les conditions de transmission du champ électromagnétique sur cette surface sont :

$$\left\{ \begin{array}{l} (\mathbf{D}_2 - \mathbf{D}_1) \cdot \mathbf{n} = \rho_\Gamma \text{ sur } \Gamma, \\ (\mathbf{H}_2 - \mathbf{H}_1) \wedge \mathbf{n} = -\mathbf{J}_\Gamma \text{ sur } \Gamma, \\ (\mathbf{B}_2 - \mathbf{B}_1) \cdot \mathbf{n} = 0 \text{ sur } \Gamma, \\ (\mathbf{E}_2 - \mathbf{E}_1) \wedge \mathbf{n} = 0 \text{ sur } \Gamma. \end{array} \right. \quad (3.12)$$

où l'indice  $i$  désigne la restriction des champs au milieu  $\Omega_i$  ( $i = 1, 2$ ). Ainsi, à la traversée de l'interface :

- La composante normale de  $\mathbf{B}$  et la composante tangentielle de  $\mathbf{E}$  sont continues à la traversée de  $\Gamma$ .
- Si  $\rho_\Gamma$  et  $\mathbf{J}_\Gamma$  sont non nuls, la composante normale de  $\mathbf{D}$  et la composante tangentielle de  $\mathbf{H}$  sont discontinues.

Les métaux possèdent des charges libres qui se déplacent vers leur surface sous l'action d'un champ électrique appliqué. Elles forment alors une couche très mince qui peut être modélisée par une densité superficielle  $\rho_\Gamma$ . Par contre, les diélectriques parfaits ne contiennent, par définition, aucune charge libre, et il n'existe a priori aucune densité superficielle de charge  $\rho_\Gamma$  ou de courant  $\mathbf{J}_\Gamma$  à l'interface entre deux tels diélectriques.

### 3.1.3 Conditions aux limites d'un conducteur parfait

On appelle conducteur parfait le cas limite où la conductivité  $\sigma$  est infinie. Les métaux permettent d'approcher la notion du conducteur parfait (la constante  $\sigma$  est très élevée). Dans un conducteur parfait, les champs électromagnétiques ne pénètrent pas et les courants sont surfaciques. En conséquence, le champ électrique et le champ magnétique sont nuls à l'intérieur. Ainsi, l'onde incidente se propageant dans l'air et tombant sur un conducteur parfait est donc totalement diffractée. Si on est dans la situation de deux milieux différents, comme vu précédemment, par exemple le cas où le milieu 1 est un conducteur parfait, le champ électromagnétique est nul dans ce milieu :

$$\mathbf{E}_1 = \mathbf{B}_1 = 0.$$

Pour simplifier, on élimine l'indice 2 et on a suivant la nature de la surface  $\Gamma$  dans ce cas :

- ou bien des conditions aux limites de type conducteur électrique parfait :

$$\mathbf{B} \cdot \mathbf{n} = 0 \text{ sur } \Gamma, \quad (3.13)$$

$$\mathbf{E} \wedge \mathbf{n} = 0 \text{ sur } \Gamma. \quad (3.14)$$

- ou bien des conditions aux limites de type conducteur magnétique parfait :

$$\mathbf{D} \cdot \mathbf{n} = 0 \text{ sur } \Gamma, \quad (3.15)$$

$$\mathbf{H} \wedge \mathbf{n} = 0 \text{ sur } \Gamma. \quad (3.16)$$

**Remarque 3.1.4.** *Les conditions d'interfaces des composantes normales se déduisent de celle sur les composantes tangentielles et l'équation de Maxwell (3.10).*

## 3.2 Problème de diffraction électromagnétique par un conducteur parfait

### 3.2.1 Problème de Maxwell classique en domaine non borné

Nous nous intéressons à la propagation d'une onde électromagnétique dans un milieu conducteur parfait. Pour cela nous considérons un obstacle parfaitement conducteur  $\Omega_i$  de  $\mathbb{R}^3$  ouvert, borné et connexe, avec une frontière  $\Gamma$  assez régulière. On note  $\Omega^e = \mathbb{R}^3 \setminus \overline{\Omega_i}$  et on désigne par  $\mathbf{n}$  la normale sortante sur  $\Gamma$ . On éclaire cet obstacle par une onde incidente plane qu'on notera dans la suite par  $\mathbf{E}^{inc}$  dont l'expression est :

$$\mathbf{E}^{inc}(x) = \mathbf{p}e^{ik_s x \cdot \mathbf{d}}$$

où  $k_s$  est le nombre d'onde tel que

$$\begin{cases} k_s^2 = \omega^2 \mu \varepsilon, \\ \mathcal{I}m k_s = 0. \end{cases}$$

$\mathbf{d}$  est un vecteur unitaire qui donne la direction de l'onde incidente plane et  $\mathbf{p} \in \mathbb{R}^3$  est la polarisation vérifiant  $\mathbf{p} \cdot \mathbf{d} = 0$ .

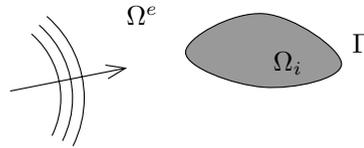
$\mathbf{E}^{inc}$  vérifie l'équation suivante

$$\mathbf{rot} \mathbf{rot} \mathbf{E}^{inc} - k_s^2 \mathbf{E}^{inc} = 0 \text{ dans } \mathbb{R}^3.$$

L'onde se propage dans le milieu extérieur. Il s'agit d'un diélectrique homogène. Notre but est de déterminer l'expression du champ diffracté  $\mathbf{E}^s$  ou du champ total  $\mathbf{E} = \mathbf{E}^{inc} + \mathbf{E}^s$ . Puisque le diffracteur est parfaitement conducteur, on a dans ce cas  $\mathbf{E}|_{\Omega_i} = \mathbf{H}|_{\Omega_i} = 0$ . Nous serons amenés alors à la résolution du problème de Maxwell classique ([13]) défini comme suit :

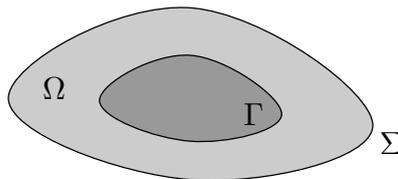
$$\begin{cases} \mathbf{rot} \mathbf{rot} \mathbf{E} - k_s^2 \mathbf{E} = 0 \text{ dans } \Omega^e, \\ \mathbf{E} \wedge \mathbf{n}_e = 0 \text{ sur } \Gamma, \\ \lim_{R \rightarrow \infty} \int_{\|x\|=R} \|\mathbf{rot} \mathbf{E}^s \wedge \mathbf{n}_e - ik_s \mathbf{E}^s\|^2 d\gamma = 0. \end{cases} \quad (3.17)$$

avec  $\mathbf{n}_e$  la normale sortante du domaine extérieur  $\Omega^e$ . La deuxième condition dite condition de Silver-Müller traduit le comportement du champ diffracté loin de l'obstacle.



### 3.2.2 Problème en domaine borné

Notre objectif est de pouvoir déterminer le champ diffracté en faisant appel à des techniques numériques. Le problème est naturellement posé sur le complémentaire du corps dans  $\mathbb{R}^3$ . Il existe plusieurs méthodes qui permettent de se ramener à un domaine de calcul borné. On se propose d'utiliser la méthode de couplage entre éléments finis et représentation intégrale (CEFRI). Elle a été introduite pour le système de Maxwell dans [12, 13, 51]. L'idée consiste à introduire un bord  $\Sigma$  qui entoure l'obstacle et sur lequel on définit une condition aux limites au moyen d'une représentation intégrale.



Le problème de Maxwell classique en domaine non borné se ramène à un problème posé sur un domaine borné  $\Omega$  compris entre  $\Gamma$  et  $\Sigma$  :

$$\begin{cases} \mathbf{rot} \mathbf{rot} \mathbf{E} - k_s^2 \mathbf{E} = 0 \text{ dans } \Omega, \\ \mathbf{E} \wedge \mathbf{n}_e = 0 \text{ sur } \Gamma, \\ \mathbf{rot} \mathbf{E} \wedge \mathbf{n}_\sigma - ik_s(\mathbf{E})_t = T(\mathbf{E}^{inc} - \mathcal{I}(\Gamma, \mathbf{E})) \text{ sur } \Sigma. \end{cases} \quad (3.18)$$

où  $\mathbf{n}_\sigma$  représente le vecteur normal unitaire à la surface  $\Sigma$  dirigé vers l'extérieur de  $\Omega$ ,  $(\mathbf{E})_t = \mathbf{n}_\sigma \wedge (\mathbf{E} \wedge \mathbf{n}_\sigma)$  désigne la composante tangentielle de  $\mathbf{E}$  sur  $\Sigma$  et l'opérateur  $T$  est défini comme suit :

$$T(\mathbf{E}) = \mathbf{rot} \mathbf{E} \wedge \mathbf{n}_\sigma - ik_s(\mathbf{E})_t$$

et pour tout  $x \in \Omega^e$ , on a

$$\mathcal{I}(\Gamma, \mathbf{E})(x) = \int_{\Gamma} (\mathcal{G}(x, y)(\mathbf{rot} \mathbf{E}(y) \wedge \mathbf{n}(y)) - \nabla_y G(x, y) \wedge (\mathbf{E}(y) \wedge \mathbf{n}(y))) d\gamma(y),$$

où  $\mathbf{n}$  représente le vecteur normal unitaire à la surface  $\Gamma$  dirigé vers l'extérieur du domaine  $\Omega_i$ .  $\mathcal{G} = GI + \frac{1}{k_s^2} \text{Hess}(G)$ ,  $I$  étant la matrice identité de  $\mathbb{R}^3$ , Hess représente l'opérateur Hessien, et  $G$  est la solution fondamentale de l'équation de Helmholtz (pour plus de détails, voir annexe A).

**Remarque 3.2.1.**  $\mathcal{I}(\Gamma, \mathbf{E}^{inc}) = 0$  où  $\mathbf{E}^{inc}$  désigne le champ incident. Par suite, le champ total  $\mathbf{E}$  peut alors s'écrire en dehors de  $\Gamma$  comme suit

$$\mathbf{E} = \mathbf{E}^{inc} - \mathcal{I}(\Gamma, \mathbf{E}).$$

### 3.2.3 Formulation variationnelle, existence et unicité

On introduit maintenant l'espace de Hilbert

$$\mathbb{H} = \{\mathbf{E} \in H(\mathbf{rot}, \Omega) \text{ tel que, } (\mathbf{E} \wedge \mathbf{n})_{\Gamma} = 0, (\mathbf{E} \wedge \mathbf{n})_{\Sigma} \in L^2(\Sigma)^3\}$$

muni du produit scalaire  $(\cdot, \cdot)$  défini par :

$$(\mathbf{E}, \mathbf{E}') = \int_{\Omega} \mathbf{E} \cdot \mathbf{E}' d\Omega + \int_{\Omega} \mathbf{rot} \mathbf{E} \cdot \mathbf{rot} \mathbf{E}' d\Omega + \int_{\Sigma} (\mathbf{E} \wedge \mathbf{n}_\sigma) \cdot (\mathbf{E}' \wedge \mathbf{n}_\sigma) d\sigma.$$

Si  $\mathbf{E} \in \mathbb{H}$ , alors le terme  $\mathbf{rot} \mathbf{E} \wedge \mathbf{n}$  est non défini. On a besoin donc d'étendre la représentation intégrale  $\mathcal{I}(\Gamma, \mathbf{E})(x)$  par :

$$\begin{aligned} \mathcal{I}^{\mathcal{R}}(\Gamma, \mathbf{E})(x) &= \int_{\Omega} \mathbf{rot}_y \mathcal{R} \mathcal{G}(x, y) \mathbf{rot} \mathbf{E}(y) d\Omega(y) - \int_{\Omega} k_s^2 \mathcal{R} \mathcal{G}(x, y) \mathbf{E}(y) d\Omega(y) \\ &\quad - \int_{\Gamma} \nabla_y G(x, y) \wedge (\mathbf{E}(y) \wedge \mathbf{n}(y)) d\gamma(y). \end{aligned}$$

$\mathcal{R}$  est un relèvement régulier de  $\Gamma$  vers  $\Omega$  qui transforme toute fonction régulière  $\phi$  définie sur  $\Gamma$  en une fonction régulière  $\mathcal{R}\phi$  définie sur  $\Omega$  vérifiant :

$$\mathcal{R}\phi = \begin{cases} \phi \text{ sur } \Gamma, \\ 0 \text{ sur } \Sigma. \end{cases}$$

On peut définir  $\mathcal{R}\phi$  comme suit

$$\mathcal{R}\phi(x, y) = \chi(y)\phi(x, y)$$

où  $\chi \in C^\infty(\Omega)$  satisfaisant  $\chi = \begin{cases} 1 & \text{sur } \Gamma, \\ 0 & \text{sur } \Sigma. \end{cases}$

La formulation variationnelle consiste à trouver  $\mathbf{E} \in \mathbb{H}$

$$\begin{aligned} \int_{\Omega} \mathbf{rot} \mathbf{E} \cdot \mathbf{rot} \mathbf{E}' - k_s^2 \mathbf{E} \cdot \mathbf{E}' d\Omega - ik_s \int_{\Sigma} (\mathbf{n}_\sigma \wedge \mathbf{E}) \cdot (\mathbf{n}_\sigma \wedge \mathbf{E}') d\sigma \\ + \int_{\Sigma} T(\mathcal{I}^{\mathcal{R}}(\Gamma, \mathbf{E})) \cdot \mathbf{E}' d\sigma = \int_{\Sigma} T(\mathbf{E}^{inc}) \cdot \mathbf{E}' d\sigma \quad \forall \mathbf{E}' \in \mathbb{H}. \end{aligned} \quad (3.19)$$

De manière équivalente, la formulation variationnelle du problème de Maxwell tronqué (3.18) consiste à

$$\begin{cases} \text{Trouver } \mathbf{E} \in \mathbb{H} \\ a(\mathbf{E}, \mathbf{E}') + c(\mathbf{E}, \mathbf{E}') = F(\mathbf{E}') \quad \forall \mathbf{E}' \in \mathbb{H}. \end{cases} \quad (3.20)$$

où  $a(\cdot, \cdot)$  et  $c(\cdot, \cdot)$  sont les formes sesquilinéaires définies sur  $\mathbb{H} \times \mathbb{H}$  par

$$a(\mathbf{E}, \mathbf{E}') = \int_{\Omega} \mathbf{rot} \mathbf{E} \cdot \mathbf{rot} \mathbf{E}' + k_s^2 \mathbf{E} \cdot \mathbf{E}' d\Omega - ik_s \int_{\Sigma} (\mathbf{n}_\sigma \wedge \mathbf{E}) \cdot (\mathbf{n}_\sigma \wedge \mathbf{E}') d\sigma$$

$$c(\mathbf{E}, \mathbf{E}') = -2k_s^2 \int_{\Omega} \mathbf{E} \cdot \mathbf{E}' d\Omega + \int_{\Sigma} T(\mathcal{I}^{\mathcal{R}}(\Gamma, \mathbf{E})) \cdot \mathbf{E}' d\sigma$$

et  $F(\cdot)$  est la forme semi-linéaire définie par

$$F(\mathbf{E}') = \int_{\Sigma} T(\mathbf{E}^{inc}) \cdot \mathbf{E}' d\sigma.$$

D'après le théorème de représentation de Riesz, le problème variationnel de (3.18) consiste à trouver  $\mathbf{E} \in \mathbb{H}$  tel que

$$(\mathcal{A}^c + \mathcal{C}^c)\mathbf{E} = \mathcal{F}. \quad (3.21)$$

Les opérateurs  $\mathcal{A}^c$  et  $\mathcal{C}^c : \mathbb{H} \rightarrow \mathbb{H}$  sont définis selon les formules suivantes :

$$(\mathcal{A}^c \mathbf{E}, \mathbf{E}') = \int_{\Omega} \mathbf{rot} \mathbf{E} \cdot \mathbf{rot} \mathbf{E}' + k_s^2 \mathbf{E} \cdot \mathbf{E}' d\Omega - ik_s \int_{\Sigma} (\mathbf{n}_\sigma \wedge \mathbf{E}) \cdot (\mathbf{n}_\sigma \wedge \mathbf{E}') d\sigma$$

$$(\mathcal{C}^c \mathbf{E}, \mathbf{E}') = -2k_s^2 \int_{\Omega} \mathbf{E} \cdot \mathbf{E}' d\Omega + \int_{\Sigma} T(\mathcal{I}^{\mathcal{R}}(\Gamma, \mathbf{E})) \cdot \mathbf{E}' d\sigma$$

et  $\mathcal{F}$  est donné par :  $\forall \mathbf{E} \in \mathbb{H}$

$$(\mathcal{F}, \mathbf{E}') = \int_{\Sigma} T(\mathbf{E}^{inc}) \cdot \mathbf{E}' d\sigma.$$

Le problème (3.18) est bien posé.  $\mathcal{A}^c$  et  $\mathcal{C}^c$  sont respectivement un automorphisme et un opérateur compact ([13]). On constate donc que le problème (3.18) s'exprime sous forme d'une équation de Fredholm. Si on étudie le problème homogène alors la solution est nécessairement triviale (voir théorème 3.1 de [13]). Par suite, on en déduit alors que le problème (3.18) admet une unique solution. Pour plus de détails, nous renvoyons le lecteur à [13].

Dans ce qui suit nous allons donner une relecture du problème résultat de l'utilisation de la méthode **CEFRI** au problème de Maxwell extérieur. En fait, la méthode **CEFRI** peut être interprétée comme une méthode de décomposition de domaine. Plus précisément, nous nous intéressons à la méthode de Schwarz multiplicative (voir annexe B).

### 3.3 Méthode de Schwarz avec recouvrement total pour le problème de Maxwell extérieur classique

On s'intéresse dans cette section à l'étude du comportement de l'algorithme de Schwarz pour le problème de Maxwell extérieur. Jin et Liu proposent dans [12] une technique qui consiste à découpler la condition exacte posée sur la frontière artificielle. Ceci est l'application de l'algorithme du point fixe pour les problèmes d'EDP. La résolution du problème en domaine borné (4.5) par la méthode de point fixe de Cauchy se ramène à déterminer à chaque itération  $\mathbf{E}_{n+1}$  solution de

$$\begin{cases} \mathbf{rot rot} \mathbf{E}_{n+1} - k_s^2 \mathbf{E}_{n+1} = 0 \text{ dans } \Omega, \\ \mathbf{E}_{n+1} \wedge \mathbf{n} = 0 \text{ sur } \Gamma, \\ \mathbf{rot} \mathbf{E}_{n+1} \wedge \mathbf{n}_\sigma - ik_s (\mathbf{E}_{n+1})_t = T(\mathbf{E}^{inc} - \mathcal{I}_t^R(\Gamma, \mathbf{E}_n)) \text{ sur } \Sigma. \end{cases} \quad (3.22)$$

Les résultats numériques encourageants réalisés dans les articles de Jin et Liu, montrent l'intérêt d'utiliser des algorithmes itératifs dans les domaines de la diffraction d'ondes électromagnétiques. On a illustré dans ces travaux le lien entre ces méthodes et l'algorithme de Schwarz. Cette idée est initialement présentée dans [14] pour le problème de Poisson posé dans un domaine non borné. Les auteurs de cet article montrent que le problème (3.22) revient simplement à une méthode de décomposition de domaine dite méthode de Schwarz où la région de recouvrement est la totalité du domaine  $\Omega$ . F. Jellassi a également étudié de nombreuses extensions de cette technique ([52]) tels que le problème des courants de Foucault, problème de Helmholtz, modèle des plaques minces et le système de Stokes posés dans le domaine extérieur. L'application du processus itératif de Schwarz avec recouvrement total sur le problème (3.22) consiste à construire une suite  $(\mathbf{E}_n)_n$  déterminée par la résolution des deux sous problèmes. Le premier est un problème de transmission

$$\begin{cases} \mathbf{rot rot} \mathbf{E}_{2n+1} - k_s^2 \mathbf{E}_{2n+1} = 0 \text{ dans } \Omega_i \cup \Omega^e, \\ \mathbf{n} \wedge [\mathbf{E}_{2n+1}] = 0 \text{ sur } \Gamma, \\ \mathbf{n} \wedge [\mathbf{rot} \mathbf{E}_{2n+1}] = -\mathbf{n} \wedge \mathbf{rot} \mathbf{E}_{2n} \text{ sur } \Gamma, \\ \lim_{R \rightarrow \infty} \int_{\|x\|=R} \|\mathbf{rot} \mathbf{E}_{2n+1}^s \wedge \mathbf{n}_e - ik_s \mathbf{E}_{2n+1}^s\|^2 d\gamma = 0. \end{cases} \quad (3.23)$$

et le deuxième problème consiste à trouver  $\mathbf{E}_{2n+2}$  tel que

$$\begin{cases} \mathbf{rot rot} \mathbf{E}_{2n+2} - k_s^2 \mathbf{E}_{2n+2} = 0 \text{ dans } \Omega, \\ \mathbf{E}_{2n+2} \wedge \mathbf{n} = 0 \text{ sur } \Gamma, \\ (\mathbf{rot} \mathbf{E}_{2n+2}) \wedge \mathbf{n}_\sigma - ik_s (\mathbf{E}_{2n+2})_t = \\ \quad (\mathbf{rot} \mathbf{E}_{2n+1}) \wedge \mathbf{n}_\sigma - ik_s (\mathbf{E}_{2n+1})_t \text{ sur } \Sigma. \end{cases} \quad (3.24)$$

La solution  $\mathbf{E}_{2n+1}$  du problème (3.23) est déterminée explicitement au moyen d'une représentation intégrale (voir Annexe A) puis on l'insère dans la deuxième condition du second problème (3.24). On retrouve ainsi le système (3.22). Le problème (3.22) est celui qu'on peut résoudre numériquement. Mais le voir comme deux problèmes séparés (3.23) et (3.24) permet d'utiliser tous les progrès réalisés dans le domaine des méthodes de décomposition de domaine et permet donc de faire l'étude de convergence dans le cas d'une géométrie simplifiée. L'algorithme de Schwarz avec recouvrement total peut s'écrire à la  $n^{\text{ième}}$  itération comme suit

$$\mathcal{A} \mathbf{E}^{n+1} = -\mathcal{C} \mathbf{E}^n + \mathcal{F} \quad (3.25)$$

où les opérateurs  $\mathcal{A}$  et  $\mathcal{C} : \mathbb{H} \rightarrow \mathbb{H}$  sont définis selon les formules suivantes :

$$(\mathcal{A} \mathbf{E}, \mathbf{E}') = \int_{\Omega} \text{rot } \mathbf{E} \cdot \text{rot } \mathbf{E}' - k_s^2 \mathbf{E} \cdot \mathbf{E}' d\Omega - ik_s \int_{\Sigma} (\mathbf{n}_{\sigma} \wedge \mathbf{E}) \cdot (\mathbf{n}_{\sigma} \wedge \mathbf{E}') d\sigma,$$

$$(\mathcal{C} \mathbf{E}, \mathbf{E}') = \int_{\Sigma} T(\mathcal{I}^{\mathcal{R}}(\Gamma, \mathbf{E})) \cdot \mathbf{E}' d\sigma.$$

**Proposition 3.3.1.** *Considérons un ouvert borné  $O$ . Le problème suivant n'admet que la solution nulle :*

$$\begin{cases} \text{rot rot } \mathbf{E} - k_s^2 \mathbf{E} = 0 \text{ dans } O, \\ \text{rot } \mathbf{E} \wedge \mathbf{n} - ik_s (\mathbf{E})_t = 0 \text{ sur } \partial O. \end{cases} \quad (3.26)$$

avec  $\mathbf{n}$  la normale sortante du domaine  $O$ .

Preuve : En effet, en écrivant sa formulation variationnelle et en utilisant la condition sur le bord  $\partial O$ , on obtient

$$\int_O \text{rot } \mathbf{E} \cdot \text{rot } \mathbf{E}' - k_s^2 \mathbf{E} \cdot \mathbf{E}' dO - ik_s \int_{\partial O} (\mathbf{n} \wedge \mathbf{E}) \cdot (\mathbf{n} \wedge \mathbf{E}') d\partial O = 0.$$

Si on choisit  $\mathbf{E}' = \mathbf{E}$  alors  $\mathbf{n} \wedge \mathbf{E} = 0$  sur  $\partial O$ . Compte tenu de la condition aux limites sur  $\partial O$ , on a  $\text{rot } \mathbf{E} \wedge \mathbf{n} = 0$ . Cela donne, d'après la formule de représentation intégrale sur  $\partial O$  (voir Annexe A), que

$$\mathbf{E}(x) = 0 \quad \forall x \in O.$$

D'après la proposition 3.3.1, l'opérateur  $\mathcal{A}$  est inversible. L'équation (3.25) est équivalente à

$$\mathbf{E}^{n+1} = -\mathcal{A}^{-1} \mathcal{C} \mathbf{E}^n + \mathcal{A}^{-1} \mathcal{F} \quad (3.27)$$

où  $\mathcal{A}$  et  $\mathcal{C}$  sont les opérateurs mentionnés précédemment. Les deux schémas itératifs (3.25) et (3.27) convergent si et seulement si le rayon spectral de  $(-\mathcal{A}^{-1} \mathcal{C})$  est strictement inférieur à l'unité. On constate alors que la méthode de Schwarz avec recouvrement total est naturellement préconditionnée par  $\mathcal{A}$ .

## 3.4 Diffraction par une sphère conductrice

Nous nous intéressons dans cette partie à mener une étude analytique en  $3D$  qui permet de mettre en évidence la convergence de l'algorithme de Schwarz avec recouvrement total pour le problème de Maxwell extérieur (3.17).

### 3.4.1 Outils mathématiques

Nous rappelons que  $j_l$  est la fonction de Bessel sphérique d'ordre  $l$  et  $h_l$  est la fonction de Hankel sphérique d'ordre  $l$  de première espèce.  $\rho, \theta$  et  $\phi$  désignent les coordonnées sphériques de  $x \in \mathbb{R}^3$  et  $\hat{x} = \frac{x}{|x|}$ .

**Définition 3.4.1.** *Les harmoniques sphériques ont pour expression*

$$Y_l^m(\hat{x}) = \sqrt{\frac{2l+1}{4\pi} \frac{(l-|m|)!}{(l+|m|)!}} P_l^{|m|}(\cos \theta) e^{im\phi}.$$

Les  $P_l^{|m|}$  sont les polynômes de Legendre modifiés

$$P_l^{|m|} = (1-t^2) \frac{d^m P_l(t)}{dt^m}$$

et les  $P_l$  sont les polynômes de Legendre

$$P_l(t) = \frac{(-1)^l}{2^l l!} \frac{d^l}{dt^l} [(1-t^2)^l].$$

**Remarque 3.4.1.** *Soit  $u$  une solution de l'équation de Helmholtz  $\Delta u + k^2 u = 0$ . La fonction définie par  $\mathbf{U}(x) = \mathbf{rot}(x u(x))$  est solution de l'équation de Maxwell de second ordre*

$$\mathbf{rot} \mathbf{rot} \mathbf{U} - k^2 \mathbf{U} = 0. \quad (3.28)$$

**Théorème 3.4.1.** *Soit  $Y_l^m$  une fonction harmonique d'ordre  $l > 0$  alors les fonctions*

$$\mathbf{M}_{lm} = \mathbf{rot} \{x j_l(k|x|) Y_l^m(\hat{x})\}, \quad \frac{1}{ik} \mathbf{rot} \mathbf{M}_{lm}$$

sont solutions de l'équation de Maxwell (3.28) dans  $\mathbb{R}^3$ . Les fonctions

$$\mathbf{N}_{lm} = \mathbf{rot} \{x h_l(k|x|) Y_l^m(\hat{x})\}, \quad \frac{1}{ik} \mathbf{rot} \mathbf{N}_{lm}$$

sont solutions radiales de l'équation de Maxwell (3.28) dans  $\mathbb{R}^3 \setminus \{0\}$ .

**Théorème 3.4.2.** *Pour toute base orthonormale  $Y_l^m, m = -l, \dots, l$  d'harmoniques sphériques d'ordre  $l > 0$ , les champs tangentiels à la sphère unité  $\mathcal{S}$  que nous définissons par*

$$\mathbf{u}_{lm} = \frac{1}{\sqrt{l(l+1)}} \nabla_{\mathcal{S}} Y_l^m \text{ et } \mathbf{v}_{lm} = \mathbf{n} \wedge \mathbf{u}_{lm}$$

sont appelés harmoniques sphériques vectorielles d'ordre  $l$  avec  $\nabla_{\mathcal{S}}$  le gradient surfacique sur  $\mathcal{S}$ . De plus,  $\mathbf{u}_{lm}$  et  $\mathbf{v}_{lm}$  forment une base orthonormale complète de

$$T^2(\mathcal{S}) = \{\mathbf{a} : \mathcal{S} \rightarrow \mathbb{C}^3 \text{ tel que } \mathbf{a} \in L^2(\mathcal{S})^3, \mathbf{a} \cdot \mathbf{n} = 0\}.$$

Pour la démonstration de ces théorèmes, on se référera à [8] et à [50].

### 3.4.2 Investigation analytique en 3D

On se place dans le cas où l'obstacle  $\Omega_i$  est une boule de rayon  $R_* = 1$  qu'on notera  $B_1$  et on suppose que  $\Sigma$  est la frontière de la boule concentrique à  $B_1$  et de rayon  $R > 1$ . Le but de cet exemple est le calcul explicite du taux de convergence de la méthode de Schwarz avec recouvrement total. La solution du problème de Maxwell en domaine non borné est donnée par ([8])

$$\mathbf{E} = \sum_{l=0}^{\infty} \sum_{m=-l}^l (\alpha_{lm} \mathbf{N}_{lm} + \beta_{lm} \mathbf{rot} \mathbf{N}_{lm}) + \mathbf{E}^{inc}.$$

On note  $(\mathbf{w}_n)_n$  la suite des erreurs :

$$\mathbf{w}_{2n} = \mathbf{E}_{2n} - \mathbf{E} \text{ dans } \Omega$$

et

$$\mathbf{w}_{2n+1} = \begin{cases} \mathbf{E}_{2n+1} - \mathbf{E} \text{ dans } \Omega^e, \\ \mathbf{E}_{2n+1} \text{ dans } \Omega_i, \end{cases}$$

$$\left\{ \begin{array}{l} \mathbf{rot} \mathbf{rot} \mathbf{w}_{2n+1} - k_s^2 \mathbf{w}_{2n+1} = 0 \text{ dans } \Omega_i \cup \Omega^e, \\ \mathbf{n} \wedge [\mathbf{w}_{2n+1}] = -\mathbf{n} \wedge \mathbf{w}_{2n} = 0 \text{ sur } \Gamma, \\ \mathbf{n} \wedge [\mathbf{rot} \mathbf{w}_{2n+1}] = -\mathbf{n} \wedge \mathbf{rot} \mathbf{w}_{2n} \text{ sur } \Gamma, \\ \lim_{R \rightarrow \infty} \int_{\|x\|=R} \|\mathbf{rot} \mathbf{w}_{2n+1} \wedge \mathbf{n}_e - ik_s(\mathbf{w}_{2n+1})_t\|^2 d\gamma = 0, \end{array} \right. \quad (3.29)$$

et  $\mathbf{w}_{2n+2}$  est solution du problème dans le domaine borné  $\Omega$

$$\left\{ \begin{array}{l} \mathbf{rot} \mathbf{rot} \mathbf{w}_{2n+2} - k_s^2 \mathbf{w}_{2n+2} = 0 \text{ dans } \Omega, \\ \mathbf{w}_{2n+2} \wedge \mathbf{n} = 0 \text{ sur } \Gamma, \\ \mathbf{rot} \mathbf{w}_{2n+2} \wedge \mathbf{n}_\sigma - ik_s(\mathbf{w}_{2n+2})_t = \\ \mathbf{rot} \mathbf{w}_{2n+1} \wedge \mathbf{n}_\sigma - ik_s(\mathbf{w}_{2n+1})_t \text{ sur } \Sigma. \end{array} \right. \quad (3.30)$$

$$\mathbf{w}_{2n+1} = \begin{cases} \sum_{l=0}^{\infty} \sum_{m=-l}^l (a_{lm}^{2n+1} \mathbf{M}_{lm} + b_{lm}^{2n+1} \mathbf{rot} \mathbf{M}_{lm}) & |x| \leq 1, \\ \sum_{l=0}^{\infty} \sum_{m=-l}^l (c_{lm}^{2n+1} \mathbf{N}_{lm} + d_{lm}^{2n+1} \mathbf{rot} \mathbf{N}_{lm}) & |x| > 1, \end{cases}$$

$$\mathbf{w}_{2n+2} = \sum_{l=0}^{\infty} \sum_{m=-l}^l (e_{lm}^{2n+2} \mathbf{N}_{lm} + f_{lm}^{2n+2} \mathbf{M}_{lm} + \tilde{e}_{lm}^{2n+2} \mathbf{rot} \mathbf{N}_{lm} + \tilde{f}_{lm}^{2n+2} \mathbf{rot} \mathbf{M}_{lm}).$$

On pose

$$H_l(k|x|) = h_l(k|x|) + k|x|h'_l(k|x|),$$

$$J_l(k|x|) = j_l(k|x|) + k|x|j'_l(k|x|).$$

On a

$$\hat{x} \wedge \mathbf{N}_{lm} = \sqrt{l(l+1)} h_l(k|x|) \mathbf{u}_{lm},$$

$$\begin{aligned}\hat{x} \wedge \mathbf{M}_{lm} &= \sqrt{l(l+1)} j_l(k|x|) \mathbf{u}_{lm}, \\ \hat{x} \wedge \mathbf{rot} \mathbf{N}_{lm} &= \frac{\sqrt{l(l+1)}}{|x|} H_l(k|x|) \mathbf{v}_{lm}, \\ \hat{x} \wedge \mathbf{rot} \mathbf{M}_{lm} &= \frac{\sqrt{l(l+1)}}{|x|} J_l(k|x|) \mathbf{v}_{lm}.\end{aligned}$$

### Prise en compte de la condition aux limites sur $\Gamma$

On a

$$\hat{x} \wedge \mathbf{w}_{2n+2} = 0 \text{ sur } \Gamma,$$

ce qui implique

$$\begin{aligned}\sum_{l=0}^{\infty} \sum_{m=-l}^l (e_{lm}^{2n+2} \hat{x} \wedge \mathbf{N}_{lm} + f_{lm}^{2n+2} \hat{x} \wedge \mathbf{M}_{lm} + \tilde{e}_{lm}^{2n+2} \hat{x} \wedge \mathbf{rot} \mathbf{N}_{lm} + \tilde{f}_{lm}^{2n+2} \hat{x} \wedge \mathbf{rot} \mathbf{M}_{lm}) &= 0. \\ \sum_{l=0}^{\infty} \sum_{m=-l}^l \left( \sqrt{l(l+1)} [e_{lm}^{2n+2} h_l(k_s R_*) + f_{lm}^{2n+2} j_l(k_s R_*)] \mathbf{u}_{lm} \right. \\ &\left. + \frac{\sqrt{l(l+1)}}{R_*} [\tilde{e}_{lm}^{2n+2} H_l(k_s R_*) + \tilde{f}_{lm}^{2n+2} J_l(k_s R_*)] \mathbf{v}_{lm} \right) = 0,\end{aligned}$$

ce qui donne

$$\begin{cases} e_{lm}^{2n+2} = -f_{lm}^{2n+2} \frac{j_l(k_s R_*)}{h_l(k_s R_*)}, \\ \tilde{e}_{lm}^{2n+2} = -\tilde{f}_{lm}^{2n+2} \frac{J_l(k_s R_*)}{H_l(k_s R_*)}. \end{cases} \quad (3.31)$$

### Prise en compte des conditions de transmissions sur $\Gamma$

Première condition :  $\hat{x} \wedge [\mathbf{w}_{2n+1}] = 0$

$$\begin{aligned}\sum_{l=0}^{\infty} \sum_{m=-l}^l (a_{lm}^{2n+1} \hat{x} \wedge \mathbf{M}_{lm} + b_{lm}^{2n+1} \hat{x} \wedge \mathbf{rot} \mathbf{M}_{lm}) \\ - \sum_{l=0}^{\infty} \sum_{m=-l}^l (c_{lm}^{2n+1} \hat{x} \wedge \mathbf{N}_{lm} + d_{lm}^{2n+1} \hat{x} \wedge \mathbf{rot} \mathbf{N}_{lm}) &= 0, \\ \sum_{l=0}^{\infty} \sum_{m=-l}^l \left( a_{lm}^{2n+1} \sqrt{l(l+1)} j_l(k_s R_*) \mathbf{u}_{lm} + b_{lm}^{2n+1} \frac{\sqrt{l(l+1)}}{R_*} J_l(k_s R_*) \mathbf{v}_{lm} \right) \\ - \sum_{l=0}^{\infty} \sum_{m=-l}^l \left( c_{lm}^{2n+1} \sqrt{l(l+1)} h_l(k_s R_*) \mathbf{u}_{lm} + d_{lm}^{2n+1} \frac{\sqrt{l(l+1)}}{R_*} H_l(k_s R_*) \mathbf{v}_{lm} \right) &= 0,\end{aligned}$$

ce qui donne

$$\begin{cases} a_{lm}^{2n+1} = c_{lm}^{2n+1} \frac{h_l(k_s R_*)}{j_l(k_s R_*)}, \\ b_{lm}^{2n+1} = d_{lm}^{2n+1} \frac{H_l(k_s R_*)}{J_l(k_s R_*)}. \end{cases} \quad (3.32)$$

Deuxième condition :  $\hat{x} \wedge [\mathbf{rot} \mathbf{w}_{2n+1}] = -\hat{x} \wedge \mathbf{rot} \mathbf{w}_{2n}$

$$\begin{aligned} & \sum_{l=0}^{\infty} \sum_{m=-l}^l [(a_{lm}^{2n+1} + f_{lm}^{2n}) \hat{x} \wedge \mathbf{rot} \mathbf{M}_{lm} + (e_{lm}^{2n} - c_{lm}^{2n+1}) \hat{x} \wedge \mathbf{rot} \mathbf{N}_{lm} \\ & + (\tilde{f}_{lm}^{2n} + b_{lm}^{2n+1}) \hat{x} \wedge \mathbf{rot} \mathbf{rot} \mathbf{M}_{lm} + (\tilde{e}_{lm}^{2n} - d_{lm}^{2n+1}) \hat{x} \wedge \mathbf{rot} \mathbf{rot} \mathbf{N}_{lm}] = 0. \end{aligned}$$

Comme  $\mathbf{M}_{lm}$  et  $\mathbf{N}_{lm}$  sont solutions de l'équation de Maxwell (3.28), on a alors

$$\begin{aligned} & \sum_{l=0}^{\infty} \sum_{m=-l}^l \left( \frac{\sqrt{l(l+1)}}{R_*} [(a_{lm}^{2n+1} + f_{lm}^{2n}) J_l(k_s R_*) + (e_{lm}^{2n} - c_{lm}^{2n+1}) H_l(k_s R_*)] \mathbf{v}_{lm} \right. \\ & \left. + k^2 \sqrt{l(l+1)} [(\tilde{f}_{lm}^{2n} + b_{lm}^{2n+1}) j_l(k_s R_*) + (\tilde{e}_{lm}^{2n} - d_{lm}^{2n+1}) h_l(k_s R_*)] \mathbf{u}_{lm} \right) = 0, \end{aligned}$$

ce qui donne

$$\begin{cases} (a_{lm}^{2n+1} + f_{lm}^{2n}) J_l(k_s R_*) + (e_{lm}^{2n} - c_{lm}^{2n+1}) H_l(k_s R_*) = 0, \\ (\tilde{f}_{lm}^{2n} + b_{lm}^{2n+1}) j_l(k_s R_*) + (\tilde{e}_{lm}^{2n} - d_{lm}^{2n+1}) h_l(k_s R_*) = 0. \end{cases} \quad (3.33)$$

### Calcul du taux de convergence de la méthode de Schwarz

Pour évaluer le taux de convergence de la méthode de Schwarz, on détermine l'opérateur des itérations  $K$  défini par la relation :

$$\Lambda^{n+1} = K \Lambda^n \quad (3.34)$$

où  $\Lambda^n = T(\mathbf{w}_{2n})$ , se développe dans la base formée par les harmoniques sphériques vectorielles (voir le théorème 3.4.2) sous la forme suivante :

$$\Lambda^n = \sum_{l=0}^{\infty} \sum_{m=-l}^l \left( e_{lm}^{2n} T(\mathbf{N}_{lm}) + f_{lm}^{2n} T(\mathbf{M}_{lm}) + \tilde{e}_{lm}^{2n} T(\mathbf{rot} \mathbf{N}_{lm}) + \tilde{f}_{lm}^{2n} T(\mathbf{rot} \mathbf{M}_{lm}) \right).$$

Par définition de  $T$  et de  $\mathbf{N}_{lm}$ , il vient :

$$\begin{aligned} T(\mathbf{N}_{lm}) &= \mathbf{rot} \mathbf{N}_{lm} \wedge \hat{x} - ik_s \hat{x} \wedge (\mathbf{N}_{lm} \wedge \hat{x}) \\ &= -\frac{\sqrt{l(l+1)}}{R} H_l(k_s R) \mathbf{v}_{lm} - ik_s \sqrt{l(l+1)} h_l(k_s R) (\mathbf{u}_{lm} \wedge \hat{x}) \\ &= -\frac{\sqrt{l(l+1)}}{R} [H_l(k_s R) - ik_s R h_l(k_s R)] \mathbf{v}_{lm}. \end{aligned}$$

Les quantités  $T(\mathbf{M}_{lm})$ ,  $T(\mathbf{rot N}_{lm})$  et  $T(\mathbf{rot M}_{lm})$  s'expriment d'une manière analogue. Par conséquent,

$$\Lambda^n = \sum_{l,m} \lambda_{lm}^{1,n} \mathbf{u}_{lm} + \lambda_{lm}^{2,n} \mathbf{v}_{lm}$$

avec

$$\left\{ \begin{array}{l} \lambda_{lm}^{1,n} = -\frac{\sqrt{l(l+1)}}{R} [\tilde{e}_{lm}^{2n} (ik_s H_l(k_s R) + k_s^2 R h_l(k_s R)) + \tilde{f}_{lm}^{2n} (ik_s J_l(k_s R) + k^2 R j_l(k_s R))], \\ \lambda_{lm}^{2,n} = -\frac{\sqrt{l(l+1)}}{R} [e_{lm}^{2n} (H_l(k_s R) - ik_s R h_l(k_s R)) + f_{lm}^{2n} (J_l(k_s R) - ik_s R j_l(k_s R))]. \end{array} \right.$$

et

$$\left\{ \begin{array}{l} \lambda_{lm}^{1,n+1} = \left(-\frac{\sqrt{l(l+1)}}{R}\right) d_{lm}^{2n+1} (ik_s H_l(k_s R) + k^2 R h_l(k_s R)), \\ \lambda_{lm}^{2,n+1} = \left(-\frac{\sqrt{l(l+1)}}{R}\right) c_{lm}^{2n+1} (H_l(k_s R) - ik_s R h_l(k_s R)). \end{array} \right.$$

Il en découle alors la relation suivante entre les coefficients du développement de  $\Lambda^n$  dans la base formée par les vecteurs  $\mathbf{u}_{lm}$  et  $\mathbf{v}_{lm}$  :

$$\left\{ \begin{array}{l} \lambda_{lm}^{1,n+1} = \tau_{1,l} \lambda_{lm}^{1,n} \\ \lambda_{lm}^{2,n+1} = \tau_{2,l} \lambda_{lm}^{2,n}. \end{array} \right.$$

où

$$\left\{ \begin{array}{l} \tau_{1,l}^{-1} = \frac{\tilde{e}_{lm}^{2n}}{d_{lm}^{2n+1}} + \frac{\tilde{f}_{lm}^{2n}}{d_{lm}^{2n+1}} \frac{ik_s J_l(k_s R) + k^2 R j_l(k_s R)}{ik_s H_l(k_s R) + k_s^2 R h_l(k_s R)}, \\ \tau_{2,l}^{-1} = \frac{e_{lm}^{2n}}{c_{lm}^{2n+1}} + \frac{f_{lm}^{2n}}{c_{lm}^{2n+1}} \frac{J_l(k_s R) - ik_s R j_l(k_s R)}{H_l(k_s R) - ik_s R h_l(k_s R)}. \end{array} \right.$$

En tenant compte des résultats (3.31) et (3.32) dans (3.33), on a

$$\left\{ \begin{array}{l} \tilde{e}_{lm}^{2n} = d_{lm}^{2n+1}, \\ e_{lm}^{2n} = c_{lm}^{2n+1}. \end{array} \right.$$

On a alors

$$\left\{ \begin{array}{l} \tau_{1,l}^{-1} = 1 - \frac{H_l(k_s R_*)}{J_l(k_s R_*)} \frac{J_l(k_s R) - ik_s R j_l(k_s R)}{H_l(k_s R) - ik_s R h_l(k_s R)}, \\ \tau_{2,l}^{-1} = 1 - \frac{h_l(k_s R_*)}{j_l(k_s R_*)} \frac{J_l(k_s R) - ik_s R j_l(k_s R)}{H_l(k_s R) - ik_s R h_l(k_s R)}. \end{array} \right. \quad (3.35)$$

Il en découle donc que l'opérateur des itérations  $K$  se diagonalise dans la base formée par les vecteurs  $\mathbf{u}_{lm}$ ,  $\mathbf{v}_{lm}$  et son spectre est formé par les valeurs propres  $\tau_{1,l}$  et  $\tau_{2,l}$ . L'algorithme de Schwarz avec recouvrement total converge lorsque  $|\tau_{1,l}|$  et  $|\tau_{2,l}|$  sont inférieurs à l'unité. L'étude asymptotique des fonctions sphériques  $j_l$  et  $h_l$  pour  $l$  grand et  $x$  fixé donne

$$j_l(x) \sim \frac{x^l}{(2l+1)!!},$$

$$h_l(x) \sim \frac{(2l-1)!!}{ix^{l+1}}.$$

Le symbole  $n!!$  désigne le produit de tous les entiers inférieurs à  $n$  et de même parité que  $n$ . On a alors

$$\tau_{1,l} \sim \tau_{2,l} \sim \frac{1}{1-R^{2l}}.$$

On conclut que la méthode de Schwarz avec recouvrement total converge si les frontières  $\Sigma$  et  $\Gamma$  sont assez éloignées. Ce résultat est confirmé par des tests numériques où on a changé l'épaisseur du domaine délimité par  $\Gamma$  et  $\Sigma$ .

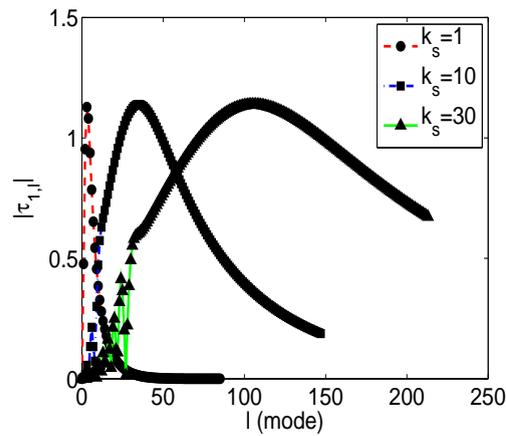


FIGURE 3.1 – Variation de  $\tau_{1,l}$  selon les modes  $l$  pour l'épaisseur  $e = \frac{\lambda}{100}$ . Cas  $k_s = 1$ ,  $k_s = 10$  et  $k_s = 30$ .

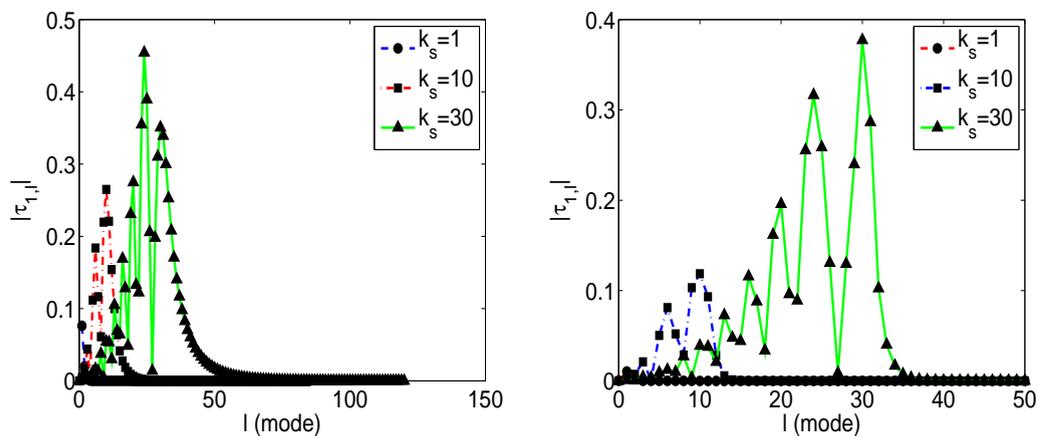


FIGURE 3.2 – Variation de  $|\tau_{1,l}|$  selon les modes  $l$  pour l'épaisseur  $e = \frac{\lambda}{4}$  (à gauche) et  $e = \lambda$  (à droite). Cas  $k_s = 1$ ,  $k_s = 10$  et  $k_s = 30$ .

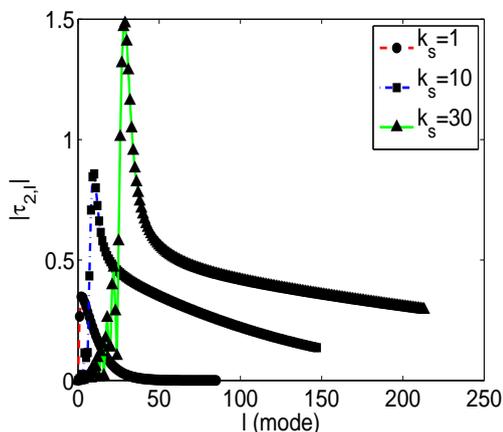


FIGURE 3.3 – Variation de  $|\tau_{2,l}|$  selon les modes  $l$  pour l'épaisseur  $e = \frac{\lambda}{100}$ . Cas  $k_s = 1$ ,  $k_s = 10$  et  $k_s = 30$ .

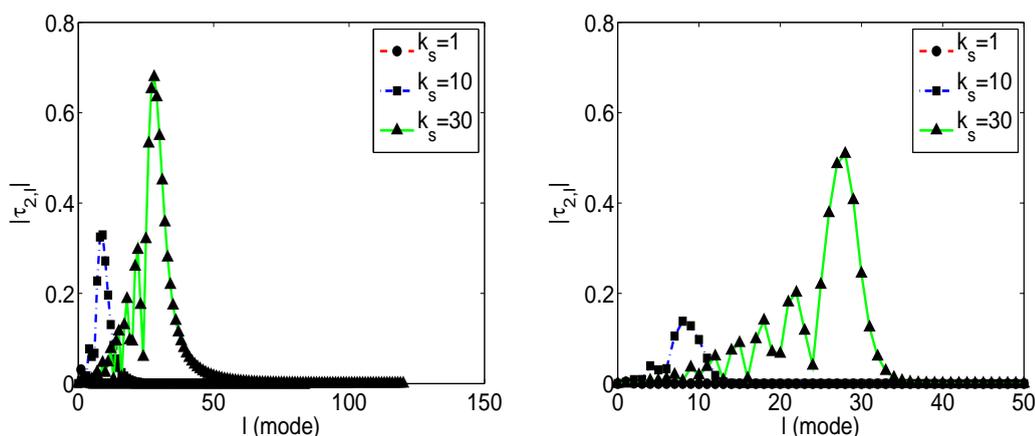


FIGURE 3.4 – Variation de  $|\tau_{2,l}|$  selon les modes  $l$  pour l'épaisseur  $e = \frac{\lambda}{4}$  (à gauche) et  $e = \lambda$  (à droite). Cas  $k_s = 1$ ,  $k_s = 10$  et  $k_s = 30$ .

Au regard de la figure 3.1, dans ce cas ( $R \simeq R_*$ ), on observe que pour un grand nombre de modes  $l$ , on a  $|\tau_{1,l}| > 1$  et plus le nombre d'onde  $k$  est grand, plus cette plage de modes est large. Par contre, en considérant une épaisseur égale à  $\frac{\lambda}{4}$  et  $\lambda$  (voir Fig. 3.2) la valeur maximale de  $|\tau_{1,l}|$  est inférieure à l'unité pour la totalité des modes. D'une manière similaire, on refait les mêmes cas tests pour étudier  $|\tau_{2,l}|$  et on observe les mêmes comportements (voir figures 3.3 et 3.4). Cette étude permet d'évaluer les cas de convergence de cet algorithme. Après cette analyse, le choix de l'épaisseur entre le bord de l'obstacle et la frontière artificielle joue un rôle important pour assurer la convergence de l'algorithme de Schwarz avec recouvrement total sans augmenter le coût de la résolution du problème. Pour cette raison, nous nous intéressons plutôt à un solveur rapide et convergeant lorsque le bord fictif est proche de l'obstacle. En fait, une seconde alternative plus efficace des

méthodes itératives consistent à utiliser des approximations polynomiales de la forme

$$\mathbf{E}^n = \mathbf{E}^0 + p_{n-1}(\mathcal{A} + \mathcal{C})r_0$$

où  $r_0 = \mathcal{F} - (\mathcal{A} + \mathcal{C})\mathbf{E}^0$  est le résidu initial et  $p_{n-1}(\mathcal{A} + \mathcal{C})$  est un polynôme de degré au plus  $n-1$ . Pour cela on peut citer par exemple les méthodes de Krylov comme la méthode GMRES ou la méthode BI-CGstab. Dans ce cas,

$$\mathbf{E}^n \in \mathbf{E}^0 + \mathcal{K}((\mathcal{A} + \mathcal{C}), r_0)$$

où  $\mathcal{K}((\mathcal{A} + \mathcal{C}), r_0) = \text{SPAN}(r_0, (\mathcal{A} + \mathcal{C})r_0, \dots, (\mathcal{A} + \mathcal{C})^{n-1}r_0)$  est l'espace de Krylov associé à  $(\mathcal{A} + \mathcal{C})$  et  $r_0$ .

L'expérimentation analytique et numérique de l'algorithme de Schwarz avec recouvrement total a orienté le choix d'un préconditionneur aux méthodes de type Krylov. Autrement dit nous nous intéressons plutôt à la résolution du problème préconditionné par  $\mathcal{A}$  :

$$(\mathcal{I} + \mathcal{A}^{-1}\mathcal{C})\mathbf{E} = \mathcal{A}^{-1}\mathcal{F}.$$

L'analyse faite sur la convergence de la méthode de Schwarz montre que dans le cas où le domaine de calcul est assez large toutes les valeurs propres de l'opérateur  $(-\mathcal{A}^{-1}\mathcal{C})$  sont dans le disque unité. Ceci induit que les méthodes de type Krylov convergent aussi. Si la distance entre les deux bords  $\Gamma$  et  $\Sigma$  n'est pas suffisamment grande alors on a un nombre fini de valeurs propres qui sont en dehors du disque unité. Ceci cause la divergence de la méthode de Schwarz. Dans ce cas, on peut toujours avoir recours à des méthodes de type Krylov qui s'adaptent à ce type de problème et assurent mieux la convergence ([53]). Nous proposons dans le chapitre suivant de montrer la convergence théorique de la méthode GMRES préconditionnée par le préconditionneur de Schwarz.



# Chapitre 4

## Analyse de la méthode de Schwarz et de la convergence de GMRES pour le problème de Maxwell régularisé

### 4.1 Equation de Maxwell régularisée

Afin de mettre en œuvre le couplage entre éléments finis et représentation intégrale pour les équations de Maxwell, nous avons fait le choix de travailler avec la librairie éléments finis MÉLINA ++. Compte tenu de sa structure et de l'état de son avancement, nous avons établi une approximation par des éléments finis de Lagrange (une analyse sera développée dans le chapitre 5). Ceci, nécessite le passage des équations de Maxwell classiques aux équations de Maxwell régularisées. En fait, pour garder l'emploi des éléments finis nodaux pour la résolution numérique de l'équation de Maxwell, de nombreux auteurs ont proposé une technique de régularisation ([13]). Cette technique consiste à ajouter un terme régularisant  $\nabla \text{div}$  à l'équation de Maxwell (3.10). On obtient une équation dite équation de Maxwell régularisée donnée par

$$\text{rot} \left( \frac{1}{\mu} \text{rot} \mathbf{E} \right) - \bar{\varepsilon} \nabla (s^{-1} \text{div} \varepsilon \mathbf{E}) - \omega^2 \varepsilon \mathbf{E} = 0. \quad (4.1)$$

En choisissant  $s^{-1} = 0$ , on se ramène à l'équation de Maxwell classique (3.10). Aussi, pour toute valeur  $s$  (réelle ou complexe), la solution  $\mathbf{E}$  de l'équation (3.10) sera toujours solution de (4.1) puisque on a  $\text{div} \varepsilon \mathbf{E} = 0$ .

**Remarque 4.1.1.** *Dans le vide, où  $\mu$  et  $\varepsilon$  sont constants et  $\sigma = 0$ . Pour un choix d'un  $s$  constant, l'équation (4.1) devient*

$$\text{rot rot} \mathbf{E} - t^{-1} \nabla (\text{div} \mathbf{E}) - k_s^2 \mathbf{E} = 0 \quad (4.2)$$

avec  $t^{-1} = s^{-1} \mu |\varepsilon|^2$ . Dans le cas particulier  $t = 1$  et par la relation vectorielle

$$\text{rot rot} \mathbf{E} - \nabla \text{div} \mathbf{E} = -\Delta \mathbf{E},$$

l'équation (4.2) n'est autre que l'équation de Helmholtz vectorielle :

$$-\Delta \mathbf{E} - k_s^2 \mathbf{E} = 0.$$

Nous allons étudier la convergence de l'algorithme de Schwarz avec recouvrement total appliqué à la formulation régularisée des équations de Maxwell, établie par M. Lenoir et C. Hazard dans [13].

## 4.2 Problème de Maxwell régularisé en domaine non borné

On se place dans le même cadre physique que le problème présenté dans le chapitre précédent. On s'intéresse à la résolution du problème de Maxwell suivant :

$$\left\{ \begin{array}{l} \mathbf{rot rot E} - k_s^2 \mathbf{E} = 0 \text{ dans } \Omega^e, \\ \mathbf{E} \wedge \mathbf{n}_e = 0 \text{ sur } \Gamma, \\ \lim_{R \rightarrow \infty} \int_{\|x\|=R} \|\mathbf{rot E}^s \wedge \mathbf{n}_e - ik_s \mathbf{E}^s\|^2 d\gamma = 0. \end{array} \right. \quad (4.3)$$

Le problème (19) est équivalent au problème régularisé suivant ([13]) :

$$\left\{ \begin{array}{l} \mathbf{rot rot E} - t^{-1} \nabla(\operatorname{div E}) - k_s^2 \mathbf{E} = 0 \text{ dans } \Omega^e, \\ \mathbf{E} \wedge \mathbf{n}_e = 0, \operatorname{div E} = 0 \text{ sur } \Gamma, \\ \lim_{R \rightarrow \infty} \int_{\|x\|=R} \|\mathbf{rot E}^s \wedge \mathbf{n}_e - ik_s \mathbf{n}_e \wedge (\mathbf{E}^s \wedge \mathbf{n}_e)\|^2 d\gamma = 0, \\ \lim_{R \rightarrow \infty} \int_{\|x\|=R} |\sqrt{t^{-1}} \operatorname{div E}^s - ik_s \mathbf{E}^s \cdot \mathbf{n}_e|^2 d\gamma = 0. \end{array} \right. \quad (4.4)$$

La première (resp. deuxième) condition à l'infini est dite condition transverse (resp. radiale). Pour résoudre le problème en domaine non borné (4.4), on se propose d'utiliser la méthode de couplage entre éléments finis et représentation intégrale. On se ramène à résoudre le problème de Maxwell tronqué posé sur le domaine borné  $\Omega$  délimité par les deux frontières  $\Sigma$  et  $\Gamma$  :

$$\left\{ \begin{array}{l} \mathbf{rot rot E} - t^{-1} \nabla(\operatorname{div E}) - k_s^2 \mathbf{E} = 0 \text{ dans } \Omega, \\ \mathbf{E} \wedge \mathbf{n}_e = 0, \operatorname{div E} = 0 \text{ sur } \Gamma, \\ \mathbf{rot E} \wedge \mathbf{n}_\sigma + \nu_1 \mathbf{n}_\sigma \wedge (\mathbf{E} \wedge \mathbf{n}_\sigma) = T_{\nu_1}(\mathbf{E}^{inc} - \mathcal{I}_t(\Gamma, \mathbf{E})) \text{ sur } \Sigma, \\ \operatorname{div E} + \nu_2 \mathbf{E} \cdot \mathbf{n}_\sigma = N_{\nu_2}(\mathbf{E}^{inc} - \mathcal{I}_t(\Gamma, \mathbf{E})) \text{ sur } \Sigma. \end{array} \right. \quad (4.5)$$

$\mathbf{n}_\sigma$  est la normale unitaire extérieure à  $\Sigma$  orientée vers l'extérieur de  $\Omega$ . Les opérateurs  $T_{\nu_1}$  et  $N_{\nu_2}$  sont définis par :

$$T_{\nu_1} \mathbf{E} = \mathbf{rot E} \wedge \mathbf{n}_\sigma + \nu_1 \mathbf{n}_\sigma \wedge (\mathbf{E} \wedge \mathbf{n}_\sigma) \quad \text{et} \quad N_{\nu_2} \mathbf{E} = \operatorname{div E} + \nu_2 \mathbf{E} \cdot \mathbf{n}_\sigma.$$

où  $\nu_1$  et  $\nu_2$  sont des paramètres complexes arbitraires. Les deux conditions aux limites sur  $\Sigma$  sont appelées "conditions de couplage". Pour tout  $x \in \Omega^e$ ,

$$\mathcal{I}_t(\Gamma, \mathbf{E})(x) = \int_{\Gamma} \mathcal{G}_t(x, y) (\mathbf{rot E}(y) \wedge \mathbf{n}(y) + t^{-1} \mathbf{n}(y) \operatorname{div E}(y)) d\gamma(y)$$

$$- \int_{\Gamma} \mathbf{rot}_y \mathcal{G}_t(x, y) (\mathbf{E}(y) \wedge \mathbf{n}(y)) d\gamma(y) - t^{-1} \int_{\Gamma} \operatorname{div}_y \mathcal{G}_t(x, y)^T (\mathbf{E}(y) \cdot \mathbf{n}(y)) d\gamma(y). \quad (4.6)$$

Nous introduisons maintenant l'espace de Hilbert

$$\mathbb{H}_t = \{ \mathbf{E} \in H(\mathbf{rot} \mathbf{E}, \Omega) \text{ tel que } \operatorname{div} \mathbf{E} \in L^2(\Omega), (\mathbf{E} \wedge \mathbf{n})_{\Gamma} = 0, \\ (\mathbf{E} \wedge \mathbf{n})_{\Sigma} \in L^2(\Sigma)^3 \text{ et } (\mathbf{E} \cdot \mathbf{n})_{\Sigma} \in L^2(\Sigma) \}$$

muni du produit scalaire  $(\cdot, \cdot)_t$  défini par :

$$(\mathbf{E}, \mathbf{E}')_t = \int_{\Omega} \mathbf{E} \cdot \mathbf{E}' d\Omega + \int_{\Omega} \mathbf{rot} \mathbf{E} \cdot \mathbf{rot} \mathbf{E}' d\Omega \\ + |t|^{-1} \int_{\Omega} \operatorname{div} \mathbf{E} \operatorname{div} \mathbf{E}' d\Omega + \int_{\Sigma} (\mathbf{E} \wedge \mathbf{n}) \cdot (\mathbf{E}' \wedge \mathbf{n}) d\sigma + |t|^{-1} \int_{\Sigma} (\mathbf{E} \cdot \mathbf{n})(\mathbf{E}' \cdot \mathbf{n}) d\sigma.$$

La formulation de représentation intégrale  $\mathcal{I}_t(\Gamma, \mathbf{E})$  n'est pas toujours définie pour tout  $\mathbf{E} \in \mathbb{H}_t$ . Pour cela, on considère un relèvement régulier  $\mathcal{R}$  de  $\Gamma$  vers  $\Omega$ . Dans la partie numérique de ce rapport, on détaillera le choix qu'on a fait pour la définition de  $\mathcal{R}\mathcal{G}_t$  dans MÉLINA++. La formule de représentation intégrale modifiée à l'extérieur de  $\Gamma$  est

$$\mathcal{I}_t^{\mathcal{R}}(\Gamma, \mathbf{E})(x) = -k_s^2 \int_{\Omega} \mathcal{R}\mathcal{G}_t(x, y) \mathbf{E}(y) d\Omega(y) + \int_{\Omega} \mathbf{rot}_y \mathcal{R}\mathcal{G}_t(x, y) \mathbf{rot} \mathbf{E}(y) d\Omega(y) \\ + t^{-1} \int_{\Omega} \operatorname{div}_y \mathcal{R}\mathcal{G}_t(x, y)^T \operatorname{div} \mathbf{E}(y) d\gamma(y) - \int_{\Gamma} \mathbf{rot}_y \mathcal{G}_t(x, y) (\mathbf{E}(y) \wedge \mathbf{n}(y)) d\gamma(y) \\ - t^{-1} \int_{\Gamma} (\operatorname{div}_y \mathcal{G}_t(x, y)^T) (\mathbf{E}(y) \cdot \mathbf{n}(y)) d\gamma(y). \quad (4.7)$$

On obtient alors le problème posé en domaine borné suivant qui consiste à trouver  $\mathbf{E}$  dans  $\mathbb{H}_t$  tel que

$$\left\{ \begin{array}{l} \mathbf{rot} \mathbf{rot} \mathbf{E} - t^{-1} \nabla(\operatorname{div} \mathbf{E}) - k_s^2 \mathbf{E} = 0 \text{ dans } \Omega, \\ \mathbf{E} \wedge \mathbf{n}_e = 0, \operatorname{div} \mathbf{E} = 0 \text{ sur } \Gamma, \\ \mathbf{rot} \mathbf{E} \wedge \mathbf{n}_{\sigma} + \nu_1 \mathbf{n}_{\sigma} \wedge (\mathbf{E} \wedge \mathbf{n}_{\sigma}) = T_{\nu_1}(\mathbf{E}^{inc} - \mathcal{I}_t^{\mathcal{R}}(\Gamma, \mathbf{E})) \text{ sur } \Sigma, \\ \operatorname{div} \mathbf{E} + \nu_2 \mathbf{E} \cdot \mathbf{n}_{\sigma} = N_{\nu_2}(\mathbf{E}^{inc} - \mathcal{I}_t^{\mathcal{R}}(\Gamma, \mathbf{E})) \text{ sur } \Sigma. \end{array} \right. \quad (4.8)$$

**Théorème 4.2.1.** : (Equivalence entre le problème initial et le problème tronqué) :

Soient  $\nu_1 \in \mathbb{C}$  choisi de telle sorte que  $\operatorname{Im} \nu_1 < 0$  et  $\nu_2 \in \mathbb{C}$  tel que  $\operatorname{Im} \nu_2 < 0$ , les problèmes (4.4) et (4.8) sont équivalents, tant pour l'existence que pour l'unicité : le problème (4.4) admet au moins (respectivement au plus) une solution si et seulement si il est de même pour le problème tronqué (4.8). La relation entre ces solutions :

- Si  $\mathbf{E}$  est solution de (4.4) alors la restriction de  $\mathbf{E}$  à  $\Omega$  est solution de (4.8).
- Si  $\mathbf{E}_{\Omega}$  est solution de (4.8) alors  $\mathbf{E}$  défini par

$$\left\{ \begin{array}{l} \mathbf{E} = \mathbf{E}_{\Omega} \text{ dans } \Omega, \\ \mathbf{E} = \mathbf{E}^{inc} - \mathcal{I}_t^{\mathcal{R}}(\Gamma, \mathbf{E}_{\Omega}) \text{ dans } \Omega^e \setminus \Omega. \end{array} \right.$$

est solution de (4.4).

Pour la démonstration, on se référera à [13].

### 4.3 Formulation variationnelle, existence et unicité

On suppose que  $\mathbf{E}$  est suffisamment régulière. A l'aide des formules de Green suivante :

$$\begin{aligned} \int_{\Omega} \nabla \operatorname{div} \mathbf{E} \cdot \mathbf{E} \, d\Omega &= - \int_{\Omega} \operatorname{div} \mathbf{E} \operatorname{div} \mathbf{E}' + \int_{\Sigma} \operatorname{div} \mathbf{E} (\mathbf{E}' \cdot \mathbf{n}_{\sigma}) \, d\sigma - \int_{\Gamma} \operatorname{div} \mathbf{E} (\mathbf{E}' \cdot \mathbf{n}) \, d\gamma, \\ \int_{\Omega} \operatorname{rot} \mathbf{E} \cdot \operatorname{rot} \mathbf{E}' \, d\Omega &= \int_{\Omega} \operatorname{rot} \operatorname{rot} \mathbf{E} \cdot \mathbf{E}' \, d\Omega + \int_{\Sigma} \operatorname{rot} \mathbf{E} \cdot (\mathbf{n}_{\sigma} \wedge \mathbf{E}') \, d\sigma \\ &\quad - \int_{\Gamma} \operatorname{rot} \mathbf{E} \cdot (\mathbf{n} \wedge \mathbf{E}') \, d\gamma, \end{aligned}$$

pour toute fonction test  $\mathbf{E}'$ , on a

$$\begin{aligned} &\int_{\Omega} \operatorname{rot} \mathbf{E} \cdot \operatorname{rot} \mathbf{E}' + t^{-1} \operatorname{div} \mathbf{E} \operatorname{div} \mathbf{E}' - k_s^2 \mathbf{E} \cdot \mathbf{E}' \, d\Omega + \int_{\Sigma} (\mathbf{n}_{\sigma} \wedge \operatorname{rot} \mathbf{E}) \cdot \mathbf{E}' \, d\sigma \\ &- \int_{\Gamma} (\mathbf{n} \wedge \operatorname{rot} \mathbf{E}) \cdot \mathbf{E}' \, d\gamma - t^{-1} \int_{\Sigma} \operatorname{div} \mathbf{E} (\mathbf{n}_{\sigma} \cdot \mathbf{E}') \, d\sigma + t^{-1} \int_{\Gamma} \operatorname{div} \mathbf{E} (\mathbf{n} \cdot \mathbf{E}') \, d\gamma = 0. \end{aligned}$$

En utilisant les conditions sur  $\Sigma$  on a

$$\begin{aligned} &\int_{\Omega} \operatorname{rot} \mathbf{E} \cdot \operatorname{rot} \mathbf{E}' + t^{-1} \operatorname{div} \mathbf{E} \operatorname{div} \mathbf{E}' - k_s^2 \mathbf{E} \cdot \mathbf{E}' \, d\Omega + \nu_1 \int_{\Sigma} (\mathbf{n}_{\sigma} \wedge \mathbf{E}) \cdot (\mathbf{n}_{\sigma} \wedge \mathbf{E}') \, d\sigma \\ &+ t^{-1} \nu_2 \int_{\Sigma} (\mathbf{n}_{\sigma} \cdot \mathbf{E}) (\mathbf{n}_{\sigma} \cdot \mathbf{E}') \, d\sigma + \int_{\Sigma} T_{\nu_1}(\mathcal{I}_t^{\mathcal{R}}(\Gamma, \mathbf{E})) \cdot \mathbf{E}' \, d\sigma + t^{-1} \int_{\Sigma} N_{\nu_2}(\mathcal{I}_t^{\mathcal{R}}(\Gamma, \mathbf{E})) (\mathbf{n}_{\sigma} \cdot \mathbf{E}') \, d\sigma \\ &= \int_{\Sigma} T_{\nu_1} \mathbf{E}^{inc} \cdot \mathbf{E}' \, d\sigma + t^{-1} \int_{\Sigma} N_{\nu_2} \mathbf{E}^{inc} (\mathbf{n}_{\sigma} \cdot \mathbf{E}') \, d\sigma. \end{aligned}$$

D'après le théorème de représentation de Riesz, le problème variationnel de (4.8) consiste à trouver  $\mathbf{E} \in \mathbb{H}_t$  tel que

$$(\mathcal{A}_t^c \mathbf{E}, \mathbf{E}')_t + (\mathcal{C}_t^c \mathbf{E}, \mathbf{E}')_t = (\mathcal{F}_t, \mathbf{E}')_t \quad \forall \mathbf{E}' \in \mathbb{H}_t. \quad (4.9)$$

Les opérateurs  $\mathcal{A}_t^c$  et  $\mathcal{C}_t^c : \mathbb{H}_t \rightarrow \mathbb{H}_t$  sont définis selon les formules suivantes :

$$\begin{aligned} (\mathcal{A}_t^c \mathbf{E}, \mathbf{E}')_t &= \int_{\Omega} \operatorname{rot} \mathbf{E} \cdot \operatorname{rot} \mathbf{E}' + t^{-1} \operatorname{div} \mathbf{E} \operatorname{div} \mathbf{E}' + k_s^2 \mathbf{E} \cdot \mathbf{E}' \, d\Omega + \nu_1 \int_{\Sigma} (\mathbf{n}_{\sigma} \wedge \mathbf{E}) \cdot (\mathbf{n}_{\sigma} \wedge \mathbf{E}') \, d\sigma \\ &+ t^{-1} \nu_2 \int_{\Sigma} (\mathbf{n}_{\sigma} \cdot \mathbf{E}) (\mathbf{n}_{\sigma} \cdot \mathbf{E}') \, d\sigma, \end{aligned}$$

$$\begin{aligned} (\mathcal{C}_t^c \mathbf{E}, \mathbf{E}')_t &= -2k_s^2 \int_{\Omega} \mathbf{E} \cdot \mathbf{E}' \, d\Omega + \int_{\Sigma} T_{\nu_1}(\mathcal{I}_t^{\mathcal{R}}(\Gamma, \mathbf{E})) \cdot \mathbf{E}' \, d\sigma \\ &+ t^{-1} \int_{\Sigma} N_{\nu_2}(\mathcal{I}_t^{\mathcal{R}}(\Gamma, \mathbf{E})) (\mathbf{n}_{\sigma} \cdot \mathbf{E}') \, d\sigma, \end{aligned}$$

et  $\mathcal{F}_t$  est donné par :  $\forall \mathbf{E}' \in \mathbb{H}_t$

$$(\mathcal{F}_t, \mathbf{E}')_t = \int_{\Sigma} T_{\nu_1} \mathbf{E}^{inc} \cdot \mathbf{E}' \, d\sigma + t^{-1} \int_{\Sigma} N_{\nu_2} \mathbf{E}^{inc} (\mathbf{n}_{\sigma} \cdot \mathbf{E}') \, d\sigma.$$

$\mathcal{A}_t^c$  et  $\mathcal{C}_t^c$  sont respectivement un automorphisme et un opérateur compact ([13]). On constate donc que le problème (4.5) s'exprime sous forme d'une équation de Fredholm. On sait que dans ce cas les questions de l'existence et de l'unicité de la solution sont équivalentes. Si on étudie le problème homogène (cas où  $F_t$  est nulle) alors la solution est nécessairement triviale (voir théorème 3.1 de [13]). Cela donne que le problème (4.5) admet une unique solution pour tout choix du second membre et elle en dépend continûment. Le théorème (4.2.1) assure que le problème en domaine non borné (4.4) est aussi bien posé.

**Remarque 4.3.1.** *En général,  $\mathbb{H}_t(\Omega) \not\subset H^1(\Omega)^3$  [54]. Si le domaine  $\Omega$  est lipschitz et ne présente pas des singularités géométriques entrantes (coins, arêtes, sommets) alors  $\mathbb{H}_t(\Omega) \cap H^1(\Omega)^3$  est dense dans  $\mathbb{H}_t(\Omega)$  [55]. On peut alors approcher la solution par des éléments finis de Lagrange.*

*Par contre, si le domaine présente des arêtes ou des coins entrants alors  $\mathbb{H}_t(\Omega) \cap H^1(\Omega)^3$  est un sous-espace fermé de  $\mathbb{H}_t(\Omega)$  ([56]). Par conséquent,  $\mathbb{H}_t(\Omega)$  n'admet pas de sous-espace dense inclus dans  $H^1(\Omega)^3$ . Dans ce cas, la discrétisation par des éléments finis de Lagrange de l'équation de Maxwell régularisée conduit à une solution qui ne converge pas vers la solution physique ([57]). De fait dans cette dernière configuration, il est conseillé de faire appel à la méthode de régularisation à poids ([58], [59]). Pour la résolution numérique du problème (4.4), nous allons nous restreindre au cas des obstacles qui ne présentent pas des singularités géométriques entrantes.*

## 4.4 Méthode de Schwarz avec recouvrement total pour le problème de Maxwell régularisé

Dans cette section, on présente l'application de la méthode de Schwarz avec recouvrement total sur le problème de Maxwell régularisé sans donner de détails. Pour plus de précisions, on invite le lecteur à lire le chapitre précédent. Le but de cette section est de comparer les résultats obtenus à ceux trouvés pour le cas du problème de Maxwell classique. La résolution du problème en domaine borné (4.5) en se basant sur l'idée du point fixe de Cauchy mène à déterminer  $\mathbf{E}_{n+1}$  solution de

$$\left\{ \begin{array}{l} \text{rot rot } \mathbf{E}_{n+1} - t^{-1} \nabla(\text{div } \mathbf{E}_{n+1}) - k_s^2 \mathbf{E}_{n+1} = 0 \text{ dans } \Omega, \\ \mathbf{E}_{n+1} \wedge \mathbf{n} = 0, \text{div } \mathbf{E}_{n+1} = 0 \text{ sur } \Gamma, \\ \text{rot } \mathbf{E}_{n+1} \wedge \mathbf{n}_\sigma + \nu_1 \mathbf{n}_\sigma \wedge (\mathbf{E} \wedge \mathbf{n}_\sigma) = T_{\nu_1}(\mathbf{E}^{inc} - \mathcal{I}_t^{\mathcal{R}}(\Gamma, \mathbf{E}_n)) \text{ sur } \Sigma, \\ \text{div } \mathbf{E}_{n+1} + \nu_2 \mathbf{E}_{n+1} \cdot \mathbf{n}_\sigma = N_{\nu_2}(\mathbf{E}^{inc} - \mathcal{I}_t^{\mathcal{R}}(\Gamma, \mathbf{E}_n)) \text{ sur } \Sigma. \end{array} \right. \quad (4.10)$$

L'algorithme précédent s'interprète comme une méthode de Schwarz avec recouvrement total appliqué au problème (4.8) et revient donc à construire une suite  $(\mathbf{E}_n)_n$  vérifiant ces

deux problèmes :

$$\left\{ \begin{array}{l} \mathbf{rot} \mathbf{rot} \mathbf{E}_{2n+1} - t^{-1} \nabla(\operatorname{div} \mathbf{E}_{2n+1}) - k_s^2 \mathbf{E}_{2n+1} = 0 \text{ dans } \Omega_i \cup \Omega^e, \\ \mathbf{n} \wedge [\mathbf{E}_{2n+1}] = 0 \text{ sur } \Gamma, \\ \mathbf{n} \wedge [\mathbf{rot} \mathbf{E}_{2n+1}] = -\mathbf{n} \wedge \mathbf{rot} \mathbf{E}_{2n} \text{ sur } \Gamma, \\ [\operatorname{div} \mathbf{E}_{2n+1}] = 0 \text{ sur } \Gamma, \\ \mathbf{n} \cdot [\mathbf{E}_{2n+1}] = -\mathbf{n} \cdot [\mathbf{E}_{2n}] \text{ sur } \Gamma, \\ \lim_{R \rightarrow \infty} \int_{\|x\|=R} \|\mathbf{rot} \mathbf{E}_{2n+1}^s \wedge \mathbf{n}_e - ik_s \mathbf{n}_e \wedge (\mathbf{E}_{2n+1}^s \wedge \mathbf{n}_e)\|^2 d\gamma = 0, \\ \lim_{R \rightarrow \infty} \int_{\|x\|=R} |\sqrt{t^{-1}} \operatorname{div} \mathbf{E}_{2n+1}^s - ik_s \mathbf{E}_{2n+1}^s \cdot \mathbf{n}_e|^2 d\gamma = 0. \end{array} \right. \quad (4.11)$$

$$\left\{ \begin{array}{l} \mathbf{rot} \mathbf{rot} \mathbf{E}_{2n+2} - t^{-1} \nabla(\operatorname{div} \mathbf{E}_{2n+2}) - k_s^2 \mathbf{E}_{2n+2} = 0 \text{ dans } \Omega, \\ \mathbf{E}_{2n+2} \wedge \mathbf{n} = 0 \text{ sur } \Gamma, \\ \operatorname{div} \mathbf{E}_{2n+2} = 0 \text{ sur } \Gamma, \\ (\mathbf{rot} \mathbf{E}_{2n+2}) \wedge \mathbf{n}_\sigma + \nu_1 (\mathbf{E}_{2n+2})_t = \\ \qquad \qquad \qquad (\mathbf{rot} \mathbf{E}_{2n+1}) \wedge \mathbf{n}_\sigma + \nu_1 (\mathbf{E}_{2n+1})_t \text{ sur } \Sigma, \\ \operatorname{div} \mathbf{E}_{2n+2} + \nu_2 \mathbf{E}_{2n+2} \cdot \mathbf{n}_\sigma = \\ \qquad \qquad \qquad \operatorname{div} \mathbf{E}_{2n+1} + \nu_2 \mathbf{E}_{2n+1} \cdot \mathbf{n}_\sigma \text{ sur } \Sigma. \end{array} \right. \quad (4.12)$$

La solution  $\mathbf{E}_{2n+1}$  du problème (4.11) équivaut à une détermination au moyen d'une représentation intégrale. Ainsi sa considération dans la deuxième condition de (4.12) ramène au problème de départ (4.10). L'algorithme de Schwarz avec recouvrement total peut s'écrire à la  $n^{\text{ième}}$  itération comme suit

$$\mathcal{A}_t \mathbf{E}^{n+1} = -\mathcal{C}_t \mathbf{E}^n + \mathcal{F}_t \quad (4.13)$$

où les opérateurs  $\mathcal{A}_t$  et  $\mathcal{C}_t : \mathbb{H}_t \rightarrow \mathbb{H}_t$  sont définis selon les formules suivantes :

$$\begin{aligned} (\mathcal{A}_t \mathbf{E}, \mathbf{E}')_t &= \int_{\Omega} \mathbf{rot} \mathbf{E} \cdot \mathbf{rot} \mathbf{E}' + t^{-1} \operatorname{div} \mathbf{E} \operatorname{div} \mathbf{E}' - k_s^2 \mathbf{E} \cdot \mathbf{E}' d\Omega \\ &\quad + \nu_1 \int_{\Sigma} (\mathbf{n}_\sigma \wedge \mathbf{E}) \cdot (\mathbf{n}_\sigma \wedge \mathbf{E}') d\sigma + t^{-1} \nu_2 \int_{\Sigma} (\mathbf{n}_\sigma \cdot \mathbf{E})(\mathbf{n}_\sigma \cdot \mathbf{E}') d\sigma, \end{aligned}$$

$$(\mathcal{C}_t \mathbf{E}, \mathbf{E}')_t = \int_{\Sigma} T_{\nu_1}(\mathcal{I}_t^{\mathcal{R}}(\Gamma, \mathbf{E})) \cdot \mathbf{E}' d\sigma + t^{-1} \int_{\Sigma} N_{\nu_2}(\mathcal{I}_t^{\mathcal{R}}(\Gamma, \mathbf{E})) (\mathbf{n}_\sigma \cdot \mathbf{E}') d\sigma.$$

$\mathcal{A}_t$  est inversible. La preuve se fait de la même façon que la proposition 3.3.1. L'équation (4.13) s'écrit d'une manière équivalente comme

$$\mathbf{E}^{n+1} = -\mathcal{A}_t^{-1} \mathcal{C}_t \mathbf{E}^n + \mathcal{A}_t^{-1} \mathcal{F}_t. \quad (4.14)$$

## 4.5 Investigation analytique en 3D :

Dans cette partie, on se propose d'étudier la convergence de l'algorithme de Schwarz avec recouvrement total pour le problème de Maxwell régularisé extérieur (4.4) . Pour cela, on se place toujours dans le cas d'une boule de rayon  $R_* = 1$  qu'on notera  $B_1$  et on suppose que  $\Sigma$  est la frontière de la boule concentrique à  $B_1$  et de rayon  $R > 1$ . Puisque le problème de Maxwell classique est équivalent au problème de Maxwell régularisé dans un domaine non borné, la solution du problème de (4.4) est donnée par ([8])

$$\mathbf{E} = \sum_{l=0}^{\infty} \sum_{m=-l}^l (\alpha_{lm} \mathbf{N}_{lm} + \beta_{lm} \mathbf{rot} \mathbf{N}_{lm}) + \mathbf{E}^{inc}.$$

On note  $(\mathbf{w}_n)_n$  la suite des erreurs :

$$\mathbf{w}_{2n} = \mathbf{E}_{2n} - \mathbf{E} \text{ dans } \Omega \text{ et } \mathbf{w}_{2n+1} = \begin{cases} \mathbf{E}_{2n+1} - \mathbf{E} \text{ dans } \Omega^e, \\ \mathbf{E}_{2n+1} \text{ dans } \Omega_i. \end{cases}$$

$$\left\{ \begin{array}{l} \mathbf{rot} \mathbf{rot} \mathbf{w}_{2n+1} - t^{-1} \nabla(\operatorname{div} \mathbf{w}_{2n+1}) - k_s^2 \mathbf{w}_{2n+1} = 0 \text{ dans } \Omega_i \cup \Omega^e, \\ \mathbf{n} \wedge [\mathbf{w}_{2n+1}] = -\mathbf{n} \wedge \mathbf{w}_{2n} = 0 \text{ sur } \Gamma, \\ \mathbf{n} \wedge [\mathbf{rot} \mathbf{w}_{2n+1}] = -\mathbf{n} \wedge \mathbf{rot} \mathbf{w}_{2n} \text{ sur } \Gamma, \\ [\operatorname{div} \mathbf{w}_{2n+1}] = 0 \text{ sur } \Gamma, \\ \mathbf{n} \cdot [\mathbf{w}_{2n+1}] = -\mathbf{n} \cdot \mathbf{w}_{2n} \text{ sur } \Gamma, \\ \lim_{R \rightarrow \infty} \int_{\|x\|=R} \|\mathbf{rot} \mathbf{w}_{2n+1} \wedge \mathbf{n}_e - ik_s \mathbf{n} \wedge (\mathbf{w}_{2n+1} \wedge \mathbf{n}_e)\|^2 d\gamma = 0, \\ \lim_{R \rightarrow \infty} \int_{\|x\|=R} |\sqrt{t^{-1}} \operatorname{div} \mathbf{w}_{2n+1} - ik_s \mathbf{w}_{2n+1} \cdot \mathbf{n}_e|^2 d\gamma = 0. \end{array} \right. \quad (4.15)$$

et  $\mathbf{w}_{2n+2}$  est solution du problème dans le domaine borné  $\Omega$  :

$$\left\{ \begin{array}{l} \mathbf{rot} \mathbf{rot} \mathbf{w}_{2n+2} - t^{-1} \nabla(\operatorname{div} \mathbf{w}_{2n+2}) - k_s^2 \mathbf{w}_{2n+2} = 0 \text{ dans } \Omega, \\ \mathbf{w}_{2n+2} \wedge \mathbf{n} = 0, \operatorname{div} \mathbf{w}_{2n+2} = 0 \text{ sur } \Gamma, \\ \mathbf{rot} \mathbf{w}_{2n+2} \wedge \mathbf{n}_\sigma + \nu_1 (\mathbf{w}_{2n+2})_t = \\ \quad \mathbf{rot} \mathbf{w}_{2n+1} \wedge \mathbf{n}_\sigma + \nu_1 (\mathbf{w}_{2n+1})_t \text{ sur } \Sigma, \\ \operatorname{div} \mathbf{w}_{2n+2} + \nu_2 \mathbf{w}_{2n+2} \cdot \mathbf{n}_\sigma = \operatorname{div} \mathbf{w}_{2n+1} + \nu_2 \mathbf{w}_{2n+1} \cdot \mathbf{n}_\sigma \text{ sur } \Sigma. \end{array} \right. \quad (4.16)$$

En comparant l'étude faite pour le problème de Maxwell classique, on remarque qu'on va prendre en considération en plus la condition aux limites sur  $\Gamma$

$$\operatorname{div} \mathbf{w}_{2n+2} = 0, \quad (4.17)$$

les conditions de transmission sur  $\Gamma$

$$[\operatorname{div} \mathbf{w}_{2n+1}] = 0 \quad (4.18)$$

et

$$\mathbf{n} \cdot [\mathbf{w}_{2n+1}] = -\mathbf{n} \cdot \mathbf{w}_{2n}. \quad (4.19)$$

On a aussi une condition supplémentaire sur  $\Sigma$

$$\operatorname{div} \mathbf{w}_{2n+2} + \nu_2 \mathbf{w}_{2n+2} \cdot \mathbf{n}_\sigma = \operatorname{div} \mathbf{w}_{2n+1} + \nu_2 \mathbf{w}_{2n+1} \cdot \mathbf{n}_\sigma. \quad (4.20)$$

Comme  $\mathbf{N}_{lm}$ ,  $\mathbf{M}_{lm}$ ,  $\operatorname{rot} \mathbf{N}_{lm}$  et  $\operatorname{rot} \mathbf{M}_{lm}$  sont des fonctions à divergence est nulle ainsi les deux conditions (4.17) et (4.18) donnent deux équations redondantes. Comme  $\mathbf{N}_{lm}$  et  $\mathbf{M}_{lm}$  sont à composantes normales nulles, on en déduit de (4.19) que

$$b_{l,m}^{2n+1} \mathbf{n} \cdot \operatorname{rot} \mathbf{M}_{lm} - d_{l,m}^{2n+1} \mathbf{n} \cdot \operatorname{rot} \mathbf{N}_{lm} + \tilde{e}_{lm}^{2n} \mathbf{n} \cdot \operatorname{rot} \mathbf{N}_{lm} + \tilde{f}_{lm}^{2n} \mathbf{n} \cdot \operatorname{rot} \mathbf{M}_{lm} = 0.$$

On obtient alors la deuxième équation de (3.33) trouvée pour l'étude du problème de Maxwell classique.

Pour déterminer le taux de convergence de la méthode de Schwarz avec recouvrement total pour le problème de Maxwell régularisé, nous considérons les deux conditions sur le bord  $\Sigma$  qui assurent l'échange entre les deux sous problèmes (4.15) et (4.16). La première condition est la même que celle étudiée pour le cas du problème de Maxwell classique. Par suite, la relation qui lie  $\Lambda^{n+1}$  et  $\Lambda^n$  dans (3.34) reste inchangée par rapport à ce qui est fait dans le chapitre précédent (voir le système d'équations (3.35) qui fait intervenir  $\tau_{1,l}$  et  $\tau_{2,l}$ ). De la deuxième condition sur  $\Sigma$ , on détermine l'opérateur des itérations vérifiant :

$$\delta^{n+1} = \tau_{3,l} \delta^n$$

avec

$$\delta^n = -ik_s \sum_{l=0}^{\infty} \sum_{m=-l}^l \frac{1}{R} l(l+1) (\tilde{e}_{lm}^{2n} h_l(k_s R) + \tilde{f}_{lm}^{2n} j_l(k_s R)) Y_l^m$$

et

$$\begin{aligned} \delta^{n+1} &= -ik_s \sum_{l=0}^{\infty} \sum_{m=-l}^l d_{lm}^{2n+1} \operatorname{rot} \mathbf{N}_{lm} \cdot \hat{x} \\ &= -ik_s \sum_{l=0}^{\infty} \sum_{m=-l}^l d_{lm}^{2n+1} \frac{1}{R} l(l+1) h_l(k_s R) Y_l^m. \end{aligned}$$

En déterminant la relation entre les coefficients du développement de  $\delta^{n+1}$  et  $\delta^n$  dans la base des harmoniques sphériques scalaires  $Y_l^m$  on a

$$\tau_{3,l}^{-1} = 1 - \frac{H_l(k_s R_*) j_l(k_s R)}{J_l(k_s R_*) h_l(k_s R)}.$$

L'algorithme de Schwarz avec recouvrement total pour le problème de Maxwell régularisé converge lorsque  $|\tau_{1,l}|$ ,  $|\tau_{2,l}|$  et  $|\tau_{3,l}|$  sont inférieurs à l'unité. On montre le comportement des taux de convergence  $|\tau_{1,l}|$  et  $|\tau_{2,l}|$  dans les figures (3.1), (3.2), (3.3) et (3.4). Il suffit donc de tracer la simulation numérique de  $|\tau_{3,l}|$  pour les mêmes cas tests. L'étude asymptotique des fonctions sphériques  $j_l$  et  $h_l$  pour  $l$  grand et  $x$  fixé donne que

$$\tau_{1,l} \sim \tau_{2,l} \sim \tau_{3,l} \sim \frac{1}{1 - R^{2l}}.$$

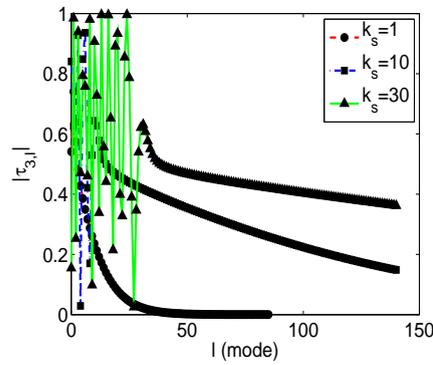


FIGURE 4.1 – Variation de  $|\tau_{3,l}|$  selon les modes  $l$  pour l'épaisseur  $e = \frac{\lambda}{100}$ . Cas  $k_s = 1$ ,  $k_s = 10$  et  $k_s = 30$ .

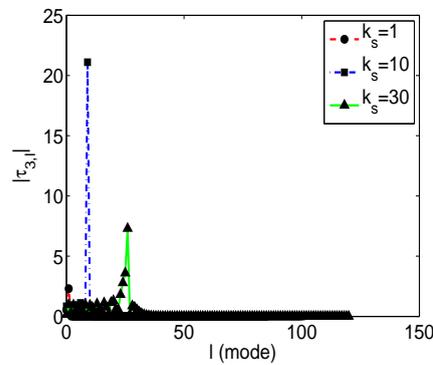


FIGURE 4.2 – Variation du  $|\tau_{3,l}|$  selon les modes  $l$  pour l'épaisseur  $e = \frac{\lambda}{2}$ . Cas  $k_s = 1$ ,  $k_s = 10$  et  $k_s = 30$ .

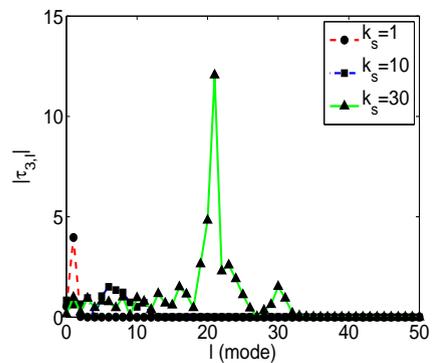


FIGURE 4.3 – Variation de  $|\tau_{3,l}|$  selon les modes  $l$  pour l'épaisseur  $e = \lambda$ . Cas  $k_s = 1$ ,  $k_s = 10$  et  $k_s = 30$ .

On conclut, que l'asymptotique du taux de convergence de la méthode de Schwarz avec recouvrement total pour le problème de Maxwell régularisé est la même que celle trouvée

pour le cas du problème de Maxwell classique.

Les conclusions tirées dans le chapitre précédent concernant la méthode de Schwarz, le choix du préconditionneur et l'utilisation des méthodes de Krylov restent valables pour la résolution du problème de Maxwell régularisé (4.8).

## 4.6 Analyse de la méthode GMRES

On se propose dans cette partie, d'étudier la convergence de la méthode GMRES pour le problème posé sur le domaine  $\Omega$

$$(\mathcal{I} + \mathcal{A}_t^{-1}\mathcal{C}_t)\mathbf{E} = \mathcal{A}_t^{-1}\mathcal{F}_t. \quad (4.21)$$

Pour cela, nous considérons un problème équivalent à celui ci et qui fait intervenir une équation posée sur le bord artificiel  $\Sigma$ . Le taux de convergence du GMRES pour ce problème intermédiaire est le même que celui du problème (4.21). La différence ne réside que dans la taille de stockage de la base d'Arnoldi pour la méthode GMRES.

### 4.6.1 Formulation du problème sur le bord artificiel

Il existe une autre manière de poser le problème (4.8). Elle consiste à se ramener à un problème posé sur le bord  $\Sigma$ . Pour cela, on introduit une nouvelle inconnue  $\Psi$  définie sur  $\Sigma$  comme suit :

$$\Psi(x) = \Lambda(x) + t^{-1} \mathbf{n}_\sigma \lambda(x) \quad \forall x \in \Sigma$$

avec

$$\Lambda = T_{\nu_1}(\mathbf{E}^{inc} - \mathcal{I}_t^{\mathcal{R}}(\Gamma, \mathbf{E}))$$

et

$$\lambda = N_{\nu_2}(\mathbf{E}^{inc} - \mathcal{I}_t^{\mathcal{R}}(\Gamma, \mathbf{E})).$$

Le problème (4.8) devient alors

$$\left\{ \begin{array}{l} \mathbf{rot} \mathbf{rot} \mathbf{E} - t^{-1} \nabla(\operatorname{div} \mathbf{E}) - k_s^2 \mathbf{E} = 0 \text{ dans } \Omega, \\ \mathbf{E} \wedge \mathbf{n} = 0, \operatorname{div} \mathbf{E} = 0 \text{ sur } \Gamma, \\ \mathbf{rot} \mathbf{E} \wedge \mathbf{n}_\sigma + \nu_1 \mathbf{n}_\sigma \wedge (\mathbf{E} \wedge \mathbf{n}_\sigma) = \Lambda \text{ sur } \Sigma, \\ \operatorname{div} \mathbf{E} + \nu_2 \mathbf{E} \cdot \mathbf{n}_\sigma = \lambda \text{ sur } \Sigma. \end{array} \right. \quad (4.22)$$

En adoptant les notations vues précédemment, le problème variationnel (4.9) devient

$$\mathcal{A}_t \mathbf{E} - \mathcal{D}_t \Psi = 0 \quad (4.23)$$

où

$$(\mathcal{D}_t \Psi, \mathbf{E}')_t = \int_{\Sigma} \Psi \cdot \mathbf{E}' d\sigma.$$

Résoudre le problème (4.8) revient à déterminer  $\Psi$  puis trouver l'inconnue  $\mathbf{E}$ .

### 4.6.2 Etude de la convergence du GMRES

On se propose dans cette section d'étudier la convergence théorique de la résolution du problème (4.23). Plus précisément, on s'intéresse à la résolution en  $\Psi$ . L'idée de l'étude est initialement présentée dans [15] pour le problème de Helmholtz. Les auteurs de cet article ont montré une convergence superlinéaire de la méthode GMRES en se basant sur les résultats de la théorie spectrale. Dans notre contexte, on considère un opérateur  $\mathcal{B}_\Gamma^R$  défini par

$$\mathcal{B}_\Gamma^R(\Psi) = T_{\nu_1}(\mathcal{I}_t^R(\Gamma, \mathbf{E})) + t^{-1}N_{\nu_2}(\mathcal{I}_t^R(\Gamma, \mathbf{E}))\mathbf{n}_\sigma \quad (4.24)$$

de sorte que la nouvelle inconnue  $\Psi$  vérifie l'équation suivante

$$\Psi(x) + \mathcal{B}_\Gamma^R(\Psi)(x) = T_{\nu_1}(\mathbf{E}^{inc}) + t^{-1}N_{\nu_2}(\mathbf{E}^{inc})\mathbf{n}_\sigma \quad \forall x \in \Sigma.$$

On s'intéresse donc à résoudre le problème suivant

$$(\mathcal{I} + \mathcal{B}_\Gamma^R)\Psi = \tilde{\mathcal{F}}_t \text{ dans } L^2(\Sigma)^3, \quad (4.25)$$

avec  $\tilde{\mathcal{F}}_t$  donné par le théorème de représentation de Riesz :  $\exists! \tilde{\mathcal{F}}_t \in L^2(\Sigma)^3$  tel que  $\forall \Psi' \in L^2(\Sigma)^3$

$$(\tilde{\mathcal{F}}_t, \Psi')_{L^2(\Sigma)^3} = \int_\Sigma T_{\nu_1}(\mathbf{E}^{inc}) \cdot \Psi' d\sigma + t^{-1} \int_\Sigma N_{\nu_2}(\mathbf{E}^{inc}) (\mathbf{n}_\sigma \cdot \Psi') d\sigma.$$

D'après [13], l'opérateur  $\mathcal{B}_\Gamma^R$  est un opérateur continu de  $L^2(\Sigma)^3$  dans  $\mathbb{H}_t$ . On rappelle que

$$\mathbb{H}_t = \left\{ \mathbf{E} \in H(\mathbf{rot} \mathbf{E}, \Omega) \text{ tel que } \operatorname{div} \mathbf{E} \in L^2(\Omega), (\mathbf{E} \wedge \mathbf{n})_\Gamma = 0, \right. \\ \left. (\mathbf{E} \wedge \mathbf{n})_\Sigma \in L^2(\Sigma)^3 \text{ et } (\mathbf{E} \cdot \mathbf{n})_\Sigma \in L^2(\Sigma) \right\}.$$

De plus, comme l'injection  $\mathbb{H}_t \hookrightarrow L^2(\Omega)^3$  est compacte et l'injection  $L^2(\Omega)^3 \hookrightarrow L^2(\Sigma)^3$  est continue. Donc, l'injection  $\mathbb{H}_t \hookrightarrow L^2(\Sigma)^3$  est compacte. Ceci implique que  $\mathcal{B}_\Gamma^R$  est un opérateur compact sur  $L^2(\Sigma)^3$ . Donc d'après le théorème de Moret ([60]), on a la convergence de GMRES pour résoudre le problème (4.25). Dans la suite on se propose d'étudier le taux de décroissance des valeurs singulières de  $\mathcal{B}_\Gamma^R$  afin de préciser le taux de convergence de la méthode GMRES.

**Proposition 4.6.1.** : *Principe min-max de Courant-Weyl ([61])*

$$\gamma_p = \min_{\mathcal{P}_p} \| \mathcal{B}_\Gamma^R - \mathcal{P}_p \|_{\mathcal{L}(L^2(\Sigma)^3)}$$

$\mathcal{P}_p$  parcourant l'ensemble des opérateurs linéaires sur  $L^2(\Sigma)^3$  de rang  $p$  et  $\{\gamma_p\}_{p \geq 0}$  désigne les valeurs singulières de l'opérateur  $\mathcal{B}_\Gamma^R$ .

La première étape consiste alors à construire l'opérateur  $(\mathcal{B}_\Gamma^R)_p$  qui approche  $\mathcal{B}_\Gamma^R$ . Comme le problème (4.8) admet une unique solution  $\mathbf{E}$  dans  $\mathbb{H}_t$ , alors  $\mathbf{E} \in L^2(\Omega)^3$ . En tenant compte de l'équation de Maxwell régularisée, le terme  $\mathbf{rot} \mathbf{E} \wedge \mathbf{n} + t^{-1}\mathbf{n} \operatorname{div} \mathbf{E}$  appartient à  $H^{-1/2}(\Gamma)^3$ . Dans ce cas, on peut utiliser la formule  $\mathcal{I}_t(\Gamma, \mathbf{E})$  sans passer par la représentation intégrale modifiée  $\mathcal{I}_t^R(\Gamma, \mathbf{E})$ . On rappelle la formule suivante ([8]) :

$$G_{k_s}(x, y) = ik_s \sum_{l=0}^{\infty} \sum_{m=-l}^l h_l(k_s|x|) Y_l^m(\hat{x}) j_l(k_s|y|) \overline{Y_l^m(\hat{y})}. \quad (4.26)$$

La série (4.26) et ses dérivées premières par rapport à  $x$  et  $y$  sont absolument et uniformément convergentes sur tout compact de  $\{(x, y) \in \mathbb{R}^3 \times \mathbb{R}^3 \text{ tel que } |x| > |y|\}$ . D'autre part, on a pour un vecteur fixe  $\mathbf{v}$

$$\begin{aligned} G_{k_s}(x, y)\mathbf{v} &= \frac{i}{k_s} \sum_{l=1}^{\infty} \frac{1}{l(l+1)} \sum_{m=-l}^l \mathbf{rot} \mathbf{N}_{lm}(x) \overline{\mathbf{rot} \mathbf{M}_{lm}(y)} \cdot \mathbf{v} \\ &+ ik_s \sum_{l=1}^{\infty} \frac{ik_s}{l(l+1)} \sum_{m=-l}^l \mathbf{N}_{lm}(x) \overline{\mathbf{M}_{lm}(y)} \cdot \mathbf{v} \\ &+ \frac{i}{k_s} \sum_{l=1}^{\infty} \sum_{m=-l}^l \nabla(h_l(k_s|x|)Y_l^m(\hat{x})) \overline{\nabla(j_l(k_s|y|)Y_l^m(\hat{y}))} \cdot \mathbf{v} \end{aligned} \quad (4.27)$$

où  $\mathbf{M}_{lm}(x) = \mathbf{rot} \{xj_l(k_s|x|)Y_l^m(\hat{x})\}$ ,  $\mathbf{N}_{lm}(x) = \mathbf{rot} \{xh_l(k_s|x|)Y_l^m(\hat{x})\}$  et  $\hat{x} = \frac{x}{|x|}$ .

La série (4.27) et ses dérivées par rapport à  $x$  ou par rapport à  $y$  sont uniformément et absolument convergentes pour tout compact de  $\{(x, y) \in \mathbb{R}^3 \times \mathbb{R}^3 \text{ tel que } |x| > |y|\}$  ([8], [50]).

D'autre part, on a  $\forall x \in \Omega^e$

$$\begin{aligned} \mathcal{I}_t(\Gamma, \mathbf{E})(x) &= \mathbf{rot}_x \mathbf{rot}_x \int_{\Gamma} G_{k_s}(x, y) (\mathbf{rot} \mathbf{E}(y) \wedge \mathbf{n}(y)) d\gamma(y) \\ &- \int_{\Gamma} \nabla_y G_{k_p}(x, y) (\mathbf{E}(y) \cdot \mathbf{n}(y)) d\gamma(y). \end{aligned}$$

Donc,

$$\begin{aligned} \mathcal{I}_t(\Gamma, \mathbf{E})(x) &= \\ &\mathbf{rot}_x \mathbf{rot}_x \int_{\Gamma} \frac{i}{k_s} \sum_{l=1}^{\infty} \frac{1}{l(l+1)} \sum_{m=-l}^l \mathbf{rot} \mathbf{N}_{lm}(x) \overline{\mathbf{rot} \mathbf{M}_{lm}(y)} \cdot (\mathbf{rot} \mathbf{E}(y) \wedge \mathbf{n}(y)) d\gamma(y) \\ &+ \mathbf{rot}_x \mathbf{rot}_x \int_{\Gamma} ik_s \sum_{l=1}^{\infty} \frac{1}{l(l+1)} \sum_{m=-l}^l \mathbf{N}_{lm}(x) \overline{\mathbf{M}_{lm}(y)} \cdot (\mathbf{rot} \mathbf{E}(y) \wedge \mathbf{n}(y)) d\gamma(y) \\ &+ ik_p \int_{\Gamma} \sum_{l=0}^{\infty} \sum_{m=-l}^l j_l(k_p|y|) \overline{Y_l^m(\hat{y})} (k_p \hat{x} h'_l(k_p|x|) Y_l^m(\hat{x}) + h_l(k_p|x|) \frac{1}{|x|} \nabla_S Y_l^m(\hat{x})) (\mathbf{E}(y) \cdot \mathbf{n}(y)) d\gamma(y). \end{aligned}$$

Pour le choix de  $\nu_1 = \nu_2 = -ik_s$ , on a  $\forall x \in \Sigma$

$$\begin{aligned} \mathcal{B}_{\Gamma}^{\mathcal{R}} \Psi(x) &= \int_{\Gamma} \sum_{l=1}^{\infty} \frac{ik_s^3}{l(l+1)} \sum_{m=-l}^l (\mathbf{N}_{lm}(x) \wedge \mathbf{n}_{\sigma}(x)) \overline{\mathbf{rot} \mathbf{M}_{lm}(y)} \cdot (\mathbf{rot} \mathbf{E}(y) \wedge \mathbf{n}(y)) d\gamma(y) \\ &+ \int_{\Gamma} \sum_{l=1}^{\infty} \frac{ik_s^3}{l(l+1)} \sum_{m=-l}^l (\mathbf{rot} \mathbf{N}_{lm}(x) \wedge \mathbf{n}_{\sigma}(x)) \overline{\mathbf{M}_{lm}(y)} \cdot (\mathbf{rot} \mathbf{E}(y) \wedge \mathbf{n}(y)) d\gamma(y) \\ &+ \int_{\Gamma} \sum_{l=1}^{\infty} \frac{k_s^2}{l(l+1)} \sum_{m=-l}^l (\mathbf{n}_{\sigma}(x) \wedge (\mathbf{rot} \mathbf{N}_{lm}(x) \wedge \mathbf{n}_{\sigma}(x))) \overline{\mathbf{rot} \mathbf{M}_{lm}(y)} \cdot (\mathbf{rot} \mathbf{E}(y) \wedge \mathbf{n}(y)) d\gamma(y) \\ &+ \int_{\Gamma} \sum_{l=1}^{\infty} \frac{k_s^4}{l(l+1)} \sum_{m=-l}^l (\mathbf{n}_{\sigma}(x) \wedge (\mathbf{N}_{lm}(x) \wedge \mathbf{n}_{\sigma}(x))) \overline{\mathbf{M}_{lm}(y)} \cdot (\mathbf{rot} \mathbf{E}(y) \wedge \mathbf{n}(y)) d\gamma(y) \end{aligned}$$

$$+ \int_{\Gamma} \frac{k_p^2}{t} \mathbf{n}_{\sigma}(x) \sum_{l=0}^{\infty} \sum_{m=-l}^l j_l(k_p|y|) \overline{Y_l^m(\hat{y})} (ik_p h_l(k_p|x|) + k_s h'_l(k_p|x|)) Y_l^m(\hat{x}) (\mathbf{E}(y) \cdot \mathbf{n}(y)) d\gamma(y).$$

**Proposition 4.6.2.** *On considère un obstacle dont le bord ne présente ni des arêtes ni des coins entrants. Le bord artificiel entourant l'obstacle est sphérique. Il existe alors deux constantes positives  $c$  et  $\tau$  tel que*

$$0 \leq \gamma_p \leq c e^{-\tau\sqrt{p}}, \quad \forall p \geq 0.$$

Preuve :

Soit  $p$  un entier donné et  $p^*$  la partie entière de  $\sqrt{p}$ . On désigne  $(\mathcal{B}_{\Gamma}^{\mathcal{R}})_p$  la troncature de l'opérateur  $\mathcal{B}_{\Gamma}^{\mathcal{R}}$ .

$$\begin{aligned} & \| (\mathcal{B}_{\Gamma}^{\mathcal{R}} - (\mathcal{B}_{\Gamma}^{\mathcal{R}})_p) \Psi \|_{L^2(\Sigma)^3} \leq \\ & \alpha_1 \sum_{l=p^*+1}^{\infty} \frac{k_s^3}{l(l+1)} \sum_{m=-l}^l \| \mathbf{n}_{\sigma} \wedge \mathbf{N}_{lm} \|_{L^2(\Sigma)^3} \| \mathbf{n} \wedge \mathbf{rot} \mathbf{M}_{lm} \|_{L^2(\Gamma)^3} \\ & + \alpha_2 \sum_{l=p^*+1}^{\infty} \frac{k_s^3}{l(l+1)} \sum_{m=-l}^l \| \mathbf{n}_{\sigma} \wedge \mathbf{rot} \mathbf{N}_{lm} \|_{L^2(\Sigma)^3} \| \mathbf{n} \wedge \mathbf{M}_{lm} \|_{L^2(\Gamma)^3} \\ & + \alpha_3 \sum_{l=p^*+1}^{\infty} \frac{k_s^2}{l(l+1)} \sum_{m=-l}^l \| \mathbf{n}_{\sigma} \wedge (\mathbf{rot} \mathbf{N}_{lm} \wedge \mathbf{n}_{\sigma}) \|_{L^2(\Sigma)^3} \| \mathbf{n} \wedge \mathbf{rot} \mathbf{M}_{lm} \|_{L^2(\Gamma)^3} \\ & + \alpha_4 \sum_{l=p^*+1}^{\infty} \frac{k_s^4}{l(l+1)} \sum_{m=-l}^l \| \mathbf{n}_{\sigma} \wedge (\mathbf{N}_{lm} \wedge \mathbf{n}_{\sigma}) \|_{L^2(\Sigma)^3} \| \mathbf{n} \wedge \mathbf{M}_{lm} \|_{L^2(\Gamma)^3} \\ & + \alpha_5 \sum_{l=p^*+1}^{\infty} \frac{k_p^2}{t} \sum_{m=-l}^l \| (ik_p h_l(k_p|\cdot|) + k_s h'_l(k_p|\cdot|)) Y_l^m(\cdot) \|_{L^2(\Sigma)^3} \| j_l(k_p|\cdot|) Y_l^m(\cdot) \|_{L^2(\Gamma)^3}. \end{aligned} \quad (4.28)$$

**Remarque 4.6.1.** *Les  $\alpha_i$  pour  $i = 1..4$  dépendent de  $\| \mathbf{rot} \mathbf{E} \|_{H^{1/2}(\Gamma)^3}$  et  $\alpha_5$  dépend de  $\| \mathbf{E} \cdot \mathbf{n} \|_{H^{-1/2}(\Gamma)}$ .*

Nous précisons dans ce qui suit comment nous avons obtenu le résultat précédent par un calcul d'un terme présent dans l'expression de  $\| (\mathcal{B}_{\Gamma}^{\mathcal{R}} - (\mathcal{B}_{\Gamma}^{\mathcal{R}})_p) \Psi \|_{L^2(\Sigma)^3}$ . Les autres termes sont traités de la même façon.

$$\begin{aligned} & \left\| \sum_{l=p^*+1}^{\infty} \frac{ik_s^3}{l(l+1)} \sum_{m=-l}^l (\mathbf{n}_{\sigma} \wedge \mathbf{N}_{lm}) \int_{\Gamma} \overline{\mathbf{rot} \mathbf{M}_{lm}(y)} \cdot (\mathbf{rot} \mathbf{E}(y) \wedge \mathbf{n}(y)) d\gamma(y) \right\|_{L^2(\Sigma)^3} \\ & \leq c \sum_{l=p^*+1}^{\infty} \frac{k_s^3}{l(l+1)} \sum_{m=-l}^l \| \mathbf{n}_{\sigma} \wedge \mathbf{N}_{lm} \|_{L^2(\Sigma)^3} \left| \int_{\Gamma} (\mathbf{n}(y) \wedge \overline{\mathbf{rot} \mathbf{M}_{lm}(y)}) \cdot \mathbf{rot} \mathbf{E}(y) d\gamma(y) \right|. \end{aligned} \quad (4.29)$$

On rappelle que par abus de notation, on a gardé l'intégrale sur  $\Gamma$  plutôt que les crochets de dualité entre  $H^{-1/2}(\Gamma)^3$  et  $H^{1/2}(\Gamma)^3$ . En utilisant l'inégalité de Cauchy-Schwarz et le fait que l'injection

$$L^2(\Gamma)^3 \hookrightarrow H^{-1/2}(\Gamma)^3$$

est continue, on alors

$$\begin{aligned} & \left\| \sum_{l=p^*+1}^{\infty} \frac{ik_s^3}{l(l+1)} \sum_{m=-l}^l (\mathbf{n}_\sigma \wedge \mathbf{N}_{lm}) \int_{\Gamma} \overline{\mathbf{rot} \mathbf{M}_{lm}(y)} \cdot (\mathbf{rot} \mathbf{E}(y) \wedge \mathbf{n}(y)) d\gamma(y) \right\|_{L^2(\Sigma)^3} \\ & \leq \alpha_1 \sum_{l=p^*+1}^{\infty} \frac{k_s^3}{l(l+1)} \sum_{m=-l}^l \left\| \mathbf{n}_\sigma \wedge \mathbf{N}_{lm} \right\|_{L^2(\Sigma)^3} \left\| \mathbf{n} \wedge \mathbf{rot} \mathbf{M}_{lm} \right\|_{L^2(\Gamma)^3} . \end{aligned}$$

L'expression (4.28) devient

$$\begin{aligned} & \left\| (\mathcal{B}_{\Gamma}^R - (\mathcal{B}_{\Gamma}^R)_p) \Psi \right\|_{L^2(\Sigma)^3} \leq \\ & \alpha_1 \sum_{l=p^*+1}^{\infty} \frac{k_s^3}{l(l+1)} \sum_{m=-l}^l \sqrt{l(l+1)} |h_l(k_s R) | R l(l+1) | j_l(k_s) + k j_l'(k_s) | \\ & + \alpha_2 \sum_{l=p^*+1}^{\infty} \frac{k_s^3}{l(l+1)} \sum_{m=-l}^l \sqrt{l(l+1)} |h_l(k_s R) + k_s R h_l'(k_s R) | l(l+1) | j_l(k_s) | \\ & + \alpha_3 \sum_{l=p^*+1}^{\infty} \frac{k_s^2}{l(l+1)} \sum_{m=-l}^l \sqrt{l(l+1)} |h_l(k_s R) + k_s R h_l'(k_s R) | l(l+1) | j_l(k_s) + k_s j_l'(k_s) | \\ & + \alpha_4 \sum_{l=p^*+1}^{\infty} \frac{k_s^4}{l(l+1)} \sum_{m=-l}^l \sqrt{l(l+1)} |h_l(k_s R) | R l(l+1) | j_l(k_s) | \\ & + \alpha_5 \sum_{l=p^*+1}^{\infty} \frac{k_p^2}{t} \sum_{m=-l}^l |i k_p h_l(k_p R) + k_s h_l'(k_p R) | R | j_l(k_p) | . \end{aligned}$$

Le comportement asymptotique des fonctions spéciales pour un  $l$  qui tend vers l'infini est :

$$|h_l(kR)| \leq c \frac{(2l-1)!!}{(kR)^{l+1}},$$

$$|h_l'(kR)| \leq c (l+1) \frac{(2l-1)!!}{(kR)^{l+2}},$$

$$|h_l(kR) + kR h_l'(kR)| \leq c (l+1) \frac{(2l-1)!!}{(kR)^{l+1}},$$

$$|j_l(k)| \leq c \frac{k^l}{(2l+1)!!},$$

$$|j_l'(k)| \leq c l \frac{k^{l-1}}{(2l+1)!!},$$

$$|j_l(k) + k j_l'(k)| \leq c l \frac{k^l}{(2l+1)!!}.$$

Il en résulte que

$$\begin{aligned} \| (\mathcal{B}_\Gamma^{\mathcal{R}} - (\mathcal{B}_\Gamma^{\mathcal{R}})_p) \Psi \|_{L^2(\Sigma)^3} \leq & c_1 \sum_{l=p^*+1}^{\infty} l \sqrt{l(l+1)} \frac{k_s^2}{R^l} + c_2 \sum_{l=p^*+1}^{\infty} \sqrt{l(l+1)}^{\frac{3}{2}} \frac{k_s^2}{R^{l+1}} \\ & + c_3 \sum_{l=p^*+1}^{\infty} (l(l+1))^{\frac{3}{2}} \frac{k_s}{R^{l+1}} + c_4 \sum_{l=p^*+1}^{\infty} \sqrt{l(l+1)} \frac{k_s^3}{R^l} \\ & + c_5 \sum_{l=p^*+1}^{\infty} (l+1) \frac{k_s}{R^{l+1}}, \end{aligned}$$

avec  $c_i$  depend de  $\alpha_i$  pour  $i = 1 \dots 5$  et les constantes induites par le comportement asymptotique des fonctions spéciales. D'où

$$\| (\mathcal{B}_\Gamma^{\mathcal{R}} - (\mathcal{B}_\Gamma^{\mathcal{R}})_p) \Psi \|_{L^2(\Sigma)^3} \leq \alpha e^{-\tau(p^*+1)} \leq c e^{-\tau\sqrt{p}},$$

avec  $\alpha, c$  et  $\tau$  sont des constantes positives.

**Proposition 4.6.3.** *La convergence du GMRES est superlinéaire.*

Preuve : On désigne par  $r_m$  le  $m^{\text{ième}}$  résidu et  $\{\mu_p\}_{p \geq 0}$  les valeurs singulières de  $(\mathcal{K})^{-1}$  avec

$$\mathcal{K} = \mathcal{I} + \mathcal{B}_\Gamma^{\mathcal{R}}.$$

$\mathcal{K}$  est continu et bijectif. On a alors  $\mathcal{K}^{-1}$  continu. Les valeurs singulières  $\mu_p$  sont bornées. D'après le théorème 1 de [60], on a

$$\| r_m \|_{L^2(\Sigma)^3} \leq \left( \prod_{p=1}^m \gamma_p \mu_p \right) \| \tilde{\mathcal{F}}_t \|_{L^2(\Sigma)^3}.$$

De la proposition 4.6.2, on a

$$\gamma_p \leq \frac{\alpha}{p^\tau} \quad \forall p \geq 1,$$

où  $\alpha$  est une constante positive. Par suite

$$\left( \| r_m \|_{L^2(\Sigma)^3} \right)^{\frac{1}{m}} \leq \left( \beta \frac{\alpha^m}{m!^\tau} \right)^{\frac{1}{m}}$$

avec  $\beta$  est une constante positive dépendant de  $\| \tilde{\mathcal{F}}_t \|_{L^2(\Sigma)^3}$ . En approchant  $m!$  par la formule de Stirling on a

$$\left( \| r_m \|_{L^2(\Sigma)^3} \right)^{\frac{1}{m}} \leq \frac{c}{m^\tau}; \quad \forall m \geq 0.$$

Par conséquent,

$$\lim_{m \rightarrow \infty} \left( \| r_m \|_{L^2(\Sigma)^3} \right)^{\frac{1}{m}} = 0$$

Cela justifie une convergence superlinéaire de la méthode GMRES.

**Remarque 4.6.2.** *Dans notre étude, on considère  $\Gamma$  et  $\Sigma$  deux sphères concentriques :  $\Gamma$  est de rayon  $l$  et  $\Sigma$  est de rayon  $R$ . Si le bord de l'obstacle est non sphérique et ne présente pas de singularités géométriques entrantes (coins, arêtes), on peut considérer une frontière intermédiaire  $\tilde{\Gamma}$  sphérique de rayon  $\tilde{R}$  entre les deux frontières  $\Gamma$  et  $\Sigma$  sur laquelle on écrit la représentation intégrale et dans ce cas on peut mener la même démonstration.*

**Remarque 4.6.3.** *Nous pouvons montrer d'une manière similaire que la convergence du GMRES est superlinéaire pour le problème de Maxwell classique.*

# Chapitre 5

## Discrétisation et mise en œuvre numérique avec MÉLINA++

### 5.1 Approximation par éléments finis de Lagrange

La méthode de couplage entre éléments finis et représentation intégrale permet de formuler le problème extérieur en un problème posé sur un domaine borné ce qui donne la possibilité d'utiliser une approximation par éléments finis du système d'équations obtenu. Nous détaillerons dans cette partie la discrétisation du problème obtenu pour définir les matrices intervenant dans le système linéaire à résoudre. Le principe de la méthode des éléments finis est de décomposer  $\Omega$  en une partition, appelée maillage ou triangulation  $\mathcal{T}_h$  et de chercher une approximation de la solution dans un sous-espace de dimension finie  $V_h$  constitué des fonctions polynomiales par morceaux sur chaque élément  $K$  de  $\mathcal{T}_h$ . Deux éléments quelconques de la triangulation  $\mathcal{T}_h$  ont soit une intersection vide, soit un sommet commun, soit une arête commune ou une surface commune. Les éléments de  $\mathcal{T}_h$  sont de diamètre inférieur ou égal à  $h$ . Le paramètre  $h$  est un paramètre destiné à tendre vers 0.  $h$  est choisi en fonction de la précision souhaitée et en cohérence avec les paramètres physiques.

Dans notre cas, nous considérons une décomposition éléments finis  $\mathcal{T}_h$  du domaine borné  $\Omega$  délimité par  $\Gamma$  et  $\Sigma$ . Notons  $V_h$  l'espace des éléments finis de dimension  $N$  défini par

$$V_h = \left\{ v_h \in C^0(\bar{\Omega})^3; v_h|_K \in P_m(K)^3, \forall K \in \mathcal{T}_h \right\}$$

$P_m$  désigne l'ensemble des polynômes de degré total  $m \geq 1$  pour les tétraèdres et de degré  $m \geq 1$  par rapport à chaque variable d'espace pour les cubes et les rectangles. On désigne par  $N$  le nombre total de degrés de liberté. La base de l'espace  $V_h$  est formée par les fonctions vectorielles  $w_j \mathbf{e}_k$ ,  $j$  variant de 1 à  $N$  et  $k$  de 1 à 3. Le choix de travailler avec une base scalaire est lié au spécificité de la librairie MELINA++. Nous nous intéressons à chercher  $\mathbf{E}_h$  aux points du maillage défini comme suit :

$$\mathbf{E}_h = \sum_{j=1}^N \mathbf{E}_j w_j,$$

où  $\mathbf{E}_j = E_j^1 \mathbf{e}_1 + E_j^2 \mathbf{e}_2 + E_j^3 \mathbf{e}_3$  avec  $\mathbf{e}_k$  le  $k^{\text{ième}}$  vecteur de la base canonique de  $\mathbb{R}^3$ . On approche la solution de la formulation variationnelle (4.9) dans l'espace discret. Cela nous ramène à résoudre le problème suivant

$$\begin{cases} \text{Trouver } \mathbf{E}_h \in V_h \\ (\mathcal{A}_t \mathbf{E}_h, \mathbf{E}'_h)_t + (\mathcal{C}_t \mathbf{E}_h, \mathbf{E}'_h)_t = (F_t, \mathbf{E}'_h)_t \quad \forall \mathbf{E}'_h \in V_h. \end{cases} \quad (5.1)$$

En décomposant  $\mathbf{E}_h$  dans la formulation variationnelle (5.1) et en choisissant  $\mathbf{E}'_h = w_i \mathbf{e}_k$ ,  $\forall k \in \{1, \dots, 3\}$ ,  $\forall i \in \{1, \dots, N\}$  il vient que :

$$\begin{aligned} & \forall i \in \{1, \dots, N\}, \forall k \in \{1, \dots, 3\}, \\ & \sum_{j=1}^N \sum_{l=1}^3 ((\mathcal{A}_t w_j \mathbf{e}_l, w_i \mathbf{e}_k)_t + (\mathcal{C}_t w_j \mathbf{e}_l, w_i \mathbf{e}_k)_t) E_j^l = (F_t, w_i \mathbf{e}_k)_t \end{aligned}$$

ce qui est équivalent à écrire le système linéaire suivant :

$$(\bar{\bar{A}} + \bar{\bar{C}}) \bar{\bar{\mathbf{E}}} = \bar{\bar{\mathbf{F}}} \quad (5.2)$$

où  $\bar{\bar{A}}$  (resp.  $\bar{\bar{C}}$ ) est définie par blocs  $A_{ij}$  (resp.  $C_{ij}$ )  $\in \mathbb{C}^3$ ,  $i = 1 \dots N$  et  $j = 1 \dots N$  tels que le bloc

$$A_{ij} = \begin{pmatrix} (\mathcal{A}_t w_j \mathbf{e}_1, w_i \mathbf{e}_1)_t & (\mathcal{A}_t w_j \mathbf{e}_2, w_i \mathbf{e}_1)_t & (\mathcal{A}_t w_j \mathbf{e}_3, w_i \mathbf{e}_1)_t \\ (\mathcal{A}_t w_j \mathbf{e}_1, w_i \mathbf{e}_2)_t & (\mathcal{A}_t w_j \mathbf{e}_2, w_i \mathbf{e}_2)_t & (\mathcal{A}_t w_j \mathbf{e}_3, w_i \mathbf{e}_2)_t \\ (\mathcal{A}_t w_j \mathbf{e}_1, w_i \mathbf{e}_3)_t & (\mathcal{A}_t w_j \mathbf{e}_2, w_i \mathbf{e}_3)_t & (\mathcal{A}_t w_j \mathbf{e}_3, w_i \mathbf{e}_3)_t \end{pmatrix}$$

et le bloc

$$C_{ij} = \begin{pmatrix} (\mathcal{C}_t w_j \mathbf{e}_1, w_i \mathbf{e}_1)_t & (\mathcal{C}_t w_j \mathbf{e}_2, w_i \mathbf{e}_1)_t & (\mathcal{C}_t w_j \mathbf{e}_3, w_i \mathbf{e}_1)_t \\ (\mathcal{C}_t w_j \mathbf{e}_1, w_i \mathbf{e}_2)_t & (\mathcal{C}_t w_j \mathbf{e}_2, w_i \mathbf{e}_2)_t & (\mathcal{C}_t w_j \mathbf{e}_3, w_i \mathbf{e}_2)_t \\ (\mathcal{C}_t w_j \mathbf{e}_1, w_i \mathbf{e}_3)_t & (\mathcal{C}_t w_j \mathbf{e}_2, w_i \mathbf{e}_3)_t & (\mathcal{C}_t w_j \mathbf{e}_3, w_i \mathbf{e}_3)_t \end{pmatrix}.$$

$\bar{\bar{\mathbf{E}}} = (\mathbf{E}_i)_{i=1, \dots, N}$  est l'inconnu vectorielle : pour chaque  $i$ ,  $\mathbf{E}_i \in \mathbb{C}^3$  et  $\bar{\bar{\mathbf{F}}} = (\mathbf{F}_i)_{i=1, \dots, N}$  tel que  $\mathbf{F}_i$  le vecteur appartenant à  $\mathbb{C}^3$  et composé des valeurs  $(F_t, w_i \mathbf{e}_k)_t$  pour  $k = 1, 2, 3$ . Si on choisit  $\bar{\bar{A}}$  comme préconditionneur, on obtient le système :

$$(\bar{\bar{I}} + \bar{\bar{A}}^{-1} \bar{\bar{C}}) \bar{\bar{\mathbf{E}}} = \bar{\bar{A}}^{-1} \bar{\bar{\mathbf{F}}}. \quad (5.3)$$

## 5.2 Tests de validation du code CEFRI Maxwell

Nous présentons des résultats numériques qui ont été obtenus à l'aide de la bibliothèque MÉLINA++. Nous nous intéressons à la résolution numérique en 3D du problème (5.3). On se place dans le cas où l'obstacle  $\Omega_i$  est une boule unité  $B_1$  et on choisit  $\Sigma$  la frontière de la boule concentrique à  $B_1$  et de rayon  $R$ . Vu la complexité de l'implémentation numérique du problème (5.3) du point de vue de la prise en compte de la condition essentielle et du choix numérique d'un relèvement. Nous avons choisi de commencer par

la résolution de quelques problèmes plus simples afin de valider les différents outils programmés dans MÉLINA++. Dans un premier temps, nous nous intéressons à l'étude du problème avec une condition de type Neumann. Dans ce cas, nous n'avons pas besoin de la pénalisation et du relèvement, ce qui permet de valider quelques intégrands éléments finis et des noyaux de Green sans relèvement que nous avons intégrés dans MÉLINA++. L'étape suivante consiste à tester le cas d'un problème avec une condition essentielle. Les maillages sont construits avec le logiciel GMSH [62]. Dans les tableaux 5.1 et 5.2, nous présentons le nombre de tétraèdres en fonction du pas de discrétisation. Nous travaillons avec des éléments finis P1. Les tests numériques sont effectués avec  $\nu_1 = \nu_2 = -ik_s$  et un paramètre de régularisation  $t = 1$ . Nous considérons un point source artificiel noté  $x_0$  appartenant à la boule  $B_1$  et égal à  $(0, 0, 0.01)$ . On désigne par  $\mathbf{E}_e$  (resp.  $\mathbf{E}_c$ ) la solution exacte (resp. calculée) du problème étudié dans la suite. Nous nous proposons d'étudier le comportement du résidu du solveur itératif et l'erreur relative entre  $\mathbf{E}_e$  et  $\mathbf{E}_c$ . De plus, une analyse est faite, pour chaque cas test, concernant la vérification de  $\mathbf{E}_e$  dans le système linéaire à résoudre. On note par  $\mathbf{F}$  le second membre du système linéaire à résoudre.

| $h$                  | 0.3  | 0.2  | 0.15 | 0.12 | 0.1  | 0.09  | 0.08  | 0.07  | 0.06  |
|----------------------|------|------|------|------|------|-------|-------|-------|-------|
| Nombre de tétraèdres | 4142 | 3832 | 4482 | 7046 | 9778 | 13006 | 18786 | 33890 | 48472 |

TABLE 5.1 – Cas  $R = 1.1$ .

| $h$                  | 0.3  | 0.2  | 0.15  | 0.12  |
|----------------------|------|------|-------|-------|
| Nombre de tétraèdres | 2787 | 8823 | 20331 | 45083 |

TABLE 5.2 – Cas  $R = 1.5$ .

### 5.2.1 Problème de Maxwell avec une condition de Neumann

Le critère d'arrêt du solveur itératif correspond à un résidu inférieur à  $10^{-6}$  et le choix du nombre d'onde à  $k_s = 6$ . Nous travaillons dans cette section avec les maillages du tableau 5.1.

#### Cas sans la représentation intégrale

Nous nous intéressons à la résolution d'un problème sans la représentation intégrale. Ce problème est artificiel et sans signification physique. L'intérêt est la validation des outils intégrés dans la librairie MÉLINA++. Nous désignons par  $\mathcal{G}_t^1$  la première colonne de la fonction de Green dyadique  $\mathcal{G}_t = G_{k_s}I + \frac{1}{k_s^2}Hess(G_{k_s} - G_{k_p})$ . Nous considérons

désormais le problème suivant :

$$\left\{ \begin{array}{l} \mathbf{rot} \mathbf{rot} \mathbf{E} - t^{-1} \nabla(\operatorname{div} \mathbf{E}) - k_s^2 \mathbf{E} = 0 \text{ dans } \Omega, \\ \mathbf{n} \wedge \mathbf{rot} \mathbf{E} = \mathbf{n} \wedge \mathbf{rot} \mathcal{G}_t^1(\cdot, x_0) \text{ sur } \Gamma, \\ \operatorname{div} \mathbf{E} = \operatorname{div} \mathcal{G}_t^1(\cdot, x_0) \text{ sur } \Gamma, \\ T_{\nu_1}(\mathbf{E}) = T_{\nu_1}(\mathcal{G}_t^1(\cdot, x_0)) \text{ sur } \Sigma, \\ N_{\nu_2}(\mathbf{E}) = N_{\nu_2}(\mathcal{G}_t^1(\cdot, x_0)) \text{ sur } \Sigma. \end{array} \right. \quad (5.4)$$

Les opérateurs  $T_{\nu_1}$  et  $N_{\nu_2}$  sont définis par

$$T_{\nu_1} \mathbf{E} = \mathbf{rot} \mathbf{E} \wedge \mathbf{n}_\sigma + \nu_1 \mathbf{n}_\sigma \wedge (\mathbf{E} \wedge \mathbf{n}_\sigma) \text{ et } N_{\nu_2} \mathbf{E} = \operatorname{div} \mathbf{E} + \nu_2 \mathbf{E} \cdot \mathbf{n}_\sigma.$$

Pour ce problème,  $\mathcal{G}_t^1$  est la solution exacte.

| $h$  | Résidu      | $\frac{\ \bar{\bar{A}}\mathbf{E}_e - \mathbf{F}\ _2}{\ \mathbf{F}\ _2}$ | $\frac{\ \mathbf{E}_c - \mathbf{E}_e\ _2}{\ \mathbf{E}_e\ _2}$ | Nombre<br>d'itérations |
|------|-------------|---|--|------------------------|
| 0.15 | $7.10^{-7}$ | 0.087   | 0.017  | 94                     |
| 0.12 | $9.10^{-7}$ | 0.052   | 0.016  | 79                     |
| 0.1  | $8.10^{-7}$ | 0.06  | 0.015  | 110                    |
| 0.09 | $9.10^{-7}$ | 0.1   | 0.015  | 120                    |
| 0.08 | $9.10^{-7}$ | 0.21  | 0.016  | 172                    |
| 0.07 | $8.10^{-7}$ | 0.38  | 0.012  | 182                    |
| 0.06 | $9.10^{-7}$ | 0.075   | 0.006  | 184                    |

TABLE 5.3 – Résolution par la méthode du gradient conjugué.

Au regard du comportement du résidu dans le tableau 5.3,  $\frac{\|\bar{\bar{A}}\mathbf{E}_c - \mathbf{F}\|_2}{\|\mathbf{F}\|_2}$ , nous constatons que le gradient conjugué converge. A partir de l'étude du terme  $\frac{\|\bar{\bar{A}}\mathbf{E}_e - \mathbf{F}\|_2}{\|\mathbf{F}\|_2}$ , nous constatons que la solution exacte vérifie le système linéaire considéré. Du point de vue précision, nous remarquons que l'erreur relative décroît en choisissant un pas de maillage de plus en plus petit.

### Cas avec la représentation intégrale

Nous considérons désormais le problème de Neumann suivant :

$$\left\{ \begin{array}{l} \mathbf{rot} \mathbf{rot} \mathbf{E} - t^{-1} \nabla(\operatorname{div} \mathbf{E}) - k_s^2 \mathbf{E} = 0 \text{ dans } \Omega, \\ \mathbf{n} \wedge \mathbf{rot} \mathbf{E} = \mathbf{n} \wedge \mathbf{rot} \mathcal{G}_t^1(\cdot, x_0) \text{ sur } \Gamma, \\ \operatorname{div} \mathbf{E} = \operatorname{div} \mathcal{G}_t^1(\cdot, x_0) \text{ sur } \Gamma, \\ T_{\nu_1}(\mathbf{E}) = -T_{\nu_1}(\mathcal{I}_t(\Gamma, \mathbf{E})) \text{ sur } \Sigma, \\ N_{\nu_2}(\mathbf{E}) = -N_{\nu_2}(\mathcal{I}_t(\Gamma, \mathbf{E})) \text{ sur } \Sigma. \end{array} \right. \quad (5.5)$$

La solution exacte  $\mathbf{E}_e$  du problème (5.5) est  $\mathcal{G}_t^1(\cdot, x_0)$ . Afin de valider la stratégie **CE-FRI**, nous menons l'étude faite dans le tableau (5.4). En étudiant le comportement du

| $h$  | Résidu      | $\frac{\ (I + \bar{A}^{-1} \bar{C}) \mathbf{E}_e - \bar{A}^{-1} \mathbf{F}\ _2}{\ \bar{A}^{-1} \mathbf{F}\ _2}$ | $\frac{\ \mathbf{E}_c - \mathbf{E}_e\ _2}{\ \mathbf{E}_e\ _2}$ | Nombre<br>d'itérations |
|------|-------------|---|--|------------------------|
| 0.15 | $7.10^{-7}$ | 0.09  | 0.09   | 22                     |
| 0.12 | $5.10^{-7}$ | 0.07  | 0.06   | 17                     |
| 0.1  | $6.10^{-7}$ | 0.06  | 0.06   | 18                     |
| 0.09 | $9.10^{-7}$ | 0.06  | 0.058  | 20                     |
| 0.08 | $4.10^{-7}$ | 0.08  | 0.086  | 19                     |
| 0.07 | $9.10^{-7}$ | 0.05  | 0.054  | 18                     |
| 0.06 | $9.10^{-7}$ | 0.05  | 0.054  | 21                     |

TABLE 5.4 – Résolution par la méthode du gradient biconjugué stabilisé.

résidu qui est égal à  $\frac{\|(I + \bar{A}^{-1} \bar{C}) \mathbf{E}_e - \bar{A}^{-1} \mathbf{F}\|_2}{\|\bar{A}^{-1} \mathbf{F}\|_2}$ , nous constatons qu'on a une bonne convergence du solveur itératif. A travers la troisième colonne du tableau (5.4), nous remarquons que la solution exacte vérifie le problème discrétisé. De plus, la quantité  $\frac{\|(I + \bar{A}^{-1} \bar{C}) \mathbf{E}_e - \bar{A}^{-1} \mathbf{F}\|_2}{\|\bar{A}^{-1} \mathbf{F}\|_2}$  devient de plus en plus petite en raffinant de plus en plus le maillage. Pour un nombre d'onde  $k_s$  fixé et en raffinant le maillage, nous remarquons que l'erreur relative entre la solution calculée et la solution exacte devient de plus en plus petite.

### 5.2.2 Problème de Maxwell avec une condition de Dirichlet

Nous nous proposons d'établir une étude numérique avec des maillages du tableau (5.2). Ce choix est fait à cause de la présence de la stratégie du relèvement dans les termes intégraux.

#### Cas sans la représentation intégrale

Nous considérons le problème suivant :

$$\left\{ \begin{array}{l} \text{rot rot } \mathbf{E} - t^{-1} \nabla(\text{div } \mathbf{E}) - k_s^2 \mathbf{E} = 0 \text{ dans } \Omega, \\ \mathbf{E} \wedge \mathbf{n} = \mathcal{G}_t^1(\cdot, x_0) \wedge \mathbf{n} \text{ sur } \Gamma, \\ \text{div } \mathbf{E} = \text{div } \mathcal{G}_t^1(\cdot, x_0) \text{ sur } \Gamma, \\ T_{\nu_1}(\mathbf{E}) = T_{\nu_1}(\mathcal{G}_t^1(\cdot, x_0)) \text{ sur } \Sigma, \\ N_{\nu_2}(\mathbf{E}) = N_{\nu_2}(\mathcal{G}_t^1(\cdot, x_0)) \text{ sur } \Sigma. \end{array} \right. \quad (5.6)$$

Pour des raisons pratiques de mise en œuvre numérique dans MELINA++, la prise en compte de la condition essentielle sur  $\Gamma$

$$\mathbf{E} \wedge \mathbf{n} = \mathcal{G}_t^1(\cdot, x_0) \wedge \mathbf{n} \quad (5.7)$$

se fait par la stratégie de pénalisation qui consiste à remplacer la condition (5.7) par

$$\varepsilon_p (\mathbf{n} \wedge \text{rot } \mathbf{E}) + (\mathbf{E} - \mathcal{G}_t^1(\cdot, x_0)) \wedge \mathbf{n} = 0,$$

où  $\varepsilon_p$  est choisi petit. Nous menons une étude du choix optimal du paramètre de pénalisation  $\varepsilon_p$ . Nous choisissons un nombre d'onde égal à 6, le test d'arrêt du solveur est  $10^{-6}$  et le nombre d'itérations maximum est 200 itérations. Nous désignons par  $\frac{\|\bar{\bar{A}}\mathbf{E}_c - \mathbf{F}\|_2}{\|\mathbf{F}\|_2}$  le résidu de la résolution itérative.

| $h$  | Résidu | $\frac{\ \bar{\bar{A}}\mathbf{E}_c - \mathbf{F}\ _2}{\ \mathbf{F}\ _2}$ | Nombre d'itérations |
|------|--------|---|---------------------|
| 0.3  | 0.25   | 0.08  | 200                 |
| 0.2  | 0.08   | 0.09  | 200                 |
| 0.15 | 0.06   | 0.08  | 200                 |
| 0.12 | 0.04   | 0.07  | 200                 |

TABLE 5.5 – Résolution par la méthode du gradient biconjugué stabilisé, cas  $\varepsilon_p = 10^{-2}$ .

| $h$  | Résidu | $\frac{\ \bar{\bar{A}}\mathbf{E}_e - \mathbf{F}\ _2}{\ \mathbf{F}\ _2}$ | Nombre d'itérations |
|------|--------|---|---------------------|
| 0.3  | 0.63   | $8.10^{-3}$   | 200                 |
| 0.2  | 1.29   | $9.10^{-3}$   | 200                 |
| 0.15 | 0.1    | $8.10^{-3}$   | 200                 |
| 0.12 | 0.36   | $7.10^{-3}$   | 200                 |

TABLE 5.6 – Résolution par la méthode du gradient biconjugué stabilisé, cas  $\varepsilon_p = 10^{-3}$ .

| $h$  | Résidu | $\frac{\ \bar{\bar{A}}\mathbf{E}_e - \mathbf{F}\ _2}{\ \mathbf{F}\ _2}$ | Nombre d'itérations |
|------|--------|---|---------------------|
| 0.3  | 267    | $8.10^{-4}$   | 200                 |
| 0.2  | 8.4    | $9.10^{-4}$   | 200                 |
| 0.15 | 3.25   | $8.10^{-4}$   | 200                 |
| 0.12 | 2.42   | $7.10^{-4}$   | 200                 |

TABLE 5.7 – Résolution par la méthode du gradient biconjugué stabilisé, cas  $\varepsilon_p = 10^{-4}$ .

D'après les tableaux 5.5, 5.6 et 5.7, nous pouvons tirer plusieurs conclusions :

- Vérification de la solution exacte : nous constatons que  $\mathbf{E}_e$  vérifie le système linéaire. Nous remarquons que la quantité  $\frac{\|\bar{\bar{A}}\mathbf{E}_e - \mathbf{F}\|_2}{\|\mathbf{F}\|_2}$  devient de plus en plus petite lorsque  $\varepsilon_p$  décroît. Par exemple, pour le cas  $h = 0.12$  si nous choisissons  $\varepsilon_p = 10^{-2}$  alors  $\frac{\|\bar{\bar{A}}\mathbf{E}_e - \mathbf{F}\|_2}{\|\mathbf{F}\|_2}$  est égale à 0.07. Par contre, cette quantité devient plus petite pour un choix de  $\varepsilon_p = 10^{-4}$ . Cela valide la stratégie de la pénalisation qui doit mieux fonctionner pour le choix de  $\varepsilon_p$  qui tend vers 0.
- Convergence du solveur itératif : nous observons que le résidu devient de plus en plus grand en choisissant  $\varepsilon_p$  de plus en plus petit. Ainsi, nous constatons l'effet de la stratégie de la pénalisation sur la résolution numérique du problème. Ce résultat est intéressant : il montre la nécessité d'établir une stratégie de préconditionnement relative au procédé de pénalisation pour résoudre le problème avec la représentation

intégrale.

### Cas avec la représentation intégrale

Nous considérons le problème de Dirichlet suivant :

$$\left\{ \begin{array}{l} \text{rot rot } \mathbf{E} - t^{-1} \nabla(\text{div } \mathbf{E}) - k_s^2 \mathbf{E} = 0 \text{ dans } \Omega, \\ \mathbf{E} \wedge \mathbf{n} = \mathcal{G}_t^1(\cdot, x_0) \wedge \mathbf{n} \text{ sur } \Gamma, \\ \text{div } \mathbf{E} = \text{div } \mathcal{G}_t^1(\cdot, x_0) \text{ sur } \Gamma, \\ T_{\nu_1}(\mathbf{E}) = -T_{\nu_1}(\mathcal{I}_t^R(\Gamma, \mathbf{E})) \text{ sur } \Sigma, \\ N_{\nu_2}(\mathbf{E}) = -N_{\nu_2}(\mathcal{I}_t^R(\Gamma, \mathbf{E})) \text{ sur } \Sigma. \end{array} \right. \quad (5.8)$$

Nous nous intéressons dans un premier temps à la validation numérique du code **CE-FRI** en fixant le nombre d'onde à  $k_s = 3$ . Le test d'arrêt du solveur est  $10^{-6}$ . Nous étudions le comportement du résidu ainsi que la précision pour différents choix du paramètre de pénalisation (cas  $\varepsilon_p = 10^{-3}, 10^{-4}, 10^{-5}$  et  $10^{-6}$ ). L'étude de l'erreur relative pondérée

| $h$  | Résidu      | $\ (\bar{\bar{A}} + \bar{\bar{C}})\mathbf{E}_e - \mathbf{F}\ _2$ | $\ (I + \bar{\bar{A}}^{-1}\bar{\bar{C}})\mathbf{E}_e - \bar{\bar{A}}^{-1}\mathbf{F}\ _2$ | Nombre<br>d'itérations |
|------|-------------|--|--|------------------------|
|      |             | $\ \mathbf{F}\ _2$   | $\ \bar{\bar{A}}^{-1}\mathbf{F}\ _2$   |                        |
| 0.3  | $4.10^{-7}$ | 0.01   | 0.77   | 11                     |
| 0.2  | $7.10^{-7}$ | 0.009  | 0.82   | 14                     |
| 0.15 | $3.10^{-7}$ | 0.007  | 0.81   | 15                     |
| 0.12 | $4.10^{-7}$ | 0.006  | 0.75   | 11                     |

TABLE 5.8 – Etude du résidu  $= \frac{\|(I + \bar{\bar{A}}^{-1}\bar{\bar{C}})\mathbf{E}_e - \bar{\bar{A}}^{-1}\mathbf{F}\|_2}{\|\bar{\bar{A}}^{-1}\mathbf{F}\|_2}$  et vérification de la solution exacte ; cas  $k_s = 3$ ,  $\varepsilon_p = 10^{-3}$ ,  $R_\Sigma = 1.5$ .

par la matrice du système linéaire montre que nous avons une bonne convergence vers la solution exacte en raffinant le maillage. Cette analyse explique aussi l'effet de préconditionner par la matrice  $\bar{\bar{A}}$  sur la résolution numérique.

| $h$  | $\ \mathbf{E}_c - \mathbf{E}_e\ _2$ | $\ (\bar{\bar{A}} + \bar{\bar{C}})(\mathbf{E}_c - \mathbf{E}_e)\ _2$ | $\ \mathbf{E}_c - \mathbf{E}_e\ _\infty$ | $\ (\bar{\bar{A}} + \bar{\bar{C}})(\mathbf{E}_c - \mathbf{E}_e)\ _\infty$ |
|------|-------------------------------------|--|--|---|
|      | $\ \mathbf{E}_e\ _2$                | $\ (\bar{\bar{A}} + \bar{\bar{C}})\mathbf{E}_e\ _2$                  | $\ \mathbf{E}_e\ _\infty$                | $\ (\bar{\bar{A}} + \bar{\bar{C}})\mathbf{E}_e\ _\infty$                  |
| 0.3  | 1.19                                | 0.01   | 2.48                                     | 0.12  |
| 0.2  | 1.61                                | 0.009  | 2.52                                     | 0.1   |
| 0.15 | 1.38                                | 0.007  | 2.16                                     | 0.09  |
| 0.12 | 1.11                                | 0.006  | 1.59                                     | 0.09  |

TABLE 5.9 – Erreurs relatives cas  $k_s = 3$ ,  $\varepsilon_p = 10^{-3}$ ,  $R_\Sigma = 1.5$ .

| $h$  | Résidu      | $\ (\bar{\bar{A}} + \bar{\bar{C}})\mathbf{E}_e - \mathbf{F}\ _2$ | $\ (I + \bar{\bar{A}}^{-1}\bar{\bar{C}})\mathbf{E}_e - \bar{\bar{A}}^{-1}\mathbf{F}\ _2$ | Nombre<br>d'itérations |
|------|-------------|--|--|------------------------|
|      |             | $\ \mathbf{F}\ _2$   | $\ \bar{\bar{A}}^{-1}\mathbf{F}\ _2$   |                        |
| 0.3  | $3.10^{-7}$ | 0.01   | 1.01   | 10                     |
| 0.2  | $6.10^{-7}$ | 0.008  | 0.98   | 16                     |
| 0.15 | $8.10^{-7}$ | 0.005  | 0.97   | 14                     |
| 0.12 | $7.10^{-7}$ | 0.003  | 0.87   | 13                     |

TABLE 5.10 – Etude du résidu et vérification de la solution exacte ; cas  $k_s = 3$ ,  $\varepsilon_p = 10^{-4}$ ,  $R_\Sigma = 1.5$ .

| $h$  | $\frac{\ \mathbf{E}_c - \mathbf{E}_e\ _2}{\ \mathbf{E}_e\ _2}$ | $\frac{\ (\bar{\bar{A}} + \bar{\bar{C}})(\mathbf{E}_c - \mathbf{E}_e)\ _2}{\ (\bar{\bar{A}} + \bar{\bar{C}})\mathbf{E}_e\ _2}$ | $\frac{\ \mathbf{E}_c - \mathbf{E}_e\ _\infty}{\ \mathbf{E}_e\ _\infty}$ | $\frac{\ (\bar{\bar{A}} + \bar{\bar{C}})(\mathbf{E}_c - \mathbf{E}_e)\ _\infty}{\ (\bar{\bar{A}} + \bar{\bar{C}})\mathbf{E}_e\ _\infty}$ |
|------|--|--|--|--|
| 0.3  | 5.8  | 0.01   | 17   | 0.11   |
| 0.2  | 3.6  | 0.008  | 14.6   | 0.09   |
| 0.15 | 2.27   | 0.005  | 8.6  | 0.08   |
| 0.12 | 1.55   | 0.003  | 4.9  | 0.06   |

TABLE 5.11 – Erreurs relatives cas  $k_s = 3$ ,  $\varepsilon_p = 10^{-4}$ ,  $R_\Sigma = 1.5$ .

| $h$  | Résidu            | $\frac{\ (\bar{\bar{A}} + \bar{\bar{C}})\mathbf{E}_e - \mathbf{F}\ _2}{\ \mathbf{F}\ _2}$ | $\frac{\ (I + \bar{\bar{A}}^{-1}\bar{\bar{C}})\mathbf{E}_e - \bar{\bar{A}}^{-1}\mathbf{F}\ _2}{\ \bar{\bar{A}}^{-1}\mathbf{F}\ _2}$ | Nombre<br>d'itérations |
|------|-------------------|---|---|------------------------|
|      |                   |   |   |                        |
| 0.3  | $7 \cdot 10^{-7}$ | 0.01  | 1.01  | 8                      |
| 0.2  | $4 \cdot 10^{-7}$ | 0.008   | 1.001   | 12                     |
| 0.15 | $5 \cdot 10^{-7}$ | 0.005   | 1.009   | 13                     |
| 0.12 | $8 \cdot 10^{-7}$ | 0.003   | 0.99  | 10                     |

TABLE 5.12 – Etude du résidu et vérification de la solution exacte ; cas  $k_s = 3$ ,  $\varepsilon_p = 10^{-5}$ ,  $R_\Sigma = 1.5$ .

| $h$  | $\frac{\ \mathbf{E}_c - \mathbf{E}_e\ _2}{\ \mathbf{E}_e\ _2}$ | $\frac{\ (\bar{\bar{A}} + \bar{\bar{C}})(\mathbf{E}_c - \mathbf{E}_e)\ _2}{\ (\bar{\bar{A}} + \bar{\bar{C}})\mathbf{E}_e\ _2}$ | $\frac{\ \mathbf{E}_c - \mathbf{E}_e\ _\infty}{\ \mathbf{E}_e\ _\infty}$ | $\frac{\ (\bar{\bar{A}} + \bar{\bar{C}})(\mathbf{E}_c - \mathbf{E}_e)\ _\infty}{\ (\bar{\bar{A}} + \bar{\bar{C}})\mathbf{E}_e\ _\infty}$ |
|------|--|--|--|--|
| 0.3  | 23   | 0.01   | 78   | 0.1  |
| 0.2  | 19.5   | 0.008  | 85   | 0.09   |
| 0.15 | 12.8   | 0.005  | 47.6   | 0.08   |
| 0.12 | 8.6  | 0.003  | 36.6   | 0.06   |

TABLE 5.13 – Erreurs relatives cas  $k_s = 3$ ,  $\varepsilon_p = 10^{-5}$ ,  $R_\Sigma = 1.5$ .

| $h$  | Résidu            | $\frac{\ (\bar{\bar{A}} + \bar{\bar{C}})\mathbf{E}_e - \mathbf{F}\ _2}{\ \mathbf{F}\ _2}$ | $\frac{\ (I + \bar{\bar{A}}^{-1}\bar{\bar{C}})\mathbf{E}_e - \bar{\bar{A}}^{-1}\mathbf{F}\ _2}{\ \bar{\bar{A}}^{-1}\mathbf{F}\ _2}$ | Nombre<br>d'itérations |
|------|-------------------|---|---|------------------------|
| 0.3  | $9 \cdot 10^{-7}$ | 0.01  | 1   | 8                      |
| 0.2  | $10^{-7}$         | 0.008   | 0.99  | 10                     |
| 0.15 | $10^{-7}$         | 0.005   | 1.002   | 11                     |
| 0.12 | $9 \cdot 10^{-7}$ | 0.003   | 0.99  | 11                     |

TABLE 5.14 – Etude du résidu et vérification de la solution exacte ; cas  $k_s = 3$ ,  $\varepsilon_p = 10^{-6}$ ,  $R_\Sigma = 1.5$

| $h$  | $\frac{\ \mathbf{E}_c - \mathbf{E}_e\ _2}{\ \mathbf{E}_e\ _2}$ | $\frac{\ (\bar{\bar{A}} + \bar{\bar{C}})(\mathbf{E}_c - \mathbf{E}_e)\ _2}{\ (\bar{\bar{A}} + \bar{\bar{C}})\mathbf{E}_e\ _2}$ | $\frac{\ \mathbf{E}_c - \mathbf{E}_e\ _\infty}{\ \mathbf{E}_e\ _\infty}$ | $\frac{\ (\bar{\bar{A}} + \bar{\bar{C}})(\mathbf{E}_c - \mathbf{E}_e)\ _\infty}{\ (\bar{\bar{A}} + \bar{\bar{C}})\mathbf{E}_e\ _\infty}$ |
|------|--|--|--|--|
| 0.3  | 75   | 0.01   | 296  | 0.1  |
| 0.2  | 61   | 0.008  | 390  | 0.09   |
| 0.15 | 58   | 0.005  | 238  | 0.08   |
| 0.12 | 37   | 0.003  | 236  | 0.06   |

TABLE 5.15 – Erreurs relatives cas  $k_s = 3$ ,  $\varepsilon_p = 10^{-6}$ ,  $R_\Sigma = 1.5$ .

A travers les tableaux 5.8, 5.10, 5.12 et 5.14, nous constatons que le solveur itératif converge pour un nombre d'itérations réduit. L'analyse du terme  $\frac{\|(\bar{\bar{A}} + \bar{\bar{C}})\mathbf{E}_e - \mathbf{F}\|_2}{\|\mathbf{F}\|_2}$  montre que la solution exacte vérifie le système linéaire considéré pour tout choix de  $\varepsilon_p$ . Or, la vérification de  $\mathbf{E}_e$  sur le même système preconditionné par la matrice  $\bar{\bar{A}}$  en calculant le terme  $\frac{\|(I + \bar{\bar{A}}^{-1}\bar{\bar{C}})\mathbf{E}_e - \bar{\bar{A}}^{-1}\mathbf{F}\|_2}{\|\bar{\bar{A}}^{-1}\mathbf{F}\|_2}$  nous mène à poser des questions sur le conditionnement de la matrice  $\bar{\bar{A}}$ . Au regard des tableaux 5.9, 5.11, 5.13 et 5.15, nous pouvons tirer la même conclusion précédente. En fait, nous remarquons une grande différence entre l'erreur relative  $\frac{\|\mathbf{E}_c - \mathbf{E}_e\|_2}{\|\mathbf{E}_e\|_2}$  et l'erreur relative pondérée par la matrice du système  $\frac{\|(\bar{\bar{A}} + \bar{\bar{C}})(\mathbf{E}_c - \mathbf{E}_e)\|_2}{\|(\bar{\bar{A}} + \bar{\bar{C}})\mathbf{E}_e\|_2}$ . Ces deux quantités sont également calculées en norme infinie. De plus, l'écart entre la solution calculée et le solution exacte devient de plus en plus grand lorsque  $\varepsilon_p$  diminue. Tous ces résultats numériques nous mènent à étudier le conditionnement de la matrice  $\bar{\bar{A}}$ . Cette matrice peut s'écrire comme une somme d'une matrice  $\bar{\bar{A}}_{EF}$  et  $\bar{\bar{A}}_{eps}$  où  $\bar{\bar{A}}_{eps}$  est la matrice résultant de la pénalisation de la condition essentielle. Nous désignons  $\lambda_g$  (resp.  $\lambda_p$ ) la plus grande (resp. petite) valeur propre d'une matrice donnée. D'après les tableaux 5.16, 5.17 et 5.18, nous constatons l'effet de la matrice  $\bar{\bar{A}}_{eps}$  sur le conditionnement de la matrice  $\bar{\bar{A}}$  lorsque  $\varepsilon_p$  tend vers 0. On observe que la plus grande valeur propre de la matrice  $\bar{\bar{A}}$  est du même ordre que celle de la matrice  $\bar{\bar{A}}_{eps}$ . De plus, elle est proportionnelle à  $\varepsilon_p^{-1}$ . Ceci, explique la perturbation observée au niveau de l'étude de la vérification de la solution exacte sur le système linéaire preconditionné ainsi que sur la précision de calcul. Nous pouvons donc conclure que la stratégie de pénalisation détériore ce conditionnement des matrices du système linéaire ce qui nous donne un résultat limité en terme de précision.

| $\lambda_g$                             | $\lambda_p$                | $cond_2$  |
|---|----------------------------|-----------|
| $(14.4865) - (7.32195 \cdot 10^{-6}) i$ | $0.0859181 - (0.087054) i$ | 118.43854 |

TABLE 5.16 – Matrice  $\bar{\bar{A}}_{EF}$ , cas  $h = 0.3$  et  $k_s = 2.1$ ,  $R_\Sigma = 1.5$ .

| $\varepsilon_p$ | $\lambda_g$ | $\lambda_p$              | $cond_2$  |
|-----------------|-------------|--------------------------|-----------|
| $10^{-2}$       | 8.84872     | $6.51797 \cdot 10^{-7}$  | 13575883. |
| $10^{-3}$       | 88.4582     | $7.66253 \cdot 10^{-7}$  | 1.154D+08 |
| $10^{-4}$       | 884.565     | $-2.04043 \cdot 10^{-7}$ | 4.335D+09 |
| $10^{-5}$       | 8845.66     | $7.35942 \cdot 10^{-7}$  | 1.202D+10 |
| $10^{-6}$       | 88456.7     | $9.1692 \cdot 10^{-7}$   | 9.647D+10 |

TABLE 5.17 –  $\bar{\bar{A}}_{eps}$ , cas  $h = 0.3$  et  $k_s = 2.1$ ,  $R_\Sigma = 1.5$ .

| $\varepsilon_p$ | $\lambda_g$                              | $\lambda_p$                 | $cond_2$ |
|-----------------|--|-----------------------------|----------|
| $10^{-2}$       | $(16.2733) - (4.46251 \cdot 10^{-6}) i$  | $(0.111321) + (0.111321) i$ | 103.36   |
| $10^{-3}$       | $(89.2469) - (1.60268 \cdot 10^{-8}) i$  | $(0.141682) - (0.123085) i$ | 475.52   |
| $10^{-4}$       | $(885.346) - (1.3209 \cdot 10^{-10}) i$  | $(0.153267) - (0.135483) i$ | 4327.96  |
| $10^{-5}$       | $(8846.45) - (1.1022 \cdot 10^{-12}) i$  | $(0.154249) - (0.135636) i$ | 43069.01 |
| $10^{-6}$       | $(88457.5) - (1.57468 \cdot 10^{-12}) i$ | $(0.154346) - (0.135646) i$ | 430489.5 |

TABLE 5.18 – Matrice  $\bar{\bar{A}} = \bar{\bar{A}}_{EF} + \bar{\bar{A}}_{eps}$ , cas  $h = 0.3$  et  $k_s = 2.1$ ,  $R_\Sigma = 1.5$ .

Un autre facteur contribuant à cette limitation est l'utilisation de la méthode LU pour préconditionner le système linéaire par la matrice  $\bar{\bar{A}}$ . En fait, nous avons étudié l'impact de  $\varepsilon_p$  sur la résolution de la méthode LU. Ceci consiste à comparer le second membre donné d'un système linéaire défini par la matrice  $\bar{\bar{A}}$ , avec le produit de  $\bar{\bar{A}}$  par la solution calculée par la méthode LU. A travers les résultats du tableau 5.19, nous constatons que le choix de  $\varepsilon_p$  de plus en plus petit génère une instabilité du solveur LU.

| $\varepsilon_p$ | $10^{-2}$    | $10^{-3}$    | $10^{-4}$    | $10^{-5}$    | $10^{-6}$   | $10^{-7}$   |
|-----------------|--------------|--------------|--------------|--------------|-------------|-------------|
| Vérification LU | $3.10^{-15}$ | $7.10^{-14}$ | $5.10^{-12}$ | $4.10^{-10}$ | $7.10^{-7}$ | $2.10^{-3}$ |

TABLE 5.19 – Vérification LU en fonction de  $\varepsilon_p$ ; cas  $h = 0.3$ ,  $k_s = 3$ .

Nous étudions dans la figure 5.1, les résultats de l'erreur relative pondérée par la matrice  $\bar{\bar{A}} + \bar{\bar{C}}$  pour différents choix de  $\varepsilon_p$  (voir Fig. 5.1). On constate que le choix de  $\varepsilon_p$  de plus en plus petit, donne une solution de plus en plus proche de la solution exacte. Nous remarquons que la courbe pour le cas  $\varepsilon_p = 10^{-4}$  est la même que celle pour  $\varepsilon_p = 10^{-5}$ . Dans notre cas, nous ne pouvons plus avoir mieux que ce résultat à cause de l'effet des

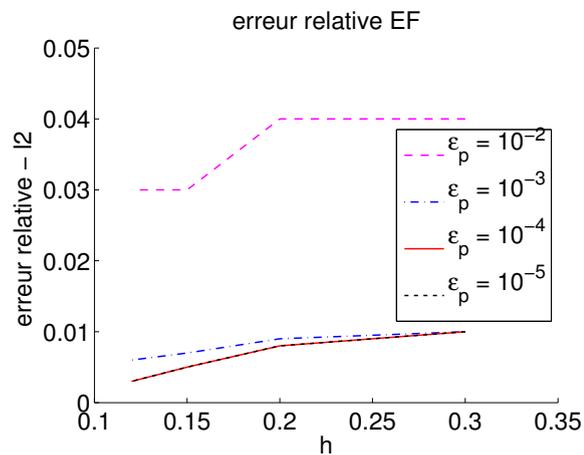


FIGURE 5.1 – Etude de l’erreur relative pondérée par la matrice  $\bar{\bar{A}} + \bar{\bar{C}}$  en fonction du paramètre de pénalisation

erreurs éléments finis sur la résolution du problème. Nous étudions aussi (figure 5.2) le comportement des résidus du gradient biconjugué stabilisé pour un choix de  $hk$  fixé (en mettant 10 points par longueur d’onde). Nous constatons que l’algorithme converge rapidement et que le nombre d’itérations ne dépasse pas les 16 itérations pour atteindre la convergence souhaitée (cas  $k_s = 3.1$ ).

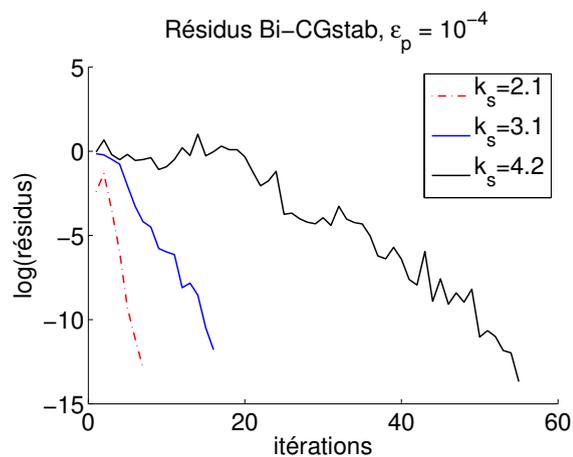


FIGURE 5.2 – Etude du comportement des résidus, cas  $hk_s$  fixé, cas  $k_s = 2.1, k_s = 3.1, k_s = 4.2, \varepsilon_p = 10^{-4}, R_\Sigma = 1.5$ .

### 5.3 Résolution du problème de Maxwell initial

Nous rappelons le problème auquel nous nous intéressons :

$$\left\{ \begin{array}{l} \text{rot rot} \mathbf{E} - t^{-1} \nabla(\text{div} \mathbf{E}) - k_s^2 \mathbf{E} = 0 \text{ dans } \Omega, \\ \mathbf{E} \wedge \mathbf{n} = 0 \text{ sur } \Gamma, \\ \text{div} \mathbf{E} = 0 \text{ sur } \Gamma, \\ T_{\nu_1}(\mathbf{E}) = T_{\nu_1}(\mathbf{E}^{inc} - \mathcal{I}_t^{\mathcal{R}}(\Gamma, \mathbf{E})) \text{ sur } \Sigma, \\ N_{\nu_2}(\mathbf{E}) = N_{\nu_2}(\mathbf{E}^{inc} - \mathcal{I}_t^{\mathcal{R}}(\Gamma, \mathbf{E})) \text{ sur } \Sigma. \end{array} \right. \quad (5.9)$$

Nous considérons le cas d'une onde incidente plane de direction  $(0, 0, 1)$  et de polarisation  $(1, 0, 0)$ . Le critère d'arrêt du gradient biconjugué stabilisé correspond à un résidu inférieur à  $10^{-6}$ . Nous choisissons  $\nu_1 = \nu_2 = -ik_s$  et un paramètre de régularisation  $t = 1$ , ainsi que le paramètre de pénalisation  $\varepsilon_p = 10^{-4}$ .

| $h$  | $\frac{\ (\bar{A} + \bar{C})\mathbf{E}_c - \mathbf{F}\ _2}{\ \mathbf{F}\ _2}$ | $\frac{\ (I + \bar{A}^{-1}\bar{C})\mathbf{E}_c - \bar{A}^{-1}\mathbf{F}\ _2}{\ A^{-1}\mathbf{F}\ _2}$ | Nombre d'itérations |
|------|---|---|---------------------|
| 0.3  | $10^{-7}$   | $10^{-7}$   | 8                   |
| 0.2  | $6.10^{-7}$   | $5.10^{-7}$   | 9                   |
| 0.15 | $3.10^{-7}$   | $3.10^{-7}$   | 11                  |

TABLE 5.20 – Etude du comportement des résidus du BI-CGstab, cas  $h$  varie et  $k_s = 3$ .

Nous nous proposons d'établir une analyse comparative entre les résultats de la résolution du problème (5.9) par deux choix de préconditionneur. En fait, le problème (5.9) peut s'écrire de deux manières équivalentes comme

$$(\mathcal{A}_t + \mathcal{C}_t)\mathbf{E} = \mathcal{F}_t. \quad (5.10)$$

ou

$$(\mathcal{A}_t^c + \mathcal{C}_t^c)\mathbf{E} = \mathcal{F}_t. \quad (5.11)$$

Les opérateurs  $\mathcal{A}_t$ ,  $\mathcal{C}_t$ ,  $\mathcal{A}_t^c$  et  $\mathcal{C}_t^c$  sont définis dans le chapitre 4. L'étude théorique que nous avons faite permet de justifier le choix du préconditionneur de la méthode de Schwarz comme un préconditionneur au problème (5.10). Dans cette partie, nous avons étudié le comportement du solveur pour la résolution de

$$(\mathcal{I} + \mathcal{A}_t^{-1}\mathcal{C}_t)\mathbf{E} = \mathcal{A}_t^{-1}\mathcal{F}_t. \quad (5.12)$$

Puis, nous avons refait l'étude en préconditionnant par l'opérateur coercif  $\mathcal{A}_t^c$  ce qui donne

$$(\mathcal{I} + (\mathcal{A}_t^c)^{-1}\mathcal{C}_t^c)\mathbf{E} = (\mathcal{A}_t^c)^{-1}\mathcal{F}_t. \quad (5.13)$$

Le système linéaire correspondant au problème (5.13) est

$$(I + (\bar{A}^c)^{-1}\bar{C}^c)\mathbf{E} = (\bar{A}^c)^{-1}\mathbf{F}.$$

Pour les tests numériques, nous avons choisi un nombre d'onde égal à  $k_s = 2.1$ ,  $k_s = 3.1$  et  $k_s = 4.2$ , avec des maillages satisfaisant la présence de 10 points par longueur d'onde.

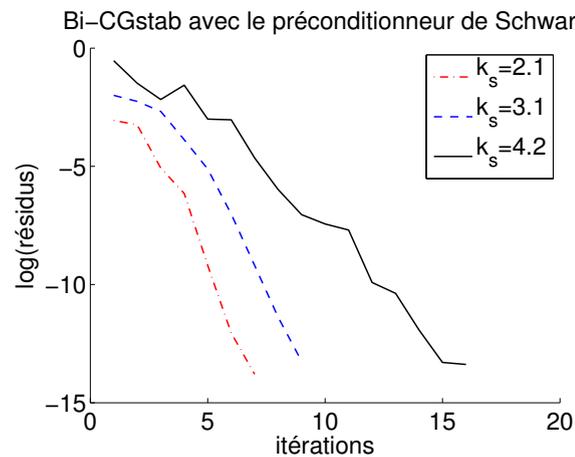


FIGURE 5.3 – Etude du comportement des résidus du BI-CGstab préconditionné par  $\bar{A}$ , cas  $hk_s$  fixé,  $\varepsilon_p = 10^{-4}$ ,  $R_\Sigma = 1.5$ .

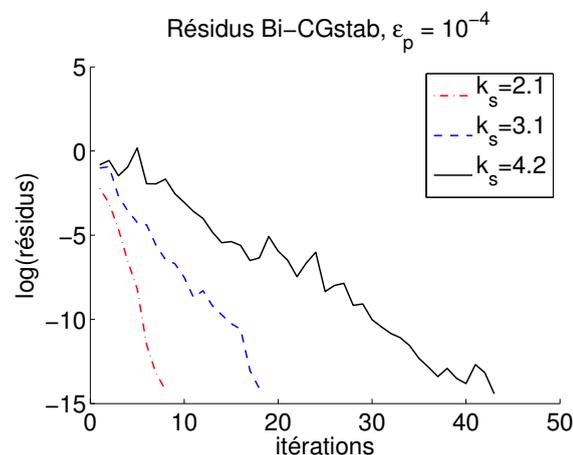


FIGURE 5.4 – Etude du comportement des résidus du BI-CGstab préconditionné par  $\bar{A}^c$ , cas  $hk_s$  fixé,  $\varepsilon_p = 10^{-4}$ ,  $R_\Sigma = 1.5$ .

Au regard des courbes des figures 5.3 et 5.4, nous constatons que le gradient biconjugué stabilisé converge rapidement. Nous remarquons l'avantage de choisir le préconditionneur de Schwarz. Par exemple, pour le cas de  $k_s = 3.1$ , le solveur itératif avec le préconditionneur de Schwarz converge après 9 itérations. Par contre en choisissant  $\bar{A}^c$ , la convergence est atteinte après 18 itérations.

A l'aide de Medit, nous visualisons la partie réelle et la partie imaginaire de la première composante de la solution calculée par le programme **CEFRI** avec le préconditionneur de

la méthode de Schwarz. Pour cela, nous avons fait une coupe selon le plan d'équation  $x = 0$  et nous présentons le résultat pour chaque cas étudié.

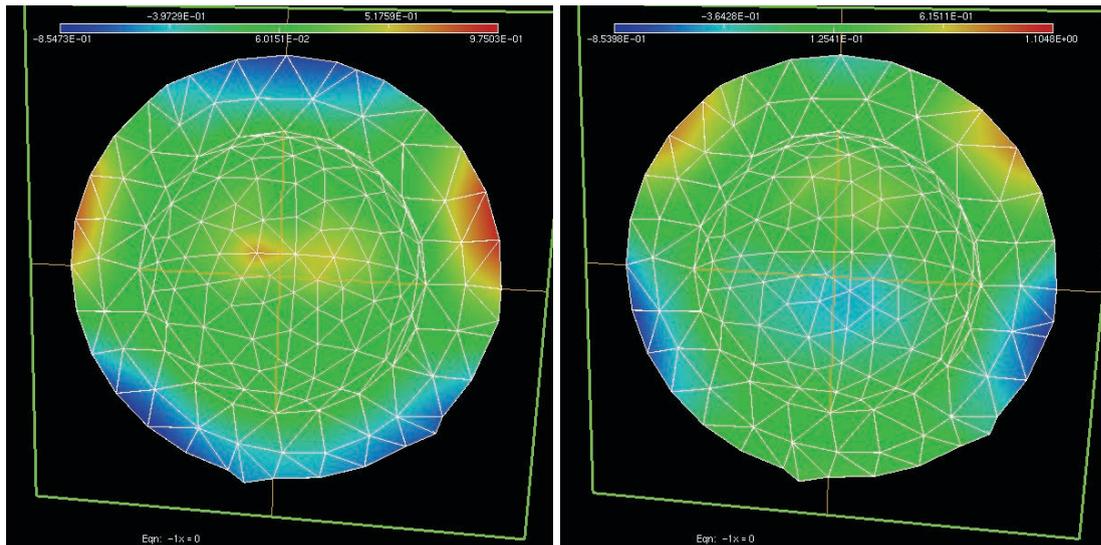


FIGURE 5.5 – Cas  $k_s = 2.1, h = 0.3$ ; à droite (respectivement à gauche) une coupe de la première composante de la partie imaginaire du champ total (respectivement la partie réelle).

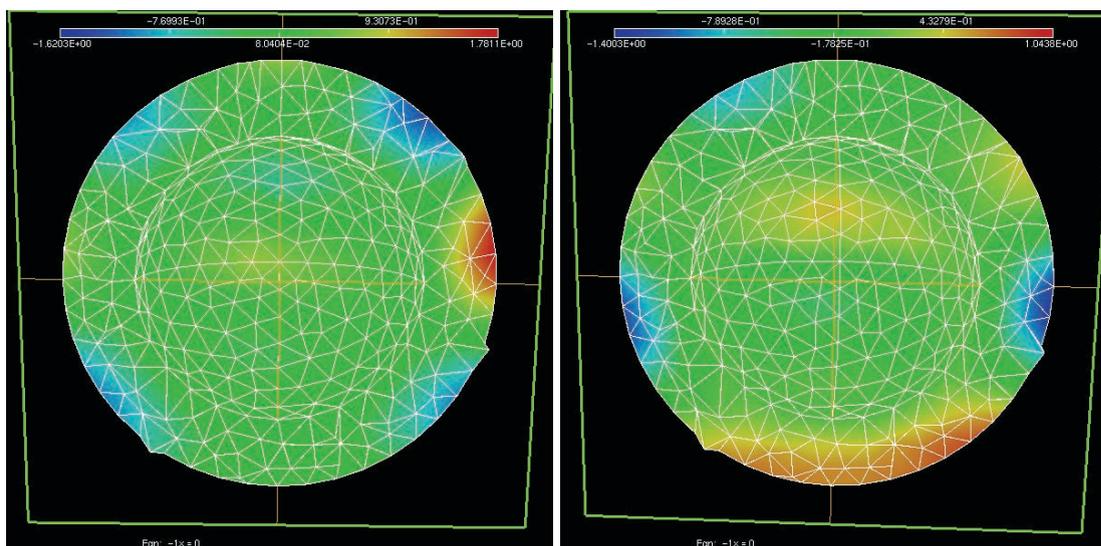


FIGURE 5.6 – Cas  $k_s = 3.1, h = 0.2$ ; à droite (respectivement à gauche) une coupe de la première composante de la partie imaginaire du champ total (respectivement la partie réelle).

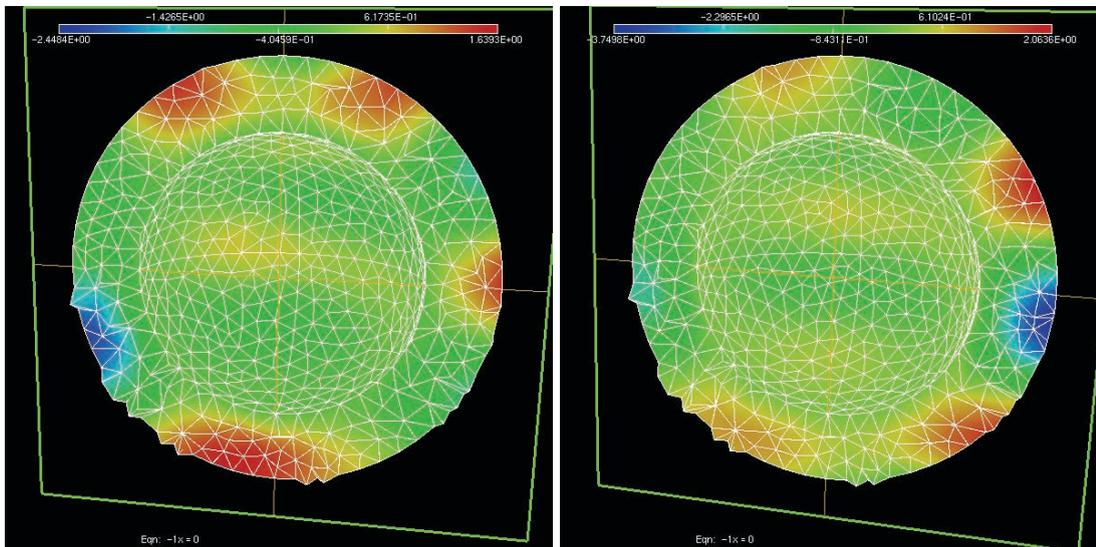


FIGURE 5.7 – Cas  $k_s = 4.2$ ,  $h = 0.15$ ; à droite (respectivement à gauche) une coupe de la première composante de la partie imaginaire du champ total (respectivement la partie réelle).

Au regard de ces figures, nous remarquons que les résultats semblent cohérents. On voit en particulier que les oscillations augmentent avec la fréquence.

Nous considérons désormais le GMRES pour les mêmes cas tests que ceux réalisés avec le gradient biconjugué stabilisé. Nous nous intéressons à l'investigation numérique de la convergence de la méthode GMRES afin de valider numériquement l'étude théorique effectuée au chapitre 4.

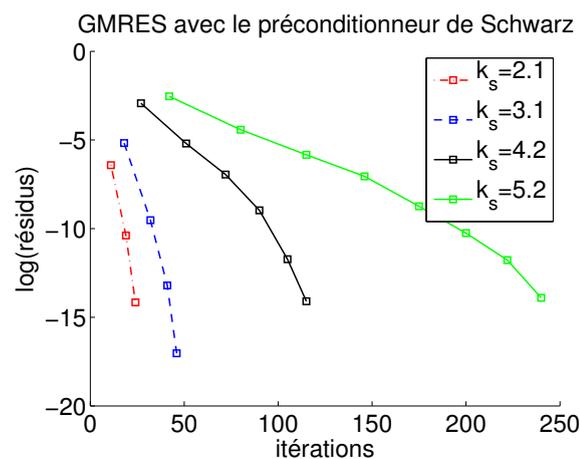


FIGURE 5.8 – Etude du comportement des résidus du GMRES préconditionné par  $\bar{\bar{A}}$ , cas  $hk_s$  fixé,  $\varepsilon_p = 10^{-4}$ ,  $R_\Sigma = 1.5$ .

Dans la figure 5.8, nous présentons les résidus du GMRES à chaque restart. En fait, le GMRES de MÉLINA ++ est programmé de telle sorte que le choix de la dimension de l'espace de Krylov est donné par l'utilisateur ou défini par défaut. Par exemple, pour le cas test  $k_s = 2.1$ , le GMRES effectue 3 restarts avec un espace de Krylov de dimension

attribuée par défaut par MÉLINA++ égale à 65. Le nombre d'itérations correspondant à ce choix est 24 itérations (voir tableau 5.21). La même étude est effectuée pour les nombres d'onde  $k_s = 3.1$ ,  $k_s = 4.2$  et  $k_s = 5.2$ . A travers ces cas tests simples, nous pouvons conclure que le comportement asymptotique du GMRES valide l'étude théorique de la vitesse de convergence du solveur.

| $k_s$ | Dimension de Krylov | Nombre de restarts | Nombre d'itérations |
|-------|---------------------|--------------------|---------------------|
| 2.1   | 65                  | 3                  | 24                  |
| 3.1   | 108                 | 4                  | 46                  |
| 4.2   | 161                 | 6                  | 115                 |
| 5.2   | 232                 | 8                  | 240                 |

TABLE 5.21 – Etude du nombre d'itérations du GMRES préconditionné par  $\bar{A}$  pour les dimensions d'espace de Krylov données par défaut par MÉLINA++, cas  $hk_s$ ,  $\varepsilon_p = 10^{-4}$ ,  $R_\Sigma = 1.5$ .

Une question que nous proposons de regarder plus précisément concerne l'impact du choix de la dimension de l'espace de Krylov sur la méthode GMRES. Pour cela, nous avons fait varier la dimension de l'espace de Krylov pour le cas test où  $k_s = 2.1$ . A travers la figure 5.9, nous remarquons qu'on n'a pas encore atteint le comportement asymptotique du GMRES pour le choix d'une dimension de l'espace de Krylov inférieure à 6. Mais au-delà de ceci, le GMRES converge d'une manière superlinéaire. Compte-tenu des résultats du tableau 5.22, nous constatons que le choix optimal de la dimension de l'espace de Krylov pour cette configuration est 4. Pour ce cas, le nombre de restarts du GMRES est égal à 3 correspondant à 12 itérations. Dans notre configuration, il semble préférable de choisir une dimension d'espace de Krylov petite pour atteindre la convergence à un temps de calcul satisfaisant.

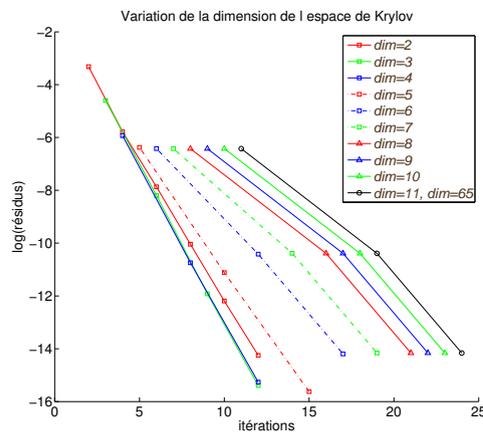


FIGURE 5.9 – Etude du comportement des résidus du GMRES préconditionné par  $\bar{A}$  en fonction du choix de la dimension de l'espace de Krylov, cas  $k_s = 2.1$  ( $\lambda = 10h$ ),  $\varepsilon_p = 10^{-4}$ .

| Dimension de Krylov | Nombre de restarts | Nombre d'itérations | CPU résolution GMRES (s) |
|---------------------|--------------------|---------------------|--------------------------|
| 2                   | 6                  | 2+2+2+2+2+2=12      | 19                       |
| 3                   | 4                  | 3+3+3+3=12          | 17                       |
| 4                   | 3                  | 4+4+4=12            | 16                       |
| 5                   | 3                  | 5+5+5=15            | 19                       |
| 6                   | 3                  | 6+6+5=17            | 21                       |
| 7                   | 3                  | 7+7+5=19            | 23                       |
| 8                   | 3                  | 8+8+5=21            | 25                       |
| 9                   | 3                  | 9+8+5=22            | 28                       |
| 10                  | 3                  | 10+8+5=23           | 27                       |
| 11                  | 3                  | 11+8+5=24           | 28                       |
| 12                  | 3                  | 11+8+5=24           | 28                       |
| 65                  | 3                  | 11+8+5=24           | 28                       |

TABLE 5.22 – Etude du nombre d'itérations du GMRES préconditionné par  $\bar{\bar{A}}$  en fonction du choix de la dimension d'espace de Krylov, cas  $k_s = 2.1$  ( $\lambda = 10h$ ),  $\varepsilon_p = 10^{-4}$ ,  $R_{\Sigma} = 1.5$ .

Nous avons mené la même étude avec une fréquence plus grande ( $k_s = 4.2$ ). Au regard de la figure 5.10 et du tableau 5.23, nous pouvons tirer les mêmes conclusions que celles citées précédemment. Dans ce cas, le choix optimal de la dimension de l'espace de Krylov est 6. Le temps de calcul de la résolution du GMRES est à peu près égal à un tiers de celui obtenu avec la dimension par défaut de l'espace de Krylov (161). Nous constatons que le choix optimal de la dimension de l'espace de Krylov croît avec le nombre d'onde.

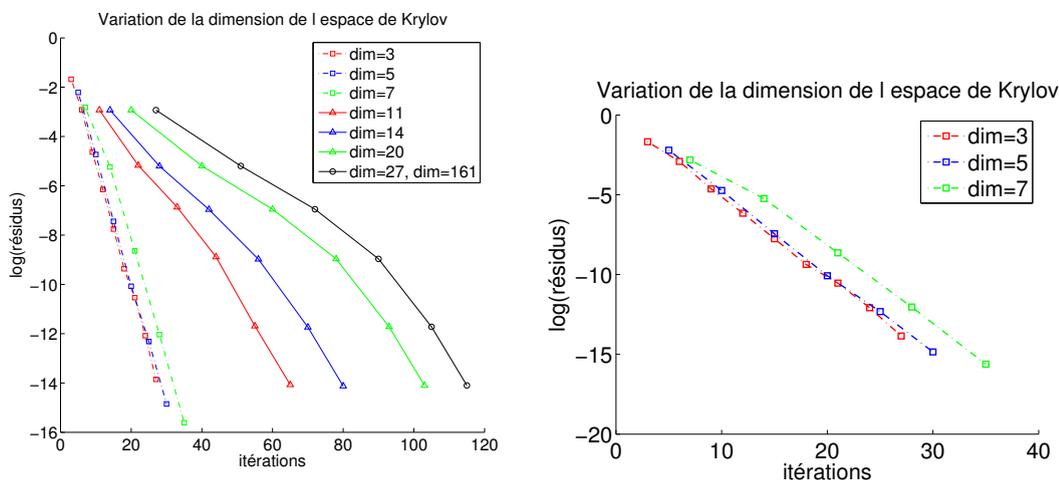


FIGURE 5.10 – Etude du comportement des résidus du GMRES préconditionné par  $\bar{\bar{A}}$  en fonction du choix de la dimension de l'espace de Krylov, cas  $k_s = 4.2$  ( $\lambda = 10h$ ),  $\varepsilon_p = 10^{-4}$  (la figure à droite est un zoom de celle qui est à gauche).

| Dimension de Krylov | Nombre de restarts | Nombre d'itérations   | CPU résolution GMRES (s) |
|---------------------|--------------------|-----------------------|--------------------------|
| 2                   | 14                 | 2+2+...+2=28          | 1003                     |
| 3                   | 9                  | 3+3+...+3=27          | 863                      |
| 4                   | 7                  | 4+4+...+4=28          | 845                      |
| 5                   | 6                  | 5+5+...+5=30          | 862                      |
| 6                   | 5                  | 6+6+...+6=30          | 844                      |
| 7                   | 5                  | 7+7+...+7=35          | 956                      |
| 11                  | 6                  | 11+11...+11+10=65     | 1689                     |
| 14                  | 6                  | 14+14...+14+10=80     | 2040                     |
| 20                  | 6                  | 20+20+20+18+15+10=103 | 2565                     |
| 27                  | 6                  | 27+24+21+18+15+10=115 | 2860                     |
| 161                 | 6                  | 27+24+21+18+15+10=115 | 2859                     |

TABLE 5.23 – Etude du nombre d'itérations de GMRES préconditionné par  $\bar{A}$  en fonction du choix de la dimension d'espace de Krylov, cas  $k_s = 4.2$  ( $\lambda = 10h$ ),  $\varepsilon_p = 10^{-4}$ ,  $R_\Sigma = 1.5$ .



## Conclusions et perspectives

Nous nous sommes intéressés à la résolution des problèmes de diffraction en domaine non borné par la méthode de couplage entre éléments finis et représentation intégrale **CEFRI**. L'étude comparative entre différentes approches pour le traitement du caractère infini du domaine établie pour le problème de Helmholtz a montré quelques avantages de la méthode **CEFRI**. Nous avons constaté que la méthode **CEFRI** n'a pas besoin de considérer un domaine de calcul très large. En fait, il suffit de placer le bord artificiel à une ou deux couches éléments finis du bord de l'obstacle pour avoir des résultats satisfaisants. Ceci, n'est pas vrai pour les méthodes à condition aux limites transparentes locales telles que la méthode à condition de rayonnement à distance finie (**CRDF**) et la méthode à condition transparente de type Bayliss-Gunzburger-Turkel (**BGT**). Mais, la méthode **CEFRI** nécessite plus de développement numérique par rapport à celui employé pour les méthodes **CRDF** et **BGT**. Le système à résoudre fait intervenir des opérateurs intégraux ce qui rend crucial le recours à des méthodes rapides pour la considération de hautes fréquences. Dans ce cadre, nous avons mis en œuvre la méthode FMM 1 niveau et nous avons constaté le gain en temps de calcul. De plus, à travers des tests numériques nous avons mis l'accent sur la mise en évidence des limitations du développement classique de la méthode FMM qui s'avère limité face au problème dit de basses fréquences.

D'ailleurs, nous avons étendu les résultats faits dans [15] pour le problème de Helmholtz posé en domaine non borné au problème de Maxwell. En fait, nous avons montré l'analogie entre la méthode de couplage entre éléments finis et représentation intégrale **CEFRI** et la méthode de Schwarz avec recouvrement total pour la résolution du problème de Maxwell en domaine non borné. Ceci nous a permis de faire une exploration analytique de la méthode de Schwarz pour une géométrie simplifiée afin d'illustrer le taux de convergence. Nous avons constaté que la méthode de Schwarz avec recouvrement total converge pour le choix d'un domaine de calcul assez large. Cette relecture de la méthode **CEFRI** offre également une technique de préconditionnement pour les solveurs de Krylov et nous a permis d'avoir une idée préliminaire sur la convergence de ces méthodes. Ainsi, nous nous intéressons plutôt à des méthodes itératives rapides. Pour cela, nous avons mené une analyse théorique afin de montrer la convergence superlinéaire du GMRES.

D'un point de vue numérique, une validation de ces aspects est établie par la librairie éléments finis **MÉLINA++**. Dans un premier volet, nous avons greffé au sein de la librairie **MÉLINA++** les intégrands et les noyaux de Green nécessaires à la résolution du problème. On a également mis en évidence quelques bugs initiaux de la librairie ce qui a permis ainsi la correction et la validation de quelques objets essentiels de la librairie. Compte-tenu de la librairie **MÉLINA++**, nous avons choisi la méthode de pénalisation pour la prise en compte de la condition essentielle du problème de diffraction électromagnétique par un conducteur parfait. Ceci a nécessité une étude minutieuse du comportement numérique.

Nous avons conclu que la stratégie de la pénalisation détériore le conditionnement des matrices du système linéaire. Ceci a un effet sur les résultats de la méthode LU utilisée pour le préconditionnement du système linéaire. Cette limitation a participé d'une manière significative à la dégradation de la précision des calculs. La résolution numérique est établie par les méthodes rapides de type Krylov : le Bi-CGstab et le GMRES. Nous avons constaté une convergence rapide de ces deux solveurs avec le préconditionneur de la méthode de Schwarz. Nous avons mis l'accent sur l'analyse numérique de la convergence de GMRES afin de justifier l'étude théorique que nous avons accomplie. A travers des cas tests simples, nous avons montré que le choix de la dimension de l'espace de Krylov impacte sur le temps de la résolution du GMRES. Nous avons vérifié que le GMRES a un comportement asymptotique superlinéaire. Ainsi, dans le cadre de notre application, le GMRES devrait être stable à une montée en fréquence.

Dans le cadre d'une application de la méthode **CEFRI** à une configuration concrète des problèmes de diffraction des ondes électromagnétiques, l'utilisation de la méthode FMM peut être une solution aux problèmes liés à l'optimisation des coûts de calcul et de la place mémoire. Pour cela, nous envisageons de mettre en œuvre la méthode FMM multiniveaux valable pour un large spectre de fréquences incluant les basses fréquences en faisant appel aux travaux de E. Darve et P. Havé [24]. Nous souhaitons aussi mettre en œuvre les travaux de O. Bruno ([43], [44], [45]) et la méthode dite des H matrices ([46], [47], [48]) pour la méthode **CEFRI**. Ces méthodes sont destinées aussi à réduire la complexité de la mise en œuvre des problèmes de grande taille contenant des blocs pleins. Pour cela, nous comptons comparer ces différentes méthodes afin d'étudier la pertinence de chacune d'elles dans le cadre de nos applications. D'autre part, nous souhaitons étendre le travail que nous avons établi au problème de diffraction électromagnétique par un obstacle qui présente des singularités géométriques entrantes (coins, arêtes, sommets). Dans ce cas, nous pouvons faire appel à la méthode de la régularisation à poids pour garder l'usage des éléments finis de Lagrange ([58], [59]).

# Annexe A :

## Définitions des espaces

$L^2(\Omega) = \{u \text{ mesurable sur } \Omega \text{ tel que } \int_{\Omega} u^2 d\Omega < \infty\}$  muni de la norme

$$\|u\|_{L^2(\Omega)} = \left( \int_{\Omega} u^2 d\Omega \right)^{1/2}.$$

$L^2(\partial\Omega) = \{u \text{ mesurable sur } \partial\Omega \text{ tel que } \int_{\partial\Omega} u^2 d\Omega < \infty\}$ .

$H^1(\Omega) = \{u \in L^2(\Omega) \text{ tel que } \nabla u \in L^2(\Omega)^3\}$  muni de la norme

$$\|u\|_{H^1(\Omega)} = \left( \|u\|_{L^2(\Omega)}^2 + \|\nabla u\|_{L^2(\Omega)}^2 \right)^{1/2}.$$

$H^{1/2}(\partial\Omega) = \{u \in L^2(\partial\Omega) \text{ tel que } \int_{\partial\Omega} \int_{\partial\Omega} \frac{(u(x) - u(y))^2}{\|x - y\|^3} d\sigma(x) d\sigma(y) < \infty\}$ .

$L^2(\Omega)^3$  muni de la norme  $\|\mathbf{v}\|_{L^2(\Omega)^3} = \left( \|\mathbf{v}_1\|_{L^2(\Omega)}^2 + \|\mathbf{v}_2\|_{L^2(\Omega)}^2 + \|\mathbf{v}_3\|_{L^2(\Omega)}^2 \right)^{1/2}$ .

$H^1(\Omega)^3$  muni de la norme  $\|\mathbf{v}\|_{H^1(\Omega)^3} = \left( \|\mathbf{v}\|_{L^2(\Omega)^3}^2 + \|\nabla \mathbf{v}\|_{L^2(\Omega)^{3 \times 3}}^2 \right)^{1/2}$ .

$H(\mathbf{rot}, \Omega) = \{\mathbf{v} \in L^2(\Omega)^3 \text{ tel que } \mathbf{rot} \mathbf{v} \in L^2(\Omega)^3\}$  muni de la norme

$$\|\mathbf{v}\|_{H(\mathbf{rot}, \Omega)} = \left( \|\mathbf{v}\|_{L^2(\Omega)^3}^2 + \|\mathbf{rot} \mathbf{v}\|_{L^2(\Omega)^3}^2 \right)^{1/2}.$$

$H_0(\mathbf{rot}, \Omega) = \{\mathbf{v} \in H(\mathbf{rot}, \Omega); \mathbf{v} \wedge \mathbf{n}|_{\Gamma} = 0\}$ .

$H(\mathbf{div}, \Omega) = \{\mathbf{v} \in L^2(\Omega)^3 \text{ tel que } \mathbf{div} \mathbf{v} \in L^2(\Omega)\}$  muni de la norme

$$\|\mathbf{v}\|_{H(\mathbf{div}, \Omega)} = \left( \|\mathbf{v}\|_{L^2(\Omega)^3}^2 + \|\mathbf{div} \mathbf{v}\|_{L^2(\Omega)}^2 \right)^{1/2}.$$

## Opérateurs de traces

On définit les opérateurs de traces suivants :

$$\begin{aligned} \gamma : H^1(\Omega) &\rightarrow H^{1/2}(\Gamma) \\ v &\mapsto v|_{\Gamma} \end{aligned}$$

$$\begin{aligned} \gamma_n : H(\text{div}, \Omega) &\rightarrow H^{-1/2}(\Gamma) \\ \mathbf{v} &\mapsto \mathbf{v} \cdot \mathbf{n}|_{\Gamma} \\ \gamma_t : H(\mathbf{rot}, \Omega) &\rightarrow H^{-1/2}(\Gamma)^3 \\ \mathbf{v} &\mapsto \mathbf{v} \wedge \mathbf{n}|_{\Gamma} \end{aligned}$$

où  $H^{-1/2}(\Gamma)$  désigne le dual de  $H^{1/2}(\Gamma) = \{v|_{\Gamma}; v \in H^1(\Omega)\}$ .

## Définitions des sauts

Si  $\Gamma$  est une surface intermédiaire entre deux milieux 1 et 2 et  $\mathbf{n}_i$  la normale unitaire sortante au milieu  $i$ . On définit le saut de  $\mathbf{v}$  pour  $x \in \Gamma$  par :

$$[\mathbf{v}] = \mathbf{v}_1 - \mathbf{v}_2,$$

avec  $\mathbf{v}_1$  (resp.  $\mathbf{v}_2$ ) est la trace de  $\mathbf{v}$  sur  $\Gamma$  du côté extérieur (resp. du côté intérieur) du milieu 1. On définit le saut tangentiel de  $\mathbf{v}$  pour  $x \in \Gamma$  par :

$$[\mathbf{v}]_t = \mathbf{n}_1 \wedge \mathbf{v}_1 + \mathbf{n}_2 \wedge \mathbf{v}_2 = \mathbf{n}_1 \wedge [\mathbf{v}].$$

et le saut normal de  $\mathbf{v}$  pour  $x \in \Gamma$  par :

$$[\mathbf{v}]_n = \mathbf{n}_1 \cdot \mathbf{v}_1 + \mathbf{n}_2 \cdot \mathbf{v}_2 = \mathbf{n}_1 \cdot [\mathbf{v}]$$

## Représentations intégrales

### Cas de l'équation de Helmholtz

Soit  $G$  la fonction de Green associée à l'équation de Helmholtz en dimension 3, elle est définie par

$$G(x, y) = \frac{e^{ik||x-y||}}{4\pi||x-y||}.$$

**Proposition .0.1.** : *Considérons un domaine  $D$  borné, de frontière  $\Gamma$  régulière et une fonction  $\psi$  définie sur  $D$  solution de l'équation de Helmholtz; on la supposera "suffisamment régulière" au voisinage de  $\Gamma$ . Si la fonction  $\psi$  est solution de l'équation de Helmholtz dans le domaine borné  $D$  alors on a la formule de représentation intégrale suivante :*

$$\int_{\Gamma} \left( \psi(y) \frac{\partial G(x-y)}{\partial \mathbf{n}_y} - \frac{\partial \psi(y)}{\partial \mathbf{n}_y} G(x-y) \right) d\gamma(y) = \begin{cases} \psi(x) & \text{pour } x \in D, \\ 0 & \text{pour } x \in D', \end{cases}$$

où  $\mathbf{n}_y$  est la normale extérieure à  $D'$  avec  $D'$  le domaine complémentaire à  $D$ .

**Proposition .0.2.** : *Si la fonction  $\psi$  est solution (supposée encore régulière) de l'équation de Helmholtz dans le domaine extérieur  $D'$  du domaine borné  $D$  et satisfait la condition de rayonnement de Sommerfeld, alors la formule de représentation intégrale suivante est valable :*

$$\int_{\Gamma} \left( \psi(y) \frac{\partial G(x-y)}{\partial \mathbf{n}_y} - \frac{\partial \psi(y)}{\partial \mathbf{n}_y} G(x-y) \right) d\gamma(y) = \begin{cases} -\psi(x) & \text{pour } x \in D', \\ 0 & \text{pour } x \in D, \end{cases}$$

où  $\mathbf{n}_y$  est la normale extérieure à  $D'$ .

**Proposition .0.3.** : Si la fonction  $\psi$  est solution de l'équation de Helmholtz séparément dans  $D$  et  $D'$ , et satisfait la condition de rayonnement de Sommerfeld, alors pour tout  $x \in D \cup D'$

$$\psi(x) = \int_{\Gamma} \left( \left[ \frac{\partial \psi(y)}{\partial \mathbf{n}_y} \right] G(x-y) - [\psi(y)] \frac{\partial}{\partial \mathbf{n}_y} G(x-y) \right) d\gamma(y).$$

## Cas de l'équation de Maxwell

Considérons un domaine  $D$  borné, de frontière  $\Gamma$  régulière et  $\mathbf{n}$  la normale unitaire orientée vers l'extérieur de  $D$ . Soit  $\mathcal{G}_t$  la fonction de Green dyadique de l'équation de Maxwell

$$\mathbf{rot} \mathbf{rot} \mathbf{E} - t^{-1} \nabla(\operatorname{div} \mathbf{E}) - k_s^2 \mathbf{E} = 0 \quad (14)$$

définie par  $\mathcal{G}_t = G_{k_s} I + \frac{1}{k_s^2} \operatorname{Hess}(G_{k_s} - G_{k_p})$ ,  $I$  est la matrice identité  $3 \times 3$ ,  $k_p = k_s \sqrt{t}$ , et  $G_k$  est la fonction de Green associée à l'équation de Helmholtz définie par

$$G_k(x, y) = \frac{e^{ik||x-y||}}{4\pi||x-y||}.$$

Si  $t = \infty$  (resp.  $t \neq \infty$ ) on retrouve alors l'équation de Maxwell classique (resp. l'équation de Maxwell régularisée). D'après [13] on a les formules de représentations intégrales suivantes :

**Proposition .0.4.** Soit  $\mathbf{E} \in H(\mathbf{rot}, D) \cap H(\operatorname{div}, D)$  vérifiant au sens de distribution l'équation (14) dans  $D$ . On a alors la formule de représentation intégrale suivante :

$$\begin{aligned} & \int_{\Gamma} \mathcal{G}_t(x, y) (\mathbf{rot} \mathbf{E}(y) \wedge \mathbf{n}(y) + t^{-1} \mathbf{n}(y) \operatorname{div} \mathbf{E}(y)) d\gamma(y) \\ & - \int_{\Gamma} \mathbf{rot}_y \mathcal{G}_t(x, y) (\mathbf{E}(y) \wedge \mathbf{n}(y)) d\gamma(y) - t^{-1} \int_{\Gamma} \operatorname{div}_y \mathcal{G}_t(x, y)^T (\mathbf{E}(y) \cdot \mathbf{n}(y)) d\gamma(y) \\ & = \begin{cases} \mathbf{E}(x) & \text{pour } x \in D, \\ 0 & \text{pour } x \in D' = \mathbb{R}^3 \setminus \overline{D}. \end{cases} \end{aligned} \quad (15)$$

**Proposition .0.5.** Soit  $\mathbf{E} \in H_{loc}(\mathbf{rot}, D') \cap H_{loc}(\operatorname{div}, D')$  vérifiant au sens de distribution l'équation (14) dans  $D'$  et satisfait les conditions de rayonnement de (4.4). On a alors la formule de représentation intégrale suivante :

$$\begin{aligned} & \int_{\Gamma} \mathcal{G}_t(x, y) (\mathbf{rot} \mathbf{E}(y) \wedge \mathbf{n}(y) + t^{-1} \mathbf{n}(y) \operatorname{div} \mathbf{E}(y)) d\gamma(y) \\ & - \int_{\Gamma} \mathbf{rot}_y \mathcal{G}_t(x, y) (\mathbf{E}(y) \wedge \mathbf{n}(y)) d\gamma(y) - t^{-1} \int_{\Gamma} \operatorname{div}_y \mathcal{G}_t(x, y)^T (\mathbf{E}(y) \cdot \mathbf{n}(y)) d\gamma(y) \\ & = \begin{cases} 0 & \text{pour } x \in D, \\ -\mathbf{E}(x) & \text{pour } x \in D' = \mathbb{R}^3 \setminus \overline{D}. \end{cases} \end{aligned} \quad (16)$$

**Remarque .0.1.** Dans ces formules, les quantités  $\mathbf{E} \cdot \mathbf{n}$  et  $\mathbf{E} \wedge \mathbf{n}$  sont définies dans  $H^{-1/2}(\Gamma)$  et en ajoutant l'hypothèse que

$$\text{rot rot } \mathbf{E} - t^{-1} \nabla(\text{div } \mathbf{E}) \in L^2(D)$$

l'expression  $\text{rot } \mathbf{E} \wedge \mathbf{n} + t^{-1} \mathbf{n} \text{ div } \mathbf{E}$  se trouve dans  $H^{-1/2}(\Gamma)$ . Par abus de notation, on garde les intégrales plutôt que les crochets de dualité entre  $H^{-1/2}(\Gamma)$  et  $H^{1/2}(\Gamma)$ .

## Fonctions spéciales

Nous introduisons dans cette partie les fonctions de Bessel et pour plus de détails, nous invitons le lecteur à consulter ([63]). On appelle équation de Bessel l'équation

$$z^2 \frac{d^2 F}{dz^2} + z \frac{dF}{dz} + (z^2 - \nu^2) F = 0, \quad (17)$$

où  $\nu$  est un nombre arbitraire réel ou complexe. On définit les fonction suivantes :

– Fonctions de Bessel de premier espèce :

$$J_\nu(z) = \sum_{p=0}^{\infty} \frac{(-1)^p}{p! \Gamma(p + \nu + 1)} \left(\frac{z}{2}\right)^{2p+\nu},$$

où  $\Gamma(z)$  est la fonction Gamma ([63]). En particulier, pour  $\nu$  entier,

$$J_\nu(z) = \sum_{p=0}^{\infty} \frac{(-1)^p}{p!(\nu + p)!} \left(\frac{z}{2}\right)^{2p+\nu}.$$

– Fonctions de Bessel de deuxième espèce pour  $\nu$  non entier :

$$Y_\nu(z) = \frac{J_\nu(z) \cos(\nu\pi) - J_{-\nu}(z)}{\sin(\nu\pi)}. \quad (18)$$

Pour  $\nu$  entier, le terme de droite de l'expression (18) est remplacé par une limite. Les fonctions de Hankel  $H_\nu^1$  (resp.  $H_\nu^2$ ) de première espèce (resp. de deuxième espèce) sont définies à partir des fonctions de Bessel :

$$H_\nu^1 = J_\nu + i Y_\nu,$$

$$H_\nu^2 = J_\nu - i Y_\nu.$$

Les fonctions  $J_\nu$ ,  $Y_\nu$ ,  $H_\nu^1$  et  $H_\nu^2$  sont solutions de l'équation (17). D'autre part, on peut définir les fonctions de Bessel sphériques qui sont introduites à partir des fonctions de Bessel classiques. On désigne par  $h_\nu^1$  (resp.  $h_\nu^2$ ) la fonction de Hankel sphérique de première espèce (resp. de deuxième espèce) et  $j_\nu$  (resp.  $y_\nu$ ) la fonction de Bessel sphérique de première espèce (resp. de deuxième espèce). La fonction  $y_\nu$  est aussi appelée fonction de Neumann.

$$j_\nu(z) = \sqrt{\frac{\pi}{2z}} J_{\nu+\frac{1}{2}} \quad \text{et} \quad y_\nu(z) = \sqrt{\frac{\pi}{2z}} Y_{\nu+\frac{1}{2}}.$$

De plus,  $\forall \nu \in \mathbb{N}$ , on a

$$j_\nu(z) = \sum_{p=0}^{\infty} \frac{(-1)^p z^{\nu+2p}}{2^p p! 1.3 \cdots (2\nu + 2p + 1)},$$

$$y_\nu(z) = -\frac{(2\nu)!}{2^\nu \nu!} \sum_{p=0}^{\infty} \frac{(-1)^p z^{2p-\nu-1}}{2^p p! (-2\nu + 1)(-2\nu + 3) \cdots (-2\nu + 2p - 1)}.$$

Les fonctions  $h_\nu^1$  et  $h_\nu^2$  sont définies comme suit :

$$h_\nu^1 = j_\nu + i y_\nu,$$

$$h_\nu^2 = j_\nu - i y_\nu.$$

Pour  $\nu = 0$ ,  $j_0(z) = \frac{\sin z}{z}$  et  $y_0(z) = -\frac{\cos(z)}{z}$ . Ceci implique que

$$h_0^1(z) = \frac{e^{iz}}{iz}.$$

Chacune de ces fonctions vérifie les relations de récurrence suivante :

$$f_{\nu+1}(z) + f_{\nu-1}(z) = \frac{2\nu + 1}{z} f_\nu(z), \quad \nu \in \mathbb{N}^*,$$

$$f_{\nu+1}(z) = -z^\nu \frac{d}{dz} (z^{-\nu} f_\nu(z)), \quad \nu \in \mathbb{N},$$

$$f'_\nu(z) = f_{\nu-1}(z) - \frac{\nu + 1}{z} f_\nu(z), \quad \nu \in \mathbb{N}^*.$$

Dans le corps de ce manuscrit, nous désignons par  $h_\nu$  la fonction de Hankel sphérique de première espèce  $h_\nu^1$ .

## Relations vectorielles

### Identités vectorielles

$$\mathbf{a} \wedge \mathbf{b} = -\mathbf{b} \wedge \mathbf{a}$$

$$\mathbf{a} \cdot (\mathbf{b} \wedge \mathbf{c}) = (\mathbf{a} \wedge \mathbf{b}) \cdot \mathbf{c} = (\mathbf{c} \wedge \mathbf{a}) \cdot \mathbf{b}$$

### Identités différentielles

$$\mathbf{rot}(\nabla u) = 0,$$

$$\mathbf{div}(\mathbf{rot} \mathbf{v}) = 0,$$

$$\mathbf{div}(u\mathbf{v}) = \nabla u \cdot \mathbf{v} + u \mathbf{div} \mathbf{v},$$

$$\mathbf{rot}(u\mathbf{v}) = u \mathbf{rot} \mathbf{v} + \nabla u \wedge \mathbf{v},$$

$$\mathbf{rot}(\mathbf{v} \wedge \mathbf{w}) = \mathbf{v}(\operatorname{div} \mathbf{w}) - \mathbf{w}(\operatorname{div} \mathbf{v}) + (\mathbf{w} \cdot \nabla) \mathbf{v} - \mathbf{v}(\nabla \cdot \mathbf{w}),$$

$$\mathbf{rot} \operatorname{rot} \mathbf{v} = \nabla(\operatorname{div} \mathbf{v}) - \Delta \mathbf{v},$$

$$\operatorname{div}(\mathbf{v} \wedge \mathbf{w}) = \mathbf{w} \cdot \operatorname{rot} \mathbf{v} - \mathbf{v} \cdot \operatorname{rot} \mathbf{w}.$$

# Annexe B :

## Introduction à la méthode de Schwarz

La méthode de décomposition de domaine repose sur l'idée que le domaine de calcul est découpé en  $n$  sous domaines, notés  $\Omega_i$ ,  $i = 1, \dots, n$ . On distingue deux types de décomposition de domaine selon la décomposition de  $\Omega$  en  $\Omega_i$ . On dit que la décomposition de domaine est sans recouvrement si l'intersection entre les sous domaines se limite aux interfaces. Elle est avec recouvrement si chaque sous-domaine recouvre une partie de ses sous-domaines voisins. Le problème initial est alors reformulé sur chaque domaine, conduisant ainsi à une famille de sous problèmes de taille plus petite. Ils sont couplés les uns aux autres par les inconnues sur les interfaces des sous-domaines. Parmi les ouvrages de référence sur la méthode de décomposition de domaine, on cite par exemple le livre [64].

Nous nous intéressons dans notre étude à la méthode de Schwarz qui est une méthode de décomposition de domaine avec recouvrement. Elle a été initialement proposée par H. A. Schwarz en 1869 dans le but de trouver la solution analytique d'un problème de Poisson sur un domaine borné  $\Omega$  réunion d'un disque et d'un rectangle qui se recouvrent (voir Fig. 11).

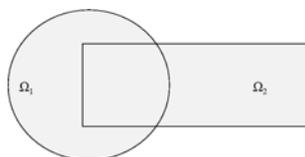


FIGURE 11 – Domaine de calcul  $\Omega = \Omega_1 \cup \Omega_2$

On considère le problème de Poisson suivant :

$$\begin{cases} -\Delta u = f \text{ dans } \Omega, \\ u = 0 \text{ sur } \partial\Omega, \end{cases} \quad (19)$$

On note dans la suite  $\Gamma_{12} = \partial\Omega_1 \setminus (\partial\Omega \cap \Omega_1)$  et  $\Gamma_{21} = \partial\Omega_2 \setminus (\partial\Omega \cap \Omega_2)$ . On fixe  $u_2^0$  sur  $\Gamma_{12}$  et pour  $n \geq 0$ , H. A. Schwarz a résolu le problème (19) en résolvant alternativement ces

deux sous problèmes

$$\begin{cases} -\Delta u_1^{n+1} = f \text{ dans } \Omega_1, \\ u_1^{n+1} = 0 \text{ sur } \partial\Omega_1 \cap \partial\Omega, \\ u_1^{n+1} = u_2^n \text{ sur } \Gamma_{12}, \end{cases} \quad (20)$$

et

$$\begin{cases} -\Delta u_2^{n+1} = f \text{ dans } \Omega_2, \\ u_2^{n+1} = 0 \text{ sur } \partial\Omega_2 \cap \partial\Omega, \\ u_2^{n+1} = u_1^{n+1} \text{ sur } \Gamma_{21}. \end{cases} \quad (21)$$

En fait, son idée consistait à découper le domaine  $\Omega$  en deux sous-domaines (un rectangle et un disque) pour lesquels il connaissait une solution analytique du problème de Laplace ((20) et (21)). En suivant l'algorithme mentionné précédemment, il a déterminé la solution analytique sur le domaine tout entier  $\Omega$ . H. A. Schwarz a montré la convergence de cet algorithme à l'aide du principe du maximum. P. L. Lions a donné une autre démonstration en se basant sur une interprétation variationnelle. Cet algorithme est dit "algorithme de Schwarz alterné" ou "algorithme de Schwarz multiplicatif". On distingue aussi une autre version de la technique de Schwarz : c'est la version additive qui consiste à fixer  $u_1^0, u_2^0$  et résoudre alternativement

$$\begin{cases} -\Delta u_1^{n+1} = f \text{ dans } \Omega_1, \\ u_1^{n+1} = 0 \text{ sur } \partial\Omega_1 \cap \partial\Omega, \\ u_1^{n+1} = u_2^n \text{ sur } \Gamma_{12}, \end{cases} \quad (22)$$

et

$$\begin{cases} -\Delta u_2^{n+1} = f \text{ dans } \Omega_2, \\ u_2^{n+1} = 0 \text{ sur } \partial\Omega_2 \cap \partial\Omega, \\ u_2^{n+1} = u_1^n \text{ sur } \Gamma_{21}. \end{cases} \quad (23)$$

La différence entre ces deux algorithmes réside dans le choix des conditions de bord. Pour l'algorithme de Schwarz multiplicatif, on prend comme condition de Dirichlet les dernières valeurs d'interface calculées par le sous domaine voisin. Par contre, pour l'algorithme de Schwarz additif, on prend comme condition de Dirichlet pour un sous domaine donné les valeurs données par le domaine voisin à l'itération précédente. Ceci donne une résolution totalement indépendante dans chaque sous domaine.

# Annexe C :

## Outils intégrés dans MÉLINA++

La mise en œuvre de la méthode **CEFRI** pour le problème auquel nous nous intéressons nécessite tout d'abord l'intégration dans MÉLINA++ des intégrands éléments finis et des noyaux de Green avec et sans relèvement. On désigne par  $i$  l'indice pour l'inconnue en ligne (fonction test) notée par  $v$  si elle est scalaire et par  $\mathbf{V}$  si elle est vectorielle. On note par  $j$  l'indice pour l'inconnue en colonne (fonction inconnue) notée par  $u$  si elle est scalaire et par  $\mathbf{U}$  si elle est vectorielle. On désigne par  $D$  un domaine d'intégration qui peut être soit un volume, soit une surface. Les intégrands éléments finis que nous avons implémenté dans MÉLINA++ sont :

- `EI_idUNCrossV` matrice élémentaire dont chaque coefficient  $ij$  a pour intégrand

$$w_i w_j \begin{pmatrix} 0 & n_3 & -n_2 \\ -n_3 & 0 & n_1 \\ n_2 & -n_1 & 0 \end{pmatrix}$$

cela correspond à  $\int_D u (\mathbf{n} \wedge \mathbf{V})$ .

- `EI_NcrossUIdV` matrice élémentaire dont chaque coefficient  $ij$  a pour intégrand

$$w_i w_j \begin{pmatrix} 0 & -n_3 & n_2 \\ n_3 & 0 & -n_1 \\ -n_2 & n_1 & 0 \end{pmatrix}$$

cela correspond à  $\int_D (\mathbf{n} \wedge \mathbf{U}) v$ .

- `EI_idUNCrossNCrossV` matrice élémentaire dont chaque coefficient  $ij$  a pour intégrand

$$w_i w_j \begin{pmatrix} -n_2^2 - n_3^2 & n_1 n_2 & n_1 n_3 \\ n_1 n_2 & -n_2^2 - n_3^2 & n_2 n_3 \\ n_1 n_3 & n_2 n_3 & -n_1^2 - n_2^2 \end{pmatrix}$$

cela correspond à  $\int_D u (\mathbf{n} \wedge (\mathbf{n} \wedge \mathbf{V}))$ .

- `EI_iduNdotV` matrice élémentaire dont chaque coefficient  $ij$  est vectoriel et a pour intégrand

$$w_i w_j \mathbf{n}$$

cela correspond à  $\int_D u (\mathbf{n} \cdot \mathbf{V})$  et à  $\int_D (\mathbf{n} \cdot \mathbf{U})v$ . Ceci provient du fait que dans MÉLINA++ il n'y a pas de distinction entre vecteur ligne et vecteur colonne.

- `EI_CurlUidV` matrice élémentaire dont chaque coefficient  $ij$  a pour intégrand

$$\begin{pmatrix} 0 & -w_i \partial_3 w_j & w_i \partial_2 w_j \\ w_i \partial_3 w_j & 0 & -w_i \partial_1 w_j \\ -w_i \partial_2 w_j & w_i \partial_1 w_j & 0 \end{pmatrix}$$

cela correspond à  $\int_D (\mathbf{rot} \mathbf{U}) v$ . La quantité  $\partial_l w_j$  désigne la dérivée partielle de  $w_j$  par rapport à la  $l^{\text{ième}}$  composante ( $l = 1, 2, 3$ ).

On approche les noyaux de Green par une interpolation sur les fonctions de base éléments finis définie sur les nœuds de  $\Sigma \times \Gamma$  ou  $\Sigma \times \Omega$ . On note  $N^\Sigma$  (resp.  $N^\Gamma$ ) le nombre de points sur le bord  $\Sigma$  (resp.  $\Gamma$ ) et  $N^\Omega$  est le nombre de points total. On désigne par  $G_\alpha$  les noyaux de Green et par  $\mathcal{R}G_\alpha$  les noyaux de Green avec relèvement. Les  $\mathcal{R}G_\alpha$  sont définis comme suit :

$$\mathcal{R}G_\alpha(x, y) = G_\alpha(x, y) f(y)$$

avec  $f$  est une fonction de relèvement. Dans MÉLINA++, on a fait le choix d'un relèvement linéaire pour un obstacle sphérique de rayon  $R_\Gamma = 1$  et de frontière artificielle sphérique de rayon  $R_\Sigma$ . Autrement dit, nous considérons dans ce cas la fonction  $f$  définie par :

$$f : \mathbb{R}^3 \rightarrow \mathbb{R}$$

$$y \mapsto \begin{cases} 1 & \text{si } \|y\| \in [R_\Gamma, R_\Gamma + \varepsilon[, \\ \frac{1}{\varepsilon - \delta} \|y\| - \frac{1}{\varepsilon - \delta} (R_\Gamma + \delta) & \text{si } \|y\| \in [R_\Gamma + \varepsilon, R_\Gamma + \delta[, \\ 0 & \text{si } \|y\| \in [R_\Gamma + \delta, R_\Sigma]. \end{cases}$$

où  $\delta$  résulte de la multiplication de la distance séparant les deux frontières  $\Sigma$  et  $\Gamma$  par un taux choisi par l'utilisateur qui fixe le domaine de relèvement. On approche les noyaux de Green définis sur les nœuds de  $\Sigma \times \Gamma$  par :

$$G(x, y) \approx \sum_{\sigma} \sum_{\gamma} G(x_\sigma, y_\gamma) w_\sigma(x) w_\gamma(y).$$

où l'indice  $\sigma$  (resp.  $\gamma$ ) est générique des nœuds de  $\Sigma$  (resp.  $\Gamma$ ). De même, on approche les noyaux de Green de relèvement définis sur les nœuds de  $\Sigma \times \Omega$  par :

$$\mathcal{R}G(x, y) \approx \sum_{\omega} \sum_{\gamma} \mathcal{R}G(x_\sigma, y_\omega) w_\sigma(x) w_\omega(y).$$

où l'indice  $\omega$  est générique des nœuds de  $\Omega$ . Les  $G(x_\sigma, y_\gamma)$  intégrés dans MÉLINA++ sont :

```
NV_Div_xDyadicDiv_yK, NV_DyadicDiv_yK,
NV_DyadicRot_yK, NV_Id3plusgrad_xgrad_yK,
NV_Rot_xDyadicRot_yK, NV_grad_xK,
NV_grad_xgrad_yK, NV_grad_yK,
NV_DyadicRot_xK.
```

Les  $\mathcal{R}G(x_\sigma, y_\omega)$  sont :

```
NV_Div_x_DyadicRot_y_Relev_yK,
NV_Div_x_Rot_yDyadicRot_y_Relev_yK,
NV_DyadicDiv_xRelev_yK,
NV_DyadicDiv_yRelev_yK,
NV_DyadicRot_xRelev_yK,
NV_DyadicRot_yRelev_yK,
NV_Id3plusgrad_x_grad_yRelev_yK,
NV_Rot_xDyadicRot_yRelev_yK,
NV_Rot_xRot_yDyadicRot_yRelev_yK,
NV_Rot_x_DyadicDiv_y_Relev_yK,
NV_Rot_yDyadicRot_yRelev_yK.
```

Les noms de  $G(x_\sigma, y_\gamma)$  et  $\mathcal{R}G(x_\sigma, y_\omega)$  sont structurés de telle sorte à définir le noyau de Green. Par exemple,  $NV\_DyadicRot\_xK$  correspond au noyau de Green  $\mathbf{rot}_x \mathcal{G}_t(x_\sigma, y_\gamma)$ .

On détaille cette interpolation pour la décomposition des termes de couplage par les deux exemples suivants :

- Premier exemple :

$$\begin{aligned} & \int_{\Sigma} \int_{\Omega} \mathbf{rot}_x(\mathcal{R}\mathcal{G}_t(x, y)) \mathbf{E}(y) d\Omega(y) \cdot (\mathbf{n}_\sigma(x) \wedge \mathbf{E}'(x)) d\sigma(x) \\ & \approx \sum_{\sigma=1}^{N^\Sigma} \left[ \left( \int_{\Sigma} w_\sigma(x) w_i(x) (\mathbf{n}_\sigma(x) \wedge \mathbf{E}_i(x)) d\sigma(x) \right)^T \right. \\ & \quad \left. \left( \sum_{\omega=1}^{N^\Omega} \mathbf{rot}_x(\mathcal{R}\mathcal{G}_t(x_\sigma, y_\omega)) \sum_{j=1}^{N^\Omega} \int_{\Omega} w_\omega(y) w_j(y) \mathbf{E}_j(y) d\gamma(y) \right) \right] \end{aligned} \quad (24)$$

Cet exemple fait intervenir l'intégrand  $EI\_idUNCrossV$  et le noyau de Green  $NV\_DyadicRot\_xRelev\_yK$ . Le facteur  $\int_{\Sigma} w_\sigma(x) w_i(x) (\mathbf{n}_\sigma(x) \wedge \mathbf{E}_i(x)) d\sigma(x)$

correspond au terme  $\int_D u (\mathbf{n} \wedge \mathbf{V})$  où  $w_\sigma$  joue le rôle de la fonction test et  $w_j$  joue le rôle de la fonction inconnue par rapport à la définition de l'intégrand.

- Deuxième exemple :

$$\begin{aligned}
& \int_{\Sigma} \int_{\Gamma} \operatorname{div}_y \mathcal{G}_t(x, y)^T \mathbf{E}(y) \cdot \mathbf{n}(y) d\gamma(y) \cdot (\mathbf{n}_{\sigma}(x) \wedge (\mathbf{n}_{\sigma}(x) \wedge \mathbf{E}'(x))) d\sigma(x) \\
& \approx \sum_{\sigma=1}^{N^{\Sigma}} \left[ \left( \int_{\Sigma} w_{\sigma}(x) w_i(x) (\mathbf{n}_{\sigma}(x) \wedge (\mathbf{n}_{\sigma}(x) \wedge \mathbf{E}_i(x))) d\sigma \right) \right. \\
& \quad \left. \left( \sum_{\gamma=1}^{N^{\Gamma}} \operatorname{div}_y \mathcal{G}_t(x_{\sigma}, y_{\gamma})^T \sum_{j=1}^{N^{\Gamma}} \int_{\Gamma} w_{\gamma}(y) w_j(y) \mathbf{E}_j(y) \cdot \mathbf{n}(y) d\gamma(y) \right) \right] \quad (25)
\end{aligned}$$

Cet exemple fait intervenir l'intégrand `EI_idUNCrossNcrossV`, le noyau de Green `NV_DyadicDiv_yK` et l'intégrand `EI_iduNdotV`. Le facteur  $\int_{\Sigma} w_{\sigma}(x) w_i(x) (\mathbf{n}_{\sigma}(x) \wedge (\mathbf{n}_{\sigma}(x) \wedge \mathbf{E}_i(x))) d\sigma(x)$  correspond au terme  $\int_D u (\mathbf{n} \wedge (\mathbf{n} \wedge \mathbf{V}))$  où  $w_{\sigma}$  joue le rôle de la fonction inconnue et  $w_i$  joue le rôle de la fonction test par rapport à la définition de l'intégrand. Le facteur  $\int_{\Gamma} w_{\gamma}(y) w_j(y) \mathbf{E}_j(y) \cdot \mathbf{n}(y) d\gamma(y)$  correspond au terme  $\int_D (\mathbf{n} \cdot \mathbf{U})v$  où  $w_{\gamma}$  joue le rôle de la fonction test et  $w_j$  joue le rôle de la fonction inconnue par rapport à la définition de l'intégrand.

Par la discrétisation éléments finis et par l'interpolation des noyaux de Green, la matrice  $\bar{C}$  peut s'écrire alors sous forme du produit suivant

$$\bar{C} = \sum_{\alpha} C_{\alpha} P_{\Sigma}^t M_{\alpha\Sigma} \mathcal{G}_{\alpha\Sigma\Gamma} M_{\alpha\Gamma} P_{\Gamma} + \sum_{\beta} d_{\beta} P_{\Sigma}^t M_{\beta\Sigma} \mathcal{R}\mathcal{G}_{\beta\Sigma\Omega} M_{\beta\Omega} \quad (26)$$

où  $P_{\Gamma}$  (resp.  $P_{\Sigma}$ ) est la matrice de projection sur  $\Gamma$  (resp. sur  $\Sigma$ ), et  $M_{\beta\Gamma}$  (resp.  $M_{\alpha\Sigma}$ ) la matrice de masse sur  $\Gamma$  (resp. sur  $\Sigma$ ).  $\mathcal{G}_{\alpha\Sigma\Gamma}$  (resp.  $\mathcal{R}\mathcal{G}_{\beta\Sigma\Gamma}$ ) les matrices noyaux de coefficients  $G_{\alpha}(x_{\sigma}, y_{\gamma})$  (resp.  $\mathcal{R}\mathcal{G}_{\beta\Sigma\Gamma}$ ).  $c_{\alpha}$  et  $d_{\beta}$  étant des constantes dépendant du nombre d'onde.

**Remarque .0.2.** *Le choix fait du relèvement dans MÉLINA++ est limité au cas d'un obstacle sphérique entouré d'un bord fictif sphérique. Il est intéressant de généraliser la formule du relèvement choisie pour n'importe quel diffracteur.*

## Programmation à l'aide de MÉLINA++

Le programme principal du code **CEFRI** est structuré comme suit :

- lire les paramètres qui sont donnés par l'utilisateur (par exemple le nombre d'onde) à partir d'un fichier de donnée,
- créer d'une liste de paramètres, objet de la classe `me_parameterList` contenant les paramètres nécessaires au calcul du second membre,
- définir la fonction de Green comme objet de la classe `GreenFunction`,
- définir le domaine, le maillage et choisir l'espace d'approximation (les éléments finis),

- déclarer les `TermMatrix` qui seront calculés par MÉLINA++ ; pour cela on utilise deux classes :  
`FETermMatrix` pour la définition des matrices élément finis, constituants de la matrice  $\bar{\bar{A}}$ ,  
`NVTermMatrix` pour définir les matrices  $\mathcal{G}_{\Sigma\Gamma}$  et  $\mathcal{R}\mathcal{G}_{\Sigma\Omega}$  contenant les valeurs des noyaux de Green,
- définir le second membre sur le domaine  $\Gamma$  en déclarant tout d'abord un objet de la classe `me_function`,
- déclarer les vecteurs de la classe `NVTermVector` qui seront utilisés pour la restriction (sur les nœuds de  $\Gamma$  ou  $\Sigma$ ) ou l'extension sur tous les nœuds du maillage d'un vecteur connu sur un bord (méthode `resphaeas`),
- calculer les termes déclarés,
- résoudre le système linéaire.

Le but est de résoudre le système (5.2) à l'aide d'une méthode itérative. La stratégie généralement utilisée est de sommer les matrices  $\bar{\bar{A}}$  et  $\bar{\bar{C}}$  puis d'utiliser un solveur itératif avec préconditionnement. Dans ce travail, nous avons résolu (5.2) sans avoir recours à cette opération. A terme, il s'agit d'appliquer la méthode FMM à la matrice  $\bar{\bar{C}}$  sans avoir besoin de modifier la matrice du système linéaire  $\bar{\bar{A}} + \bar{\bar{C}}$ . Le calcul du produit de la matrice  $\bar{\bar{C}}$  par un vecteur donné est réalisé sans stocker la matrice complète. Ceci est fait en suivant la stratégie (26). Par exemple, pour le cas d'une matrice contenant un noyau de Green sans relèvement, on applique la procédure suivante :

- restreindre l'inconnue  $\bar{\bar{\mathbf{E}}}$  à  $\Gamma$  :  $\bar{\bar{\mathbf{E}}}_{\Gamma} = P_{\Gamma} \bar{\bar{\mathbf{E}}}$  (fonction `reshapeas`),
- multiplier la matrice  $M_{\alpha\Gamma}$  par  $\bar{\bar{\mathbf{E}}}_{\Gamma}$ ,
- multiplier la matrice des noyaux de Green  $\mathcal{G}_{\alpha\Sigma\Gamma}$  par le vecteur résultat ( $M_{\beta\Gamma} P_{\Gamma} \bar{\bar{\mathbf{E}}}$ ),
- multiplier la matrice  $M_{\alpha\Sigma}$  au vecteur résultat ( $\mathcal{G}_{\alpha\Sigma\Gamma} M_{\alpha\Gamma} P_{\Gamma} \bar{\bar{\mathbf{E}}}$ ),
- étendre le vecteur  $M_{\alpha\Sigma} \mathcal{G}_{\alpha\Sigma\Gamma} M_{\beta\Gamma} P_{\Gamma} \bar{\bar{\mathbf{E}}}$  sur le domaine  $\Omega$  (fonction `reshapeas`).



# Bibliographie

- [1] A. Bendali and L. Halpern. Approximation par troncature de domaine de la solution du problème aux limites extérieur en régime sinusoïdal. *C. R. Acad. Sci. Paris Ser.I, Math.*, 294 :557–560, 1982.
- [2] A. Bayliss, M. Gunzburger, and E. Turkel. Boundary conditions for the numerical solution of elliptic equations in exterior regions. *SIAM J. Numer. Anal.*, 42(2) :430–451, 1982.
- [3] X. Antoine, H. Barucq, and A. Bendali. Bayliss-turkel-like radiation condition on surfaces of arbitrary shape. *Journal of Mathematical Analysis and Applications*, 229 :184–211, 1999.
- [4] D. Rabia, F. Charbel, M. Antonini, and T. Radek. Finite element solution of two-dimensional acoustic scattering problems using arbitrarily shaped convex artificial boundaries. *J. Comput. Acoust.*, 8(1) :81–99, 2000. Finite elements for wave problems (Trieste, 1999).
- [5] R. Djellouli, C. Farhat, A. Macedo, and R. Tezaur. Three-dimensional finite element calculations in acoustic scattering using arbitrary shaped convex artificial boundaries. *Int. J. Numer. Meth.Engng.*, 53 :1461–1476, 2002.
- [6] J-P. Berenger. A perfectly matched layer for the absorption of electromagnetic waves. *Journal of Computational Physics*, 114(2) :185–200, 1994.
- [7] A. J. Burton and G. F Miller. The application of integral equation methods to the numerical solution of some exterior boundary-value problems. *Proc. Roy. Soc. London. Ser. A*, 323 :201–210, 1971. A discussion on numerical analysis of partial differential equations (1970).
- [8] D. Colton and R. Kress. *Inverse acoustic and electromagnetic scattering theory*, volume 93 of *Applied Mathematical Sciences*. Springer, New York, third edition, 2013.
- [9] J-C. Nédélec. *Acoustic and electromagnetic equations*, volume 144 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 2001. Integral representations for harmonic problems.
- [10] C. Johnson and J.C. Nédélec. On the coupling of boundary integral and finite element methods. *Math. Comp.*, 35(152) :1063–1079, 1980.
- [11] A. Jami and M. Lenoir. Formulation variationnelle pour le couplage entre une méthode éléments finis et une représentation intégrale. *C. R. Acad. Sci.*, 285(4) :269272, 1977.

- [12] J-M. Jin and J. Liu. A novel hybridization of higher order finite element and boundary integral methods for electromagnetic scattering and radiation problems. *IEEE Transactions on antennas and propagation*, 49(12) :1794–1806, 2001.
- [13] C. Hazard and M. Lenoir. On the solution of time-harmonic scattering problems for Maxwell’s equations. *SIAM J. Math. Anal.*, 27(6) :1597–1630, 1996.
- [14] F. Ben Belgacem, M. Fournié, N. Gmati, and F. Jelassi. On the Schwarz algorithms for the elliptic exterior boundary value problems. *M2AN Math. Model. Numer. Anal.*, 39(4) :693–714, 2005.
- [15] F. Ben Belgacem, N. Gmati, and F. Jelassi. Convergence bounds of GMRES with Schwarz’ preconditioner for the scattering problem. *Internat. J. Numer. Methods Engrg.*, 80(2) :191–209, 2009.
- [16] Y. Saad and M. H. Schultz. GMRES : a generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM J. Sci. Statist. Comput.*, 7(3) :856–869, 1986.
- [17] D. Martin. [http://anum-maths.univ-rennes1.fr/melina/melina++\\_distrib/](http://anum-maths.univ-rennes1.fr/melina/melina++_distrib/). Documentation melina++, Université Rennes 1, last update : Jul. 2012.
- [18] D. Martin. <http://anum-maths.univ-rennes1.fr/melina/>. Documentation melina, Université Rennes 1.
- [19] V. Rokhlin. Rapid solution of integral equations of scattering theory in two dimensions. *J. Comput. Phys.*, 86(2) :414–439, 1990.
- [20] V. Rokhlin. Diagonal forms of translation operators for the Helmholtz equation in three dimensions. *Appl. Comput. Harmon. Anal.*, 1(1) :82–93, 1993.
- [21] E. Darve. *Méthodes multipôles rapides : Résolution des équations de Maxwell par formulations intégrales*. Thèse de doctorat, Université Paris 6, 1999.
- [22] E. Darrigrand. *Couplage méthodes mutipôles-Discrétisation microlocale pour les équations intégrales de l’électromagnétisme*. Thèse de doctorat, Université Bordeaux I, 2002.
- [23] E. Darve and P. Havé. Efficient fast multipole method for low-frequency scattering. *J. Comput. Phys.*, 197(1) :341–363, 2004.
- [24] E. Darve and P. Havé. A fast multipole method for Maxwell equations stable at all frequencies. *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 362(1816) :603–628, 2004.
- [25] E. Darrigrand and Y. Lafranche. <https://sourcesup.cru.fr/projects/fastmmlib/> (accès restreint). Documentation fastMMLib, Université Rennes 1, in progress.
- [26] C. Hazard and M. Lenoir. Modélisation et résolution des problèmes de diffraction, cours de l’ensta et de dea de mécanique. *Cours de L’ENSTA et de DEA de Mécanique, Paris VI, ENSTA SMP, centre de l’Yvette, Palaiseau*, 1995.
- [27] L. Greengard and V. Rokhlin. The rapid evaluation of potential fields in three dimensions. In *Vortex methods (Los Angeles, CA, 1987)*, volume 1360 of *Lecture Notes in Math.*, pages 121–141.
- [28] L. Greengard and V. Rokhlin. A new version of the fast multipole method for the Laplace equation in three dimensions. In *Acta numerica, 1997*, volume 6 of *Acta Numer.*, pages 229–269.

- [29] H. Cheng, L. Greengard, and V. Rokhlin. A fast adaptative multipole algorithm in three dimensions. *J. Comput. Phys.*, 155(2) :468–498, 1999.
- [30] R.L. Wagner and W.C. Chew. A ray-propagation fast multipole algorithm. *Micro-wave Opt. Tech. Letter*, 7(10) :435–438, 1994.
- [31] S. Koc, J. Song, and W. C. Chew. Error analysis for the numerical evaluation of the diagonal forms of the scalar spherical addition theorem. *SIAM J. Numer. Anal.*, 36(3) :906–921, 1999.
- [32] G. Arfken and H. Weber. *Mathematical Methods for Physicists*. Academic Press, New York, 1985.
- [33] B. Albert and R. Jakob-Chien. A fast spherical filter with uniform resolution. *J. Comput. Phys.*, 136(7) :480–484, 1997.
- [34] A. Dutt, M. Gu, and V. Rokhlin. Fast algorithms for polynomial interpolation, integration, and differentiation. *SIAM J. Numer. Anal.*, 33(5) :1689–1711, 1996.
- [35] N. Yarvin and V. Rokhlin. A generalized one-dimensional fast multipole method with application to filtering of spherical harmonics. *J. Comput. Phys.*, 147(2) :594–609, 1998.
- [36] N. Yarvin and V. Rokhlin. An improved fast multipole algorithm for potential fields on the line. *SIAM J. Numer. Anal.*, 36(2) :629–666 (electronic), 1999.
- [37] H. Wallén and J. Sarvas. Translation procedures for broadband mlfma. *Progress In Electromagnetics Research.*, 55 :47–78, 2005.
- [38] I. Bogaert and F. Olyslager. A low frequency stable plane wave addition theorem. *Commun. Comput. Phys.*, 228 :1000–1016, 2009.
- [39] Y. G. Liu and W. C. Chew. A low frequency vector fast multipole algorithm with vector addition theorem. *Commun. Comput. Phys.*, 8(5) :1183–1207, 2010.
- [40] L. Greengard, J. Huang, and V. Rokhlin. Accelerating fast multipole methods for the helmholtz equation at low frequencies. In *IEEE Comput. Sci. Engng*, volume 5, pages 32–38.
- [41] G. H. Golub and C. F. Van Loan. *Matrix computations*. Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, Baltimore, MD, fourth edition, 2013.
- [42] N. Yarvin and V. Rokhlin. Generalized Gaussian quadratures and singular value decompositions of integral operators. *SIAM J. Sci. Comput.*, 20(2) :699–718 (electronic), 1998.
- [43] O.P. Bruno and L. A. Kunyansky. High-order algorithm for the solution of surface scattering problems : basic implementation, tests, and applications. *J. Comput. Phys.*, 169 :80–110, 2001.
- [44] O. P. Bruno and L. A. Kunyansky. Surface scattering in 3-d : an accelerated high order solver. *JProc. Roy. Soc. London Ser. A Comput. Phys.*, 457 :2921–2934, 2001.
- [45] C. Geuzaine, O. P. Bruno, and F. Reitich. On the  $o(1)$  solution of multiple-scattering problems. *IEEE Trans. Magnetics*, 41 :1488–1491, 2005.
- [46] W. Hackbusch. A sparse matrix arithmetic based on h-matrices. i. introduction to h-matrices. *Computing*, 62 :89–108, 1999.

- [47] L. Grasedyck and W. Hackbusch. Construction and arithmetics of h-matrices. *Computing*, 70 :295–334, 2003.
- [48] J. Ostrowski, M. Bebendorf, R. Hiptmair, and F. Kramer. H-matrix-based operator preconditioning for full maxwell at low frequencies. *IEEE Trans. Magn.*, 46(8) :3193–3196, 2010.
- [49] B. Khoromskij and V. Khoromskaia. <http://jsa2013.sciencesconf.org/resource/page/id/39>). slides, Max-Planck-Institute for Mathematics in the Sciences, Leipzig, 2013.
- [50] P. Monk. *Finite element methods for Maxwell's equations*. Numerical Mathematics and Scientific Computation. Oxford University Press, New York, 2003.
- [51] N. Zerbib, F. Collino, and F. Millot. Etude tridimensionnelle de la condition absorbante adaptative proposée par liu et jin pour la résolution des problèmes de diffraction. *CERFACS, Report TR/EMC/05/81*, 2005.
- [52] F. Jelassi. *Méthodes de Schwarz pour les problèmes extérieurs. Application au calcul des courants de Foucault en électrotechnique*. Thèse de doctorat, Université de Tunis El Manar et de l'Université de Toulouse III, 2006.
- [53] N. Gmati and B. Philippe. Comments on the GMRES convergence for preconditioned systems. In *Large-scale scientific computing*, volume 4818 of *Lecture Notes in Comput. Sci.*, pages 40–51. Springer, Berlin, 2008.
- [54] V. Girault and P.A. Raviart. *Finite element methods for Navier-Stokes equations*, volume 5 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 1986. Theory and algorithms.
- [55] M. Costabel. A coercive bilinear form for Maxwell's equations. *J. Math. Anal. Appl.*, 157(2) :527–541, 1991.
- [56] M. Costabel, M. Dauge, and S. Nicaise. Singularities of Maxwell interface problems. *M2AN Math. Model. Numer. Anal.*, 33(3) :627–649, 1999.
- [57] M. Costabel and M. Dauge. Maxwell and Lamé eigenvalues on polyhedra. *Math. Methods Appl. Sci.*, 22(3) :243–258, 1999.
- [58] M. Costabel, M. Dauge, D. Martin, and G. Vial. Weighted regularization of Maxwell equations : computations in curvilinear polygons. In *Numerical mathematics and advanced applications*, pages 273–280. Springer Italia, Milan, 2003.
- [59] M. Costabel and M. Dauge. Weighted regularization of Maxwell equations in polyhedral domains. A rehabilitation of nodal finite elements. *Numer. Math.*, 93(2) :239–277, 2002.
- [60] I. Moret. A note on the superlinear convergence of GMRES. *SIAM J. Numer. Anal.*, 34(2) :513–516, 1997.
- [61] H. Weyl. Das asymptotische Verteilungsgesetz der Eigenwerte linearer partieller Differentialgleichungen (mit einer Anwendung auf die Theorie der Hohlraumstrahlung). *Math. Ann.*, 71(4) :441–479, 1912.
- [62] C. Geuzaine and J.F. Remacle. A three-dimensional finite element mesh generator with built-in pre- and post-processing facilities. *International Journal for Numerical Methods in Engineering*, Issue 11(79) :1309–1331, 2009.

- [63] M. Abramowitz and IA. Stegun. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, volume 55 of *National Bureau of Standards Applied Mathematics Series*. For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C., 1964.
- [64] A. Quarteroni and A. Valli. *Domain decomposition methods for partial differential equations*. Numerical Mathematics and Scientific Computation. The Clarendon Press Oxford University Press, New York, 1999. Oxford Science Publications.

## **Résumé :**

Le travail effectué dans cette thèse a consisté à analyser différents aspects mathématiques et numériques d'une stratégie de résolution des problèmes de propagation d'onde acoustique et électromagnétique en domaine extérieur. Nous nous intéressons plus particulièrement à la méthode de couplage entre éléments finis et représentation intégrale (CEFRI) où nous analysons un algorithme de résolution itérative par analogie avec une méthode de décomposition de domaine ainsi que l'utilisation de la méthode multipôles rapide (FMM). Le système à résoudre fait intervenir des opérateurs intégraux ce qui rend crucial le recours à des méthodes rapides telles que la FMM. L'analogie avec une méthode de décomposition de domaine s'obtient par extension au problème de Maxwell des résultats établis par F. Ben Belgacem et al. pour le problème de Helmholtz posé en domaine non borné. Pour cela, nous avons montré le lien entre la méthode CEFRI et la méthode de Schwarz avec recouvrement total pour la résolution du problème de Maxwell en domaine non borné. Cette relecture de la méthode CEFRI offre également une technique de préconditionnement pour les solveurs de Krylov et nous a permis d'avoir une idée préliminaire sur la convergence de ces méthodes. Ainsi, nous nous intéressons plutôt à des méthodes itératives rapides. Pour cela, nous avons mené une analyse théorique afin de montrer la convergence superlinéaire du GMRES dans une configuration sphérique. La validation de ces aspects a été réalisée par l'enrichissement de nombreux intégrands de la librairie éléments finis MÉLINA++, en C++.

**Mots clés :** Problèmes de diffraction, Acoustique, Electromagnétisme, Représentation intégrale, Élément finis, Méthode de Schwarz, Méthodes multipôles rapides.

## **Abstract :**

We are concerned with the study of different aspects of a numerical strategy for the resolution of acoustic and electromagnetic scattering problems. We focus more particularly on a coupling of finite element and integral representation (CEFRI) : we study an iterative algorithm by analogy with a domain decomposition method, and consider the use of the Fast Multipole Method (FMM). The system to be solved involves integral operators which requires the use of fast methods such as the FMM. The correspondence with a domain decomposition method is obtained by extending to the exterior Maxwell problem the results derived by F. Ben Belgacem et al. for the Helmholtz problem posed in unbounded domain. To this aim, we show the analogy to the Schwarz method with total overlap. This interpretation of CEFRI suggests a preconditioner for Krylov solvers and enables us to have a preliminary idea of their convergence. We derive in this context an analytical proof of a superlinear convergence of GMRES in a spherical configuration. The validation of these aspects has been achieved by the enrichment of the finite element library MÉLINA++ in C++.

**Key words :** Diffraction problems, Acoustic, Electromagnetism, Integral representation, Finite element, Schwarz method, Fast Multipole Method.