

UNIVERSITÉ DE NICE-SOPHIA ANTIPOLIS - UFR Sciences
Ecole Doctorale Sciences Fondamentales et Appliquées

THÈSE

pour obtenir le titre de
Docteur en Sciences
de l'UNIVERSITÉ de Nice-Sophia Antipolis

Discipline : Mathématiques

présentée et soutenue par
HUGO FOL

MÉTHODES DE TYPE GALERKIN DISCONTINU POUR LA RÉOLUTION NUMÉRIQUE DES ÉQUATIONS DE MAXWELL 3D EN RÉGIME HARMONIQUE

Thèse dirigée par Stéphane LANTERI
soutenue le 22 Décembre 2006

Jury :

LUC GIRAUD,	professeur,	examineur
FRÉDÉRIC MAGOULES,	professeur,	rapporteur
BONIFACE NKONGA,	maître de conférence,	rapporteur
JEAN VIRIEUX,	professeur,	examineur
STÉPHANE LANTERI,	directeur de recherche,	directeur
CLAUDE DEDEBAN,	ingénieur,	invité

A mon épouse Sophie,

Ce qui ne tue pas rend plus fort,
extrait de *Ecce homo*, Friedrich NIETZSCHE

Remerciements

Je tiens en tout premier lieu à remercier les membres du jury. A ce titre, je remercie chaleureusement Messieurs Giraud, Virieux et Nkonga pour avoir participé à ma soutenance orale. J'ai eu le plaisir à cette occasion de faire la connaissance de Monsieur Luc Giraud, de retrouver Monsieur Jean Virieux en tant que président de jury, et de retrouver Monsieur Boniface Nkonga en tant que rapporteur après l'avoir eu comme enseignant à plusieurs reprises au cours de mon cursus universitaire. Je remercie aussi Monsieur Frédéric Magoulès pour avoir rapporté sur ma thèse, et Monsieur Claude Dedeban pour avoir accepté l'invitation à ma soutenance.

Je ne peux achever cette thèse sans exprimer ma gratitude envers Madame Katherine Nkonga et Monsieur Alain Bachelot pour leurs encouragements permanents et pour avoir toujours cru en moi.

Je pense ensuite aux membres du projet CAIMAN pour ces trois années passées ensemble, et en particulier à Grégory et Hassan qui sont devenus des amis. Hassan, par son appui et son amitié, m'a beaucoup aidé à surmonter les moments difficiles. Je suis aussi extrêmement reconnaissant envers Ronan Perrussel pour le temps passé sur ma thèse durant la dernière année à proposer des améliorations, à me conseiller judicieusement, et à la relire avec minutie à plusieurs reprises.

J'ai également une petite pensée envers mes parents pour leur confiance et leur soutien constant durant toutes mes années d'études.

J'aimerais enfin remercier mon directeur de thèse Stéphane Lanteri et mon épouse Sophie sans qui cette thèse n'aurait pu s'achever. Stéphane, par son investissement considérable et sa disponibilité permanente, m'a énormément aidé dans sa réalisation. Quant à Sophie, son réconfort, son appui et son comportement de tous les jours ont été un facteur déterminant dans la réussite de cette thèse. C'est un juste retour des choses si elle lui est dédiée.

Je souhaite bonne chance à tous les membres du projet CAIMAN pour la suite.

Table des matières

Introduction générale	9
L'électromagnétisme numérique et ses applications	9
Les difficultés de la modélisation numérique	9
Résolution numérique des équations de Maxwell	10
Résolution des systèmes linéaires	11
Objectif et plan de la thèse	12
1 Cadres mathématique et numérique	15
1.1 Modélisation de la propagation d'ondes électromagnétiques	15
1.1.1 Les équations de Maxwell	15
1.1.2 Relations phénoménologiques	17
1.1.3 L'équation des ondes	18
1.1.4 Symétrisation du système de Maxwell	18
1.1.5 Conditions aux limites et conditions d'interface	19
1.2 Cadre mathématique	21
1.2.1 Problème de la diffraction d'ondes	21
1.2.2 Hyperbolicité du système de Maxwell	21
1.2.3 Le régime harmonique	24
1.3 La discrétisation des équations de Maxwell	25
2 Méthodes de type Galerkin discontinu en régime harmonique	27
2.1 Problème aux limites et analyse préliminaire	28
2.1.1 Méthode de volumes finis	28
2.1.2 Méthode de type Galerkin discontinu	30
2.2 Équations de Maxwell 1D	34
2.2.1 Méthode de volumes finis	36
2.2.2 Méthode de type Galerkin discontinu	43
2.2.3 Résultats numériques	52
2.3 Équations de Maxwell 3D	57
2.3.1 Méthode de volumes finis	62
2.3.2 Méthode de type Galerkin discontinu	70
2.3.3 Étude du problème discret	76
2.4 Annexes	79
2.4.1 Équivalences des conditions aux limites 1D	79
2.4.2 Propriétés de la matrice N_{jk}	81

2.4.3	Décomposition de la matrice des flux	82
2.4.4	Équivalences des conditions aux limites 3D	83
2.4.5	Quelques propriétés des matrices antisymétriques	85
2.4.6	Blocs matriciels du schéma GD-P0	87
2.4.7	Blocs matriciels du schéma GD-P1	87
3	Méthodes DD pour les équations de Maxwell en régime harmonique	93
3.1	Introduction	93
3.2	Conditions d'interface naturelles	96
3.2.1	Équations de Maxwell 2D en mode TM	96
3.2.2	Algorithme de Schwarz	98
3.2.3	Mise en œuvre de l'algorithme de Schwarz sans recouvrement	111
3.2.4	Formulation d'un système interface	114
3.2.5	Résultats numériques	116
3.3	Conditions d'interface optimisées en 2D	122
3.3.1	Conditions absorbantes exactes	122
3.3.2	Algorithme de Schwarz avec conditions généralisées	123
3.3.3	Conditions d'interface approchées	125
3.3.4	Mise en œuvre de l'algorithme de Schwarz avec conditions optimisées	133
3.3.5	Résultats numériques	142
3.4	Annexe	142
3.4.1	Algorithme de Schwarz pour les équations de Maxwell 3D	142
4	Résultats numériques en 3D	151
4.1	Validation numérique	151
4.1.1	Propagation d'une onde plane dans le vide	151
4.1.2	Diffraction d'une onde plane par une sphère	152
4.1.3	Diffraction d'une onde plane par un cube	166
4.2	Résolution numérique en mode séquentiel	183
4.2.1	Résolution itérative non-préconditionnée	183
4.2.2	Résolution itérative préconditionnée	184
4.2.3	Résolution directe	188
4.3	Résolution numérique en mode parallèle	190
4.3.1	Propagation d'une onde plane dans le vide	190
4.3.2	Diffraction par une sphère parfaitement conductrice	193
	Conclusion générale	194
	Références bibliographiques	201

Introduction générale

L'électromagnétisme numérique et ses applications

La modélisation numérique est un outil essentiel pour l'ingénieur et son importance n'a cessé de croître avec l'augmentation des capacités de calcul et de stockage des ordinateurs et notamment l'avènement des calculateurs parallèles. Ce phénomène a été d'autant plus ressenti dans le domaine de l'électromagnétisme qu'il s'est accompagné de la mise au point de nouvelles méthodes numériques adaptées à ce type d'architecture de calcul, et que la simulation numérique dans ce domaine, traditionnellement réservée aux applications militaires comme la furtivité radar ou la vulnérabilité des systèmes d'arme, s'est aussi ouverte à un large spectre d'applications civiles.

Ainsi, l'électromagnétisme numérique est aujourd'hui une discipline en plein essor et force est de reconnaître que de nombreux défis qui paraissaient inaccessibles il y a encore dix ans sont maintenant relevés. On constate que le champ d'application de l'électromagnétisme numérique s'est élargi à des contextes aussi variés que l'électronique, les accélérateurs de particules, la magnétohydrodynamique, l'optimisation de forme d'antennes, la conception de dispositifs hyperfréquences ou la compatibilité électromagnétique. Ce dernier est particulièrement crucial car les nouveaux matériaux qui composent les avions, les bâtiments ou les voitures, n'assurent plus la protection des nombreux composants électroniques contre la foudre ou d'autres sources électromagnétiques. La santé au sens large est un autre contexte d'application de l'électromagnétisme numérique. Il s'agit essentiellement de quantifier numériquement l'absorption d'un champ électromagnétique dans des tissus biologiques, soit pour évaluer les éventuels effets nocifs de l'exposition à ces champs [Bernardi et al., 2001]-[Clatz et al., 2000], soit pour des besoins de planification du traitement par hyperthermie¹ de tumeurs cancéreuses [Lin et al., 2000]-[Siauve et al., 2003].

Les équations qui modélisent les problèmes de propagation d'ondes électromagnétiques, à savoir les équations de Maxwell, suscitent un engouement tant de la part des physiciens que de la part des mathématiciens, en raison notamment des applications industrielles pressenties. La simulation numérique a contribué à une meilleure compréhension des phénomènes physiques sous-jacents, motivant de nombreuses recherches afin de parvenir à des méthodes qui restituent au mieux les caractéristiques mathématiques de ces équations et l'universalité de leur application, avec à la clef un souci de performance.

Les difficultés de la modélisation numérique

Plusieurs difficultés émergent lorsque l'on tente de modéliser et de développer des méthodes de calcul pour simuler des phénomènes de propagation d'ondes électromagnétiques :

¹L'hyperthermie est un échauffement (local) des tissus biologiques.

- la plupart des phénomènes électromagnétiques modélisés requièrent un domaine de calcul non borné,
- la précision des méthodes d'approximation des équations d'une part, et la discrétisation du domaine de calcul d'autre part, doivent être compatibles avec le caractère ondulatoire des phénomènes de propagation d'ondes électromagnétiques,
- les nouveaux matériaux constituant les objets étudiés possèdent des caractéristiques électromagnétiques de plus en plus complexes où, par exemple, les hypothèses de linéarité et d'isotropie ne s'appliquent pas. Ces matériaux nécessitent donc des algorithmes robustes pour la résolution des équations de Maxwell,
- les applications d'intérêt industriel conduisent à la résolution numérique de systèmes discrets dont la taille, évaluée en termes du nombre d'inconnues pour atteindre une précision donnée, est très grande. Par exemple, les longueurs caractéristiques sont généralement de l'ordre de quelques dizaines de longueur d'onde mais peuvent dépasser la centaine, conduisant à des maillages qui peuvent contenir jusqu'à dix millions de maille pour des maillages volumiques. La résolution numérique de problèmes de cette taille ne peut se faire qu'en exploitant pleinement les possibilités des calculateurs parallèles. L'algorithme généré par la méthode doit donc avoir un haut degré de parallélisme.

Résolution numérique des équations de Maxwell

De nombreuses méthodes ont été développées pour la résolution numérique des équations de Maxwell mais il semble qu'aucune méthode ne soit prédominante, le choix étant déterminé essentiellement par le type d'application considéré. Il ne s'agit pas ici de proposer une revue exhaustive de ces méthodes mais plutôt de préciser quelques caractéristiques importantes qui motivent le choix d'une méthode particulière.

Ces méthodes peuvent être regroupées en plusieurs classes, suivant que l'on souhaite traiter les équations de Maxwell en domaine temporel en regardant l'évolution en temps du champ électromagnétique, ou les équations de Maxwell en domaine fréquentiel (ou en régime harmonique) en regardant cette fois-ci le comportement du champ électromagnétique lorsque le terme source suit une dépendance harmonique en temps. Bien que ces deux formulations aient un lien physique étroit (voir par exemple [Helluy, 1994]), les méthodes numériques développées pour leur approximation peuvent s'attaquer indistinctement aux deux systèmes d'équations ou au contraire être spécifiques à chacun. Les méthodes mises au point pour la résolution des équations en domaine temporel utilisent le plus souvent une formulation des équations au premier ordre, alors que les méthodes conçues pour la résolution des équations de Maxwell en domaine fréquentiel s'appuient habituellement sur la formulation des équations du deuxième ordre dans laquelle on élimine le champ électrique ou le champ magnétique, afin de réduire la taille des systèmes algébriques résultants. Pour compléter et conclure ce paragraphe, nous discutons brièvement ci-dessous de deux familles de méthodes qui ont plus particulièrement retenu notre attention en regard des objectifs de notre étude.

La reformulation du système de Maxwell à l'aide d'équations intégrales peut permettre, en particulier pour les problèmes de diffraction d'ondes en domaine extérieur, de ramener un problème extérieur volumique à un problème surfacique. Cette approche présente un certain nombre d'avantages. D'une part, seule la surface des objets considérés est maillée et on peut alors résoudre en utilisant la méthode des éléments finis de frontière. On obtient ainsi une réduction significative du nombre d'inconnues par rapport à une discrétisation volumique. D'autre part, le traitement des conditions aux limites étant

intégré à la formulation, il n'est pas nécessaire de mailler le vide autour de l'objet, ni même une frontière fictive simulant l'infini. Par ailleurs, cette condition limite correspond à une condition de radiation exacte, ce qui rend l'approximation plus précise. En revanche, ces méthodes restent peu adaptées aux calculs en milieux très hétérogènes et la matrice résultant de la mise en place du système linéaire est pleine, ce qui alourdit considérablement les coûts de calcul et d'occupation mémoire. Ces questions de coûts sont en grande partie résolues en faisant appel aux méthodes rapides multipôles [Darrigrand, 2002]-[Sylvand, 2002]. Cependant, l'utilisation des équations intégrales pour la formulation temporelle à l'aide des potentiels retardés reste très coûteuse puisqu'elle nécessite le stockage de plusieurs matrices pour suivre l'évolution en temps des courants.

Les méthodes qui s'appuient sur une discrétisation du volume complet du domaine de calcul permettent à la fois de traiter le problème de la diffraction d'onde en domaine extérieur, mais aussi de tenir compte avec précision des fortes hétérogénéités du milieu de propagation. On peut envisager des méthodes de discrétisation de type éléments finis, différences finies ou volumes finis qui ont chacune des caractéristiques que nous détaillerons par la suite. Un autre avantage de cette classe de méthodes est que les formulations associées conduisent généralement à des systèmes linéaires dont la matrice est creuse. Cela dit, dans la grande majorité des cas, ces systèmes linéaires sont aussi difficiles à résoudre, et nécessitent par conséquent d'avoir recours à des méthodes de résolution robustes et performantes.

Résolution des systèmes linéaires

La discrétisation des équations de Maxwell en domaine fréquentiel conduit pour toutes les méthodes de discrétisation à la résolution d'un système linéaire. On distingue classiquement deux grandes classes de solveurs pour la résolution de ces systèmes linéaires : les solveurs directs et les solveurs itératifs.

Les solveurs directs utilisent des variantes de la méthode de Gauss, plus couramment la factorisation LU, et consistent en deux étapes principales : la construction des facteurs triangulaires L et U et la résolution des systèmes triangulaires associés. C'est un processus robuste, puisque les besoins en puissance de calcul et en mémoire sont connus à l'avance avec exactitude. De plus, pour une résolution multi-secondes membres comme on en rencontre dans les applications de calcul de diagramme de rayonnement d'antennes, le surcoût en temps de calcul des résolutions avec les différents seconds membres est faible devant le temps requis pour la factorisation. Malgré ces avantages, ce type de solveur est rapidement inutilisable dans un contexte industriel en raison de son coût mémoire et du nombre d'opérations à effectuer. Une autre difficulté avec ce type de méthode tient à l'adaptation aux architectures de calcul parallèle, en particulier pour les systèmes linéaires mettant en jeu des matrices creuses. Dans ce contexte, il y a deux stratégies principales pour obtenir une méthode de résolution directe extensible : l'approche supernodale et l'approche multifrontale. PaStiX [Hénon et al., 2002] développé au LABRI et à l'INRIA Bordeaux, et MUMPS [Amestoy et al., 2000] développé à l'ENSEEIH, au CERFACS à l'INRIA Rhône-Alpes et à l'ENS Lyon, sont deux exemples de solveurs s'appuyant sur ces approches.

Les méthodes itératives constituent l'autre famille de solveurs. Elles ont toutes en commun de construire une suite de vecteurs convergeant vers la solution du problème. La famille des méthodes par sous-espaces de Krylov [Saad, 1996] dont fait partie la méthode du gradient conjugué, la méthode de minimisation des résidus généralisés (GMRES) [Saad and Schultz, 1986] ou la méthode du bi-gradient conjugué stabilisé (Bi-CGStab) [van der Vorst, 1992], en est l'exemple le plus populaire. Le noyau algorithmique principal de toutes ces méthodes est le produit d'une matrice par un vecteur. Ces méthodes sont donc théoriquement facilement adaptables aux plates-formes de calcul parallèle. Elles sont en

outre moins robustes mais nettement plus économiques pour ce qui est de la capacité mémoire requise. Néanmoins, lorsque l'on cherche à traiter des cas avec de grands nombres d'inconnues ou pour des problèmes mettant en jeu des modélisations complexes, les méthodes par sous-espaces de Krylov convergent difficilement ou pas du tout sans la prise en compte d'une technique de préconditionnement. Toute la difficulté réside alors dans la recherche d'une technique de préconditionnement qui améliore sensiblement la convergence de la méthode sans pour autant augmenter dramatiquement le coût de réalisation d'une itération et aboutissant à un solveur parallélisable. Les techniques de préconditionnement de type factorisation approchée comme la méthode ILUT [Saad, 1994] sont parmi les plus utilisées. Elles présentent essentiellement les mêmes avantages et inconvénients que les méthodes de factorisation LU. D'autres techniques de préconditionnement sont présentées dans [Saad, 1996].

Les systèmes linéaires résultant de la discrétisation des équations de Maxwell en domaine fréquentiel par une méthode volumique (différences finies, éléments finis ou volumes finis) sont non seulement de très grande taille dans le cas d'une application industrielle, mais ils sont également difficiles à résoudre par une méthode itérative car ils impliquent des matrices à coefficients complexes et généralement, non-hermitiennes. Une manière de construire un préconditionneur efficace et parallèle pour un tel système linéaire, est d'adopter une approche par décomposition de domaine [Smith et al., 1996]. Le problème global est alors décomposé en sous-problèmes reliés entre eux par des conditions d'interface spécifiques. Cette approche qui se prête parfaitement au parallélisme, permet de résoudre chaque problème local par une méthode directe, et préconditionne le système global pour une résolution effectuée par une méthode itérative. On peut alors tirer profit à la fois du parallélisme, de la robustesse des méthodes directes, et de la flexibilité des méthodes itératives.

Objectif et plan de la thèse

L'objectif général de notre étude est le développement et l'évaluation de méthodes de type Galerkin discontinu en maillages tétraédriques non-structurés pour la résolution numérique des équations de Maxwell en formulation du premier ordre et en régime harmonique. Le point de départ de cette étude est l'application des méthodes de ce type à la résolution numérique des équations de Maxwell en domaine temporel, qui ont été mises au point ces dernières années dans l'équipe Caiman de l'INRIA Sophia Antipolis [Remaki, 2000]-[Piperno et al., 2002]-[Piperno and Fezoui, 2003]- [Fezoui et al., 2005]-[Bernacki et al., 2006]. Ces méthodes s'appuient sur des approximations centrées d'ordre 0 (méthode de volumes finis ou GD-P0) ou d'ordre 1 (méthode de type Galerkin discontinu linéaire ou GD-P1) combinées à un schéma d'intégration en temps explicite centré de type saute-mouton. Il s'agit donc de méthodes non-dissipatives dont on démontre qu'elles conservent une forme discrète de l'énergie électromagnétique, ce qui permet de garantir la stabilité sous des contraintes sur le pas de temps (conditions de CFL). La stabilité de ces méthodes est étudiée au moyen d'une approche énergétique valable pour tout type de maillage [Fezoui et al., 2005]. Dans cette thèse, nous nous intéressons à l'adaptation de ces méthodes temporelles au contexte des équations de Maxwell en régime harmonique, en préservant l'utilisation d'approximations centrées pour le calcul des flux numériques aux interfaces entre éléments voisins du maillage. Nous nous limitons à des approximations de bas ordre (méthodes GD-P0 et GD-P1) bien que des travaux récents ont clairement démontré l'intérêt d'adopter des méthodes d'interpolation d'ordre élevé, notamment dans le cadre des équations de Maxwell en domaine temporel [Hesthaven and Warburton, 2004a] et pour la résolution du problème aux valeurs propres associé aux équations de Maxwell en régime harmonique [Hesthaven and Warburton, 2004b]. De ce point de vue, cette étude constitue une première étape vers la mise au point de méthodes GD-P k ($k > 0$ quelconque)

en maillages tétraédriques pour la résolution numérique des équations de Maxwell en régime harmonique.

La réalisation de cette étude nous a conduit à viser trois contributions que l'on retrouve dans la structuration de ce manuscrit qui comporte :

- la formulation et l'analyse de méthodes de type Galerkin discontinu d'ordre 0 et 1 sont au cœur du chapitre 2. Nous avons choisi de donner un style didactique à ce chapitre en considérant successivement les situations 1D et 3D et en présentant séparément les formulations GD-P0 et GD-P1. En ce qui concerne l'analyse de ces méthodes, nous étudions d'une part, l'inversibilité des systèmes linéaires résultant des formulations GD-P0 et GD-P1 en 1D et 3D et, d'autre part, la dispersion numérique des ces formulations en 3D. On conclut ce chapitre par une évaluation numérique préliminaire des méthodes GD-P0 et GD-P1 en 1D.
- dans le chapitre 3, on étudie des méthodes de décomposition de domaine pour la résolution des équations de Maxwell en régime harmonique. On considère tout d'abord le système continu et on analyse la convergence d'algorithmes de Schwarz avec ou sans recouvrement basés sur des conditions d'interface naturelles. Ces conditions consistent à imposer aux interfaces les variables caractéristiques associées aux ondes entrantes dans un domaine. On s'intéresse ensuite au cas discret sur la base de la méthode d'approximation volume fini (méthode GD-P0) introduite au chapitre 2 mais formulée sur un maillage quadrangulaire en vue d'une mise en œuvre en Matlab pour une évaluation préliminaire en 2D. On étudie ensuite des conditions d'interface optimisées ayant pour but d'accélérer la convergence de l'algorithme de Schwarz sans recouvrement. Des tests préliminaires en 2D à l'aide de Matlab permettent de montrer clairement les gains résultant de l'utilisation de ces conditions.
- le chapitre 4 est consacré à l'évaluation numérique des méthodes d'approximation GD-P0 et GD-P1 en maillages tétraédriques. On considère pour cela une série de cas tests de complexité croissante portant sur des problèmes de propagation en milieux homogènes et hétérogènes. Les méthodes GD-P0 et GD-P1 sont tout d'abord évaluées sur le plan de la précision des calculs en lien avec la finesse de la discrétisation. On s'intéresse ensuite à la résolution des systèmes linéaires associés aux formulations GD-P0 et GD-P1 en mode d'exécution séquentiel (monoprocasseur). Enfin, on évalue en détail les performances parallèles d'un algorithme de Schwarz avec recouvrement basé sur des conditions d'interface naturelles pour la résolution des systèmes linéaires en question. On présente notamment les résultats de calculs portant sur plusieurs millions d'inconnues.

En amont de ces chapitres, le chapitre 1 précise les cadres mathématique et numérique de notre étude. Pour finir, le chapitre 4.3.2 fait un bilan de cette étude et identifie quelques pistes pour des travaux futurs.

Chapitre 1

Cadres mathématique et numérique

1.1 Modélisation de la propagation d'ondes électromagnétiques

La modélisation de phénomènes électromagnétiques est basée sur les équations de Maxwell. Elles regroupent les lois formulées par un ensemble de scientifiques et complétées par J.C. Maxwell qui en a donné un exposé complet dans son ouvrage [Maxwell, 1987].

1.1.1 Les équations de Maxwell

La définition des champs électromagnétiques résulte principalement des expériences sur les interactions mutuelles des charges électriques. L'idée de base est de supposer que chaque charge est plongée dans un champ électrique \mathbf{E} . Si une charge q subit une force \mathbf{F}_E , le champ \mathbf{E} à la position de la charge est défini par la relation $\mathbf{F}_E = q\mathbf{E}$. De plus, on peut observer que le déplacement d'une charge peut être affecté par une force proportionnelle à sa vitesse. Un autre champ doit donc être défini, l'induction magnétique \mathbf{B} , de telle sorte que cette force s'exprime par la relation $\mathbf{F}_B = q\mathbf{v} \times \mathbf{B}$. Ainsi, la présence d'un champ magnétique et d'un champ électrique appliquent à une particule de charge q une force globale dite force de Lorentz qui s'écrit $\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$. L'interdépendance des champs se révèle au travers des équations de Maxwell.

Loi d'induction de Faraday. Une importante contribution de Faraday en électromagnétisme a été l'établissement de cette loi : une variation en temps du flux magnétique passant à travers une boucle conductrice génère un courant dans cette boucle. Soit une surface ouverte A reposant sur une boucle conductrice C , cette loi s'écrit alors pour un chemin d'intégration fixe :

$$\oint_C \mathbf{E} \cdot d\mathbf{l} = - \iint_A \frac{\partial \mathbf{B}}{\partial t} \cdot d\mathbf{S}. \quad (1.1)$$

Cette loi indique qu'un champ électrique peut non seulement être généré par des charges électriques mais aussi par une variation du flux magnétique.

Loi de Gauss électrique. Cette loi relie le flux du champ électrique à travers une surface fermée, toujours notée A , aux charges se trouvant à l'intérieur de cette surface de la façon suivante :

$$\oiint_A \mathbf{E} \cdot d\mathbf{S} = \frac{1}{\varepsilon} \iiint_V \rho dV, \quad (1.2)$$

où V est le volume délimité par la surface A et ρ représente une densité volumique de charge. ε est la permittivité électrique du milieu. Cette loi indique la présence de sources électriques à l'intérieur de la surface A pour peu que l'on ait une différence entre les flux entrants et sortants à travers cette surface.

Loi de Gauss magnétique. Malgré toutes les interdépendances qui existent entre champs électrique et magnétique, aucune charge magnétique (un monopôle), analogue à la charge électrique, n'a pu être mise en évidence et ce en dépit de nombreuses investigations. Le champ d'induction magnétique peut toutefois être décrit en termes de courants et l'équivalent d'un dipôle magnétique serait une boucle infiniment petite. L'absence de charge magnétique entraîne la nullité du flux d'induction magnétique à travers toute surface fermée A :

$$\oiint_A \mathbf{B} \cdot d\mathbf{S} = 0. \quad (1.3)$$

Loi de Maxwell-Ampère. Cette loi relie la circulation de l'induction magnétique \mathbf{B} le long d'une courbe fermée C à l'intensité du courant i qui parcourt cette boucle, ce dernier étant défini par :

$$i = \iint_A \mathbf{j} \cdot d\mathbf{S},$$

où \mathbf{j} est la densité de courant par unité de surface (A étant une surface ouverte délimitée par C). Une contribution majeure de Maxwell fut de corriger cette loi puisque le déplacement de charges n'est pas la seule source de champ magnétique. Cette loi s'exprime par :

$$\oint_C \mathbf{B} \cdot d\mathbf{l} = \mu \iint_A \left(\frac{\varepsilon \partial \mathbf{E}}{\partial t} + \mathbf{j} \right) \cdot d\mathbf{S}, \quad (1.4)$$

où μ correspond à la perméabilité magnétique du milieu. La variation d'un champ électrique \mathbf{E} dépendant du temps génère un champ d'induction magnétique \mathbf{B} même en l'absence de courant source, c'est-à-dire lorsque $\mathbf{j} = 0$.

Dans le vide, les équations intégrales (1.1) à (1.4) deviennent :

$$\left\{ \begin{array}{l} \oint_C \mathbf{E} \cdot d\mathbf{l} = - \iint_A \frac{\partial \mathbf{B}}{\partial t} \cdot d\mathbf{S}, \\ \oint_C \mathbf{B} \cdot d\mathbf{l} = \varepsilon_0 \mu_0 \iint_A \frac{\partial \mathbf{E}}{\partial t} \cdot d\mathbf{S}, \\ \oiint_A \mathbf{E} \cdot d\mathbf{S} = 0, \\ \oiint_A \mathbf{B} \cdot d\mathbf{S} = 0, \end{array} \right. \quad (1.5)$$

puisque le vide est un milieu exempt de charges électriques et de courants sources. La permittivité électrique et la perméabilité magnétique du vide valent respectivement :

$$\left\{ \begin{array}{l} \varepsilon_0 = 8,85 \cdot 10^{-12} \text{ F.m}^{-1}, \\ \mu_0 = 1,26 \cdot 10^{-6} \text{ H.m}^{-1}. \end{array} \right.$$

Le théorème de la divergence de Gauss ainsi que le théorème de Stokes permettent alors d'écrire le système (1.5) en coordonnées locales sous la forme d'équations aux dérivées partielles. Pour un milieu quelconque :

$$\left\{ \begin{array}{l} \frac{\partial \mathbf{B}}{\partial t} + \text{rot } \mathbf{E} = 0, \\ \varepsilon \mu \frac{\partial \mathbf{E}}{\partial t} - \text{rot } \mathbf{B} = -\mu \mathbf{j}, \\ \text{div } \mathbf{E} = \frac{\rho}{\varepsilon}, \\ \text{div } \mathbf{B} = 0. \end{array} \right. \quad (1.6)$$

Les champs vectoriels \mathbf{E} , \mathbf{B} et \mathbf{j} sont des fonctions de $\mathbf{R} \times \mathbf{R}^3$ à valeurs dans \mathbf{R}^3 . \mathbf{E} est exprimé en V.m^{-1} , \mathbf{B} en T, et \mathbf{j} en A.m^{-2} . La densité de charge ρ est une fonction scalaire variable en espace et en temps et est exprimée en C.m^{-3} .

1.1.2 Relations phénoménologiques

Nous n'avons pas jusqu'à maintenant complètement considéré les caractéristiques du milieu de propagation. La permittivité électrique ε et la perméabilité magnétique μ du milieu ont été en effet considérées constantes.

La distribution interne de charge de certains matériaux se modifie cependant sous l'influence d'un champ électromagnétique ce qui entraîne notamment des effets de dispersion dont la modélisation conduit entre autres à considérer que la permittivité électrique dépend de la fréquence. La prise en compte de lois de comportement des matériaux ou lois constitutives est donc nécessaire. Des vecteurs de polarisation \mathbf{P} et \mathbf{M} respectivement électrique et magnétique sont introduits pour tenir compte des phénomènes de dispersion. La modification du champ électrique et de l'induction magnétique internes est alors modélisée à l'aide de la définition de deux nouveaux champs, le champ \mathbf{D} appelé déplacement électrique et le champ \mathbf{H} appelé champ magnétique. Ces deux nouveaux champs vérifient les relations :

$$\left\{ \begin{array}{l} \mathbf{D} = \varepsilon_0 \mathbf{E} + \mathbf{P}, \\ \mathbf{H} = \frac{1}{\mu_0} (\mathbf{B} - \mathbf{M}). \end{array} \right.$$

Pour des matériaux homogènes linéaires, isotropes, diélectriques, les champs \mathbf{E} et \mathbf{D} sont proportionnels. De même, les vecteurs \mathbf{B} et \mathbf{H} sont parallèles et proportionnels. On obtient donc les relations :

$$\left\{ \begin{array}{l} \mathbf{D} = \varepsilon \mathbf{E}, \\ \mathbf{B} = \mu \mathbf{H}. \end{array} \right.$$

Pour des matériaux hétérogènes isotropes, la permittivité électrique et la perméabilité magnétique dépendent de la variable d'espace et sont alors des fonctions scalaires $\varepsilon(\mathbf{x})$ et $\mu(\mathbf{x})$. Pour des matériaux anisotropes et hétérogènes, les fonctions scalaires $\varepsilon(\mathbf{x})$ et $\mu(\mathbf{x})$ sont remplacées par des tenseurs symétriques et définis positifs pour tenir compte du second principe de la thermodynamique [Dautray and Lions, 1985].

Outre ces lois de comportement, on ajoute d'autres lois constitutives pour des matériaux possédant d'autres caractéristiques physiques. Par exemple, pour les conducteurs à température constante, le flux de charge est conditionné par l'intensité du champ électrique suivant la loi d'Ohm :

$$\mathbf{j} = \sigma \mathbf{E},$$

où σ représente la conductivité électrique du milieu exprimée en S.m^{-1} . Dans le cas général, σ est un tenseur. A partir de cette loi et de la deuxième équation du système (1.6), on peut retrouver la loi de conservation de la charge électrique :

$$\frac{\partial \rho}{\partial t} + \text{div } \mathbf{j} = 0, \quad (1.7)$$

qui implique que les charges intérieures suivent la relation de relaxation suivante dans les conducteurs :

$$\rho(\mathbf{x}, t) = \rho(\mathbf{x}, 0) e^{-\frac{t}{\tilde{\tau}}},$$

avec $\tilde{\tau} = \varepsilon/\sigma$. Ce temps de relaxation est extrêmement faible dans les bons conducteurs et les charges se portent très rapidement à la surface.

Remarque 1 Une simplification importante mais généralement suffisante dans les applications relatives aux conducteurs métalliques est le modèle du conducteur parfait. La conductivité est supposée infinie et les charges se portent instantanément à la surface du conducteur.

1.1.3 L'équation des ondes

Les équations de Maxwell (1.6) considérées dans le vide impliquent deux équations des ondes vectorielles :

$$\begin{cases} \varepsilon_0 \mu_0 \frac{\partial^2 \mathbf{E}}{\partial t^2} = \Delta \mathbf{E}, \\ \varepsilon_0 \mu_0 \frac{\partial^2 \mathbf{B}}{\partial t^2} = \Delta \mathbf{B}. \end{cases}$$

Dans certains milieux et pour certains dispositifs, comme ici dans le vide, les champs électrique et magnétique peuvent ainsi être découplés et chacune des composantes vérifie une équation des ondes scalaire. Ces champs se propagent donc par ondes. L'équation des ondes a été étudiée bien avant l'établissement des équations de Maxwell, et J.C. Maxwell en a déduit la valeur théorique de la vitesse de la lumière dans le vide :

$$c_0 = \frac{1}{\sqrt{\varepsilon_0 \mu_0}} \approx 3.10^8 \text{ m.s}^{-1}.$$

1.1.4 Symétrisation du système de Maxwell

L'introduction du champ de déplacement électrique permet d'écrire les équations de Maxwell sous une forme conservative (y compris pour des matériaux hétérogènes) qui possèdent des propriétés intéressantes d'un point de vue mathématique. La formulation conservative la plus souvent rencontrée dans la littérature fait intervenir les quatre différents champs de vecteurs de la théorie de l'électromagnétisme,

soient \mathbf{E} , \mathbf{H} , \mathbf{D} et \mathbf{B} . Les équations de Maxwell révèlent alors un comportement symétrique de ces quatre champs. Seule l'absence de sources magnétiques rompt cette symétrie en ce qui concerne les sources. L'introduction d'une densité volumique de courant magnétique \mathbf{m} et d'une densité volumique de charge magnétique ϱ permet de symétriser totalement le système de Maxwell qui s'écrit dans ces conditions :

$$\left\{ \begin{array}{l} \frac{\partial \mathbf{D}}{\partial t} - \operatorname{rot} \mathbf{H} = -\mathbf{j}, \\ \frac{\partial \mathbf{B}}{\partial t} + \operatorname{rot} \mathbf{E} = -\mathbf{m}, \\ \operatorname{div} \mathbf{D} = \rho, \\ \operatorname{div} \mathbf{B} = \varrho. \end{array} \right. \quad (1.8)$$

On ajoute également à ce système symétrisé une loi de conservation de la charge magnétique analogue à (1.7) :

$$\frac{\partial \varrho}{\partial t} + \operatorname{div} \mathbf{m} = 0.$$

Cette symétrisation n'a de sens que d'un point de vue mathématique puisqu'il n'existe pas réellement de courants et de sources magnétiques.

Remarque 2 On peut également introduire une loi d'Ohm magnétique $\mathbf{m} = \zeta \mathbf{H}$ afin de prendre en compte les pertes dans des milieux magnétiques.

1.1.5 Conditions aux limites et conditions d'interface

L'écriture classique des équations de Maxwell suppose implicitement que sources et champs ont une régularité qui assure un sens aux diverses dérivations. Or, des répartitions de charges (éventuellement ponctuelles) ou de courants localisés sur des courbes ou des surfaces se rencontrent fréquemment dans les applications. L'écriture des équations de Maxwell au sens des distributions permet alors de traiter des singularités et d'expliciter de manière générale les conditions aux limites ou d'interface associées à ce type de problème.

1.1.5.1 Equations au sens des distributions

Soit une distribution \mathbf{T}_A associée au champ de vecteur \mathbf{A} dérivable au sens des fonctions dans le complémentaire d'une surface de discontinuité S orientée et de normale \mathbf{n} . La divergence et le rotationnel de \mathbf{T}_A vérifient tous deux :

$$\left\{ \begin{array}{l} \operatorname{div}(\mathbf{T}_A) = \operatorname{div} \mathbf{A} + \mathbf{n} \cdot [\mathbf{A}]_S \delta_S, \\ \operatorname{rot}(\mathbf{T}_A) = \operatorname{rot} \mathbf{A} + \mathbf{n} \times [\mathbf{A}]_S \delta_S, \end{array} \right.$$

où δ_S est la mesure de Dirac portée par la surface S et $[\mathbf{A}]_S$ est le saut de \mathbf{A} à travers la surface S . Si on se sert de ces relations ainsi que des distributions associées aux grandeurs intervenant dans les

équations de Maxwell, on aboutit en prenant le système (1.8) au sens des distributions à :

$$\left\{ \begin{array}{l} \frac{\partial \mathbf{D}}{\partial t} - \text{rot } \mathbf{H} - \mathbf{n} \times [\mathbf{H}]_S \delta_S = -\mathbf{j} - \mathbf{j}_S \delta_S, \\ \frac{\partial \mathbf{B}}{\partial t} + \text{rot } \mathbf{E} + \mathbf{n} \times [\mathbf{E}]_S \delta_S = -\mathbf{m} - \mathbf{m}_S \delta_S, \\ \text{div } \mathbf{D} + \mathbf{n} \cdot [\mathbf{D}]_S \delta_S = \rho + \rho_S \delta_S, \\ \text{div } \mathbf{B} + \mathbf{n} \cdot [\mathbf{B}]_S \delta_S = \varrho + \varrho_S \delta_S, \end{array} \right. \quad (1.9)$$

où les densités de charges ρ et ϱ prises elles aussi au sens des distributions peuvent se décomposer comme la somme d'une densité volumique et d'une densité superficielle répartie sur la surface S : $T_\rho = \rho + \rho_S \delta_S$ et $T_\varrho = \varrho + \varrho_S \delta_S$.

On peut alors déduire de (1.9) par identification les relations de sauts normaux et tangentiels en se servant des équations de Maxwell prises au sens des fonctions :

$$\left\{ \begin{array}{l} \mathbf{n} \times [\mathbf{H}]_S = \mathbf{j}_S, \\ \mathbf{n} \times [\mathbf{E}]_S = -\mathbf{m}_S, \\ \mathbf{n} \cdot [\mathbf{D}]_S = \rho_S, \\ \mathbf{n} \cdot [\mathbf{B}]_S = \varrho_S. \end{array} \right. \quad (1.10)$$

1.1.5.2 Conditions d'interface

Des relations (1.10) nous pouvons déduire les conditions d'interface à travers une surface de discontinuité qui peut comporter des sources de courants. En l'absence de courants, les composantes tangentielles des champs électrique et magnétique vérifient respectivement :

$$\left\{ \begin{array}{l} \mathbf{n} \times [\mathbf{E}]_S = 0, \\ \mathbf{n} \times [\mathbf{H}]_S = 0. \end{array} \right.$$

On remarque que les composantes tangentielles sont continues quelles que soient la permittivité et la perméabilité du milieu. Quant aux composantes normales, elles ne sont en général pas continues même en l'absence de charges, puisqu'elles vérifient :

$$\left\{ \begin{array}{l} \mathbf{n} \cdot [\varepsilon \mathbf{E}]_S = 0, \\ \mathbf{n} \cdot [\mu \mathbf{H}]_S = 0. \end{array} \right.$$

Nous pouvons utiliser également les relations (1.10) pour exprimer les conditions limites.

1.1.5.3 Conditions aux limites pour une frontière métallique

Nous assimilerons dans cette étude les parois métalliques au modèle idéal et fictif du conducteur parfait. Nous considérons alors que les charges magnétiques ϱ_S et les courants \mathbf{m}_S sont nuls, et que le champ électromagnétique est nul à l'intérieur du métal. Soient \mathbf{E}_S et \mathbf{H}_S les champs électrique et magnétique à la surface extérieure du métal. Les relations de saut (1.10) deviennent :

$$\begin{cases} \mathbf{n} \times \mathbf{H}_S = \mathbf{j}_S, \\ \mathbf{n} \times \mathbf{E}_S = 0, \\ \mathbf{n} \cdot \mathbf{E}_S = \frac{\rho_S}{\varepsilon}, \\ \mathbf{n} \cdot \mathbf{H}_S = 0. \end{cases}$$

et on en déduit qu'à la surface d'un conducteur parfait, le champ électrique est normal et le champ magnétique est tangent.

1.1.5.4 Conditions aux limites pour une frontière artificielle

Pour être en mesure de procéder à des simulations numériques, nous devons restreindre le domaine de calcul à un domaine borné. On se trouve obligé de poser une condition aux limites totalement absorbante ou transparente qui ne génère pas d'ondes parasites se réfléchissant vers l'intérieur du domaine. Néanmoins, ces conditions non-locales en temps et en espace restent difficiles à implémenter. On a donc généralement recours à des conditions locales approximativement transparentes moins précises mais plus facilement implémentables. Un exemple de ces conditions est la condition absorbante d'ordre un de Silver-Müller :

$$\mathbf{n} \times \mathbf{E}_S = -\sqrt{\frac{\mu_0}{\varepsilon_0}} \mathbf{n} \times (\mathbf{n} \times \mathbf{H}_S),$$

pour une frontière dans le vide placée assez loin de l'objet diffractant afin d'éviter les réflexions parasites.

1.2 Cadre mathématique

1.2.1 Problème de la diffraction d'ondes

Le problème de la diffraction d'une onde par un obstacle est un problème dont la résolution sert par exemple au calcul de la surface équivalente radar (SER). On considère pour ce faire un obstacle de \mathbf{R}^3 possédant une frontière Γ^m métallique, considérée comme parfaitement conductrice. On envoie une onde électromagnétique incidente appelée *champ incident*, se propageant dans le domaine extérieur Ω en direction de l'obstacle. Cette onde incidente étant connue, l'obstacle induit une perturbation de cette dernière en créant une onde diffractée que l'on nommera *champ diffracté*. Il s'agit alors de déterminer le *champ diffracté* $\mathbf{W}^{\text{dif}} = (\mathbf{E}^{\text{dif}}, \mathbf{H}^{\text{dif}})^t$ considéré comme étant la différence entre le *champ électromagnétique total* $\mathbf{W}^{\text{tot}} = (\mathbf{E}^{\text{tot}}, \mathbf{H}^{\text{tot}})^t$ et le *champ incident* $\mathbf{W}^{\text{inc}} = (\mathbf{E}^{\text{inc}}, \mathbf{H}^{\text{inc}})^t$: $\mathbf{W}^{\text{dif}} = \mathbf{W}^{\text{tot}} - \mathbf{W}^{\text{inc}}$.

1.2.2 Hyperbolicité du système de Maxwell

Le caractère hyperbolique est intrinsèque au système de Maxwell et a une interprétation physique : les ondes et l'énergie associée se propagent en temps fini suivant des directions particulières. Cette propriété n'a été jusqu'à maintenant que très peu exploitée pour la résolution du système de Maxwell alors qu'elle l'a été largement pour le système d'Euler par exemple. La principale application de cette propriété pour les numériciens est la construction de schémas décentrés qui tiennent naturellement compte de la direction de propagation des ondes.

Soit Ω un ouvert régulier borné de \mathbf{R}^3 de frontière $\Gamma = \partial\Omega = \Gamma^m \cup \Gamma^a$ (avec $\Gamma^m \cap \Gamma^a = \emptyset$) comme représentée sur la figure 2.1.

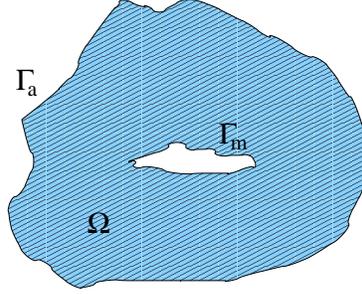


FIG. 2.1 – Domaine de calcul

Dans ce qui suit, on considère le cas où $\varepsilon(\mathbf{x})$, $\mu(\mathbf{x})$, $\sigma(\mathbf{x})$ et $\varsigma(\mathbf{x})$ sont des variables scalaires (matériaux isotropes) et indépendantes du temps. Les vecteurs "déplacement électrique", "induction magnétique" et "densité de courant" peuvent donc se mettre sous la forme :

$$\mathbf{D} = \varepsilon\mathbf{E}, \quad \mathbf{B} = \mu\mathbf{H} \quad \text{et} \quad \mathbf{J} = (\sigma\mathbf{E}, \varsigma\mathbf{H})^t \quad (2.11)$$

Les quantités $\varepsilon(\mathbf{x})$ et $\mu(\mathbf{x})$ s'expriment en fonction des indices du vide ε_0 et μ_0 , et des valeurs adimensionnées relatives au milieu considéré ε_r et μ_r . On introduit enfin une constante appelée l'impédance du vide notée z_0 . En résumé :

$$\left\{ \begin{array}{l} \varepsilon(\mathbf{x}) = \varepsilon_0\varepsilon_r(\mathbf{x}), \\ \mu(\mathbf{x}) = \mu_0\mu_r(\mathbf{x}), \\ c_0 = \frac{1}{\sqrt{\varepsilon_0\mu_0}}, \quad c_r(\mathbf{x}) = \frac{1}{\sqrt{\varepsilon_r(\mathbf{x})\mu_r(\mathbf{x})}} \quad \text{et} \quad c(\mathbf{x}) = c_0c_r(\mathbf{x}), \\ z_0 = \sqrt{\frac{\mu_0}{\varepsilon_0}}, \quad z_r(\mathbf{x}) = \sqrt{\frac{\mu_r(\mathbf{x})}{\varepsilon_r(\mathbf{x})}} \quad \text{et} \quad z(\mathbf{x}) = z_0z_r(\mathbf{x}). \end{array} \right. \quad (2.12)$$

Si on note :

$$N_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \quad N_y = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad N_z = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

et :

$$G_l = \begin{pmatrix} \mathbf{0}_3 & N_l \\ -N_l & \mathbf{0}_3 \end{pmatrix} \quad \text{avec} \quad l \in \{x, y, z\},$$

alors les lois d'Ampère et de Faraday peuvent se réécrire sous la forme matricielle suivante :

$$\partial_t(G_0\mathbf{W}) + G_x\partial_x\mathbf{W} + G_y\partial_y\mathbf{W} + G_z\partial_z\mathbf{W} + K\mathbf{W} = 0, \quad (2.13)$$

où on a adopté la notation classique $\partial_{\{t,x,y,z\}} = \frac{\partial}{\partial\{t,x,y,z\}}$ et où :

$$\mathbf{W} = \begin{pmatrix} \mathbf{E} \\ \mathbf{H} \end{pmatrix}, \quad G_0 = \begin{pmatrix} \varepsilon(\mathbf{x}) & 0 \\ 0 & \mu(\mathbf{x}) \end{pmatrix} \quad \text{et} \quad K = \begin{pmatrix} \sigma(\mathbf{x}) & 0 \\ 0 & \varsigma(\mathbf{x}) \end{pmatrix}.$$

La solution de ces équations est le champ électromagnétique total. L'intérêt de l'écriture des équations de Maxwell comme (2.13) est que cette forme est celle des systèmes hyperboliques symétriques du premier ordre ou systèmes de Friedrichs.

Remarque 3 *Les lois de Gauss électrique et de Gauss magnétique sont redondantes pour une donnée initiale vérifiant les contraintes (2.11) [Dautray and Lions, 1985]. C'est pour cette raison qu'il est inutile de les prendre en compte.*

On suppose que le champ incident vérifie l'équation (2.13) dans le vide si bien que le champ diffracté (qu'on notera abusivement \mathbf{W}) vérifie l'équation (voir par exemple Helluy [Helluy, 1994] pour plus de détails) :

$$\begin{aligned} \partial_t(G_0\mathbf{W}) + G_x\partial_x\mathbf{W} + G_y\partial_y\mathbf{W} + G_z\partial_z\mathbf{W} + K\mathbf{W} &= -K\mathbf{W}^{\text{inc}} + (\mathbf{I}_6 - G_0)\partial_t\mathbf{W}^{\text{inc}} \\ &\equiv S(\mathbf{W}^{\text{inc}}). \end{aligned} \quad (2.14)$$

où \mathbf{I}_6 désigne la matrice identité de l'espace $\mathbf{M}_6(\mathbf{R}^6)$. Si on note par $G_{\mathbf{n}} = n^x G_x + n^y G_y + n^z G_z$ où \mathbf{n} désigne un vecteur non-nul de \mathbf{R}^3 , le fait que le système des équations de Maxwell (2.14) est hyperbolique signifie que la matrice de flux $\bar{G}_{\mathbf{n}} = G_{\mathbf{n}} G_0^{-1}$ est diagonalisable dans \mathbf{R}^6 et on a :

$$\begin{cases} \bar{G}_{\mathbf{n}} &= T_{\mathbf{n}} \Lambda_{\mathbf{n}} T_{\mathbf{n}}^{-1} \\ \bar{G}_{\mathbf{n}}^{\pm} &= T_{\mathbf{n}} \Lambda_{\mathbf{n}}^{\pm} T_{\mathbf{n}}^{-1} \\ \Lambda_{\mathbf{n}}^+ &= \text{diag}(\max(\lambda_{\mathbf{n},k}, 0)) \\ \Lambda_{\mathbf{n}}^- &= \text{diag}(\min(\lambda_{\mathbf{n},k}, 0)) \end{cases}$$

où $\lambda_{\mathbf{n},k}$, $k = 1, \dots, 6$ désignent les valeurs propres de la matrice de flux $\bar{G}_{\mathbf{n}}$.

On discute maintenant des conditions limites imposées sur la frontière de Ω .

Condition réfléchissante. On suppose que la partie $\Gamma^m = \partial\Omega_m$ de la frontière du domaine de calcul est la surface d'un obstacle métallique. On y impose que la trace tangentielle du champ électrique est nulle c'est-à-dire $\mathbf{E} \times \mathbf{n} = 0$ où \mathbf{n} désigne la normale à Γ^m .

Condition absorbante. Sur la partie $\Gamma^a = \partial\Omega_a$, qui est en fait la frontière artificielle du domaine, on va imposer une condition absorbante qui vise à minimiser la réflexion dans le domaine d'ondes passant par cette frontière. Puisque c'est la matrice $\bar{G}_{\mathbf{n}} = G_{\mathbf{n}} G_0^{-1}$ qui est diagonalisable, cette condition s'écrit $\bar{G}_{\mathbf{n}}^- G_0 \mathbf{W} = 0$ ¹ On vérifie aisément que $\bar{G}_{\mathbf{n}}^- = G_{\mathbf{n}}^- G_0^{-1}$ si bien que la condition à imposer s'écrit aussi $G_{\mathbf{n}}^- \mathbf{W} = 0$ où \mathbf{n} désigne la normale sortante à Γ^a . On constate que celle-ci est équivalente à la condition de Silver-Müller d'ordre 1 :

¹Le système des équations de Maxwell (2.14) est hyperbolique ce qui signifie que la diagonalisation s'opère sur la matrice de flux obtenue lorsque le vecteur d'état est $\mathbf{Q} = G_0 \mathbf{W}$. En d'autres termes, le système de référence pour la diagonalisation est le système (2.14) sous forme conservative : $\partial_t(\mathbf{Q}) + \partial_x(\bar{G}_x \mathbf{Q}) + \partial_y(\bar{G}_y \mathbf{Q}) + \partial_z(\bar{G}_z \mathbf{Q}) + K G_0^{-1} \mathbf{Q} = \bar{S}(\mathbf{Q}^{\text{inc}})$.

$$\mathbf{n} \times \mathbf{E} + \mathbf{n} \times \mathbf{n} \times \mathbf{H} = 0.$$

Nous allons maintenant voir le lien entre les équations de Maxwell en domaine temporel définies jusqu'alors et les équations de Maxwell en domaine fréquentiel qui donnent lieu au *problème de Maxwell harmonique*.

1.2.3 Le régime harmonique

De nombreuses applications font intervenir des grandeurs et des champs qui varient sinusoidalement en fonction du temps avec la même pulsation $\omega = 2\pi f$ s'exprimant en rad.s^{-1} où la fréquence f s'exprime en Hz. En l'occurrence, si le champ incident suit cette dépendance harmonique en temps, c'est-à-dire s'il s'exprime sous la forme $\mathbf{W}^{\text{inc}}(\mathbf{x}, t) = \mathbf{W}^{\text{inc}}(\mathbf{x})e^{i\omega t}$, le comportement asymptotique en temps du champ diffracté sera déterminé en supposant le *principe d'amplitude limite* vrai (voir notamment Rauch [Rauch, 1978] pour les limites d'applicabilité de ce principe). Ce qui veut dire que le champ diffracté est la superposition d'une onde harmonique d'amplitude $\mathbf{W}(\mathbf{x})$ avec la même pulsation que le champ incident, d'ondes harmoniques correspondant aux fréquences propres ω_k du problème (d'amplitudes \mathbf{W}_k) et d'une onde évanescence. Si le principe d'amplitude limite est vrai, les \mathbf{W}_k sont à support compact et l'unique inconnue du problème asymptotique est \mathbf{W} . On rappelle dans ce qui suit des résultats énoncés dans Helluy [Helluy, 1994], qui montrent le lien entre les équations en domaine temporel et les équations harmoniques :

Théorème 1 *Principe d'amplitude limite.*

(i) Si on note $p = i\omega + \nu$, $\nu > 0$ et $M_\nu(u) = \nu L_p(u)$ où $L_p(u)$ est la transformée de Laplace partielle par rapport au temps au point p , et si :

$$\mathbf{W}(\mathbf{x}, t) = \mathbf{W}(\mathbf{x})e^{i\omega t} + \sum_{k=1}^m \mathbf{W}_k e^{i\omega_k t} + \phi(\mathbf{x}, t) \quad \text{tel que} \quad \lim_{t \rightarrow \infty} \phi(\mathbf{x}, t) = 0, \quad (2.15)$$

alors $\lim_{\nu \rightarrow 0} M_\nu(\mathbf{W}(\mathbf{x}, t)) = \mathbf{W}(\mathbf{x})$ dans $L^2(\Omega)$.

(ii) Si $\mathbf{W}(\mathbf{x}, t)$ est solution des équations (2.14) alors la quantité $\mathbf{W}_\nu(\mathbf{x}) = M_\nu(\mathbf{W}(\mathbf{x}, t))$ vérifie les équations harmoniques perturbées :

$$pG_0 \mathbf{W}_\nu + G_x \partial_x \mathbf{W}_\nu + G_y \partial_y \mathbf{W}_\nu + G_z \partial_z \mathbf{W}_\nu + K \mathbf{W}_\nu = S_\nu(\mathbf{W}^{\text{inc}}), \quad (2.16)$$

avec $S_\nu(\mathbf{W}^{\text{inc}}) = -K \mathbf{W}^{\text{inc}} + (I_6 - G_0)p \mathbf{W}^{\text{inc}}$.

Remarque 4 *Dans une cavité métallique, la décomposition (2.15) est fautive si ω est une fréquence propre du domaine.*

La limite au sens des distributions (quand ν tend vers zero) des équations (2.16) donne les équations harmoniques de solution $\mathbf{W}(\mathbf{x})$ qui est l'amplitude de l'onde harmonique de pulsation ω (principe de l'amplitude limite). Par la suite, si nécessaire (par exemple pour démontrer que le problème discret est bien posé) on étudiera plutôt le problème (2.16) en admettant les principes d'absorption limite et d'amplitude limite.

1.3 La discrétisation des équations de Maxwell

Indépendamment des formulations temporelles et fréquentielles des équations de Maxwell, on peut distinguer trois grandes classes de méthodes numériques lorsqu'il s'agit de traiter de problèmes de propagation d'ondes dans des milieux fortement hétérogènes : les méthodes de différences finies (DF), les méthodes d'éléments finis (EF), et enfin les méthodes de volumes finis (VF). La première méthode numérique efficace développée pour les équations de Maxwell a été une méthode DF due à Yee [Yee, 1966]. Cette méthode DFDT (différences finies dans le domaine temporel) conserve l'énergie, propriété que l'on cherche pour les méthodes numériques appliquées aux problèmes de propagation d'ondes en domaine temporel. De plus, sa grande simplicité d'implémentation la rend peu coûteuse en taille mémoire et en temps de calcul et on peut, par exemple pour le problème de la diffraction, effectuer un calcul pour un signal large bande, et étendre par transformation de Fourier le résultat à toutes les fréquences du spectre. Malheureusement, l'utilisation d'un maillage cartésien contraint le pas d'espace à être petit pour tenir compte de la géométrie des objets représentés dès lors que ceux-ci ont des formes complexes, ce qui réduit par la même occasion le pas de temps pour les schémas explicites comme le schéma de Yee (condition de stabilité de type CFL). D'autre part, la représentation en maillages cartésiens non-uniformes engendre des ondes parasites qui perturbent les solutions numériques obtenues.

L'autre méthode importante développée une quinzaine d'années plus tard fut la méthode des EF d'arêtes [Nedelec, 1980] qui fut déclinée en plusieurs versions [Monk, 2003]-[Nedelec, 1986]. Cette méthode présente beaucoup d'avantages, avec entre autres la faculté de traiter des maillages non-structurés et donc des géométries complexes, et de conserver l'énergie. Elle peut même être utilisée en tant que méthode d'ordre élevé [Demkowicz and Xue, 2005]-[Hiptmair, 2001]-[Monk, 2003]. Un point faible majeur réside dans l'inversion de la matrice de masse obtenue à chaque pas de temps en domaine temporel, ce qui amène à envisager des techniques de condensation de masse [Cohen and Monk, 1998]. Cet inconvénient a également conduit à l'étude des schémas de type VF pour les problèmes de propagation d'ondes en domaine temporel en raison de leur caractère hyperbolique, en s'inspirant des méthodes développées pour les problèmes issus de la mécanique des fluides. Tout comme les EF, les méthodes VF ont l'avantage de pouvoir s'adapter facilement à des maillages non-structurés. Leur particularité réside dans la formulation du flux numérique, qui peut être décentré, ce qui pour les équations de Maxwell rend le schéma diffusif [Piperno, 2000] et peu dispersif, ou centré mais dans ce cas le schéma est non-diffusif mais davantage dispersif [Remaki, 2000].

Une classe de méthodes numériques relativement récente, appelée méthodes de type Galerkin discontinu (GD), combine les avantages des méthodes d'éléments finis (EF) et de volumes finis (VF) puisqu'elle traite la discontinuité entre cellules voisines par une approximation VF sur les flux et qu'elle approche le champ dans chaque cellule par une base locale de fonctions. Cela permet une représentation plus riche des solutions recherchées, mais les méthodes GD sont à ce titre plus gourmandes en espace mémoire que les méthodes EF standards du fait de l'augmentation du nombre de degrés de liberté. Dans ce formalisme, la méthode GD d'ordre 0 est considérée comme une méthode de type VF étant donné que les fonctions de base utilisées sont constantes. Ces méthodes ont été proposées initialement pour résoudre l'équation scalaire du transport des neutrons [Hill and Reed, 1973]. Dans le domaine temporel, ces méthodes permettent d'approcher efficacement plusieurs classes de problèmes hyperboliques lorsqu'on les combine à une intégration temporelle de type Runge-Kutta [Cockburn and Shu, 1989].

Dans le domaine de la propagation d'ondes, et plus particulièrement pour l'équation de Helmholtz, les méthodes EF ne peuvent éviter un effet de pollution numérique lorsque le nombre d'onde augmente

[Babuska and Ihlenburg, 1995]-[Gerdes and Ihlenburg, 1999] et ce quel que soit le type des fonctions de base utilisées. Pour ce qui concerne les équations de Maxwell en domaine temporel, de nombreux schémas reposent soit sur une formulation à flux centrés [Piperno and Fezoui, 2003]-[Fezoui et al., 2005], soit sur une formulation à flux décentrés [Hesthaven and Warburton, 2004a]. Elles ont même été utilisées avec une interpolation d'ordre très élevé (polynômes du dixième degré) pour approcher les équations de Maxwell en domaine temporel [Hesthaven and Warburton, 1999]. Le principal problème récurrent à ces travaux est que l'intégration en temps explicite contraint à prendre une condition CFL restrictive sur des maillages fins. Ce constat a conduit certains auteurs à développer des méthodes spatio-temporelles qui reposent sur l'utilisation d'un schéma semi-explicite pour les systèmes hyperboliques linéaires en milieu inhomogène [Monk and Richter, 2005]. La condition CFL devient alors une CFL locale qui ne dépend plus que d'un petit nombre d'éléments, mais le décentrage du flux introduit malheureusement de la diffusion numérique. La solution est donc altérée après un nombre de pas de temps significatif.

Toujours en électromagnétisme mais cette fois-ci dans l'étude du problème aux valeurs propres discret, les méthodes GD ont aussi montré leur efficacité [Hesthaven and Warburton, 2004b]. L'approximation discontinue et la nature non-conforme de l'approximation permettent en l'occurrence de résoudre un problème aux valeurs propres standard puisque la matrice de masse est alors diagonale par blocs, plutôt que de résoudre un problème aux valeurs propres généralisé. Cependant, bien que des exemples 2D montrent que l'on n'obtient pas de valeurs propres parasites grâce à l'équivalence en 2D entre les solutions du problème aux valeurs propres discret pour les équations de Maxwell du premier ordre et les solutions de l'équation de Helmholtz, les exemples 3D ne peuvent éviter cet écueil à cause de la mauvaise représentation du noyau de l'opérateur rot-rot [Hesthaven and Warburton, 2006].

Les méthodes GD ont également connu un développement très important pour des équations de nature elliptique et parabolique [Arnold, 1982]-[J. Douglas Jr. and Dupont, 1976]-[Wheeler, 1978] et cela indépendamment de leur développement pour les équations hyperboliques. Elles ont toutes la particularité de reposer sur l'inclusion d'un terme de pénalisation au sein du flux numérique sans lequel la convergence ne serait pas assurée. Toutes ces méthodes répertoriées et étudiées dans [Arnold et al., 2002] ont des avantages et des inconvénients, mais restent supérieures en terme de souplesse aux méthodes EF classiques sur le plan de l'adaptativité *hp* et de leur utilisation sur des maillages non-conformes. Un autre avantage est de pouvoir utiliser une base de fonctions nodales qui peuvent approcher des solutions dans des zones du domaine fortement singulières, contrairement aux éléments finis classiques [Assous et al., 1999]-[Monk, 2003].

Bien que les méthodes EF pour traiter les équations de Maxwell en domaine fréquentiel considèrent pour la majorité la formulation du second ordre [Perugia et al., 2002]-[Perugia and Schötzau, 2003]-[Houston et al., 2004], certaines études se basent sur la formulation du premier ordre comme dans [Helluy, 1994] et [Bonnet, 1998]. Dans [Helluy, 1994], le schéma utilisé est un schéma GD d'ordre 0, basé sur une fonction de flux numérique décentré. C'est grâce à cette formulation du flux numérique, rendant coercive la forme bilinéaire associée, que des résultats d'existence et d'unicité de la solution discrète sur des domaines bornés et non-bornés peuvent être établis [Helluy and Dayma, 1994]-[Helluy et al., 1994]. Des preuves de convergence ont même été établies lorsque la perturbation est nulle, mais cela nécessite de repasser par la formulation continue de manière à exploiter des résultats provenant de la théorie du contrôle.

Chapitre 2

Méthodes de type Galerkin discontinu en régime harmonique

Ce chapitre est organisé en trois sections. Dans la section 2.1 on définit le problème aux limites associé aux équations de Maxwell en domaine fréquentiel et on réalise une analyse préliminaire des problèmes discrets qui caractérisent des méthodes GD d'ordre 0 (GD-P0 ou méthode de volumes finis) et d'ordre 1 (GD-P1) centrées. La section 2.2 est consacrée à l'étude de ces formulations GD-P0 et GD-P1 appliquées aux équations de Maxwell 1D. On s'intéresse notamment à l'inversibilité des systèmes matriciels résultant des formulations GD-P0 et GD-P1 en 1D. Enfin, la section 2.3 traite des formulations GD-P0 et GD-P1 pour les équations de Maxwell 3D.

Notations

\mathbb{K} représente le corps des réels \mathbb{R} ou celui des complexes \mathbb{C} .

$|\cdot|$ désigne simultanément la norme euclidienne d'un élément de $\mathbb{K}^{m \times n}$ et le cardinal d'un ensemble fini.

$\mathbf{M}_n(\mathbb{K})$ est l'espace vectoriel des matrices carrées sur \mathbb{K} de taille $n \times n$.

0_n est la matrice nulle de l'espace $\mathbf{M}_n(\mathbb{K})$.

I_n est la matrice identité de l'espace $\mathbf{M}_n(\mathbb{K})$.

$\mathbf{M}_{n,m}(\mathbb{K})$ est l'espace vectoriel des matrices rectangulaires sur \mathbb{K} de taille $n \times m$.

$\mathbf{M}_{m \times n}(\mathbb{K})$ est l'espace vectoriel des matrices carrées sur \mathbb{K} de $n \times n$ blocs de taille m .

V_j est le volume occupé par la cellule C_j .

dx et $d\sigma$ sont respectivement les mesures de volume et de surface induite dans \mathbb{R}^3 .

\mathcal{I} est l'ensemble des indices des cellules purement internes à Ω_h .

\mathcal{I}_b est l'ensemble des indices des cellules dont au moins un des bords s'appuie sur Γ_h .

$\mathcal{V}_j = \{k \in \mathcal{I} \cup \mathcal{I}_b / \partial C_j \cap \partial C_k \neq \emptyset\}$ est l'ensemble des indices k pour lesquels C_k est une cellule voisine de C_j .

$\Sigma_{jk} = \partial C_j \cap \partial C_k$, $k \in \mathcal{V}_j$.

S_{jk} est l'aire de la surface Σ_{jk} .

$\mathcal{V}_j^d = \{k \in \mathcal{I} \cap \mathcal{V}_j / \Sigma_{jk} \subset \Omega_h\}$ l'ensemble des indices des voisins internes à C_j .

$\mathcal{V}_j^a = \{k \in \mathcal{I}_b \cap \mathcal{V}_j / \Sigma_{jk} \subset \Gamma_h^a\}$ l'ensemble des indices des voisins absorbants à C_j .

$\mathcal{V}_j^m = \{k \in \mathcal{I}_b \cap \mathcal{V}_j / \Sigma_{jk} \subset \Gamma_h^m\}$ l'ensemble des indices des voisins métalliques à C_j .

$\mathcal{I}_m = \{j \in \mathcal{I}_b / \mathcal{V}_j^m \neq \emptyset\}$ est l'ensemble des indices des cellules qui possèdent au moins un côté appartenant à Γ_h^m .

$\mathcal{I}_a = \{j \in \mathcal{I}_b / \mathcal{V}_j^a \neq \emptyset\}$ est l'ensemble des indices des cellules qui possèdent au moins un côté appartenant à Γ_h^a .

$$\mathcal{V}_j^d \cup \mathcal{V}_j^m \cup \mathcal{V}_j^a = \mathcal{V}_j.$$

$$\mathcal{V}_j^d \cap \mathcal{V}_j^m = \mathcal{V}_j^d \cap \mathcal{V}_j^a = \mathcal{V}_j^a \cap \mathcal{V}_j^m = \emptyset.$$

$\mathbf{n}_j \in \mathbf{R}^3$ est la normale unitaire sortante à la surface ∂C_j de C_j .

$\mathbf{n}_{jk} = (n_{jk}^{x_1}, n_{jk}^{x_2}, n_{jk}^{x_3})^t \in \mathbf{R}^3$ est la normale unitaire sortante à Σ_{jk} allant de C_j vers C_k .

$$\forall (j, k) \in \mathcal{I} \times (\mathcal{I} \cup \mathcal{I}_b), \mathbf{n}_{kj} = -\mathbf{n}_{jk}.$$

2.1 Problème aux limites et analyse préliminaire

On considère maintenant l'équation perturbée (2.16) et on suppose $\Gamma = \Gamma^a$. On doit alors résoudre le problème aux limites suivant :

$$\begin{cases} pG_0 \mathbf{W} + G_x \partial_x \mathbf{W} + G_y \partial_y \mathbf{W} + G_z \partial_z \mathbf{W} + K \mathbf{W} = \mathbf{S} & \text{dans } \Omega, \\ G_{\bar{\mathbf{n}}}^- \mathbf{W} = 0 & \text{sur } \Gamma, \end{cases} \quad (1.1)$$

où $p = \nu + i\omega$ avec $\nu > 0$, qui peut aussi se mettre sous la forme :

$$\begin{cases} pG_0 \mathbf{W} + \operatorname{div} F(\mathbf{W}) + K \mathbf{W} = \mathbf{S} & \text{dans } \Omega, \\ G_{\bar{\mathbf{n}}}^- \mathbf{W} = 0 & \text{sur } \Gamma. \end{cases} \quad (1.2)$$

avec $F = (G_x, G_y, G_z)^t$ le flux de taille 18×6 .

On désigne par Ω_h une approximation polyédrique du domaine Ω et on suppose avoir découpé Ω_h en une partition finie $\mathcal{I} \subset \mathbb{N}$ de cellules volumiques C_j (pouvant être par exemple tétraédriques ou hexaédriques) : $\bar{\Omega}_h = \bigcup_{j \in \mathcal{I}} C_j$. Le maillage correspondant est désigné par \mathcal{T}_h . Tout comme pour Ω , Ω_h possède des frontières de type métallique Γ_h^m et de type absorbante Γ_h^a , vérifiant $\Gamma_h = \partial\Omega_h = \Gamma_h^m \cup \Gamma_h^a$ avec $\Gamma_h^m \cap \Gamma_h^a = \emptyset$.

Dans chaque élément, la restriction de \mathbf{W} à la cellule C_j appartient à l'espace :

$$V_h = \{\mathbf{W} \in L^2(\Omega)^6 \mid \forall j \in \mathcal{I}, \mathbf{W}|_{C_j} \in \mathbb{P}_m(C_j)^6\}.$$

2.1.1 Méthode de volumes finis

On considère ici le cas $m = 0$ qui correspond à l'approximation volumes finis (cette méthode sera notamment étudiée plus en détail dans la section 2.2.1). L'approximation est donc discontinue, aucune relation de continuité n'étant imposée entre les éléments. On note \mathbf{W}_j la valeur moyenne du champ \mathbf{W} sur la cellule C_j :

$$\mathbf{W}_j = \frac{1}{V_j} \int_{C_j} \mathbf{W} dx.$$

Si on intègre l'équation (1.2) sur la cellule C_j on obtient :

$$\int_{C_j} (pG_0 + K)\mathbf{W} dx + \int_{C_j} \operatorname{div} F(\mathbf{W}) dx = \int_{C_j} \mathbf{S} dx,$$

ou bien encore :

$$\int_{C_j} (pG_0 + K)\mathbf{W} dx + \int_{\partial C_j} F(\mathbf{W}) \cdot \mathbf{n} ds = \int_{C_j} \mathbf{S} dx. \quad (1.3)$$

On utilise une fonction de flux numérique centré pour les interfaces internes et décentré pour les interfaces frontières absorbantes et on note par $G_{jk} = \sum_{l=1}^3 n_{jk}^{x_l} G_l$ où $\mathbf{n}_{jk} = (n_{jk}^{x_1}, n_{jk}^{x_2}, n_{jk}^{x_3})^t$ est la normale unitaire à l'interface entre les cellules C_j et C_k , dirigée de C_j vers C_k . On obtient alors :

$$\begin{aligned} (1.3) &\Leftrightarrow \int_{C_j} (pG_0 + K)\mathbf{W} dx + \sum_{k \in \mathcal{V}_j} \int_{\Sigma_{jk}} F(\mathbf{W}) \cdot \mathbf{n} ds = \int_{C_j} \mathbf{S} dx \\ &\Leftrightarrow \int_{C_j} (pG_0 + K)\mathbf{W} dx + \sum_{k \in \mathcal{V}_j^d} \int_{\Sigma_{jk}} G_{jk} \left(\frac{\mathbf{W}_j + \mathbf{W}_k}{2} \right) ds \\ &\quad + \sum_{k \in \mathcal{V}_j^a} \int_{\Sigma_{jk}} (G_{jk}^+ \mathbf{W}_j + G_{jk}^- \mathbf{W}_k) ds = \int_{C_j} \mathbf{S} dx. \end{aligned} \quad (1.4)$$

Comme nous le verrons plus loin, la condition absorbante correspond à $G_{jk}^- \mathbf{W}_k = 0$. Prenant cela en compte dans (1.4), on obtient :

$$\begin{aligned} (1.4) &\Leftrightarrow \int_{C_j} (pG_0 + K)\mathbf{W} dx + \sum_{k \in \mathcal{V}_j^d} \int_{\Sigma_{jk}} G_{jk} \left(\frac{\mathbf{W}_j + \mathbf{W}_k}{2} \right) ds \\ &\quad + \sum_{k \in \mathcal{V}_j^a} \int_{\Sigma_{jk}} G_{jk}^+ \mathbf{W}_j ds = \int_{C_j} \mathbf{S} dx. \end{aligned} \quad (1.5)$$

On peut alors démontrer le lemme suivant :

Lemme 1 *La solution du problème homogène associé à (1.2) vérifie la relation suivante :*

$$\frac{1}{2} \int_{\Gamma_h} (|G\mathbf{n}| \mathbf{W}, \mathbf{W}) ds + \int_{\Omega_h} ((\nu G_0 + K)\mathbf{W}, \mathbf{W}) dx = 0, \quad (1.6)$$

et on peut conclure que l'unique solution du problème discret homogène est la solution nulle.

Preuve Comme \mathbf{W} est constant sur la cellule C_j en prenant le produit scalaire de l'équation homogène associée à (1.5) avec \mathbf{W}_j on obtient :

$$\begin{aligned} \int_{C_j} ((pG_0 + K)\mathbf{W}_j, \mathbf{W}_j) dx + \sum_{k \in \mathcal{V}_j^d} \int_{\Sigma_{jk}} (G_{jk} \left(\frac{\mathbf{W}_j + \mathbf{W}_k}{2} \right), \mathbf{W}_j) ds \\ + \sum_{k \in \mathcal{V}_j^a} \int_{\Sigma_{jk}} (G_{jk}^+ \mathbf{W}_j, \mathbf{W}_j) ds = 0. \end{aligned}$$

Si C_j ne s'appuie pas sur la frontière Γ_h , la relation précédente s'écrit (en tenant compte du fait que $\sum_k G_{jk} = 0$) :

$$\int_{C_j} ((pG_0 + K)\mathbf{W}_j, \mathbf{W}_j)dx + \sum_{k \in \mathcal{V}_j^d} \int_{\Sigma_{jk}} (G_{jk} \frac{\mathbf{W}_k}{2}, \mathbf{W}_j)ds = 0, \quad (1.7)$$

et si C_j a une intersection non nulle avec la frontière Γ_h alors :

$$\begin{aligned} \int_{C_j} ((pG_0 + K)\mathbf{W}_j, \mathbf{W}_j)dx &+ \sum_{k \in \mathcal{V}_j^d} \int_{\Sigma_{jk}} (G_{jk} \frac{\mathbf{W}_k}{2}, \mathbf{W}_j)ds \\ &+ \frac{1}{2} \sum_{k \in \mathcal{V}_j^a} \int_{\Sigma_{jk}} (|G_{jk}| \mathbf{W}_j, \mathbf{W}_j)ds = 0. \end{aligned} \quad (1.8)$$

sachant que $G_{jk} = G_{jk}^+ + G_{jk}^-$ et $|G_{jk}| = G_{jk}^+ - G_{jk}^-$.

En sommant les relations (1.7) et (1.8) sur toutes les cellules C_j on obtient :

$$\begin{aligned} \frac{1}{2} \int_{\Gamma_h} (|\mathbf{G}_n| \mathbf{W}, \mathbf{W})ds + \int_{\Omega_h} ((\nu G_0 + K)\mathbf{W}, \mathbf{W})dx &+ \\ \frac{1}{2} \sum_{\Sigma_{jk} \in \mathcal{F}_h^{int}} \int_{\Sigma_{jk}} ((G_{jk} \mathbf{W}_k, \mathbf{W}_j) + (G_{kj} \mathbf{W}_j, \mathbf{W}_k)) ds &= 0, \end{aligned} \quad (1.9)$$

où on a noté par \mathcal{F}_h^{int} l'ensemble des faces internes. En tenant compte du fait que G_{jk} est hermitienne et que $G_{kj} = -G_{jk}$:

$$\begin{aligned} (G_{jk} \mathbf{W}_k, \mathbf{W}_j) + (G_{kj} \mathbf{W}_j, \mathbf{W}_k) &= (G_{jk} \mathbf{W}_k, \mathbf{W}_j) - \overline{(G_{jk} \mathbf{W}_k, \mathbf{W}_j)} \\ &= 2\Im((G_{jk} \mathbf{W}_k, \mathbf{W}_j)), \end{aligned}$$

et puisque les matrices $|\mathbf{G}_n|$ et $(\nu G_0 + K)$ sont hermitiennes (car symétriques réelles), on constate que la partie réelle de l'équation (1.9) est exactement la relation recherchée. Or, dans la mesure où les matrices $|\mathbf{G}_n|$ et $(\nu G_0 + K)$ sont respectivement positive et définie positive (car $\nu > 0$), on conclut que la seule solution du problème homogène est la solution nulle. ■

2.1.2 Méthode de type Galerkin discontinu

On considère ici le cas m quelconque. Cette méthode, plus particulièrement la variante $m = 1$, est étudiée plus en détail dans la section 2.2.2. Comme précédemment, l'approximation est donc discontinue, aucune relation de continuité n'étant imposée entre les éléments.

Si on introduit les espaces PH^s de fonctions H^s par morceaux, définis comme les fonctions qui sont localement H^s sur une partition de Ω (par exemple sur des ouverts Ω_j où les tenseurs ε et μ sont constants), en imposant que $\mathbf{W} \in (PH^s)^6$, alors on garantit l'existence des traces sur l'ensembles des faces frontières ou internes.

Lemme 2 Pour chaque cellule C_j et pour chaque fonction test $\mathbf{U} \in V_h$, la solution $\mathbf{W}_h \in V_h$ du problème homogène associé à (1.2) vérifie la relation suivante :

$$\begin{aligned} \int_{C_j} ((pG_0 + K)\mathbf{W}_j, \mathbf{U}_j) dx &+ \int_{C_j} (\operatorname{div} F(\mathbf{W}_j), \mathbf{U}_j) - (F(\mathbf{W}_j), \nabla \mathbf{U}_j) dx \\ &+ \sum_{k \in \mathcal{V}_j^d} \int_{\Sigma_{jk}} (G_{jk} \mathbf{W}_k, \mathbf{U}_j) ds \\ &+ \sum_{k \in \mathcal{V}_j^a} \int_{\Sigma_{jk}} (|G_{jk}| \mathbf{W}_j, \mathbf{U}_j) ds = 0. \end{aligned} \quad (1.10)$$

Preuve Si on intègre (1.2) sur une cellule C_j du maillage après avoir fait le produit scalaire avec une fonction test \mathbf{U} , on obtient :

$$\int_{C_j} ((pG_0 + K)\mathbf{W}_j, \mathbf{U}_j) dx + \int_{C_j} (\operatorname{div} F(\mathbf{W}_j), \mathbf{U}_j) dx = 0.$$

Le théorème de Green-Riemann nous permet alors d'écrire :

$$\begin{aligned} \int_{C_j} ((pG_0 + K)\mathbf{W}_j, \mathbf{U}_j) dx &- \int_{C_j} (F(\mathbf{W}_j), \nabla \mathbf{U}_j) dx \\ &+ \int_{\partial C_j} (F(\mathbf{W}_j) \cdot \mathbf{n}, \mathbf{U}_j) ds = 0. \end{aligned} \quad (1.11)$$

Comme dans le cas de la méthode volumes finis, on utilise une fonction de flux numérique centré pour les interfaces internes et décentré pour les interfaces frontières absorbantes. De même, les définitions de la matrice G_{jk} et de la normale unitaire \mathbf{n}_{jk} sont identiques, si bien qu'on obtient :

$$\begin{aligned} \int_{C_j} ((pG_0 + K)\mathbf{W}_j, \mathbf{U}_j) dx &- \int_{C_j} (F(\mathbf{W}_j), \nabla \mathbf{U}_j) dx \\ &+ \sum_{k \in \mathcal{V}_j^d} \int_{\Sigma_{jk}} (G_{jk} \left(\frac{\mathbf{W}_j + \mathbf{W}_k}{2} \right), \mathbf{U}_j) ds \\ &+ \sum_{k \in \mathcal{V}_j^a} \int_{\Sigma_{jk}} ((G_{jk}^+ \mathbf{W}_j + G_{jk}^- \mathbf{W}_k), \mathbf{U}_j) ds = 0. \end{aligned} \quad (1.12)$$

La condition absorbante correspond à $G_{jk}^- \mathbf{W}_k = 0$. Prenant cela en compte dans l'équation (1.12), on obtient :

$$\begin{aligned} \int_{C_j} ((pG_0 + K)\mathbf{W}_j, \mathbf{U}_j) dx &- \int_{C_j} (F(\mathbf{W}_j), \nabla \mathbf{U}_j) dx \\ &+ \sum_{k \in \mathcal{V}_j^d} \int_{\Sigma_{jk}} (G_{jk} \left(\frac{\mathbf{W}_j + \mathbf{W}_k}{2} \right), \mathbf{U}_j) ds \\ &+ \sum_{k \in \mathcal{V}_j^a} \int_{\Sigma_{jk}} (G_{jk}^+ \mathbf{W}_j, \mathbf{U}_j) ds = 0. \end{aligned} \quad (1.13)$$

Par ailleurs, le bilan des flux sur chaque cellule s'écrit :

$$\int_{\partial C_j} G \mathbf{n}_j \mathbf{W}_j ds = \sum_{k \in \mathcal{V}_j^a} \int_{\Sigma_{jk}} G_{jk} \mathbf{W}_j ds + \sum_{k \in \mathcal{V}_j^d} \int_{\Sigma_{jk}} G_{jk} \mathbf{W}_j ds,$$

qui, utilisé dans (1.13), permet d'écrire :

$$\begin{aligned} \int_{C_j} ((pG_0 + K) \mathbf{W}_j, \mathbf{U}_j) dx &- \int_{C_j} (F(\mathbf{W}_j), \nabla \mathbf{U}_j) dx \\ &+ \frac{1}{2} \int_{\partial C_j} (G \mathbf{n}_j \mathbf{W}_j, \mathbf{U}_j) ds \\ &+ \frac{1}{2} \sum_{k \in \mathcal{V}_j^d} \int_{\Sigma_{jk}} (G_{jk} \mathbf{W}_k, \mathbf{U}_j) ds \\ &+ \sum_{k \in \mathcal{V}_j^a} \int_{\Sigma_{jk}} (G_{jk}^+ - \frac{G_{jk}}{2}) \mathbf{W}_j, \mathbf{U}_j) ds = 0, \end{aligned} \quad (1.14)$$

ce qui équivaut à :

$$\begin{aligned} 2 \int_{C_j} ((pG_0 + K) \mathbf{W}_j, \mathbf{U}_j) dx &- 2 \int_{C_j} (F(\mathbf{W}_j), \nabla \mathbf{U}_j) dx \\ &+ \int_{\partial C_j} (G \mathbf{n}_j \mathbf{W}_j, \mathbf{U}_j) ds \\ &+ \sum_{k \in \mathcal{V}_j^d} \int_{\Sigma_{jk}} (G_{jk} \mathbf{W}_k, \mathbf{U}_j) ds \\ &+ \sum_{k \in \mathcal{V}_j^a} \int_{\Sigma_{jk}} (|G_{jk}| \mathbf{W}_j, \mathbf{U}_j) ds = 0. \end{aligned} \quad (1.15)$$

Ensuite, en réutilisant le théorème de Green-Riemann, mais cette fois-ci dans l'autre sens, on trouve :

$$\begin{aligned} 2 \int_{C_j} ((pG_0 + K) \mathbf{W}_j, \mathbf{U}_j) dx &- 2 \int_{C_j} (F(\mathbf{W}_j), \nabla \mathbf{U}_j) dx \\ &+ \int_{C_j} (\operatorname{div} F(\mathbf{W}_j), \mathbf{U}_j) + (F(\mathbf{W}_j), \nabla \mathbf{U}_j) dx \\ &+ \sum_{k \in \mathcal{V}_j^d} \int_{\Sigma_{jk}} (G_{jk} \mathbf{W}_k, \mathbf{U}_j) ds \\ &+ \sum_{k \in \mathcal{V}_j^a} \int_{\Sigma_{jk}} (|G_{jk}| \mathbf{W}_j, \mathbf{U}_j) ds = 0. \end{aligned} \quad (1.16)$$

qui, en réduisant les termes, donne la relation désirée. ■

Si on somme les relations (1.16) sur toutes les cellules C_j on obtient la relation :

$$a(\mathbf{W}, \mathbf{U}) + b(\mathbf{W}, \mathbf{U}) + c(\mathbf{W}, \mathbf{U}) = 0,$$

où a , b et c sont les formes sesquilinéaires suivantes :

$$\begin{aligned}
a(\mathbf{W}, \mathbf{U}) &= 2 \int_{\Omega_h} ((pG_0 + K)\mathbf{W}, \mathbf{U}) dx, \\
b(\mathbf{W}, \mathbf{U}) &= \int_{\Omega_h} (\operatorname{div} F(\mathbf{W}), \mathbf{U}) - (F(\mathbf{W}), \nabla \mathbf{U}) dx, \\
c(\mathbf{W}, \mathbf{U}) &= \sum_{\Sigma_{jk} \subset \mathcal{F}^{int}} \int_{\Sigma_{jk}} ((G_{jk}\mathbf{W}_k, \mathbf{U}_j) + (G_{kj}\mathbf{W}_j, \mathbf{U}_k)) ds \\
&\quad + \int_{\partial\Omega_h} (|G_n|\mathbf{W}, \mathbf{U}) ds.
\end{aligned} \tag{1.17}$$

La proposition qui suit va nous permettre de prouver que notre problème discret est bien posé.

Proposition 1 *La forme sesquilinéaire b définie dans (1.17) vérifie :*

$$\Re(b(\mathbf{W}, \mathbf{W})) = 0$$

Preuve Grâce à la proposition 16 de l'annexe 2.4.7.4 qui permet d'établir que :

$$(F(\mathbf{W}), \nabla \mathbf{U}) = \overline{(\operatorname{div} F(\mathbf{U}), \mathbf{W})},$$

b peut se réécrire comme :

$$b(\mathbf{W}, \mathbf{U}) = \int_{\Omega_h} (\operatorname{div} F(\mathbf{W}), \mathbf{U}) - \overline{(\operatorname{div} F(\mathbf{U}), \nabla \mathbf{W})} dx$$

En passant à la forme quadratique associée, nous obtenons :

$$\begin{aligned}
b(\mathbf{W}, \mathbf{W}) &= \int_{\Omega_h} (\operatorname{div} F(\mathbf{W}), \mathbf{W}) - \overline{(\operatorname{div} F(\mathbf{W}), \nabla \mathbf{W})} dx \\
&= 2i \int_{\Omega_h} \Im((\operatorname{div} F(\mathbf{W}), \mathbf{W})) dx
\end{aligned}$$

ce qui implique bien évidemment $\Re(b(\mathbf{W}, \mathbf{W})) = 0$. ■

Proposition 2 *Les formes sesquilinéaires a , b et c définies dans (1.17) vérifient :*

$$\begin{aligned}
\Re(a(\mathbf{W}, \mathbf{W}) + b(\mathbf{W}, \mathbf{W}) + c(\mathbf{W}, \mathbf{W})) &= 2 \int_{\Omega_h} ((\nu G_0 + K)\mathbf{W}, \mathbf{W}) dx \\
&\quad + \int_{\Gamma_h} (|G_n|\mathbf{W}, \mathbf{W}) ds = 0,
\end{aligned} \tag{1.18}$$

et on peut conclure que l'unique solution du problème discret homogène est la solution nulle.

Preuve Dans un premier temps, sachant que $p = i\omega + \nu$ et que $(\nu G_0 + K)$ est une matrice symétrique réelle définie positive :

$$\Re(a(\mathbf{W}, \mathbf{W})) = \int_{\Omega_h} ((\nu G_0 + K)\mathbf{W}, \mathbf{W}) dx. \tag{1.19}$$

En outre, par le même biais que pour la méthode volumes finis, on trouve que :

$$(G_{jk} \mathbf{W}_k, \mathbf{W}_j) - (G_{jk} \mathbf{W}_j, \mathbf{W}_k) = (G_{jk} \mathbf{W}_k, \mathbf{W}_j) - \overline{(G_{jk} \mathbf{W}_k, \mathbf{W}_j)} = 2\Im((G_{jk} \mathbf{W}_k, \mathbf{W}_j)),$$

et on obtient donc :

$$\Re(c(\mathbf{W}, \mathbf{W})) = \int_{\Gamma_h} (|G_n| \mathbf{W}, \mathbf{W}) ds. \quad (1.20)$$

En sommant (1.19) et (1.20) et en tenant compte de la proposition 1, on retrouve l'égalité (1.18). Or, dans la mesure où $|G_n|$ et $(\nu G_0 + K)$ sont des matrices hermitiennes respectivement positive et définie positive (car $\nu > 0$), ceci entraîne que $\mathbf{W} = 0$ et la seule solution du problème homogène est donc la solution nulle. ■

2.2 Équations de Maxwell 1D

Pour dériver la version monodimensionnelle du système de Maxwell on prend comme direction de propagation de l'onde $\mathbf{k} = (k, 0, 0)^t$ ce qui correspond à la direction x , et la polarisation du champ électrique telle que $\mathbf{E} = (0, 0, E_z)^t$. La polarisation du champ magnétique est déduite du produit vectoriel $\mathbf{k} \times \mathbf{E}$ ce qui donne $\mathbf{H} = (0, H_y, 0)^t$. En supposant en plus que E_z et H_y sont des fonctions de x et t uniquement on obtient les équations de Maxwell en une dimension (on notera simplement E et H à la place de E_z et H_y). On suppose aussi que $\sigma = 0$ ce qui entraîne $J = 0$:

$$\begin{cases} \frac{\partial}{\partial t}(\varepsilon E) - \frac{\partial H}{\partial x} = 0, \\ \frac{\partial}{\partial t}(\mu H) - \frac{\partial E}{\partial x} = 0. \end{cases} \quad (2.21)$$

On souhaite étudier le système (2.21) sur l'intervalle $[a, b]$. On doit donc définir les conditions aux limites en $x = a$ et $x = b$. Rappelons qu'il est possible de résoudre les équations de Maxwell en champ total ou en champ diffracté. Pour formaliser simultanément les deux formulations, on introduit un paramètre ξ valant alternativement 0 si on souhaite résoudre (2.21) en champ diffracté (dans ce cas E désignera implicitement E^{dif}) ou 1 si on souhaite résoudre ce système en champ total (et E désignera E^{tot}).

En ce qui concerne les conditions aux limites, nous choisissons de considérer une condition métallique en $x = b$ vérifiant $E^{\text{tot}}(b) = 0$ et une condition absorbante de type Silver-Müller d'ordre 1 en $x = a$ vérifiant $[E^{\text{dif}} - zH^{\text{dif}}](a) = 0^1$. Si on écrit ces conditions de manière à concilier les deux formulations, on a :

$$\begin{aligned} \begin{cases} E^{\text{tot}}(b) = 0 \\ [E^{\text{dif}} - zH^{\text{dif}}](a) = 0 \end{cases} &\Leftrightarrow \begin{cases} E(b) + (1 - \xi)E^{\text{inc}}(b) = 0 \\ [E - \xi E^{\text{inc}}](a) - z[H - \xi H^{\text{inc}}](a) = 0 \end{cases} \\ &\Leftrightarrow \begin{cases} E(b) = (\xi - 1)E^{\text{inc}}(b) \\ [E - zH](a) = \xi[E^{\text{inc}} - zH^{\text{inc}}](a). \end{cases} \end{aligned}$$

¹Nous reviendrons plus loin sur l'origine de cette formulation

On procède maintenant à l'adimensionnement des équations (2.21) en effectuant les changements de variables suivants :

$$\left\{ \begin{array}{l} \tilde{x} = \frac{x}{L_0} \text{ et } \tilde{t} = \frac{c_0 t}{L_0} \\ \tilde{c} = \frac{c}{c_0} = c_r \\ \tilde{E} = \frac{E}{z_0 H_0} \\ \tilde{H} = \frac{H}{H_0} \end{array} \right. \quad \begin{array}{l} \text{et idem pour } E^{\text{inc}}, \\ \text{et idem pour } H^{\text{inc}}. \end{array}$$

où $c_0 = 1/\sqrt{\varepsilon_0 \mu_0}$, $z_0 = \sqrt{\mu_0/\varepsilon_0}$, H_0 est l'intensité du champ magnétique de référence et L_0 la longueur de référence. Dans ce cas, nous avons :

$$\left\{ \begin{array}{l} \partial_t(\varepsilon E) - \partial_x H = 0 \\ \partial_t(\mu H) - \partial_x E = 0 \\ E(b) = (\xi - 1)E^{\text{inc}}(b) \\ [E - zH](a) = \xi[E^{\text{inc}} - zH^{\text{inc}}](a) \end{array} \right. \Leftrightarrow \left\{ \begin{array}{l} \partial_t(\varepsilon_0 \varepsilon_r z_0 H_0 \tilde{E}) - \partial_x(H_0 \tilde{H}) = 0 \\ \partial_t(\mu_0 \mu_r H_0 \tilde{H}) - \partial_x(z_0 H_0 \tilde{E}) = 0 \\ [z_0 H_0 \tilde{E}](b) = (\xi - 1)[z_0 H_0 \tilde{E}^{\text{inc}}](b) \\ [z_0 H_0 \tilde{E} - z_0 z_r H_0 \tilde{H}](a) = \xi[z_0 H_0 \tilde{E}^{\text{inc}} - z_0 z_r H_0 \tilde{H}^{\text{inc}}](a) \end{array} \right.$$

soit encore, en simplifiant par H_0 et en tenant compte de l'adimensionnement de t et x :

$$\left\{ \begin{array}{l} \partial_{\tilde{t}}\left(\frac{\varepsilon_0 \varepsilon_r z_0 c_0}{L_0} \tilde{E}\right) - \partial_{\tilde{x}}\left(\frac{\tilde{H}}{L_0}\right) = 0 \\ \partial_{\tilde{t}}\left(\frac{\mu_0 \mu_r c_0}{L_0} \tilde{H}\right) - \partial_{\tilde{x}}\left(\frac{z_0 \tilde{E}}{L_0}\right) = 0 \\ [z_0 H_0 \tilde{E}](b) = (\xi - 1)[z_0 H_0 \tilde{E}^{\text{inc}}](b) \\ [z_0 H_0 \tilde{E} - z_0 z_r H_0 \tilde{H}](a) = \xi[z_0 H_0 \tilde{E}^{\text{inc}} - z_0 z_r H_0 \tilde{H}^{\text{inc}}](a), \end{array} \right.$$

avec les notations classiques $\partial_t = \frac{\partial}{\partial t}$ et $\partial_x = \frac{\partial}{\partial x}$. Pour conclure, après simplification, on obtient :

$$\left\{ \begin{array}{l} \partial_{\tilde{t}}(\varepsilon_r \tilde{E}) - \partial_{\tilde{x}}(\tilde{H}) = 0 \\ \partial_{\tilde{t}}(\mu_r \tilde{H}) - \partial_{\tilde{x}}(\tilde{E}) = 0 \\ \tilde{E}(b) = (\xi - 1)\tilde{E}^{\text{inc}}(b) \\ [\tilde{E} - z_r \tilde{H}](a) = \xi[\tilde{E}^{\text{inc}} - z_r \tilde{H}^{\text{inc}}](a), \end{array} \right. \quad (2.22)$$

Dans la suite, pour alléger les notations, nous abandonnons le symbole $\tilde{}$ pour toutes les variables et nous notons simplement ε au lieu de ε_r (et idem pour μ_r). On obtient alors un système formellement équivalent au système initial (2.21). On introduit les quantités suivantes :

$$W = \begin{pmatrix} E \\ H \end{pmatrix}, \quad \theta = \begin{pmatrix} \varepsilon & 0 \\ 0 & \mu \end{pmatrix}, \quad A = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}, \quad F(W) = AW. \quad (2.23)$$

Alors, les deux premières équations du système (2.22) peuvent se réécrire sous la forme pseudo-conservative suivante :

$$\theta \frac{\partial W(t, x)}{\partial t} + \frac{\partial F(W(t, x))}{\partial x} = 0. \quad (2.24)$$

Nous supposons maintenant que le champ W suit une dépendance harmonique en temps de pulsation $\tilde{\omega} \in \mathbf{R}$ vérifiant $\tilde{\omega} = \omega/c_0$. Tout comme pour les autres variables, on assimile $\tilde{\omega}$ à ω , si bien que (2.24) devient :

$$i\omega\theta W(x) + \frac{\partial F(W(x))}{\partial x} = S, \quad (2.25)$$

où on rappelle que le terme source S est nul si on travaille en champ total et non nul dans le cas contraire.

2.2.1 Méthode de volumes finis

2.2.1.1 Formulation du schéma

On discrétise l'intervalle $[a, b]$ en N cellules C_j de taille Δx_j : $C_j = [x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}]$ avec $x_{\frac{1}{2}} = a$ et $x_{N+\frac{1}{2}} = b$ (voir la figure 2.1).

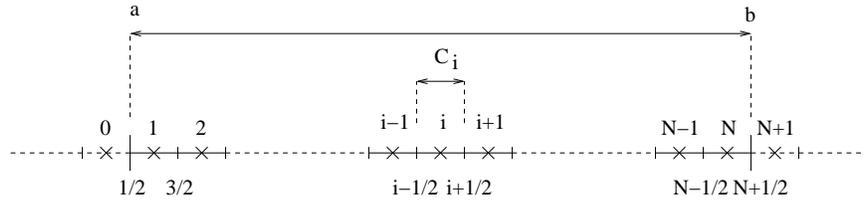


FIG. 2.1 – Discrétisation de l'intervalle $[a, b]$

Les cellules C_0 et C_{N+1} représentent des cellules fictives qui serviront dans ce qui suit à calculer les flux aux points $x = a$ et $x = b$.

Pour une fonction vectorielle W on note W_j la valeur moyenne de W sur la cellule C_j :

$$W_j = \frac{1}{\Delta x_j} \int_{C_j} W(x) dx.$$

On notera également $W_{j\pm\frac{1}{2}}$ la valeur approchée de W à l'interface entre les cellules C_j et $C_{j\pm 1}$ ($W_{j\pm\frac{1}{2}} \approx W(x_{j\pm\frac{1}{2}})$). En intégrant l'équation (2.25) sur la cellule $\{C_j\}_{1 \leq j \leq N}$, on obtient :

$$\begin{aligned} & \int_{C_j} [i\omega\theta W(x) + \partial_x F(W(x))] dx = \int_{C_j} S(x) dx \\ \Leftrightarrow & i\omega \int_{C_j} \theta W(x) dx + \int_{C_j} \partial_x F(W(x)) dx = \int_{C_j} S(x) dx \\ \Leftrightarrow & i\omega \Delta x_j \theta_j W_j + \left(F(W(x_{j+\frac{1}{2}})) - F(W(x_{j-\frac{1}{2}})) \right) = \Delta x_j S_j \\ \Leftrightarrow & i\omega \Delta x_j \theta_j W_j + \left(F_{j+\frac{1}{2}} - F_{j-\frac{1}{2}} \right) = \Delta x_j S_j, \end{aligned} \quad (2.26)$$

où ε_j et μ_j sont les caractéristiques électromagnétiques associées à la cellule C_j , supposées constantes sur C_j . Par ailleurs, $F_{j\pm\frac{1}{2}}$ est une approximation du flux $F(W_{j\pm\frac{1}{2}})$ aux frontières des cellules. Pour les faces internes $\{x_{j+\frac{1}{2}}\}_{1\leq j\leq N-1}$, ces flux sont calculés de manière centrée [Piperno et al., 2002] :

$$\begin{cases} F_{j-\frac{1}{2}} = \frac{1}{2}(F_j + F_{j-1}) = A \left(\frac{W_j + W_{j-1}}{2} \right) \\ F_{j+\frac{1}{2}} = \frac{1}{2}(F_j + F_{j+1}) = A \left(\frac{W_j + W_{j+1}}{2} \right) \end{cases}$$

Si on injecte ces expressions dans (2.26), on obtient pour $j \in \llbracket 2; N-1 \rrbracket$:

$$\begin{aligned} i\omega\Delta x_j\theta_j W_j + F_{j+\frac{1}{2}} - F_{j-\frac{1}{2}} &= \Delta x_j S_j \\ \Leftrightarrow i\omega\Delta x_j\theta_j W_j + \frac{1}{2}(AW_{j+1} + AW_j) - \frac{1}{2}(AW_j + AW_{j-1}) &= \Delta x_j S_j, \end{aligned} \quad (2.27)$$

ce qui donne :

$$-AW_{j-1} + 2i\omega\Delta x_j\theta_j W_j + AW_{j+1} = 2\Delta x_j S_j. \quad (2.28)$$

Il s'agit maintenant de prendre en compte les conditions aux limites. Mais nous devons au préalable réaliser un calcul préliminaire qui nous sera utile pour le traitement de la condition absorbante de type Silver-Müller.

2.2.1.2 Diagonalisation de la matrice de flux

Le système de Maxwell temporel (2.22) est de nature hyperbolique donc diagonalisable. Cependant, cette diagonalisation doit s'opérer sur le système de Maxwell écrit sous forme conservative :

$$\frac{\partial Q}{\partial t} + \frac{\partial G(Q)}{\partial x} = 0, \quad (2.29)$$

avec :

$$Q = \begin{pmatrix} D \\ B \end{pmatrix} = \begin{pmatrix} \varepsilon & 0 \\ 0 & \mu \end{pmatrix} \begin{pmatrix} E \\ H \end{pmatrix} \quad \text{et} \quad G(Q) = MQ \quad \text{avec} \quad M = \begin{pmatrix} 0 & -\frac{1}{\mu} \\ -\frac{1}{\varepsilon} & 0 \end{pmatrix}$$

La diagonalisation de M s'écrit alors $M = P\Lambda P^{-1}$ avec :

$$\Lambda = \begin{pmatrix} -c & 0 \\ 0 & c \end{pmatrix}, \quad P = \begin{pmatrix} 1 & 1 \\ z & -z \end{pmatrix} \quad \text{et} \quad P^{-1} = \frac{1}{2} \begin{pmatrix} 1 & z^{-1} \\ 1 & -z^{-1} \end{pmatrix}.$$

D'autre part, si introduit les notations $x^+ = \max(x, 0)$ et $x^- = \min(x, 0)$ pour $x \in \mathbf{R}$, on a que $x = x^+ + x^-$ et que $|x| = x^+ - x^-$. En appliquant ces notations aux éléments des matrices ci-dessus, on définit :

$$\Lambda^- = \begin{pmatrix} -c & 0 \\ 0 & 0 \end{pmatrix}, \quad \Lambda^+ = \begin{pmatrix} 0 & 0 \\ 0 & c \end{pmatrix} \quad \text{et} \quad |\Lambda| = \begin{pmatrix} c & 0 \\ 0 & c \end{pmatrix}.$$

De là, on définit aussi les matrices de flux positifs et négatifs, et on retrouve alors les relations exprimées plus haut. On a en effet : $M^\pm = P\Lambda^\pm P^{-1}$ et $|M| = P|\Lambda|P^{-1}$. Ces matrices vérifient bien $M = M^+ + M^-$ et $|M| = M^+ - M^-$. Elles valent respectivement :

$$M^+ = \frac{1}{2} \begin{pmatrix} c & -\frac{1}{\mu} \\ -\frac{1}{\varepsilon} & c \end{pmatrix}, \quad M^- = \frac{1}{2} \begin{pmatrix} -c & -\frac{1}{\mu} \\ -\frac{1}{\varepsilon} & -c \end{pmatrix} \quad \text{et } |M| = |\Lambda| = cI_2.$$

2.2.1.3 Traitement des conditions aux limites

Le cadre est le suivant : on impose une condition absorbante au point $x = a$, une condition métallique au point $x = b$, et on souhaite envoyer une onde incidente W^{inc} qui se propage de $x = a$ vers l'intérieur de $[a, b]$. Comme nous l'avons précisé précédemment, cette onde incidente devra être nécessairement solution des équations de Maxwell. Pour satisfaire cette condition, nous pouvons prendre une onde plane de la forme $W^{\text{inc}}(x) = (\eta_e, \eta_h)e^{-kx}$ avec une condition sur le nombre d'onde k .

Pour traiter ces conditions aux limites, il convient de les respecter au sens faible. Ainsi, on approche les états aux points $x = a$ et $x = b$ par les formules suivantes :

$$\begin{aligned} - W(a) &\approx W_{\frac{1}{2}} = (W_0 + W_1)/2 \\ - W(b) &\approx W_{N+\frac{1}{2}} = (W_N + W_{N+1})/2 \end{aligned}$$

Néanmoins, les cellules C_0 et C_{N+1} sont fictives et on considère que les paramètres électromagnétiques ε et μ correspondent à ceux de leurs cellules voisines. De plus, il faut choisir pour W_0 et W_{N+1} des états fictifs respectivement associés aux cellules C_0 et C_{N+1} de sorte que $W_{\frac{1}{2}}$ et $W_{N+\frac{1}{2}}$ satisfassent respectivement les conditions absorbantes et métalliques. Ainsi, si on pose :

$$W_0 = \begin{pmatrix} 0 & z_1 \\ z_1^{-1} & 0 \end{pmatrix} W_1 + \xi \begin{pmatrix} 1 & -z_1 \\ -z_1^{-1} & 1 \end{pmatrix} W^{\text{inc}}, \quad (2.30)$$

$$W_{N+1} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} W_N + 2(\xi - 1) \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} W^{\text{inc}}, \quad (2.31)$$

et :

$$A_1 = \begin{pmatrix} 1 & -z_1 \\ -z_1^{-1} & 1 \end{pmatrix} \quad \text{et} \quad A_N = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},$$

alors :

$$\begin{aligned} A_1 \frac{W_0 + W_1}{2} &= \frac{1}{2} A_1 \left[\begin{pmatrix} 1 & z_1 \\ z_1^{-1} & 1 \end{pmatrix} W_1 + \xi \begin{pmatrix} 1 & -z_1 \\ -z_1^{-1} & 1 \end{pmatrix} W^{\text{inc}} \right] \\ &= \frac{1}{2} A_1 \underbrace{\begin{pmatrix} 1 & z_1 \\ z_1^{-1} & 1 \end{pmatrix} W_1}_{=0} + \frac{\xi}{2} \underbrace{A_1^2}_{=2A_1} W^{\text{inc}} \\ &= \xi A_1 W^{\text{inc}}, \end{aligned}$$

$$\begin{aligned} A_N \frac{W_N + W_{N+1}}{2} &= \frac{1}{2} A_N \left[\begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix} W_N + 2(\xi - 1) \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} W^{\text{inc}} \right] \\ &= \frac{1}{2} A_N \underbrace{\begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix} W_N}_{=0} + (\xi - 1) \underbrace{A_N^2}_{=A_N} W^{\text{inc}} \\ &= (\xi - 1) A_N W^{\text{inc}}. \end{aligned}$$

Compte tenu des propositions 4 et 5 énoncées dans l'annexe 2.4.1, les conditions limites sont bien respectées faiblement. Remarquons au passage que $Z_1^{-1}M_1^+\theta_1 = A_1/2$ si bien que la condition au bord absorbant s'écrit aussi :

$$A_1 \frac{W_0 + W_1}{2} = \xi A_1 W^{\text{inc}} = 2\xi Z_1^{-1} M_1^+ \theta_1 W^{\text{inc}}$$

avec $Z_1 = \begin{pmatrix} z_1^{-1} & 0 \\ 0 & z_1 \end{pmatrix}$.

Nous pouvons passer à l'approximation des flux $F_{\frac{1}{2}}$ et $F_{N+\frac{1}{2}}$ associés respectivement aux points $x = a$ et $x = b$. Tout comme pour les cellules internes, le calcul des flux aux frontières métallique et absorbante repose sur un schéma centré. Néanmoins, nous allons voir que dans le cas de la frontière absorbante, la définition de W_{N+1} (2.30) va en fait conduire à l'utilisation d'un schéma décentré pour le calcul du flux en $x = b$.

Pour $j = \frac{1}{2}$ le flux correspondant s'exprime par $F_{\frac{1}{2}} = (AW_0 + AW_1)/2$. Compte tenu de la condition (2.30) et des paramètres électromagnétiques de C_0 identifiés à ceux de C_1 , on obtient :

$$\begin{aligned} F_{\frac{1}{2}} &= \frac{1}{2} A \left[\begin{pmatrix} 0 & z_1 \\ z_1^{-1} & 0 \end{pmatrix} W_1 + \xi \begin{pmatrix} 1 & -z_1 \\ -z_1^{-1} & 1 \end{pmatrix} W^{\text{inc}} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} W_1 \right] \\ &= \frac{1}{2} A \left[\begin{pmatrix} 1 & z_1 \\ z_1^{-1} & 1 \end{pmatrix} W_1 + \xi \begin{pmatrix} 1 & -z_1 \\ -z_1^{-1} & 1 \end{pmatrix} W^{\text{inc}} \right] \\ &= \frac{1}{2} \left[\begin{pmatrix} -z_1^{-1} & -1 \\ -1 & -z_1 \end{pmatrix} W_1 + \xi \begin{pmatrix} z_1^{-1} & -1 \\ -1 & z_1 \end{pmatrix} W^{\text{inc}} \right] \\ &= M_1^- \theta_1 W_1 + \xi M_1^+ \theta_1 W^{\text{inc}} = M_1^- Q_1 + \xi M_1^+ Q^{\text{inc}} \end{aligned}$$

On constate que le flux $F_{\frac{1}{2}}$ est finalement évalué au moyen d'un schéma décentré similaire au schéma de Steger et Warming [Steger and Warming, 1981] souvent adopté pour le traitement des conditions aux limites à l'infini dans la résolution numérique des systèmes d'EDPs de la mécanique des fluides compressibles (voir par exemple [Lanteri, 1996]). Ce schéma permet de différencier les ondes entrantes des ondes sortantes à la frontière d'une cellule. Ainsi, la matrice des flux positifs (construite à l'aide des valeurs propres positives) permet de considérer les ondes allant de $\{a\}$ vers $\{b\}$, c'est-à-dire allant de l'extérieur vers l'intérieur du domaine. Quant à la matrice des flux négatifs (construite à l'aide des valeurs propres négatives), elle permet de considérer les ondes allant de $\{b\}$ vers $\{a\}$, c'est-à-dire allant de l'intérieur vers l'extérieur du domaine. La condition absorbante au point $x = a$ exprime la non-réflexion des ondes provenant de la cellule C_1 vers l'intérieur du domaine. Les seules ondes entrantes sont les ondes incidentes.

Le terme $M_1^+ \theta_1 W^{\text{inc}}$ est donc une donnée du problème et, par souci de commodité, nous posons $\gamma_1^{\text{inc}} = M_1^+ \theta_1 W^{\text{inc}}$. Par suite :

$$F_{\frac{1}{2}} = M_1^- \theta_1 W_1 + \xi \gamma_1^{\text{inc}}. \quad (2.32)$$

En injectant (2.32) dans la dernière équation de (2.26), on obtient :

$$\begin{aligned}
& i\omega\Delta x_1\theta_1W_1 + F_{\frac{3}{2}} - F_{\frac{1}{2}} = \Delta x_1S_1 \\
\Leftrightarrow & i\omega\Delta x_1\theta_1W_1 + \frac{1}{2}(AW_1 + AW_2) - (M_1^-\theta_1W_1 + \xi\gamma_1^{\text{inc}}) = \Delta x_1S_1 \\
\Leftrightarrow & \left(i\omega\Delta x_1l_2 + \frac{1}{2}A\theta_1^{-1} - M_1^- \right) \theta_1W_1 + \frac{1}{2}AW_2 = \Delta x_1S_1 + \xi\gamma_1^{\text{inc}} \\
\Leftrightarrow & (2i\omega\Delta x_1l_2 + M_1 - 2M_1^-) \theta_1W_1 + AW_2 = 2(\Delta x_1S_1 + \xi\gamma_1^{\text{inc}}).
\end{aligned}$$

Or $M_1 - 2M_1^- = (M_1^+ + M_1^-) - 2M_1^- = |M_1| = c_1l_2$ et on obtient donc pour la première cellule :

$$(2i\omega\Delta x_1 + c_1) \theta_1W_1 + AW_2 = 2(\Delta x_1S_1 + \xi\gamma_1^{\text{inc}}). \quad (2.33)$$

Similairement, le flux pour $j = \frac{N}{2}$ s'écrit $F_{N+\frac{1}{2}} = (AW_N + AW_{N+1})/2$. La cellule C_{N+1} étant fictive, nous allons imposer à cette cellule l'état miroir (2.31) pour la condition réfléchissante au point $x = b$. De plus, on identifie les paramètres électromagnétiques de C_{N+1} à ceux de C_N . On obtient alors :

$$\begin{aligned}
F_{N+\frac{1}{2}} &= \frac{1}{2}(AW_N + AW_{N+1}) \\
&= \frac{1}{2}A \left[W_N + \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} W_N + 2(\xi - 1) \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} W^{\text{inc}} \right] \\
&= \frac{1}{2}A \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix} W_N + (\xi - 1)A \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} W^{\text{inc}} \\
&= \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix} W_N + (\xi - 1) \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix} W^{\text{inc}}.
\end{aligned}$$

Si on note $P_N = \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix}$ et $\gamma_N^{\text{inc}} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} W^{\text{inc}}$, nous obtenons :

$$F_{N+\frac{1}{2}} = P_NW_N + (1 - \xi)\gamma_N^{\text{inc}}. \quad (2.34)$$

En injectant (2.34) dans (2.26), on en déduit :

$$\begin{aligned}
& i\omega\Delta x_N\theta_NW_N + F_{N+\frac{1}{2}} - F_{N-\frac{1}{2}} = \Delta x_NS_N \\
\Leftrightarrow & i\omega\Delta x_N\theta_NW_N + \left(P_NW_N + (1 - \xi)\gamma_N^{\text{inc}} \right) - \frac{1}{2}(AW_{N-1} + AW_N) = \Delta x_NS_N \\
\Leftrightarrow & -\frac{1}{2}AW_{N-1} + \left(i\omega\Delta x_N\theta_N + P_N - \frac{1}{2}A \right) W_N = \Delta x_NS_N + (\xi - 1)\gamma_N^{\text{inc}} \\
\Leftrightarrow & -AW_{N-1} + (2i\omega\Delta x_Nl_2 + (2P_N - A)\theta_N^{-1}) \theta_NW_N = 2(\Delta x_NS_N + (\xi - 1)\gamma_N^{\text{inc}}).
\end{aligned}$$

Or, si on note $R_N = \begin{pmatrix} 0 & -\mu_N^{-1} \\ \varepsilon_N^{-1} & 0 \end{pmatrix}$, alors $(2P_N - A)\theta_N^{-1} = R_N$ et donc :

$$-AW_{N-1} + (2i\omega\Delta x_Nl_2 + R_N) \theta_NW_N = 2(\Delta x_NS_N + (\xi - 1)\gamma_N^{\text{inc}}). \quad (2.35)$$

Nous avons finalement évalué les flux aux interfaces de chaque cellule du maillage. En rassemblant les expressions (2.28), (2.33), et (2.35), nous nous ramenons à la formulation du système d'équations suivant :

$$\begin{cases} (2i\omega\Delta x_1 + c_1)\theta_1 W_1 + AW_2 = 2(\Delta x_1 S_1 + \xi\gamma_1^{\text{inc}}), \\ -AW_{j-1} + 2i\omega\Delta x_j\theta_j W_j + AW_{j+1} = 2\Delta x_j S_j, \text{ pour } 2 \leq j \leq N-1, \\ -AW_{N-1} + (2i\omega\Delta x_N l_2 + R_N)\theta_N W_N = 2(\Delta x_N S_N + (\xi-1)\gamma_N^{\text{inc}}). \end{cases} \quad (2.36)$$

2.2.1.4 Étude du problème discret

Dans cette section, on va d'abord écrire sous forme matricielle le système (2.36) en exhibant les propriétés des matrices qui interviennent. On va ensuite démontrer l'inversibilité du système linéaire résultant.

Avant d'exprimer le système (2.36) sous forme matricielle, nous allons le modifier de façon à obtenir des propriétés remarquables sur la matrice correspondante. On définit tout d'abord :

$$\Theta = \begin{pmatrix} \theta_1 & 0_2 & \dots & \dots & 0_2 \\ 0_2 & \theta_2 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \theta_{N-1} & 0_2 \\ 0_2 & \dots & \dots & 0_2 & \theta_N \end{pmatrix} \quad \text{avec} \quad \theta_j = \begin{pmatrix} \varepsilon_j & 0 \\ 0 & \mu_j \end{pmatrix}.$$

Comme $\forall j \in \llbracket 1; N \rrbracket$, $\varepsilon_j > 0$ et $\mu_j > 0$, il est clair que Θ est une matrice définie positive. Si on pose :

$$D_j = 2\omega\Delta x_j\theta_j, \quad Z_j = c_j\theta_j = \begin{pmatrix} z_j^{-1} & 0 \\ 0 & z_j \end{pmatrix}, \quad \text{et si on note que } \bar{R}_N \equiv R_N\theta_N = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

alors on peut réécrire le système (2.36) sous la forme :

$$\begin{cases} (iD_1 + Z_1)W_1 + AW_2 = 2(\Delta x_1 S_1 + \xi\gamma_1^{\text{inc}}), \\ -AW_{j-1} + iD_j W_j + AW_{j+1} = 2\Delta x_j S_j, \text{ pour } 2 \leq j \leq N-1, \\ -AW_{N-1} + (iD_N + \bar{R}_N)W_N = 2(\Delta x_N S_N + (\xi-1)\gamma_N^{\text{inc}}), \end{cases} \quad (2.37)$$

et, sous forme matricielle :

$$\mathcal{A}W = b \quad (2.38)$$

où \mathcal{A} , \mathbf{W} et \mathbf{b} sont donnés par :

$$\mathcal{A} = \begin{pmatrix} iD_1 + Z_1 & A & 0_2 & \dots & \dots & 0_2 \\ -A & iD_2 & A & \ddots & & \vdots \\ 0_2 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0_2 \\ \vdots & & \ddots & \ddots & iD_{N-1} & A \\ 0_2 & \dots & \dots & 0_2 & -A & iD_N + \bar{R}_N \end{pmatrix},$$

$$\mathbf{W} = \begin{pmatrix} W_1 \\ W_2 \\ \vdots \\ \vdots \\ \vdots \\ W_{N-1} \\ W_N \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 2(\Delta x_1 S_1 + \xi \gamma_1^{\text{inc}}) \\ 2\Delta x_2 S_2 \\ \vdots \\ \vdots \\ \vdots \\ 2\Delta x_{N-1} S_{N-1} \\ 2(\Delta x_N S_N + (\xi - 1)\gamma_N^{\text{inc}}) \end{pmatrix}.$$

Nous allons maintenant montrer que le système linéaire (2.38) est toujours inversible, ce qui revient à démontrer l'existence et l'unicité de la solution du problème discrétisé. Compte tenu des propriétés de la matrice Θ , on est ramené à démontrer l'inversibilité de la matrice \mathcal{A} . Pour faciliter cette démonstration, on introduit les matrices \mathcal{D} , \mathcal{Z} et \mathcal{R} , définies par :

$$\mathcal{D} = \begin{pmatrix} D_1 & 0_2 & \dots & 0_2 \\ 0_2 & D_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0_2 \\ 0_2 & \dots & 0_2 & D_N \end{pmatrix}, \quad \mathcal{Z} = \begin{pmatrix} Z_1 & 0_2 & \dots & \dots & 0_2 \\ 0_2 & 0_2 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0_2 \\ 0_2 & \dots & \dots & 0_2 & 0_2 \end{pmatrix},$$

$$\mathcal{R} = \begin{pmatrix} 0_2 & A & 0_2 & \dots & 0_2 \\ -A & 0_2 & \ddots & \ddots & \vdots \\ 0_2 & \ddots & \ddots & \ddots & 0_2 \\ \vdots & \ddots & \ddots & 0_2 & A \\ 0_2 & \dots & 0_2 & -A & \bar{R}_N \end{pmatrix},$$

de telle sorte que $\mathcal{A} = i\mathcal{D} + \mathcal{Z} + \mathcal{R}$. On notera que \mathcal{Z} est une matrice réelle symétrique positive et que \mathcal{R} est une matrice antisymétrique. Nous pouvons alors démontrer le lemme suivant :

Lemme 3 *La matrice \mathcal{A} du système linéaire (2.38) est inversible.*

Preuve Supposons que $\mathcal{A}\mathbf{W} = 0$ pour $\mathbf{W} \in \mathbb{C}^{2 \times N}$. Montrons que $W_j = 0, \forall j \in \llbracket 1; N \rrbracket$. On constate dans un premier temps que :

$$\mathcal{A}\mathbf{W} = 0 \Rightarrow \mathbf{W}^* \mathcal{A}\mathbf{W} = 0 \Rightarrow \mathbf{W}^* (i\mathcal{D} + \mathcal{Z} + \mathcal{R})\mathbf{W} = 0 \Rightarrow \mathbf{W}^* (i\mathcal{D} + \mathcal{R})\mathbf{W} + \mathbf{W}^* \mathcal{Z}\mathbf{W} = 0.$$

où $i\mathcal{D} + \mathcal{R}$ est une matrice antihermitienne. Par suite, d'après la proposition 12 de l'annexe 2.4.5, $i\mathcal{D} + \mathcal{R} \in i\mathbb{R}$. D'autre part, puisque \mathcal{Z} est une matrice réelle positive, on a que $W^* \mathcal{Z} W \in \mathbb{R}_+$. En résumé, $\Re(W^* \mathcal{A} W) = W^* \mathcal{Z} W$ et $i\Im(W^* \mathcal{A} W) = W^* (i\mathcal{D} + \mathcal{R}) W$. Donc $W^* \mathcal{A} W = 0 \Rightarrow W^* \mathcal{Z} W = W^* (i\mathcal{D} + \mathcal{R}) W = 0$. Comme $W^* \mathcal{Z} W = W_1^* Z_1 W_1$, et comme Z_1 est une matrice définie positive, alors nécessairement $W_1 = 0$. Par ailleurs :

$$\mathcal{A} W = 0 \Rightarrow \forall j \in \llbracket 1; N \rrbracket, (\mathcal{A} W)_j = 0$$

En particulier, pour $j = 1$, on : $(iD_1 + Z_1)W_1 + AW_2 = 0$. Or, $W_1 = 0$, donc $AW_2 = 0$ et $W_2 = 0$ car A est inversible. On procède ensuite par récurrence pour $j \in \llbracket 3; N - 1 \rrbracket$. Supposons donc que $W_{j-1} = W_j = 0$. On a alors :

$$\mathcal{A} W = 0 \Rightarrow (\mathcal{A} W)_j = 0 \Rightarrow -AW_{j-1} + iD_j W_j + AW_{j+1} = 0 \Rightarrow AW_{j+1} = 0 \Rightarrow W_{j+1} = 0.$$

En conclusion, l'hypothèse de récurrence est vraie et donc :

$$\mathcal{A} W = 0 \Rightarrow \forall j \in \llbracket 1; N \rrbracket, W_j = 0 \Rightarrow W = 0$$

et \mathcal{A} est inversible. ■

Remarque 5 Dans la démonstration précédente, on peut se rendre compte que c'est la matrice \mathcal{Z} dont les termes correspondent aux flux associés aux frontières absorbantes (le point $x = a$ dans le cas présent) qui permet en partie de prouver l'inversibilité de la matrice \mathcal{A} . D'un point de vue purement algébrique, en l'absence de la matrice \mathcal{Z} , la partie réelle de la matrice \mathcal{A} serait antisymétrique et l'inversibilité de \mathcal{A} ne serait donc pas toujours garantie. D'un point de vue physique, les frontières absorbantes empêchent le champ électromagnétique d'être piégé et donc confronté à des modes de résonance selon que l'on choisisse une fréquence ω proche d'une fréquence propre en rapport avec la taille du domaine Ω . Cela signifie que le problème discret est bien posé s'il comporte une condition absorbante, ce qui est mathématiquement vérifié dans le cadre continu [Ihlenburg, 1998].

2.2.2 Méthode de type Galerkin discontinu

2.2.2.1 Formulation du schéma

On se donne donc exactement le même cadre que celui adopté pour la méthode de volumes finis. Nous partons donc de l'équation :

$$i\omega\theta W(x) + \frac{\partial F(W(x))}{\partial x} = S. \quad (2.39)$$

Une formulation de type Galerkin discontinu (GD) repose sur l'introduction d'un ensemble de fonctions de base locales $(\varphi_{jk})_{1 \leq k \leq d_j}$ pour chaque cellule $(C_j)_{1 \leq j \leq N}$ où d_j représente le nombre de degrés de liberté caractérisant l'approximation des variables d'état dans la cellule C_j . Dans chaque cellule, le champ $W(x)$ est approché par une combinaison linéaire des fonctions de base locales (par exemple polynômes de degré au plus m), supposées linéairement indépendantes qui engendrent un espace noté $\mathcal{P}_j = \mathbb{P}_m[C_j]$:

$$W_j(x) \equiv W(x)|_{C_j} = \sum_{k=1}^{d_j} W_{jk} \varphi_{jk}(x), \quad (2.40)$$

où l'on a implicitement considéré ici le cas de fonctions de base scalaires. Les degrés de liberté locaux sont $W_{jk} \in \mathbb{C}^2$. Soit φ_j une fonction quelconque de \mathcal{P}_j . On multiplie (2.39) par φ_j et on intègre sur la cellule C_j :

$$\begin{aligned} & \int_{C_j} [i\omega\theta W(x) + \partial_x F(W(x))] \varphi_j dx = \int_{C_j} S \varphi_j dx \\ \Leftrightarrow & i\omega \int_{C_j} \theta W(x) \varphi_j dx - \int_{C_j} F(W(x)) \partial_x \varphi_j dx + \\ & F(W(x_{j+\frac{1}{2}})) \varphi_j(x_{j+\frac{1}{2}}) - F(W(x_{j-\frac{1}{2}})) \varphi_j(x_{j-\frac{1}{2}}) = \int_{C_j} S \varphi_j dx. \end{aligned} \quad (2.41)$$

Posons $F_{j\pm\frac{1}{2}} = F(W(x_{j\pm\frac{1}{2}}))$ et $\varphi_j^\pm = \varphi_j(x_{j\pm\frac{1}{2}})$. Avec ces notations, l'équation (2.41) devient :

$$i\omega \int_{C_j} \theta W(x) \varphi_j dx - \int_{C_j} F(W(x)) \partial_x \varphi_j dx + F_{j+\frac{1}{2}} \varphi_j^+ - F_{j-\frac{1}{2}} \varphi_j^- = \int_{C_j} S \varphi_j dx. \quad (2.42)$$

On pose $\phi_j = (\varphi_{j1}, \varphi_{j2}, \dots, \varphi_{jd_j})$ la base locale de \mathcal{P}_j . Nous allons remplacer sur C_j W par son approximation $W_j \equiv W_j(x)$ défini par (2.40). Il s'agit maintenant d'injecter W_j dans l'équation (2.42). Notre problème est alors de chercher $W_j \in \mathcal{P}_j$ tel que $\forall \varphi_j \in \mathcal{P}_j$:

$$i\omega \int_{C_j} \theta W_j \varphi_j dx - \int_{C_j} F(W(x))_j \partial_x \varphi_j dx + F_{j+\frac{1}{2}} \varphi_j^+ - F_{j-\frac{1}{2}} \varphi_j^- = \int_{C_j} S_j \varphi_j dx. \quad (2.43)$$

Comme pour la méthode VF, le calcul des flux $F_{j\pm\frac{1}{2}}$ repose sur un schéma centré :

$$F_{j+\frac{1}{2}} = A \left(\frac{W_j^+ + W_{j+1}^-}{2} \right), \quad \text{avec } W_j^\pm = W_j(x_{j\pm\frac{1}{2}}), \quad (2.44)$$

ce qui conduit à :

$$\begin{aligned} \forall \varphi_i \in \mathcal{P}_j : \quad & i\omega \int_{C_j} \theta W_j \varphi_j dx - \int_{C_j} A W_j \partial_x \varphi_j dx \\ & + \frac{1}{2} (A W_j^+ + A W_{j+1}^-) \varphi_j^+ \\ & - \frac{1}{2} (A W_{j-1}^+ + A W_j^-) \varphi_j^- = \int_{C_j} S_j \varphi_j dx, \end{aligned} \quad (2.45)$$

c'est-à-dire :

$$\begin{aligned} \forall \varphi_j \in \mathcal{P}_j : \quad & i\omega \int_{C_j} \theta W_j \varphi_j dx - \int_{C_j} A W_j \partial_x \varphi_j dx \\ & + \frac{1}{2} A W_{j+1}^- \varphi_j^+ + \frac{1}{2} [A W_j \varphi_j]_{j-\frac{1}{2}}^{j+\frac{1}{2}} \\ & - \frac{1}{2} A W_{j-1}^+ \varphi_j^- = \int_{C_j} S_j \varphi_j dx. \end{aligned} \quad (2.46)$$

En utilisant la formule d'intégration par parties :

$$[AW_j\varphi_j]_{j-\frac{1}{2}}^{j+\frac{1}{2}} = \int_{C_j} AW_j\partial_x\varphi_j dx + \int_{C_j} A(W_j)_x\varphi_j dx,$$

on obtient, $\forall \varphi_j \in \mathcal{P}_j$:

$$\begin{aligned} i\omega\theta_j \int_{C_j} W_j\varphi_j dx + \frac{1}{2} \int_{C_j} (A(\partial_x W_j)\varphi_j - AW_j\partial_x\varphi_j) dx \\ + \frac{1}{2} AW_{j+1}^-\varphi_j^+ - \frac{1}{2} AW_{j-1}^+\varphi_j^- = \int_{C_j} S_j\varphi_j dx. \end{aligned} \quad (2.47)$$

où on a de plus supposé que les caractéristiques électromagnétiques ε_j et μ_j sont constantes sur C_j .

2.2.2.2 Traitement des conditions aux limites

Tout comme dans le cas de la méthode de volumes finis, nous utilisons les états auxiliaires suivants :

$$W_0(x_{\frac{1}{2}}) = \begin{pmatrix} 0 & z_1 \\ z_1^{-1} & 0 \end{pmatrix} W_1(x_{\frac{1}{2}}) + \xi \begin{pmatrix} 1 & -z_1 \\ -z_1^{-1} & 1 \end{pmatrix} W^{\text{inc}}(x_{\frac{1}{2}}), \quad (2.48)$$

$$W_{N+1}(x_{N+\frac{1}{2}}) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} W_N(x_{N+\frac{1}{2}}) + 2(\xi - 1) \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} W^{\text{inc}}(x_{N+\frac{1}{2}}), \quad (2.49)$$

qui sont telles que les conditions limites sont bien vérifiées de manière faible.

Remarque 6 Il faut ici lever une ambiguïté concernant W^{inc} en définissant ce champ incident vis-à-vis des cellules frontières C_1 et C_N . Ainsi, pour une formulation écrite par rapport à la cellule C_1 , on notera $(W^{\text{inc}})^-$ le champ W^{inc} pris au point $\frac{1}{2}$, et $(W^{\text{inc}})^+$ le champ W^{inc} pris au point $N + \frac{1}{2}$ pour une formulation écrite cette fois-ci par rapport à la cellule C_N .

Condition absorbante au point $x = a$. En utilisant (2.48) et en procédant comme pour la méthode de volumes finis :

$$\begin{aligned} AW_0^+ &= A \left(\begin{pmatrix} 0 & z_1 \\ z_1^{-1} & 0 \end{pmatrix} W_1^- + \xi \begin{pmatrix} 1 & -z_1 \\ -z_1^{-1} & 1 \end{pmatrix} (W^{\text{inc}})^- \right) \\ &= \begin{pmatrix} -z_1^{-1} & 0 \\ 0 & -z_1 \end{pmatrix} W_1^- + \xi \begin{pmatrix} z_1^{-1} & -1 \\ -1 & z_1 \end{pmatrix} (W^{\text{inc}})^- \\ &= -c_1\theta_1 W_1^- + 2\xi M_1^+ \theta_1 (W^{\text{inc}})^- \\ &= -c_1 Q_1^- + 2\xi M_1^+ (Q^{\text{inc}})^-. \end{aligned} \quad (2.50)$$

En réécrivant la formule (2.47) pour $j = 1$ et en se servant de (2.50), on trouve :

$$\begin{aligned}
& i\omega\theta_1 \int_{C_1} W_1\varphi_1 dx + \frac{1}{2} \int_{C_1} (A(\partial_x W_1)\varphi_1 - AW_1\partial_x\varphi_1) dx \\
& + \frac{1}{2}AW_2^-\varphi_1^+ - \frac{1}{2}AW_0^+\varphi_1^- = \int_{C_1} S_1\varphi_1 dx \\
\Leftrightarrow & i\omega\theta_1 \int_{C_1} W_1\varphi_1 dx + \frac{1}{2} \int_{C_1} (A(\partial_x W_1)\varphi_1 - AW_1\partial_x\varphi_1) dx \\
& + \frac{1}{2}AW_2^-\varphi_1^+ - \frac{1}{2}(-c_1Q_1^- + 2\xi M_1^+(Q^{\text{inc}})^-)\varphi_1^- = \int_{C_1} S_1\varphi_1 dx.
\end{aligned}$$

On obtient donc :

$$\begin{aligned}
\forall \varphi_1 \in \mathcal{P}_1, \quad i\omega\theta_1 \int_{C_1} W_1\varphi_1 dx + \frac{1}{2} \int_{C_1} (A(\partial_x W_1)\varphi_1 - AW_1\partial_x\varphi_1) dx \\
+ \frac{1}{2}c_1Q_1^-\varphi_1^- + \frac{1}{2}AW_2^-\varphi_1^+ \\
= \int_{C_1} S_1\varphi_1 dx + \xi M_1^+(Q^{\text{inc}})^-\varphi_1^-.
\end{aligned} \tag{2.51}$$

Condition métallique au point $x = b$. En utilisant (2.49) et en notant :

$$\rho_N = (\xi - 1) \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \quad \bar{R}_N = R_N\theta_N = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad \text{et} \quad R_N = \begin{pmatrix} 0 & -\mu_N^{-1} \\ \varepsilon_N^{-1} & 0 \end{pmatrix}$$

alors :

$$\begin{aligned}
AW_{N+1}^- &= A \left(\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} W_N^+ + 2(\xi - 1) \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} (W^{\text{inc}})^+ \right) \\
&= \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} W_N^+ + 2(\xi - 1) \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix} (W^{\text{inc}})^+ \\
&= \bar{R}_N W_N^+ - 2\rho_N (W^{\text{inc}})^+.
\end{aligned} \tag{2.52}$$

En réécrivant cette fois-ci la formule (2.47) pour $j = N$ et en se servant de (2.52), on trouve :

$$\begin{aligned}
& i\omega\theta_N \int_{C_N} W_N\varphi_N dx + \frac{1}{2} \int_{C_N} (A(\partial_x W_N)\varphi_N - AW_N\partial_x\varphi_N) dx \\
& - \frac{1}{2}AW_{N-1}^+\varphi_N^- + \frac{1}{2}AW_{N+1}^-\varphi_N^+ = \int_{C_N} S_N\varphi_N dx \\
\Leftrightarrow & i\omega\theta_N \int_{C_N} W_N\varphi_N dx + \frac{1}{2} \int_{C_N} (A(\partial_x W_N)\varphi_N - AW_N\partial_x\varphi_N) dx \\
& - \frac{1}{2}AW_{N-1}^+\varphi_N^- + \frac{1}{2}(\bar{R}_N W_N^+ - 2\rho_N (W^{\text{inc}})^+)\varphi_N^+ = \int_{C_N} S_N\varphi_N dx.
\end{aligned}$$

On obtient donc pour tout $\varphi_N \in \mathcal{P}_N$:

$$\begin{aligned} & i\omega\theta_N \int_{C_N} W_N \varphi_N dx + \frac{1}{2} \int_{C_N} (A(\partial_x W_N)\varphi_N - AW_N \partial_x \varphi_N) dx \\ & + \frac{1}{2} \bar{R}_N W_N^+ \varphi_N^+ - \frac{1}{2} AW_{N-1}^+ \varphi_N^- = \int_{C_N} S_N \varphi_N dx + \rho_N (W^{\text{inc}})^+ \varphi_N^+. \end{aligned} \quad (2.53)$$

2.2.2.3 Formulation matricielle

Posons $\mathbb{W}_j = (W_{j1}, W_{j2}, \dots, W_{jd_j})$ une matrice complexe de taille $2 \times d_j$. La relation (2.40) s'écrit $W_j = \mathbb{W}_j \phi_j^t$ où $\phi_j = (\varphi_{j1}, \varphi_{j2}, \dots, \varphi_{jd_j})^t$. En injectant cette expression dans (2.47) qui a été préalablement multipliée par 2, on obtient :

$$\begin{aligned} \forall \varphi_j \in \mathcal{P}_j : \quad & 2i\omega \int_{C_j} \theta_j \mathbb{W}_j \phi_j^t \varphi_j dx + \int_{C_j} (A \mathbb{W}_j (\partial_x \phi_j)^t \varphi_j - A \mathbb{W}_j \phi_j^t \partial_x \varphi_j) dx \\ & + A \mathbb{W}_{j+1} (\phi_{j+1}^-)^t \varphi_j^+ - A \mathbb{W}_{j-1} (\phi_{j-1}^+)^t \varphi_j^- = 2 \int_{C_j} \mathbb{S}_j \phi_j^t \varphi_j dx, \end{aligned} \quad (2.54)$$

Introduisons maintenant les notations suivantes :

$$\begin{cases} \Phi_j & = \int_{C_j} \phi_j^t \phi_j dx, \\ \Phi'_j & = \int_{C_j} (\phi_j^t \partial_x \phi_j - (\partial_x \phi_j)^t \phi_j) dx, \\ \Phi_j^\pm & = (\phi_j^\pm)^t \phi_j^\pm = \phi_j^t(x_{j\pm\frac{1}{2}}) \phi_j^t(x_{j\pm\frac{1}{2}}), \\ \Phi_{j\pm 1, j} & = (\phi_{j\pm 1}^\mp)^t \phi_j^\pm = \phi_{j\pm 1}^t(x_{j\pm\frac{1}{2}}) \phi_j^t(x_{j\pm\frac{1}{2}}). \end{cases} \quad (2.55)$$

On peut vérifier que les relations ci-dessous s'appliquent :

$$\begin{cases} \Phi_j^t & = \Phi_j, \\ (\Phi'_j)^t & = -\Phi'_j, \\ (\Phi_j^\pm)^t & = \Phi_j^\pm, \\ (\Phi_{j, j\pm 1})^t & = \Phi_{j\pm 1, j}. \end{cases} \quad (2.56)$$

Clairement la matrice de masse locale Φ_j est symétrique définie positive, Φ_j^\pm est symétrique, alors que Φ'_j est antisymétrique. Maintenant, en prenant pour φ_j successivement toutes les fonctions de base φ_{jk} pour $1 \leq k \leq d_j$, on peut reformuler les équations (2.51), (2.47) et (2.53) comme suit :

$$2i\omega\theta_1 \mathbb{W}_1 \Phi_1 + A \mathbb{W}_1 (\Phi'_1)^t + A \mathbb{W}_2 \Phi_{2,1} + Z_1 \mathbb{W}_1 \Phi_1^- = 2 \left(\mathbb{S}_1 \Phi_1 + \xi M_1^+ \theta_1 \mathbb{W}^{\text{inc}} \Phi_1^- \right), \quad (2.57)$$

$$2 \leq j \leq N-1, \quad 2i\omega\theta_j \mathbb{W}_j \Phi_j + A \mathbb{W}_j (\Phi'_j)^t + A \mathbb{W}_{j+1} \Phi_{j+1, j} - A \mathbb{W}_{j-1} \Phi_{j-1, j} = 2 \mathbb{S}_j \Phi_j, \quad (2.58)$$

$$2i\omega\theta_N \mathbb{W}_N \Phi_N + A \mathbb{W}_N (\Phi'_N)^t - A \mathbb{W}_{N-1} \Phi_{N-1, N} + \bar{R}_N \mathbb{W}_N \Phi_N^+ = 2 \left(\mathbb{S}_N \Phi_N + \rho_N \mathbb{W}^{\text{inc}} \Phi_N^+ \right), \quad (2.59)$$

$$\text{où } Z_1 = c_1 \theta_1 = \begin{pmatrix} z_1^{-1} & 0 \\ 0 & z_1 \end{pmatrix}.$$

Pour aboutir à un système linéaire à partir des équations (2.57), (2.58), et (2.59) nous voyons qu'il faut convertir notre inconnue matricielle $\mathbb{W} \in \mathbf{M}_{2,d_j}(\mathbb{C})$ pour tout j dans $\llbracket 1; N \rrbracket$ en un vecteur de $\mathbb{C}^{d_j \times 2}$ ou de $\mathbb{C}^{2 \times d_j}$. Nous utiliserons pour ce faire, un procédé de vectorisation, noté \aleph , qui transforme la matrice $\mathbb{W}_j = (W_{j1}, W_{j2}, \dots, W_{jd_j})$ en un vecteur colonne $\mathbf{W}_j = (W_{j1}, W_{j2}, \dots, W_{jd_j})^t \in \mathbb{C}^{2 \times d_j}$. L'application \aleph répond à la proposition 14 de l'annexe 2.4.7.1. Elle fait intervenir un produit de Kronecker \otimes , lui-même explicité dans cette proposition, et vérifie alors $\aleph(\mathbb{W}_j) = \mathbf{W}_j$. D'autre part, sachant que l'on considère une formulation où m est fixé pour toutes les cellules, nous pouvons écrire que $d_j = d$, $\forall j \in \llbracket 1; N \rrbracket$.

En appliquant maintenant \aleph à l'équation (2.58), nous obtenons par linéarité que (2.58) est successivement équivalente à :

$$2i\omega \aleph(\theta_j \mathbb{W}_j \Phi_j) + \aleph(A \mathbb{W}_j (\Phi'_j)^t) + \aleph(A \mathbb{W}_{j+1} \Phi_{j+1,j}^+) - \aleph(A \mathbb{W}_{j-1} \Phi_{j-1,j}^-) = 2\aleph(S_j \Phi_j) \Leftrightarrow$$

$$(2i\omega \Phi_j \otimes \theta_j + \Phi'_j \otimes A) \mathbf{W}_j + (\Phi_{j,j+1}^+ \otimes A) \mathbf{W}_{j+1} - (\Phi_{j,j-1}^- \otimes A) \mathbf{W}_{j-1} = 2(\Phi_j \otimes I_2) \mathbf{S}_j. \quad (2.60)$$

En procédant de même pour les équations (2.57) (2.59), on trouve que (2.57) est équivalente à :

$$\begin{aligned} (2i\omega \Phi_1 \otimes \theta_1 + \Phi'_1 \otimes A + \Phi_1^- \otimes Z_1) \mathbf{W}_1 + (\Phi_{1,2} \otimes A) \mathbf{W}_2 &= 2(\Phi_1 \otimes I_2) \mathbf{S}_j \\ &+ 2\xi(\Phi_1^- \otimes M_1^+) \theta_1 W^{\text{inc}}, \end{aligned} \quad (2.61)$$

et (2.59) est équivalente à :

$$\begin{aligned} (2i\omega(\Phi_N \otimes \theta_N) + (\Phi'_N \otimes A) + (\Phi_N^+ \otimes \bar{R}_N)) \mathbf{W}_N \\ - (\Phi_{N,N-1} \otimes A) \mathbf{W}_{N-1} = 2(\Phi_N \otimes I_2) \mathbf{S}_N + 2(\Phi_N^+ \otimes \rho_N) W^{\text{inc}} \end{aligned} \quad (2.62)$$

A première vue, il n'est pas évident que (2.60), (2.61) et (2.62) permettent d'aboutir à une matrice qui conserve la structure de la matrice issue du schéma volumes finis. Cela dit, grâce à certaines propriétés conférées par le produit de Kronecker, nous verrons que les deux matrices possèdent en effet la même structure. Ces propriétés sont énoncées dans l'annexe 2.4.7.2. Nous définissons ainsi les matrices suivantes :

$$\mathcal{D} = \begin{pmatrix} 2\omega \Phi_1 \otimes \theta_1 & 0_{2d} & \dots & 0_{2d} \\ 0_{2d} & 2\omega \Phi_2 \otimes \theta_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0_{2d} \\ 0_{2d} & \dots & 0_{2d} & 2\omega \Phi_N \otimes \theta_N \end{pmatrix}, \quad \mathcal{Z} = \begin{pmatrix} \Phi_1^- \otimes Z_1 & 0_{2d} & \dots & 0_{2d} \\ 0_{2d} & 0_{2d} & \ddots & \vdots \\ \vdots & & \ddots & 0_{2d} \\ 0_{2d} & \dots & 0_{2d} & 0_{2d} \end{pmatrix},$$

$$\mathcal{R} = \begin{pmatrix} \Phi'_1 \otimes A & \Phi_{1,2} \otimes A & 0_{2d} & \dots & 0_{2d} \\ -\Phi_{2,1} \otimes A & \Phi'_2 \otimes A & \Phi_{2,3} \otimes A & \ddots & \vdots \\ 0_{2d} & \ddots & \ddots & \ddots & 0_{2d} \\ \vdots & \ddots & -\Phi_{N-1,N-2} \otimes A & \Phi'_{N-1} \otimes A & \Phi_{N-1,N} \otimes A \\ 0_{2d} & \dots & 0_{2d} & -\Phi_{N,N-1} \otimes A & \Phi'_N \otimes A + \Phi_N^+ \otimes \bar{R}_N \end{pmatrix}.$$

La matrice \mathcal{A} résultant du schéma (2.60)-(2.61)-(2.62) peut encore s'écrire encore sous la forme $\mathcal{A} = i\mathcal{D} + \mathcal{Z} + \mathcal{R}$, comme dans le cas de la méthode de volumes finis.

Dans ce qui suit, on explicite ces matrices dans le cas particulier où l'on prend des fonctions de base \mathbb{P}_1 puis on démontre l'inversibilité du système linéaire résultant.

2.2.2.4 Formulation P_1

On se place maintenant dans le cas concret des fonctions affines par morceaux, i.e. $\mathcal{P}_i = \mathbb{P}_1[C_j]$. Dans ce cas $\phi_j = (\varphi_{j1}, \varphi_{j2})$. Pour le calcul numérique des différents termes de (2.55), on se place dans le contexte habituel des éléments finis, c'est-à-dire que l'on raisonne sur une cellule de référence \hat{C} qui est ici l'intervalle $[0, 1]$. On prend sur cet élément les fonctions de base $\hat{\varphi}_1(\hat{x}) = 1 - \hat{x}$ et $\hat{\varphi}_2(\hat{x}) = \hat{x}$. Les fonctions de base sur C_j sont alors $\varphi_{j1}(x) = \hat{\varphi}_1\left[\frac{x - x_{j-\frac{1}{2}}}{\Delta x_j}\right]$ et $\varphi_{j2}(x) = \hat{\varphi}_2\left[\frac{x - x_{j-\frac{1}{2}}}{\Delta x_j}\right]$. On a, pour $1 \leq j \leq N$:

$$\left\{ \begin{array}{l} \Phi_j = \int_{C_j} \phi_j^t \phi_j dx = \Delta x_j \int_{\hat{C}} \hat{\phi}^t \hat{\phi} d\hat{x} = \frac{\Delta x_j}{6} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \\ \Phi'_j = \int_{C_j} (\phi_j^t \partial_x \phi_j - (\partial_x \phi_j)^t \phi_j) dx = \Delta x_j \int_{\hat{C}} \frac{1}{\Delta x_j} (\hat{\phi}^t \partial_x \hat{\phi} - (\partial_x \hat{\phi})^t \hat{\phi}) d\hat{x} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \\ \Phi_j^+ = (\phi_j^+)^t \phi_j^+ = \phi_j^t(x_{j+\frac{1}{2}}) \phi_j(x_{j+\frac{1}{2}}) = \hat{\phi}^t(1) \hat{\phi}(1) = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \\ \Phi_j^- = (\phi_j^-)^t \phi_j^- = \phi_j^t(x_{j-\frac{1}{2}}) \phi_j(x_{j-\frac{1}{2}}) = \hat{\phi}^t(0) \hat{\phi}(0) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \end{array} \right.$$

D'autre part, on trouve aussi que :

$$\begin{aligned} \forall j \in [1; N-1], \quad \Phi_{j,j+1} &= (\phi_j^+)^t \phi_{j+1}^- = \phi_j^t(x_{j+\frac{1}{2}}) \phi_{j+1}(x_{j+\frac{1}{2}}) = \hat{\phi}^t(1) \hat{\phi}(0) = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \\ \forall j \in [2; N], \quad \Phi_{j,j-1} &= (\phi_j^-)^t \phi_{j-1}^+ = \phi_j^t(x_{j-\frac{1}{2}}) \phi_{j-1}(x_{j-\frac{1}{2}}) = \hat{\phi}^t(0) \hat{\phi}(1) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}. \end{aligned}$$

Enfin, si on multiplie (2.60) par 3 pour éviter les termes fractionnaires, et si on formule un système linéaire $\mathcal{A}\mathbf{W} = \mathbf{b}$ avec $(\mathcal{A}, \mathbf{b}) \in \mathbf{M}_{4 \times N}(\mathbb{C}) \times \mathbb{C}^{4 \times N}$, alors la matrice \mathcal{A} est définie par blocs par :

$$\begin{aligned}
\bullet \mathcal{A}_{1,1} &= \left(\begin{array}{cc|cc} 2i\omega\Delta x_1\varepsilon_1 + 3z_1^{-1} & 0 & i\omega\Delta x_1\varepsilon_1 & -3 \\ 0 & 2i\omega\Delta x_1\mu_1 + 3z_1 & -3 & i\omega\Delta x_1\mu_1 \\ \hline i\omega\Delta x_1\varepsilon_1 & 3 & 2i\omega\Delta x_1\varepsilon_1 & 0 \\ 3 & i\omega\Delta x_1\mu_1 & 0 & 2i\omega\Delta x_1\mu_1 \end{array} \right), \\
\bullet \mathcal{A}_{1,2} &= \left(\begin{array}{cc|c} 0_2 & & 0_2 \\ \hline 0 & -3 & 0_2 \\ -3 & 0 & 0_2 \end{array} \right), \\
\bullet \mathcal{A}_{j,j-1} &= \left(\begin{array}{cc|cc} 0_2 & 0 & 3 & \\ \hline 0_2 & 3 & 0 & \\ 0_2 & & 0_2 & \end{array} \right), \quad \forall j \in \llbracket 2; N-1 \rrbracket, \\
\bullet \mathcal{A}_{jj} &= \left(\begin{array}{cc|cc} 2i\omega\Delta x_j\varepsilon_j & 0 & i\omega\Delta x_j\varepsilon_j & -3 \\ 0 & 2i\omega\Delta x_j\mu_j & -3 & i\omega\Delta x_j\mu_j \\ \hline i\omega\Delta x_j\varepsilon_j & 3 & 2i\omega\Delta x_j\varepsilon_j & 0 \\ 3 & i\omega\Delta x_j\mu_j & 0 & 2i\omega\Delta x_j\mu_j \end{array} \right), \quad \forall j \in \llbracket 2; N-1 \rrbracket,
\end{aligned}$$

$$\begin{aligned}
\bullet \mathcal{A}_{j,j+1} &= \left(\begin{array}{cc|c} 0_2 & & 0_2 \\ \hline 0 & -3 & 0_2 \\ -3 & 0 & 0_2 \end{array} \right), \quad \forall j \in \llbracket 2; N-1 \rrbracket, \\
\bullet \mathcal{A}_{N,N-1} &= \left(\begin{array}{cc|cc} 0_2 & 0 & 3 & \\ \hline 0_2 & 3 & 0 & \\ 0_2 & & 0_2 & \end{array} \right), \\
\bullet \mathcal{A}_{N,N} &= \left(\begin{array}{cc|cc} 2i\omega\Delta x_N\varepsilon_N & 0 & i\omega\Delta x_N\varepsilon_N & -3 \\ 0 & 2i\omega\Delta x_N\mu_N & -3 & i\omega\Delta x_N\mu_N \\ \hline i\omega\Delta x_N\varepsilon_N & 3 & 2i\omega\Delta x_N\varepsilon_N & -3 \\ 3 & i\omega\Delta x_N\mu_N & 3 & 2i\omega\Delta x_N\mu_N \end{array} \right).
\end{aligned}$$

Ces matrices ont les mêmes caractéristiques que celles du schéma VF, à savoir :

- \mathcal{D} est diagonale par blocs et symétrique définie positive,
- \mathcal{Z} est diagonale positive,
- \mathcal{R} est antisymétrique.

Il est montré dans la proposition 15 de l'annexe 2.4.7.4 qu'une matrice résultant de l'assemblage de fonctions \mathbb{P}_m possède la même structure. Pour autant, l'inversibilité ne sera pas forcément assurée. Néanmoins, nous allons voir dans ce qui suit que pour notre formulation d'ordre 1, l'inversibilité de la matrice est vérifiée.

2.2.2.5 Étude du problème discret

Nous constatons malheureusement que $\forall j \in \llbracket 1; N-1 \rrbracket$ la matrice $\mathcal{A}_{j(j+1)}$ n'est pas inversible. Ainsi, il n'est pas possible de reproduire exactement la démonstration de l'inversibilité de la matrice associée à la méthode de volumes finis. En revanche, nous allons faire une démonstration similaire dont l'idée sera

de perpétuer une relation de récurrence non plus par cellule mais par degré de liberté. Nous considérons que $\mathbf{W} = (W_1^t, W_2^t, \dots, W_N^t)^t$ où pour $j \in \llbracket 1; N \rrbracket$ $W_j^t = (W_j^1, W_j^2)^t$ avec $W_j^{1,2} \in \mathbb{C}^2$.

Lemme 4 La matrice \mathcal{A} définie par $\mathcal{A} = i\mathcal{D} + \mathcal{Z} + \mathcal{R}$, caractérisant le schéma Galerkin Discontinu 1D en formulation \mathbb{P}_1 , est inversible.

Preuve Montrons par récurrence sur j pour $j \in \llbracket 1; N-1 \rrbracket$ que si $\mathcal{A}\mathbf{W} = 0$ pour $\mathbf{W} \in \mathbb{C}^{4 \times N}$ alors $W_j = W_{j+1} = 0$.

1. Pour le premier rang de la récurrence, nous avons :

$$\mathcal{A}\mathbf{W} = 0 \Rightarrow \mathbf{W}^* \mathcal{A}\mathbf{W} = 0 \Rightarrow \mathbf{W}^*(i\mathcal{D} + \mathcal{Z} + \mathcal{R})\mathbf{W} = 0 \Rightarrow \mathbf{W}^*(i\mathcal{D} + \mathcal{R})\mathbf{W} + \mathbf{W}^* \mathcal{Z}\mathbf{W} = 0.$$

La matrice \mathcal{Z} est hermitienne donc $\mathbf{W}^* \mathcal{Z}\mathbf{W} \in \mathbb{R}$. D'autre part, reprenant les arguments de la démonstration réalisée pour la méthode de volumes finis, $(i\mathcal{D} + \mathcal{R})$ est antihermitienne et donc $\mathbf{W}^*(i\mathcal{D} + \mathcal{R})\mathbf{W} \in i\mathbb{R}$. Ceci entraîne donc que $\Re(\mathbf{W}^* \mathcal{A}\mathbf{W}) = \mathbf{W}^* \mathcal{Z}\mathbf{W}$ et $\Im(\mathbf{W}^* \mathcal{A}\mathbf{W}) = \mathbf{W}^*(i\mathcal{D} + \mathcal{R})\mathbf{W}$.

En conséquence, $\mathbf{W}^* \mathcal{A}\mathbf{W} = 0 \Rightarrow \mathbf{W}^* \mathcal{Z}\mathbf{W} = \mathbf{W}^*(i\mathcal{D} + \mathcal{R})\mathbf{W} = 0$. Si on s'intéresse plus particulièrement à $\mathbf{W}^* \mathcal{Z}\mathbf{W}$, nous obtenons en développant que $\mathbf{W}^* \mathcal{Z}\mathbf{W} = W_1^* \mathcal{Z}_{11} W_1 = W_1^*(\Phi_1^- \otimes Z_1) W_1 = 0$. De là on tire $W_1^1 = 0$.

Maintenant, nous avons aussi :

$$\begin{aligned} \mathcal{A}\mathbf{W} = 0 &\Rightarrow \forall j \in \llbracket 1; N \rrbracket, (\mathcal{A}\mathbf{W})_j = 0 \quad (\text{et en particulier pour } j = 1) \\ &\Rightarrow A_{11}W_1 + A_{12}W_2 = 0 \\ &\Rightarrow A_{11}(0, (W_1^1)^t)^t + A_{12}W_2 = 0 \\ &\Rightarrow \begin{cases} \begin{pmatrix} i\omega\Delta x_1 \varepsilon_1 & -3 \\ -3 & i\omega\Delta x_1 \mu_1 \end{pmatrix} W_1^2 = 0 \\ \begin{pmatrix} i\omega\Delta x_1 \varepsilon_1 & 0 \\ 0 & i\omega\Delta x_1 \mu_1 \end{pmatrix} W_1^2 + \begin{pmatrix} 0 & -3 \\ -3 & 0 \end{pmatrix} W_2^1 = 0. \end{cases} \end{aligned}$$

Notons $C = \begin{pmatrix} i\omega\Delta x_1 \varepsilon_1 & -3 \\ -3 & i\omega\Delta x_1 \mu_1 \end{pmatrix}$. Puisque $\det C = -(\frac{\omega^2 \Delta x_1^2}{c_1^2} + 9) < 0$, alors C est inversible si bien que $W_1^2 = 0$. Combinant le fait que $W_1^2 = 0$ et l'autre égalité, alors il est évident que $W_2^1 = 0$.

Nous avons donc montré que $\mathcal{A}\mathbf{W} = 0 \Rightarrow W_1 = 0$ et $W_2^1 = 0$.

2. Supposons maintenant que la relation est vérifiée au rang j , c'est-à-dire que $W_{j-1} = 0$ pour $j \in \llbracket 2; N-1 \rrbracket$ et que $W_j^1 = 0$. Nous devons montrer que $W_j = 0$ et que $W_{j+1}^1 = 0$. Mais vu que $W_j^1 = 0$, alors il faut juste prouver que $W_j^2 = 0$ et $W_{j+1}^1 = 0$ sachant que :

$$\begin{aligned}
\mathcal{A}W = 0 &\Rightarrow \forall j \in \llbracket 1; N \rrbracket, (\mathcal{A}W)_j = 0 \quad (\text{et en particulier pour } j \in \llbracket 2; N-1 \rrbracket) \\
&\Rightarrow \mathcal{A}_{j(j-1)}W_{j-1} + \mathcal{A}_{jj}W_j + \mathcal{A}_{j(j+1)}W_{j+1} = 0 \\
&\Rightarrow \mathcal{A}_{jj}(0, (W_j^2)^t)^t + \mathcal{A}_{j(j+1)}W_{j+1} = 0 \\
&\Rightarrow \begin{cases} \begin{pmatrix} 2i\omega\Delta x_j \varepsilon_j & -3 \\ -3 & 2i\omega\Delta x_j \mu_j \end{pmatrix} W_j^2 = 0 \\ \begin{pmatrix} i\omega\Delta x_j \varepsilon_j & 0 \\ 0 & i\omega\Delta x_j \mu_j \end{pmatrix} W_j^2 + \begin{pmatrix} 0 & -3 \\ -3 & 0 \end{pmatrix} W_{j+1}^1 = 0. \end{cases}
\end{aligned}$$

Avec ce système d'égalités, on reprend à la lettre le cheminement du 1. pour trouver au final que $W_j = 0$ et que $W_{j+1}^1 = 0$.

3. Il ne reste plus qu'à montrer que $W_N^2 = 0$ sachant que pour $j \in \llbracket 1; N-1 \rrbracket$ $W_j = 0$ et que $W_N^1 = 0$. En développant la dernière ligne du système et en tenant compte de ces affirmations, on trouve :

$$\mathcal{A}W = 0 \Rightarrow \begin{cases} \begin{pmatrix} i\omega\Delta x_N \varepsilon_N & -3 \\ -3 & i\omega\Delta x_N \mu_N \end{pmatrix} W_N^2 = 0, \\ \begin{pmatrix} 2i\omega\Delta x_N \varepsilon_N & -3 \\ 3 & 2i\omega\Delta x_N \mu_N \end{pmatrix} W_N^2 = 0. \end{cases} \quad (2.63)$$

On aboutit comme précédemment à $W_N^2 = 0$. ■

2.2.3 Résultats numériques

On considère le domaine $[a, b] = [0, 1]$. Conformément aux conditions que nous avons choisies dans les paragraphes précédents, nous optons pour une condition absorbante aux points $\{a\}$ et $\{b\}$. Dans le cas test considéré, une onde plane se dirige du point $\{a\}$ vers le point $\{b\}$, et on observe la réponse obtenue pour différentes caractéristiques du maillage et du milieu. L'onde plane est de la forme $W^{\text{inc}}(x) = (\eta_e, \eta_h)e^{-ikx}$ où $k = \omega/c_1$ et $\eta_h = z_1\eta_e^2$, de sorte qu'elle satisfait bien les équations de Maxwell 1D. Le maillage est toujours construit de façon à ce que l'on ait un minimum de 10 points par longueur d'onde. Pour chaque test, on compare les solutions fournies par les versions GD-P0 et GD-P1. On choisit de visualiser la partie réelle du champ électrique.

2.2.3.1 Maillage uniforme, milieu homogène

Dans ce test, on pose pour $\forall j : \varepsilon_j = 1$ et $\mu_j = 1$. Les conditions en $\{a\}$ et $\{b\}$ permettent de simuler la propagation d'une onde plane dans le vide et donc de pouvoir comparer la solution obtenue à la solution de référence. Pour l'onde plane choisie, cette solution est $\Re(e^{-ikx}) = \cos(kx)$. On utilise un maillage uniforme de $N = 200$ cellules pour les fréquences $f = 100$ MHz, $f = 500$ MHz et $f = 1$ GHz. Rappelons au passage les relations de base :

² c_1 et z_1 sont les valeurs de c et z dans la cellule C_1 .

$$\begin{cases} k &= \frac{\omega}{c}, \\ \omega &= 2\pi f, \\ \lambda &= \frac{c}{f}. \end{cases}$$

dont on tire $\lambda = 2\pi/k$. Les solutions obtenues sont montrées sur la figure 2.2. En comparant ces graphiques, on constate que la solution P1 est plus précise que la solution P0, en particulier pour les hautes fréquences.

On souhaite maintenant évaluer l'impact de la dispersion numérique, toujours en maillage uniforme et en milieu homogène. Ces conditions permettent de considérer la relation de dispersion trouvée pour la méthode GD-P1, qui s'exprime à partir de la relation (3.107) qui conduit à :

$$|\omega^2 - \omega_h^2| \sim \omega_h^2 \frac{k_x^2 \Delta x^2}{24},$$

qui montre que plus le nombre de points par longueur d'onde ($\frac{\lambda}{\Delta x} = \frac{2\pi}{k_x \Delta x}$) est petit, plus l'erreur de dispersion sera importante. Pour garantir une précision acceptable, il faut donc s'assurer que la taille de maille Δx corresponde à un certain nombre de points par longueur d'onde. Dans le cas présent, nous avons fixé ce nombre à 10. Pour noter cette différence, on peut par exemple fixer une taille de maille et augmenter la fréquence le plus possible tout en satisfaisant le critère de 10 points par longueur d'onde. Ainsi, en prenant une fréquence $f = 1$ GHz et un maillage de $N = 50$ cellules, on obtient la courbe de la figure 2.3.

La solution résultant de la méthode GD-P1 permet de noter clairement l'erreur de dispersion pour ce jeu de données, alors que la solution associée à la méthode GD-P0 est trop altérée pour s'en rendre compte. Ici, cette erreur est caractérisée comme prévu par un déphasage entre les deux courbes. Si on souhaite comparer les deux méthodes, il faut considérer le cas d'un maillage de $N = 500$ cellules et faire un zoom (cf. figure 2.4). On constate une erreur de dispersion pour les deux méthodes bien que la différence ne soit pas vraiment significative. Pourtant, nous avons zoomé sur une zone où la solution associée à la méthode GD-P0 était la plus proche de la solution exacte. Si on se réfère aux formules qui expriment la relation de dispersion pour la méthode GD-P0 et la relation de dispersion pour la méthode GD-P1, nous constatons que cette dernière est moins dispersive d'un facteur 8. Mais pour bien se rendre compte de cette erreur, il faudrait comparer les courbes en prenant un maillage très grossier, chose qui est irréalisable compte tenu de la mauvaise qualité de la solution obtenue avec la méthode GD-P0.

2.2.3.2 Maillage non uniforme, milieu homogène

Dans cette section, le but est de voir comment se comporte la méthode GD-P1 en présence d'un maillage non uniforme. Les conditions aux limites sont identiques à celles prises dans la section 2.2.3.1. Le maillage comporte 200 cellules mais le pas d'espace est défini de façon aléatoire en vérifiant toujours le critère $\max_{1 \leq j \leq N} (\Delta x_j) \leq \lambda/10$. Les cellules sont de trois tailles différentes bien que la longueur de la plus grande maille reste égale à trois fois celle de la plus petite. Des tests sont réalisés pour des fréquences de 100 MHz, 500 MHz et 1 GHz. Les solutions obtenues sont montrées sur la figure 2.5. Là où la méthode GD-P0 pêche par manque de précision, la solution résultant de la méthode GD-P1 est correcte même sur un maillage aléatoire grossier. On peut donc prendre un maillage aléatoire de $N = 200$ cellules et observer que la solution se calque sur la solution de référence.

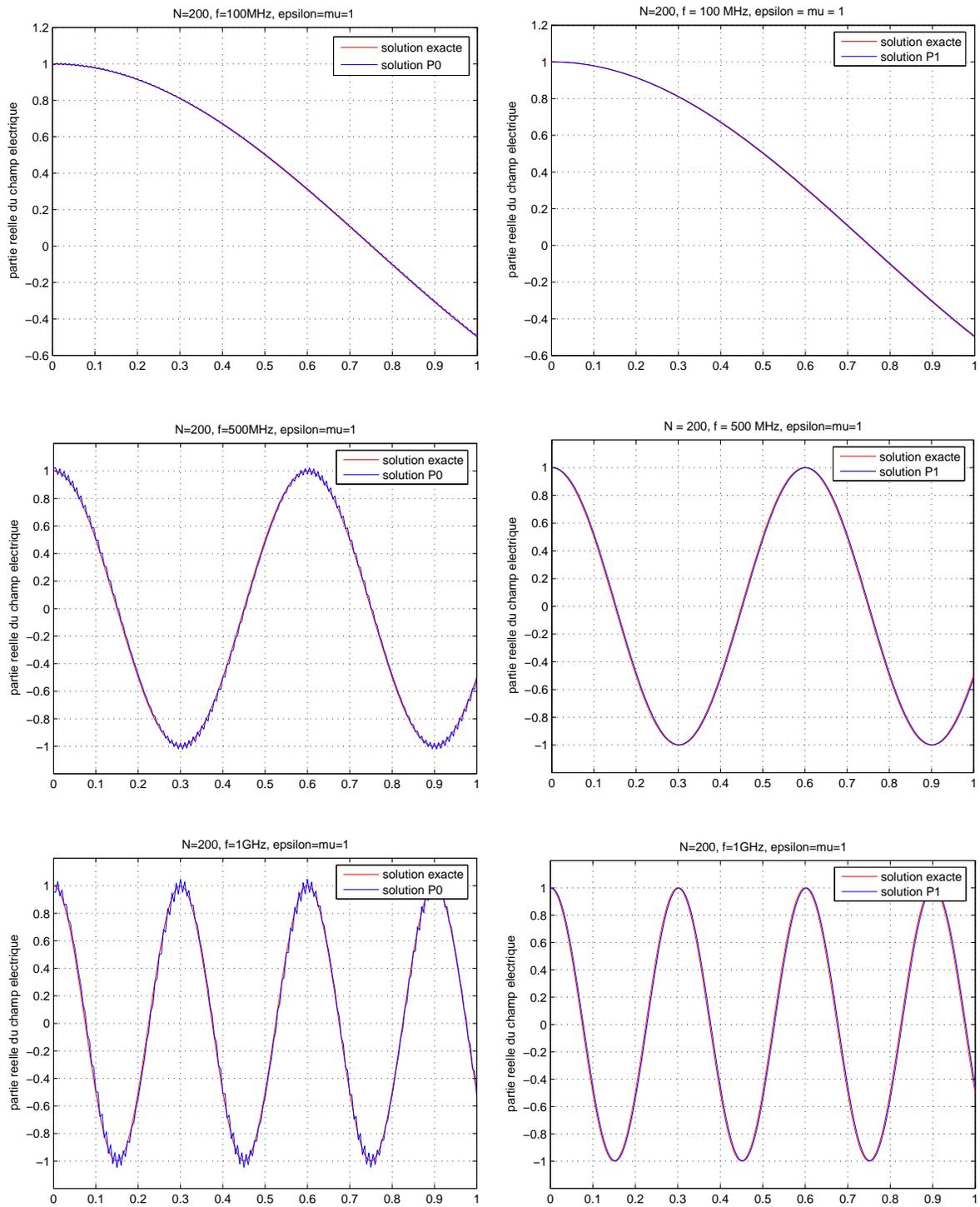


FIG. 2.2 – Equations de Maxwell 1D : milieu homogène et maillage uniforme

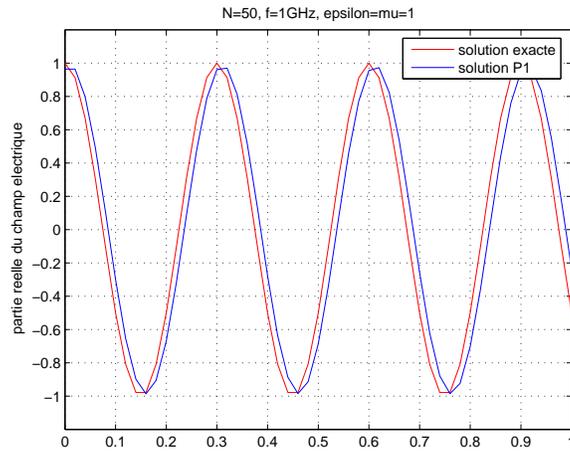


FIG. 2.3 – Equations de Maxwell 1D : milieu homogène et maillage uniforme
 Erreur de dispersion pour un maillage de $N = 50$ cellules et $f = 1$ GHz

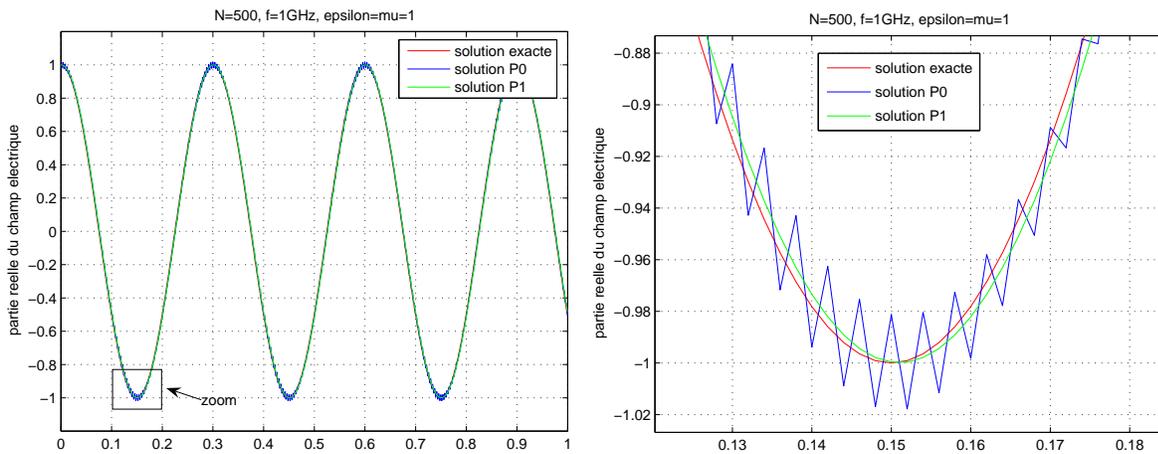


FIG. 2.4 – Equations de Maxwell 1D : milieu homogène et maillage uniforme
 Erreur de dispersion pour un maillage de $N = 500$ cellules et $f = 1$ GHz

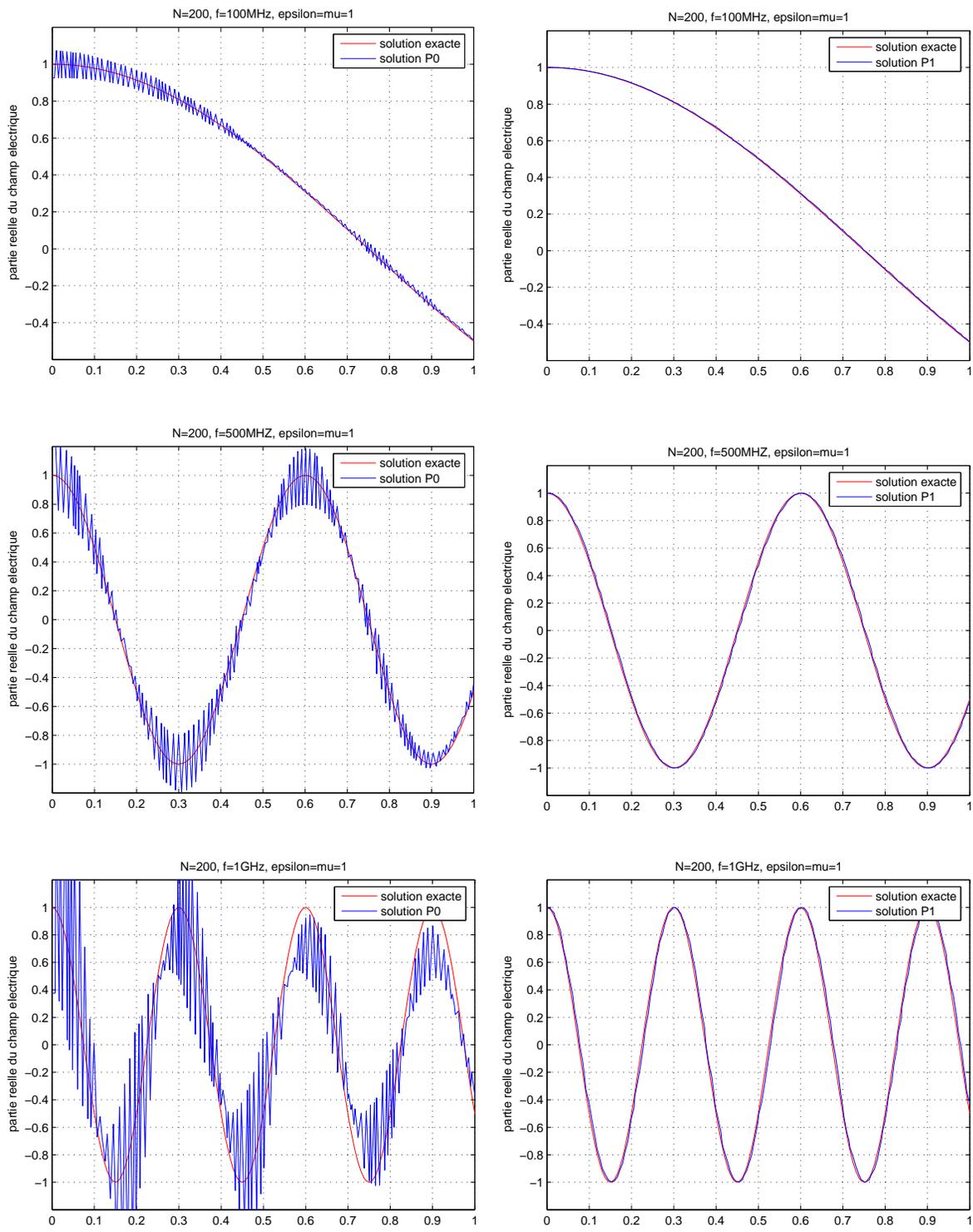


FIG. 2.5 – Equations de Maxwell 1D : milieu homogène et maillage non uniforme

Ceci dit, un test intéressant consiste à comparer les solutions P0 et P1 obtenues en considérant un maillage aléatoire fin pour la méthode GD-P0, et un maillage aléatoire grossier pour la méthode GD-P1. Le but est bien entendu d'évaluer la qualité de la solution obtenue, et de voir s'il est plus avantageux de raffiner le maillage ou d'augmenter l'ordre d'interpolation (cf. figure 2.6). Il est assez clair qu'il est beaucoup plus avantageux d'augmenter l'ordre d'interpolation de la méthode GD plutôt que de raffiner le maillage lorsqu'on recherche une solution précise. En dépit d'un maillage 6 fois plus fin dans les tests utilisant la méthode P0, la solution reste non seulement moins précise, mais elle est également plus dispersive. Néanmoins, ces remarques ne se soucient point de l'aspect résolution : l'opérateur matriciel issu de la méthode GD-P1 est peut être moins bien conditionné que celui issu de la méthode GD-P0. Cet aspect n'est pas abordé ici.

2.2.3.3 Maillage uniforme, milieu hétérogène

Dans cette section, toutes les conditions des essais numériques des sections 2.2.3.1 et 2.2.3.2 sont reproduites, sauf deux. D'une part, sur la section $[0.4, 0.6]$ du domaine de calcul, on pose $\varepsilon_j = 4$ et $\mu_j = 1$. Pour le reste du domaine, nous conserverons $\varepsilon = \mu = 1$. D'autre part, nous avons fixé un nombre de cellules à $N = 400$ de sorte que la solution fournie par la méthode GD-P0 ne subisse pas les inconvénients d'un maillage grossier. Ces test sont réalisés pour des fréquences de 100 MHz, 500 MHz et 1 GHz. Les solutions obtenues sont montrées sur la figure 2.7. On observe que les solutions associées aux méthodes GD-P0 et GD-P1 ne coïncident absolument pas. Cette différence est d'autant plus notable que la fréquence est élevée.

2.2.3.4 Maillage non uniforme, milieu hétérogène

Les conditions de simulation sont strictement les mêmes que pour la section 2.2.3.3 hormis le fait que le maillage est non uniforme. Le nombre de cellules est $N = 400$. Les solutions obtenues sont montrées sur la figure 2.8. Nous constatons que les solutions associées aux méthodes GD-P0 et GD-P1 ont un comportement et un profil identique à leur homologue en maillage uniforme. On retrouve les oscillations de la solution P0 en milieu homogène, mais la solution P1 est quasi-identique à la solution P1 en maillage uniforme. Ces résultats confirment que la méthode GD-P1 est bien moins sensible aux irrégularités de maillage et permet visiblement de mieux prendre en compte l'hétérogénéité. On remarque notamment que pour une fréquence $f = 100$ MHz, le profil de la solution P1 change lorsqu'elle se situe dans la zone hétérogène $[0.4, 0.6]$, alors que la solution P0 ne semble pas prendre en compte cette transition.

2.3 Équations de Maxwell 3D

Pour des soucis d'implémentation, nous souhaitons traiter de front la résolution du système en champ diffracté et en champ total en introduisant le paramètre ξ qui permet de distinguer un calcul en champ diffracté ($\xi = 0$) d'un calcul en champ total ($\xi = 1$). Rappelons que :

$$\begin{cases} \mathbf{W} = \mathbf{W}^{\text{dif}} + \xi \mathbf{W}^{\text{inc}}, \\ \mathbf{W} \equiv \mathbf{W}^{\text{dif}} & \text{si } \xi = 0, \\ \mathbf{W} \equiv \mathbf{W}^{\text{tot}} & \text{si } \xi = 1. \end{cases}$$

Si on note :

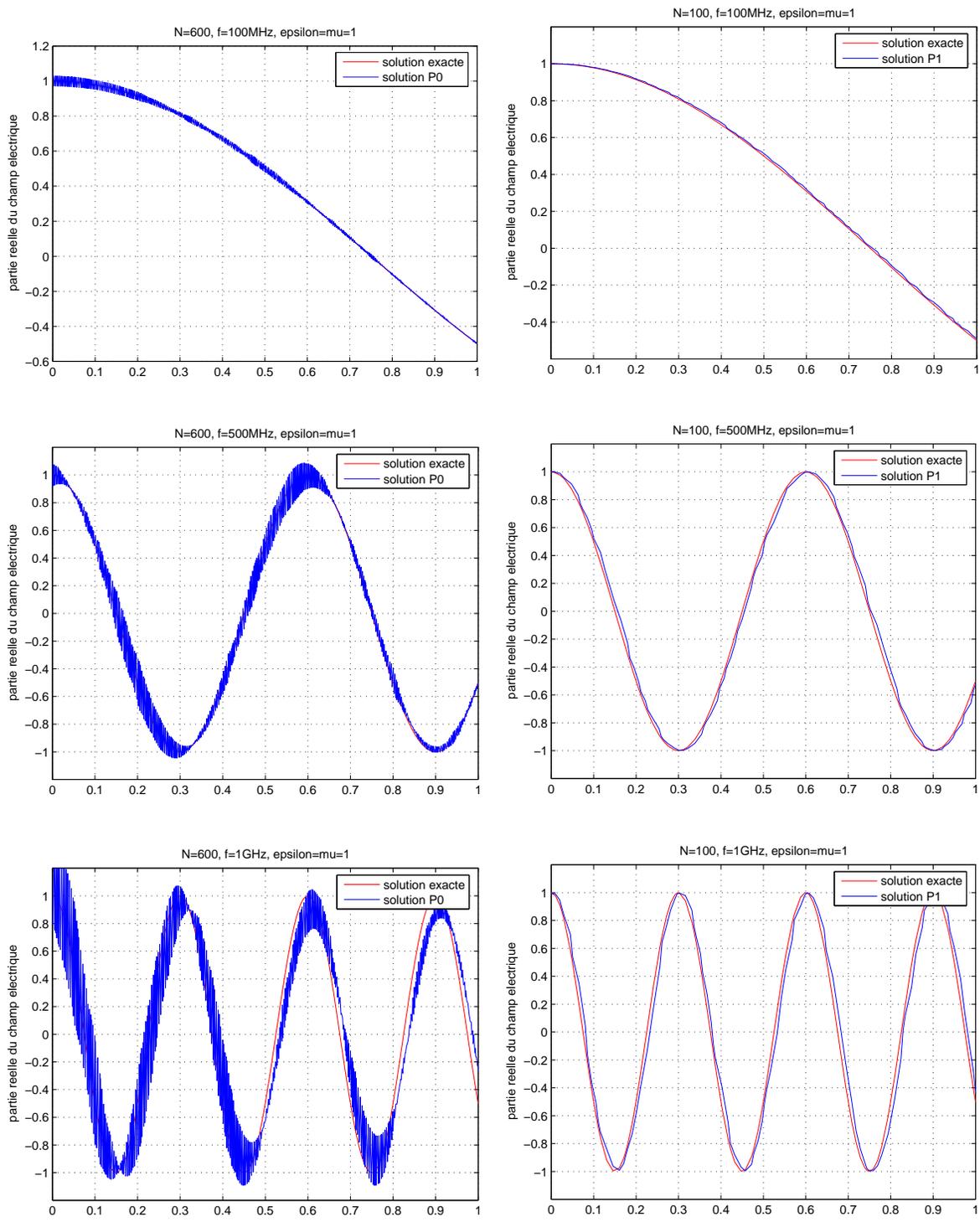


FIG. 2.6 – Equations de Maxwell 1D : milieu homogène et maillage non uniforme

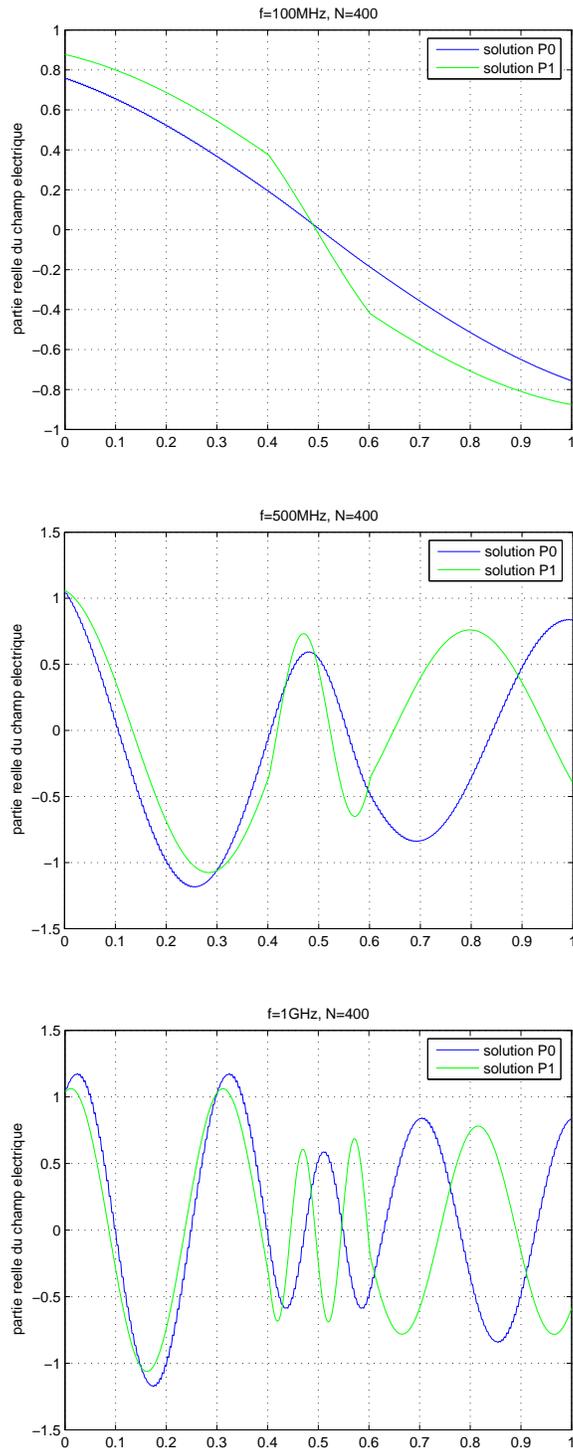


FIG. 2.7 – Equations de Maxwell 1D : milieu hétérogène et maillage uniforme

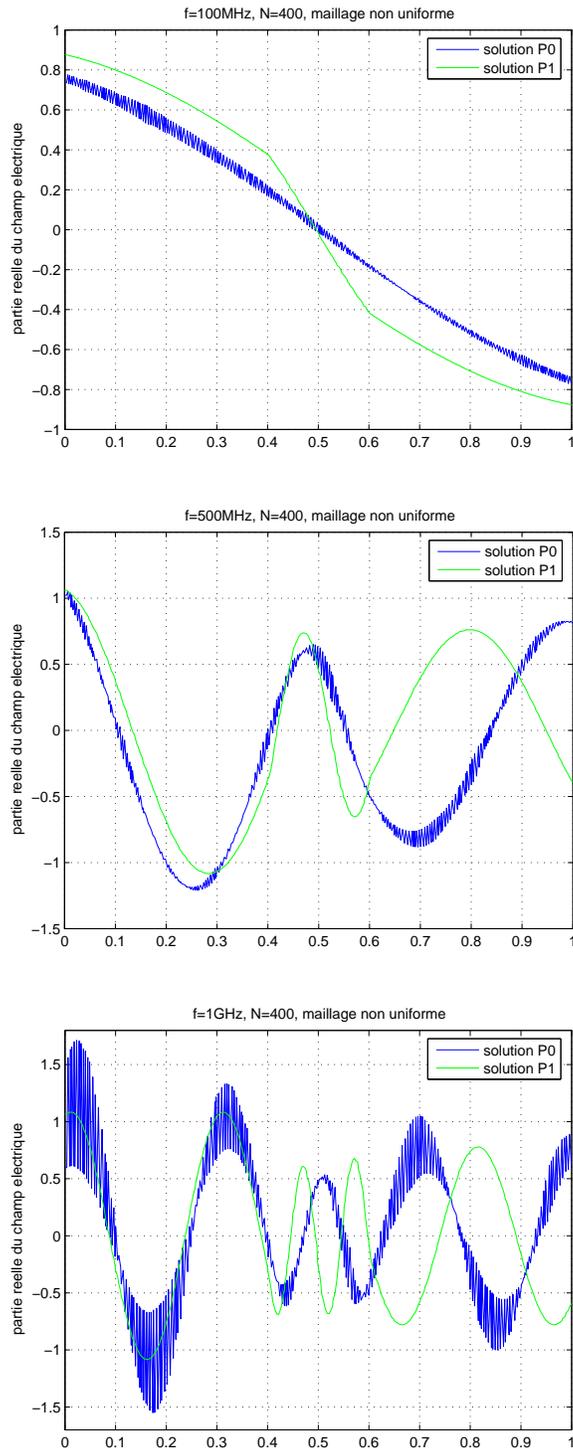


FIG. 2.8 – Equations de Maxwell 1D : milieu hétérogène et maillage non uniforme

$$\begin{cases} \varepsilon(\mathbf{x}) &= \varepsilon_0 \varepsilon_r(\mathbf{x}), \\ \mu(\mathbf{x}) &= \mu_0 \mu_r(\mathbf{x}), \end{cases}$$

où ε_0 et μ_0 désignent respectivement la permittivité et la perméabilité du vide, alors on peut adimensionner le système (2.16) grâce au changement de variables suivant :

$$\begin{cases} \tilde{\mathbf{E}} &= \frac{\mathbf{E}}{z_0 \mathbf{H}_0}, \\ \tilde{\mathbf{H}} &= \frac{\mathbf{H}}{\mathbf{H}_0}, \\ \tilde{\omega} &= \frac{\omega}{c_0}, \end{cases}$$

où z_0 est l'impédance du vide (d'une manière générale on notera par z l'impédance d'un milieu quelconque $z = \sqrt{\mu/\varepsilon}$). Il faut aussi modifier les conditions aux bords :

– sur la *frontière métallique* Γ^m , on a :

$$\begin{aligned} \mathbf{n} \times \mathbf{E}^{\text{tot}} = 0 &\Leftrightarrow \mathbf{n} \times (\mathbf{E} + (1 - \xi)\mathbf{E}^{\text{inc}}) = 0 \\ &\Leftrightarrow \mathbf{n} \times \mathbf{E} = (\xi - 1)\mathbf{n} \times \mathbf{E}^{\text{inc}} \\ &\Leftrightarrow \mathbf{n} \times (z_0 \mathbf{H}_0 \tilde{\mathbf{E}}) = (\xi - 1)\mathbf{n} \times (z_0 \mathbf{H}_0 \tilde{\mathbf{E}}^{\text{inc}}). \end{aligned}$$

– sur la *frontière absorbante* Γ^a , on a :

$$\begin{aligned} \mathbf{n} \times \mathbf{E}^{\text{dif}} + z\mathbf{n} \times (\mathbf{n} \times \mathbf{H}^{\text{dif}}) &= 0 \\ \Leftrightarrow \mathbf{n} \times (\mathbf{E} - \xi\mathbf{E}^{\text{inc}}) + z\mathbf{n} \times (\mathbf{n} \times (\mathbf{H} - \xi\mathbf{H}^{\text{inc}})) &= 0 \\ \Leftrightarrow \mathbf{n} \times \mathbf{E} + z\mathbf{n} \times (\mathbf{n} \times \mathbf{H}) &= \xi(\mathbf{n} \times \mathbf{E}^{\text{inc}} + z\mathbf{n} \times (\mathbf{n} \times \mathbf{H}^{\text{inc}})) \\ \Leftrightarrow \mathbf{n} \times \tilde{\mathbf{E}} + z_r \mathbf{n} \times (\mathbf{n} \times \tilde{\mathbf{H}}) &= \xi(\mathbf{n} \times \tilde{\mathbf{E}}^{\text{inc}} + z_r \mathbf{n} \times (\mathbf{n} \times \tilde{\mathbf{H}}^{\text{inc}})). \end{aligned}$$

Pour alléger les notations, on abandonne les symboles $\tilde{\cdot}$ et on fait les substitutions $\varepsilon_r \rightarrow \varepsilon$, $\mu_r \rightarrow \mu$, $c_r \rightarrow c$ et $z_r \rightarrow z$, ce qui donne finalement après simplification :

$$\begin{cases} \mathbf{n} \times \mathbf{E} = (\xi - 1)\mathbf{n} \times \mathbf{E}^{\text{inc}} & \text{sur } \Gamma^m, \\ \mathbf{n} \times \mathbf{E} + z\mathbf{n} \times (\mathbf{n} \times \mathbf{H}) = \xi(\mathbf{n} \times \mathbf{E}^{\text{inc}} + z\mathbf{n} \times (\mathbf{n} \times \mathbf{H}^{\text{inc}})) & \text{sur } \Gamma^a. \end{cases} \quad (3.64)$$

Par ailleurs, si pour $l \in \{x, y, z\}$ on note $F_l(\mathbf{W}) = G_l \mathbf{W}$ avec :

$$G_l = \begin{pmatrix} 0_3 & N_l \\ -N_l & 0_3 \end{pmatrix},$$

alors le système des équations de Maxwell harmoniques s'écrit :

$$i\omega\theta\mathbf{W} + \text{div}F(\mathbf{W}) + K\mathbf{W} = \mathbf{S}, \quad (3.65)$$

ou bien encore :

$$i\omega\theta\mathbf{W} + G_x \partial_x \mathbf{W} + G_y \partial_y \mathbf{W} + G_z \partial_z \mathbf{W} + K\mathbf{W} = \mathbf{S}. \quad (3.66)$$

où :

$$\mathbf{W} = \begin{pmatrix} \mathbf{E} \\ \mathbf{H} \end{pmatrix} \quad \text{et} \quad \theta = \begin{pmatrix} \varepsilon_{l_3} & 0_3 \\ 0_3 & \mu_{l_3} \end{pmatrix}$$

On supposera par la suite que $\sigma = 0$ ce qui conduit à $K = 0$. Rappelons aussi que si on résoud les équations en champ total, le second membre \mathbf{S} est nul (il est égal à $(I_6 - G_0)p\mathbf{W}^{\text{inc}}$ pour un calcul en champ diffracté). Pour la prise en compte des conditions aux limites, on procède comme en 1D c'est-à-dire que l'on associe à Γ_h une partition $\mathcal{I}_b \subset \mathbb{N}$ de cellules fictives. On suppose que les paramètres ε et μ sont constants par cellule : $\varepsilon|_{C_j} = \varepsilon_j$ et $\mu|_{C_j} = \mu_j$.

2.3.1 Méthode de volumes finis

Dans cette section, on décrit en détail la méthode de volumes finis introduite dans la section 2.1.1. Pour une fonction vectorielle \mathbf{W} , nous noterons \mathbf{W}_j sa valeur moyenne sur C_j , c'est-à-dire :

$$\mathbf{W}_j = \frac{1}{V_j} \int_{C_j} \mathbf{W} dx.$$

En intégrant l'équation (3.65) sur une cellule C_j , on obtient :

$$i\omega V_j \theta_j \mathbf{W}_j + \int_{C_j} \text{div} F(\mathbf{W}) dx = V_j \mathbf{S}_j.$$

Grâce au théorème de Stokes :

$$\begin{aligned} i\omega V_j \theta_j \mathbf{W}_j + \int_{C_j} F(\mathbf{W}) \cdot \mathbf{n} ds &= V_j \mathbf{S}_j \\ \Leftrightarrow i\omega V_j \theta_j \mathbf{W}_j + \sum_{k \in \mathcal{V}_j} \int_{\Sigma_{jk}} F(\mathbf{W}) \cdot \mathbf{n}_{jk} ds &= V_j \mathbf{S}_j, \end{aligned}$$

qui est équivalent au système :

$$\begin{cases} i\omega V_j \varepsilon_j \mathbf{E}_j + \sum_{k \in \mathcal{V}_j} \int_{\Sigma_{ij}} N_{jk} \mathbf{H} ds &= -V_j \mathbf{S}_j, \\ i\omega V_j \mu_j \mathbf{H}_j - \sum_{k \in \mathcal{V}_j} \int_{\Sigma_{jk}} N_{jk} \mathbf{E} ds &= 0. \end{cases} \quad (3.67)$$

où $N_{jk} = \sum_{l=1}^3 n_{jk}^{x_l} N_l$ est basé sur des normales unitaires. On introduit les flux numériques suivants :

$$\begin{cases} \Phi_{H,jk} \equiv \Phi_H(\mathbf{H}_j, \mathbf{H}_k, \mathbf{n}_{jk}) \approx \int_{\Sigma_{jk}} N_{jk} \mathbf{H} ds, \\ \Phi_{E,jk} \equiv \Phi_E(\mathbf{E}_j, \mathbf{E}_k, \mathbf{n}_{jk}) \approx \int_{\Sigma_{jk}} N_{jk} \mathbf{E} ds, \end{cases} \quad (3.68)$$

et $\Phi_{jk} = (\Phi_{H,jk}, -\Phi_{E,jk})^t$. Ainsi, (3.67) s'écrit aussi :

$$i\omega V_j \theta_j \mathbf{W}_j + \sum_{k \in \mathcal{V}_j} \Phi_{jk} = V_j \mathbf{S}_j. \quad (3.69)$$

Les flux numériques (3.68) sont calculés de manière centrée [Piperno et al., 2002] :

$$\begin{cases} \Phi_{H,jk} = \frac{S_{jk}}{2} [N_{jk}\mathbf{H}_j + N_{jk}\mathbf{H}_k], \\ \Phi_{E,jk} = \frac{S_{jk}}{2} [N_{jk}\mathbf{E}_j + N_{jk}\mathbf{E}_k]. \end{cases} \quad (3.70)$$

On introduit $M_{jk}^j = M_{jk}\theta_j^{-1}$ avec $M_{jk} = \begin{pmatrix} 0_3 & N_{jk} \\ -N_{jk} & 0_3 \end{pmatrix}$.

En annexe 2.4.2, on montre un certain nombre de propriétés assez simples relatives à la matrice N_{jk} . Il en découle bien entendu certaines propriétés pour M_{jk}^j :

- $M_{kj}^j = M_{kj}\theta_j^{-1} = -M_{jk}\theta_j^{-1} = -M_{jk}^j$.
- $\sum_{k \in \mathcal{V}_j} S_{jk} M_{jk}^j = \left(\sum_{k \in \mathcal{V}_j} S_{jk} M_{jk} \right) \theta_j^{-1} = 0$.

Enfin, on définit pour $j \in \mathcal{I}_a$, la matrice d'impédance $Z_j = c_j \theta_j = \begin{pmatrix} z_j^{-1} \mathbf{l}_3 & 0_3 \\ 0_3 & z_j \mathbf{l}_3 \end{pmatrix}$. Pour revenir à notre schéma, nous devons donc étudier :

$$i\omega V_j \theta_j \mathbf{W}_j + \sum_{k \in \mathcal{V}_j} S_{jk} \mathbf{F}_{jk} = V_j \mathbf{S}_j, \quad (3.71)$$

avec $\mathbf{F}_{jk} = M_{jk} \left(\frac{\mathbf{W}_j + \mathbf{W}_k}{2} \right)$, pour $j \in \mathcal{I}$ et $k \in \mathcal{V}_j$, sachant que $\Phi_{jk} = S_{jk} \mathbf{F}_{jk}$.

2.3.1.1 Traitement des conditions aux limites

La démarche est strictement identique à celle suivie dans le cas 1D. Dans la mesure où les cellules $(C_j)_{j \in \mathcal{I}_b}$ sont fictive, nous identifierons leurs paramètres à ceux de leurs cellules voisines réelles, si bien que pour $k \in \mathcal{V}_j^m \cup \mathcal{V}_j^a$, $(\varepsilon_k, \mu_k) = (\varepsilon_{|\Sigma_{ij}}, \mu_{|\Sigma_{ij}}) = (\varepsilon_j, \mu_j)$, d'où $M_{jk}^k = M_{jk}^j$. Il nous faut également définir les états fictifs \mathbf{W}_k de sorte que les conditions limites soient respectées faiblement. En particulier, si nous prenons :

$$\forall (j, k) \in \mathcal{I}_a \times \mathcal{V}_j^a : \mathbf{W}_k = \begin{pmatrix} 0_3 & z_j N_{jk} \\ -z_j^{-1} N_{jk} & 0_3 \end{pmatrix} \mathbf{W}_j + \xi \begin{pmatrix} \mathbf{l}_3 & -z_j N_{jk} \\ z_j^{-1} N_{jk} & \mathbf{l}_3 \end{pmatrix} \mathbf{W}^{\text{inc}}, \quad (3.72)$$

$$\forall (j, k) \in \mathcal{I}_m \times \mathcal{V}_j^m : \mathbf{W}_k = \begin{pmatrix} -\mathbf{l}_3 & 0_3 \\ 0_3 & \mathbf{l}_3 \end{pmatrix} \mathbf{W}_j + 2(\xi - 1) \begin{pmatrix} \mathbf{l}_3 & 0_3 \\ 0_3 & 0_3 \end{pmatrix} \mathbf{W}^{\text{inc}}, \quad (3.73)$$

alors le champ $(\mathbf{W}_j + \mathbf{W}_k)/2 \approx W_{|\Sigma_{jk}}$ satisfiera faiblement les conditions limites.

Comme dans le cas 1D (voir la section 2.2.1.2), la prise en compte de la condition aux limites imposée sur la frontière absorbante nécessite la diagonalisation de la matrice de flux associée à la forme conservative du système d'équations de Maxwell (2.14). Dans le cas présent, il s'agit de la matrice $M_{jk}^j = M_{jk}\theta_j^{-1}$. La diagonalisation de cette matrice est décrite dans l'annexe 2.4.3.

On utilise les conditions énoncées en annexe 2.4.4 équivalentes aux conditions limites.

Frontières absorbantes : $\forall (j, k) \in \mathcal{I}_a \times \mathcal{V}_j^a :$

$$\begin{aligned}
(M_{jk}^j)^{-\theta_j} \frac{\mathbf{W}_j + \mathbf{W}_k}{2} &= \frac{1}{2} (M_{jk}^j)^{-} \begin{pmatrix} \varepsilon_j & 0 \\ 0 & \mu_j \end{pmatrix} \begin{pmatrix} \mathbf{l}_3 & z_j^{-1} N_{jk} \\ -z_j N_{jk} & \mathbf{l}_3 \end{pmatrix} \mathbf{W}_j \\
&+ \frac{\xi}{2} (M_{jk}^j)^{-} \begin{pmatrix} \varepsilon_j & 0 \\ 0 & \mu_j \end{pmatrix} \begin{pmatrix} \mathbf{l}_3 & -z_j N_{jk} \\ z_j^{-1} N_{jk} & \mathbf{l}_3 \end{pmatrix} \mathbf{W}^{\text{inc}} \\
&= \frac{c_j}{4} \begin{pmatrix} N_{jk}^2 & z_j^{-1} N_{jk} \\ -z_j N_{jk} & N_{jk}^2 \end{pmatrix} \begin{pmatrix} \varepsilon_j & \frac{N_{jk}}{c_j} \\ -\frac{N_{jk}}{c_j} & \mu_j \end{pmatrix} \mathbf{W}_j \\
&+ \xi \frac{c_j}{4} \begin{pmatrix} N_{jk}^2 & z_j^{-1} N_{jk} \\ -z_j N_{jk} & N_{jk}^2 \end{pmatrix} \begin{pmatrix} \varepsilon_j & -\frac{N_{jk}}{c_j} \\ \frac{N_{jk}}{c_j} & \mu_j \end{pmatrix} \mathbf{W}^{\text{inc}} \\
&= \xi \frac{c_j}{2} \begin{pmatrix} N_{jk}^2 \varepsilon_j & \frac{N_{jk}}{c_j} \\ -\frac{N_{jk}}{c_j} & N_{jk}^2 \mu_j \end{pmatrix} \mathbf{W}^{\text{inc}} = \xi (M_{jk}^i)^{-\theta_j} \mathbf{W}^{\text{inc}}.
\end{aligned}$$

Frontières métalliques : $\forall (j, k) \in \mathcal{I}_m \times \mathcal{V}_j^m$ avec $\Upsilon_{jk} = \begin{pmatrix} -N_{jk} & 0_3 \\ 0_3 & 0_3 \end{pmatrix}$:

$$\begin{aligned}
\Upsilon_{jk} \frac{\mathbf{W}_j + \mathbf{W}_k}{2} &= \Upsilon_{jk} \begin{pmatrix} 0_3 & 0_3 \\ 0_3 & \mathbf{l}_3 \end{pmatrix} \mathbf{W}_j + (1 - \xi) \Upsilon_{jk} \begin{pmatrix} -\mathbf{l}_3 & 0_3 \\ 0_3 & 0_3 \end{pmatrix} \mathbf{W}^{\text{inc}} \\
&= \begin{pmatrix} -N_{jk} & 0_3 \\ 0_3 & 0_3 \end{pmatrix} \begin{pmatrix} 0_3 & 0_3 \\ 0_3 & \mathbf{l}_3 \end{pmatrix} \mathbf{W}_j \\
&+ (1 - \xi) \begin{pmatrix} -N_{jk} & 0_3 \\ 0_3 & 0_3 \end{pmatrix} \begin{pmatrix} -\mathbf{l}_3 & 0_3 \\ 0_3 & 0_3 \end{pmatrix} \mathbf{W}^{\text{inc}} \\
&= (1 - \xi) \begin{pmatrix} N_{jk} & 0_3 \\ 0_3 & 0_3 \end{pmatrix} \mathbf{W}^{\text{inc}} = (\xi - 1) \Upsilon_{jk} \mathbf{W}^{\text{inc}}.
\end{aligned}$$

Nous obtenons des résultats qui confirment que ces conditions sont vérifiées faiblement. On peut maintenant calculer le flux correspondant à chaque condition aux limites.

Flux aux frontières absorbantes : $\forall (j, k) \in \mathcal{I}_a \times \mathcal{V}_j^a$:

$$\begin{aligned}
\mathbf{F}_{jk} &= \frac{1}{2} (M_{jk} \mathbf{W}_j + M_{jk} \mathbf{W}_k) = \frac{1}{2} M_{jk} (\mathbf{W}_j + \mathbf{W}_k) \\
&= \frac{1}{2} M_{jk} \left[\begin{pmatrix} \mathbf{l}_3 & z_j N_{jk} \\ -z_j^{-1} N_{jk} & \mathbf{l}_3 \end{pmatrix} \mathbf{W}_j + \xi \begin{pmatrix} \mathbf{l}_3 & -z_j N_{jk} \\ z_j^{-1} N_{jk} & \mathbf{l}_3 \end{pmatrix} \mathbf{W}^{\text{inc}} \right] \\
&= \frac{1}{2} \left[\begin{pmatrix} -z_j^{-1} N_{jk}^2 & N_{jk} \\ -N_{jk} & -z_j N_{jk}^2 \end{pmatrix} \mathbf{W}_j + \xi \begin{pmatrix} z_j^{-1} N_{jk}^2 & N_{jk} \\ -N_{jk} & z_j N_{jk}^2 \end{pmatrix} \mathbf{W}^{\text{inc}} \right] \\
&= \frac{c_j}{2} \left[\begin{pmatrix} -\varepsilon_j N_{jk}^2 & z_j^{-1} \mu_j N_{jk} \\ -z_j \varepsilon_j N_{jk} & -\mu_j N_{jk}^2 \end{pmatrix} \mathbf{W}_j + \xi \begin{pmatrix} \varepsilon_j N_{jk}^2 & z_j^{-1} \mu_j N_{jk} \\ -z_j \varepsilon_j N_{jk} & \mu_j N_{jk}^2 \end{pmatrix} \mathbf{W}^{\text{inc}} \right] \\
&= (M_{jk}^j)^{+\theta_j} \mathbf{W}_j + \xi (M_{jk}^j)^{-\theta_j} \mathbf{W}^{\text{inc}}.
\end{aligned}$$

Comme pour le flux absorbant monodimensionnel, notre schéma à flux centrés muni de la condition (3.72) se ramène à un schéma à flux décentrés de type Steger et Warming [Steger and Warming, 1981]. Les expressions des matrices $(M_{jk}^j)^+$ et $(M_{jk}^j)^-$ sont rappelées en annexe 2.4.3.

Flux aux frontières métalliques : $\forall (j, k) \in \mathcal{I}_m \times \mathcal{V}_j^m :$

$$\begin{aligned}
\mathbf{F}_{jk} &= \frac{1}{2}(M_{jk}\mathbf{W}_j + M_{jk}\mathbf{W}_k) = \frac{1}{2}M_{jk}(\mathbf{W}_j + \mathbf{W}_k) \\
&= \frac{1}{2}M_{jk} \left[2 \begin{pmatrix} 0_3 & 0_3 \\ 0_3 & l_3 \end{pmatrix} \mathbf{W}_j + 2(1 - \xi) \begin{pmatrix} -l_3 & 0_3 \\ 0_3 & 0_3 \end{pmatrix} \mathbf{W}^{\text{inc}} \right] \\
&= \begin{pmatrix} 0_3 & N_{jk} \\ -N_{jk} & 0_3 \end{pmatrix} \begin{pmatrix} 0_3 & 0_3 \\ 0_3 & l_3 \end{pmatrix} \mathbf{W}_j + (1 - \xi) \begin{pmatrix} 0_3 & N_{jk} \\ -N_{jk} & 0_3 \end{pmatrix} \begin{pmatrix} -l_3 & 0_3 \\ 0_3 & 0_3 \end{pmatrix} \mathbf{W}^{\text{inc}} \\
&= \begin{pmatrix} 0_3 & N_{jk} \\ 0_3 & 0_3 \end{pmatrix} \mathbf{W}_j + (1 - \xi) \begin{pmatrix} 0_3 & 0_3 \\ N_{jk} & 0_3 \end{pmatrix} \mathbf{W}^{\text{inc}}.
\end{aligned}$$

Par conséquent, si on pose $Q_{jk}^m = \begin{pmatrix} 0_3 & N_{jk} \\ 0_3 & 0_3 \end{pmatrix}$, alors :

$$\begin{aligned}
\forall (j, k) \in \mathcal{I}_a \times \mathcal{V}_j^a : \quad \mathbf{F}_{jk}^a &= (M_{jk}^j)^+ \theta_j \mathbf{W}_j + \xi (M_{jk}^j)^- \theta_j \mathbf{W}^{\text{inc}}, \\
\forall (j, k) \in \mathcal{I}_m \times \mathcal{V}_j^m : \quad \mathbf{F}_{jk}^m &= Q_{jk}^m \mathbf{W}_j + (\xi - 1)(Q_{jk}^m)^t \mathbf{W}^{\text{inc}}.
\end{aligned} \tag{3.74}$$

sachant que Q_{jk}^m est antisymétrique.

2.3.1.2 Formulation matricielle

Maintenant que nous avons exprimé le flux élémentaire associé à chaque type de face du maillage \mathcal{T}_h , nous allons en écrire le bilan de manière à formuler un système matriciel similaire à celui que nous avons trouvé dans le cadre monodimensionnel. Plus précisément, pour une cellule $\{C_j\}_{j \in \mathcal{I}}$, on obtient :

$$\begin{aligned}
\sum_{k \in \mathcal{V}_j} S_{jk} \mathbf{F}_{jk} &= \sum_{k \in \mathcal{V}_j^d} S_{jk} \mathbf{F}_{jk} + \sum_{k \in \mathcal{V}_j^m} S_{jk} \mathbf{F}_{jk}^m + \sum_{k \in \mathcal{V}_j^a} S_{jk} \mathbf{F}_{jk}^a \\
&= \frac{1}{2} \sum_{k \in \mathcal{V}_j^d} S_{jk} (M_{jk} \mathbf{W}_j + M_{jk} \mathbf{W}_k) \\
&\quad + \sum_{k \in \mathcal{V}_j^m} S_{jk} \left(Q_{jk}^m \mathbf{W}_j + (\xi - 1)(Q_{jk}^m)^t \mathbf{W}^{\text{inc}} \right) \\
&\quad + \sum_{k \in \mathcal{V}_j^a} S_{jk} \left((M_{jk}^j)^+ \theta_j \mathbf{W}_j + \xi (M_{jk}^j)^- \theta_j \mathbf{W}^{\text{inc}} \right)
\end{aligned}$$

Comme dans le cas 1D, les termes qui contiennent le champ incident seront assimilés à des termes sources. Nous pouvons donc les regrouper en un seul et même terme que nous noterons de la façon suivante :

$$\gamma_j^{\text{inc}} = \xi \sum_{k \in \mathcal{V}_j^a} S_{jk} (M_{jk}^j)^- \theta_j \mathbf{W}_{|\Gamma_h^a}^{\text{inc}} + (\xi - 1) \sum_{k \in \mathcal{V}_j^m} S_{jk} (Q_{jk}^m)^t \mathbf{W}_{|\Gamma_h^m}^{\text{inc}}.$$

On obtient alors :

$$\begin{aligned} \sum_{k \in \mathcal{V}_j} S_{jk} \mathbf{F}_{jk} &= \frac{1}{2} \sum_{k \in \mathcal{V}_j^d} S_{jk} M_{jk} \mathbf{W}_j + \sum_{k \in \mathcal{V}_j^m} S_{jk} Q_{jk}^m \mathbf{W}_j \\ &+ \sum_{k \in \mathcal{V}_j^a} S_{jk} (M_{jk}^j)^+ \theta_j \mathbf{W}_j + \frac{1}{2} \sum_{k \in \mathcal{V}_j^d} S_{jk} M_{jk} \mathbf{W}_k + \gamma_j^{\text{inc}} \end{aligned}$$

ce qui nous conduit à :

$$\begin{aligned} \sum_{k \in \mathcal{V}_j} S_{jk} \mathbf{F}_{jk} &= \frac{1}{2} \left(\sum_{k \in \mathcal{V}_j^d} S_{jk} M_{jk} + \sum_{k \in \mathcal{V}_j^m} 2S_{jk} Q_{jk}^m + \sum_{k \in \mathcal{V}_j^a} 2S_{jk} (M_{jk}^j)^+ \theta_j \right) \mathbf{W}_j \\ &+ \frac{1}{2} \sum_{k \in \mathcal{V}_j^d} S_{jk} M_{jk} \mathbf{W}_k + \gamma_j^{\text{inc}}. \end{aligned} \quad (3.75)$$

La deuxième assertion de la proposition 6 de l'annexe 2.4.2 donne :

$$\begin{aligned} \sum_{k \in \mathcal{V}_j} S_{jk} M_{jk} = 0 &\Rightarrow \sum_{k \in \mathcal{V}_j^d} S_{jk} M_{jk} + \sum_{k \in \mathcal{V}_j^m} S_{jk} M_{jk} + \sum_{k \in \mathcal{V}_j^a} S_{jk} M_{jk} = 0 \\ &\Rightarrow \sum_{k \in \mathcal{V}_j^d} S_{jk} M_{jk} = - \sum_{k \in \mathcal{V}_j^m} S_{jk} M_{jk} - \sum_{k \in \mathcal{V}_j^a} S_{jk} M_{jk}. \end{aligned} \quad (3.76)$$

En injectant (3.76) dans (3.75), nous trouvons :

$$\begin{aligned} \sum_{k \in \mathcal{V}_j} S_{jk} \mathbf{F}_{jk} &= \frac{1}{2} \left(\sum_{k \in \mathcal{V}_j^m} S_{jk} (2Q_{jk}^m - M_{jk}) + \sum_{k \in \mathcal{V}_j^a} S_{jk} (2(M_{jk}^j)^+ \theta_j - M_{jk}) \right) \mathbf{W}_j \\ &+ \frac{1}{2} \sum_{k \in \mathcal{V}_j^d} S_{jk} M_{jk} \mathbf{W}_k + \gamma_j^{\text{inc}}, \end{aligned}$$

Or des calculs simples donnent :

$$\begin{aligned} 2Q_{jk}^m - M_{jk} &= \begin{pmatrix} \mathbf{0}_3 & N_{jk} \\ N_{jk} & \mathbf{0}_3 \end{pmatrix}, \\ 2(M_{jk}^j)^+ \theta_j - M_{jk} &= \begin{pmatrix} -z_j^{-1} N_{jk}^2 & \mathbf{0}_3 \\ \mathbf{0}_3 & -z_j N_{jk}^2 \end{pmatrix}. \end{aligned}$$

Si on pose $P_{jk}^m = \begin{pmatrix} 0_3 & N_{jk} \\ N_{jk} & 0_3 \end{pmatrix}$ et $P_{jk}^a = \begin{pmatrix} -z_j^{-1}N_{jk}^2 & 0_3 \\ 0_3 & -z_j N_{jk}^2 \end{pmatrix}$, alors :

$$\sum_{k \in \mathcal{V}_j} S_{jk} \mathbf{F}_{jk} = \frac{1}{2} \left(\sum_{k \in \mathcal{V}_j^m} S_{jk} P_{jk}^m + \sum_{k \in \mathcal{V}_j^a} S_{jk} P_{jk}^a \right) \mathbf{W}_j + \frac{1}{2} \sum_{k \in \mathcal{V}_j^d} S_{jk} M_{jk} \mathbf{W}_k + \gamma_j^{\text{inc}}. \quad (3.77)$$

La dernière étape consiste à injecter (3.77) dans (3.71). On obtient :

$$i\omega V_j \theta_j \mathbf{W}_j + \frac{1}{2} \left(\sum_{k \in \mathcal{V}_j^m} S_{jk} P_{jk}^m + \sum_{k \in \mathcal{V}_j^a} S_{jk} P_{jk}^a \right) \mathbf{W}_j + \frac{1}{2} \sum_{k \in \mathcal{V}_j^d} S_{jk} M_{jk} \mathbf{W}_k + \gamma_j^{\text{inc}} = V_j \mathbf{S}_j,$$

pour aboutir au schéma suivant où $\forall j \in \mathcal{I}$:

$$\left(2i\omega V_j \theta_j + \sum_{k \in \mathcal{V}_j^m} S_{jk} P_{jk}^m + \sum_{k \in \mathcal{V}_j^a} S_{jk} P_{jk}^a \right) \mathbf{W}_j + \sum_{k \in \mathcal{V}_j^d} S_{jk} M_{jk} \mathbf{W}_k = 2(V_j \mathbf{S}_j - \gamma_j^{\text{inc}}). \quad (3.78)$$

Il s'agit maintenant d'écrire une formulation matricielle du schéma (3.78). Ce faisant, nous voulons montrer que la matrice globale du système possède une structure identique à celle du cas 1D. On désigne par $\mathbf{b} \in \mathbb{C}^{6 \times N}$ le second membre et par $\mathcal{A} \in \mathbf{M}_{6 \times N}(\mathbb{C})$ la matrice globale résultant de l'assemblage du schéma (3.78).

Pour $\mathbf{W} \in \mathbb{C}^{6 \times N}$, on trouve grâce au bilan des flux pour la cellule C_j que (3.78) s'écrit sous la forme $(\mathcal{A}\mathbf{W})_j = \mathbf{b}_j$ avec $\mathcal{A} = i\mathcal{D} + \mathcal{Z} + \mathcal{R}$ et :

$$\begin{cases} (\mathcal{D}\mathbf{W})_j = 2\omega V_j \theta_j \mathbf{W}_j, \\ (\mathcal{Z}\mathbf{W})_j = \sum_{k \in \mathcal{V}_j^a} S_{jk} P_{jk}^a \mathbf{W}_j, \\ (\mathcal{R}\mathbf{W})_j = \sum_{k \in \mathcal{V}_j^m} S_{jk} P_{jk}^m \mathbf{W}_j + \sum_{k \in \mathcal{V}_j^d} S_{jk} M_{jk} \mathbf{W}_k, \\ \mathbf{b}_j = 2(V_j \mathbf{S}_j - \gamma_j^{\text{inc}}). \end{cases} \quad (3.79)$$

Proposition 3 Les matrices \mathcal{Z} , \mathcal{R} et \mathcal{D} ont les propriétés suivantes :

- (i) \mathcal{D} est réelle symétrique définie positive,
- (ii) \mathcal{Z} est réelle symétrique et positive,
- (iii) \mathcal{R} est réelle antisymétrique.

Preuve

(i) \mathcal{D} est clairement diagonale par blocs, avec $\mathcal{D}_{jj} = 2\omega V_j \theta_j$ qui est symétrique définie positive. On a donc que \mathcal{D} est symétrique définie positive.

(ii) \mathcal{Z} est aussi diagonale par blocs et $\mathcal{Z}_{jj} = \sum_{k \in \mathcal{V}_j^a} S_{jk} P_{jk}^a$. De plus, P_{jk}^a est symétrique positive (voir proposition 13). Donc $\sum_{j \in \mathcal{V}_j^a} S_{jk} P_{jk}^a$ est symétrique positive. On en déduit que \mathcal{Z} est symétrique et positive.

(iii) D'après son développement :

$$\forall j \in \mathcal{I}, \mathcal{R}_{jj} = \sum_{k \in \mathcal{V}_j^m} S_{jk} P_{jk}^m \text{ et } \mathcal{R}_{jk} = \begin{cases} S_{jk} M_{jk} & \text{pour } k \in \mathcal{V}_j^d, \\ 0 & \text{pour } k \in \mathcal{I} \setminus (\{j\} \cup \mathcal{V}_j^d). \end{cases}$$

Nous avons prouvé en proposition 13 de l'annexe 2.4.6 que P_{jk}^m et M_{jk} sont respectivement anti-symétrique et symétrique, on en déduit que :

$$\forall j \in \mathcal{I}, \mathcal{R}_{jj} \text{ est antisymétrique et } \forall (j, k) \in \mathcal{I}^2, j \neq k, \mathcal{R}_{jk} \text{ est symétrique}$$

D'autre part, nous avons aussi que $\mathcal{R}_{kj} = -\mathcal{R}_{jk}$ puisque $M_{kj} = -M_{jk}$. On est en mesure d'appliquer la proposition 11 de l'annexe 2.4.6, et donc de conclure que \mathcal{R} est antisymétrique. ■

En résumé, nous avons une matrice ayant une structure identique à la matrice obtenue en 1D.

2.3.1.3 Étude de la dispersion numérique

Dans ce paragraphe, nous allons étudier la dispersion numérique de notre schéma. L'étude de la dispersion d'un schéma permet de connaître sa capacité à propager correctement une onde. En effet, l'ordre de précision de la relation de dispersion discrète est souvent mis en avant pour évaluer la précision d'un schéma numérique. L'étude de dispersion consiste à s'intéresser à des solutions particulières des équations de Maxwell, appelées ondes planes, qui sont de la forme :

$$\mathbf{W}(x, y, z) = \mathbf{W}_0 e^{-i\mathbf{k} \cdot (x, y, z)^t}.$$

En injectant cette onde dans les EM continues, on trouve la relation de dispersion exacte $\omega_{ex}^2 = c^2 |\mathbf{k}|^2$.

En considérant une grille cartésienne dont les cellules sont de taille $\Delta x \times \Delta y \times \Delta z$ et dont les caractéristiques du milieu ε et μ sont constantes, nous pouvons rechercher les solutions particulières de notre schéma, de type ondes planes numériques, qui sont de la forme suivante : pour chaque cellule $C_{j,l,m}$, dont le centre de gravité a pour coordonnées $(j\Delta x, l\Delta y, m\Delta z)$:

$$\mathbf{W}_{j,l,m} = \mathbf{W}_0 e^{-i(jk_x \Delta x + lk_y \Delta y + mk_z \Delta z)},$$

ce qui nous amènera à connaître la relation de dispersion discrète entre ω , \mathbf{k} le nombre d'onde et c la vitesse de l'onde dans le milieu de propagation. Sous les hypothèses précédentes, notre schéma (3.78) s'écrit :

$$\begin{aligned} i\omega \Delta x \Delta y \Delta z \theta_{j,l,m} \mathbf{W}_{j,l,m} &+ \Delta y \Delta z \frac{M_{j+1,l,m} \mathbf{W}_{j+1,l,m} + M_{j-1,l,m} \mathbf{W}_{j-1,l,m}}{2} \\ &+ \Delta x \Delta z \frac{M_{j,l+1,m} \mathbf{W}_{j,l+1,m} + M_{j,l-1,m} \mathbf{W}_{j,l-1,m}}{2} \\ &+ \Delta x \Delta y \frac{M_{j,l,m+1} \mathbf{W}_{j,l,m+1} + M_{j,l,m-1} \mathbf{W}_{j,l,m-1}}{2} = 0, \end{aligned}$$

et conduit au schéma différences finies centrées suivant :

$$i\omega \begin{pmatrix} E_{j,l,m}^x \\ E_{j,l,m}^y \\ E_{j,l,m}^z \end{pmatrix} = \frac{1}{\mu} \begin{pmatrix} \frac{H_{j,l,m+1}^y - H_{j,l,m-1}^y}{2\Delta z} - \frac{H_{j,l+1,m}^z - H_{j,l-1,m}^z}{2\Delta y} \\ \frac{H_{j+1,l,m}^z - H_{j-1,l,m}^z}{2\Delta x} - \frac{H_{j,l,m+1}^x - H_{j,l,m-1}^x}{2\Delta y} \\ \frac{H_{j,l+1,m}^x - H_{j,l-1,m}^x}{2\Delta y} - \frac{H_{j+1,l,m}^y - H_{j-1,l,m}^y}{2\Delta x} \end{pmatrix} \quad (3.80)$$

et :

$$i\omega \begin{pmatrix} H_{j,l,m}^x \\ H_{j,l,m}^y \\ H_{j,l,m}^z \end{pmatrix} = -\frac{1}{\varepsilon} \begin{pmatrix} \frac{E_{j,l,m+1}^y - E_{j,l,m-1}^y}{2\Delta z} - \frac{E_{j,l+1,m}^z - E_{j,l-1,m}^z}{2\Delta y} \\ \frac{E_{j+1,l,m}^z - E_{j-1,l,m}^z}{2\Delta x} - \frac{E_{j,l,m+1}^x - E_{j,l,m-1}^x}{2\Delta y} \\ \frac{E_{j,l+1,m}^x - E_{j,l-1,m}^x}{2\Delta y} - \frac{E_{j+1,l,m}^y - E_{j-1,l,m}^y}{2\Delta x} \end{pmatrix}. \quad (3.81)$$

En injectant (3.81) dans (3.80) et en supposant vérifié au niveau discret la relation $\nabla(\text{div}\mathbf{E}) = 0^3$, (3.80) devient :

$$\omega^2 \mathbf{E}_{j,l,m} = c^2 \left(\frac{-\mathbf{E}_{j-2,l,m} + 2\mathbf{E}_{j,l,m} - \mathbf{E}_{j+2,l,m}}{4\Delta x^2} + \frac{-\mathbf{E}_{j,l-2,m} + 2\mathbf{E}_{j,l,m} - \mathbf{E}_{j,l+2,m}}{4\Delta y^2} + \frac{-\mathbf{E}_{j,l,m-2} + 2\mathbf{E}_{j,l,m} - \mathbf{E}_{j,l,m+2}}{4\Delta z^2} \right) \quad (3.82)$$

Il se trouve que (3.82) correspond exactement à une discrétisation de l'équation de Helmholtz par un schéma différences finies centrées. En effet, si on part des équations de Maxwell en domaine fréquentiel, on obtient sous l'hypothèse que ε et μ sont constantes :

$$\begin{cases} i\omega \mathbf{E} - \nabla \times (\mu^{-1} \mathbf{H}) = 0 \\ i\omega \mathbf{H} + \nabla \times (\varepsilon^{-1} \mathbf{E}) = 0 \end{cases} \Leftrightarrow \begin{cases} i\omega \mathbf{E} - \mu^{-1} \text{rot} \mathbf{H} = 0 \\ \mathbf{H} = \frac{i}{\varepsilon \omega} \text{rot} \mathbf{E} \end{cases}$$

$$\Leftrightarrow \omega^2 \mathbf{E} + c^2 (\Delta \mathbf{E} - \nabla(\text{div} \mathbf{E})) = 0,$$

$$\Leftrightarrow \omega^2 \mathbf{E} + c^2 \Delta \mathbf{E} = 0,$$

dont la discrétisation en différences centrées aboutit à (3.82), ce qui montre le lien très étroit entre l'équation de Helmholtz et les équations de Maxwell en domaine fréquentiel. Nous allons maintenant évaluer les termes du second membre de (3.82) en y injectant l'onde plane numérique définie auparavant.

³En l'absence de densité de charge volumique la loi de Gauss donne $\text{div} \mathbf{E} = 0$.

On trouve :

$$\begin{aligned}
\frac{-\mathbf{E}_{j-2,l,m} + 2\mathbf{E}_{j,l,m} - \mathbf{E}_{j+2,l,m}}{4\Delta x^2} &= \frac{-e^{j2k_x\Delta x} + 2 - e^{-j2k_x\Delta x}}{4\Delta x^2} \mathbf{E}_{j,l,m} \\
&= \frac{2 - 2\cos(2k_x\Delta x)}{4\Delta x^2} \mathbf{E}_{j,l,m} \\
&= \frac{\sin^2(k_x\Delta x)}{\Delta x^2} \mathbf{E}_{j,l,m}.
\end{aligned}$$

On en déduit donc que, pour une solution discrète non identiquement nulle, ω et \mathbf{k} sont liés par la relation de dispersion discrète suivante :

$$\frac{\omega^2}{c^2} = \frac{\sin^2(k_x\Delta x)}{\Delta x^2} + \frac{\sin^2(k_y\Delta y)}{\Delta y^2} + \frac{\sin^2(k_z\Delta z)}{\Delta z^2}. \quad (3.83)$$

Cette relation signifie que les ondes planes ne se propagent pas exactement avec la bonne pulsation ou la bonne vitesse. On remarque néanmoins qu'on retrouve la relation de dispersion exacte quand les trois pas d'espace Δx , Δy et Δz tendent vers 0.

En effectuant un développement limité en Δx , Δy et Δz en 0 d'ordre 5 dans (3.83), on trouve, en notant $h = \max(\Delta x, \Delta y, \Delta z)$, la relation de dispersion approchée :

$$\frac{\omega^2}{\omega_{ex}^2} = 1 - \frac{k_x^4\Delta x^2 + k_y^4\Delta y^2 + k_z^4\Delta z^2}{3|k|^2} + O(|k|^4h^4). \quad (3.84)$$

Il est intéressant de comparer cette relation à celle obtenue dans le domaine temporel. En effet, en se référant à [Remaki, 2000], l'erreur de dispersion dans le domaine temporel pour ce même schéma en espace et pour un schéma explicite de type saute-mouton pour l'intégration en temps est donnée par :

$$\frac{\omega^2}{\omega_{ex}^2} \left(1 - \frac{\omega_{ex}^2\Delta t^2}{12} + O(\omega_{ex}^4\Delta t^4) \right) = 1 - \frac{k_x^4\Delta x^2 + k_y^4\Delta y^2 + k_z^4\Delta z^2}{3|k|^2} + O(|k|^4h^4).$$

Nous remarquons que notre relation de dispersion est la même que la relation en temporel lorsque le pas de temps tend vers 0. Mais nous constatons aussi que notre schéma est également plus dispersif que le schéma en temporel, cette différence diminuant lorsque l'on prend un pas de temps de plus en plus petit.

2.3.2 Méthode de type Galerkin discontinu

Dans cette section, on décrit en détail la méthode de type Galerkin discontinu introduite dans la section 2.1.2. Nous considérons de nouveau l'équation :

$$i\omega\theta\mathbf{W} + \text{div}F(\mathbf{W}) = \mathbf{S}. \quad (3.85)$$

On multiplie (3.85) par une fonction test scalaire φ_j et on intègre sur une cellule C_j :

$$\begin{aligned}
\int_{C_j} (i\omega\theta\mathbf{W} + \text{div}F(\mathbf{W}))\varphi_j dx &= \int_{C_j} \mathbf{S}\varphi_j dx \\
\Leftrightarrow i\omega\theta_j \int_{C_j} \mathbf{W}\varphi_j dx + \int_{C_j} (\text{div}F(\mathbf{W}))\varphi_j dx &= \int_{C_j} \mathbf{S}\varphi_j dx.
\end{aligned}$$

En appliquant le théorème de Green-Riemann, on obtient :

$$i\omega\theta_j \int_{C_j} \mathbf{W} \varphi_j dx - \int_{C_j} F(\mathbf{W}) \cdot \nabla \varphi_j dx + \int_{\partial C_j} (F(\mathbf{W}) \cdot \mathbf{n}) \varphi_j ds = \int_{C_j} \mathbf{S} \varphi_j dx. \quad (3.86)$$

Comme dans le cas 1D, on pose $\phi_j = (\varphi_{j1}, \varphi_{j2}, \dots, \varphi_{jd_j})$ la base locale de $\mathcal{P}_j = \mathbb{P}_m[C_j]$ et on note \mathbf{W}_j la projection de \mathbf{W} sur \mathcal{P}_j (de même, \mathbf{S}_j est la projection de \mathbf{S} sur \mathcal{P}_j). Il s'agit maintenant d'injecter \mathbf{W}_j dans l'équation (3.86). Notre problème est alors de chercher $\mathbf{W}_j \in \mathcal{P}_j$ tel que :

$$\begin{aligned} i\omega\theta_j \int_{C_j} \mathbf{W}_j \varphi_j dx - \int_{C_j} F(\mathbf{W})_j \cdot \nabla \varphi_j dx \\ + \int_{\partial C_j} (F(\mathbf{W}) \cdot \mathbf{n}) \varphi_j ds &= \int_{C_j} \mathbf{S} \varphi_j dx \\ \Leftrightarrow i\omega\theta_j \int_{C_j} \mathbf{W}_j \varphi_j dx - \int_{C_j} F(\mathbf{W})_j \cdot \nabla \varphi_j dx \\ + \sum_{k \in \mathcal{V}_j} \int_{\Sigma_{jk}} (F(\mathbf{W}) \cdot \mathbf{n}_{jk}) \varphi_j ds &= \int_{C_j} \mathbf{S}_j \varphi_j dx. \end{aligned} \quad (3.87)$$

Il reste à approcher les flux sur C_j ainsi que sur Σ_{jk} . Nous choisissons d'approcher le flux sur Σ_{jk} par un schéma centré :

$$F(\mathbf{W})|_{\Sigma_{jk}} \approx \frac{F(\mathbf{W})_j + F(\mathbf{W})_k}{2}.$$

Par conséquent :

$$\begin{aligned} \sum_{k \in \mathcal{V}_j} \int_{\Sigma_{jk}} (F(\mathbf{W}) \cdot \mathbf{n}_{jk}) \varphi_j ds &= \sum_{k \in \mathcal{V}_j} \int_{\Sigma_{jk}} \left(\frac{F(\mathbf{W})_j + F(\mathbf{W})_k}{2} \cdot \mathbf{n}_{jk} \right) \varphi_j ds \\ &= \frac{1}{2} \sum_{k \in \mathcal{V}_j} \int_{\Sigma_{jk}} (F(\mathbf{W})_j \cdot \mathbf{n}_{jk}) \varphi_j ds \\ &+ \frac{1}{2} \sum_{k \in \mathcal{V}_j} \int_{\Sigma_{jk}} (F(\mathbf{W})_k \cdot \mathbf{n}_{jk}) \varphi_j ds \\ &= \frac{1}{2} \int_{\partial C_j} (F(\mathbf{W})_j \cdot \mathbf{n}_{jk}) \varphi_j ds \\ &+ \frac{1}{2} \int_{\partial C_j} (F(\mathbf{W})_k \cdot \mathbf{n}_{jk}) \varphi_j ds. \end{aligned}$$

En appliquant dans l'autre sens le théorème de Green-Riemann au premier terme du membre de droite, on obtient :

$$\begin{aligned} \sum_{k \in \mathcal{V}_j} \int_{\Sigma_{jk}} (F(\mathbf{W}) \cdot \mathbf{n}_{jk}) \varphi_j ds &= \frac{1}{2} \int_{C_j} ((\operatorname{div} F(\mathbf{W}))_j \varphi_j + F(\mathbf{W})_j \cdot \nabla \varphi_j) dx \\ &+ \frac{1}{2} \sum_{k \in \mathcal{V}_j} \int_{\Sigma_{jk}} (F(\mathbf{W})_k \cdot \mathbf{n}_{jk}) \varphi_j ds. \end{aligned} \quad (3.88)$$

En injectant (3.88) dans la dernière équation de (3.87), on aboutit à :

$$\begin{aligned}
i\omega\theta_j \int_{C_j} \mathbf{W}_j \varphi_j dx &- \int_{C_j} F(\mathbf{W})_j \cdot \nabla \varphi_j dx \\
&+ \frac{1}{2} \int_{C_j} ((\operatorname{div} F(\mathbf{W}))_j) \varphi_j + F(\mathbf{W})_j \cdot \nabla \varphi_j dx \\
&+ \frac{1}{2} \sum_{k \in \mathcal{V}_j} \int_{\Sigma_{jk}} (F(\mathbf{W})_k \cdot \mathbf{n}_{jk}) \varphi_j ds = \int_{C_j} \mathbf{S}_j \varphi_j dx.
\end{aligned} \tag{3.89}$$

On pose :

$$F(\mathbf{W})_j \equiv F(\mathbf{W})|_{C_j} \approx F_j(\mathbf{W}_j).$$

Par suite, l'équation (3.89) devient :

$$\begin{aligned}
i\omega\theta_j \int_{C_j} \mathbf{W}_j \varphi_j dx &- \int_{C_j} F_j(\mathbf{W}_j) \cdot \nabla \varphi_j dx \\
&+ \frac{1}{2} \int_{C_j} ((\operatorname{div} F_j(\mathbf{W}_j))) \varphi_j + F_j(\mathbf{W}_j) \cdot \nabla \varphi_j dx \\
&+ \frac{1}{2} \sum_{k \in \mathcal{V}_j} \int_{\Sigma_{jk}} (F_k(\mathbf{W}_k) \cdot \mathbf{n}_{jk}) \varphi_j ds = \int_{C_j} \mathbf{S}_j \varphi_j dx.
\end{aligned} \tag{3.90}$$

En regroupant les termes de gradients et de divergence, (3.90) devient :

$$\begin{aligned}
i\omega\theta_j \int_{C_j} \mathbf{W}_j \varphi_j dx &+ \frac{1}{2} \int_{C_j} ((\operatorname{div} F_j(\mathbf{W}_j))) \varphi_j - F_j(\mathbf{W}_j) \cdot \nabla \varphi_j dx \\
&+ \frac{1}{2} \sum_{k \in \mathcal{V}_j} \int_{\Sigma_{jk}} (F_k(\mathbf{W}_k) \cdot \mathbf{n}_{jk}) \varphi_j ds = \int_{C_j} \mathbf{S}_j \varphi_j dx,
\end{aligned} \tag{3.91}$$

qui, après développement, donne :

$$\begin{aligned}
i\omega\theta_j \int_{C_j} \mathbf{W}_j \varphi_j dx &+ \frac{1}{2} \int_{C_j} \left(\sum_{l=1}^3 \partial_{x_l} (G_{x_l})_j \mathbf{W}_j \right) \varphi_j dx \\
&- \frac{1}{2} \int_{C_j} \sum_{l=1}^3 (\partial_{x_l} \varphi_j) (G_{x_l})_j \mathbf{W}_j dx \\
&+ \frac{1}{2} \sum_{k \in \mathcal{V}_j} \int_{\Sigma_{jk}} \left(\sum_{l=1}^3 n_{jk}^{x_l} (G_{x_l})_k \mathbf{W}_k \right) \varphi_j ds = \int_{C_j} \mathbf{S}_j \varphi_j dx,
\end{aligned} \tag{3.92}$$

avec $\partial_{x_l} (G_{x_l})_j \mathbf{W}_j \equiv \frac{\partial (G_{x_l})_j \mathbf{W}_j}{\partial x_l}$, $\partial_{x_l} \varphi_j \equiv \frac{\partial \varphi_j}{\partial x_l}$ et $\mathbf{n}_{jk} = (n_{jk}^{x_1}, n_{jk}^{x_2}, n_{jk}^{x_3})^t$. L'équation (3.92) peut donc se réécrire :

$$\begin{aligned}
i\omega\theta_j \int_{C_j} \mathbf{W}_j \varphi_j dx &+ \frac{1}{2} \int_{C_j} \left(\sum_{l=1}^3 (\partial_{x_l} (G_{x_l})_j \mathbf{W}_j) \varphi_j - \sum_{l=1}^3 (\partial_{x_l} \varphi_j) (G_{x_l})_j \mathbf{W}_j \right) dx \\
&+ \frac{1}{2} \sum_{k \in \mathcal{V}_j} \int_{\Sigma_{jk}} M_{jk} \mathbf{W}_k \varphi_j ds = \int_{C_j} \mathbf{S}_j \varphi_j dx,
\end{aligned} \tag{3.93}$$

où $M_{jk}^k = M_{jk}\theta_k^{-1}$ si $M_{jk} = \begin{pmatrix} 0_3 & N_{jk} \\ -N_{jk} & 0_3 \end{pmatrix}$ et $\theta_k = \begin{pmatrix} \varepsilon_k & 0_3 \\ 0_3 & \mu_k \end{pmatrix}$.

2.3.2.1 Traitement des conditions aux limites

On approche les flux sur les cellules qui s'appuient sur les frontières physiques comme pour le schéma volumes finis. On utilise donc les relations (3.72) et (3.73), ainsi que les matrices qui seront utiles pour les prochains calculs :

$$Q_{jk}^m = \begin{pmatrix} 0_3 & N_{jk} \\ 0_3 & 0_3 \end{pmatrix}, \quad Z_j = \begin{pmatrix} z_j^{-1}l_3 & 0_3 \\ 0_3 & z_jl_3 \end{pmatrix},$$

$$P_{jk}^a = \begin{pmatrix} -z_j^{-1}N_{jk}^2 & 0_3 \\ 0_3 & -z_jN_{jk}^2 \end{pmatrix}, \quad P_{jk}^m = \begin{pmatrix} 0_3 & N_{jk} \\ N_{jk} & 0_3 \end{pmatrix}.$$

Nous pouvons alors exprimer les flux frontières. Pour tout $j \in \mathcal{I}_a$ et tout $k \in \mathcal{V}_i^a$:

$$\begin{aligned} M_{jk}\mathbf{W}_k &= M_{jk} \left(\begin{pmatrix} 0_3 & z_jN_{jk} \\ -z_j^{-1}N_{jk} & 0_3 \end{pmatrix} \mathbf{W}_j + \xi \begin{pmatrix} l_3 & -z_jN_{jk} \\ z_j^{-1}N_{jk} & l_3 \end{pmatrix} \mathbf{W}^{\text{inc}} \right) \\ &= M_{jk}^2 Z_j \mathbf{W}_j + \xi M_{jk} (l_6 - M_{jk} Z_j) \mathbf{W}^{\text{inc}} \\ &= M_{jk}^2 Z_j \mathbf{W}_j + \xi (M_{jk} - M_{jk}^2 Z_j) \mathbf{W}^{\text{inc}} \\ &= P_{jk}^a \mathbf{W}_j + 2\xi (M_{jk}^j)^- \theta_j \mathbf{W}^{\text{inc}}. \end{aligned} \tag{3.94}$$

Pour tout $j \in \mathcal{I}_m$ et tout $k \in \mathcal{V}_j^m$:

$$\begin{aligned} M_{jk}\mathbf{W}_k &= M_{jk} \mathbf{W}_k \\ &= M_{jk} \left(\begin{pmatrix} -l_3 & 0_3 \\ 0_3 & l_3 \end{pmatrix} \mathbf{W}_j + 2(1 - \xi) \begin{pmatrix} -l_3 & 0_3 \\ 0_3 & 0_3 \end{pmatrix} \mathbf{W}^{\text{inc}} \right) \\ &= P_{jk}^m \mathbf{W}_j + 2(\xi - 1)(Q_{jk}^m)^t \mathbf{W}^{\text{inc}}. \end{aligned} \tag{3.95}$$

Maintenant, si on se sert des relations (3.94) et (3.95) dans la quantité ci-dessous :

$$\begin{aligned} \sum_{k \in \mathcal{V}_j} \int_{\Sigma_{jk}} M_{jk} \mathbf{W}_k \varphi_j ds &= \sum_{k \in \mathcal{V}_j^d} \int_{\Sigma_{jk}} M_{jk} \mathbf{W}_k \varphi_j ds + \sum_{k \in \mathcal{V}_j^a} \int_{\Sigma_{jk}} M_{jk} \mathbf{W}_k \varphi_j ds \\ &+ \sum_{k \in \mathcal{V}_j^m} \int_{\Sigma_{jk}} M_{jk} \mathbf{W}_k \varphi_j ds, \end{aligned} \tag{3.96}$$

on trouve que les flux aux cellules voisines s'écrivent :

$$\begin{aligned} \sum_{k \in \mathcal{V}_j} \int_{\Sigma_{jk}} M_{jk} \mathbf{W}_k \varphi_j ds &= \sum_{k \in \mathcal{V}_j^d} \int_{\Sigma_{jk}} M_{jk} \mathbf{W}_k \varphi_j ds \\ &+ \sum_{k \in \mathcal{V}_j^a} \int_{\Sigma_{jk}} \left(P_{jk}^a \mathbf{W}_j + 2\xi (M_{jk}^j)^- \theta_j \mathbf{W}^{\text{inc}} \right) \varphi_j ds \\ &+ \sum_{k \in \mathcal{V}_j^m} \int_{\Sigma_{jk}} \left(P_{jk}^m \mathbf{W}_j + 2(\xi - 1)(Q_{jk}^m)^t \mathbf{W}^{\text{inc}} \right) \varphi_j ds, \end{aligned}$$

soit encore :

$$\begin{aligned} \sum_{k \in \mathcal{V}_j} \int_{\Sigma_{jk}} M_{jk} \mathbf{W}_k \varphi_j ds &= \sum_{k \in \mathcal{V}_j^d} \int_{\Sigma_{jk}} M_{jk} \mathbf{W}_k \varphi_j ds + \sum_{k \in \mathcal{V}_j^a} \int_{\Sigma_{jk}} P_{jk}^a \mathbf{W}_j \varphi_j ds \\ &+ \sum_{k \in \mathcal{V}_j^m} \int_{\Sigma_{jk}} P_{jk}^m \mathbf{W}_j \varphi_j ds + 2\gamma_j^{\text{inc}}, \end{aligned} \quad (3.97)$$

avec :

$$\gamma_j^{\text{inc}} = \xi \sum_{k \in \mathcal{V}_j^a} \int_{\Sigma_{jk}} (M_{jk}^j)^{-1} \theta_j \mathbf{W}_{|\Gamma_h^a}^{\text{inc}} \varphi_j ds + (\xi - 1) \sum_{k \in \mathcal{V}_j^m} \int_{\Sigma_{jk}} (Q_{jk}^m)^t \mathbf{W}_{|\Gamma_h^m}^{\text{inc}} \varphi_j ds.$$

Comme dans le cas 1D, nous supposons maintenant que $\mathbf{W}_j(\mathbf{x}) = \sum_{k=1}^{d_j} \mathbf{W}_{jk} \varphi_{jk}(\mathbf{x})$, où $\mathcal{P}_j = \mathbb{P}_m[C_j]$ est pris comme l'espace engendré par les fonctions scalaires polynomiales de degré au plus égal à m sur la cellule C_j et les degrés de liberté locaux sont $\mathbf{W}_{jk} \in \mathbb{C}^6$. Posons $\mathbb{W}_j = (\mathbf{W}_{j1}, \mathbf{W}_{j2}, \dots, \mathbf{W}_{jd_j})$ (\mathbb{W}_j est une matrice rectangle $6 \times d_j$) et $\phi_j = (\varphi_{j1}, \varphi_{j2}, \dots, \varphi_{jd_j})$. On a $\mathbf{W}_j = \mathbb{W}_j \phi_j^t$. On introduit alors les matrices suivantes :

$$\left\{ \begin{array}{ll} \Phi_j &= \int_{C_j} \phi_j^t \phi_j dx & (\Phi_j \text{ est symétrique définie positive}), \\ \Phi_j^{x_l} &= \int_{C_j} (\phi_j^t (\partial_{x_l} \phi_j) - (\partial_{x_l} \phi_j)^t \phi_j) dx & (\Phi_j^{x_l} \text{ est antisymétrique}), \\ \Phi_{jk} &= \int_{\Sigma_{jk}} \phi_j^t \phi_k ds & (\Phi_{jk} \text{ est quelconque de taille } d_j \times d_k). \end{array} \right.$$

Maintenant, en prenant pour φ_j successivement toutes les fonctions de base φ_{jk} pour $1 \leq k \leq d_j$, on peut reformuler l'équation (3.93) sous la forme, $\forall C_j \in \mathcal{T}_h$:

$$\begin{aligned} 2i\omega \theta_j \mathbb{W}_j \Phi_j^t + \sum_{l=1}^3 G_{x_l} \mathbb{W}_j (\Phi_j^{x_l})^t &+ \sum_{k \in \mathcal{V}_j^a} P_{jk}^a \mathbb{W}_j (\Phi_{jk})^t \\ &+ \sum_{k \in \mathcal{V}_j^m} P_{jk}^m \mathbb{W}_j (\Phi_{jk})^t \\ &+ \sum_{k \in \mathcal{V}_j^d} M_{jk} \mathbb{W}_k (\Phi_{jk})^t = \mathbf{b}_j, \end{aligned} \quad (3.98)$$

où :

$$\mathbf{b}_j = 2 \left(\mathbb{S}_j \Phi_j^t - \left(\xi \sum_{k \in \mathcal{V}_j^a} (M_{jk}^j)^{-1} \theta_j \mathbb{W}_{|\Gamma_h^a}^{\text{inc}} (\Phi_{jk})^t + (\xi - 1) \sum_{k \in \mathcal{V}_j^m} (Q_{jk}^m)^t \mathbb{W}_{|\Gamma_h^m}^{\text{inc}} (\Phi_{jk})^t \right) \right)$$

On constate que cette formulation se rapproche de celle établie dans le cas 1D, ce qui va nous amener à réutiliser le procédé de vectorisation qui transforme la matrice $\mathbb{W}_j = (\mathbf{W}_{j1}, \mathbf{W}_{j2}, \dots, \mathbf{W}_{jd_j})$ en un

vecteur colonne $(\mathbf{W}_{j1}, \mathbf{W}_{j2}, \dots, \mathbf{W}_{jd_j})^t \in \mathbb{C}^{6 \times d_j}$. Ce rangement correspond toujours à un empilement par coordonnées selon un même degré de liberté. Nous obtenons alors grâce au produit de Kronecker :

$$\begin{aligned}
2i\omega(\Phi_j \otimes \theta_j)\aleph(\mathbb{W}_j) + \sum_{l=1}^3 (\Phi_j^{x_l} \otimes G_{x_l})\aleph(\mathbb{W}_j) &+ \sum_{k \in \mathcal{V}_j^a} (\Phi_{jk} \otimes P_{jk}^a)\aleph(\mathbb{W}_j) \\
&+ \sum_{k \in \mathcal{V}_j^m} (\Phi_{jk} \otimes P_{jk}^m)\aleph(\mathbb{W}_j) \\
&+ \sum_{k \in \mathcal{V}_j^d} (\Phi_{jk} \otimes M_{jk})\aleph(\mathbb{W}_k) = \aleph(\mathbf{b}_j),
\end{aligned} \tag{3.99}$$

qui, après factorisation, donne :

$$\begin{aligned}
\left[2i\omega(\Phi_j \otimes \theta_j) + \sum_{l=1}^3 (\Phi_j^{x_l} \otimes G_{x_l}) \right. &+ \sum_{k \in \mathcal{V}_j^a} (\Phi_{jk} \otimes P_{jk}^a) \\
&+ \left. \sum_{k \in \mathcal{V}_j^m} (\Phi_{jk} \otimes P_{jk}^m) \right] \aleph(\mathbb{W}_j) \\
&+ \sum_{k \in \mathcal{V}_j^d} (\Phi_{jk} \otimes M_{jk})\aleph(\mathbb{W}_k) = \aleph(\mathbf{b}_j).
\end{aligned} \tag{3.100}$$

On voit donc que l'on obtient une formulation matricielle dont la matrice diagonale est constituée de d_j^2 blocs de taille 6×6 , et dont les matrices extradiagonales sont constituées de $d_j \times d_k$ blocs de taille 6×6 . On peut aussi ordonner les inconnues d'une autre manière, c'est-à-dire les ranger en vecteur où cette fois-ci ce sont les lignes de la matrice \mathbb{W} qui ont été empilées plutôt que ses colonnes. C'est alors un rangement par degrés de liberté successifs pour une même coordonnée, ce qui est en quelque sorte le rangement alternatif au précédent. En particulier, ce vecteur correspond à $\aleph(\mathbb{W}^t)$ et le schéma relatif à cet empilement s'écrit :

$$\begin{aligned}
\left[2i\omega(\theta_j \otimes \Phi_j) + \sum_{l=1}^3 (G_{x_l} \otimes \Phi_j^{x_l}) \right. &+ \sum_{k \in \mathcal{V}_j^a} (P_{jk}^a \otimes \Phi_{jk}) \\
&+ \left. \sum_{k \in \mathcal{V}_j^m} (P_{jk}^m \otimes \Phi_{jk}) \right] \aleph(\mathbb{W}_j^t) \\
&+ \sum_{k \in \mathcal{V}_j^d} (M_{jk} \otimes \Phi_{jk})\aleph(\mathbb{W}_k^t) = \mathbf{b}_j.
\end{aligned}$$

La matrice diagonale est ici formée de 6×6 blocs de taille d_j^2 et la matrice extradiagonale est formée de 6×6 blocs de taille $d_j \times d_k$. Nous préférons utiliser la première formulation qui utilise un rangement

par coordonnées puis par degré de liberté local. En posant $\mathbf{W}_j = \aleph(\mathbb{W}_j)$, nous devons étudier le schéma :

$$\left[2i\omega(\Phi_j \otimes \theta_j) + \sum_{l=1}^3 (\Phi_j^{x_l} \otimes G_{x_l}) + \sum_{k \in \mathcal{V}_j^a} (\Phi_{jk} \otimes P_{jk}^a) + \sum_{k \in \mathcal{V}_j^m} (\Phi_{jk} \otimes P_{jk}^m) \right] \mathbf{W}_j + \sum_{k \in \mathcal{V}_j^d} (\Phi_{jk} \otimes M_{jk}) \mathbf{W}_k = \mathbf{b}_j. \quad (3.101)$$

2.3.3 Étude du problème discret

2.3.3.1 Formulation P_1

On se place maintenant dans le cas où $\mathcal{P}_j = \mathbb{P}_1[C_j]$ et où les cellules C_j sont des tétraèdres (i.e. $C_j = \tau_j$). Dans ce cas, on prend $\phi_j = (\varphi_{j1}, \varphi_{j2}, \varphi_{j3}, \varphi_{j4}) = (\lambda_{j1}, \lambda_{j2}, \lambda_{j3}, \lambda_{j4})$, où les λ_{jk} sont les coordonnées barycentriques relatives aux quatre sommets du tétraèdre τ_j . En utilisant la formule :

$$\int_{\hat{\tau}} \hat{\lambda}_1^{e_1} \hat{\lambda}_2^{e_2} \hat{\lambda}_3^{e_3} \hat{\lambda}_4^{e_4} d\hat{x} = \frac{e_1!e_2!e_3!e_4!}{(e_1 + e_2 + e_3 + e_4 + 3)!},$$

où $\hat{\tau}$ désigne le tétraèdre de référence de volume $\hat{V} = 1/6$ et les $(\hat{\lambda}_j)_{1 \leq j \leq 4}$ les coordonnées barycentriques associées à $\hat{\tau}$, on obtient :

$$\Phi_j = \int_{\tau_j} \phi_l^t \phi_m dx = \frac{V_j}{\hat{V}} \int_{\hat{\tau}} \hat{\phi}_l^t \hat{\phi}_m d\hat{x} = \frac{V_j}{120\hat{V}} \begin{pmatrix} 2 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 2 \end{pmatrix}, \quad (3.102)$$

où V_j désigne le volume du tétraèdre τ_j . De même :

$$\Phi_j^{x_k} = \int_{\tau_j} (\phi_j^t (\partial_{x_k} \phi_j) - (\partial_{x_k} \phi_j)^t \phi_j) dx = \left(\frac{n_{mk} - n_{lk}}{12} \right)_{lm},$$

où n_{lk} est la composante selon e_k de la normale intérieure à τ_j intégrée sur la face opposée au sommet numéro l dans τ_j . Par ailleurs :

$$\Phi_{jk} = \int_{\Sigma_{jk}} \phi_j^t \phi_k ds$$

est une matrice 4×4 dont le (lm) ^{ème} terme vaut $|\Sigma_{jk}| \alpha_{lm} / 12$, où α_{lm} dépend des degrés de liberté associés (si le sommet l dans τ_j n'est pas dans τ_k ou réciproquement, $\alpha_{lm} = 0$; s'ils sont égaux, $\alpha_{lm} = 2$; sinon $\alpha_{lm} = 1$). On a aussi :

$$\Phi_j \otimes \theta_j = \frac{V_j}{120\hat{V}} \begin{pmatrix} 2\theta_j & \theta_j & \theta_j & \theta_j \\ \theta_j & 2\theta_j & \theta_j & \theta_j \\ \theta_j & \theta_j & 2\theta_j & \theta_j \\ \theta_j & \theta_j & \theta_j & 2\theta_j \end{pmatrix} \text{ avec } \theta_j = \begin{pmatrix} \varepsilon_j |_3 & 0_3 \\ 0_3 & \mu_j |_3 \end{pmatrix}. \quad (3.103)$$

La matrice $\sum_{l=1}^3 \Phi_j^{x_l} \otimes G_{x_l}$ est explicitée dans l'annexe 2.4.7.3.

2.3.3.2 Étude de la dispersion numérique

Tout comme pour le schéma VF (GD-P0), nous allons procéder à l'étude de dispersion du schéma GD-P1 sur une grille cartésienne (et non pas tétraédrique). La différence avec l'étude concernant le schéma GD-P0 est que nous allons être contraints d'étudier un problème aux valeurs propres afin de trouver la relation de dispersion du schéma GD-P1. Pour ce faire, nous considérerons les mêmes hypothèses que pour l'étude du schéma GD-P0 à savoir : une grille régulière homogène dont les cellules sont de taille $\Delta x \times \Delta y \times \Delta z$, et une cellule $C_{j,l,m}$ sur laquelle nous allons appliquer notre schéma et dont le centre de gravité G a pour coordonnées $(j\Delta x, l\Delta y, m\Delta z)$. Pour appliquer des fonctions \mathbb{P}_1 sur un maillage hexaédrique orthogonal, nous considérons les fonctions de base suivantes :

$$\left\{ \begin{array}{lll} \varphi_1 = (1, 0, 0)^t, & \varphi_2 = (0, 1, 0)^t, & \varphi_3 = (0, 0, 1)^t, \\ \varphi_4 = \left(\frac{x - x_G}{\Delta x}, 0, 0\right)^t, & \varphi_5 = \left(0, \frac{x - x_G}{\Delta x}, 0\right)^t, & \varphi_6 = \left(0, 0, \frac{x - x_G}{\Delta x}\right)^t, \\ \varphi_7 = \left(\frac{y - y_G}{\Delta y}, 0, 0\right)^t, & \varphi_8 = \left(0, \frac{y - y_G}{\Delta y}, 0\right)^t, & \varphi_9 = \left(0, 0, \frac{y - y_G}{\Delta y}\right)^t, \\ \varphi_{10} = \left(\frac{z - z_G}{\Delta z}, 0, 0\right)^t, & \varphi_{11} = \left(0, \frac{z - z_G}{\Delta z}, 0\right)^t, & \varphi_{12} = \left(0, 0, \frac{z - z_G}{\Delta z}\right)^t, \end{array} \right.$$

De ce fait, tout champ $U_{j,l,m}$ désignant indifféremment $E_{j,l,m}$ ou $H_{j,l,m}$ sur une cellule $C_{j,l,m}$ du maillage s'écrit :

$$U_{j,l,m} = \begin{pmatrix} U_{j,l,m}^1 + U_{j,l,m}^4 \left(\frac{x - x_G}{\Delta x}\right) + U_{j,l,m}^7 \left(\frac{y - y_G}{\Delta y}\right) + U_{j,l,m}^{10} \left(\frac{z - z_G}{\Delta z}\right) \\ U_{j,l,m}^2 + U_{j,l,m}^5 \left(\frac{x - x_G}{\Delta x}\right) + U_{j,l,m}^8 \left(\frac{y - y_G}{\Delta y}\right) + U_{j,l,m}^{11} \left(\frac{z - z_G}{\Delta z}\right) \\ U_{j,l,m}^3 + U_{j,l,m}^6 \left(\frac{x - x_G}{\Delta x}\right) + U_{j,l,m}^9 \left(\frac{y - y_G}{\Delta y}\right) + U_{j,l,m}^{12} \left(\frac{z - z_G}{\Delta z}\right) \end{pmatrix}.$$

Nous partons donc de l'équation (3.93) en exploitant les hypothèses précédentes :

$$\begin{aligned} i\omega \int_{C_{i,j,k}} \varphi \mathbf{W}_{j,l,m} dx &+ \frac{1}{2}\theta^{-1} \int_{C_{j,l,m}} \left(\varphi \left(\sum_{r=1}^3 \partial_{x_r} G_{x_r} \mathbf{W}_{j,l,m} \right) - \sum_{r=1}^3 \partial_{x_r} \varphi G_{x_r} \mathbf{W}_{j,l,m} \right) dx \\ &+ \frac{1}{2}\theta^{-1} \left(\int_{\Sigma_{i+1}} \varphi M_{j+1,l,m} \mathbf{W}_{j+1,l,m} ds + \right. \\ &\quad \int_{\Sigma_{i-1}} \varphi M_{j-1,l,m} \mathbf{W}_{j-1,l,m} ds + \\ &\quad \int_{\Sigma_{j+1}} \varphi M_{j,l+1,m} \mathbf{W}_{j,l+1,m} ds + \\ &\quad \int_{\Sigma_{j-1}} \varphi M_{j,l-1,m} \mathbf{W}_{j,l-1,m} ds + \\ &\quad \int_{\Sigma_{l+1}} \varphi M_{j,l,m+1} \mathbf{W}_{j,l,m+1} ds + \\ &\quad \left. \int_{\Sigma_{l-1}} \varphi M_{j,l,m-1} \mathbf{W}_{j,l,m-1} ds \right) = 0, \end{aligned}$$

qui donne les équations suivantes :

$$\begin{aligned}
i\omega \int_{C_{i,j,k}} \varphi \mathbf{E}_{j,l,m} dx &+ \frac{1}{2\mu} \int_{C_{j,l,m}} \left(\varphi \left(\sum_{r=1}^3 \partial_{x_r} N^{x_r} \mathbf{H}_{j,l,m} \right) - \sum_{r=1}^3 \partial_{x_r} \varphi N^{x_r} \mathbf{H}_{j,l,m} \right) dx \\
&+ \frac{1}{2\mu} \left(\int_{\Sigma_{i+1}} \varphi N_{j+1,l,m} \mathbf{H}_{j+1,l,m} ds + \right. \\
&\quad \int_{\Sigma_{i-1}} \varphi N_{j-1,l,m} \mathbf{H}_{j-1,l,m} ds + \\
&\quad \int_{\Sigma_{j+1}} \varphi N_{j,l+1,m} \mathbf{H}_{j,l+1,m} ds + \\
&\quad \int_{\Sigma_{j-1}} \varphi N_{j,l-1,m} \mathbf{H}_{j,l-1,m} ds + \\
&\quad \int_{\Sigma_{l+1}} \varphi N_{j,l,m+1} \mathbf{H}_{j,l,m+1} ds + \\
&\quad \left. \int_{\Sigma_{l-1}} \varphi N_{j,l,m-1} \mathbf{H}_{j,l,m-1} ds \right) = 0.
\end{aligned}$$

On remplace alors les champs du précédent système par leurs solutions particulières qui sont de la forme $\mathbf{U}_{j,l,m} = \sum_{p=1}^{12} U_0^p e^{-i\mathbf{k} \cdot (i\Delta x, j\Delta y, l\Delta z)} \varphi_p$. Si on note $\tilde{\mathbf{U}} = (U_0^1, U_0^2, \dots, U_0^{12})^t$ on obtient le système d'équations :

$$\begin{cases} i\omega \tilde{\mathbf{E}} = \mu^{-1} \tilde{\mathbf{T}} \tilde{\mathbf{H}} \\ i\omega \tilde{\mathbf{H}} = -\varepsilon^{-1} \tilde{\mathbf{T}} \tilde{\mathbf{E}} \end{cases} \Rightarrow \omega^2 \tilde{\mathbf{E}} = c^2 \tilde{\mathbf{T}}^2 \tilde{\mathbf{E}}$$

avec :

$$\tilde{\mathbf{T}} = \left(\begin{array}{ccc|ccc|ccc|ccc}
0 & -s_z & s_y & 0 & 0 & 0 & 0 & 0 & r_y & 0 & -r_z & 0 \\
s_z & 0 & -s_x & 0 & 0 & -r_x & 0 & 0 & 0 & r_z & 0 & 0 \\
-s_y & s_x & 0 & 0 & r_x & 0 & -r_y & 0 & 0 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & -s_z & s_y & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 12r_x & s_z & 0 & 3s_x & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -12r_x & 0 & -s_y & -3s_x & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
0 & 0 & -12r_y & 0 & 0 & 0 & 0 & -s_z & -3s_y & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & s_z & 0 & -s_x & 0 & 0 & 0 \\
12r_y & 0 & 0 & 0 & 0 & 0 & 3s_y & s_x & 0 & 0 & 0 & 0 \\
\hline
0 & 12r_z & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3s_z & s_y \\
-12r_z & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -3s_z & 0 & -s_x \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -s_y & s_x & 0
\end{array} \right),$$

où on a posé, pour $v \in \{x, y, z\}$:

$$\begin{cases} s_v = \frac{i \sin(k_v \Delta v)}{\Delta v}, \\ r_v = \frac{1 - \cos(k_v \Delta v)}{2\Delta v}. \end{cases}$$

On est donc amené à trouver les valeurs propres λ de la matrice \tilde{T} . Ensuite, ces valeurs propres vérifient :

$$\omega^2 = c^2 \lambda^2, \quad (3.104)$$

ce qui donne une relation de dispersion. On introduit ensuite un paramètre h qui nous permet de poser :

$$\begin{cases} \Delta x = \alpha_x h, \\ \Delta y = \alpha_y h, \\ \Delta z = \alpha_z h, \end{cases}$$

et on suppose que les paramètres α_x , α_y et α_z restent bornés. Les valeurs propres de \tilde{T} sont recherchées sous la forme :

$$\lambda(\tilde{T}) = |\mathbf{k}|(1 + ah + bh^2) \text{ avec } (a, b) \in \mathbf{R}^2.$$

Le but est de trouver a et b pour connaître un développement limité de λ quand $h \rightarrow 0$, et donc une relation de dispersion asymptotique quand $h \rightarrow 0$. Tout d'abord, on effectue un développement limité des coefficients de la matrice \tilde{T} par rapport à h :

$$\begin{cases} s_v = ik_v \left(1 - \frac{k_v^2 (\alpha_v h)^2}{6} + O(k_v^4 h^4) \right) \\ r_v = \frac{k_v^2 (\alpha_v h)}{4} + O(k_v^4 h^4) \end{cases}$$

Maintenant, nous recherchons les valeurs de a et de b pour lesquelles le polynôme caractéristique de \tilde{T} s'annule pour tout h . λ sera de la forme :

$$\lambda = |k| \left(1 + \frac{1}{48} \frac{k_x^4 \alpha_x^2 + k_y^4 \alpha_y^2 + k_z^4 \alpha_z^2}{|k|^2} h^2 \right). \quad (3.105)$$

Ainsi, en injectant cette valeur propre dans l'équation (3.104), nous obtenons la relation de dispersion discrète :

$$\frac{\omega^2}{\omega_{ex}^2} = 1 + \frac{k_x^4 \Delta x^2 + k_y^4 \Delta y^2 + k_z^4 \Delta z^2}{24|k|^2} + O(|k|^4 h^4), \quad (3.106)$$

qui reste une dispersion d'ordre deux, huit fois moindre que celle qui caractérise le schéma volumes finis (3.84). Pour connaître par exemple la relation de dispersion en 1D, il suffit de prendre une onde qui se propage dans la direction $\mathbf{k} = |\mathbf{k}|(1, 0, 0)^t$, ce qui donne :

$$\frac{\omega^2}{\omega_{ex}^2} = 1 + \frac{|k|^2 \Delta x^2}{24} + O(|k|^4 h^4). \quad (3.107)$$

2.4 Annexes

2.4.1 Équivalences des conditions aux limites 1D

Proposition 4 *Pour imposer faiblement la condition au bord absorbant 1D, les deux égalités suivantes sont équivalentes :*

$$(i) E_{\frac{1}{2}} - z_1 H_{\frac{1}{2}} = \xi(E^{inc} - z_1 H^{inc}),$$

$$(ii) A_1 \frac{W_0 + W_1}{2} = \xi A_1 W^{inc},$$

$$\text{où } A_1 = \begin{pmatrix} 1 & -z_1 \\ -z_1^{-1} & 1 \end{pmatrix}.$$

Preuve On part dans un premier temps de (i) que l'on multiplie par $-z_1^{-1}$ de manière à obtenir :

$$\begin{cases} E_{\frac{1}{2}} - z_1 H_{\frac{1}{2}} = \xi(E^{inc} - z_1 H^{inc}) \\ -z_1^{-1} E_{\frac{1}{2}} + H_{\frac{1}{2}} = \xi(-z_1^{-1} E^{inc} + H^{inc}) \end{cases}$$

$$\Leftrightarrow \begin{pmatrix} 1 & -z_1 \\ -z_1^{-1} & 1 \end{pmatrix} \begin{pmatrix} E_{\frac{1}{2}} \\ H_{\frac{1}{2}} \end{pmatrix} = \xi \begin{pmatrix} 1 & -z_1 \\ -z_1^{-1} & 1 \end{pmatrix} \begin{pmatrix} E^{inc} \\ H^{inc} \end{pmatrix},$$

ce qui correspond à :

$$A_1 \frac{W_0 + W_1}{2} = \xi A_1 W^{inc}.$$

■

Proposition 5 Pour imposer faiblement la condition au bord métallique 1D, les deux égalités suivantes sont équivalentes :

$$(i) E_{N+\frac{1}{2}} = (\xi - 1)E^{inc},$$

$$(ii) A_N \frac{W_N + W_{N+1}}{2} = (\xi - 1)A_N W^{inc},$$

$$\text{où } A_N = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Preuve Si on utilise le fait que $E_{N+\frac{1}{2}} = \frac{E_N + E_{N+1}}{2}$, on a en partant de (i) :

$$E_{N+\frac{1}{2}} = (\xi - 1)E^{inc} \Leftrightarrow \frac{E_N + E_{N+1}}{2} = (\xi - 1)E^{inc}$$

$$\Leftrightarrow \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{E_N + E_{N+1}}{2} \\ \frac{H_N + H_{N+1}}{2} \end{pmatrix} = (\xi - 1) \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} E^{inc} \\ H^{inc} \end{pmatrix},$$

ce qui exprime exactement :

$$A_N \frac{W_N + W_{N+1}}{2} = (\xi - 1)A_N W^{inc}.$$

■

2.4.2 Propriétés de la matrice N_{jk}

Proposition 6 La matrice N_{jk} possède les propriétés suivantes :

(i) N_{jk} est antisymétrique et $N_{jk} = -N_{kj}$.

(ii) $\sum_{k \in \mathcal{V}_j} S_{jk} N_{jk} = 0$.

(iii) Soit $\mathbf{X} \in \mathbb{C}^3$. Si on note respectivement \mathbf{X}_\perp et \mathbf{X}_\parallel les parties normales et tangentielles de \mathbf{X} relativement à la surface Σ_{jk} :

$$\mathbf{X}_\perp = (\mathbf{n}_{jk} \mathbf{n}_{jk}^t) \mathbf{X} = (\mathbf{n}_{jk} \cdot \mathbf{X}) \mathbf{n}_{jk} \quad \text{et} \quad \mathbf{X}_\parallel = \mathbf{n}_{jk} \times (\mathbf{X} \times \mathbf{n}_{jk}) = \mathbf{X} - (\mathbf{n}_{jk} \cdot \mathbf{X}) \mathbf{n}_{jk},$$

alors on a que $\mathbf{X}_\parallel = -N_{jk}^2 \mathbf{X}$ et donc que $\mathbf{X}_\perp = (N_{jk}^2 + I_3) \mathbf{X}$.

(iv) $N_{jk}^3 = -N_{jk}$.

(v) N_{jk}^2 est symétrique et négative et elle admet les valeurs propres $\{0, -1, -1\}$.

Preuve

(i) Tout d'abord, $N_{kj} \mathbf{X} = \mathbf{X} \times \mathbf{n}_{kj} = -\mathbf{X} \times \mathbf{n}_{jk} = -N_{jk} \mathbf{X} \Rightarrow N_{kj} = -N_{jk}$.

(ii) En faisant les calculs on obtient :

$$\begin{aligned} \sum_{k \in \mathcal{V}_j} S_{jk} \mathbf{n}_{jk} = 0 &\Rightarrow \sum_{k \in \mathcal{V}_j} S_{jk} N_{jk} \mathbf{X} = -\left(\sum_{k \in \mathcal{V}_j} S_{jk} \mathbf{n}_{jk} \right) \times \mathbf{X} = 0 \\ &\Rightarrow \sum_{k \in \mathcal{V}_j} S_{jk} N_{jk} = 0. \end{aligned}$$

(iv) En faisant les calculs on obtient :

$$\begin{aligned} N_{jk}^3 \mathbf{X} &= ((\mathbf{X} \times \mathbf{n}_{jk}) \times \mathbf{n}_{jk}) \times \mathbf{n}_{jk} \\ &= -(\mathbf{n}_{jk} \times (\mathbf{X} \times \mathbf{n}_{jk})) \times \mathbf{n}_{jk} \\ &= -(\mathbf{X} - (\mathbf{n}_{jk} \cdot \mathbf{X}) \mathbf{n}_{jk}) \times \mathbf{n}_{jk} \\ &= -\mathbf{X} \times \mathbf{n}_{jk} \\ &= -N_{jk} \mathbf{X} \\ \Rightarrow N_{jk}^3 &= -N_{jk}. \end{aligned}$$

■

2.4.3 Décomposition de la matrice des flux

On cherche dans ce paragraphe à décomposer la matrice M_{jk}^j en sa partie positive et négative. Nous avons déjà vu que :

$$M_{jk}^j = M_{jk}\theta_j^{-1} \text{ avec } M_{jk} = \begin{pmatrix} 0_3 & N_{jk} \\ -N_{jk} & 0_3 \end{pmatrix} \text{ et } \theta_j = \begin{pmatrix} \varepsilon_j \mathbf{l}_3 & 0_3 \\ 0_3 & \mu_j \mathbf{l}_3 \end{pmatrix}.$$

L'hyperbolicité du système d'équations de Maxwell signifie que M_{jk}^j est diagonalisable. On montre que M_{jk}^j admet les valeurs propres réelles doubles $\{\lambda_0 = 0, \lambda_+ = c_j, \lambda_- = -c_j\}$. Nous allons d'abord trouver une paire de vecteurs propres orthonormaux associés à la valeur propre λ_0 . Si on choisit la paire de vecteurs $(\mathbf{W}_0^a, \mathbf{W}_0^b) \in \mathbf{R}^{6 \times 2}$ telle que $\mathbf{W}_0^a = (\mathbf{n}_{jk}, 0_3)^t$ et $\mathbf{W}_0^b = (0_3, \mathbf{n}_{jk})^t$, on s'aperçoit qu'ils sont orthonormaux entre eux, et qu'ils correspondent bien à des vecteurs propres de M_{jk}^j associés à λ_0 . En effet :

$$\begin{cases} M_{jk}^j \mathbf{W}_0^a = M_{jk}\theta_j^{-1} \mathbf{W}_0^a = \begin{pmatrix} 0_3 & N_{jk} \\ -N_{jk} & 0_3 \end{pmatrix} \begin{pmatrix} \varepsilon_j^{-1} \mathbf{n}_{jk} \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ -\varepsilon_j^{-1} N_{jk} \mathbf{n}_{jk} \end{pmatrix} = 0 \\ M_{jk}^j \mathbf{W}_0^b = M_{jk}\theta_j^{-1} \mathbf{W}_0^b = \begin{pmatrix} 0_3 & N_{jk} \\ -N_{jk} & 0_3 \end{pmatrix} \begin{pmatrix} 0 \\ \mu_j^{-1} \mathbf{n}_{jk} \end{pmatrix} = \begin{pmatrix} -\mu_j^{-1} N_{jk} \mathbf{n}_{jk} \\ 0 \end{pmatrix} = 0. \end{cases}$$

On peut par conséquent définir le projecteur orthogonal Π_{jk}^0 sur le sous-espace propre $\text{Ker}(M_{jk}^j)$ par :

$$\Pi_{jk}^0 = \mathbf{W}_0^a (\mathbf{W}_0^a)^t + \mathbf{W}_0^b (\mathbf{W}_0^b)^t = \begin{pmatrix} \mathbf{n}_{jk} \mathbf{n}_{jk}^t & 0_3 \\ 0_3 & \mathbf{n}_{jk} \mathbf{n}_{jk}^t \end{pmatrix} = \begin{pmatrix} N_{jk}^2 + \mathbf{l}_3 & 0_3 \\ 0_3 & N_{jk}^2 + \mathbf{l}_3 \end{pmatrix}.$$

On note Π_{jk}^+ et Π_{jk}^- les projecteurs orthogonaux sur les sous-espaces propres respectifs $\text{Ker}((M_{jk}^j - c_j \mathbf{l}_6))$ et $\text{Ker}((M_{jk}^j + c_j \mathbf{l}_6))$. En raison de la diagonalisabilité de M_{jk}^j , on a :

$$\begin{cases} \mathbf{l}_6 & = \Pi_{jk}^0 + \Pi_{jk}^+ + \Pi_{jk}^- \\ M_{jk}^j & = \lambda_0 \Pi_{jk}^0 + \lambda_+ \Pi_{jk}^+ + \lambda_- \Pi_{jk}^- \end{cases} \Leftrightarrow \begin{cases} \mathbf{l}_6 & = \Pi_{jk}^0 + \Pi_{jk}^+ + \Pi_{jk}^- \\ M_{jk}^j & = c_j \Pi_{jk}^+ - c_j \Pi_{jk}^- \end{cases}$$

$$\begin{aligned} & \Updownarrow \\ & \begin{cases} \Pi_{jk}^+ & = \frac{1}{2}(\mathbf{l}_6 - \Pi_{jk}^0 + c_j^{-1} M_{jk}^j) \\ \Pi_{jk}^- & = \frac{1}{2}(\mathbf{l}_6 - \Pi_{jk}^0 - c_j^{-1} M_{jk}^j). \end{cases} \end{aligned}$$

On se sert maintenant de la matrice d'impédance Z_j si bien que l'on a :

$$c_j^{-1} \theta_j^{-1} = \begin{pmatrix} (c_j \varepsilon_j)^{-1} \mathbf{l}_3 & 0_3 \\ 0_3 & (c_j \mu_j)^{-1} \mathbf{l}_3 \end{pmatrix} = Z_j^{-1},$$

et de ce fait :

$$\begin{aligned}
\Pi_{jk}^+ &= \frac{1}{2}(\mathbf{l}_6 - \Pi_{jk}^0 + c_j^{-1}M_{jk}\theta_j^{-1}) \\
&= \frac{1}{2} \left[\begin{pmatrix} \mathbf{l}_3 & \mathbf{0}_3 \\ \mathbf{0}_3 & \mathbf{l}_3 \end{pmatrix} - \begin{pmatrix} N_{jk}^2 + \mathbf{l}_3 & \mathbf{0}_3 \\ \mathbf{0}_3 & N_{jk}^2 + \mathbf{l}_3 \end{pmatrix} + M_{jk}Z_j^{-1} \right] \\
&= \frac{1}{2} \left[\begin{pmatrix} -N_{jk}^2 & \mathbf{0}_3 \\ \mathbf{0}_3 & -N_{jk}^2 \end{pmatrix} + \begin{pmatrix} \mathbf{0}_3 & N_{jk} \\ -N_{jk} & \mathbf{0}_3 \end{pmatrix} \begin{pmatrix} z_j \mathbf{l}_3 & \mathbf{0}_3 \\ \mathbf{0}_3 & z_j^{-1} \mathbf{l}_3 \end{pmatrix} \right] \\
&= \frac{1}{2} \begin{pmatrix} -N_{jk}^2 & z_j^{-1}N_{jk} \\ -z_j N_{jk} & -N_{jk}^2 \end{pmatrix},
\end{aligned}$$

et :

$$\begin{aligned}
\Pi_{jk}^- &= \frac{1}{2}(\mathbf{l}_6 - \Pi_{jk}^0 - c_j^{-1}M_{jk}\theta_j^{-1}) \\
&= \frac{1}{2} \left[\begin{pmatrix} \mathbf{l}_3 & \mathbf{0}_3 \\ \mathbf{0}_3 & \mathbf{l}_3 \end{pmatrix} - \begin{pmatrix} N_{jk}^2 + \mathbf{l}_3 & \mathbf{0}_3 \\ \mathbf{0}_3 & N_{jk}^2 + \mathbf{l}_3 \end{pmatrix} - M_{jk}Z_j^{-1} \right] \\
&= \frac{1}{2} \left[\begin{pmatrix} -N_{jk}^2 & \mathbf{0}_3 \\ \mathbf{0}_3 & -N_{jk}^2 \end{pmatrix} + \begin{pmatrix} \mathbf{0}_3 & -N_{jk} \\ N_{jk} & \mathbf{0}_3 \end{pmatrix} \begin{pmatrix} z_j \mathbf{l}_3 & \mathbf{0}_3 \\ \mathbf{0}_3 & z_j^{-1} \mathbf{l}_3 \end{pmatrix} \right] \\
&= \frac{1}{2} \begin{pmatrix} -N_{jk}^2 & -z_j^{-1}N_{jk} \\ z_j N_{jk} & -N_{jk}^2 \end{pmatrix},
\end{aligned}$$

et par suite : $\Pi_{jk}^+ = \frac{1}{2} \begin{pmatrix} -N_{jk}^2 & z_j^{-1}N_{jk} \\ -z_j N_{jk} & -N_{jk}^2 \end{pmatrix}$ et $\Pi_{jk}^- = \frac{1}{2} \begin{pmatrix} -N_{jk}^2 & -z_j^{-1}N_{jk} \\ z_j N_{jk} & -N_{jk}^2 \end{pmatrix}$.

Maintenant, on utilise l'expression de ces projecteurs pour en déduire les matrices $(M_{jk}^j)^\pm$ avec, par définition $M_{jk}^j = (M_{jk}^j)^+ + (M_{jk}^j)^-$ et vu que $M_{jk}^j = c_j \Pi_{jk}^+ - c_j \Pi_{jk}^-$, alors $(M_{jk}^j)^+ = c_j \Pi_{jk}^+$ et $(M_{jk}^j)^- = -c_j \Pi_{jk}^-$. Donc :

$$(M_{jk}^j)^+ = \frac{c_j}{2} \begin{pmatrix} -N_{jk}^2 & z_j^{-1}N_{jk} \\ -z_j N_{jk} & -N_{jk}^2 \end{pmatrix} \text{ et } (M_{jk}^j)^- = \frac{c_j}{2} \begin{pmatrix} N_{jk}^2 & z_j^{-1}N_{jk} \\ -z_j N_{jk} & N_{jk}^2 \end{pmatrix}.$$

On notera par ailleurs que $(M_{jk}^j)^\pm = M_{jk}^\pm \theta_j^{-14}$ avec :

$$M_{jk}^+ = \frac{1}{2} \begin{pmatrix} -z_j^{-1}N_{jk}^2 & N_{jk} \\ -N_{jk} & -z_j N_{jk}^2 \end{pmatrix} \text{ et } M_{jk}^- = \frac{1}{2} \begin{pmatrix} z_j^{-1}N_{jk}^2 & N_{jk} \\ -N_{jk} & z_j N_{jk}^2 \end{pmatrix}.$$

et $M_{jk} = M_{jk}^+ + M_{jk}^-$.

2.4.4 Équivalences des conditions aux limites 3D

Proposition 7 *Pour imposer faiblement la condition aux bords absorbants 3D, les deux égalités suivantes sont équivalentes pour $k \in \mathcal{V}_j^a$:*

⁴ Il s'agit ici d'une relation purement algébrique plus précisément, les matrices M_{jk}^\pm ne résultent pas de la diagonalisation de M_{jk} .

$$(i) \mathbf{n}_{jk} \times \mathbf{E}_{|\Sigma_{jk}} + z_j \mathbf{n}_{jk} \times (\mathbf{n}_{jk} \times \mathbf{H}_{|\Sigma_{jk}}) = \xi[\mathbf{n}_{jk} \times \mathbf{E}^{inc} + z_j \mathbf{n}_{jk} \times (\mathbf{n}_{jk} \times \mathbf{H}^{inc})],$$

$$(ii) (M_{jk}^j)^{-\theta_j} \left(\frac{\mathbf{W}_j + \mathbf{W}_k}{2} \right) = \xi (M_{jk}^j)^{-\theta_0} \mathbf{W}^{inc}.$$

où $\theta_0 = \begin{pmatrix} \varepsilon_0 & 0 \\ 0 & \mu_0 \end{pmatrix}$ avec ε_0 et μ_0 les caractéristiques électromagnétiques du vide.

Preuve Il est important ici de travailler en variables conservatives. On introduit donc $\mathbf{Q}_{E,jk} \equiv \mathbf{Q}_{E,|\Sigma_{jk}} = \varepsilon_j \mathbf{E}_{|\Sigma_{jk}}$ et $\mathbf{Q}_{H,jk} \equiv \mathbf{Q}_{H,|\Sigma_{jk}} = \mu_j \mathbf{H}_{|\Sigma_{jk}}$. Notons que l'on a supposé $\varepsilon_k = \varepsilon_j$ et $\mu_k = \mu_j$. On commence tout d'abord par ré-écrire (i) à partir de ces variables. On obtient :

$$\mathbf{n}_{jk} \times \mathbf{Q}_{E,jk} + z_j^{-1} \mathbf{n}_{jk} \times (\mathbf{n}_{jk} \times \mathbf{Q}_{H,jk}) = \xi[\mathbf{n}_{jk} \times \mathbf{Q}_E^{inc} + z_j^{-1} \mathbf{n}_{jk} \times (\mathbf{n}_{jk} \times \mathbf{Q}_H^{inc})] \quad (4.108)$$

On considère deux équations : la première est donnée par $\mathbf{n}_{jk} \times (4.108)$ et la seconde par $z_j \cdot (4.108)$. On obtient :

$$\begin{cases} \mathbf{n}_{jk} \times [\mathbf{n}_{jk} \times \mathbf{Q}_{E,jk} + z_j^{-1} \mathbf{n}_{jk} \times (\mathbf{n}_{jk} \times \mathbf{Q}_{H,jk})] = \xi \mathbf{n}_{jk} \times [\mathbf{n}_{jk} \times \mathbf{Q}_E^{inc} + z_j^{-1} \mathbf{n}_{jk} \times (\mathbf{n}_{jk} \times \mathbf{Q}_H^{inc})] \\ z_j (\mathbf{n}_{jk} \times \mathbf{Q}_{E,jk}) + \mathbf{n}_{jk} \times (\mathbf{n}_{jk} \times \mathbf{Q}_{H,jk}) = \xi [z_j (\mathbf{n}_{jk} \times \mathbf{Q}_E^{inc}) + \mathbf{n}_{jk} \times (\mathbf{n}_{jk} \times \mathbf{Q}_H^{inc})] \end{cases}$$

$$\Leftrightarrow \begin{cases} \mathbf{n}_{jk} \times (\mathbf{n}_{jk} \times \mathbf{Q}_{E,jk}) - z_j^{-1} \mathbf{n}_{jk} \times \mathbf{Q}_{H,jk} = \xi [\mathbf{n}_{jk} \times (\mathbf{n}_{jk} \times \mathbf{Q}_E^{inc}) - z_j^{-1} \mathbf{n}_{jk} \times \mathbf{Q}_H^{inc}] \\ z_j (\mathbf{n}_{jk} \times \mathbf{Q}_{E,jk}) + \mathbf{n}_{jk} \times (\mathbf{n}_{jk} \times \mathbf{Q}_{H,jk}) = \xi [z_j (\mathbf{n}_{jk} \times \mathbf{Q}_E^{inc}) + \mathbf{n}_{jk} \times (\mathbf{n}_{jk} \times \mathbf{Q}_H^{inc})]. \end{cases}$$

En utilisant cette fois-ci les notations matricielles, ce système donne :

$$\begin{cases} N_{jk}^2 \left(\frac{\mathbf{Q}_{E,j} + \mathbf{Q}_{E,k}}{2} \right) + z_j^{-1} N_{jk} \left(\frac{\mathbf{Q}_{H,j} + \mathbf{Q}_{H,k}}{2} \right) = \xi (N_{jk}^2 \mathbf{Q}_E^{inc} + z_j^{-1} N_{jk} \mathbf{Q}_H^{inc}) \\ -z_j N_{jk} \left(\frac{\mathbf{Q}_{E,j} + \mathbf{Q}_{E,k}}{2} \right) + N_{jk}^2 \left(\frac{\mathbf{Q}_{H,j} + \mathbf{Q}_{H,k}}{2} \right) = \xi (-z_j N_{jk} \mathbf{Q}_E^{inc} + N_{jk}^2 \mathbf{Q}_H^{inc}) \end{cases}$$

$$\Leftrightarrow \begin{pmatrix} N_{jk}^2 & z_j^{-1} N_{jk} \\ -z_j N_{jk} & N_{jk}^2 \end{pmatrix} \begin{pmatrix} \frac{\mathbf{Q}_{E,j} + \mathbf{Q}_{E,k}}{2} \\ \frac{\mathbf{Q}_{H,j} + \mathbf{Q}_{H,k}}{2} \end{pmatrix} = \xi \begin{pmatrix} N_{jk}^2 & z_j^{-1} N_{jk} \\ -z_j N_{jk} & N_{jk}^2 \end{pmatrix} \begin{pmatrix} \mathbf{Q}_E^{inc} \\ \mathbf{Q}_H^{inc} \end{pmatrix}.$$

Il se trouve que cette égalité correspond à :

$$(M_{jk}^j)^{-\theta_j} \left(\frac{\mathbf{W}_j + \mathbf{W}_k}{2} \right) = \xi (M_{jk}^j)^{-\theta_0} \mathbf{W}^{inc}.$$

■

Proposition 8 Pour exprimer faiblement la condition aux bords métalliques 3D, les deux égalités sont équivalentes pour $k \in \mathcal{V}_j^m$:

$$(i) \mathbf{n}_{jk} \times (\mathbf{E}_{|\Sigma_{jk}}) = (\xi - 1) \mathbf{n}_{jk} \times (\mathbf{E}^{inc})$$

$$(ii) \Upsilon_{jk} \frac{\mathbf{W}_j + \mathbf{W}_k}{2} = (\xi - 1) \Upsilon_{jk} \mathbf{W}^{inc},$$

$$\text{où } \Upsilon_{jk} = \begin{pmatrix} -N_{jk} & 0_3 \\ 0_3 & 0_3 \end{pmatrix}.$$

Preuve On part toujours de (i) en ajoutant une égalité nulle sur le champ $H_{|\Sigma_{jk}}$, ce qui donne l'égalité matricielle suivant :

$$\begin{pmatrix} -N_{jk} & 0_3 \\ 0_3 & 0_3 \end{pmatrix} \begin{pmatrix} \mathbf{E}_{|\Sigma_{jk}} \\ \mathbf{H}_{|\Sigma_{jk}} \end{pmatrix} = (\xi - 1) \begin{pmatrix} -N_{jk} & 0_3 \\ 0_3 & 0_3 \end{pmatrix} \begin{pmatrix} \mathbf{E}^{inc} \\ \mathbf{H}^{inc} \end{pmatrix}.$$

On se sert maintenant de la moyenne des champs pour exprimer ces derniers sur l'interface Σ_{jk} :

$$\begin{pmatrix} -N_{jk} & 0_3 \\ 0_3 & 0_3 \end{pmatrix} \begin{pmatrix} \frac{\mathbf{E}_j + \mathbf{E}_k}{2} \\ \frac{\mathbf{H}_j + \mathbf{H}_k}{2} \end{pmatrix} = (\xi - 1) \begin{pmatrix} -N_{jk} & 0_3 \\ 0_3 & 0_3 \end{pmatrix} \begin{pmatrix} \mathbf{E}^{inc} \\ \mathbf{H}^{inc} \end{pmatrix},$$

et aboutir à (ii) :

$$\Upsilon_{jk} \left(\frac{\mathbf{W}_j + \mathbf{W}_k}{2} \right) = (\xi - 1) \Upsilon_{jk} \mathbf{W}^{inc}.$$

■

2.4.5 Quelques propriétés des matrices antisymétriques

Ce paragraphe a pour but de fournir quelques propriétés des matrices antisymétriques, et de donner aussi certains critères permettant de prouver l'antisymétrie de matrices ayant une structure par blocs bien précise.

Proposition 9 Soit $(A, B) \in (\mathbf{M}_{m \times n}(\mathbf{R}))^2$.

Alors $(\forall (\mathbf{X}, \mathbf{Y}) \in \mathbf{R}^{2(m \times n)}, \mathbf{X}^t \mathbf{A} \mathbf{Y} = \mathbf{X}^t \mathbf{B} \mathbf{Y}) \Leftrightarrow A = B$.

Proposition 10 Soit $R \in \mathbf{M}_{m \times n}(\mathbf{R})$. R est antisymétrique $\Leftrightarrow (\forall \mathbf{X} \in \mathbf{R}^{m \times n}, \mathbf{X}^t R \mathbf{X} = 0)$.

Preuve \Rightarrow) Supposons R est antisymétrique et soit $\mathbf{X} \in \mathbf{R}^{m \times n}$. On a que :

$$\mathbf{X}^t R \mathbf{X} = (R^t \mathbf{X})^t \mathbf{X} = \mathbf{X}^t (R^t \mathbf{X}) = -\mathbf{X}^t R \mathbf{X},$$

ce qui prouve le résultat.

\Leftarrow) Supposons que $\forall \mathbf{X} \in \mathbf{R}^{m \times n} \mathbf{X}^t R \mathbf{X} = 0$. On a $\forall (\mathbf{X}, \mathbf{Y}) \in \mathbf{R}^{2(m \times n)}$:

$$(\mathbf{X} + \mathbf{Y})^t R (\mathbf{X} + \mathbf{Y}) = 0 \Rightarrow \mathbf{X}^t (R + R^t) \mathbf{Y} + \mathbf{X}^t R \mathbf{X} + \mathbf{Y}^t R \mathbf{Y} = 0 \Rightarrow \mathbf{X}^t (R + R^t) \mathbf{Y} = 0.$$

D'après la proposition 9, on en déduit que $R^t = -R$, et par conséquent que R est antisymétrique. ■

Proposition 11 Soit $M \in \mathbf{M}_{m \times n}(\mathbf{R})$ telle que $\forall (j, k) \in \llbracket 1; n \rrbracket$, $j \neq k$, (M_{jj}, M_{jk}) soient respectivement antisymétrique et symétrique réelles et $M_{kj} = -M_{jk}$. Alors M est antisymétrique.

Preuve Il nous suffit de montrer d'après la proposition 10 que $\forall \mathbf{X} \in \mathbf{R}^{m \times n}$ on ait $\mathbf{X}^t M \mathbf{X} = 0$. Si on pose alors $N = M^t$:

$$\begin{aligned} \mathbf{X}^t M \mathbf{X} &= \frac{1}{2} \mathbf{X}^t (M + N) \mathbf{X} \\ &= \frac{1}{2} \sum_{j=1}^n \sum_{k=1}^n \mathbf{X}_j^t (M_{jk} + N_{jk}) \mathbf{X}_j \\ &= \frac{1}{2} \sum_{j=1}^n \left(\mathbf{X}_j^t (M_{jj} + N_{jj}) \mathbf{X}_j + \sum_{\substack{k=1 \\ k \neq j}}^n \mathbf{X}_k^t (M_{jk} + N_{jk}) \mathbf{X}_k \right). \end{aligned}$$

Or, $N = M^t \Rightarrow N_{jj} = M_{jj}^t$ et $N_{jk} = M_{kj}^t$. Mais les caractères antisymétrique de M_{jj} et symétrique de M_{jk} impliquant respectivement $M_{jj}^t = -M_{jj}$ et $M_{kj}^t = M_{kj}$, nous obtenons donc :

$$\begin{aligned} \mathbf{X}^t M \mathbf{X} &= \frac{1}{2} \sum_{j=1}^n \left(\mathbf{X}_j^t (M_{jj} - M_{jj}) \mathbf{X}_j + \sum_{\substack{k=1 \\ k \neq j}}^n \mathbf{X}_k^t (M_{jk} + M_{kj}) \mathbf{X}_k \right) \\ &= \frac{1}{2} \sum_{j=1}^n \sum_{\substack{k=1 \\ k \neq j}}^n \mathbf{X}_j^t (M_{jk} + M_{kj}) \mathbf{X}_k. \end{aligned}$$

Enfin, $M_{kj} = -M_{jk} \Rightarrow \mathbf{X}^t M \mathbf{X} = 0$. ■

Proposition 12 Si A est antihermitienne (soit $A^* = -A$) alors $\forall \mathbf{X} \in \mathbf{C}^n$, $\mathbf{X}^* A \mathbf{X} \in i\mathbf{R}$.

Preuve On déduit facilement que :

$$\begin{aligned} 2\Re(\mathbf{X}^* A \mathbf{X}) &= \mathbf{X}^* A \mathbf{X} + \overline{\mathbf{X}^* A \mathbf{X}} \\ &= \mathbf{X}^* A \mathbf{X} + \overline{\mathbf{X}^t A^t \overline{\mathbf{X}}} \\ &= \mathbf{X}^* A \mathbf{X} + \mathbf{X}^* A^* \mathbf{X} \\ &= \mathbf{X}^* (A + A^*) \mathbf{X} = 0. \end{aligned}$$

Par conséquent, $\mathbf{X}^* A \mathbf{X} \in i\mathbf{R}$. ■

2.4.6 Blocs matriciels du schéma GD-P0

Proposition 13 Les blocs matriciels ont les propriétés suivantes :

(i) $P_{jk}^a = \begin{pmatrix} -z_j^{-1}N_{jk}^2 & 0_3 \\ 0_3 & -z_j N_{jk}^2 \end{pmatrix}$ est symétrique et positive.

(ii) $M_{jk} = \begin{pmatrix} 0_3 & N_{jk} \\ -N_{jk} & 0_3 \end{pmatrix}$ est symétrique.

(iii) $P_{jk}^m = \begin{pmatrix} 0_3 & N_{jk} \\ N_{jk} & 0_3 \end{pmatrix}$ est antisymétrique.

Preuve

(i) Avec la symétrie de N_{jk}^2 (voir la proposition 6) on a :

$$(P_{jk}^a)^t = \begin{pmatrix} -z_j^{-1}(N_{jk}^2)^t & 0_3 \\ 0_3 & -z_j(N_{jk}^2)^t \end{pmatrix} = \begin{pmatrix} -z_j^{-1}N_{jk}^2 & 0_3 \\ 0_3 & -z_j N_{jk}^2 \end{pmatrix} = P_{jk}^a.$$

De plus, $\forall X \in \mathbb{C}^{3 \times 2}$: $X^* P_{jk}^a X = -z_j^{-1} X_1^* N_{jk}^2 X_1 - z_j X_2^* N_{jk}^2 X_2$ et N_{jk}^2 est réelle symétrique (ou encore, hermitienne) négative donc :

$$X_k^* N_{jk}^2 X_k \leq 0 \text{ pour } k \in \{1; 2\} \Rightarrow X^* P_{jk}^a X \geq 0 \text{ (car } z_j > 0).$$

(ii) La matrice N_{jk} est antisymétrique (voir la proposition 6) et donc :

$$M_{jk}^t = \begin{pmatrix} 0_3 & -N_{jk}^t \\ N_{jk}^t & 0_3 \end{pmatrix} = \begin{pmatrix} 0_3 & N_{jk} \\ -N_{jk} & 0_3 \end{pmatrix} = M_{jk}.$$

(iii) De même, nous avons aussi en raison l'antisymétrie de N_{jk} :

$$(P_{jk}^m)^t = \begin{pmatrix} 0_3 & N_{jk}^t \\ N_{jk}^t & 0_3 \end{pmatrix} = \begin{pmatrix} 0_3 & -N_{jk} \\ -N_{jk} & 0_3 \end{pmatrix} = -P_{jk}^m.$$

■

2.4.7 Blocs matriciels du schéma GD-P1

2.4.7.1 Procédé de vectorisation

Proposition 14 Pour $(n, m, p, q) \in \mathbb{N}^4$, soient $(A, B, W) \in \mathbf{M}_{n,m}(\mathbb{C}) \times \mathbf{M}_{p,q}(\mathbb{C}) \times \mathbf{M}_{q,m}(\mathbb{C})$. Si $\mathbb{W} = (W_1, W_2, \dots, W_m)$ où $(W_i)_{1 \leq i \leq m} \subset \mathbb{C}^q$,

$$\mathfrak{N}(\mathbb{W}) = \begin{pmatrix} W_1 \\ W_2 \\ \vdots \\ W_m \end{pmatrix} \in \mathbb{C}^{qm} \text{ et } A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B & \dots & a_{1m}B \\ a_{21}B & a_{22}B & \dots & a_{2m}B \\ \vdots & \vdots & \dots & \vdots \\ a_{n1}B & a_{n2}B & \dots & a_{nm}B \end{pmatrix} \in \mathbf{M}_{np,mq}(\mathbb{C})$$

alors : $\mathfrak{N}(B\mathbb{W}A^t) = (A \otimes B)\mathfrak{N}(\mathbb{W})$

Preuve

$$\begin{aligned}
 (A \otimes B)\aleph(\mathbb{W}) &= \begin{pmatrix} a_{11}B & a_{12}B & \dots & a_{1m}B \\ a_{21}B & a_{22}B & \dots & a_{2m}B \\ \vdots & \vdots & & \vdots \\ a_{n1}B & a_{n2}B & \dots & a_{nm}B \end{pmatrix} \begin{pmatrix} W_1 \\ W_2 \\ \vdots \\ W_m \end{pmatrix} \\
 &= \begin{pmatrix} \sum_{j=1}^m a_{1j}BW_j \\ \sum_{j=1}^m a_{2j}BW_j \\ \vdots \\ \sum_{j=1}^m a_{nj}BW_j \end{pmatrix} \\
 &= \aleph\left(\sum_{j=1}^m a_{1j}BW_j, \sum_{j=1}^m a_{2j}BW_j, \dots, \sum_{j=1}^m a_{nj}BW_j\right)
 \end{aligned}$$

De plus, étant donné que $BW_j \in \mathbb{C}^p$, alors pour $i \in \llbracket 1; p \rrbracket$, $(BW_j)_i = \sum_{k=1}^q b_{ik}w_{kj} = (BW)_{ij}$.

Maintenant, prenons $l \in \llbracket 1; n \rrbracket$, pour chaque élément scalaire nous avons l'égalité suivante :

$$(B\mathbb{W}A^t)_{il} = \sum_{j=1}^m (BW)_{ij}a_{lj} = \sum_{j=1}^m (BW_j)_i a_{lj}$$

De ce fait, si on désigne par $(B\mathbb{W}A^t)_l$ la $l^{\text{ème}}$ colonne de la matrice $B\mathbb{W}A^t$ où $l \in \llbracket 1; n \rrbracket$, l'égalité scalaire précédente nous permet d'obtenir l'égalité vectorielle suivante :

$$\forall l \in \llbracket 1; n \rrbracket, (B\mathbb{W}A^t)_l = \sum_{j=1}^m BW_j a_{lj}$$

Ce qui implique :

$$\aleph\left(\sum_{j=1}^m a_{1j}BW_j, \sum_{j=1}^m a_{2j}BW_j, \dots, \sum_{j=1}^m a_{nj}BW_j\right) = \aleph((B\mathbb{W}A^t)_1, (B\mathbb{W}A^t)_2, \dots, (B\mathbb{W}A^t)_n)$$

et permet ainsi de conclure que :

$$\aleph(B\mathbb{W}A^t) = (A \otimes B)\aleph(\mathbb{W})$$

■

2.4.7.2 Propriétés du produit de Kronecker

Pour l'occasion, on note dans ce paragraphe :

- $\mathbf{GL}_n(\mathbb{C})$ l'ouvert des matrices inversibles de $\mathbf{M}_n(\mathbb{C})$,
- $\mathbf{S}_n(\mathbb{C})$ le sous-espace vectoriel des matrices symétriques de $\mathbf{M}_n(\mathbb{C})$,
- $\mathbf{A}_n(\mathbb{C})$ le sous-espace vectoriel des matrices antisymétriques de $\mathbf{M}_n(\mathbb{C})$.

Nous avons dès lors :

- $(A \otimes B)^t = A^t \otimes B^t$
- $(A \otimes B)(A' \otimes B') = (AA') \otimes (BB')$
- Si $(A, B) \in \mathbf{GL}_n(\mathbb{C}) \times \mathbf{GL}_m(\mathbb{C})$, alors $(A \otimes B) \in \mathbf{GL}_{np}(\mathbb{C})$ et on a :

$$(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$$

- $(A, B) \in \mathbf{S}_n(\mathbb{C}) \times \mathbf{S}_m(\mathbb{C})$ ou $(A, B) \in \mathbf{A}_n(\mathbb{C}) \times \mathbf{A}_m(\mathbb{C})$, alors $(A \otimes B) \in \mathbf{S}_{nm}(\mathbb{C})$
- $(A, B) \in \mathbf{S}_n(\mathbb{C}) \times \mathbf{A}_m(\mathbb{C})$ ou $(A, B) \in \mathbf{A}_n(\mathbb{C}) \times \mathbf{S}_m(\mathbb{C})$, alors $(A \otimes B) \in \mathbf{A}_{nm}(\mathbb{C})$

2.4.7.3 Matrice 3D des gradients des fonctions de base P1

Il s'agit de la matrice $P = \sum_{l=1}^3 \Phi_l^{x_l} \otimes G_{x_l}$ où :

$$P = \begin{pmatrix} 0_6 & P_{12} & P_{13} & P_{14} \\ P_{21} & 0_6 & P_{23} & P_{24} \\ P_{31} & P_{32} & 0_6 & P_{34} \\ P_{41} & P_{42} & P_{43} & 0_6 \end{pmatrix}$$

avec :

$$P_{12} = \begin{pmatrix} 0 & 0 & 0 & 0 & -\alpha_z & \alpha_y \\ 0 & 0 & 0 & \alpha_z & 0 & -\alpha_x \\ 0 & 0 & 0 & -\alpha_y & \alpha_x & 0 \\ 0 & \alpha_z & -\alpha_y & 0 & 0 & 0 \\ -\alpha_z & 0 & \alpha_x & 0 & 0 & 0 \\ \alpha_y & -\alpha_x & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$P_{13} = \begin{pmatrix} 0 & 0 & 0 & 0 & -\beta_z & \beta_y \\ 0 & 0 & 0 & \beta_z & 0 & -\beta_x \\ 0 & 0 & 0 & -\beta_y & \beta_x & 0 \\ 0 & \beta_z & -\beta_y & 0 & 0 & 0 \\ -\beta_z & 0 & \beta_x & 0 & 0 & 0 \\ \beta_y & -\beta_x & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$P_{14} = \begin{pmatrix} 0 & 0 & 0 & 0 & -\gamma_z & \gamma_y \\ 0 & 0 & 0 & \gamma_z & 0 & -\gamma_x \\ 0 & 0 & 0 & -\gamma_y & \gamma_x & 0 \\ 0 & \gamma_z & -\gamma_y & 0 & 0 & 0 \\ -\gamma_z & 0 & \gamma_x & 0 & 0 & 0 \\ \gamma_y & -\gamma_x & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$P_{21} = -P_{12}$$

$$P_{23} = \begin{pmatrix} 0 & 0 & 0 & 0 & -\delta_z & \delta_y \\ 0 & 0 & 0 & \delta_z & 0 & -\delta_x \\ 0 & 0 & 0 & -\delta_y & \delta_x & 0 \\ 0 & \delta_z & -\delta_y & 0 & 0 & 0 \\ -\delta_z & 0 & \delta_x & 0 & 0 & 0 \\ \delta_y & -\delta_x & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$P_{24} = \begin{pmatrix} 0 & 0 & 0 & 0 & -\epsilon_z & \epsilon_y \\ 0 & 0 & 0 & \epsilon_z & 0 & -\epsilon_x \\ 0 & 0 & 0 & -\epsilon_y & \epsilon_x & 0 \\ 0 & \epsilon_z & -\epsilon_y & 0 & 0 & 0 \\ -\epsilon_z & 0 & \epsilon_x & 0 & 0 & 0 \\ \epsilon_y & -\epsilon_x & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$P_{31} = -P_{13}$$

$$P_{32} = -P_{23}$$

$$P_{34} = \begin{pmatrix} 0 & 0 & 0 & 0 & -F_z & F_y \\ 0 & 0 & 0 & F_z & 0 & -F_x \\ 0 & 0 & 0 & -F_y & F_x & 0 \\ 0 & F_z & -F_y & 0 & 0 & 0 \\ -F_z & 0 & F_x & 0 & 0 & 0 \\ F_y & -F_x & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$P_{41} = -P_{14}$$

$$P_{42} = -P_{24}$$

$$P_{43} = -P_{34}$$

et :

$$\left\{ \begin{array}{l} \alpha_x = n_{1x} - n_{2x} \\ \beta_x = n_{1x} - n_{3x} \\ \gamma_x = n_{1x} - n_{4x} \\ \delta_x = n_{2x} - n_{3x} \\ \varepsilon_x = n_{2x} - n_{4x} \\ F_x = n_{3x} - n_{4x} \\ \alpha_y = n_{1y} - n_{2y} \\ \beta_y = n_{1y} - n_{3y} \\ \gamma_y = n_{1y} - n_{4y} \\ \delta_y = n_{2y} - n_{3y} \\ \varepsilon_y = n_{2y} - n_{4y} \\ F_y = n_{3y} - n_{4y} \\ \alpha_z = n_{1z} - n_{2z} \\ \beta_z = n_{1z} - n_{3z} \\ \gamma_z = n_{1z} - n_{4z} \\ \delta_z = n_{2z} - n_{3z} \\ \varepsilon_z = n_{2z} - n_{4z} \\ F_z = n_{3z} - n_{4z} \end{array} \right.$$

2.4.7.4 Propriétés matricielles

Proposition 15 *Les matrices \mathcal{D} , \mathcal{Z} et \mathcal{R} ont les propriétés suivantes :*

- (i) \mathcal{D} est définie positive.
- (ii) \mathcal{Z} est symétrique et positive.
- (iii) \mathcal{R} est antisymétrique.

Preuve

(i) Tout d'abord, il est évident que \mathcal{D} est symétrique réelle.

De plus, $\forall j \in \llbracket 1; N \rrbracket$ Φ_j et θ_j étant symétriques, alors $\Phi_j \otimes \theta_j$ l'est aussi compte tenu de 2.4.7.2 et par conséquent \mathcal{D} est symétrique diagonale par blocs.

(ii) Ensuite, $(\Phi_1^+$ et $Z_1)$ sont toutes deux symétriques positives, si bien que nous pouvons en déduire que \mathcal{Z} est symétrique diagonale par blocs.

(iii) Enfin,

$$- \forall j \in \llbracket 1; N-1 \rrbracket (\Phi_{j,j+1}^+ \otimes A)^t = (\Phi_{j,j+1}^+)^t \otimes A^t = \Phi_{j+1,j}^+ \otimes A.$$

Ce qui signifie donc que pour $(j, k) \in \llbracket 1; N \rrbracket^2$, $j \neq k$, $\mathcal{R}_{jk} = -\mathcal{R}_{kj}$.

- $\forall j \in \llbracket 1; N \rrbracket$ Φ'_j est antisymétrique, A est symétrique, d'où $\Phi'_j \otimes A$ est antisymétrique vu 2.4.7.2.
- Φ_N^+ est symétrique, \bar{Q} est antisymétrique, d'où $\Phi_N^+ \otimes \bar{Q}$ est antisymétrique toujours grâce à 2.4.7.2, et par conséquent $(\Phi'_N \otimes A + \Phi_N^+ \otimes \bar{Q})$ est antisymétrique par linéarité.

Ces trois points permettent d'appliquer la proposition 11 et de conclure que \mathcal{R} est antisymétrique. ■

Proposition 16 Soit $G = (G_{x_1}, G_{x_2}, G_{x_3})^t$ une matrice de taille 18×6 où pour $j \in \{1, 2, 3\}$, G_{x_j} est une matrice hermitienne de taille 6×6 . Considérons que l'opérateur ∇ agisse sur un vecteur $\mathbf{V} \in \mathbb{C}^6$ de sorte que $\nabla \mathbf{V} = (\partial_{x_1} \mathbf{V}, \partial_{x_2} \mathbf{V}, \partial_{x_3} \mathbf{V})^t$ soit un vecteur de \mathbb{C}^{18} . Dans ces conditions et $\forall \mathbf{V}, \mathbf{W} \in \mathbb{C}^6$:

$$(G(\mathbf{V}), \nabla \mathbf{W}) = \overline{(\operatorname{div} G(\mathbf{W}), \mathbf{V})}$$

où (\cdot, \cdot) désigne le produit scalaire canonique de \mathbb{C}^6 .

Preuve Il est plus simple de raisonner sur G en terme de matrice plutôt qu'en terme d'opérateur. Partons du membre de gauche de l'équation. En faisant passer G de l'autre côté du produit scalaire, on obtient :

$$\begin{aligned} (G(\mathbf{V}), \nabla \mathbf{W}) &= (G\mathbf{V}, \nabla \mathbf{W}) \\ &= (\mathbf{V}, G^* \nabla \mathbf{W}) \end{aligned}$$

En développant maintenant le second membre et en utilisant le fait que G est un opérateur constant, on trouve :

$$\begin{aligned} (\mathbf{V}, G^* \nabla \mathbf{W}) &= (\mathbf{V}, \sum_{j=1}^3 G_{x_j}^* \partial_{x_j} \mathbf{W}) \\ &= (\mathbf{V}, \sum_{j=1}^3 \partial_{x_j} G_{x_j}^* \mathbf{W}) \end{aligned}$$

Ce qui, grâce au caractère hermitien des $(G_{x_j})_{j=1,2,3}$, correspond directement à :

$$\begin{aligned} (\mathbf{V}, \sum_{j=1}^3 \partial_{x_j} G_{x_j}^* \mathbf{W}) &= (\mathbf{V}, \sum_{j=1}^3 \partial_{x_j} G_{x_j} \mathbf{W}) \\ &= (\mathbf{V}, \operatorname{div} G(\mathbf{W})) \\ &= \overline{(\operatorname{div} G(\mathbf{W}), \mathbf{V})} \end{aligned}$$

■

Chapitre 3

Méthodes de décomposition de domaine pour les équations de Maxwell en régime harmonique

3.1 Introduction

Dans la résolution numérique d'un problème modélisé par un système d'équations aux dérivées partielles (EDP), on est généralement confronté, après une étape de discrétisation, à la résolution d'un système linéaire. La taille de ce système ou bien le fait que celui-ci soit mal conditionné, sont deux raisons pour lesquelles une approche globale et une résolution par une méthode directe ne sont plus appropriées. L'arrivée des calculateurs parallèles et leur capacité à résoudre numériquement des problèmes requérant une grande quantité de calculs et une grande capacité de stockage en mémoire ont conduit à une recherche très active d'algorithmes de résolution adaptés à ce type d'architecture. Les méthodes de décomposition de domaine (DD) en sont un des exemples les plus populaires. De nos jours, les méthodes de décomposition de domaine sont très souvent utilisées pour construire des préconditionneurs parallèles pour des méthodes itératives de type Krylov [Saad, 1996] comme la méthode du gradient conjugué ou les méthodes de minimisation de résidus telle que GMRES [Saad and Schultz, 1986].

Les premières méthodes apparues dans la littérature visaient surtout la possibilité de résoudre des problèmes de grande taille par décomposition en problèmes de tailles réduites pouvant être traités par des calculateurs à faible capacité de stockage en mémoire. Dans ce cas, la solution d'un sous-problème permettait de définir les conditions aux limites du sous-problème suivant et on procédait donc à un traitement en séquence des différents sous-problèmes. L'algorithme de Schwarz, initialement conçu pour prouver l'existence de la solution du problème de Poisson posé sur un domaine quelconque [Schwarz, 1870] est la toute première illustration de cette approche. Il repose sur une décomposition spatiale du domaine de calcul en sous-domaines se recouvrant, la zone de recouvrement ainsi introduite permettant une définition simple des conditions aux limites sur les frontières artificielles (aussi nommées *interfaces*). La forme originale de cet algorithme est souvent qualifiée de *multiplicative* parce que le traitement en séquence des sous-problèmes se traduit par un opérateur itératif s'exprimant comme le produit d'opérateurs induits par les résolutions locales (voir aussi [Lions, 1988]-[Lions, 1989]-[Lions, 1990]).

Les méthodes de décomposition de domaine peuvent être classifiées suivant quelques critères (on se référera notamment aux ouvrages récents [Smith et al., 1996] et [Quarteroni and Valli, 1999] qui constituent deux références incontournables sur le sujet) :

- le type de décomposition en régions sur lesquelles on définit les sous-problèmes, on distingue les méthodes *avec recouvrement* ou *sans recouvrement*. Comme souligné plus haut, les premières permettent une actualisation très simple des solutions locales sur la base de conditions aux limites de type Dirichlet ou Neumann (ou une combinaison des deux) aux frontières artificielles. Elles se caractérisent aussi par une complexité arithmétique accrue du fait de la multiple définition des degrés de liberté aux interfaces. Leur inconvénient principal est donc de compliquer quelque peu la mise en œuvre numérique, surtout lors de la résolution de problèmes tridimensionnels sur des géométries de formes complexes ;
- l'interdépendance des résolutions locales à chaque itération, on aboutit à des formes *multiplicatives* (équivalentes à des méthodes itératives du type Gauss-Seidel par bloc) ou *additives* (équivalentes à des méthodes itératives de type Jacobi par bloc). Il est clair que les méthodes multiplicatives (comme l'algorithme de Schwarz sous sa forme originale) sont moins adaptées aux architectures de calcul parallèle. Ces dernières années ont donc vu la mise au point de nombreuses méthodes additives caractérisées par un traitement simultané des sous-problèmes et donc plus à même d'exploiter efficacement les possibilités des calculateurs parallèles.

Les méthodes de décomposition de domaine sans recouvrement sont principalement de deux types : les méthodes de sous-structuration (ou méthodes du complément de Schur) et les méthodes de type Schwarz sans recouvrement. Les méthodes de sous-structuration s'apparentent aux méthodes d'élimination de Gauss par bloc. Elles consistent à ramener la résolution du problème portant sur l'ensemble des degrés de liberté issus de la discrétisation de l'équation aux dérivées partielles sur le domaine de calcul, à la résolution d'un problème de taille moindre posé sur les degrés de liberté attachés aux interfaces délimitant les sous-domaines voisins dans la partition du domaine de calcul. Le problème à l'interface ainsi posé est alors résolu par une méthode itérative adaptée (généralement une méthode de Krylov). L'opérateur à l'interface résultant du processus d'élimination de Gauss par bloc, est appelé complément de Schur. Du point de vue du problème continu, cela revient à la définition d'un nouvel opérateur, l'opérateur de Steklov-Poincaré, qui agit sur les variables à l'interface [Quarteroni and Valli, 1991]-[Quarteroni and Valli, 1999]. Les méthodes de type Schwarz sans recouvrement peuvent aussi être vues comme des cas particuliers de méthodes de sous-structuration. On peut notamment montrer que l'itération de base d'une méthode de Schwarz sans recouvrement peut être reformulée comme une itération appliquée aux seuls degrés de liberté localisés sur les interfaces [Magoulès et al., 2004]. Cependant par rapport au cas avec recouvrement, on est ici amené à formuler des conditions aux limites plus complexes pour retrouver un niveau d'efficacité similaire à celui caractérisant une méthode de Schwarz avec recouvrement, ou même tout simplement assurer la convergence. La variante sans recouvrement de l'algorithme de Schwarz original a été introduite par P.L. Lions [Lions, 1990] et ensuite étudiée par Després [Després, 1991] et Nataf *et al.* [Nataf et al., 1994].

Notons aussi que le terme décomposition de domaine a selon [Smith et al., 1996] quelques significations sensiblement différentes comme par exemple le processus de distribution des données entre les processeurs d'un ordinateur à mémoire distribuée, la séparation du domaine physique dans des régions où les phénomènes physiques sont modélisés par des EDP différentes (raccord de modèles), ou encore, du point de vue du préconditionnement, la séparation du problème initial en sous-problèmes de dimensions plus petites et donc plus faciles à résoudre. Dans ce cas, chaque résolution locale définissant une contribution à un inverse approché de l'opérateur global. D'une façon générale, les méthodes de décomposition de domaine permettent une parallélisation plus facile et aussi plus efficace, en favorisant explicitement la localisation des données. Elles peuvent aussi conduire à des algorithmes présentant des propriétés de convergence supérieures. Concrètement, les méthodes de

décomposition de domaine ont été utilisées dans des secteurs aussi variés que la mécanique des structures [Champaney et al., 1997]-[Farhat and Roux, 1991]-[Le Tallec, 1994]-[Pierson, 2000], la mécanique des fluides [Bourgat et al., 1994]-[Clerc, 1998]- [Fatone et al., 1999]-[Paglieri et al., 1997]-[Dolean et al., 2004], l'élasticité linéaire [Smith, 1992] ou encore la propagation d'ondes en régime harmonique. C'est à ce dernier secteur que nous allons nous intéresser plus particulièrement.

Comme nous l'avons précisé ci-dessus, les méthodes de décomposition de domaines ont été introduites pour résoudre le problème de Poisson, qui rentre dans la classe des problèmes elliptiques coercifs. Cette propriété permet alors d'appliquer des conditions d'interface classiques de type Dirichlet ou Neumann dans les contextes des méthodes de Schwarz avec recouvrement. Elle est malheureusement perdue pour les équations de Helmholtz et de Maxwell en régime harmonique et, dans ce cas, des conditions d'interface plus complexes pour les algorithmes de Schwarz sans recouvrement doivent être appliquées car les conditions d'interface naturelles fonctionnent pour les modes évanescents mais ne traitent pas les modes propagatifs. Després [Després, 1991] a été le premier à utiliser des conditions d'interface de type Robin pour l'algorithme de Schwarz sans recouvrement, conditions qui peuvent être améliorées en raffinant le maillage pour les modes propagatifs [Cai and Widlund, 1992]. Cela a été le point de départ de nombreuses études sur les méthodes de décomposition de domaine pour l'équations d'Helmholtz et les équations de Maxwell harmoniques en formulation du second ordre. Dans [Després et al., 1992] les conditions de Robin de [Després, 1991] sont réutilisées en ajoutant une condition sur la régularité de la trace tangentielle du champ électrique pour montrer la convergence de l'algorithme par une méthode de décroissance de l'énergie discrète à l'interface. Cet algorithme engendre cependant des solutions numériques instables qui disparaissent en introduisant un paramètre de sous-relaxation. Par la suite, une généralisation de cet algorithme a été mise en œuvre où les conditions d'interface utilisées dérivent des conditions de Robin en utilisant des conditions absorbantes mixtes du premier ordre puis du second ordre. La convergence a été améliorée en introduisant deux paramètres de pénalisation qui permettent non seulement d'obtenir un opérateur défini positif mais surtout, de diminuer le saut de la composante tangentielle du champ électrique. D'autres études ont été menées en s'inspirant cette fois-ci de la méthode FETI pour traiter le problème de Helmholtz extérieur dans le cadre de l'acoustique [de La Bourdonnaye et al., 1998]. On construit ici un problème à l'interface à l'aide de multiplieurs de Lagrange. La méthode repose essentiellement sur trois idées clés. Tout d'abord, l'élimination de résonances locales via une stabilisation de chaque opérateur de sous-domaine par une matrice de masse complexe à l'interface associée à des conditions de radiation entre chaque sous-domaines. Un préconditionneur local est ensuite utilisé pour filtrer les hautes-fréquences et accélérer la convergence en présence de maillages fins, et enfin un préconditionneur global permet quant à lui de filtrer les basses fréquences de l'erreur.

La technique employée dans [Gander et al., 2002] reprend les conditions d'interface de type Robin et repose sur une analyse de Fourier pour déterminer l'opérateur à l'interface qui entraîne la convergence immédiate de l'algorithme de Schwarz pour le système de Helmholtz. Cet opérateur étant non-local, on ne peut l'implémenter tel quel, et une optimisation est donc effectuée afin de déterminer le meilleur opérateur local qui approche les conditions absorbantes exactes posées sur les interfaces. Ces opérateurs locaux sont des approximations polynomiales des conditions de transmission exactes. La démarche effectuée dans [Gander et al., 2002] a récemment été adaptée aux équations de Maxwell harmoniques du second ordre dans [Alonso-Rodriguez and Gerardo-Giorda, 2004] et [Lee et al., 2005]. On obtient alors un taux de convergence dont l'expression est analogue à celui que l'on trouve à l'issue de l'analyse de l'équation de Helmholtz. Dans [Lee et al., 2005], les auteurs font aussi appel à une technique de raccord particulière qui permet de faire coïncider à l'interface des grilles non-conformes, tout en utili-

sant une condition de transmission du premier ordre pour renforcer la continuité des champs entre les sous-domaines.

L'objectif ultime de notre étude est de mettre au point des méthodes de décomposition de domaine pour la résolution numérique parallèle des équations de Maxwell harmoniques en formulation du premier ordre, discrétisées par des méthodes de type Galerkin discontinu en maillages tétraédriques. Pour ce faire, nous commençons par considérer le cas des équations de Maxwell 2D en mode TM et nous formulons et analysons des algorithmes de Schwarz avec et sans recouvrement en nous inspirant des démarches de Gander *et al.* [Gander et al., 2002] et Alonso et Gerardo [Alonso-Rodriguez and Gerardo-Giorda, 2004]. Nous appliquons tout d'abord à ces équations les conditions d'interface utilisées dans [Després, 1991] et évaluons l'efficacité des algorithmes de Schwarz avec et sans recouvrement en calculant les taux de convergence continus et discrets par une méthode de Fourier. Nous concluons alors que ces conditions ne permettent pas d'assurer la convergence de l'algorithme de Schwarz sans recouvrement. Dans une seconde phase, nous nous intéressons à des conditions d'interface qui approchent les conditions absorbantes exactes. Différentes approximations sont considérées et elles aussi étudiées par une méthode de Fourier avant de procéder à l'optimisation des conditions d'interface résultantes et à l'illustration des résultats théoriques par des expériences numériques 2D sur un cas test simplifié. Des expériences numériques 3D sont présentées dans le chapitre 4.

3.2 Conditions d'interface naturelles

Dans cette section, nous nous concentrons sur les équations de Maxwell 2D en mode TM (transverse magnétique). Nous formulons et analysons des algorithmes de Schwarz avec et sans recouvrement basés sur des conditions d'interface d'ordre zéro (conditions de flux entrant). Ces conditions d'interface sont aussi qualifiées de *naturelles* ou *classiques* car elles sont déduites de la formulation faible associée au problème aux limites défini pour un système hyperbolique du premier ordre (on peut aussi se placer dans le cadre plus général des systèmes de Friedrichs dont le système des équations de Maxwell est un exemple ; voir notamment [Dolean, 2001] et [Helluy, 1994]). Dans une seconde étape nous considérons exclusivement une décomposition sans recouvrement et nous cherchons à construire des conditions d'interface optimisées pour accélérer la convergence de l'algorithme de Schwarz.

3.2.1 Équations de Maxwell 2D en mode TM

Dans ce cas, la composante verticale du champ magnétique est nulle : $H_z = 0$. Cela implique pour le champ électrique : $E_x = 0$ et $E_y = 0$. On applique un terme source $\mathbf{f} = (f_1, f_2, f_3)^t$ sur le domaine de calcul noté Ω . Enfin, on suppose qu'on applique sur la frontière du domaine $\partial\Omega$ une onde incidente notée $(E_z^{\text{inc}}, H_x^{\text{inc}}, H_y^{\text{inc}})^t$ grâce à des conditions absorbantes de type flux entrant. Ces équations qui couplent les inconnues E_z , H_x et H_y en variables adimensionnées sont alors les suivantes en domaine

fréquentiel (écrites en champ total) :

$$\left\{ \begin{array}{ll} i\omega E_z - \partial_x \left(\frac{H_y}{\mu} \right) + \partial_y \left(\frac{H_x}{\mu} \right) = f_1 & \text{dans } \Omega, \\ i\omega H_x + \partial_y \left(\frac{E_z}{\varepsilon} \right) = f_2 & \text{dans } \Omega, \\ i\omega H_y - \partial_x \left(\frac{E_z}{\varepsilon} \right) = f_3 & \text{dans } \Omega, \\ E_z - z^{-1}(n_y H_x - n_x H_y) = (E_z^{\text{inc}} - z^{-1}(n_y H_x^{\text{inc}} - n_x H_y^{\text{inc}})) & \text{sur } \partial\Omega. \end{array} \right. \quad (2.1)$$

Le terme $\mathbf{f} = (f_1, f_2, f_3)^t$ est nul en l'absence de courants sources. Par ailleurs, les variables ε et μ sont respectivement la permittivité électrique et la perméabilité magnétique qui caractérisent le milieu. L'impédance z s'exprime comme $z = \sqrt{\mu/\varepsilon}$.

Remarque 7 On a adopté ci-dessus des notations et des hypothèses physiques similaires à celles adoptées au chapitre 2 pour les équations de Maxwell 1D et 3D, à l'adimensionnement près. En effet, on choisit ici de manipuler une forme conservative des équations de Maxwell bien que cela ne soit pas indispensable à notre étude.

En mettant les trois équations de (2.1) qui portent sur Ω sous la forme d'un système on obtient :

$$i\omega \mathbf{W} + \partial_x(M_x \mathbf{W}) + \partial_y(M_y \mathbf{W}) = \mathbf{f}, \quad (2.2)$$

où :

$$\mathbf{W} = \begin{pmatrix} E \\ \mathbf{H} \end{pmatrix} = \begin{pmatrix} E_z \\ H_x \\ H_y \end{pmatrix}, \quad M_x = \begin{pmatrix} 0 & 0 & -\mu^{-1} \\ 0 & 0 & 0 \\ -\varepsilon^{-1} & 0 & 0 \end{pmatrix} \quad \text{et} \quad M_y = \begin{pmatrix} 0 & \mu^{-1} & 0 \\ \varepsilon^{-1} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

On note $\mathbf{n} = (n_x, n_y)^t$ un vecteur normal unitaire à une surface Σ donnée et $\mathbf{v} = (n_y, -n_x)^t$ le vecteur tangent associé. Ces deux vecteurs forment une base orthonormale de \mathbb{R}^2 et par conséquent, si on désigne par N la matrice $\mathbf{n}\mathbf{n}^t$ on peut écrire :

$$\mathbf{v}\mathbf{v}^t = \mathbf{I}_2 - N, \quad N\mathbf{n} = \mathbf{n}, \quad N\mathbf{v} = \mathbf{0}.$$

La condition aux limites de (2.1) peut se réécrire identiquement aux cas 1D et 3D en utilisant :

$$M_{\mathbf{n}}^+ = \frac{c}{2} \begin{pmatrix} 1 & z^{-1}\mathbf{v}^t \\ z\mathbf{v} & \mathbf{I}_2 - N \end{pmatrix} \quad \text{et} \quad M_{\mathbf{n}}^- = \frac{c}{2} \begin{pmatrix} -1 & z^{-1}\mathbf{v}^t \\ z\mathbf{v} & N - \mathbf{I}_2 \end{pmatrix}, \quad (2.3)$$

à partir de :

$$M_{\mathbf{n}} = n_x M_x + n_y M_y = \begin{pmatrix} 0 & \mu^{-1}\mathbf{v}^t \\ \varepsilon^{-1}\mathbf{v} & \mathbf{0}_2 \end{pmatrix}. \quad (2.4)$$

En effet, sur $\partial\Omega$:

$$\begin{aligned}
E_z - z^{-1}(n_y H_x - n_x H_y) &= (E_z^{\text{inc}} - z^{-1}(n_y H_x^{\text{inc}} - n_x H_y^{\text{inc}})) \\
\Leftrightarrow \begin{pmatrix} -(E - z^{-1} \mathbf{H} \cdot \mathbf{v}) \\ (E - z^{-1} \mathbf{H} \cdot \mathbf{v}) \mathbf{v} \end{pmatrix} &= \begin{pmatrix} -(E^{\text{inc}} - z^{-1} \mathbf{H}^{\text{inc}} \cdot \mathbf{v}) \\ (E^{\text{inc}} - z^{-1} \mathbf{H}^{\text{inc}} \cdot \mathbf{v}) \mathbf{v} \end{pmatrix} \\
\Leftrightarrow M_{\mathbf{n}}^- \mathbf{W} &= M_{\mathbf{n}}^- \mathbf{W}^{\text{inc}},
\end{aligned}$$

Si on introduit l'opérateur $\mathcal{L} = i\omega + M_x \partial_x + M_y \partial_y$, alors on peut réécrire (2.1) de la façon suivante :

$$\begin{cases} \mathcal{L} \mathbf{W} &= \mathbf{f} & \text{sur } \Omega, \\ M_{\mathbf{n}}^- \mathbf{W} &= M_{\mathbf{n}}^- \mathbf{W}^{\text{inc}} & \text{sur } \partial\Omega. \end{cases} \quad (2.5)$$

Il s'agit maintenant de définir un algorithme de Schwarz avec ou sans recouvrement appliqué au système (2.5).

3.2.2 Algorithme de Schwarz

3.2.2.1 Formulation de l'algorithme

Si on découpe Ω en m sous-domaines $\Omega = \bigcup_{j=1}^m \Omega_j$, nous pouvons avoir un découpage recouvrant ou non-recouvrant, ce qui implique que pour $j \neq l$, $\Omega_j \cap \Omega_l \neq \emptyset$ ou $\Omega_j \cap \Omega_l = \emptyset$. En démarrant avec une solution initiale $\mathbf{W}^{j,0}$, l'approximation de $\mathbf{W}|_{\Omega_j}$ à l'itération $p+1$ notée $\mathbf{W}^{j,p+1}$, est calculée à partir de la solution à l'itération p notée $\mathbf{W}^{j,p}$, en résolvant les problèmes locaux suivants [Després, 1991] :

$$\begin{cases} \mathcal{L} \mathbf{W}^{j,p+1} &= \mathbf{f}^j & \text{dans } \Omega_j, \\ M_{\mathbf{n}_{jl}}^- \mathbf{W}^{j,p+1} &= M_{\mathbf{n}_{jl}}^- \mathbf{W}^{l,p} & \text{sur } \Gamma_{jl} = \partial\Omega_j \cap \partial\Omega_l, \\ M_{\mathbf{n}}^- \mathbf{W}^{j,p+1} &= M_{\mathbf{n}}^- \mathbf{W}^{\text{inc}} & \text{sur } \Gamma_j = \partial\Omega \cap \partial\Omega_j, \end{cases} \quad (2.6)$$

où \mathbf{n}_{jl} est la normale sortante à Γ_{jl} et \mathbf{n} est la normale sortante à la frontière du domaine global. Pour l'analyse de convergence de cet algorithme, on se limite au cas de deux sous-domaines comme représenté sur la figure 2.1 dans le cas de sous-domaines recouvrants.

Dans ce cas l'algorithme (2.6) devient :

$$\begin{cases} \mathcal{L} \mathbf{W}^{1,p+1} &= \mathbf{f}^1 & \text{sur } \Omega_1, \\ M_{\mathbf{n}_{12}}^- \mathbf{W}^{1,p+1} &= M_{\mathbf{n}_{12}}^- \mathbf{W}^{2,p} & \text{sur } \Gamma_{12}, \\ M_{\mathbf{n}}^- \mathbf{W}^{1,p+1} &= M_{\mathbf{n}}^- \mathbf{W}^{\text{inc}} & \text{sur } \Gamma_1, \end{cases} \quad \begin{cases} \mathcal{L} \mathbf{W}^{2,p+1} &= \mathbf{f}^2 & \text{sur } \Omega_2, \\ M_{\mathbf{n}_{21}}^- \mathbf{W}^{2,p+1} &= M_{\mathbf{n}_{21}}^- \mathbf{W}^{1,p} & \text{sur } \Gamma_{21}, \\ M_{\mathbf{n}}^- \mathbf{W}^{2,p+1} &= M_{\mathbf{n}}^- \mathbf{W}^{\text{inc}} & \text{sur } \Gamma_2, \end{cases} \quad (2.7)$$

avec $\mathbf{n}_{12} = -\mathbf{n}_{21}$ ce qui conduit à $M_{\mathbf{n}_{21}}^- = -M_{\mathbf{n}_{12}}^+$. On procède dès maintenant à une analyse dont on peut retrouver la démarche dans [Alonso-Rodriguez and Gerardo-Giorda, 2004], [Dolean et al., 2004], et [Gander et al., 2002].

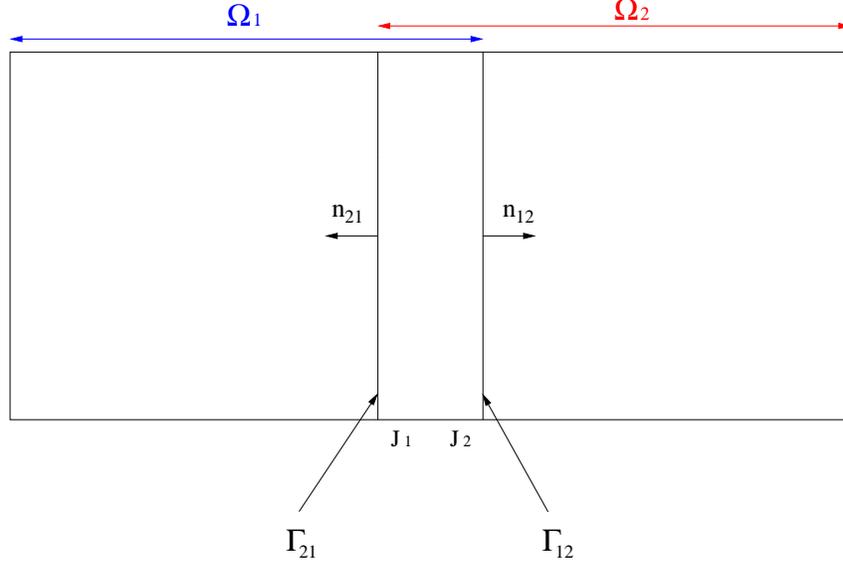


FIG. 2.1 – Décomposition en deux sous-domaines recouvrants.

3.2.2.2 Analyse de convergence de l'algorithme en continu

Pour pouvoir réaliser l'étude de convergence en utilisant l'analyse de Fourier on se place dans le cas d'une décomposition du domaine $\Omega = \mathbf{R}^2$ en deux sous-domaines infinis :

$$\Omega_1 =] - \infty, b[\times \mathbf{R} \quad \text{et} \quad \Omega_2 =]a, +\infty[\times \mathbf{R},$$

avec $a \leq b$ et en supposant que les paramètres ε et μ sont constants, pour pouvoir effectuer une transformée de Fourier sans précautions supplémentaires. Dans ce cas, les interfaces (frontières artificielles entre les sous-domaines) sont les droites $x = a$ et $x = b$ de normales extérieures $\mathbf{n}_{12} = -\mathbf{n}_{21} = (1, 0)^t$. Il est profitable de convertir notre système (2.5) en un système qui porte sur les variables caractéristiques suivant la variable x (dans la direction normale à l'interface). Notons que $M\mathbf{n}_{12} = M_x$ est diagonalisable en $M_x = T\Lambda T^{-1}$ où :

$$T = \begin{pmatrix} c^{-1} & 0 & -c^{-1} \\ 0 & 1 & 0 \\ \mu & 0 & \mu \end{pmatrix}, \quad \Lambda = \begin{pmatrix} -c & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & c \end{pmatrix} \quad \text{et} \quad T^{-1} = \frac{1}{2} \begin{pmatrix} c & 0 & \mu^{-1} \\ 0 & 2 & 0 \\ -c & 0 & \mu^{-1} \end{pmatrix}.$$

On introduit par la même occasion :

$$A_y = T^{-1}M_yT = \frac{1}{2} \begin{pmatrix} 0 & c\mu^{-1} & 0 \\ 2z & 0 & -2z \\ 0 & -c\mu^{-1} & 0 \end{pmatrix}.$$

et on pose $\mathbf{U} = T^{-1}\mathbf{W}$. On a alors :

$$\mathcal{L}\mathbf{W} = \tilde{\mathcal{L}}\mathbf{U} \quad \text{avec} \quad \mathbf{U} = \frac{1}{2} \begin{pmatrix} \mu^{-1}w_3 + cw_1 \\ 2w_2 \\ \mu^{-1}w_3 - cw_1 \end{pmatrix} \quad \text{et} \quad \tilde{\mathcal{L}} = i\omega l_3 + \Lambda \partial_x + A_y \partial_y. \quad (2.8)$$

On a $\Lambda^+ = T^{-1}M_x^+T$ et $\Lambda^- = T^{-1}M_x^-T$ et les conditions aux limites de (2.7) pour $\mathbf{n} = (1, 0)^t$ sont transformées en :

$$\Lambda^+ \mathbf{U} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & c \end{pmatrix} \mathbf{U} = \begin{pmatrix} 0 \\ 0 \\ cu_3 \end{pmatrix} \quad \text{et} \quad \Lambda^- \mathbf{U} = \begin{pmatrix} -c & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \mathbf{U} = \begin{pmatrix} -cu_1 \\ 0 \\ 0 \end{pmatrix}.$$

et $\tilde{\mathbf{f}} = T^{-1}\mathbf{f}$. A l'aide de ces nouvelles variables, l'algorithme (2.7) peut se réécrire sur \mathbf{R}^2 comme :

$$\begin{cases} \tilde{\mathcal{L}}\mathbf{U}^{1,p+1} = \tilde{\mathbf{f}}^1 & \text{dans } \Omega_1, \\ u_1^{1,p+1} = u_1^{2,p} & \text{sur } x = b, \end{cases} \quad \begin{cases} \tilde{\mathcal{L}}\mathbf{U}^{2,p+1} = \tilde{\mathbf{f}}^2 & \text{dans } \Omega_2, \\ u_3^{2,p+1} = u_3^{1,p} & \text{sur } x = a, \end{cases} \quad (2.9)$$

où on a noté par $u_l^{j,p}$ la $l^{\text{ème}}$ composante du vecteur $\mathbf{U}^{j,p}$ qui représente l'approximation de la solution à l'itération p dans le domaine Ω_j . Afin d'exprimer le taux de convergence de l'algorithme, nous allons étudier l'erreur locale $\mathbf{E}^{j,p} = \mathbf{U}^{j,p} - \mathbf{U}_{|\Omega_j}$ commise entre la solution approchée $\mathbf{U}^{j,p}$ à l'itération p et la solution exacte $\mathbf{U}_{|\Omega_j}$ pour $j = 1, 2$. Par linéarité, le système (2.9) vérifie aussi :

$$\begin{cases} \tilde{\mathcal{L}}\mathbf{E}^{1,p+1} = 0 & \text{dans } \Omega_1, \\ e_1^{1,p+1} = e_1^{2,p} & \text{sur } x = b, \end{cases} \quad \begin{cases} \tilde{\mathcal{L}}\mathbf{E}^{2,p+1} = 0 & \text{dans } \Omega_2, \\ e_3^{2,p+1} = e_3^{1,p} & \text{sur } x = a. \end{cases} \quad (2.10)$$

Il s'agit maintenant d'exprimer le taux de convergence de l'algorithme (2.10). Pour cela, on fait appel à l'analyse de Fourier. Plus précisément, dans ce qui suit, on va effectuer une transformée de Fourier partielle suivant la variable y :

$$\hat{\mathbf{E}}^{j,p}(x, k) = (\mathcal{F}_y \mathbf{E}^{j,p})(x, k) = \int_{\mathbf{R}} e^{-iky} \mathbf{E}^{j,p}(x, y) dy.$$

Nous avons :

$$\tilde{\mathcal{L}}\mathbf{E}^{j,p} = 0 \Leftrightarrow \mathcal{F}_y \tilde{\mathcal{L}}\mathbf{E}^{j,p} = 0 \Leftrightarrow \Lambda \partial_x \hat{\mathbf{E}}^{j,p} + i(\omega l_3 + k A_y) \hat{\mathbf{E}}^{j,p} = 0. \quad (2.11)$$

Il se trouve que l'opérateur Λ qui intervient dans la définition de l'équation différentielle ordinaire (2.11) comporte une valeur propre nulle. Comme nous sommes en variables caractéristiques suivant la direction de propagation en x , cela signifie que la composante H_x du champ ne se propage pas dans cette direction. Il est donc naturel de réduire notre système aux composantes qui interviennent dans la transmission de l'information. En laissant alors de côté les indices correspondant aux sous-domaines et à l'itération, on obtient que localement l'équation (2.11) s'écrit :

$$\begin{cases} i\omega \hat{e}_1 - c\partial_x \hat{e}_1 + \frac{ikc\mu^{-1}}{2} \hat{e}_2 = 0, \\ i\omega \hat{e}_2 + ikz \hat{e}_1 - ikz \hat{e}_3 = 0, \\ i\omega \hat{e}_3 + c\partial_x \hat{e}_3 - \frac{ikc\mu^{-1}}{2} \hat{e}_2 = 0. \end{cases}$$

La deuxième équation permet d'exprimer \hat{e}_2 en fonction des deux autres variables :

$$\hat{e}_2 = \frac{kz}{\omega} (-\hat{e}_1 + \hat{e}_3),$$

et, en remplaçant dans les deux autres équations, on obtient :

$$(2.11) \Leftrightarrow \begin{pmatrix} -c & 0 \\ 0 & c \end{pmatrix} \partial_x \begin{pmatrix} \hat{e}_1 \\ \hat{e}_3 \end{pmatrix} + \begin{pmatrix} i \left(\omega - \frac{k^2 c^2}{2\omega} \right) & i \frac{k^2 c^2}{2\omega} \\ i \frac{k^2 c^2}{2\omega} & i \left(\omega - \frac{k^2 c^2}{2\omega} \right) \end{pmatrix} \begin{pmatrix} \hat{e}_1 \\ \hat{e}_3 \end{pmatrix} = 0 \quad (2.12)$$

$$\Leftrightarrow \partial_x \begin{pmatrix} \hat{e}_1 \\ \hat{e}_3 \end{pmatrix} + \hat{M} \begin{pmatrix} \hat{e}_1 \\ \hat{e}_3 \end{pmatrix} = 0 \text{ où } \hat{M} = \frac{i}{2\omega c} \begin{pmatrix} -2\omega^2 + k^2 c^2 & -k^2 c^2 \\ k^2 c^2 & 2\omega^2 - k^2 c^2 \end{pmatrix}.$$

Les solutions de (2.12) s'expriment comme (ici par abus de notation on a désigné par $\hat{\mathbf{E}}$ le vecteur des composantes de l'erreur réduit aux variables \hat{e}_1 et \hat{e}_3) :

$$\hat{\mathbf{E}}(x, k) = \begin{pmatrix} \hat{e}_1 \\ \hat{e}_3 \end{pmatrix} = \gamma_1 e^{-\lambda_1(k)x} \mathbf{V}_1(k) + \gamma_2 e^{-\lambda_2(k)x} \mathbf{V}_2(k) \text{ avec } (\gamma_1, \gamma_2) \in \mathbb{C}^2. \quad (2.13)$$

où :

$$\begin{aligned} \lambda_1(k) &= -\sqrt{k^2 - \tilde{\omega}^2} & , & \quad \mathbf{V}_1(k) = (\sqrt{k^2 - \tilde{\omega}^2} + i\tilde{\omega}, \sqrt{k^2 - \tilde{\omega}^2} - i\tilde{\omega})^t, \\ \lambda_2(k) &= \sqrt{k^2 - \tilde{\omega}^2} & , & \quad \mathbf{V}_2(k) = (\sqrt{k^2 - \tilde{\omega}^2} - i\tilde{\omega}, \sqrt{k^2 - \tilde{\omega}^2} + i\tilde{\omega})^t, \end{aligned} \quad (2.14)$$

sont les valeurs propres de \hat{M} et leurs vecteurs propres associés, et $\tilde{\omega} = \omega/c$. Dans la suite, on abandonne le symbole $\tilde{}$ pour simplifier l'écriture.

Revenons maintenant au contexte de l'algorithme (2.10) où le vecteur erreur s'exprime localement dans chaque sous-domaine comme :

$$\hat{\mathbf{E}}^{j,p} = (\gamma_1)^{j,p} e^{-\lambda_1(k)x} \mathbf{V}_1(k) + (\gamma_2)^{j,p} e^{-\lambda_2(k)x} \mathbf{V}_2(k).$$

Les composantes de l'erreur doivent être bornées quand x tend respectivement vers $\pm\infty$ pour toutes les valeurs de k . Il faut donc étudier le signe de la partie réelle des valeurs propres $\lambda_{1,2}$:

- Pour $|k| \leq \omega$, on a :

$$\lambda_2(k) = -\lambda_1(k) = i\sqrt{\omega^2 - k^2} \Rightarrow \Re\lambda_1 = \Re\lambda_2 = 0,$$

ce qui prouve que $\hat{\mathbf{E}}^{j,p}(x, k)$ est bornée pour toute valeur de k et de x .

- Pour $|k| > \omega$, on a :

$$\lambda_2(k) = -\lambda_1(k) = \sqrt{k^2 - \omega^2} \Rightarrow \Re\lambda_1 < 0, \Re\lambda_2 > 0,$$

et il est donc nécessaire que $(\gamma_2)^{1,p} = 0$ et $(\gamma_1)^{2,p} = 0$ afin que les solutions demeurent bornées.

En résumé, on déduit que les solutions locales s'écrivent sous la forme suivante :

$$\begin{cases} \hat{\mathbf{E}}^{1,p}(x, k) = \gamma_1^p e^{-\lambda_1(k)x} \mathbf{V}_1(k) = \gamma_1^p e^{-\lambda_1 x} (\sqrt{k^2 - \omega^2} + i\omega, \sqrt{k^2 - \omega^2} - i\omega)^t, \\ \hat{\mathbf{E}}^{2,p}(x, k) = \gamma_2^p e^{-\lambda_2(k)x} \mathbf{V}_2(k) = \gamma_2^p e^{-\lambda_2 x} (\sqrt{k^2 - \omega^2} - i\omega, \sqrt{k^2 - \omega^2} + i\omega)^t. \end{cases} \quad (2.15)$$

L'algorithme (2.10) après avoir effectué la transformée de Fourier partielle suivant y devient :

$$\begin{cases} \partial_x \hat{\mathbf{E}}^{1,p+1} + \hat{M} \hat{\mathbf{E}}^{1,p+1} = 0 & \text{dans } \Omega_1, \\ \hat{e}_1^{1,p+1} = \hat{e}_1^{2,p} & \text{sur } x = b, \end{cases} \quad \begin{cases} \partial_x \hat{\mathbf{E}}^{2,p+1} + \hat{M} \hat{\mathbf{E}}^{2,p+1} = 0 & \text{dans } \Omega_2, \\ \hat{e}_3^{2,p+1} = \hat{e}_3^{1,p} & \text{sur } x = a. \end{cases} \quad (2.16)$$

En formulant les conditions d'interfaces dans (2.16) à l'aide des solutions (2.15) on obtient :

$$\begin{cases} \hat{e}_1^{1,p+1} = \hat{e}_1^{2,p} \\ \hat{e}_3^{2,p+1} = \hat{e}_3^{1,p} \end{cases} \Leftrightarrow \begin{cases} \gamma_1^{p+1} e^{-\lambda_1 b (\sqrt{k^2 - \omega^2} + i\omega)} = \gamma_2^p e^{-\lambda_2 b (\sqrt{k^2 - \omega^2} - i\omega)} \\ \gamma_2^{p+1} e^{-\lambda_2 a (\sqrt{k^2 - \omega^2} + i\omega)} = \gamma_1^p e^{-\lambda_1 a (\sqrt{k^2 - \omega^2} - i\omega)} \end{cases} \quad (2.17)$$

Des relations (2.17), on déduit :

$$\begin{cases} \gamma_1^{p+1} = \gamma_2^p \left(\frac{\sqrt{k^2 - \omega^2} - i\omega}{\sqrt{k^2 - \omega^2} + i\omega} \right) e^{-(\lambda_2 - \lambda_1)b}, \\ \gamma_2^{p+1} = \gamma_1^p \left(\frac{\sqrt{k^2 - \omega^2} - i\omega}{\sqrt{k^2 - \omega^2} + i\omega} \right) e^{(\lambda_2 - \lambda_1)a}, \end{cases} \quad (2.18)$$

et ensuite :

$$\frac{\gamma_1^{p+1}}{\gamma_1^{p-1}} = \frac{\gamma_2^{p+1}}{\gamma_2^{p-1}} = \left(\frac{\sqrt{k^2 - \omega^2} - i\omega}{\sqrt{k^2 - \omega^2} + i\omega} \right)^2 e^{-2\delta\sqrt{k^2 - \omega^2}}, \quad (2.19)$$

où on a noté $\delta = b - a$ la taille du recouvrement entre les sous-domaines. Si on note par $\rho(k, \delta)$ le taux de réduction de l'erreur entre deux itérations successives on déduit de (2.19) que :

$$\rho(k, \delta) = \left| \left(\frac{\sqrt{k^2 - \omega^2} - i\omega}{\sqrt{k^2 - \omega^2} + i\omega} \right) e^{-\delta\sqrt{k^2 - \omega^2}} \right|. \quad (2.20)$$

L'algorithme de Schwarz sera convergent si et seulement si $\forall k \in \mathbf{R}, \rho(k, \delta) < 1$. Or, on a que :

$$\rho(k, \delta) = \begin{cases} \left| \frac{\sqrt{\omega^2 - k^2} - \omega}{\sqrt{\omega^2 - k^2} + \omega} \right| & \text{si } |k| < \omega, \\ e^{-\delta\sqrt{k^2 - \omega^2}} & \text{si } |k| \geq \omega. \end{cases} \quad (2.21)$$

On en déduit que l'algorithme est convergent pour les modes propagatifs et ceci indépendamment de la présence du recouvrement (on voit même ce celui-ci ne joue aucun rôle dans la convergence pour ces modes) mais il converge pour les modes évanescents seulement en présence du recouvrement. En même temps, il existe un mode $|k| = \omega$ pour lequel on a toujours $\rho(k, \delta) = 1$. En particulier, pour $\delta \geq 0$ le taux de convergence global de l'algorithme, qui est donné par la norme infinie de ρ , vaut :

$$\rho_\infty = \sup_{k \in \mathbf{R}} \rho(k, \delta) = 1. \quad (2.22)$$

On en conclut que la méthode itérative seule ne converge pas strictement et qu'il sera préférable au niveau discret d'accélérer l'algorithme par une méthode de Krylov telle que la méthode de minimisation des résidus généralisés GMRES [Saad and Schultz, 1986]. Ce type de méthode projette la solution itérative orthogonalement à l'espace vectoriel engendré par les directions de descente qui nuisent à la convergence.

La convergence de l'algorithme de Schwarz (2.7) appliqué aux équations de Maxwell 3D a récemment été étudiée dans [Dolean et al., 2006]. On constate alors que le taux qui exprime la vitesse de convergence de l'algorithme de Schwarz pour les équations de Maxwell 3D est identique à celui obtenu pour les équations de Maxwell 2D (2.20). La différence réside dans la variable de Fourier qui est étendue à deux dimensions pour le traitement des équations tridimensionnelles. Par souci de complétude, les principales étapes de cette analyse sont données dans l'annexe 3.4.1.

Pour conclure cette analyse, observons les courbes du taux de convergence $\rho(k, \delta)$ en fonction du nombre d'onde adimensionné k et évaluons l'influence du recouvrement pour deux fréquences distinctes. Les paramètres ε et μ sont tous deux fixés à 1. De cette façon, nous obtenons $\omega \approx 8,38$ pour $f = 400$ MHz et $\omega \approx 20,94$ pour $f = 1$ GHz (rappelons que ω correspond ici à $\tilde{\omega} = \omega/c_0$ où $c_0 = 3.10^8$ m/s). En prévision de comparaisons avec le comportement de l'algorithme de Schwarz en discret étudié dans la section qui suit, on suppose que la taille du recouvrement δ est un multiple d'un pas d'espace pour l'instant artificiel, noté Δx et dont la valeur est fixée à $1/50$. Les courbes du taux de convergence sont représentées sur les figures 2.2 et 2.3 pour différentes tailles de recouvrement. Sur ces courbes nous voyons clairement que l'algorithme ne converge pas strictement pour $k \geq \omega$ si la taille du recouvrement est nulle. Quant aux autres courbes caractérisant différentes tailles de recouvrement non nulles, on s'aperçoit effectivement que $\rho(k, \delta) = 1$ pour $k = \omega$ et que $\rho(k, \delta) < 1$ pour toutes les autres valeurs de k .

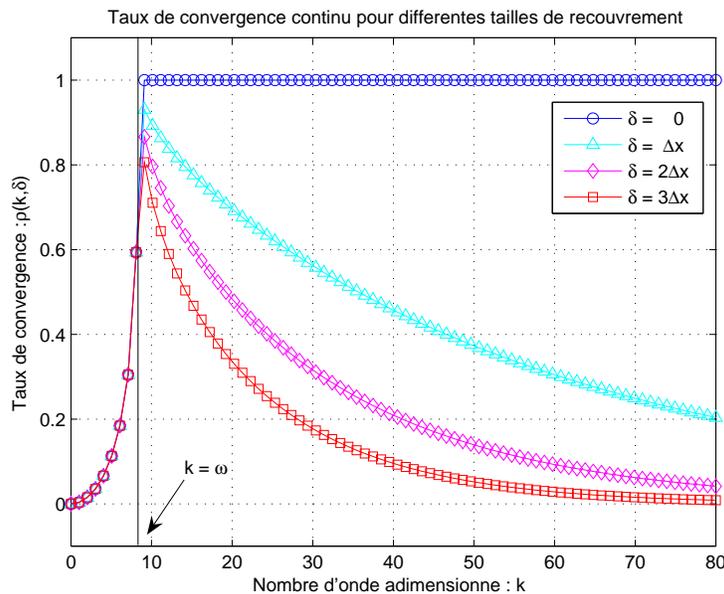


FIG. 2.2 – Evolution du taux de convergence à $f = 400$ MHz pour différentes tailles de recouvrement. Algorithme de Schwarz en continu.

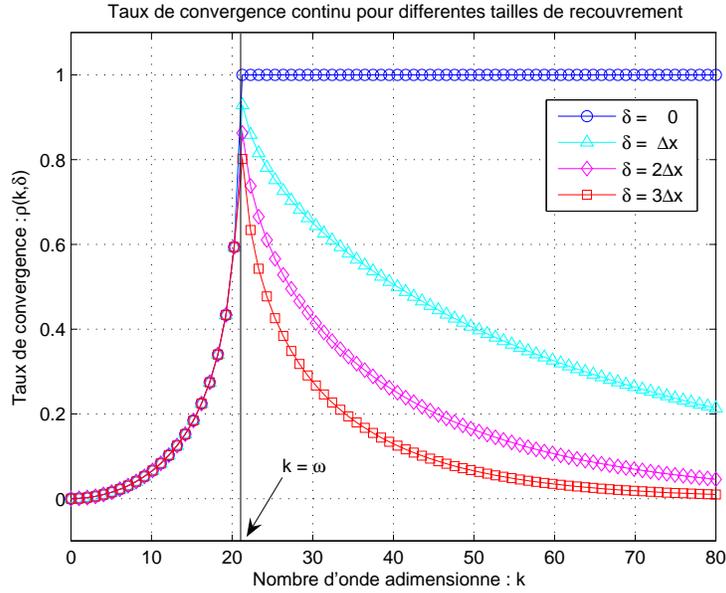


FIG. 2.3 – Evolution du taux de convergence à $f = 1$ GHz pour différentes tailles de recouvrement. Algorithme de Schwarz en continu.

3.2.2.3 Analyse de convergence de l'algorithme en discret

On étudie ici la convergence de l'algorithme de Schwarz avec ou sans recouvrement appliqué aux équations de Maxwell 2D discrétisées par le schéma volumes finis centrés décrit dans les sous-sections 2.2.1 et 2.3.1 du chapitre 2 et adapté au cas de maillages quadrangulaires. Commençons par préciser les éléments nécessaires à la discrétisation. Le domaine $\Omega = \mathbf{R}^2$ est partitionné en cellules $(C_{jl})_{(j,l) \in \mathbb{Z}^2}$ homogènes de taille $\Delta x \times \Delta y$ données par $[(j-1)\Delta x, j\Delta x] \times [(l-1)\Delta y, l\Delta y]$. Les normales unitaires sortantes à chaque cellule auront par conséquent pour coordonnées $(\pm 1, 0)^t$ ou $(0, \pm 1)^t$.

On considère une décomposition en deux sous-domaines Ω_1 et Ω_2 tels que :

$$\bar{\Omega}_1 = \overline{\bigcup_{\substack{j \leq J_2 \\ l \in \mathbb{Z}}} (C_{jl})} \quad \text{et} \quad \bar{\Omega}_2 = \overline{\bigcup_{\substack{j \geq J_1 \\ l \in \mathbb{Z}}} (C_{jl})}$$

avec $J_1 \leq J_2$. On suppose que $a = x_{J_1}$ et $b = x_{J_2}$ où x_{J_1} et x_{J_2} sont respectivement les abscisses des centres de gravité des cellules $C_{J_1 l}$ et $C_{J_2 l}$. Cette décomposition est représentée sur la figure 2.4 pour des sous-domaines recouvrants. En lien avec la section 2.3.1 du chapitre 2, la discrétisation des équations de Maxwell 2D utilise une formulation volumes finis reposant sur une fonction de flux centrés appliquée aux faces des cellules C_{jl} internes au domaine Ω . Cependant, la formulation d'une version discrète de l'algorithme de Schwarz (2.10) impose l'utilisation d'une fonction de flux décentrés à certaines des faces des cellules interfaces. On obtient ainsi l'algorithme de Schwarz discret suivant, écrit en termes du

vecteur erreur associé aux variables caractéristiques :

$$\begin{aligned}
\Omega_1 : \left\{ \begin{array}{l}
i\omega \mathbf{E}_{jl}^{1,p+1} + \Lambda \frac{\mathbf{E}_{j+1,l}^{1,p+1} - \mathbf{E}_{j-1,l}^{1,p+1}}{2\Delta x} + A_y \frac{\mathbf{E}_{j,l+1}^{1,p+1} - \mathbf{E}_{j,l-1}^{1,p+1}}{2\Delta y} = 0, j < J_2 - 1 \\
\left(i\omega l_3 + \frac{|\Lambda|}{2\Delta x} \right) \mathbf{E}_{jl}^{1,p+1} - \frac{\Lambda}{2\Delta x} \mathbf{E}_{j-1,l}^{1,p+1} + \frac{\Lambda^-}{\Delta x} \mathbf{E}_{j+1,l}^{1,p+1} \\
+ A_y \frac{\mathbf{E}_{j,l+1}^{1,p+1} - \mathbf{E}_{j,l-1}^{1,p+1}}{2\Delta y} = 0, j = J_2 - 1 \\
\Lambda^- \mathbf{E}_{jl}^{1,p+1} = \Lambda^- \mathbf{E}_{jl}^{2,p}, j = J_2,
\end{array} \right. \\
\Omega_2 : \left\{ \begin{array}{l}
i\omega \mathbf{E}_{jl}^{2,p+1} + \Lambda \frac{\mathbf{E}_{j+1,l}^{2,p+1} - \mathbf{E}_{j-1,l}^{2,p+1}}{2\Delta x} + A_y \frac{\mathbf{E}_{j,l+1}^{2,p+1} - \mathbf{E}_{j,l-1}^{2,p+1}}{2\Delta y} = 0, j > J_1 + 1 \\
\left(i\omega l_3 + \frac{|\Lambda|}{2\Delta x} \right) \mathbf{E}_{jl}^{2,p+1} + \frac{\Lambda}{2\Delta x} \mathbf{E}_{j+1,l}^{2,p+1} - \frac{\Lambda^+}{\Delta x} \mathbf{E}_{j-1,l}^{2,p+1} \\
+ A_y \frac{\mathbf{E}_{j,l+1}^{2,p+1} - \mathbf{E}_{j,l-1}^{2,p+1}}{2\Delta y} = 0, j = J_1 + 1 \\
\Lambda^+ \mathbf{E}_{jl}^{2,p+1} = \Lambda^+ \mathbf{E}_{jl}^{1,p}, j = J_1.
\end{array} \right.
\end{aligned} \tag{2.23}$$

Malheureusement, il est assez fastidieux d'exprimer le taux de convergence de l'algorithme (2.23). En effet, le schéma appliqué au voisinage de chaque interface n'étant pas symétrique, on est amené à résoudre un problème aux valeurs propres généralisé dont la solution est trop complexe. Aussi, dans l'optique d'obtenir une approximation du taux de convergence de l'algorithme discret, on choisit d'appliquer un flux centré pour toutes les cellule de Ω , ce qui conduit à l'algorithme de Schwarz modifié suivant :

$$\begin{aligned}
\Omega_1 : \left\{ \begin{array}{l}
i\omega \mathbf{E}_{jl}^{1,p+1} + \Lambda \frac{\mathbf{E}_{j+1,l}^{1,p+1} - \mathbf{E}_{j-1,l}^{1,p+1}}{2\Delta x} + A_y \frac{\mathbf{E}_{j,l+1}^{1,p+1} - \mathbf{E}_{j,l-1}^{1,p+1}}{2\Delta y} = 0, j < J_2 \\
e_{1,jl}^{1,p+1} = e_{1,jl}^{2,p}, j = J_2,
\end{array} \right. \\
\Omega_2 : \left\{ \begin{array}{l}
i\omega \mathbf{E}_{jl}^{2,p+1} + \Lambda \frac{\mathbf{E}_{j+1,l}^{2,p+1} - \mathbf{E}_{j-1,l}^{2,p+1}}{2\Delta x} + A_y \frac{\mathbf{E}_{j,l+1}^{2,p+1} - \mathbf{E}_{j,l-1}^{2,p+1}}{2\Delta y} = 0, j > J_1 \\
e_{3,jl}^{2,p+1} = e_{3,jl}^{1,p}, j = J_1,
\end{array} \right.
\end{aligned} \tag{2.24}$$

En s'inspirant maintenant de l'étude réalisée en continu, nous allons supposer que la transformée de Fourier discrète suivant la direction y de l'erreur satisfaisant (2.24), et lorsque $x \in \mathbb{R}$, est de la forme :

$$\begin{aligned}
(\hat{\mathbf{E}}_j)^{p,m} &= (\gamma_1)^{p,m} e^{-x_j \lambda_1^h(k)} \mathbf{V}_1^k(k) + (\gamma_2)^{p,m} e^{-x_j \lambda_2^h(k)} \mathbf{V}_2^h(k) \\
&= (\gamma_1)^{p,m} e^{-(j-\frac{1}{2})\Delta x \lambda_1^h(k)} \mathbf{V}_1^h(k) + (\gamma_2)^{p,m} e^{-(j-\frac{1}{2})\Delta x \lambda_2^h(k)} \mathbf{V}_2^h(k),
\end{aligned} \tag{2.25}$$

où la paire de vecteurs propres $(\mathbf{V}_1^h, \mathbf{V}_2^h)$ est à déterminer, et où la fréquence spatiale k dans la direction y est une suite réelle indexée implicitement par un entier q que l'on omettra pour alléger l'expression.

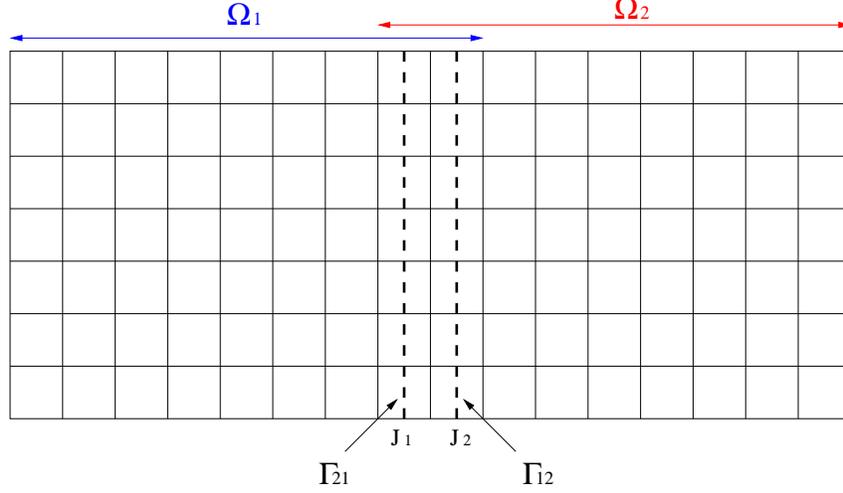


FIG. 2.4 – Décomposition de Ω en deux domaines recouvrants.
Cas d'un recouvrement de taille $\delta_h = \Delta x$.

On effectue alors une transformée de Fourier inverse de l'expression (2.25) si bien que la solution de notre schéma (2.24) doit nécessairement vérifier :

$$\begin{aligned}
(\mathbf{E}_{jl})^{p,m} &= \sum_{q \in \mathbb{Z}} e^{iky_l} (\hat{\mathbf{E}}_{jq})^{p,m} \\
&= \sum_{q \in \mathbb{Z}} e^{ik(l-\frac{1}{2})\Delta y} (\hat{\mathbf{E}}_{jq})^{p,m} \\
&= \sum_{q \in \mathbb{Z}} \left((\gamma_1)^{p,m} e^{-(j-\frac{1}{2})\Delta x \lambda_1^h(k)} e^{ik(l-\frac{1}{2})\Delta y} \mathbf{V}_1^h(k) + \right. \\
&\quad \left. (\gamma_2)^{p,m} e^{-(j-\frac{1}{2})\Delta x \lambda_2^h(k)} e^{ik(l-\frac{1}{2})\Delta y} \mathbf{V}_2^h(k) \right).
\end{aligned} \tag{2.26}$$

Par la suite, et comme nous l'avons déjà fait dans la sous-section précédente, $\tilde{\omega} = \omega/c$ sera désormais noté ω par abus de notation. Maintenant, en injectant la solution (2.26) dans le schéma principal de (2.24), nous devons satisfaire :

$$\left(i\omega l_3 - \Lambda \frac{e^{\Delta x \lambda_m^h(k)} - e^{-\Delta x \lambda_m^h(k)}}{2\Delta x} + A_y \frac{e^{ik\Delta y} - e^{-ik\Delta y}}{2\Delta y} \right) \mathbf{V}_m^h(k) = 0, \tag{2.27}$$

où $m = 1, 2$. Si on pose $L_m(k) = \frac{e^{\Delta x \lambda_m^h(k)} - e^{-\Delta x \lambda_m^h(k)}}{2\Delta x}$ et $\xi_k = \frac{\sin(k\Delta y)}{k\Delta y}$, alors (2.27) devient :

$$(i\omega l_3 - L(k)\Lambda + ik\xi_k A_y) \mathbf{V}^h(k) = 0, \tag{2.28}$$

où $L(k)$ peut être $L_1(k)$ ou $L_2(k)$ et $\mathbf{V}^h(k)$, $\mathbf{V}_1^h(k)$ ou $\mathbf{V}_2^h(k)$. Ensuite, toujours en procédant de manière analogue aux calculs effectués dans le cadre continu, on trouve en explicitant la deuxième équation du système (2.28) que :

$$V_2(k) = \frac{k\mu\xi_k}{\omega} (V_3(k) - V_1(k)), \tag{2.29}$$

où $\mathbf{V}^h = (V_1(k), V_2(k), V_3(k))^t$. Nous poserons, par abus de notation :

$$\mathbf{V}^h(k) = \begin{pmatrix} V_1(k) \\ V_3(k) \end{pmatrix}.$$

En se servant de (2.29) dans les premières et dernières relations de (2.28) on doit finalement résoudre :

$$\hat{M}_h \mathbf{V}^h(k) = L(k) \mathbf{V}^h(k) \text{ avec } \hat{M}_h = \frac{i}{2\omega} \begin{pmatrix} -2\omega^2 + k^2 \xi_k^2 & -k^2 \xi_k^2 \\ k^2 \xi_k^2 & 2\omega^2 - k^2 \xi_k^2 \end{pmatrix}. \quad (2.30)$$

Remarque 8 On rappelle que ω correspond en toute rigueur à $\tilde{\omega}$, et c'est pour cette raison que la composante $V_2(k)$ du vecteur propre $\mathbf{V}^h(k)$ ainsi que la matrice \hat{M}_h ont des expressions sensiblement différentes de leurs analogues du paragraphe 3.2.2.2.

Il se trouve que le problème (2.30) correspond à un problème aux valeurs propres pour lequel :

$$\lim_{\Delta x \rightarrow 0} L_m(k) = \lambda_m(k), \quad m = 1, 2 \quad \text{et} \quad \lim_{\Delta y \rightarrow 0} \hat{M}_h = \hat{M}, \quad (2.31)$$

où $\lambda_m(k)$ sont les valeurs propres trouvées dans le cas continu. Notre paire de vecteurs propres $\mathbf{V}_m^h(k)$ doit donc non seulement être solution de (2.30) mais aussi vérifier :

$$\lim_{\Delta y \rightarrow 0} \mathbf{V}_m^h(k) = \mathbf{V}_m(k),$$

où la paire de vecteurs $\mathbf{V}_m(k)$ correspond aux vecteurs propres exprimés en (2.14) lors de l'analyse en continu. Ces critères étant fixés, nous trouvons que les valeurs propres et les vecteurs propres associés sont :

$$\begin{aligned} \lambda_1^h(k) &= -\sqrt{k^2 \xi_k^2 - \omega^2} & , & \quad \mathbf{V}_1^h(k) = (\sqrt{k^2 \xi_k^2 - \omega^2} + i\omega, \sqrt{k^2 \xi_k^2 - \omega^2} - i\omega)^t, \\ \lambda_2^h(k) &= \sqrt{k^2 \xi_k^2 - \omega^2} & , & \quad \mathbf{V}_2^h(k) = (\sqrt{k^2 \xi_k^2 - \omega^2} - i\omega, \sqrt{k^2 \xi_k^2 - \omega^2} + i\omega)^t. \end{aligned} \quad (2.32)$$

Maintenant que les vecteurs propres ont été trouvés, on réutilise l'expression (2.25), et en invoquant la condition selon laquelle les solutions locales doivent être bornées à l'infini, nous trouvons que :

$$\left\{ \begin{aligned} \hat{\mathbf{E}}_j^{1,p} &= \gamma_1^p e^{-(j-\frac{1}{2})\Delta x \lambda_1^h(k)} \mathbf{V}_1^h(k) \\ &= \gamma_1^p e^{-(j-\frac{1}{2})\Delta x \lambda_1^h(k)} (\sqrt{k^2 \xi_k^2 - \omega^2} + i\omega, \sqrt{k^2 \xi_k^2 - \omega^2} - i\omega)^t, \\ \hat{\mathbf{E}}_j^{2,p} &= \gamma_2^p e^{-(j-\frac{1}{2})\Delta x \lambda_2^h(k)} \mathbf{V}_2^h(k) \\ &= \gamma_2^p e^{-(j-\frac{1}{2})\Delta x \lambda_2^h(k)} (\sqrt{k^2 \xi_k^2 - \omega^2} - i\omega, \sqrt{k^2 \xi_k^2 - \omega^2} + i\omega)^t. \end{aligned} \right. \quad (2.33)$$

En partant des conditions aux interfaces de (2.24) on obtient dans un premier temps :

$$\left\{ \begin{aligned} \gamma_1^{p+1} (\sqrt{k^2 \xi_k^2 - \omega^2} + i\omega) e^{(J_2 - \frac{1}{2})\Delta x \sqrt{k^2 \xi_k^2 - \omega^2}} &= \gamma_2^p (\sqrt{k^2 \xi_k^2 - \omega^2} - i\omega) e^{-(J_2 - \frac{1}{2})\Delta x \sqrt{k^2 \xi_k^2 - \omega^2}}, \\ \gamma_2^{p+1} (\sqrt{k^2 \xi_k^2 - \omega^2} + i\omega) e^{-(J_1 - \frac{1}{2})\Delta x \sqrt{k^2 \xi_k^2 - \omega^2}} &= \gamma_1^p (\sqrt{k^2 \xi_k^2 - \omega^2} - i\omega) e^{(J_1 - \frac{1}{2})\Delta x \sqrt{k^2 \xi_k^2 - \omega^2}}, \end{aligned} \right.$$

et ensuite :

$$\begin{cases} \gamma_1^{p+1} = \frac{\sqrt{k^2 \xi_k^2 - \omega^2 - i\omega}}{\sqrt{k^2 \xi_k^2 - \omega^2 + i\omega}} e^{-2(J_2 - \frac{1}{2})\Delta x \sqrt{k^2 \xi_k^2 - \omega^2}} \gamma_2^p, \\ \gamma_2^{p+1} = \frac{\sqrt{k^2 \xi_k^2 - \omega^2 - i\omega}}{\sqrt{k^2 \xi_k^2 - \omega^2 + i\omega}} e^{2(J_1 - \frac{1}{2})\Delta x \sqrt{k^2 \xi_k^2 - \omega^2}} \gamma_1^p, \end{cases} \quad (2.34)$$

dont on déduit :

$$\begin{cases} \gamma_1^{p+1} = \left(\frac{\sqrt{k^2 \xi_k^2 - \omega^2 - i\omega}}{\sqrt{k^2 \xi_k^2 - \omega^2 + i\omega}} \right)^2 e^{-2(J_2 - J_1)\Delta x \sqrt{k^2 \xi_k^2 - \omega^2}} \gamma_1^{p-1}, \\ \gamma_2^{p+1} = \left(\frac{\sqrt{k^2 \xi_k^2 - \omega^2 - i\omega}}{\sqrt{k^2 \xi_k^2 - \omega^2 + i\omega}} \right)^2 e^{-2(J_2 - J_1)\Delta x \sqrt{k^2 \xi_k^2 - \omega^2}} \gamma_2^{p-1}. \end{cases} \quad (2.35)$$

Si nous définissons un recouvrement discret $\delta_h = (J_2 - J_1)\Delta x$, alors le taux de convergence est finalement donné par :

$$\rho^h(k, \delta_h) = \left| \frac{\gamma_1^{p+1}}{\gamma_1^{p-1}} \right|^{1/2} = \left| \frac{\gamma_2^{p+1}}{\gamma_2^{p-1}} \right|^{1/2} = \left| \left(\frac{\sqrt{k^2 \xi_k^2 - \omega^2 - i\omega}}{\sqrt{k^2 \xi_k^2 - \omega^2 + i\omega}} \right) e^{-\delta_h \sqrt{k^2 \xi_k^2 - \omega^2}} \right|. \quad (2.36)$$

On se rend compte que l'expression du taux de convergence pour l'algorithme discret diffère sensiblement de l'expression du taux de convergence obtenu dans le cadre continu. Dans cette version, il dépend du signe de $k^2 \xi_k^2 - \omega^2$ suivant les valeurs de k et de Δy .

D'une façon similaire au cas continu, notre algorithme est convergent si et seulement si $\forall k \in (\pi\mathbb{Z}/L)$, $\rho^h(k, \delta_h) < 1$. Or, on a que :

$$\rho^h(k, \delta_h) = \begin{cases} \left| \frac{\sqrt{\omega^2 - k^2 \xi_k^2 - \omega}}{\sqrt{\omega^2 - k^2 \xi_k^2 + \omega}} \right|, & \text{si } |k \xi_k| < \omega, \\ e^{-\delta_h \sqrt{k^2 \xi_k^2 - \omega^2}}, & \text{si } |k \xi_k| \geq \omega. \end{cases} \quad (2.37)$$

On en déduit que l'algorithme est convergent pour les modes tels que $|k \xi_k| < \omega$ et ceci indépendamment du recouvrement (on voit même que celui-ci ne joue aucun rôle dans la convergence pour ces modes) mais il converge pour les modes qui vérifient $|k \xi_k| \geq \omega$ seulement en présence du recouvrement. Le seul mode parasite que l'on note \bar{k} qui peut poser problème est celui pour lequel $k \xi_k = \omega$, c'est-à-dire :

$$\bar{k} = \frac{\arcsin(\omega \Delta y)}{\Delta y}$$

Or, on ne considère en pratique que les modes supportés par la discrétisation. En effet, supposons dans ce qui suit que $\Omega = [0, 1] \times [0, L]$ et que la discrétisation de ce domaine est caractérisée par $\Delta x = 1/N$ et $\Delta y = h = L/N$ alors les fréquences k discrètes vérifient :

$$k \in \left(\frac{q\pi}{L} \right)_{q \in \llbracket 1, N \rrbracket},$$

et donc, on peut introduire k_{\min} et k_{\max} tels que :

$$\frac{\pi}{L} = k_{\min} \leq k \leq k_{\max} = \frac{\pi}{h}. \quad (2.38)$$

De plus, il existe un indice $\bar{q} \in \llbracket 1; N \rrbracket$ tel que :

$$\frac{\bar{q}\pi}{L} = k_- < \bar{k} < k_+ = \frac{(\bar{q} + 1)\pi}{L}. \quad (2.39)$$

Par conséquent et identiquement au cas continu, si on cherche à déterminer le taux de convergence global qui est donné par la norme infinie de ρ^h , on obtient :

$$\rho_{\infty}^h = \sup_k \rho(k, \delta_h) = \max \left(\left| \frac{\sqrt{\omega^2 - k_-^2 \xi_{k_-}^2} - \omega}{\sqrt{\omega^2 - k_-^2 \xi_{k_-}^2} + \omega} \right|, e^{-\delta_h \sqrt{k_+^2 \xi_{k_+}^2 - \omega^2}} \right). \quad (2.40)$$

Afin d'illustrer ce résultat, on propose de tracer les valeurs de ρ^h en fonction du nombre d'onde k et de la taille δ_h du recouvrement. On se limite aux modes appartenant à $[0, \frac{\pi}{2h}]$ qui est l'intervalle d'intérêt pour une comparaison avec les courbes obtenues dans le cas continu 2.2 et 2.3. En prenant un maillage de $\Omega = [0, 1] \times [0, L]$ avec $L = 1$ comprenant $N = 50$ points dans les deux directions d'espace et dont les cellules ont des paramètres caractéristiques $\varepsilon = \mu = 1$, la fréquence maximale f qui peut être représentée par le maillage est 1.5 GHz selon le critère $h \leq \lambda/10$ où λ est la longueur d'onde (ici, on admet que la discrétisation doit comporter au moins 10 cellules par longueur d'onde). Pour lever toute ambiguïté, reprenons dès à présent notre ancienne notation où $\tilde{\omega} = \omega/c$. Les relations nécessaires pour évaluer cette fréquence sont exprimées par :

$$h = \frac{L}{N}, \quad \lambda = \frac{c}{f}, \quad \omega = 2\pi f,$$

en variables non adimensionnées de sorte que :

$$h \leq \frac{\lambda}{10} \Rightarrow f \leq \frac{cN}{10L} \Rightarrow \tilde{\omega} \leq \frac{2\pi N}{10L}.$$

On propose donc de représenter les courbes du taux de convergence pour les mêmes fréquences que dans le cas continu afin de montrer l'analogie des différentes courbes. Pour $f = 400$ MHz nous avons $\tilde{\omega} \approx 8,38$ et $\bar{k} \approx 8,42$, et pour $f = 1$ GHz, $\tilde{\omega} \approx 20,94$ et $\bar{k} \approx 21,61$. Les courbes du taux de convergence pour l'algorithme discret sont représentées sur les figures 2.5 et 2.6.

On note un comportement identique entre les courbes 2.2 et 2.3 représentant le taux de convergence pour l'algorithme continu et les courbes 2.5 et 2.6. Par ailleurs, on constate effectivement que la valeur critique $k = \bar{k}$ pour laquelle l'algorithme ne convergerait pas n'est jamais atteinte, puisque $\bar{k} \notin (\pi\mathbb{Z}/L)$. Enfin, on peut vérifier graphiquement que pour une taille de recouvrement δ_h non-nulle, $\rho_{\infty}^h = \rho^h(3\pi, \delta_h)$ si $f = 400$ MHz et $\rho_{\infty}^h = \rho^h(7\pi, \delta_h)$ si $f = 1$ GHz.

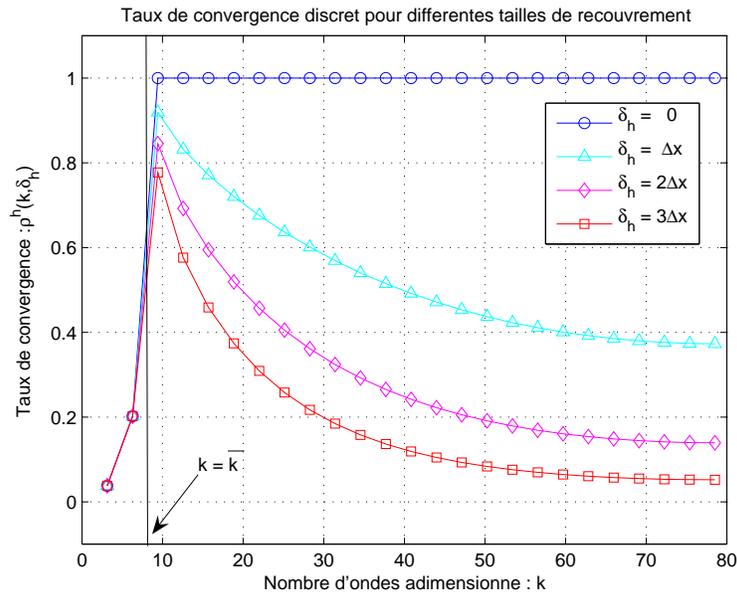


FIG. 2.5 – Evolution du taux de convergence à $f = 400$ MHz et $N = 50$ points pour différentes tailles de recouvrement. Algorithme de Schwarz en discret.

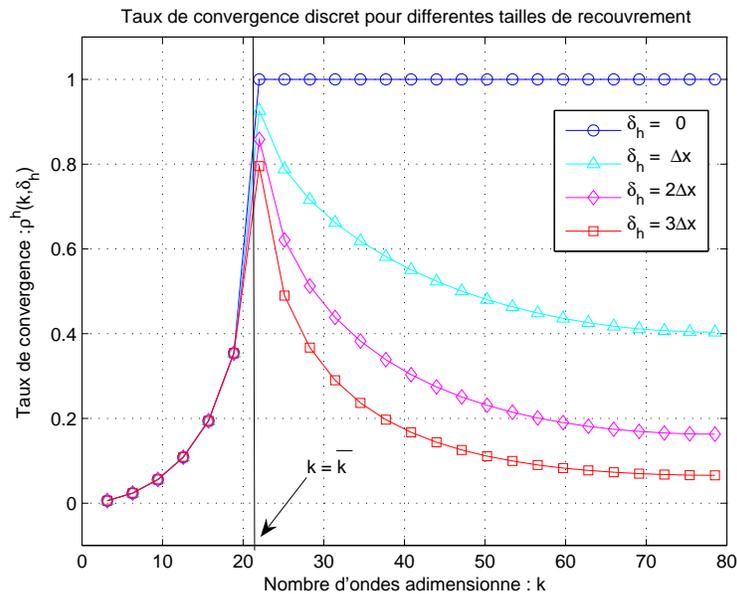


FIG. 2.6 – Evolution du taux de convergence à $f = 1$ GHz et $N = 50$ points pour différentes tailles de recouvrement. Algorithme de Schwarz en discret.

3.2.3 Mise en œuvre de l'algorithme de Schwarz sans recouvrement

Nous avons vu d'après l'expression du taux de convergence discret (2.37), que l'algorithme (2.23) ne converge pas pour un recouvrement nul. Cependant, il est possible d'assurer la convergence en faisant appel à une méthode par sous-espaces de Krylov [Saad, 1996] pour accélérer la résolution itérative. Nous montrerons aussi dans la section 3.3 qu'il est possible d'obtenir un algorithme de Schwarz sans recouvrement convergeant en utilisant des conditions d'interface optimisées. Dans ce qui suit, on propose donc de détailler la mise en œuvre de l'algorithme de Schwarz sans recouvrement. Plus précisément, nous allons montrer comment aboutir à un algorithme sans recouvrement en partant de l'algorithme de Schwarz (2.23). On se place ici dans le cadre discret du paragraphe 3.2.2.3. Ainsi, puisque nous considérons une formulation volumes finis en maillage quadrangulaire, on oriente par convention les normales suivant les directions positives des axes, si bien que si $M_{\mathbf{n}}$, $M_{\mathbf{n}}^+$ et $M_{\mathbf{n}}^-$ ont les expressions générales des relations (2.3) et (2.4), alors nous pouvons nous appuyer sur les matrices élémentaires suivantes :

$$M_r = M_{(1,0)} = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad M_u = M_{(0,1)} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$M_r^+ = M_{(1,0)}^+ = \frac{1}{2} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{pmatrix}, \quad M_r^- = M_{(1,0)}^- = \frac{1}{2} \begin{pmatrix} -1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & -1 \end{pmatrix},$$

$$|M_r| = M_r^+ - M_r^- = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

L'algorithme de Schwarz avec recouvrement (2.23) en termes du vecteur solution \mathbf{W} s'écrit pour des indices intérieurs J_1 et J_2 distincts (voir la figure 2.4) :

$$\Omega_1 : \begin{cases} i\omega \mathbf{W}_{jl}^{1,p+1} + M_r \frac{\mathbf{W}_{j+1,l}^{1,p+1} - \mathbf{W}_{j-1,l}^{1,p+1}}{2\Delta x} + M_u \frac{\mathbf{W}_{j,l+1}^{1,p+1} - \mathbf{W}_{j,l-1}^{1,p+1}}{2\Delta y} = \mathbf{f}_{jl}^1, j < J_2 - 1 \\ \left(i\omega \mathbf{l}_3 + \frac{|M_r|}{2\Delta x} \right) \mathbf{W}_{jl}^{1,p+1} - \frac{M_r}{2\Delta x} \mathbf{W}_{j-1,l}^{1,p+1} + \frac{M_r^-}{\Delta x} \mathbf{W}_{j+1,l}^{1,p+1} \\ \quad + M_u \frac{\mathbf{W}_{j,l+1}^{1,p+1} - \mathbf{W}_{j,l-1}^{1,p+1}}{2\Delta y} = \mathbf{f}_{jl}^1, j = J_2 - 1 \\ M_r^- \mathbf{W}_{jl}^{1,p+1} = M_r^- \mathbf{W}_{jl}^{2,p}, j = J_2, \end{cases} \quad (2.41)$$

$$\Omega_2 : \begin{cases} i\omega \mathbf{W}_{jl}^{2,p+1} + M_r \frac{\mathbf{W}_{j+1,l}^{2,p+1} - \mathbf{W}_{j-1,l}^{2,p+1}}{2\Delta x} + M_u \frac{\mathbf{W}_{j,l+1}^{2,p+1} - \mathbf{W}_{j,l-1}^{2,p+1}}{2\Delta y} = \mathbf{f}_{jl}^2, j > J_1 + 1 \\ \left(i\omega \mathbf{l}_3 + \frac{|M_r|}{2\Delta x} \right) \mathbf{W}_{jl}^{2,p+1} + \frac{M_r}{2\Delta x} \mathbf{W}_{j+1,l}^{2,p+1} - \frac{M_r^+}{\Delta x} \mathbf{W}_{j-1,l}^{2,p+1} \\ \quad + M_u \frac{\mathbf{W}_{j,l+1}^{2,p+1} - \mathbf{W}_{j,l-1}^{2,p+1}}{2\Delta y} = \mathbf{f}_{jl}^2, j = J_1 + 1 \\ M_r^+ \mathbf{W}_{jl}^{2,p+1} = M_r^+ \mathbf{W}_{jl}^{1,p}, j = J_1. \end{cases}$$

L'obtention d'un algorithme sans recouvrement passe par une manipulation purement algébrique dont le point de départ consiste à supposer que les cellules J_1 et J_2 sont confondues. Notons $J = J_1 = J_2$. Sur le plan strictement géométrique, on obtient la décomposition de la figure 2.7. L'interface Γ est la droite $x = x_J$ où x_J est l'abscisse du centre de gravité des cellules C_{Jl} . Donc, du point de vue du maillage, cette interface est la bande de cellules C_{Jl} pour $1 \leq l \leq N$. Nous traduisons cela par l'algorithme suivant :

$$\Omega_1 : \begin{cases} i\omega \mathbf{W}_{jl}^{1,p+1} + M_r \frac{\mathbf{W}_{j+1,l}^{1,p+1} - \mathbf{W}_{j-1,l}^{1,p+1}}{2\Delta x} + M_u \frac{\mathbf{W}_{j,l+1}^{1,p+1} - \mathbf{W}_{j,l-1}^{1,p+1}}{2\Delta y} = \mathbf{f}_{jl}^1, j < J \\ \left(i\omega l_3 + \frac{|M_r|}{2\Delta x} \right) \mathbf{W}_{jl}^{1,p+1} - \frac{M_r}{2\Delta x} \mathbf{W}_{j-1,l}^{1,p+1} + \frac{M_r^-}{\Delta x} \mathbf{W}_{j+1,l}^{1,p+1} \\ + M_u \frac{\mathbf{W}_{j,l+1}^{1,p+1} - \mathbf{W}_{j,l-1}^{1,p+1}}{2\Delta y} = \mathbf{f}_{jl}^1, j = J \\ M_r^- \mathbf{W}_{jl}^{1,p+1} = M_r^- \mathbf{W}_{jl}^{2,p}, j = J, \end{cases} \quad (2.42)$$

$$\Omega_2 : \begin{cases} i\omega \mathbf{W}_{jl}^{2,p+1} + M_r \frac{\mathbf{W}_{j+1,l}^{2,p+1} - \mathbf{W}_{j-1,l}^{2,p+1}}{2\Delta x} + M_u \frac{\mathbf{W}_{j,l+1}^{2,p+1} - \mathbf{W}_{j,l-1}^{2,p+1}}{2\Delta y} = \mathbf{f}_{jl}^2, j > J \\ \left(i\omega l_3 + \frac{|M_r|}{2\Delta x} \right) \mathbf{W}_{jl}^{2,p+1} + \frac{M_r}{2\Delta x} \mathbf{W}_{j+1,l}^{2,p+1} - \frac{M_r^+}{\Delta x} \mathbf{W}_{j-1,l}^{2,p+1} \\ + M_u \frac{\mathbf{W}_{j,l+1}^{2,p+1} - \mathbf{W}_{j,l-1}^{2,p+1}}{2\Delta y} = \mathbf{f}_{jl}^2, j = J \\ M_r^+ \mathbf{W}_{jl}^{2,p+1} = M_r^+ \mathbf{W}_{jl}^{1,p}, j = J. \end{cases}$$

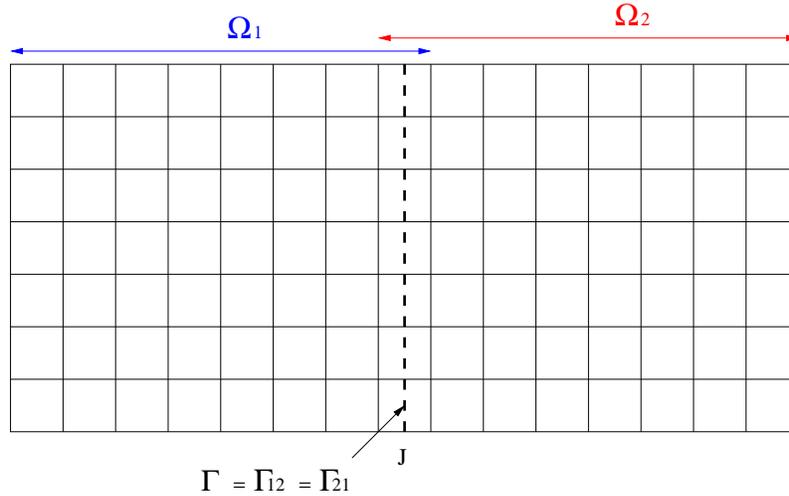


FIG. 2.7 – Algorithme de Schwarz discret sans recouvrement.

Concentrons-nous sur les relations associées aux cellules C_{Jl} que nous explicitons composante par composante, c'est-à-dire :

$$\Omega_1 : \left\{ \begin{array}{l} (i\omega + \frac{1}{2\Delta x})(E_z)_{Jl}^{1,p+1} + \frac{1}{2\Delta x}(H_y)_{J-1,l}^{1,p+1} - \frac{1}{\Delta x}[(E_z)_{J+1,l}^{1,p+1} + (H_y)_{J+1,l}^{1,p+1}] \\ \quad + \frac{(H_x)_{J,l+1}^{1,p+1} - (H_x)_{J,l-1}^{1,p+1}}{2\Delta y} = f_{E_z, Jl}^1, \quad (1.a) \\ i\omega(H_x)_{Jl}^{1,p+1} + \frac{(E_z)_{J,l+1}^{1,p+1} - (E_z)_{J,l-1}^{1,p+1}}{2\Delta y} = f_{H_x, Jl}^1, \quad (1.b) \\ (i\omega + \frac{1}{2\Delta x})(H_y)_{Jl}^{1,p+1} + \frac{1}{2\Delta x}(E_z)_{J-1,l}^{1,p+1} - \frac{1}{\Delta x}[(E_z)_{J+1,l}^{1,p+1} + (H_y)_{J+1,l}^{1,p+1}] = f_{H_y, Jl}^1, \quad (1.c) \\ (E_z)_{Jl}^{1,p+1} + (H_y)_{Jl}^{1,p+1} = (E_z)_{Jl}^{2,p} + (H_y)_{Jl}^{2,p}, \quad (1.d) \end{array} \right.$$

$$\Omega_2 : \left\{ \begin{array}{l} (i\omega + \frac{1}{2\Delta x})(E_z)_{Jl}^{2,p+1} - \frac{1}{2\Delta x}(H_y)_{J+1,l}^{2,p+1} - \frac{1}{\Delta x}[(E_z)_{J-1,l}^{2,p+1} - (H_y)_{J-1,l}^{2,p+1}] \\ \quad + \frac{(H_x)_{J,l+1}^{2,p+1} - (H_x)_{J,l-1}^{2,p+1}}{2\Delta y} = f_{E_z, Jl}^2, \quad (2.a) \\ i\omega(H_x)_{Jl}^{2,p+1} + \frac{(E_z)_{J,l+1}^{2,p+1} - (E_z)_{J,l-1}^{2,p+1}}{2\Delta y} = f_{H_x, Jl}^2, \quad (2.b) \\ (i\omega + \frac{1}{2\Delta x})(H_y)_{Jl}^{2,p+1} - \frac{1}{2\Delta x}(E_z)_{J+1,l}^{2,p+1} + \frac{1}{\Delta x}[(E_z)_{J-1,l}^{2,p+1} - (H_y)_{J-1,l}^{2,p+1}] = f_{H_y, Jl}^2, \quad (2.c) \\ (E_z)_{Jl}^{2,p+1} - (H_y)_{Jl}^{2,p+1} = (E_z)_{Jl}^{1,p} - (H_y)_{Jl}^{1,p}. \quad (2.d) \end{array} \right.$$

On constate que, compte tenu de la décomposition adoptée, les références aux termes $(E_z)_{J+1,l}^{1,p+1} + (H_y)_{J+1,l}^{1,p+1}$ dans le sous-domaine Ω_1 et $(E_z)_{J-1,l}^{2,p+1} - (H_y)_{J-1,l}^{2,p+1}$ dans le sous-domaine Ω_2 posent problème. On cherche donc à les éliminer. Pour cela, on procède aux manipulations algébriques suivantes :

$$\left\{ \begin{array}{l} (1.a) \\ (1.b) \\ (1.c) \\ (1.d) \end{array} \right. \rightarrow \left\{ \begin{array}{l} (1.a) - (1.c) \\ (1.b) \\ (1.d) \end{array} \right. \quad \text{et} \quad \left\{ \begin{array}{l} (2.a) \\ (2.b) \\ (2.c) \\ (2.d) \end{array} \right. \rightarrow \left\{ \begin{array}{l} (2.a) + (2.c) \\ (2.b) \\ (2.d) \end{array} \right.$$

qui conduisent aux équations suivantes :

$$\begin{aligned}
\Omega_1 : \left\{ \begin{array}{l}
(i\omega + \frac{1}{2\Delta x})[(E_z)_{Jl}^{1,p+1} - (H_y)_{Jl}^{1,p+1}] - \frac{1}{2\Delta x}[(E_z)_{J-1,l}^{1,p+1} - (H_y)_{J-1,l}^{1,p+1}] \\
\quad + \frac{(H_x)_{J,l+1}^{1,p+1} - (H_x)_{J,l-1}^{1,p+1}}{2\Delta y} = f_{E_z, Jl}^1 - f_{H_y, Jl}^1, \\
i\omega(H_x)_{Jl}^{1,p+1} + \frac{(E_z)_{J,l+1}^{1,p+1} - (E_z)_{J,l-1}^{1,p+1}}{2\Delta y} = f_{H_x, Jl}^1, \\
(E_z)_{Jl}^{1,p+1} + (H_y)_{Jl}^{1,p+1} = (E_z)_{Jl}^{2,p} + (H_y)_{Jl}^{2,p},
\end{array} \right. \\
\Omega_2 : \left\{ \begin{array}{l}
(i\omega + \frac{1}{2\Delta x})[(E_z)_{Jl}^{2,p+1} + (H_y)_{Jl}^{2,p+1}] - \frac{1}{2\Delta x}[(E_z)_{J+1,l}^{2,p+1} + (H_y)_{J+1,l}^{2,p+1}] \\
\quad + \frac{(H_x)_{J,l+1}^{2,p+1} - (H_x)_{J,l-1}^{2,p+1}}{2\Delta y} = f_{E_z, Jl}^2 - f_{H_y, Jl}^2, \\
i\omega(H_x)_{Jl}^{2,p+1} + \frac{(E_z)_{J,l+1}^{2,p+1} - (E_z)_{J,l-1}^{2,p+1}}{2\Delta y} = f_{H_x, Jl}^2, \\
(E_z)_{Jl}^{2,p+1} - (H_y)_{Jl}^{2,p+1} = (E_z)_{Jl}^{1,p} - (H_y)_{Jl}^{1,p}.
\end{array} \right.
\end{aligned} \tag{2.43}$$

On aboutit donc à des équations discrètes qui permettent de mettre à jour les composantes $\mathbf{W}(C_{Jl})$ dans le sous-domaine Ω_1 (respectivement dans le sous-domaine Ω_2) sans avoir recours à des valeurs associées aux cellules $\mathbf{W}(C_{J+1,l})$ (respectivement $\mathbf{W}(C_{J-1,l})$).

3.2.4 Formulation d'un système interface

Nous présentons ici une formulation alternative de l'algorithme de Schwarz qui conduit à la résolution d'un système linéaire dont les inconnues sont les degrés de liberté localisés sur les interfaces entre sous-domaines voisins. Pour cela, nous adoptons une démarche inspirée de [Magoulès et al., 2004] qui consiste à introduire des multiplicateurs de Lagrange et à résoudre le problème sur ces mêmes multiplicateurs induisant un opérateur matriciel augmenté.

Suivant la décomposition canonique en deux sous-domaines Ω_1 et Ω_2 de Ω considérée jusqu'à présent, l'algorithme continu (2.7) est reformulé comme suit :

Etape 1

$$\left\{ \begin{array}{ll} \mathcal{L}\mathbf{W}^{1,p+1} = \mathbf{f}^1 & \text{sur } \Omega_1, \\ \mathcal{B}_1(\mathbf{W}^{1,p+1}) = \lambda^{1,p} & \text{sur } \Gamma_{12}, \\ M_{\mathbf{n}}^-\mathbf{W}^{1,p+1} = M_{\mathbf{n}}^-\mathbf{W}^{\text{inc}} & \text{sur } \Gamma_1, \end{array} \right. \quad (2.44)$$

$$\left\{ \begin{array}{ll} \mathcal{L}\mathbf{W}^{2,p+1} = \mathbf{f}^2 & \text{sur } \Omega_2, \\ \mathcal{B}_2(\mathbf{W}^{2,p+1}) = \lambda^{2,p} & \text{sur } \Gamma_{21}, \\ M_{\mathbf{n}}^-\mathbf{W}^{2,p+1} = M_{\mathbf{n}}^-\mathbf{W}^{\text{inc}} & \text{sur } \Gamma_2. \end{array} \right.$$

Etape 2

$$\left\{ \begin{array}{ll} \lambda^{1,p+1} = \mathcal{B}_1(\mathbf{W}^{2,p+1}) & \text{sur } \Gamma_{12}, \\ \lambda^{2,p+1} = \mathcal{B}_2(\mathbf{W}^{1,p+1}) & \text{sur } \Gamma_{21}. \end{array} \right. \quad (2.45)$$

où les opérateurs interface \mathcal{B}_1 et \mathcal{B}_2 sont définis par :

$$\left\{ \begin{array}{ll} \mathcal{B}_1(\mathbf{W}) = E_z + H_y \\ \mathcal{B}_2(\mathbf{W}) = E_z - H_y \end{array} \right. \quad (2.46)$$

L'algorithme de Schwarz (2.44)-(2.45) peut ainsi se réinterpréter comme un algorithme de point fixe suivant les variables (λ^1, λ^2) en résolvant le système :

$$\left\{ \begin{array}{ll} \lambda^1 = \mathcal{B}_1(\mathbf{W}^2(\lambda^2, \mathbf{f}^2)), \\ \lambda^2 = \mathcal{B}_2(\mathbf{W}^1(\lambda^1, \mathbf{f}^1)). \end{array} \right. \quad (2.47)$$

où $\mathbf{W}^j = \mathbf{W}^j(\lambda^j, \mathbf{f}^j)$ est solution du problème :

$$\left\{ \begin{array}{ll} \mathcal{L}\mathbf{W}^j = \mathbf{f}^j & \text{dans } \Omega_j, \\ \mathcal{B}_j(\mathbf{W}^j) = \lambda^j & \text{sur } \Gamma_{jl}, \\ M_{\mathbf{n}}^-\mathbf{W}^j = M_{\mathbf{n}}^-\mathbf{W}^{\text{inc}} & \text{sur } \Gamma_j, \end{array} \right.$$

pour $j, l \in \{1, 2\}$ mais $l \neq j$.

Utilisant les matrices B_j et R_j respectivement relatives aux opérateurs de trace \mathcal{B}_j et aux relèvements

continus, la discrétisation du système (2.47) à partir du schéma volumes finis centrés, comme au paragraphe 3.2.2.3, conduit à la formulation du système matriciel :

$$\begin{pmatrix} A_1 & 0 & R_1 & 0 \\ 0 & A_2 & 0 & R_2 \\ 0 & -B_2 & I & 0 \\ -B_1 & 0 & 0 & I \end{pmatrix} \begin{pmatrix} \mathbf{W}^1 \\ \mathbf{W}^2 \\ \boldsymbol{\lambda}^1 \\ \boldsymbol{\lambda}^2 \end{pmatrix} = \begin{pmatrix} \mathbf{F}^1 \\ \mathbf{F}^2 \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \quad (2.48)$$

où, par abus de notation, $\mathbf{W}^{1,2}$ et $\boldsymbol{\lambda}^{1,2}$ désignent les vecteurs des inconnues discrètes respectivement internes aux sous-domaines Ω_1 et Ω_2 et interfaces. On peut appliquer au système (2.48) une technique de sous-structuration pour obtenir un système portant uniquement sur les variables interface $\boldsymbol{\lambda}$ (appelé système du complément de Schur) :

$$T\boldsymbol{\lambda} = \mathbf{d}, \quad (2.49)$$

où $\boldsymbol{\lambda} = (\boldsymbol{\lambda}^1, \boldsymbol{\lambda}^2)$ et la matrice T ainsi que le second membre \mathbf{d} sont donnés par :

$$T = \begin{pmatrix} I & B_2 A_2^{-1} R_2 \\ B_1 A_1^{-1} R_1 & I \end{pmatrix} \quad \text{et} \quad \mathbf{d} = \begin{pmatrix} B_2 A_2^{-1} \mathbf{F}^2 \\ B_1 A_1^{-1} \mathbf{F}^1 \end{pmatrix}. \quad (2.50)$$

Remarque 9 Dans le cas de sous-domaines non-recouvrants, on aurait bien entendu $\Gamma_{12} = \Gamma_{21} = \Gamma_{int}$, mais on aurait surtout que $R_j = B_j^t$. Nous appliquerons cette propriété par la suite, lors de l'étude de conditions d'interface optimisées.

On vérifie alors qu'une itération de l'algorithme de Schwarz (2.7) est équivalente à l'itération suivante agissant sur les variables interface :

$$\boldsymbol{\lambda}^{p+1} = (I - T)\boldsymbol{\lambda}^p + \mathbf{d}. \quad (2.51)$$

où la matrice T n'est jamais assemblée explicitement. Notons enfin que l'on peut accélérer la résolution du système (2.49) en lui appliquant une méthode par sous-espaces de Krylov comme GMRES [Saad, 1996].

3.2.5 Résultats numériques

Les résultats numériques qui suivent ont été obtenus avec le logiciel Matlab. Dans un premier temps et dans le but d'illustrer les résultats trouvés précédemment lors de l'étude 2D, nous choisissons de nouveau les paramètres $\varepsilon = \mu = 1$, $f = 400$ MHz et $f = 1$ GHz. Le domaine de calcul Ω est le carré unité $[0, 1]^2$ avec des conditions absorbantes posées sur $\partial\Omega$. On étudie ici le comportement de l'algorithme de Schwarz avec un recouvrement $\Delta x \leq \delta_h \leq 3\Delta x$. Le cas test modèle considéré est la propagation suivant la droite (Oy) d'une onde plane dans le vide définie par :

$$\begin{pmatrix} E \\ \mathbf{H} \end{pmatrix} = \begin{pmatrix} E_0 \\ \mathbf{H}_0 \end{pmatrix} e^{i\mathbf{k} \cdot \mathbf{X}},$$

où

$$\mathbf{X} = \begin{pmatrix} x \\ y \end{pmatrix}, \quad \mathbf{k} = \frac{2\pi f}{c} \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad E = E_0, \quad \mathbf{H} = \begin{pmatrix} H_0 \\ 0 \end{pmatrix} \quad \text{avec} \quad H_0 = \frac{E_0}{z_0} \quad \text{où} \quad z_0 = \sqrt{\frac{\mu_0}{\varepsilon_0}},$$

et on vérifie que cette onde plane est solution des équations de Maxwell.

Dans les test numériques, on utilise le critère d'arrêt suivant :

$$\frac{\|\boldsymbol{\lambda}^p - \boldsymbol{\lambda}^{p-1}\|_\infty}{\|\boldsymbol{\lambda}^1 - \boldsymbol{\lambda}^0\|_\infty} \leq \epsilon. \quad (2.52)$$

3.2.5.1 Validations préliminaires

Afin de valider l'algorithme de Schwarz, on propose d'approcher la solution nulle sur un maillage de $N = 50$ points et pour une fréquence $f = 400$ MHz. Pour cela, on initialise le second membre à 0, de même que les solutions locales W_1 et W_2 , et on applique l'algorithme de Schwarz en initialisant W_Γ par les modes qui déterminent la convergence théorique illustrée sur la figure 2.5. Plus précisément, ces modes sont les harmoniques $\sin(j\pi y_j)$ où $j \in \llbracket 1; N \rrbracket$, et on vise à retrouver une convergence dont le comportement est similaire à celui de la figure 2.5 pour un recouvrement donné. En prenant par exemple $\delta_h = 2h$, on obtient le profil de la figure 2.8 qui, bien que basé sur des valeurs différentes du taux de convergence effectif, est directement comparable au profil théorique.

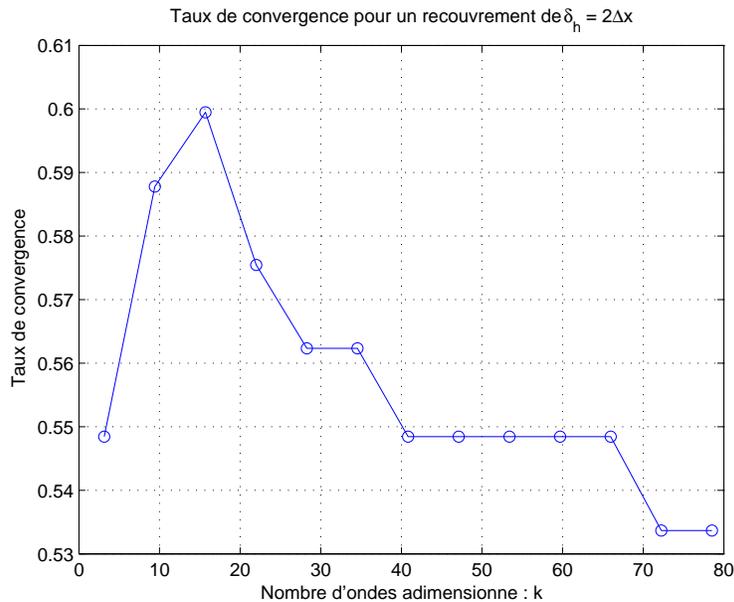


FIG. 2.8 – Taux de convergence de l'algorithme de Schwarz en discret. Nombre d'onde $k \in (j\pi)_{j \in \llbracket 1; \lfloor \frac{N}{2} \rrbracket \rrbracket}$, $f = 400$ MHz, $N = 50$ et $\delta_h = 2\Delta x$.

3.2.5.2 Etude numérique de la convergence

On se place dorénavant dans le cas d'un second membre correspondant à une onde plane. Pour les différents maillages considérés, on obtient les nombres d'itérations effectifs de la table 3.1. Ces données sont complétées par les courbes de convergence des figures 2.9 à 2.12 (on se limite au cas du maillage défini par $N = 200$). On observe que l'algorithme est très sensible à la finesse du maillage pour $\delta_h = \Delta x$ alors que cette dépendance est moins notable pour $\delta_h = 2\Delta x$ ou $3\Delta x$. Ces résultats montrent aussi

qu'il n'y a pas grand intérêt à utiliser un recouvrement de taille $\delta_h = 3\Delta x$. Enfin, on constate que l'algorithme se comporte mieux lorsque la fréquence augmente.

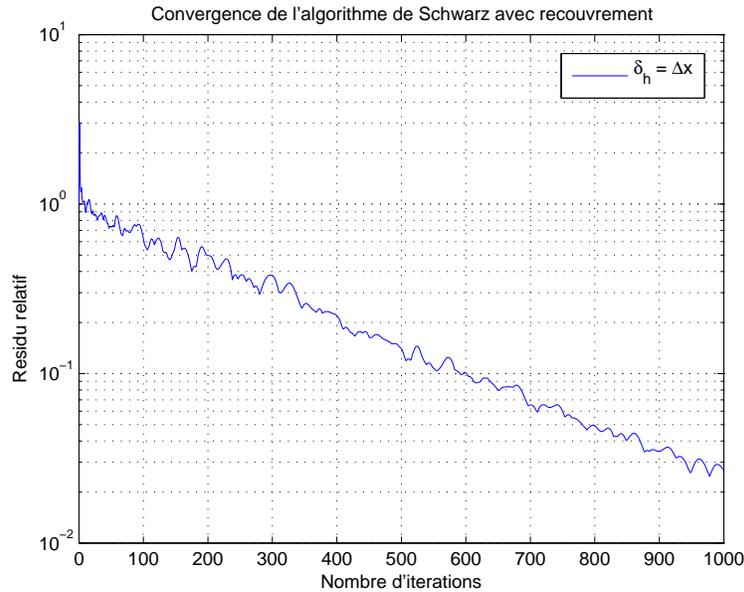


FIG. 2.9 – Evolution du résidu relatif (2.52) en fonction du nombre d'itérations. Algorithme de Schwarz en discret pour $\delta_h = \Delta x$ ($N = 200$ et $f = 400$ MHz).

δ_h		Δx	$2\Delta x$	$3\Delta x$
$f = 400$ MHz	$N = 50$	270	25	24
	$N = 100$	973	38	38
	$N = 200$	$\gg 1000$	60	60
$f = 1$ GHz	$N = 50$	101	23	23
	$N = 100$	199	30	30
	$N = 200$	655	43	43

TAB. 3.1 – Nombre d'itérations effectif en fonction du recouvrement pour différentes tailles de maillage et différentes fréquences.

3.2.5.3 Accélération par la méthode GMRES

On se propose ici d'évaluer le gain lié à l'utilisation de la méthode GMRES [Saad and Schultz, 1986] pour l'algorithme de Schwarz basé sur une taille de recouvrement égale à une cellule. On considère les maillages et les paramètres physiques énoncés plus haut, et on utilise la variante réinitialisée de la méthode GMRES avec une taille de l'espace de Krylov fixée à 5. On obtient les nombres d'itérations effectifs de la table 3.2 et les courbes de convergence des figures 2.13 et 2.14 (pour le maillage défini par $N = 200$). On conclut bien sûr au vu de ces courbes et de la table 3.2 la nette amélioration ainsi que la robustesse de l'algorithme de Schwarz accéléré par l'algorithme GMRES. On remarque encore une fois que cet algorithme converge mieux pour des fréquences élevées.

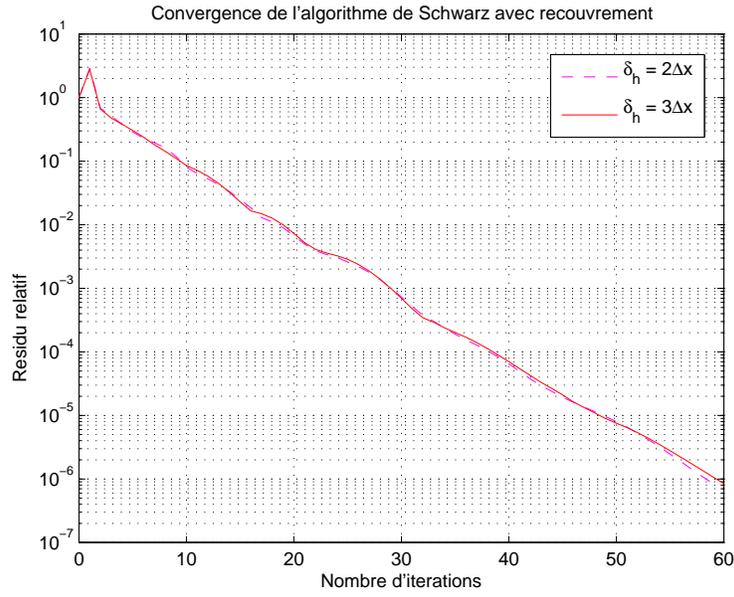


FIG. 2.10 – Evolution du résidu relatif (2.52) en fonction du nombre d'itérations. Algorithme de Schwarz en discret pour $\delta_h = 2\Delta x$ et $\delta_h = 3\Delta x$ ($N = 200$ et $f = 400$ MHz).

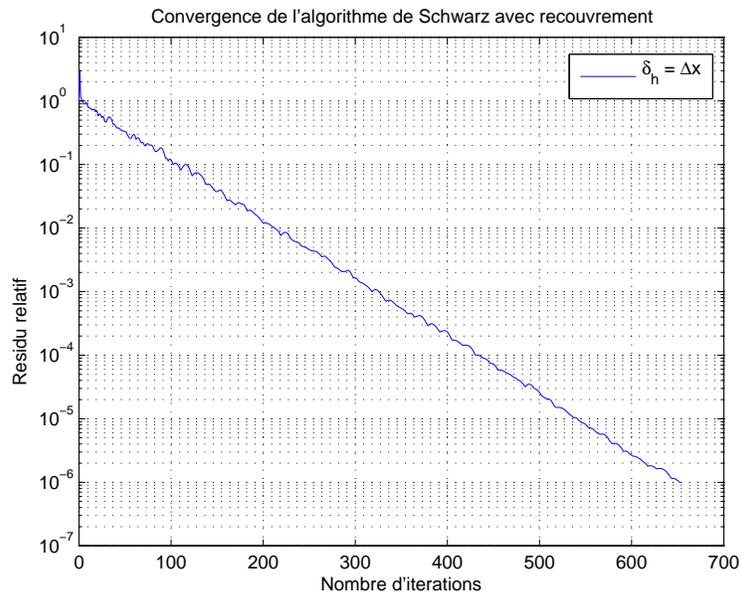


FIG. 2.11 – Evolution du résidu relatif (2.52) en fonction du nombre d'itérations. Algorithme de Schwarz en discret pour $\delta_h = \Delta x$ ($N = 200$ et $f = 1$ GHz).

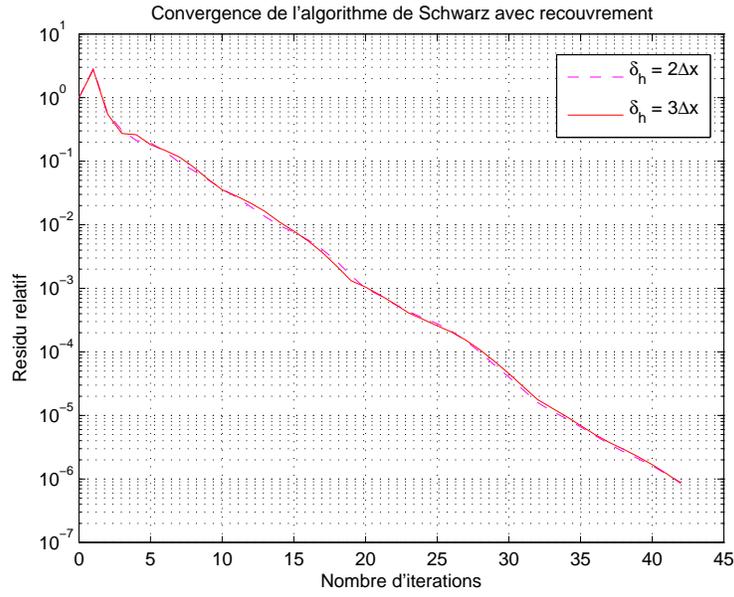


FIG. 2.12 – Evolution du résidu relatif (2.52) en fonction du nombre d'itérations. Algorithme de Schwarz en discret pour $\delta_h = 2\Delta x$ et $\delta_h = 3\Delta x$ ($N = 200$ et $f = 1$ GHz).

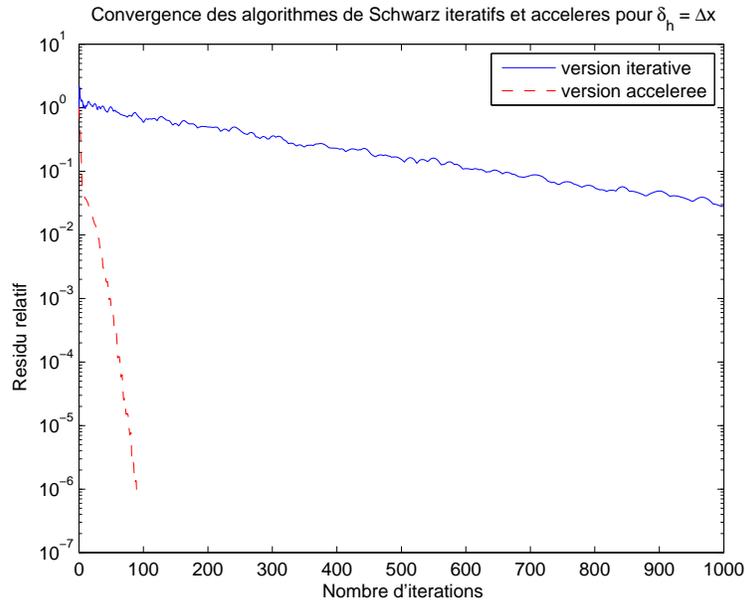


FIG. 2.13 – Résolution du système interface Evolution du résidu relatif (2.52) en fonction du nombre d'itérations, $N = 200$ et $f = 400$ MHz

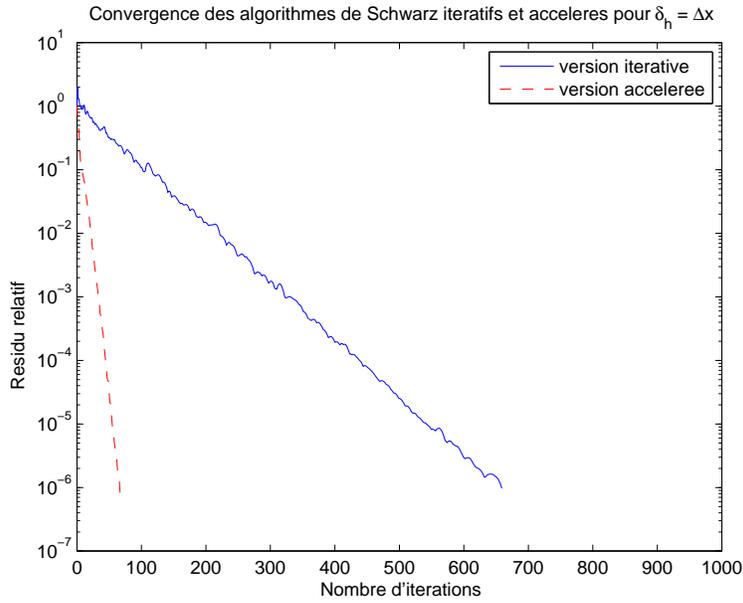


FIG. 2.14 – Résolution du système interface

Norme $\|\cdot\|_\infty$ du résidu relatif (2.52) en fonction du nombre d'itérations, $N = 200$ et $f = 1$ GHz.

		Version itérative	Version accélérée
$f = 400$ MHz	$N = 50$	268	41
	$N = 100$	989	63
	$N = 200$	$\gg 1000$	90
$f = 1$ GHz	$N = 50$	102	30
	$N = 100$	197	42
	$N = 200$	659	67

TAB. 3.2 – Nombre d'itérations effectifs par les versions itératives et accélérées pour différentes tailles de maillages et différentes fréquences ($\delta_h = \Delta x$).

3.3 Conditions d'interface optimisées en 2D

Dans l'optique d'accélérer la résolution itérative, nous nous intéressons maintenant à la formulation de conditions d'interface optimisées plus performantes que les conditions d'interface naturelles. Ces conditions d'interface optimisées approchent de manière optimale dans un sens à préciser, les conditions absorbantes exactes. Une première étape est donc de déterminer la forme des conditions absorbantes exactes, de les appliquer en tant que condition d'interface dans un algorithme de Schwarz, et d'en déduire le taux de convergence par une analyse de Fourier. Ensuite, ces conditions d'interface étant non-locales, nous étudions des approximations polynomiales que l'on optimise en vue d'accélérer la convergence de l'algorithme de Schwarz. Ce faisant, nous suivons et adaptions à notre contexte une démarche que l'on retrouve en particulier dans [Gander et al., 2002], [Alonso-Rodriguez and Gerardo-Giorda, 2004] et [Gander, 2006].

3.3.1 Conditions absorbantes exactes

L'objectif est ici d'exprimer la forme des conditions absorbantes exactes. Pour cela, considérons les équations de Maxwell (2.2) exprimées en variables caractéristiques (2.8) sur un domaine $\Omega =]a, b[\times \mathbf{R}$ et munies de conditions aux limites posées sur $\Gamma_1 = \{b\} \times \mathbf{R}$ et $\Gamma_2 = \{a\} \times \mathbf{R}$:

$$\begin{cases} \tilde{\mathcal{L}}\mathbf{U} = \tilde{\mathbf{f}} & \text{sur } \Omega, \\ u_1 + \mathcal{S}_1 u_3 = 0 & \text{sur } \Gamma_1, \\ u_3 + \mathcal{S}_2 u_1 = 0 & \text{sur } \Gamma_2. \end{cases} \quad (3.53)$$

Les conditions aux limites sont cherchées sous cette forme sachant que S_1 et S_2 sont deux opérateurs différentiels agissant dans la direction tangentielle à l'interface. Pour que ces conditions expriment des conditions absorbantes exactes, il faut déterminer S_1 et S_2 de sorte que la solution \mathbf{U} sur Ω soit la restriction à Ω de la solution $\bar{\mathbf{U}}$ sur \mathbf{R}^2 : $\mathbf{U} = \bar{\mathbf{U}}|_{\Omega}$. Ceci nous amène à démontrer le lemme suivant :

Lemme 5 *Si les opérateurs S_1 et S_2 ont le symbole de Fourier :*

$$\hat{S}_1 = \hat{S}_2 = -\frac{\sqrt{k^2 - \omega^2} - i\omega}{\sqrt{k^2 - \omega^2} + i\omega} := \hat{S}, \quad (3.54)$$

alors la solution du problème aux limites (3.53) sur Ω coïncide avec la restriction à Ω de la solution sur \mathbf{R}^2 .

Preuve Si on note $\mathbf{E} = \mathbf{U} - \bar{\mathbf{U}}|_{\Omega}$ le vecteur erreur, celui-ci vérifie par linéarité le système :

$$\begin{cases} \tilde{\mathcal{L}}\mathbf{E} = 0 & \text{sur } \Omega, \\ e_1 + \mathcal{S}_1 e_3 = 0 & \text{sur } \Gamma_1, \\ e_3 + \mathcal{S}_2 e_1 = 0 & \text{sur } \Gamma_2. \end{cases} \quad (3.55)$$

mais $\mathbf{E} = 0$ sur Ω si les conditions limites de (3.55) sont les conditions absorbantes exactes. En passant en variables de Fourier suivant y , privons \mathbf{E} de sa composante e_2 et notons en conséquence $\hat{\mathbf{E}} = (\hat{e}_1, \hat{e}_3)^t$ qui doit vérifier sur Ω selon (2.13) :

$$\hat{\mathbf{E}}(x, k) = \gamma_1 e^{x\sqrt{k^2 - \bar{\omega}^2}} \mathbf{V}_1(k) + \gamma_2 e^{-x\sqrt{k^2 - \bar{\omega}^2}} \mathbf{V}_2(k) \text{ avec } (\gamma_1, \gamma_2) \in \mathbb{C}^2, \quad (3.56)$$

où :

$$\begin{cases} \mathbf{V}_1(k) &= (\sqrt{k^2 - \tilde{\omega}^2} + i\tilde{\omega}, \sqrt{k^2 - \tilde{\omega}^2} - i\tilde{\omega})^t, \\ \mathbf{V}_2(k) &= (\sqrt{k^2 - \tilde{\omega}^2} - i\tilde{\omega}, \sqrt{k^2 - \tilde{\omega}^2} + i\tilde{\omega})^t. \end{cases} \quad (3.57)$$

Dans un souci de clarté, nous omettrons par la suite le symbôle $\tilde{\omega}$ sur $\tilde{\omega} = \omega/c$ et garderons à l'esprit que \mathbf{V}_1 et \mathbf{V}_2 sont fonctions de k . En se servant maintenant de la relation (3.56) pour exprimer les conditions limites de (3.55), nous devons satisfaire les conditions :

$$\begin{cases} \hat{e}_3(a, k) + \hat{S}_2 \hat{e}_1(a, k) &= 0, \\ \hat{e}_1(b, k) + \hat{S}_1 \hat{e}_3(b, k) &= 0, \end{cases}$$

qui deviennent après développement :

$$\begin{cases} \gamma_1 e^{a\sqrt{k^2 - \omega^2}} (\sqrt{k^2 - \omega^2} - i\omega) + \gamma_2 e^{-a\sqrt{k^2 - \omega^2}} (\sqrt{k^2 - \omega^2} + i\omega) + \\ \hat{S}_2 \left(\gamma_1 e^{a\sqrt{k^2 - \omega^2}} (\sqrt{k^2 - \omega^2} + i\omega) + \gamma_2 e^{-a\sqrt{k^2 - \omega^2}} (\sqrt{k^2 - \omega^2} - i\omega) \right) = 0, \\ \gamma_1 e^{b\sqrt{k^2 - \omega^2}} (\sqrt{k^2 - \omega^2} + i\omega) + \gamma_2 e^{-b\sqrt{k^2 - \omega^2}} (\sqrt{k^2 - \omega^2} - i\omega) + \\ \hat{S}_1 \left(\gamma_1 e^{b\sqrt{k^2 - \omega^2}} (\sqrt{k^2 - \omega^2} - i\omega) + \gamma_2 e^{-b\sqrt{k^2 - \omega^2}} (\sqrt{k^2 - \omega^2} + i\omega) \right) = 0, \end{cases} \quad (3.58)$$

$$\Leftrightarrow \begin{cases} \gamma_1 e^{a\sqrt{k^2 - \omega^2}} \left(\sqrt{k^2 - \omega^2} - i\omega + \hat{S}_2 (\sqrt{k^2 - \omega^2} + i\omega) \right) + \\ \gamma_2 e^{-a\sqrt{k^2 - \omega^2}} \left(\sqrt{k^2 - \omega^2} + i\omega + \hat{S}_2 (\sqrt{k^2 - \omega^2} - i\omega) \right) = 0, \quad (*) \\ \gamma_1 e^{b\sqrt{k^2 - \omega^2}} \left(\sqrt{k^2 - \omega^2} + i\omega + \hat{S}_1 (\sqrt{k^2 - \omega^2} - i\omega) \right) + \\ \gamma_2 e^{-b\sqrt{k^2 - \omega^2}} \left(\sqrt{k^2 - \omega^2} - i\omega + \hat{S}_1 (\sqrt{k^2 - \omega^2} + i\omega) \right) = 0. \quad (**) \end{cases}$$

Deux choix sont alors possibles pour obtenir la solution triviale. On opte pour le choix suivant :

$$\hat{S}_1 = -\frac{\sqrt{k^2 - \omega^2} - i\omega}{\sqrt{k^2 - \omega^2} + i\omega},$$

et on a nécessairement grâce à (***) que $\gamma_1 = 0$. Ensuite, si on choisit de prendre $\hat{S}_1 = \hat{S}_2$, alors (*) nous permet de conclure que $\gamma_1 = \gamma_2 = 0$. ■

Notre but dans la sous-section suivante est d'utiliser ces conditions absorbantes en tant que conditions d'interface pour l'algorithme de Schwarz.

3.3.2 Algorithme de Schwarz avec conditions généralisées

Le cadre de travail est strictement le même que celui du paragraphe 3.2.2.2. L'algorithme de Schwarz écrit en variables caractéristiques suivant la direction normale à l'interface et utilisant les conditions

d'interface généralisées exprimées à l'aide des opérateurs \mathcal{S}_1 et \mathcal{S}_2 s'écrit donc :

$$\begin{cases} \tilde{\mathcal{L}}\mathbf{U}^{1,p+1} = 0 & \text{dans } \Omega_1, \\ u_1^{1,p+1} + \mathcal{S}_1 u_3^{1,p+1} = u_1^{2,p} + \mathcal{S}_1 u_3^{2,p} \text{ sur } x = b, \end{cases} \quad (3.59)$$

$$\begin{cases} \tilde{\mathcal{L}}\mathbf{U}^{2,p+1} = 0 & \text{dans } \Omega_2, \\ u_3^{2,p+1} + \mathcal{S}_2 u_1^{2,p+1} = u_3^{1,p} + \mathcal{S}_2 u_1^{1,p} \text{ sur } x = a. \end{cases}$$

Suivant le même principe que pour (2.10), cet algorithme s'écrit en terme d'erreur :

$$\begin{cases} \tilde{\mathcal{L}}\mathbf{E}^{1,p+1} = 0 & \text{dans } \Omega_1, \\ e_1^{1,p+1} + \mathcal{S}_1 e_3^{1,p+1} = e_1^{2,p} + \mathcal{S}_1 e_3^{2,p} \text{ sur } x = b, \end{cases} \quad (3.60)$$

$$\begin{cases} \tilde{\mathcal{L}}\mathbf{E}^{2,p+1} = 0 & \text{dans } \Omega_2 \\ e_3^{2,p+1} + \mathcal{S}_2 e_1^{2,p+1} = e_3^{1,p} + \mathcal{S}_2 e_1^{1,p} \text{ sur } x = a. \end{cases}$$

Maintenant, en reprenant à l'identique la démarche adoptée pour exprimer le taux de convergence de l'algorithme basé sur des conditions d'interface naturelles, nous effectuons une transformation de Fourier dans la direction parallèle à l'interface puis supposons que la solution est bornée à l'infini de sorte qu'on obtient identiquement à (2.15) :

$$\begin{cases} \hat{\mathbf{E}}^{1,p}(x, k) = \gamma_1^p e^{x\sqrt{k^2-\omega^2}} (\sqrt{k^2-\omega^2} + i\omega, \sqrt{k^2-\omega^2} - i\omega)^t, \\ \hat{\mathbf{E}}^{2,p}(x, k) = \gamma_2^p e^{-x\sqrt{k^2-\omega^2}} (\sqrt{k^2-\omega^2} - i\omega, \sqrt{k^2-\omega^2} + i\omega)^t. \end{cases} \quad (3.61)$$

Les conditions d'interface de (3.60) vérifient alors :

$$\begin{cases} \hat{e}_1^{1,p+1} + \hat{\mathcal{S}}_1 \hat{e}_3^{1,p+1} = \hat{e}_1^{2,p} + \hat{\mathcal{S}}_1 \hat{e}_3^{2,p}, \\ \hat{e}_3^{2,p+1} + \hat{\mathcal{S}}_2 \hat{e}_1^{2,p+1} = \hat{e}_3^{1,p} + \hat{\mathcal{S}}_2 \hat{e}_1^{1,p}, \end{cases}$$

$$\Leftrightarrow \begin{cases} \gamma_1^{p+1} e^{b\sqrt{k^2-\omega^2}} (\sqrt{k^2-\omega^2} + i\omega + \hat{\mathcal{S}}_1(\sqrt{k^2-\omega^2} - i\omega)) \\ = \gamma_2^p e^{-b\sqrt{k^2-\omega^2}} (\sqrt{k^2-\omega^2} - i\omega + \hat{\mathcal{S}}_1(\sqrt{k^2-\omega^2} + i\omega)), \\ \gamma_2^{p+1} e^{-a\sqrt{k^2-\omega^2}} (\sqrt{k^2-\omega^2} + i\omega + \hat{\mathcal{S}}_2(\sqrt{k^2-\omega^2} - i\omega)) \\ = \gamma_1^p e^{a\sqrt{k^2-\omega^2}} (\sqrt{k^2-\omega^2} - i\omega + \hat{\mathcal{S}}_2(\sqrt{k^2-\omega^2} + i\omega)). \end{cases} \quad (3.62)$$

Des relations (3.62), on déduit :

$$\begin{cases} \gamma_1^{p+1} = \gamma_2^p \left(\frac{\sqrt{k^2 - \omega^2} - i\omega + \hat{\mathcal{S}}_1(\sqrt{k^2 - \omega^2} + i\omega)}{\sqrt{k^2 - \omega^2} + i\omega + \hat{\mathcal{S}}_1(\sqrt{k^2 - \omega^2} - i\omega)} \right) e^{-2b\sqrt{k^2 - \omega^2}}, \\ \gamma_2^{p+1} = \gamma_1^p \left(\frac{\sqrt{k^2 - \omega^2} - i\omega + \hat{\mathcal{S}}_2(\sqrt{k^2 - \omega^2} + i\omega)}{\sqrt{k^2 - \omega^2} + i\omega + \hat{\mathcal{S}}_2(\sqrt{k^2 - \omega^2} - i\omega)} \right) e^{2a\sqrt{k^2 - \omega^2}}, \end{cases} \quad (3.63)$$

et ensuite, en posant $\delta = b - a$ et :

$$\begin{aligned} A &= \frac{\sqrt{k^2 - \omega^2} - i\omega + \hat{\mathcal{S}}_1(\sqrt{k^2 - \omega^2} + i\omega)}{\sqrt{k^2 - \omega^2} + i\omega + \hat{\mathcal{S}}_1(\sqrt{k^2 - \omega^2} - i\omega)}, \\ B &= \frac{\sqrt{k^2 - \omega^2} - i\omega + \hat{\mathcal{S}}_2(\sqrt{k^2 - \omega^2} + i\omega)}{\sqrt{k^2 - \omega^2} + i\omega + \hat{\mathcal{S}}_2(\sqrt{k^2 - \omega^2} - i\omega)}, \end{aligned}$$

on obtient :

$$\frac{\gamma_1^{p+1}}{\gamma_1^{p-1}} = \frac{\gamma_2^{p+1}}{\gamma_2^{p-1}} = AB e^{-2\delta\sqrt{k^2 - \omega^2}}.$$

Alors le taux de convergence ρ vérifie :

$$\rho(k, \delta) = \sqrt{|AB e^{-2\delta\sqrt{k^2 - \omega^2}}|}. \quad (3.64)$$

Si les opérateurs \mathcal{S}_1 et \mathcal{S}_2 correspondent à l'opérateur des conditions absorbantes exactes dont le symbole de Fourier est donné par (3.54), on remarque que l'algorithme converge en une seule itération puisque $\rho(k, \delta) \equiv 0$. Malheureusement, ce choix est difficilement exploitable dans la pratique puisque la racine carrée dans le symbole de Fourier de l'opérateur \mathcal{S} induit un opérateur non-local. Notre but est maintenant d'étudier plusieurs approximations de \mathcal{S} par des opérateurs locaux.

3.3.3 Conditions d'interface approchées

Dans la suite de notre étude, nous considérerons des conditions d'interface pour l'utilisation ultérieure d'un algorithme de Schwarz sans recouvrement qui est donc caractérisé par le taux de convergence :

$$\rho(k) = \sqrt{|AB|}. \quad (3.65)$$

Nous supposons que $\mathcal{S}_1 = \mathcal{S}_2$ et notre objectif est de trouver un opérateur \mathcal{S}_{app} local basé sur une approximation polynomiale de $\sqrt{k^2 - \omega^2}$ dont nous choisirons les coefficients afin de minimiser le taux de convergence (3.65). Notons toutefois que la situation où $\mathcal{S}_1 \neq \mathcal{S}_2$ est aussi envisageable (voir notamment [Gander et al., 2002]).

3.3.3.1 Approximation d'ordre 0

On introduit dans ce paragraphe un complexe q indépendant de k qui sera une approximation de $\sqrt{k^2 - \omega^2}$. Minimiser le taux de convergence consiste à trouver la solution ρ_{opt} du problème suivant :

$$\rho_{opt} = \min_{q \in \mathbb{C}} \left(\max_{k \in]k_{\min}, k_-[\cup]k_+, k_{\max}[} \rho(k, q) \right). \quad (3.66)$$

Puisque $\sqrt{k^2 - \omega^2} \sim q$, l'opérateur \mathcal{S}_{app} aura pour symbole de Fourier :

$$\hat{\mathcal{S}}_{app} = -\frac{q - i\omega}{q + i\omega}. \quad (3.67)$$

En remplaçant alors \hat{S} par $\hat{\mathcal{S}}_{app}$, nous trouvons que le taux de convergence s'écrit :

$$\rho(k, q) = \left| \frac{\sqrt{k^2 - \omega^2} - q}{\sqrt{k^2 - \omega^2} + q} \right|. \quad (3.68)$$

Or, le problème (3.66) a déjà été résolu dans [Gander et al., 2002] moyennant les hypothèses (où k_{\min} , k_{\max} et k_{\pm} sont donnés par (2.38) et (2.39)) :

$$\begin{cases} 2\omega^2 \leq k_-^2 + k_+^2, & k_- < \omega \\ 2\omega^2 > k_{\min}^2 + k_+^2, \\ 2\omega^2 < k_{\min}^2 + k_{\max}^2, \end{cases} \quad (3.69)$$

et sa solution a même été réutilisée dans [Alonso-Rodriguez and Gerardo-Giorda, 2004]. Si on pose $q = r + is$ où $(r, s) \in \mathbb{C}^2$ alors la solution (r^*, s^*) du problème (3.66) est donnée par :

$$r^* = s^* = \sqrt{\frac{\sqrt{\omega^2 - k_-^2} \sqrt{k_{\max}^2 - \omega^2}}{2}}, \quad (3.70)$$

et le taux de convergence correspondant s'écrit :

$$\rho_{opt} = \left(\frac{1 - \sqrt{2} \left(\frac{\omega^2 - k_-^2}{k_{\max}^2 - \omega^2} \right)^{\frac{1}{4}} + \sqrt{\frac{\omega^2 - k_-^2}{k_{\max}^2 - \omega^2}}}{1 + \sqrt{2} \left(\frac{\omega^2 - k_-^2}{k_{\max}^2 - \omega^2} \right)^{\frac{1}{4}} + \sqrt{\frac{\omega^2 - k_-^2}{k_{\max}^2 - \omega^2}}} \right)^{\frac{1}{2}}. \quad (3.71)$$

Pour confirmer l'optimalité de ce jeu de paramètres, nous réalisons une étude numérique du taux de convergence sur un domaine $\Omega = [0, 1] \times [0, 2]$ discrétisé avec un pas $h = 1/50$, la fréquence étant fixée $f = 400$ MHz. On considère donc un cadre numérique quasiment identique à celui adopté pour l'étude des conditions d'interface naturelles hormis la longueur du domaine suivant l'axe (Oy) qui ici est égale à 2 de manière à respecter la première hypothèse de (3.69). Les paramètres optimaux correspondants et le taux de convergence optimal ont pour valeurs :

$$r^* = s^* \approx 10,669 \quad \text{et} \quad \rho_{opt} \approx 0,583.$$

On peut alors par exemple représenter graphiquement les lignes de contour de

$$\rho_+(r, s) = \max_{k \in]k_{\min}, k_-[\cup]k_+, k_{\max}[} \rho(k, r, s),$$

pour différentes valeurs de (r, s) comme illustré sur la figure 3.15. Par ailleurs, on peut aussi représenter graphiquement les évolutions de $\rho(k, r^*, s)$ et $\rho(k, r, s^*)$ (voir les figures 3.16 et 3.17) pour différentes valeurs de k et constater effectivement l'optimalité des valeurs de (r^*, s^*) obtenues.

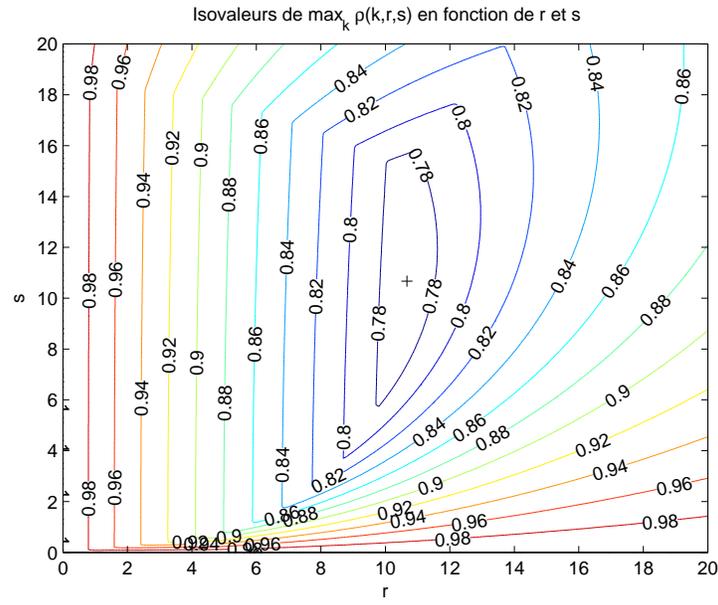


FIG. 3.15 – Isovaleurs de ρ_+ en fonction de (r, s) . Le symbole + identifie le point de coordonnées (r^*, s^*) .

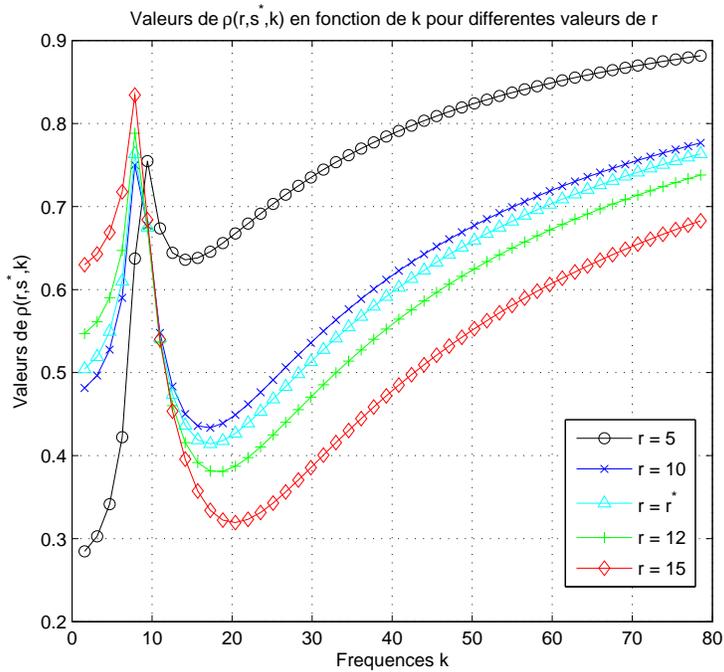


FIG. 3.16 – Evolution de $\rho(k, r, s^*)$ en fonction de k pour différentes valeurs de r .

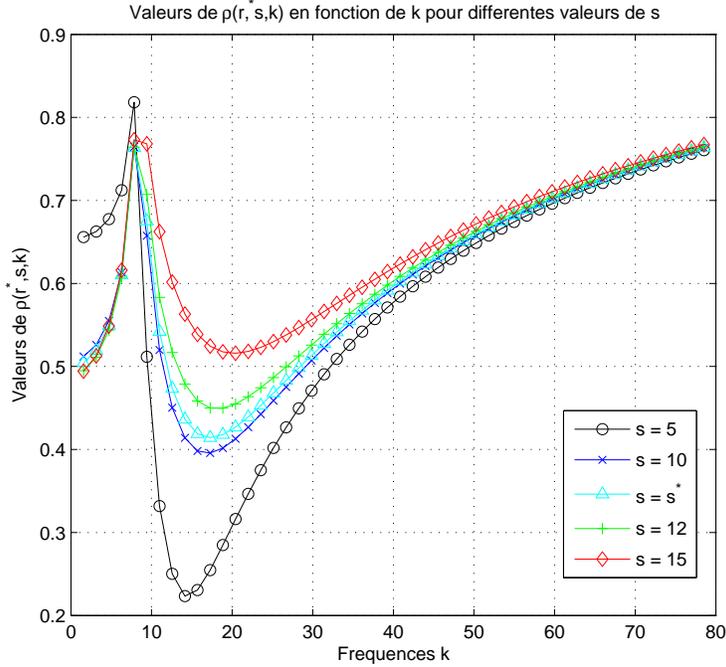


FIG. 3.17 – Evolution de $\rho(k, r^*, s)$ en fonction de k pour différentes valeurs de s .

3.3.3.2 Approximation d'ordre 2

Cette fois-ci, on étudie le choix $\sqrt{k^2 - \omega^2} \sim q_0 + q_2 k^2$ où $(q_0, q_2) \in \mathbb{C}^2$. Cette approximation implique :

$$\hat{S}_{app} = -\frac{q_0 + q_2 k^2 - i\omega}{q_0 + q_2 k^2 + i\omega}, \quad (3.72)$$

et le taux de convergence devient :

$$\rho(k, q_0, q_2) = \left| \frac{\sqrt{k^2 - \omega^2} - q_0 - q_2 k^2}{\sqrt{k^2 - \omega^2} + q_0 + q_2 k^2} \right|. \quad (3.73)$$

Nous devons donc résoudre cette fois-ci sur l'ouvert $\mathcal{C}^* = \{(z_1, z_2) \in \mathbb{C}^2 / z_2 \neq 0\}$ le problème d'optimisation suivant :

$$\rho_{opt} = \min_{(q_0, q_2) \in \mathcal{C}^*} \left(\max_{k \in]k_{\min}, k_-[\cup]k_+, k_{\max}[} \rho(k, q_0, q_2) \right). \quad (3.74)$$

Il est en effet inutile de chercher la solution sur \mathbb{C}^2 tout entier dans la mesure où le cas $q_2 = 0$ nous ramène au jeu de données étudié précédemment. Nous introduisons maintenant l'ouvert $\mathcal{U} = \{(z_1, z_2) \in \mathbb{C}^2 / z_1 + z_2 \neq 0\}$, ainsi que la fonction :

$$F : \mathcal{U} \longrightarrow \mathcal{C}^* \\ (\alpha, \beta) \longmapsto \left(\frac{\alpha\beta - \omega^2}{\alpha + \beta}, \frac{1}{\alpha + \beta} \right), \quad (3.75)$$

qui est non seulement holomorphe puisque c'est une fonction de fractions rationnelles dont les pôles n'appartiennent pas à \mathcal{U} , mais elle est également bijective puisque :

$$\forall (q_0, q_2) \in \mathcal{C}^*, F^{-1}(q_0, q_2) = \left(\frac{1 - \sqrt{1 - 4q_0q_2 - 4\omega^2q_2^2}}{2q_2}, \frac{1 + \sqrt{1 - 4q_0q_2 - 4\omega^2q_2^2}}{2q_2} \right).$$

De ce fait, en posant $(q_0, q_2) = F(\alpha, \beta)$, on trouve que le taux de convergence (3.65) s'écrit :

$$\rho(k, F(\alpha, \beta)) = \left| \left(\frac{\sqrt{k^2 - \omega^2} - \alpha}{\sqrt{k^2 - \omega^2} + \alpha} \right) \left(\frac{\sqrt{k^2 - \omega^2} - \beta}{\sqrt{k^2 - \omega^2} + \beta} \right) \right|.$$

Or, le problème :

$$\min_{(\alpha, \beta) \in \mathcal{U}} \left(\max_{k \in]k_{\min}, k_-[\cup]k_+, k_{\max}[} \rho(k, F(\alpha, \beta)) \right), \quad (3.76)$$

a déjà été résolu dans [Gander et al., 2002] et a inspiré [Alonso-Rodriguez and Gerardo-Giorda, 2004]. La solution est unique et les paramètres optimaux donnés par :

$$\alpha^* = i((\omega^2 - k_{\min}^2)(\omega^2 - k_-^2))^{\frac{1}{4}} \in i\mathbf{R} \quad \text{et} \quad \beta^* = ((k_{\max}^2 - \omega^2)(k_+^2 - \omega^2))^{\frac{1}{4}} \in \mathbf{R}, \quad (3.77)$$

permettent l'obtention du taux de convergence solution du problème (3.76) :

$$\max_{k \in]k_{\min}, k_-[\cup]k_+, k_{\max}[} \rho(k, F(\alpha^*, \beta^*)). \quad (3.78)$$

Le taux de convergence optimal a deux expressions suivant que k appartienne aux modes propagatifs ou aux modes évanescents :

$$\rho_{opt} = \begin{cases} \frac{(\omega^2 - k_-^2)^{\frac{1}{4}} - (\omega^2 - k_{\min}^2)^{\frac{1}{4}}}{(\omega^2 - k_-^2)^{\frac{1}{4}} + (\omega^2 - k_{\min}^2)^{\frac{1}{4}}} & \text{si } k \in]k_{\min}, k_-[, \\ \frac{(k_{\max}^2 - \omega^2)^{\frac{1}{4}} - (k_+^2 - \omega^2)^{\frac{1}{4}}}{(k_{\max}^2 - \omega^2)^{\frac{1}{4}} + (k_+^2 - \omega^2)^{\frac{1}{4}}} & \text{si } k \in]k_+, k_{\max}[. \end{cases} \quad (3.79)$$

Comme au paragraphe 3.3.3.1, on valide ces résultats en reproduisant les mêmes tests numériquement, sauf que maintenant l'optimisation porte sur les variables auxiliaires α et β si bien que nous représentons graphiquement la quantité :

$$\rho_+(\alpha, \beta) = \max_{k \in]k_{\min}, k_-[\cup]k_+, k_{\max}[} \rho(k, F(\alpha, \beta)),$$

Dans le cas présent, les paramètres optimaux prennent les valeurs suivantes :

$$\alpha^* \approx i4,898, \quad \beta^* \approx 18,362, \quad \text{et} \quad \rho_{opt} \approx 0,619,$$

et nous obtenons les représentations graphiques des figures 3.18, 3.19 et 3.20.

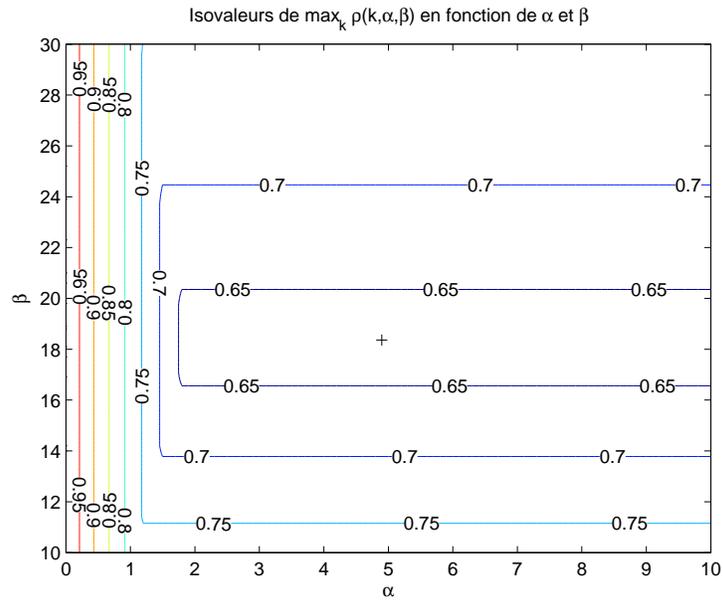


FIG. 3.18 – Isovaleurs de ρ_+ en fonction de (α, β) . Le symbole $+$ identifie le point de coordonnées (α^*, β^*) .

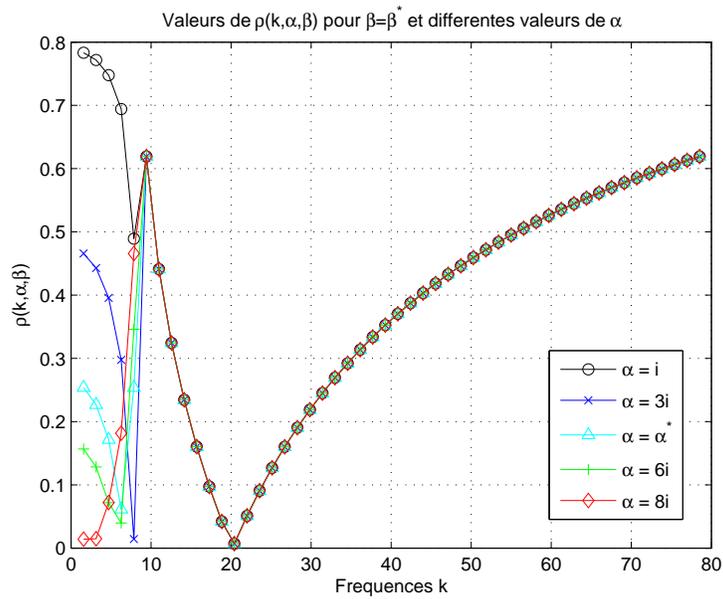


FIG. 3.19 – Evolution de $\rho(k, F(\alpha, \beta^*))$ en fonction de k pour différentes valeurs de α .

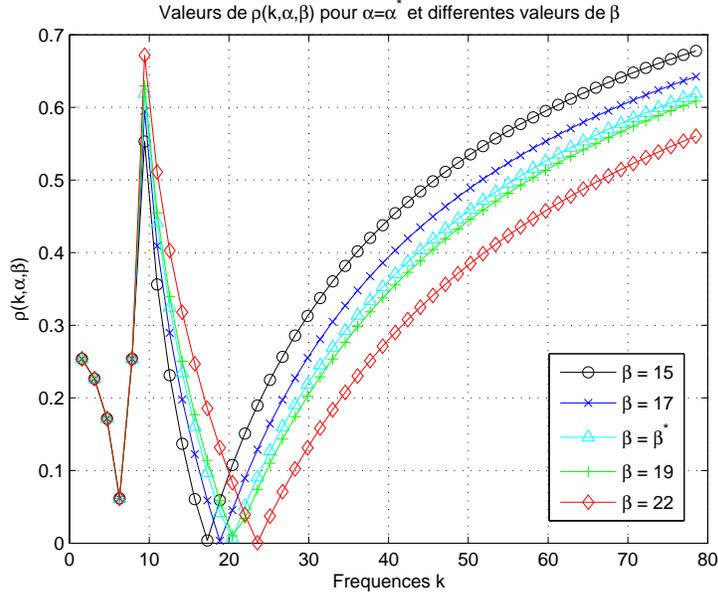


FIG. 3.20 – Evolution de $\rho(k, F(\alpha^*, \beta))$ en fonction de k pour différentes valeurs de β .

3.3.3.3 Approximations d'ordre 0 et d'ordre 2 modifiées

On peut aussi modifier l'expression du symbole de Fourier de l'opérateur \mathcal{S} (3.54) de la manière suivante :

$$\hat{\mathcal{S}} = - \left(\frac{\sqrt{k^2 - \omega^2} - i\omega}{\sqrt{k^2 - \omega^2} + i\omega} \right) \left(\frac{\sqrt{k^2 - \omega^2} + i\omega}{\sqrt{k^2 - \omega^2} + i\omega} \right) = - \frac{k^2}{(\sqrt{k^2 - \omega^2} + i\omega)^2}. \quad (3.80)$$

En cherchant une approximation d'ordre pair $\sqrt{k^2 - \omega^2} \sim q$ mais en partant cette fois-ci de l'opérateur de transmission écrit sous la forme (3.80), ceci nous amène à étudier le cas pour lequel :

$$\hat{\mathcal{S}}_{app} = - \frac{k^2}{k^2 - 2\omega^2 + 2i\omega q}, \quad (3.81)$$

ce qui donne au taux de convergence l'expression :

$$\rho(k, q) = \left| \frac{\omega(q - \sqrt{k^2 - \omega^2}) + i(q\sqrt{k^2 - \omega^2} - (k^2 - \omega^2))}{\omega(q + \sqrt{k^2 - \omega^2}) - i(q\sqrt{k^2 - \omega^2} + (k^2 - \omega^2))} \right|,$$

qui devient après quelques simplifications :

$$\rho(k, q) = \left| \left(\frac{\sqrt{k^2 - \omega^2} - i\omega}{\sqrt{k^2 - \omega^2} + i\omega} \right) \left(\frac{\sqrt{k^2 - \omega^2} - q}{\sqrt{k^2 - \omega^2} + q} \right) \right|. \quad (3.82)$$

Si nous cherchons une approximation d'ordre 0, alors $q = q_0$ pour $q_0 \in \mathbb{C}$, et le problème d'optimi-

sation qui en résulte s'écrira sous la forme :

$$\min_{q \in \mathbb{C}} \left(\max_{k \in]k_{\min}, k_- [\cup] k_+, k_{\max}[} \rho(k, q) \right). \quad (3.83)$$

avec $\rho(k, q)$ donné par l'expression (3.82). En revanche, si on cherche une approximation d'ordre 2, alors $q = q_0 + q_2 k^2$ et moyennant le changement de variables donné par (3.75) qui a été fait dans le paragraphe 3.3.3.2, le taux de convergence prend la forme suivante :

$$\rho(k, F(\alpha, \beta)) = \left| \left(\frac{\sqrt{k^2 - \omega^2} - i\omega}{\sqrt{k^2 - \omega^2} + i\omega} \right) \left(\frac{\sqrt{k^2 - \omega^2} - \alpha}{\sqrt{k^2 - \omega^2} + \alpha} \right) \left(\frac{\sqrt{k^2 - \omega^2} - \beta}{\sqrt{k^2 - \omega^2} + \beta} \right) \right|, \quad (3.84)$$

et dans ce cas, le problème de minimisation qui en résulte devient pour $\mathcal{E} = (i\mathbb{R} \times \mathbb{R}) \setminus \{(0, 0)\}$:

$$\min_{(\alpha, \beta) \in \mathcal{E}} \left(\max_{k \in]k_{\min}, k_- [\cup] k_+, k_{\max}[} \rho(k, F(\alpha, \beta)) \right). \quad (3.85)$$

Il n'est malheureusement pas possible de procéder à une étude analytique des problèmes d'optimisation (3.83) et (3.85). Ces problèmes sont étudiés dans [Dolean et al., 2006] au moyen d'une analyse asymptotique. Ici, nous avons choisi de procéder à une optimisation numérique. Il faut cependant s'attendre à ce que les paramètres numériques diffèrent des paramètres exacts selon l'algorithme de minimisation choisi.

3.3.3.4 Optimisation numérique des conditions d'interface

Pour l'optimisation numérique des conditions d'interface associées aux approximations d'ordre 0 et d'ordre 2 modifiées abordées au paragraphe 3.3.3.3, nous avons utilisé un algorithme relativement simpliste qui n'est pas forcément le mieux adapté à notre problématique mais suffit à illustrer notre propos comme nous le verrons dans la suite. Pour valider cet algorithme d'obtention des paramètres optimaux, nous avons tout d'abord vérifié que les paramètres obtenus dans un cadre d'expérimentation numérique concret se rapprochent de ceux trouvés aux paragraphes 3.3.3.1 et 3.3.3.2. Identiquement à la sous-section 3.2.5, le domaine de calcul Ω est le carré unité $[0, 1]^2$ avec des conditions absorbantes posées sur $\partial\Omega$. Nous choisissons de nouveau les paramètres $\varepsilon = \mu = 1$, $f = 400$ MHz et $f = 1$ GHz. Le pas de discrétisation est fixé à $h = 1/100$. Nous avons alors évalué l'erreur $|\rho_{opt} - \rho_{opt}^h|$ où ρ_{opt} est le taux de convergence théorique fourni par les formules (3.71) et (3.79), et ρ_{opt}^h le taux de convergence obtenu grâce aux paramètres approchés calculés numériquement. La table 3.3 résume les résultats obtenus lors des tests numériques.

	$f = 400$ MHz		$f = 1$ GHz	
	Ordre 0	Ordre 2	Ordre 0	Ordre 2
ρ_{opt}	0.825644	0.715396	0.785027	0.656262
ρ_{opt}^h	0.825866	0.715396	0.785119	0.656262
$ \rho_{opt} - \rho_{opt}^h $	$< 10^{-3}$	$< 10^{-6}$	$< 10^{-3}$	$< 10^{-6}$

TAB. 3.3 – Approximations d'ordre 0 et d'ordre 2 modifiées.
Validation numérique de l'algorithme d'optimisation.

Nous allons maintenant comparer les conditions d'interfaces utilisées dans les paragraphes 3.3.3.1 et 3.3.3.2 avec celles introduites dans le paragraphe 3.3.3.3. Nous introduisons pour la suite les abréviations suivantes :

- conditions d'interface d'ordre 0 du paragraphe 3.3.3.1 : CI0
- conditions d'interface d'ordre 2 du paragraphe 3.3.3.2 : CI2
- conditions d'interface modifiées d'ordre 0 du paragraphe 3.3.3.3 : CI0-m
- conditions d'interface modifiées d'ordre 2 du paragraphe 3.3.3.3 : CI2-m

Attardons-nous tout d'abord au taux de convergence noté $(\rho_{opt}^h)_m$ obtenu avec les conditions CI0-m et CI2-m et comparons-le au taux de convergence théorique noté ρ_{opt} qui caractérise les conditions CI0 et CI2. On s'aperçoit compte tenu des données de la table 3.4 que le taux de convergence CI0-m est très proche du taux CI2, et cela grâce à l'opérateur *dérivée tangentielle d'ordre 2* qui intervient dans la CI0-m mais qui n'intervient pas dans la CI0. En revanche, le résultat est moins satisfaisant pour la CI2-m puisque l'on n'obtient aucun gain sur le taux de convergence par rapport à la CI0-m. Il faut donc s'attendre au niveau de la résolution itérative à avoir un nombre d'itérations quasiment identique pour les conditions qui affichent des taux de convergence proches. Ce défaut est là encore attribuable au manque de précision de notre algorithme d'optimisation numérique.

	$f = 400 \text{ MHz}$		$f = 1 \text{ GHz}$	
	Ordre 0	Ordre 2	Ordre 0	Ordre 2
ρ_{opt}	0.825644	0.715396	0.785027	0.656262
$(\rho_{opt}^h)_m$	0.715623	0.715396	0.656262	0.656262

TAB. 3.4 – Approximations d'ordre 0 et d'ordre 2 modifiées.

Comparaison des taux de convergence associés aux conditions CI0, CI2, CI0-m et CI2-m.

3.3.4 Mise en œuvre de l'algorithme de Schwarz avec conditions optimisées

A l'instar de la sous-section 3.2.4, nous nous intéressons ici à la discrétisation de l'algorithme de Schwarz muni des conditions d'interface optimisées.

3.3.4.1 Conditions approchées d'ordre 0 et 2

Nous considérons ici les conditions approchées étudiées dans les paragraphes 3.3.3.1 et 3.3.3.2. Nous devons tout d'abord obtenir une forme appropriée de ces conditions d'interface en appliquant une transformée de Fourier inverse. Nous partons donc d'un domaine Ω de frontière extérieure $\partial\Omega$, décomposé en deux sous-domaines non-recouvrants Ω_1 et Ω_2 dont l'interface sera notée Γ_{int} . Les conditions d'interface optimisées posées sur Γ_{int} s'écrivent :

$$u_1^{1,p+1} + \mathcal{S}_{app} u_3^{1,p+1} = u_1^{2,p} + \mathcal{S}_{app} u_3^{2,p} \quad \text{et} \quad u_3^{2,p+1} + \mathcal{S}_{app} u_1^{2,p+1} = u_3^{1,p} + \mathcal{S}_{app} u_1^{1,p}, \quad (3.86)$$

où $u_1^{j,p}$ et $u_3^{j,p}$ sont les variables caractéristiques (2.8) associées au sous-domaine Ω_j , $j = 1, 2$ dont on rappelle les expressions :

$$u_1^{j,p} = \frac{1}{2}(E_z^{j,p} + H_y^{j,p}) \quad \text{et} \quad u_3^{j,p} = \frac{1}{2}(H_y^{j,p} - E_z^{j,p}), \quad (3.87)$$

et \mathcal{S}_{app} l'opérateur de transmission approché ayant pour symbole de Fourier :

$$\hat{\mathcal{S}}_{app} = -\frac{q_0 + q_2 k^2 - i\omega}{q_0 + q_2 k^2 + i\omega}, \quad (q_0, q_2) \in \mathbb{C}^2. \quad (3.88)$$

On se propose de démontrer le lemme suivant.

Lemme 6 *Les conditions d'interface d'interface posées sur Γ_{int} et données par (3.86) se mettent sous la forme :*

$$\begin{cases} q_0 E_z^{1,p+1} + i\omega H_y^{1,p+1} + i\omega q_2 \partial_y H_x^{1,p+1} &= q_0 E_z^{2,p} + i\omega H_y^{2,p} + i\omega q_2 \partial_y H_x^{2,p}, \\ -q_0 E_z^{2,p+1} + i\omega H_y^{2,p+1} - i\omega q_2 \partial_y H_x^{2,p+1} &= -q_0 E_z^{1,p} + i\omega H_y^{1,p} - i\omega q_2 \partial_y H_x^{1,p}. \end{cases} \quad (3.89)$$

Preuve En utilisant les expressions (3.87) et (3.88) et en simplifiant, nous trouvons donc pour la première condition de (3.86) :

$$\begin{aligned} \hat{u}_1^{1,p+1} + \hat{\mathcal{S}}_{app} \hat{u}_3^{1,p+1} &= \hat{u}_1^{2,p} + \hat{\mathcal{S}}_{app} \hat{u}_3^{2,p} \\ \Leftrightarrow (q_0 + q_2 k^2) \hat{E}_z^{1,p+1} + i\omega \hat{H}_y^{1,p+1} &= (q_0 + q_2 k^2) \hat{E}_z^{2,p} + i\omega \hat{H}_y^{2,p}. \end{aligned} \quad (3.90)$$

Et en appliquant à (3.90) la transformée de Fourier partielle inverse qui vérifie pour une fonction U régulière :

$$U^{j,p}(x, y) = (\mathcal{F}_y^{-1} \hat{U}^{j,p})(x, y) = \int_{\mathbf{R}} e^{iky} \hat{U}^{j,p}(x, k) dk,$$

on obtient :

$$q_0 E_z^{1,p+1} + i\omega H_y^{1,p+1} - q_2 \partial_y^2 E_z^{1,p+1} = q_0 E_z^{2,p} + i\omega H_y^{2,p} - q_2 \partial_y^2 E_z^{2,p} \quad \text{sur } \Gamma_{int}. \quad (3.91)$$

Or, en appliquant ∂_y à la deuxième égalité des équations de Maxwell (2.1), on obtient pour $j = 1, 2$ et $m = p, p + 1$:

$$\partial_y^2 E_z^{j,m} = \partial_y f_2^j - i\omega \partial_y H_x^{j,m} \quad \text{sur } \Omega_j, \quad (3.92)$$

si bien qu'en appliquant un opérateur de trace à l'expression (3.92) et en supposant f suffisamment régulière pour que l'on ait $(\partial_y f_2^1)|_{\Gamma_{int}} = (\partial_y f_2^2)|_{\Gamma_{int}}$, on peut écrire :

$$\begin{aligned} & q_0 E_z^{1,p+1} + i\omega H_y^{1,p+1} = q_0 E_z^{2,p} + i\omega H_y^{2,p} \\ (3.91) \Rightarrow & \quad + i\omega q_2 \partial_y H_x^{1,p+1} \quad \quad \quad + i\omega q_2 \partial_y H_x^{2,p} \\ & \quad - q_2 \partial_y f_2^1 \quad \quad \quad - q_2 \partial_y f_2^2 \\ \Rightarrow & q_0 E_z^{1,p+1} + i\omega H_y^{1,p+1} = q_0 E_z^{2,p} + i\omega H_y^{2,p} \\ & \quad + i\omega q_2 \partial_y H_x^{1,p+1} \quad \quad \quad + i\omega q_2 \partial_y H_x^{2,p} \quad \text{sur } \Gamma_{int}. \end{aligned} \quad (3.93)$$

La deuxième condition de (3.86) se déduit en reprenant la même démarche. ■

Grâce au lemme 6 et à l'aide des opérateurs différentiels \mathcal{B}_j allant de Ω_j vers Γ_{int} et tels que $\mathcal{B}_j(\mathbf{W}^j) = (-1)^{j-1}q_0E_z^j + i\omega H_y^j + (-1)^{j-1}i\omega q_2\partial_y H_x^j$, $j = 1, 2$, l'algorithme de Schwarz muni des conditions d'interface (3.89) et utilisant des multiplicateurs de Lagrange s'écrit :

$$\begin{cases} \mathcal{L}\mathbf{W}^{1,p+1} = \mathbf{f}^1 & \text{sur } \Omega_1, \\ \mathcal{B}_1(\mathbf{W}^{1,p+1}) = \lambda^{1,p} & \text{sur } \Gamma_{\text{int}}, \\ M_{\mathbf{n}}^-\mathbf{W}^{1,p+1} = M_{\mathbf{n}}^-\mathbf{W}^{\text{inc}} & \text{sur } \Gamma_1, \end{cases} \quad (3.94)$$

$$\begin{cases} \mathcal{L}\mathbf{W}^{2,p+1} = \mathbf{f}^2 & \text{sur } \Omega_2, \\ \mathcal{B}_2(\mathbf{W}^{2,p+1}) = \lambda^{2,p} & \text{sur } \Gamma_{\text{int}}, \\ M_{\mathbf{n}}^-\mathbf{W}^{2,p+1} = M_{\mathbf{n}}^-\mathbf{W}^{\text{inc}} & \text{sur } \Gamma_2, \end{cases}$$

sachant que :

$$\begin{cases} \lambda^{1,p} = \mathcal{B}_1(\mathbf{W}^{2,p}) = q_0E_z^{2,p} + i\omega H_y^{2,p} + i\omega q_2\partial_y H_x^{2,p}, \\ \lambda^{2,p} = \mathcal{B}_2(\mathbf{W}^{1,p}) = -q_0E_z^{1,p} + i\omega H_y^{1,p} - i\omega q_2\partial_y H_x^{1,p}. \end{cases} \quad (3.95)$$

De là, en itérant sur (3.95) et en se servant de l'égalité sur les multiplicateurs de Lagrange des systèmes de (3.94), on aboutit à :

$$\begin{cases} \lambda^{1,p+1} = -\lambda^{2,p} + 2i\omega H_y^{2,p+1}, \\ \lambda^{2,p+1} = -\lambda^{1,p} + 2i\omega H_y^{1,p+1}, \end{cases}$$

si bien que l'algorithme de Schwarz devient :

Etape 1

$$\begin{cases} \mathcal{L}\mathbf{W}^{1,p+1} = \mathbf{f}^1 & \text{sur } \Omega_1, \\ \mathcal{B}_1(\mathbf{W}^{1,p+1}) = \lambda^{1,p} & \text{sur } \Gamma_{\text{int}}, \\ M_{\mathbf{n}}^-\mathbf{W}^{1,p+1} = M_{\mathbf{n}}^-\mathbf{W}^{\text{inc}} & \text{sur } \Gamma_1, \end{cases} \quad (3.96)$$

$$\begin{cases} \mathcal{L}\mathbf{W}^{2,p+1} = \mathbf{f}^2 & \text{sur } \Omega_2, \\ \mathcal{B}_2(\mathbf{W}^{2,p+1}) = \lambda^{2,p} & \text{sur } \Gamma_{\text{int}}, \\ M_{\mathbf{n}}^-\mathbf{W}^{2,p+1} = M_{\mathbf{n}}^-\mathbf{W}^{\text{inc}} & \text{sur } \Gamma_2, \end{cases}$$

Etape 2

$$\begin{cases} \lambda^{1,p+1} = -\lambda^{2,p} + 2i\omega H_y^{2,p+1} & \text{sur } \Gamma_{\text{int}}, \\ \lambda^{2,p+1} = -\lambda^{1,p} + 2i\omega H_y^{1,p+1} & \text{sur } \Gamma_{\text{int}}. \end{cases} \quad (3.97)$$

L'algorithme de Schwarz (3.96)-(3.97) peut se réinterpréter comme un algorithme de point fixe aux nouvelles variables λ^j , $j = 1, 2$, solutions du problème :

$$\begin{cases} \lambda^1 = -\lambda^2 + 2i\omega H_y^2(\lambda^2, \mathbf{f}^2) & \text{sur } \Gamma_{\text{int}}, \\ \lambda^2 = -\lambda^1 + 2i\omega H_y^1(\lambda^1, \mathbf{f}^1) & \text{sur } \Gamma_{\text{int}}, \end{cases} \quad (3.98)$$

et où $H_y^j(\lambda^j, \mathbf{f}^j)$ sont les deuxièmes composantes des vecteurs $\mathbf{W}^j(\lambda^j, \mathbf{f}^j)$ eux-mêmes solutions de :

$$\begin{cases} \mathcal{L}\mathbf{W}^j = \mathbf{f}^j & \text{sur } \Omega_j, \\ (-1)^{j-1}q_0 E_z^j + i\omega H_y^j + (-1)^{j-1}i\omega q_2 \partial_y H_x^j = \lambda^j & \text{sur } \Gamma_{\text{int}}, \\ M_{\bar{\mathbf{n}}}\mathbf{W}^j = M_{\bar{\mathbf{n}}}\mathbf{W}^{\text{inc}} & \text{sur } \Gamma_j. \end{cases} \quad (3.99)$$

Tout comme au paragraphe 3.2.4, cet algorithme peut se reformuler sous la forme d'une itération portant sur les variables interfaces uniquement après avoir formulé le complément de Schur associé (cf. (2.49)). Cependant, nous devons auparavant décrire comment discrétiser et implémenter le système (3.99). Pour cela, on s'inspire de la sous-section 3.2.3 en reprenant les mêmes notations et la même démarche, la seule différence notable se trouvant au niveau de la condition d'interface. Au niveau discret, dans la mesure où $\Gamma_{\text{int}} = \bigcup_{l=1}^{N_y} (C_{Jl})$, on se propose d'intégrer la condition d'interface sur une cellule C_{Jl} pour un l fixé, ce qui donne lieu aux conditions d'interface discrètes suivantes :

$$(-1)^{j-1}q_0(E_z)_{Jl}^j + i\omega(H_y)_{Jl}^j + (-1)^{j-1}i\omega q_2 \frac{(H_x)_{J,l+1}^j - (H_x)_{J,l-1}^j}{2\Delta y} = \lambda_l^j. \quad (3.100)$$

De ce fait, le schéma numérique pour $j = J$ s'écrit comme en (2.43) hormis pour les conditions

d'interface où l'on se sert de (3.100) :

$$\begin{aligned}
\Omega_1 : \left\{ \begin{aligned}
& \left(i\omega + \frac{1}{2\Delta x} \right) [(E_z)_{Jl}^{1,p+1} - (H_y)_{Jl}^{1,p+1}] - \frac{1}{2\Delta x} [(E_z)_{J-1,l}^{1,p+1} - (H_y)_{J-1,l}^{1,p+1}] \\
& \quad + \frac{(H_x)_{J,l+1}^{1,p+1} - (H_x)_{J,l-1}^{1,p+1}}{2\Delta y} = f_{E_z, Jl}^1 - f_{H_y, Jl}^1, \\
& i\omega (H_x)_{Jl}^{1,p+1} + \frac{(E_z)_{J,l+1}^{1,p+1} - (E_z)_{J,l-1}^{1,p+1}}{2\Delta y} = f_{H_x, Jl}^1, \\
& q_0 (E_z)_{Jl}^{1,p+1} + i\omega (H_y)_{Jl}^{1,p+1} + i\omega q_2 \frac{(H_x)_{J,l+1}^{1,p+1} - (H_x)_{J,l-1}^{1,p+1}}{2\Delta y} = \lambda_l^{1,p},
\end{aligned} \right. \\
\Omega_2 : \left\{ \begin{aligned}
& \left(i\omega + \frac{1}{2\Delta x} \right) [(E_z)_{Jl}^{2,p+1} + (H_y)_{Jl}^{2,p+1}] - \frac{1}{2\Delta x} [(E_z)_{J+1,l}^{2,p+1} + (H_y)_{J+1,l}^{2,p+1}] \\
& \quad + \frac{(H_x)_{J,l+1}^{2,p+1} - (H_x)_{J,l-1}^{2,p+1}}{2\Delta y} = f_{E_z, Jl}^2 - f_{H_y, Jl}^2, \\
& i\omega (H_x)_{Jl}^{2,p+1} + \frac{(E_z)_{J,l+1}^{2,p+1} - (E_z)_{J,l-1}^{2,p+1}}{2\Delta y} = f_{H_x, Jl}^2, \\
& -q_0 (E_z)_{Jl}^{2,p+1} + i\omega (H_y)_{Jl}^{2,p+1} - i\omega q_2 \frac{(H_x)_{J,l+1}^{2,p+1} - (H_x)_{J,l-1}^{2,p+1}}{2\Delta y} = \lambda_l^{2,p}.
\end{aligned} \right.
\end{aligned} \tag{3.101}$$

Remarque 10 Pour implémenter les conditions d'interface aux cellules $C_{J,1}$ et C_{J,N_y} , on a besoin des valeurs du champ \mathbf{W} pris aux cellules fictives $C_{J,0}$ et C_{J,N_y+1} . On décide alors de prendre $\mathbf{W} = W^{inc}$ pour $l = 0, N_y + 1$.

La discrétisation du système (3.98) à partir du schéma volumes finis centrés aboutit au système linéaire :

$$\begin{cases}
\lambda^1 = -\lambda^2 + 2i\omega B_2 \mathbf{W}^2, \\
\lambda^2 = -\lambda^1 + 2i\omega B_1 \mathbf{W}^1, \\
A_1 \mathbf{W}^1 = \mathbf{F}^1 + B_1^t \lambda^1, \\
A_2 \mathbf{W}^2 = \mathbf{F}^2 + B_2^t \lambda^2.
\end{cases} \tag{3.102}$$

où les matrices B_1 et B_2 respectivement de taille $N_y \times 3JN_y$ et $N_y \times 3(N_x - J)N_y$ sont les matrices des opérateurs de trace discrétisés vérifiant cette fois-ci $\mathbf{W}_{|\Omega_j}^j \mapsto (H_y)_{|\Gamma_{int}}^j$. De même, la matrice de l'opérateur de relèvement continu vérifie cette fois-ci : $R_j = B_j^t$. Dans ces conditions, le système relatif au complément de Schur et provenant de (3.102) est donc :

$$T\lambda = \mathbf{d}, \tag{3.103}$$

où $\lambda = (\lambda^1, \lambda^2)$ et où la matrice T ainsi que le second membre \mathbf{d} sont donnés par :

$$T = \begin{pmatrix} \mathbf{I} & \mathbf{I} - 2i\omega B_2 A_2^{-1} B_2^t \\ \mathbf{I} - 2i\omega B_1 A_1^{-1} B_1^t & \mathbf{I} \end{pmatrix} \quad \text{et} \quad \mathbf{d} = \begin{pmatrix} 2i\omega B_2 A_2^{-1} \mathbf{F}^2 \\ 2i\omega B_1 A_1^{-1} \mathbf{F}^1 \end{pmatrix}. \tag{3.104}$$

3.3.4.2 Conditions approchées d'ordre 0 et 2 modifiées

Dans la mesure où ce paragraphe reprend la démarche effectuée dans le paragraphe 3.3.4.1, nous irons assez rapidement sur certains points. Les conditions d'interface optimisées posées sur Γ_{int} s'écrivent toujours :

$$u_1^{1,p+1} + \mathcal{S}_{app} u_3^{1,p+1} = u_1^{2,p} + \mathcal{S}_{app} u_3^{2,p} \quad \text{et} \quad u_3^{2,p+1} + \mathcal{S}_{app} u_1^{1,p+1} = u_3^{2,p} + \mathcal{S}_{app} u_1^{2,p}, \quad (3.105)$$

sauf que l'opérateur de transmission approché \mathcal{S}_{app} a désormais pour symbole de Fourier :

$$\hat{\mathcal{S}}_{app} = -\frac{k^2}{(1 + 2i\omega q_2)k^2 + 2i\omega q_0 - 2\omega^2}, \quad (q_0, q_2) \in \mathbb{C}^2.$$

Le lemme suivant se démontre alors comme le lemme 6.

Lemme 7 *A l'aide des opérateurs différentiels \mathcal{B}_j allant de Ω_j vers Γ_{int} et vérifiant :*

$$\mathcal{B}_j(\mathbf{W}^j) = (-1)^j(\omega^2 - i\omega q_0)E_z^j + (-1)^j(\omega^2 q_2 - i\omega)\partial_y H_x^j - (\omega^2 - i\omega q_0)H_y^j - i\omega q_2 \partial_y^2 H_y^j, \quad j = 1, 2$$

les conditions d'interface d'interface posées sur Γ_{int} et données par (3.105) se mettent sous la forme :

$$\begin{cases} \mathcal{B}_1(\mathbf{W}^{1,p+1}) &= \mathcal{B}_1(\mathbf{W}^{2,p}), \\ \mathcal{B}_2(\mathbf{W}^{2,p+1}) &= \mathcal{B}_2(\mathbf{W}^{1,p}). \end{cases} \quad (3.106)$$

Preuve Identique au lemme 6. ■

De ce fait, l'algorithme de Schwarz muni des conditions d'interface (3.105) et utilisant des multiplieurs de Lagrange s'écrit :

$$\begin{cases} \mathcal{L}\mathbf{W}^{1,p+1} &= \mathbf{f}^1 & \text{sur } \Omega_1, \\ \mathcal{B}_1(\mathbf{W}^{1,p+1}) &= \lambda^{1,p} & \text{sur } \Gamma_{\text{int}}, \\ M_{\bar{\mathbf{n}}}\mathbf{W}^{1,p+1} &= M_{\bar{\mathbf{n}}}\mathbf{W}^{\text{inc}} & \text{sur } \Gamma_1, \\ \mathcal{L}\mathbf{W}^{2,p+1} &= \mathbf{f}^2 & \text{sur } \Omega_2, \\ \mathcal{B}_2(\mathbf{W}^{2,p+1}) &= \lambda^{2,p} & \text{sur } \Gamma_{\text{int}}, \\ M_{\bar{\mathbf{n}}}\mathbf{W}^{2,p+1} &= M_{\bar{\mathbf{n}}}\mathbf{W}^{\text{inc}} & \text{sur } \Gamma_2, \end{cases} \quad (3.107)$$

sachant que :

$$\begin{cases} \lambda^{1,p} &= \mathcal{B}_1(\mathbf{W}^{2,p}) &= -(\omega^2 - i\omega q_0)E_z^{2,p} - (\omega^2 q_2 - i\omega)\partial_y H_x^{2,p} \\ & & - (\omega^2 - i\omega q_0)H_y^{2,p} - i\omega q_2 \partial_y^2 H_y^{2,p}, \\ \lambda^{2,p} &= \mathcal{B}_2(\mathbf{W}^{1,p}) &= (\omega^2 - i\omega q_0)E_z^{1,p} + (\omega^2 q_2 - i\omega)\partial_y H_x^{1,p} \\ & & - (\omega^2 - i\omega q_0)H_y^{1,p} - i\omega q_2 \partial_y^2 H_y^{1,p}. \end{cases} \quad (3.108)$$

En itérant sur (3.108) et en se servant de l'égalité sur les multiplicateurs de Lagrange des systèmes de (3.107), on aboutit à :

$$\begin{cases} \lambda^{1,p+1} = \lambda^{2,p} - 2(\omega^2 - i\omega q_0)E_z^{2,p+1} - 2(\omega^2 q_2 - i\omega)\partial_y H_x^{2,p+1}, \\ \lambda^{2,p+1} = \lambda^{1,p} + 2(\omega^2 - i\omega q_0)E_z^{1,p+1} - 2(\omega^2 q_2 - i\omega)\partial_y H_x^{1,p+1}, \end{cases}$$

si bien que l'algorithme de Schwarz devient :

Etape 1

$$\begin{cases} \mathcal{L}\mathbf{W}^{1,p+1} = \mathbf{f}^1 & \text{sur } \Omega_1, \\ \mathcal{B}_1(\mathbf{W}^{1,p+1}) = \lambda^{1,p} & \text{sur } \Gamma_{\text{int}}, \\ M_{\mathbf{n}}^- \mathbf{W}^{1,p+1} = M_{\mathbf{n}}^- \mathbf{W}^{\text{inc}} & \text{sur } \Gamma_1, \\ \mathcal{L}\mathbf{W}^{2,p+1} = \mathbf{f}^2 & \text{sur } \Omega_2, \\ \mathcal{B}_2(\mathbf{W}^{2,p+1}) = \lambda^{2,p} & \text{sur } \Gamma_{\text{int}}, \\ M_{\mathbf{n}}^- \mathbf{W}^{2,p+1} = M_{\mathbf{n}}^- \mathbf{W}^{\text{inc}} & \text{sur } \Gamma_2, \end{cases} \quad (3.109)$$

Etape 2

$$\begin{cases} \lambda^{1,p+1} = \lambda^{2,p} - 2(\omega^2 - i\omega q_0)E_z^{2,p+1} \\ \quad - 2(\omega^2 q_2 - i\omega)\partial_y H_x^{2,p+1} & \text{sur } \Gamma_{\text{int}}, \\ \lambda^{2,p+1} = \lambda^{1,p} + 2(\omega^2 - i\omega q_0)E_z^{1,p+1} \\ \quad - 2(\omega^2 q_2 - i\omega)\partial_y H_x^{1,p+1} & \text{sur } \Gamma_{\text{int}}. \end{cases} \quad (3.110)$$

Cet algorithme peut se réinterpréter comme un algorithme de point fixe appliqué aux nouvelles variables λ^j , $j = 1, 2$, solutions du problème :

$$\begin{cases} \lambda^1 = \lambda^2 - 2(\omega^2 - i\omega q_0)E_z^2(\lambda^2, \mathbf{f}^2) \\ \quad - 2(\omega^2 q_2 - i\omega)\partial_y H_x^2(\lambda^2, \mathbf{f}^2) & \text{sur } \Gamma_{\text{int}}, \\ \lambda^2 = \lambda^1 + 2(\omega^2 - i\omega q_0)E_z^1(\lambda^1, \mathbf{f}^1) \\ \quad - 2(\omega^2 q_2 - i\omega)\partial_y H_x^1(\lambda^1, \mathbf{f}^1) & \text{sur } \Gamma_{\text{int}}, \end{cases} \quad (3.111)$$

d'interface où l'on se sert de (3.114) :

$$\left\{ \begin{array}{l} \text{Dans } \Omega_1 \\ (i\omega + \frac{1}{2\Delta x})[(E_z)_{Jl}^{1,p+1} - (H_y)_{Jl}^{1,p+1}] - \frac{1}{2\Delta x}[(E_z)_{J-1,l}^{1,p+1} - (H_y)_{J-1,l}^{1,p+1}] \\ \quad + \frac{(H_x)_{J,l+1}^{1,p+1} - (H_x)_{J,l-1}^{1,p+1}}{2\Delta y} = f_{E_z, Jl}^1 - f_{H_y, Jl}^1, \\ i\omega(H_x)_{Jl}^{1,p+1} + \frac{(E_z)_{J,l+1}^{1,p+1} - (E_z)_{J,l-1}^{1,p+1}}{2\Delta y} = f_{H_x, Jl}^1, \\ -(\omega^2 - i\omega q_0)(E_z)_{Jl}^{1,p+1} - (\omega^2 q_2 - i\omega) \frac{(H_x)_{J,l+1}^{1,p+1} - (H_x)_{J,l-1}^{1,p+1}}{2\Delta y} - \\ \left(\omega^2 - i\omega \left(q_0 + \frac{q_2}{2\Delta y^2} \right) \right) (H_y)_{Jl}^{1,p+1} - i\omega q_2 \frac{(H_y)_{J,l+2}^{1,p+1} + (H_y)_{J,l-2}^{1,p+1}}{4\Delta y^2} = \lambda_l^{1,p}, \end{array} \right. \quad (3.115)$$

$$\left\{ \begin{array}{l} \text{Dans } \Omega_2 \\ (i\omega + \frac{1}{2\Delta x})[(E_z)_{Jl}^{2,p+1} + (H_y)_{Jl}^{2,p+1}] - \frac{1}{2\Delta x}[(E_z)_{J+1,l}^{2,p+1} + (H_y)_{J+1,l}^{2,p+1}] \\ \quad + \frac{(H_x)_{J,l+1}^{2,p+1} - (H_x)_{J,l-1}^{2,p+1}}{2\Delta y} = f_{E_z, Jl}^2 - f_{H_y, Jl}^2, \\ i\omega(H_x)_{Jl}^{2,p+1} + \frac{(E_z)_{J,l+1}^{2,p+1} - (E_z)_{J,l-1}^{2,p+1}}{2\Delta y} = f_{H_x, Jl}^2, \\ (\omega^2 - i\omega q_0)(E_z)_{Jl}^{2,p+1} + (\omega^2 q_2 - i\omega) \frac{(H_x)_{J,l+1}^{2,p+1} - (H_x)_{J,l-1}^{2,p+1}}{2\Delta y} - \\ \left(\omega^2 - i\omega \left(q_0 + \frac{q_2}{2\Delta y^2} \right) \right) (H_y)_{Jl}^{2,p+1} - i\omega q_2 \frac{(H_y)_{J,l+2}^{2,p+1} + (H_y)_{J,l-2}^{2,p+1}}{4\Delta y^2} = \lambda_l^{2,p}. \end{array} \right.$$

La discrétisation du système (3.115) aboutit au système linéaire :

$$\left\{ \begin{array}{l} \lambda^1 = \lambda^2 - B_2 \mathbf{W}^2, \\ \lambda^2 = \lambda^1 - B_1 \mathbf{W}^1, \\ A_1 \mathbf{W}^1 = \mathbf{F}^1 + B_1^t \lambda^1, \\ A_2 \mathbf{W}^2 = \mathbf{F}^2 + B_2^t \lambda^2, \end{array} \right. \quad (3.116)$$

où les matrices B_1 et B_2 sont de taille respective $N_y \times 3JN_y$ et $N_y \times 3(N_x - J)N_y$.

Le système relatif au complément de Schur issu de (3.116) est donc :

$$T\lambda = \mathbf{d}, \quad (3.117)$$

où $\lambda = (\lambda^1, \lambda^2)$ et où la matrice T ainsi que le second membre \mathbf{d} sont donnés par :

$$T = \begin{pmatrix} \mathbf{I} & -\mathbf{I} + B_2 A_2^{-1} B_2^t \\ -\mathbf{I} + B_1 A_1^{-1} B_1^t & \mathbf{I} \end{pmatrix} \quad \text{et} \quad \mathbf{d} = \begin{pmatrix} -B_2 A_2^{-1} \mathbf{F}^2 \\ -B_1 A_1^{-1} \mathbf{F}^1 \end{pmatrix}. \quad (3.118)$$

3.3.5 Résultats numériques

Notre objectif est d'évaluer les performances de l'algorithme de Schwarz muni de conditions optimisées et de comparer celles-ci aux performances de l'algorithme de Schwarz muni de conditions classiques. Nous adoptons ici les conditions d'expérimentation numérique de la sous-section 3.2.5. La seule différence est au niveau du domaine de calcul, puisque l'on choisit de tester ces algorithmes sur $\Omega = [0, 1] \times [0, 2]$ pour pouvoir respecter la première condition des hypothèses (3.69). Nous considérons des discrétisations uniformes de Ω caractérisées par les nombres de points $N = 50, 100, 200$, et choisissons les fréquences $f = 400$ MHz et $f = 1$ GHz. Enfin, précisons que l'on a opté pour une onde incidente se propageant suivant l'axe (Ox) de sorte que les conditions d'interface soient le maximum sollicitées.

Tout d'abord, rappelons que l'algorithme de Schwarz utilisant des conditions classiques ne converge pas pour une décomposition sans recouvrement. En revanche, l'étude du taux de convergence de l'algorithme de Schwarz basé sur des conditions optimisées a montré que cette méthode converge. Ceci est confirmé par les données de la table 3.5 qui répertorie les nombres d'itérations obtenus dans chaque cas de conditions d'interfaces.

CI0	CI2	CI0-m	CI2-m
33	22	26	26

TAB. 3.5 – Résolution itérative du système interface.
Nombres d'itérations pour les différentes conditions d'interfaces approchées.
Décomposition sans recouvrement, $f = 400$ MHz et $N = 50$.

Ensuite, on répertorie dans la table 3.6, les résultats obtenus en accélérant l'algorithme de Schwarz par la méthode GMRES. Les courbes de convergence correspondantes sont montrées sur les figures 3.21 et 3.22 pour le maillage défini par $N = 200$. On remarque tout d'abord que la convergence de l'algorithme de Schwarz basé sur les conditions d'interface classiques est assez sensible à la fois au pas de discrétisation et à la fréquence de l'onde incidente. L'algorithme de Schwarz utilisant des conditions optimisées est quant à lui plus robuste vis-à-vis de ces paramètres. Par ailleurs, le gain obtenu en utilisant les conditions optimisées augmente avec la finesse du maillage. Enfin, les observations faites dans le paragraphe 3.3.4.1 sont encore de mise. En effet, bien qu'il résulte un gain d'une dizaine d'itérations de l'utilisation des conditions CI2/CI0-m/CI2-m par rapport à l'algorithme de Schwarz basé sur les conditions CI0, on n'observe aucun gain entre les algorithmes de Schwarz respectivement basés sur les conditions CI0-m et CI2-m. Il est donc nécessaire d'utiliser un algorithme de minimisation plus robuste, ou de procéder à l'étude analytique des conditions d'interfaces modifiées bien que cette étude semble bien plus ardue que celles effectuées pour les conditions CI0 et CI2.

3.4 Annexe

3.4.1 Algorithme de Schwarz pour les équations de Maxwell 3D

Nous étudions ici la convergence de l'algorithme de Schwarz appliqué aux équations de Maxwell 3D en nous limitant au cas continu. L'objectif est d'identifier la dépendance de la dimension du problème

	$f = 400$ MHz			$f = 1$ GHz		
	$N = 50$	$N = 100$	$N = 200$	$N = 50$	$N = 100$	$N = 200$
CI classiques	60	85	122	42	62	91
CI0	27	32	38	27	29	33
CI2	17	18	19	18	17	18
CI0-m	17	17	19	18	18	18
CI2-m	16	18	20	19	18	19

TAB. 3.6 – Résolution du système interface par la méthode GMRES(5).
 Nombres d'itérations pour les différentes conditions d'interfaces approchées.
 Décomposition sans recouvrement, $f = 400$ MHz et $N = 50$.

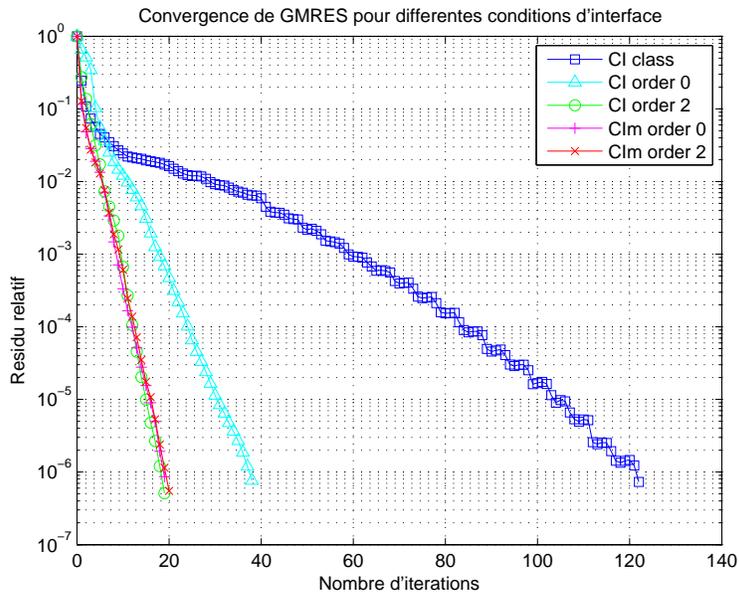


FIG. 3.21 – Evolution du résidu relatif (2.52) en fonction du nombre d'itérations.
 Algorithme de Schwarz non-recouvrant accéléré par GMRES pour $N = 200$ et $f = 400$ MHz.

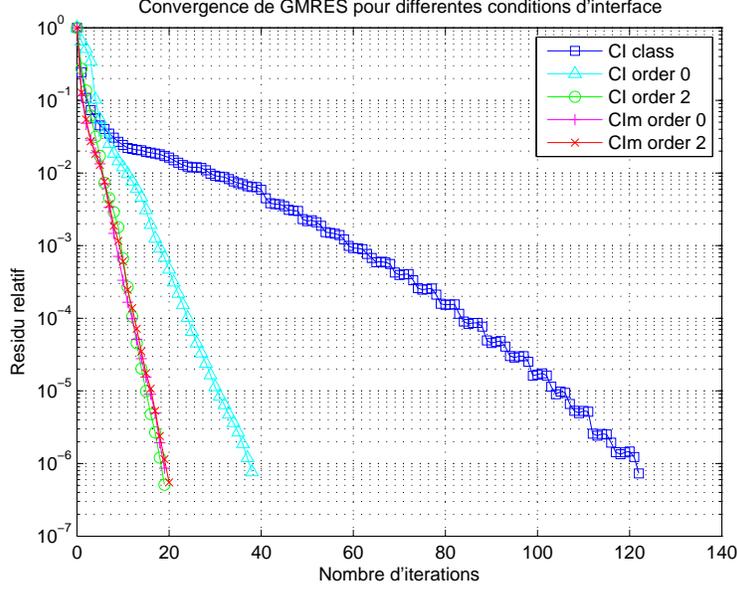


FIG. 3.22 – Evolution du résidu relatif (2.52) en fonction du nombre d'itérations. Algorithme de Schwarz non-recouvrant accéléré par GMRES pour $N = 200$ et $f = 1$ GHz.

sur le taux de convergence de l'algorithme de Schwarz. Cette étude comportant les mêmes étapes que l'étude réalisée en 2D, nous passerons outre la description de certains calculs. On note :

$$N_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \quad N_y = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad \text{et} \quad N_z = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

et on introduit :

$$M_x = \begin{pmatrix} 0_3 & N_x \mu^{-1} \\ -N_x \varepsilon^{-1} & 0_3 \end{pmatrix}, \quad M_y = \begin{pmatrix} 0_3 & N_y \mu^{-1} \\ -N_y \varepsilon^{-1} & 0_3 \end{pmatrix} \quad \text{et} \quad M_z = \begin{pmatrix} 0_3 & N_z \mu^{-1} \\ -N_z \varepsilon^{-1} & 0_3 \end{pmatrix},$$

ainsi que :

$$N = n_x N_x + n_y N_y + n_z N_z \quad \text{et} \quad M_{\bar{\mathbf{n}}} = \frac{1}{2} \begin{pmatrix} cN^2 & \mu^{-1}N \\ -\varepsilon^{-1}N & cN^2 \end{pmatrix}.$$

A partir de ces matrices, nous pouvons reprendre le cheminement de la sous-section 3.2.2.1 qui reste valable, et exploiter la formulation de l'algorithme de Schwarz (2.7) qui s'écrit :

$$\begin{cases} \mathcal{L} \mathbf{W}^{1,p+1} = \mathbf{f}^1 & \text{sur } \Omega_1, \\ M_{\bar{\mathbf{n}}_{12}}^- \mathbf{W}^{1,p+1} = M_{\bar{\mathbf{n}}_{12}}^- \mathbf{W}^{2,p} & \text{sur } \Gamma_{12}, \\ M_{\bar{\mathbf{n}}}^- \mathbf{W}^{1,p+1} = M_{\bar{\mathbf{n}}}^- \mathbf{W}^{\text{inc}} & \text{sur } \Gamma, \end{cases} \quad \begin{cases} \mathcal{L} \mathbf{W}^{2,p+1} = \mathbf{f}^2 & \text{sur } \Omega_2, \\ M_{\bar{\mathbf{n}}_{21}}^- \mathbf{W}^{2,p+1} = M_{\bar{\mathbf{n}}_{21}}^- \mathbf{W}^{1,p} & \text{sur } \Gamma_{21}, \\ M_{\bar{\mathbf{n}}}^- \mathbf{W}^{2,p+1} = M_{\bar{\mathbf{n}}}^- \mathbf{W}^{\text{inc}} & \text{sur } \Gamma, \end{cases} \quad (4.119)$$

où $\mathcal{L} = i\omega \mathbf{l}_6 + M_x \partial_x + M_y \partial_y + M_z \partial_z$.

Toujours dans l'optique de faire une analyse de Fourier, on considère le problème (4.119) sur le domaine $\Omega = \mathbf{R}^3$ qui est ici décomposé en deux sous-domaines non-bornés en y et en z , $\Omega_1 =]-\infty, b[\times \mathbf{R}^2$ et $\Omega_2 =]a, +\infty[\times \mathbf{R}^2$ avec toujours $a \leq b$, ε et μ constants, pour faciliter l'analyse.

Cette fois-ci, les interfaces (frontières artificielles entre les sous-domaines) sont des plans $\Gamma_1 = \mathbf{R} \times \{b\}$ et $\Gamma_2 = \mathbf{R} \times \{a\}$ où les normales extérieures sont données par $\mathbf{n}_{12} = -\mathbf{n}_{21} = (1, 0, 0)^t$. On convertit notre système (4.119) en un système qui porte sur les variables caractéristiques suivant la variable x (dans la direction normale à l'interface). Dans ce cas, $M_{\mathbf{n}_{12}} = M_x$ est diagonalisable en $M_x = T\Lambda T^{-1}$ où :

$$T = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & -c^{-1} & 0 & 0 & 0 & c^{-1} \\ c^{-1} & 0 & 0 & 0 & -c^{-1} & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ \mu & 0 & 0 & 0 & \mu & 0 \\ 0 & \mu & 0 & 0 & 0 & \mu \end{pmatrix}, \quad \Lambda = \begin{pmatrix} -c & 0 & 0 & 0 & 0 & 0 \\ 0 & -c & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & c & 0 \\ 0 & 0 & 0 & 0 & 0 & c \end{pmatrix},$$

$$\text{et } T^{-1} = \frac{1}{2} \begin{pmatrix} 0 & 0 & c & 0 & \mu^{-1} & 0 \\ 0 & -c & 0 & 0 & 0 & \mu^{-1} \\ 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & -c & 0 & \mu^{-1} & 0 \\ 0 & c & 0 & 0 & 0 & \mu^{-1} \end{pmatrix}.$$

On introduit :

$$A_y = T^{-1}M_yT = \frac{1}{2} \begin{pmatrix} 0 & 0 & 0 & c\mu^{-1} & 0 & 0 \\ 0 & 0 & -c^2 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 & 0 & -2 \\ 2z & 0 & 0 & 0 & -2z & 0 \\ 0 & 0 & 0 & -c\mu^{-1} & 0 & 0 \\ 0 & 0 & -c^2 & 0 & 0 & 0 \end{pmatrix},$$

$$\text{et } A_z = T^{-1}M_zT = \frac{1}{2} \begin{pmatrix} 0 & 0 & c^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & c\mu^{-1} & 0 & 0 \\ 2 & 0 & 0 & 0 & 2 & 0 \\ 0 & 2z & 0 & 0 & 0 & -2z \\ 0 & 0 & c^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & -c\mu^{-1} & 0 & 0 \end{pmatrix}.$$

Si on pose enfin $\mathbf{U} = T^{-1}\mathbf{W}$, alors :

$$\mathcal{L}\mathbf{W} = \tilde{\mathcal{L}}\mathbf{U} \text{ avec } \mathbf{U} = \frac{1}{2} \begin{pmatrix} cw_3 + \mu^{-1}w_5 \\ cw_2 + \mu^{-1}w_6 \\ 2w_1 \\ 2w_4 \\ -cw_3 + \mu^{-1}w_5 \\ -cw_2 + \mu^{-1}w_6 \end{pmatrix} \text{ et } \tilde{\mathcal{L}} = i\omega l_3 + \Lambda \partial_x + A_y \partial_y + A_z \partial_z. \quad (4.120)$$

De même, on peut extraire $\Lambda^+ = T^{-1}M_x^+T$ et $\Lambda^- = T^{-1}M_x^-T$. Avec $\tilde{\mathbf{f}} = T^{-1}\mathbf{f}$, l'algorithme (4.119) peut se réécrire sur \mathbf{R}^3 comme :

$$\left\{ \begin{array}{l} \tilde{\mathcal{L}}\mathbf{U}^{1,p+1} = \tilde{\mathbf{f}}^1 \quad \text{dans } \Omega_1, \\ u_1^{1,p+1} = u_1^{2,p} \\ u_2^{1,p+1} = u_2^{2,p} \end{array} \right. \quad \text{sur } x = b, \quad \left\{ \begin{array}{l} \tilde{\mathcal{L}}\mathbf{U}^{2,p+1} = \tilde{\mathbf{f}}^2 \quad \text{dans } \Omega_2, \\ u_5^{2,p+1} = u_5^{1,p} \\ u_6^{2,p+1} = u_6^{1,p} \end{array} \right. \quad \text{sur } x = a, \quad (4.121)$$

où on a noté par $u_l^{j,p}$ la $l^{\text{ème}}$ composante du vecteur $\mathbf{U}^{j,p}$ qui représente l'approximation de la solution à l'itération p dans le domaine j . Etudions l'erreur locale $\mathbf{E}^{j,p} = \mathbf{U}^{j,p} - \mathbf{U}|_{\Omega_j}$ commise entre la solution approchée $\mathbf{U}^{j,p}$ à l'itération p et la solution exacte $\mathbf{U}|_{\Omega_j}$ pour $j = 1, 2$. Par linéarité, le système (4.121) vérifie aussi :

$$\left\{ \begin{array}{l} \tilde{\mathcal{L}}\mathbf{E}^{1,p+1} = \tilde{\mathbf{f}}^1 \quad \text{dans } \Omega_1, \\ e_1^{1,p+1} = e_1^{2,p} \\ e_2^{1,p+1} = e_2^{2,p} \end{array} \right. \quad \text{sur } x = b, \quad \left\{ \begin{array}{l} \tilde{\mathcal{L}}\mathbf{E}^{2,p+1} = \tilde{\mathbf{f}}^2 \quad \text{dans } \Omega_2, \\ e_5^{2,p+1} = e_5^{1,p} \\ e_6^{2,p+1} = e_6^{1,p} \end{array} \right. \quad \text{sur } x = a. \quad (4.122)$$

Il s'agit maintenant d'exprimer le taux de convergence de l'algorithme (4.122). Pour ce faire, on effectue une transformée de Fourier partielle suivant les variables y et z :

$$\hat{\mathbf{E}}^{j,p}(x, \mathbf{k}) = (\mathcal{F}_{(y,z)}\mathbf{E}^{j,p})(x, \mathbf{k}) = \int_{\mathbf{R}^2} e^{-i(k_1y+k_2z)} \mathbf{E}^{j,p}(x, y, z) dy dz \quad \text{où } \mathbf{k} = (k_1, k_2)^t.$$

Nous avons :

$$\tilde{\mathcal{L}}\mathbf{E}^{j,p} = 0 \Leftrightarrow \mathcal{F}_{(y,z)}\tilde{\mathcal{L}}\mathbf{E}^{j,p} = 0 \Leftrightarrow \Lambda \partial_x \hat{\mathbf{E}}^{j,p} + i(\omega \hat{\epsilon}_6 + k_1 A_y + k_2 A_z) \hat{\mathbf{E}}^{j,p} = 0. \quad (4.123)$$

En laissant de côté les indices correspondant aux sous-domaines et à l'itération, on obtient que l'équation (4.123) s'écrit localement :

$$-c\partial_x \hat{e}_1 + i \left(\omega \hat{e}_1 + \frac{1}{2}(k_2 c^2 \hat{e}_3 + k_1 c \mu^{-1} \hat{e}_4) \right) = 0, \quad (4.124a)$$

$$-c\partial_x \hat{e}_2 + i \left(\omega \hat{e}_2 + \frac{1}{2}(k_2 c \mu^{-1} \hat{e}_4 - k_1 c^2 \hat{e}_3) \right) = 0, \quad (4.124b)$$

$$\omega \hat{e}_3 - k_1(\hat{e}_2 + \hat{e}_6) + k_2(\hat{e}_1 + \hat{e}_5) = 0, \quad (4.124c)$$

$$\omega \hat{e}_4 + z(k_1(\hat{e}_1 - \hat{e}_5) + k_2(\hat{e}_2 - \hat{e}_6)) = 0, \quad (4.124d)$$

$$c\partial_x \hat{e}_5 + i \left(\omega \hat{e}_5 + \frac{1}{2}(k_2 c^2 \hat{e}_3 - k_1 c \mu^{-1} \hat{e}_4) \right) = 0, \quad (4.124e)$$

$$c\partial_x \hat{e}_6 + i \left(\omega \hat{e}_6 - \frac{1}{2}(k_2 c \mu^{-1} \hat{e}_4 + k_1 c^2 \hat{e}_3) \right) = 0. \quad (4.124f)$$

Comme dans le cas 2D (2.11), les variables \hat{e}_3 et \hat{e}_4 associées à la valeur propre 0 ne rentrent pas en compte dans la transmission des données pour les conditions d'interface, pour ce système de coordonnées. De ce fait, en remplaçant ces variables par leur expressions obtenues grâce à (4.124c) et (4.124d), les autres équations permettent d'aboutir à l'équation différentielle ordinaire suivante :

$$\partial_x \hat{\mathbf{E}} + \hat{M} \hat{\mathbf{E}} = 0, \quad (4.125)$$

avec :

$$\hat{M} = \frac{i}{2\tilde{\omega}} \begin{pmatrix} k_1^2 + k_2^2 - 2\tilde{\omega}^2 & 0 & -(k_1^2 - k_2^2) & -2k_1 k_2 \\ 0 & k_1^2 + k_2^2 - 2\tilde{\omega}^2 & -2k_1 k_2 & k_1^2 - k_2^2 \\ k_1^2 - k_2^2 & 2k_1 k_2 & -(k_1^2 + k_2^2 - 2\tilde{\omega}^2) & 0 \\ 2k_1 k_2 & -(k_1^2 - k_2^2) & 0 & -(k_1^2 + k_2^2 - 2\tilde{\omega}^2) \end{pmatrix},$$

où $\tilde{\omega} = \omega/c$ et où $\hat{\mathbf{E}} = (\hat{e}_1, \hat{e}_2, \hat{e}_5, \hat{e}_6)^t$ a été privé des variables \hat{e}_3 et \hat{e}_4 par abus de notation puisque celles-ci ne rentrent pas en compte dans le calcul de l'erreur. Les solutions de (4.125) s'expriment alors de la façon suivante :

$$\hat{\mathbf{E}} = [\gamma_1^1 \mathbf{V}_1^1(\mathbf{k}) + \gamma_2^1 \mathbf{V}_2^1(\mathbf{k})] e^{-\lambda_1(\mathbf{k})x} + [\gamma_1^2 \mathbf{V}_1^2(\mathbf{k}) + \gamma_2^2 \mathbf{V}_2^2(\mathbf{k})] e^{-\lambda_2(\mathbf{k})x}, \quad (4.126)$$

avec $(\gamma_1^1, \gamma_2^1, \gamma_1^2, \gamma_2^2) \in \mathbb{C}^4$ et où $\lambda_2(\mathbf{k}) = -\lambda_1(\mathbf{k}) = \sqrt{|\mathbf{k}|^2 - \tilde{\omega}^2}$ sont les valeurs propres de \hat{M} et :

$$\mathbf{V}_1^1(\mathbf{k}) = \begin{pmatrix} (\sqrt{|\mathbf{k}|^2 - \tilde{\omega}^2} + i\tilde{\omega})(k_1^2 - k_2^2) \\ (\sqrt{|\mathbf{k}|^2 - \tilde{\omega}^2} + i\tilde{\omega})2k_1 k_2 \\ (\sqrt{|\mathbf{k}|^2 - \tilde{\omega}^2} - i\tilde{\omega})|\mathbf{k}|^2 \\ 0 \end{pmatrix}, \quad \mathbf{V}_2^1(\mathbf{k}) = \begin{pmatrix} (\sqrt{|\mathbf{k}|^2 - \tilde{\omega}^2} + i\tilde{\omega})2k_1 k_2 \\ (\sqrt{|\mathbf{k}|^2 - \tilde{\omega}^2} + i\tilde{\omega})(k_2^2 - k_1^2) \\ 0 \\ (\sqrt{|\mathbf{k}|^2 - \tilde{\omega}^2} - i\tilde{\omega})|\mathbf{k}|^2 \end{pmatrix},$$

$$\mathbf{V}_1^2(\mathbf{k}) = \begin{pmatrix} (\sqrt{|\mathbf{k}|^2 - \tilde{\omega}^2} - i\tilde{\omega})(k_1^2 - k_2^2) \\ (\sqrt{|\mathbf{k}|^2 - \tilde{\omega}^2} - i\tilde{\omega})2k_1 k_2 \\ (\sqrt{|\mathbf{k}|^2 - \tilde{\omega}^2} + i\tilde{\omega})|\mathbf{k}|^2 \\ 0 \end{pmatrix}, \quad \mathbf{V}_2^2(\mathbf{k}) = \begin{pmatrix} (\sqrt{|\mathbf{k}|^2 - \tilde{\omega}^2} - i\tilde{\omega})2k_1 k_2 \\ (\sqrt{|\mathbf{k}|^2 - \tilde{\omega}^2} - i\tilde{\omega})(k_2^2 - k_1^2) \\ 0 \\ (\sqrt{|\mathbf{k}|^2 - \tilde{\omega}^2} + i\tilde{\omega})|\mathbf{k}|^2 \end{pmatrix},$$

sont les vecteurs propres de \hat{M} . Dans la suite, on remplace $\tilde{\omega}$ par ω pour simplifier l'écriture. On revient au contexte de de l'algorithme de Schwarz (4.122) où le vecteur erreur s'exprime dans chaque sous-domaine comme :

$$\hat{\mathbf{E}}^{j,p} = [(\gamma_1^1)^{j,p} \mathbf{V}_1^1(\mathbf{k}) + (\gamma_2^1)^{j,p} \mathbf{V}_2^1(\mathbf{k})] e^{-\lambda_1(\mathbf{k})x} + [(\gamma_1^2)^{j,p} \mathbf{V}_1^2(\mathbf{k}) + (\gamma_2^2)^{j,p} \mathbf{V}_2^2(\mathbf{k})] e^{-\lambda_2(\mathbf{k})x}.$$

En suivant exactement la même démarche que pour les équations 2D, c'est-à-dire en invoquant le fait que les solutions sont bornées pour x tendant vers $\pm\infty$, on déduit que les solutions locales s'écrivent :

$$\begin{cases} \hat{\mathbf{E}}^{1,p}(x, \mathbf{k}) &= e^{x\sqrt{|\mathbf{k}|^2 - \omega^2}} ((\gamma_1^1)^{1,p} \mathbf{V}_1^1(\mathbf{k}) + (\gamma_2^1)^{1,p} \mathbf{V}_2^1(\mathbf{k})), \\ \hat{\mathbf{E}}^{2,p}(x, \mathbf{k}) &= e^{-x\sqrt{|\mathbf{k}|^2 - \omega^2}} ((\gamma_1^2)^{2,p} \mathbf{V}_1^2(\mathbf{k}) + (\gamma_2^2)^{2,p} \mathbf{V}_2^2(\mathbf{k})), \end{cases} \quad (4.127)$$

L'algorithme (4.122) devient dans ces conditions :

$$\left\{ \begin{array}{l} \partial_x \hat{\mathbf{E}}^{1,p+1} + \hat{M} \hat{\mathbf{E}}^{1,p+1} = 0 \quad \text{dans } \Omega_1 \\ \hat{e}_1^{1,p+1} = \hat{e}_1^{2,p} \\ \hat{e}_2^{1,p+1} = \hat{e}_2^{2,p} \end{array} \right. \quad \text{sur } x = b \quad \left\{ \begin{array}{l} \partial_x \hat{\mathbf{E}}^{2,p+1} + \hat{M} \hat{\mathbf{E}}^{2,p+1} = 0 \quad \text{dans } \Omega_2 \\ \hat{e}_5^{1,p+1} = \hat{e}_5^{2,p} \\ \hat{e}_6^{1,p+1} = \hat{e}_6^{2,p} \end{array} \right. \quad \text{sur } x = a \quad (4.128)$$

En introduisant dans les conditions interface de chaque sous-domaine les solutions (4.127) on obtient pour le premier et le deuxième sous-domaine :

$$\Leftrightarrow \left\{ \begin{array}{l} \left(\begin{array}{l} \hat{e}_1 \\ \hat{e}_2 \end{array} \right)^{1,p+1} = \left(\begin{array}{l} \hat{e}_1 \\ \hat{e}_2 \end{array} \right)^{2,p} \\ \left(\begin{array}{l} \hat{e}_5 \\ \hat{e}_6 \end{array} \right)^{2,p+1} = \left(\begin{array}{l} \hat{e}_5 \\ \hat{e}_6 \end{array} \right)^{1,p} \\ e^{b\sqrt{|\mathbf{k}|^2 - \omega^2}} (\sqrt{|\mathbf{k}|^2 - \omega^2} + i\omega) \begin{pmatrix} (\gamma_1^1)^{1,p+1}(k_1^2 - k_2^2) + 2(\gamma_2^1)^{1,p+1}k_1k_2 \\ 2(\gamma_1^1)^{1,p+1}k_1k_2 - (\gamma_2^1)^{1,p+1}(k_1^2 - k_2^2) \end{pmatrix} \\ = e^{-b\sqrt{|\mathbf{k}|^2 - \omega^2}} (\sqrt{|\mathbf{k}|^2 - \omega^2} - i\omega) \begin{pmatrix} (\gamma_1^2)^{2,p}(k_1^2 - k_2^2) + 2(\gamma_2^2)^{2,p}k_1k_2 \\ 2(\gamma_1^2)^{2,p}k_1k_2 - (\gamma_2^2)^{2,p}(k_1^2 - k_2^2) \end{pmatrix}, \\ e^{-a\sqrt{|\mathbf{k}|^2 - \omega^2}} (\sqrt{|\mathbf{k}|^2 - \omega^2} + i\omega) \begin{pmatrix} (\gamma_1^2)^{2,p+1}|\mathbf{k}|^2 \\ (\gamma_2^2)^{2,p+1}|\mathbf{k}|^2 \end{pmatrix} \\ = e^{a\sqrt{|\mathbf{k}|^2 - \omega^2}} (\sqrt{|\mathbf{k}|^2 - \omega^2} - i\omega) \begin{pmatrix} (\gamma_1^1)^{1,p}|\mathbf{k}|^2 \\ (\gamma_2^1)^{1,p}|\mathbf{k}|^2 \end{pmatrix}. \end{array} \right. \quad (4.129)$$

On pose $K = \begin{pmatrix} k_1^2 - k_2^2 & 2k_1k_2 \\ 2k_1k_2 & k_2^2 - k_1^2 \end{pmatrix}$ dont le déterminant $\det K = -|\mathbf{k}|^4$. Si par ailleurs, $\gamma^j = (\gamma_1^j, \gamma_2^j)^t$, les expressions (4.129) deviennent :

$$\left\{ \begin{array}{l} e^{b\sqrt{|\mathbf{k}|^2 - \omega^2}} (\sqrt{|\mathbf{k}|^2 - \omega^2} + i\omega) K (\gamma^1)^{1,p+1} = e^{-b\sqrt{|\mathbf{k}|^2 - \omega^2}} (\sqrt{|\mathbf{k}|^2 - \omega^2} - i\omega) K (\gamma^2)^{2,p}, \\ e^{-a\sqrt{|\mathbf{k}|^2 - \omega^2}} (\sqrt{|\mathbf{k}|^2 - \omega^2} + i\omega) |\mathbf{k}|^2 \mathbf{l}_2 (\gamma^2)^{2,p+1} = e^{a\sqrt{|\mathbf{k}|^2 - \omega^2}} (\sqrt{|\mathbf{k}|^2 - \omega^2} - i\omega) |\mathbf{k}|^2 \mathbf{l}_2 (\gamma^1)^{1,p}. \end{array} \right.$$

Puque le déterminant de K est non-nul $\forall \mathbf{k} \neq 0$, nous pouvons simplifier ces expressions et les factoriser pour obtenir :

$$\left\{ \begin{array}{l} (\gamma^1)^{1,p+1} = (\gamma^2)^{2,p} \left(\frac{\sqrt{|\mathbf{k}|^2 - \omega^2} - i\omega}{\sqrt{|\mathbf{k}|^2 - \omega^2} + i\omega} \right) e^{-2b\sqrt{|\mathbf{k}|^2 - \omega^2}}, \\ (\gamma^2)^{2,p+1} = (\gamma^1)^{1,p} \left(\frac{\sqrt{|\mathbf{k}|^2 - \omega^2} - i\omega}{\sqrt{|\mathbf{k}|^2 - \omega^2} + i\omega} \right) e^{2a\sqrt{|\mathbf{k}|^2 - \omega^2}}, \end{array} \right. \quad (4.130)$$

et ensuite :

$$\frac{|(\gamma^1)^{1,p+1}|}{|(\gamma^1)^{1,p-1}|} = \frac{|(\gamma^2)^{2,p+1}|}{|(\gamma^2)^{2,p-1}|} = \left| \frac{\sqrt{|\mathbf{k}|^2 - \omega^2} - i\omega}{\sqrt{|\mathbf{k}|^2 - \omega^2} + i\omega} \right|^2 |e^{-2\delta\sqrt{|\mathbf{k}|^2 - \omega^2}}|, \quad (4.131)$$

où on a noté $\delta = b - a$ la taille du recouvrement entre les sous-domaines. Si on désigne par $\rho(k, \delta)$ le taux de réduction de l'erreur entre deux itérations successives on déduit de (4.131) que :

$$\rho(k, \delta) = \left| \left(\frac{\sqrt{|\mathbf{k}|^2 - \omega^2} - i\omega}{\sqrt{|\mathbf{k}|^2 - \omega^2} + i\omega} \right) e^{-\delta\sqrt{|\mathbf{k}|^2 - \omega^2}} \right|. \quad (4.132)$$

Nous constatons alors que le taux qui exprime la vitesse de convergence de l'algorithme de Schwarz pour les équations de Maxwell 3D est identique à celui obtenu pour les équations de Maxwell 2D (2.20). La différence réside dans la variable de Fourier qui a été étendue à deux dimensions pour le traitement des équations tridimensionnelles. On peut donc s'attendre à retrouver des résultats de convergence identiques à ceux qui ont été obtenus en 2D.

Chapitre 4

Résultats numériques en 3D

Les formulations de type Galerkin discontinu GD-P0 et GD-P1 étudiées au chapitre 2 sont ici appliquées à la résolution numérique des équations de Maxwell 3D en régime fréquentiel. Ce chapitre est structuré en trois sections. Dans la section 4.1, on considère des cas de calcul de complexité croissante afin de valider les méthodes GD-P0 et GD-P1. La section 4.2 traite de la résolution en mode séquentiel des systèmes linéaires résultant des formulations GD-P0 et GD-P1, par une méthode directe adaptée aux matrices creuses et par une méthode itérative par sous-espace de Krylov préconditionnée. Enfin, la section 4.3 est consacrée à la résolution numérique en mode parallèle avec notamment une étude détaillée des performances d'une méthode de décomposition de domaine proposée au chapitre 3.

4.1 Validation numérique

4.1.1 Propagation d'une onde plane dans le vide

Le premier cas test considéré consiste en la propagation d'une onde plane dans le vide. Ce cas test n'a clairement pas d'intérêt physique mais permet une validation préliminaire des formulations GD-P0 et GD-P1 développées dans notre étude. En outre, la génération du maillage de calcul associé à ce cas test est relativement directe et de ce fait, on peut aisément étudier l'influence de la taille du problème (i.e. de la finesse de la discrétisation) sur la précision des calculs avec les formulations GD-P0 et GD-P1, ainsi que les performances des algorithmes de résolution (objet de la section 4.3). Le domaine de calcul est un cube unitaire. Sur chacune des faces de ce cube, on applique une condition absorbante de type Silver-Müller. Les calculs sont effectués en champ total pour des fréquences de 300 MHz, 600 MHz et 900 MHz. Ils reposent sur des maillages tétraédriques obtenus par subdivision d'un maillage cartésien caractérisés par les nombres N_x , N_y et N_z de points suivants les axes (Ox) , (Oy) et (Oz) . Chaque cellule de ce maillage cartésien est ensuite décomposée en six tétraèdres et le maillage tétraédrique résultant est uniforme. Un exemple de tel maillage est montré sur la figure 1.1 (coupe dans le plan $Y = 0$).

Les solutions résultant des méthodes GD-P0 et GD-P1 sont montrées sur les figures 1.3 (méthode GD-P0, maillage $N_x = 41$, $N_y = N_z = 15$) et 1.4 (méthode GD-P1, maillage $N_x = 31$, $N_y = N_z = 15$) pour une fréquence $F = 600$ MHz, sous la forme des lignes de contour de la partie réelle de E_z dans le plan $Y = 0$. Pour comparaison, la solution exacte est représentée sur la figure 1.2 (maillage $N_x = 41$, $N_y = N_z = 15$). De plus, nous avons représenté sur les figures 1.5 à 1.8 des coupes 1D de la partie réelle de E_z (évolution de E_z en fonction de X pour $Y = Z = 0$), pour différentes

fréquences et différents maillages. Les solutions obtenues pour une fréquence $F = 900$ MHz permettent ici de comparer la précision des méthodes GD-P0 et GD-P1. A cette fréquence, la longueur d'onde est égale à $\lambda = 0.3333$ m. Ainsi, le maillage tel que $N_x = 41$ conduit à une discrétisation suivant (Ox) avec environ 14 points par longueur d'onde. Sur ce maillage, la méthode GD-P1 fournit une solution tout à fait acceptable ce qui n'est pas le cas de la méthode GD-P0. La figure 1.8 ne concerne que cette dernière méthode et montre qu'un maillage comprenant près de 24 points par longueur ($N_x = 71$) d'onde suivant (Ox) ne suffit pas à obtenir une solution correcte. Face à ce type de comportement, il est naturel de se demander si la méthode GD-P0 converge et à quelle vitesse. Pour répondre à cette question, tout au moins sur le plan numérique, nous avons représenté sur les figures 1.9 et 1.10 les courbes de convergence numérique des méthodes GD-P0 et GD-P1 dans le cas d'un maillage uniforme (comme ceux utilisés dans les expériences considérées dans cette section de résultats) et d'un maillage non-uniforme. Ces courbes ont été obtenues dans le cadre de la résolution numérique des équations de Maxwell fréquentielles 2D (formulation TM). Elles montrent que la convergence de la méthode GD-P0 est très sensible à la régularité du maillage, ce qui a déjà été observé dans la résolution numérique des équations de Maxwell en domaine temporel [Fezoui et al., 2005]. Elles confirment aussi que dans le cas d'un maillage uniforme, l'utilisation d'un maillage suffisamment fin peut s'avérer nécessaire dans des situations comme celles de la figure 1.8.

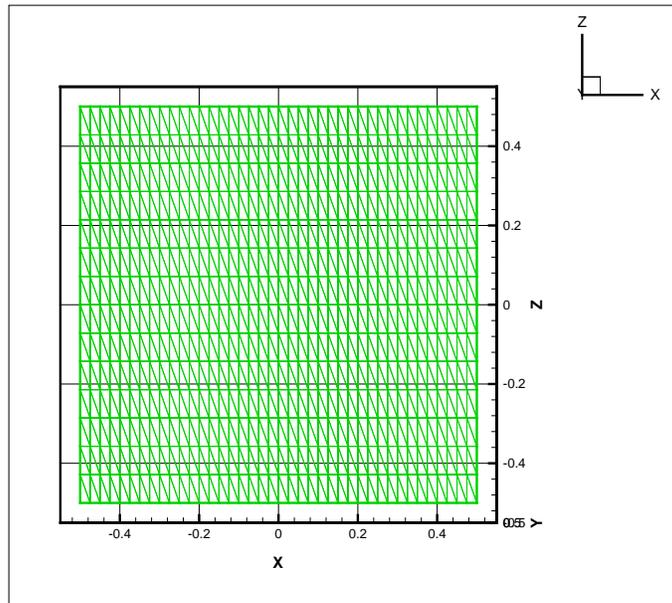


FIG. 1.1 – Equations de Maxwell 3D : propagation d'une onde plane dans le vide.
 $N_x = 41$, $N_y = N_z = 15$: maillage dans le plan $Y = 0$.

4.1.2 Diffraction d'une onde plane par une sphère

On considère ici le problème de la diffraction d'une onde plane par une sphère de rayon $R = 1$ m parfaitement conductrice. La fréquence F est fixée à 600 MHz et l'onde plane est définie par :

$$\mathbf{k} = (0, 0, k_z)^t, \quad \mathbf{E} = (E_x, 0, 0)^t \quad \text{et} \quad \mathbf{H} = (0, H_y, 0)^t$$

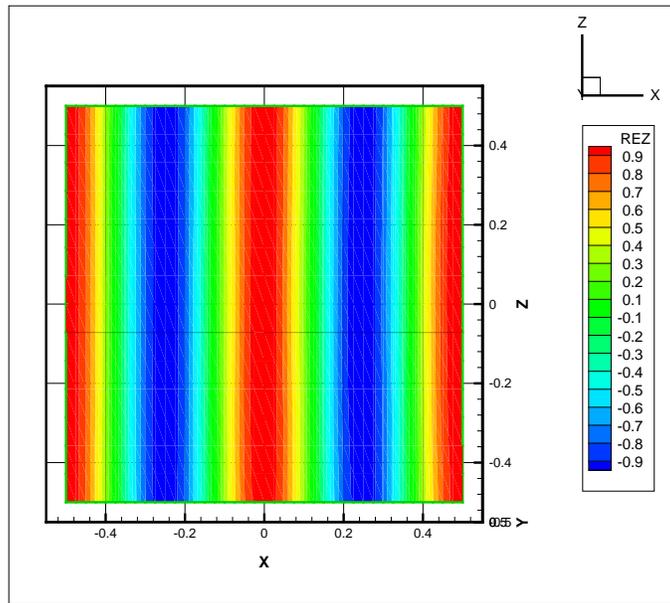


FIG. 1.2 – Equations de Maxwell 3D : propagation d’une onde plane dans le vide.
 $N_x = 41$, $N_y = N_z = 15$, solution exacte, $F = 600$ MHz.
 Partie réelle de E_z : lignes de contour dans le plan $Y = 0$.

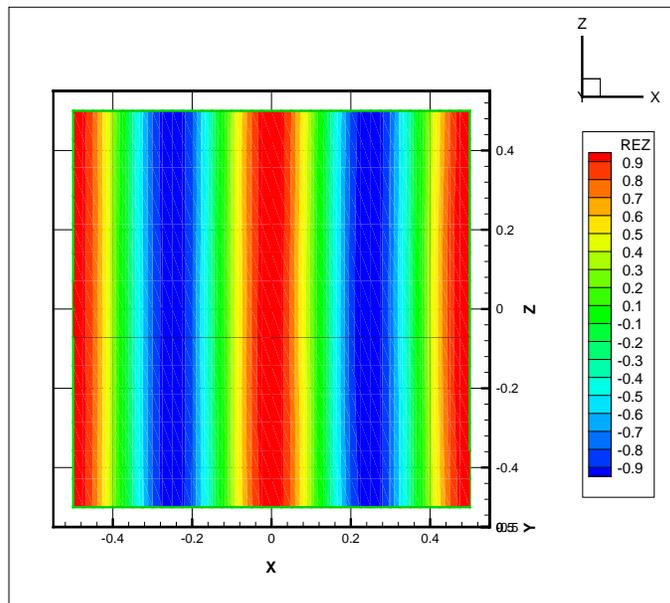


FIG. 1.3 – Equations de Maxwell 3D : propagation d’une onde plane dans le vide.
 $N_x = 41$, $N_y = N_z = 15$, méthode GD-P0, $F = 600$ MHz.
 Partie réelle de E_z : lignes de contour dans le plan $Y = 0$.

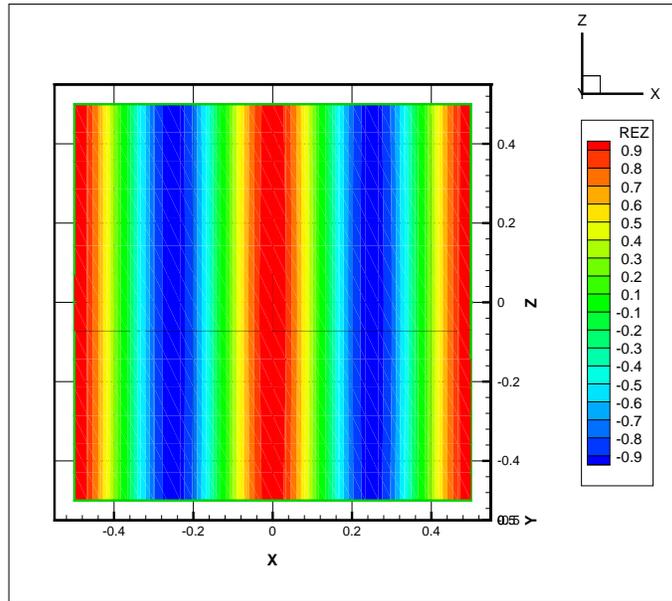


FIG. 1.4 – Equations de Maxwell 3D : propagation d’une onde plane dans le vide.
 $N_x = 31$, $N_y = N_z = 15$, méthode GD-P1, $F = 600$ MHz.
 Partie réelle de E_z : lignes de contour dans le plan $Y = 0$.

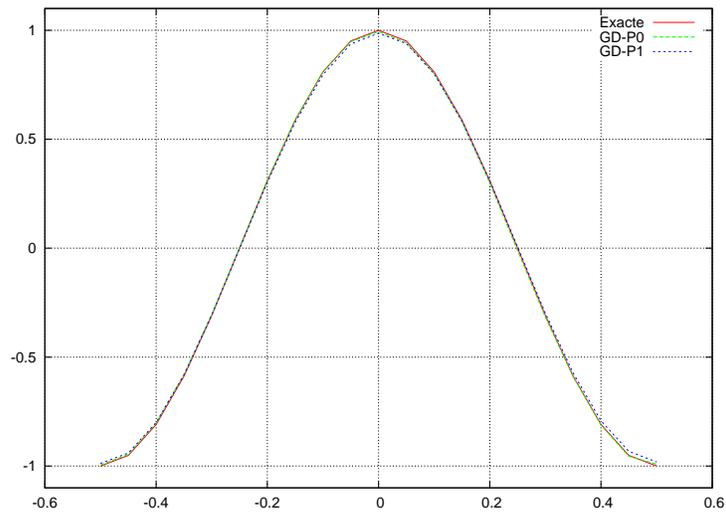


FIG. 1.5 – Equations de Maxwell 3D : propagation d’une onde plane dans le vide
 $N_x = 21$, $N_y = N_z = 11$, $F = 300$ MHz
 Partie réelle de E_z : coupe 1D en $Y = Z = 0$

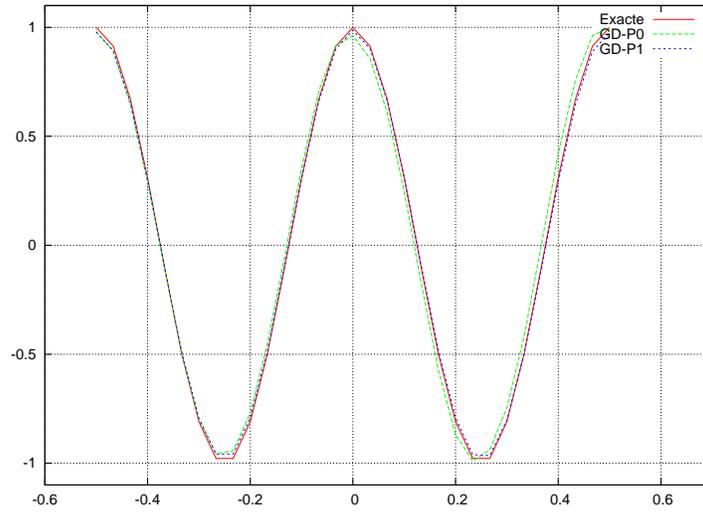


FIG. 1.6 – Equations de Maxwell 3D : propagation d’une onde plane dans le vide.
 $N_x = 31, N_y = N_z = 15, F = 600$ MHz.
 Partie réelle de E_z : coupe 1D en $Y = Z = 0$.

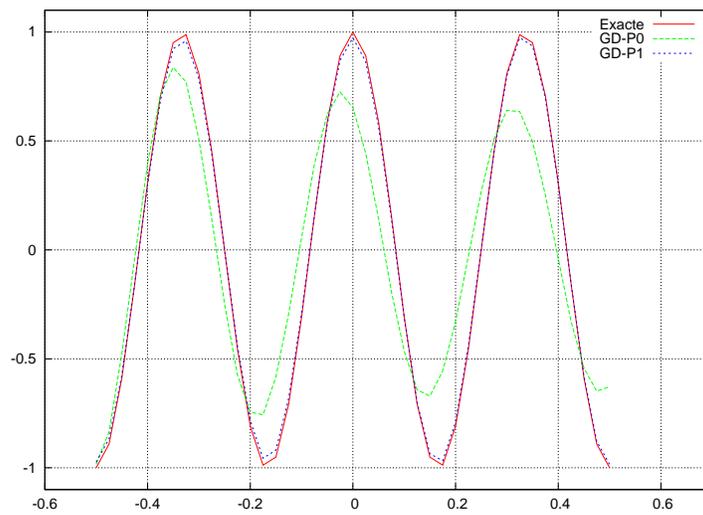


FIG. 1.7 – Equations de Maxwell 3D : propagation d’une onde plane dans le vide.
 $N_x = 41, N_y = N_z = 15, F = 900$ MHz.
 Partie réelle de E_z : coupe 1D en $Y = Z = 0$.

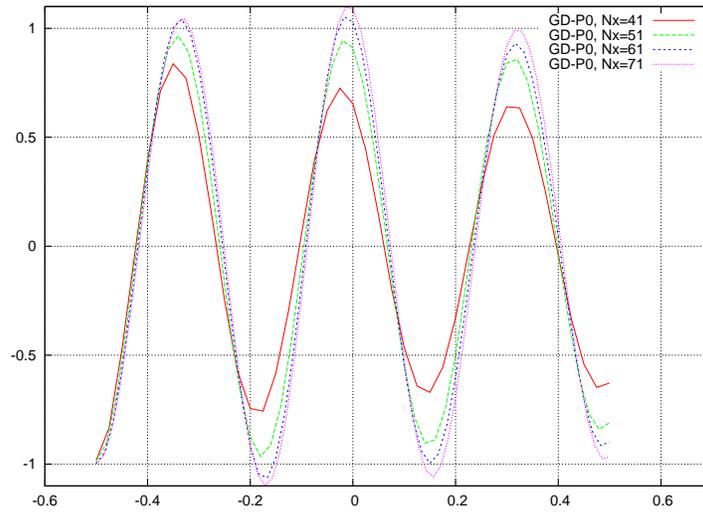


FIG. 1.8 – Equations de Maxwell 3D : propagation d’une onde plane dans le vide.
Méthode GD-P0, $N_y = N_z = 15$, $F = 900$ MHz.
Partie réelle de E_z : coupe 1D en $Y = Z = 0$.

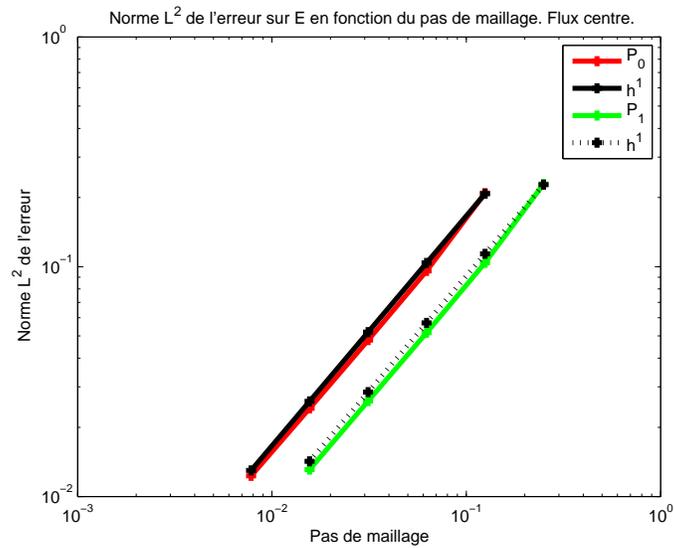


FIG. 1.9 – Equations de Maxwell 3D : propagation d’une onde plane dans le vide.
Convergence numérique des méthodes GD-P0 et GD-P1.
Cas d’un maillage uniforme.

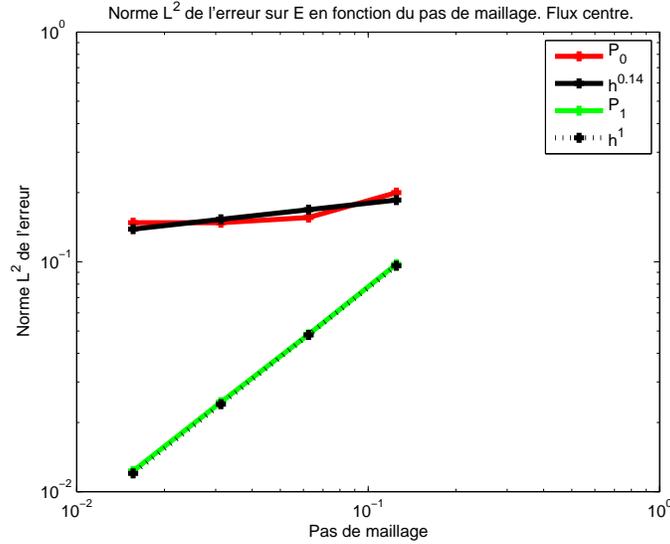


FIG. 1.10 – Equations de Maxwell 3D : propagation d’une onde plane dans le vide.
Convergence numérique des méthodes GD-P0 et GD-P1.
Cas d’un maillage non-uniforme.

avec $E_x = e^{-i\mathbf{k}\cdot\mathbf{x}}$ et $\mathbf{H} = (\mathbf{k} \times \mathbf{E})/z_0$ où $z_0 = \sqrt{\mu_0/\varepsilon_0}$ est l’impédance du vide. La longueur d’onde correspondant à cette fréquence est $\lambda = 0.5$ m. Les simulations numériques reposent sur trois maillages tétraédriques dont les caractéristiques sont résumées dans la table 4.1. Dans cette table, les quantités L_{\min} , L_{\max} et L_{avg} désignent respectivement les longueurs minimale, maximale et moyenne des arêtes des maillages en question. Notons aussi que dans ces maillages, la frontière absorbante sur laquelle on applique une condition de type Silver-Müller est fixée à $R + \lambda = 1.5$ m. On présente ici des comparaisons avec la solution analytique du problème exprimée à l’aide des potentiels de Debye [Duruflé, 2006].

Maillage	# sommets	# tétraèdres	L_{\min} (en m)	L_{\max} (en m)	L_{avg} (en m)
M1	70,422	384,000	0.039267	0.118029	0.066279
M2	151,452	843,648	0.030206	0.091805	0.051038
M3	244,834	1,382,400	0.025665	0.078819	0.043431

TAB. 4.1 – Equations de Maxwell 3D : diffraction d’une onde plane par une sphère.
Caractéristiques des maillages tétraédriques, $F = 600$ MHz.

Les solutions obtenues sont représentées graphiquement sous la forme des lignes de contour de la partie réelle des composantes E_x , E_y et E_z dans le plan $Z = 0$. La solution exacte est montrée sur les figures 1.12, 1.17 et 1.22. Les figures 1.13 à 1.15, 1.18 à 1.20 et 1.23 à 1.25 correspondent aux solutions fournies par la méthode GD-P0 appliquée aux maillages M1 à M3, pour la partie réelle des composantes E_x , E_y et E_z . Enfin, les figures 1.16, 1.21 et 1.26 correspondent à la solution fournie par la méthode GD-P1 appliquée au maillage M1 uniquement. On constate que cette dernière combinaison (méthode GD-P1, maillage M1) conduit à une solution très proche de la solution exacte en dépit d’une discrétisation relativement grossière (les longueurs L_{\min} et L_{\max} équivalent dans ce cas à environ 13 points et 4 points par longueur d’onde respectivement).

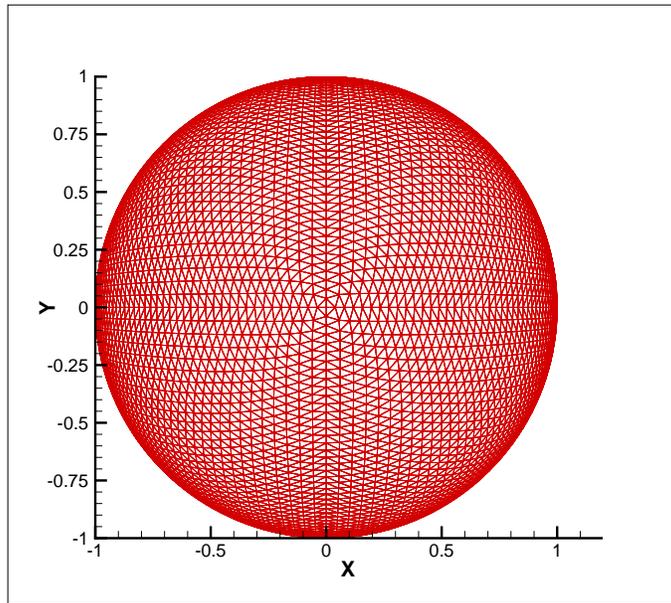


FIG. 1.11 – Equations de Maxwell 3D : diffraction d’une onde plane par une sphère parfaitement conductrice.

Maillage M1 : triangulation de la surface de la sphère.

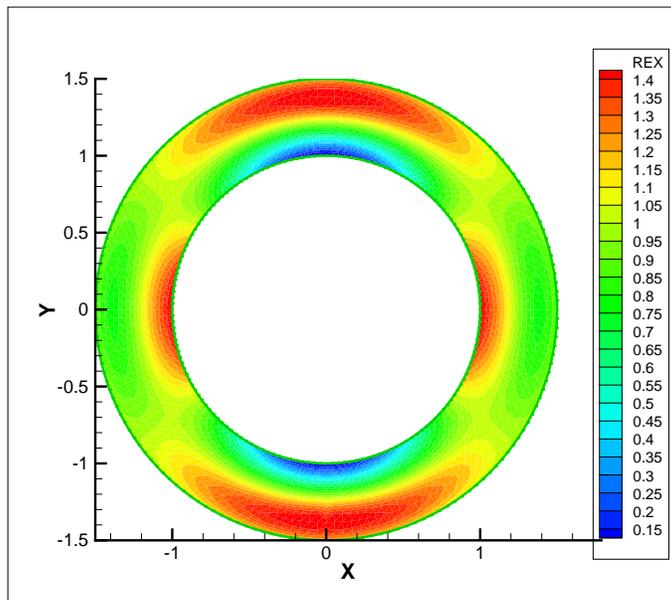


FIG. 1.12 – Equations de Maxwell 3D : diffraction d’une onde plane par une sphère parfaitement conductrice.

Maillage M1, solution exacte, $F = 600$ MHz.
Partie réelle de E_x : lignes de contour dans le plan $Z = 0$.

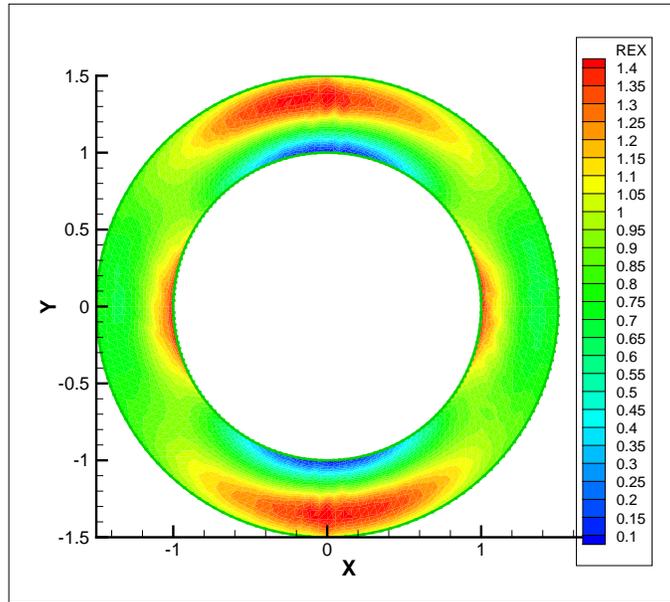


FIG. 1.13 – Equations de Maxwell 3D : diffraction d’une onde plane par une sphère parfaitement conductrice.

Maillage M1, méthode GD-P0, $F = 600$ MHz.
Partie réelle de E_x : lignes de contour dans le plan $Z = 0$.

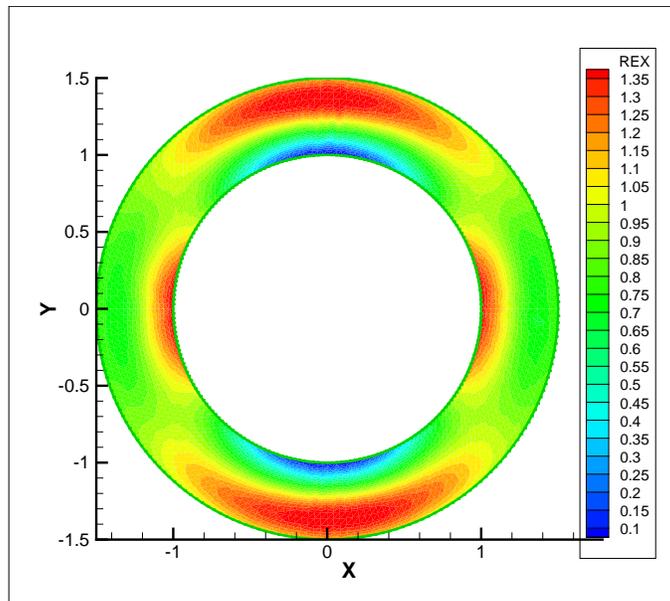


FIG. 1.14 – Equations de Maxwell 3D : diffraction d’une onde plane par une sphère parfaitement conductrice.

Maillage M2, méthode GD-P0, $F = 600$ MHz.
Partie réelle de E_x : lignes de contour dans le plan $Z = 0$.

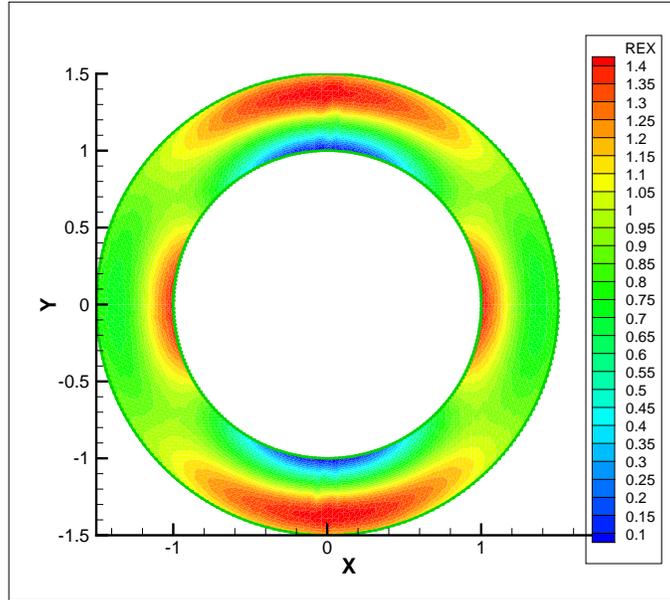


FIG. 1.15 – Equations de Maxwell 3D : diffraction d’une onde plane par une sphère parfaitement conductrice.

Maillage M3, méthode GD-P0, $F = 600$ MHz.
Partie réelle de E_x : lignes de contour dans le plan $Z = 0$.

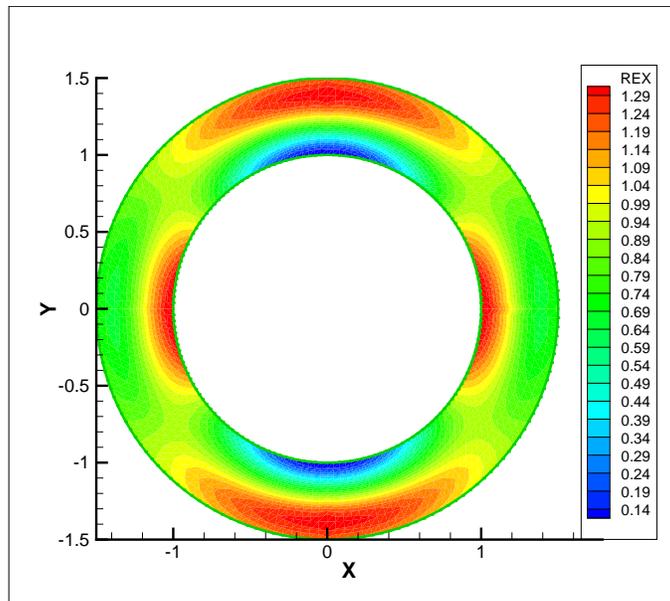


FIG. 1.16 – Equations de Maxwell 3D : diffraction d’une onde plane par une sphère parfaitement conductrice.

Maillage M1, méthode GD-P1, $F = 600$ MHz.
Partie réelle de E_x : lignes de contour dans le plan $Z = 0$.

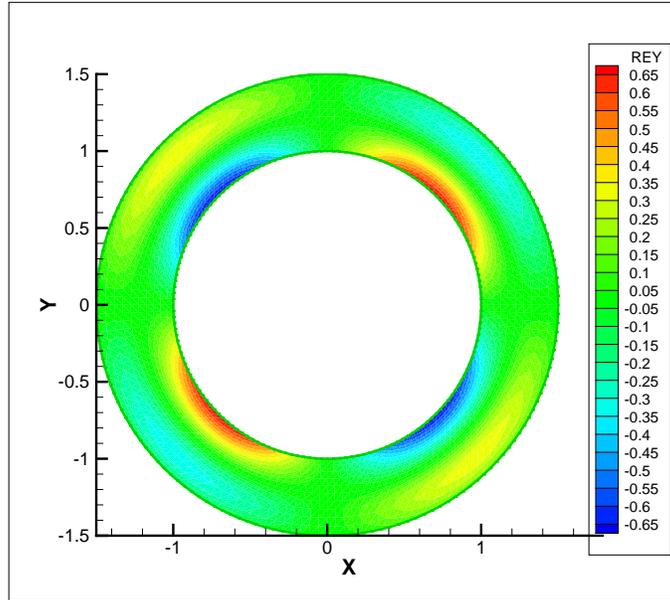


FIG. 1.17 – Equations de Maxwell 3D : diffraction d’une onde plane par une sphere parfaitement conductrice.

Maillage M1, solution exacte, $F = 600$ MHz.
Partie réelle de E_y : lignes de contour dans le plan $Z = 0$.

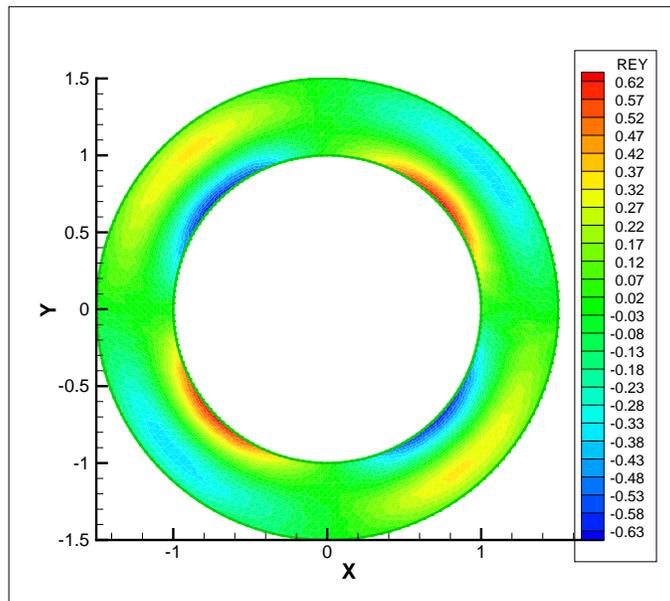


FIG. 1.18 – Equations de Maxwell 3D : diffraction d’une onde plane par une sphere parfaitement conductrice.

Maillage M1, méthode GD-P0, $F = 600$ MHz.
Partie réelle de E_y : lignes de contour dans le plan $Z = 0$.

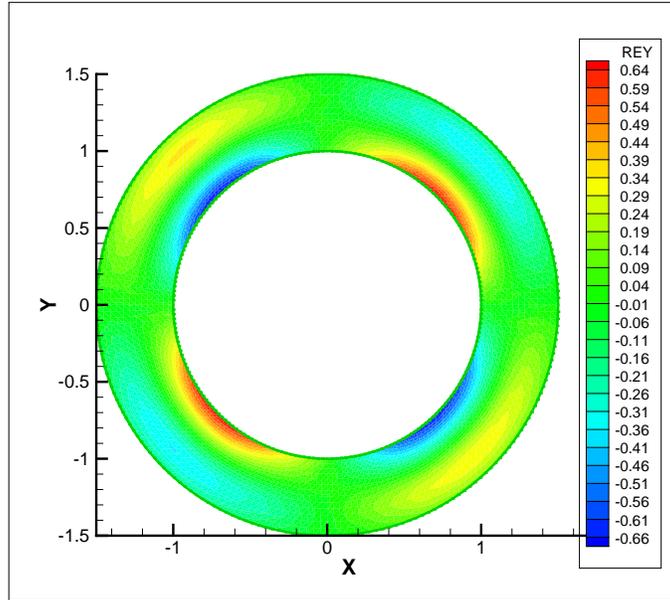


FIG. 1.19 – Equations de Maxwell 3D : diffraction d’une onde plane par une sphère parfaitement conductrice.

Maillage M2, méthode GD-P0, $F = 600$ MHz.
Partie réelle de E_y : lignes de contour dans le plan $Z = 0$.

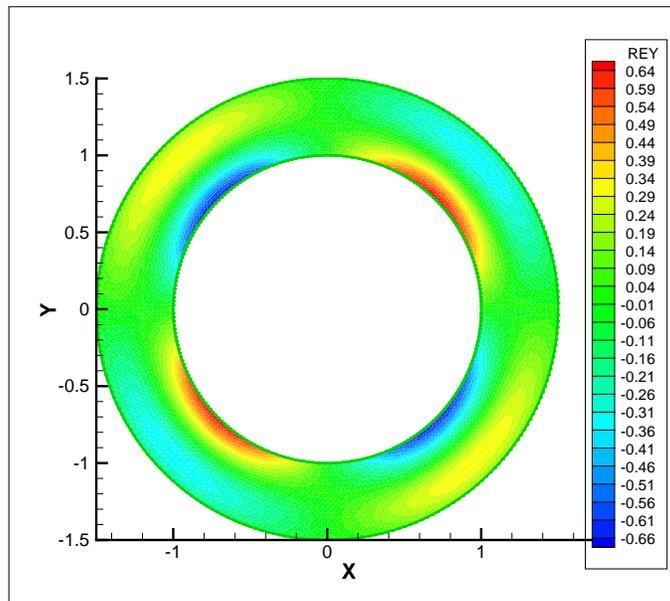


FIG. 1.20 – Equations de Maxwell 3D : diffraction d’une onde plane par une sphère parfaitement conductrice.

Maillage M3, méthode GD-P0, $F = 600$ MHz.
Partie réelle de E_y : lignes de contour dans le plan $Z = 0$.

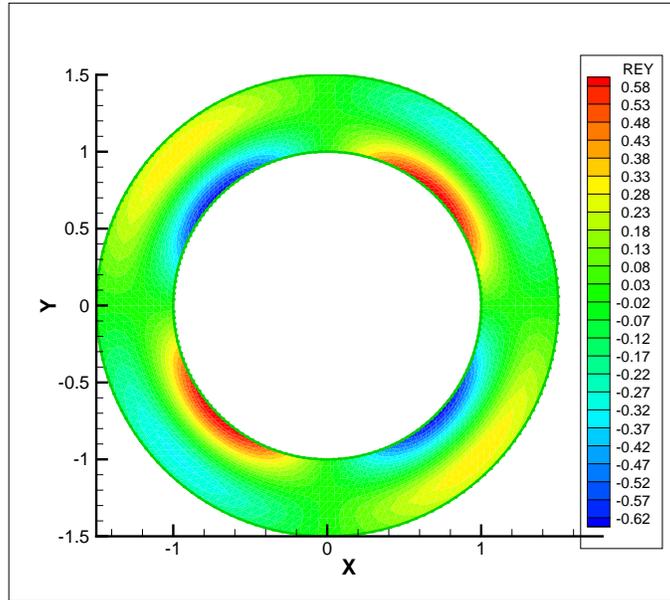


FIG. 1.21 – Equations de Maxwell 3D : diffraction d’une onde plane par une sphère parfaitement conductrice.

Maillage M1, méthode GD-P1, $F = 600$ MHz.
Partie réelle de E_y : lignes de contour dans le plan $Z = 0$.

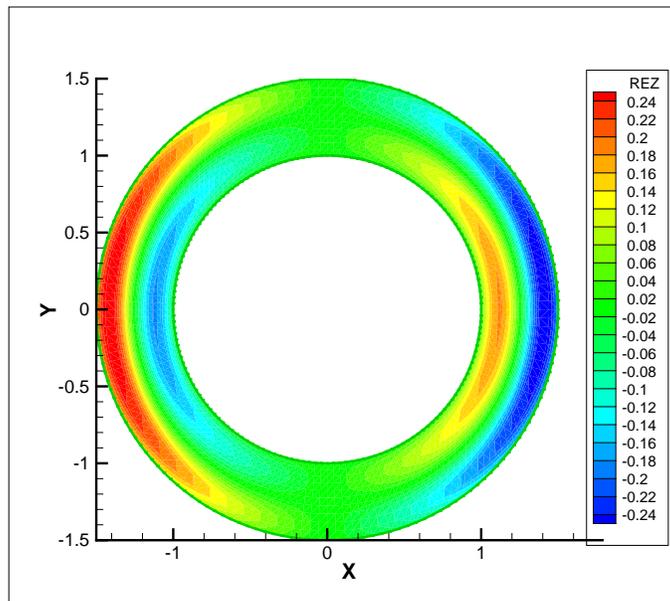


FIG. 1.22 – Equations de Maxwell 3D : diffraction d’une onde plane par une sphère parfaitement conductrice.

Maillage M1, solution exacte, $F = 600$ MHz.
Partie réelle de E_z : lignes de contour dans le plan $Z = 0$.

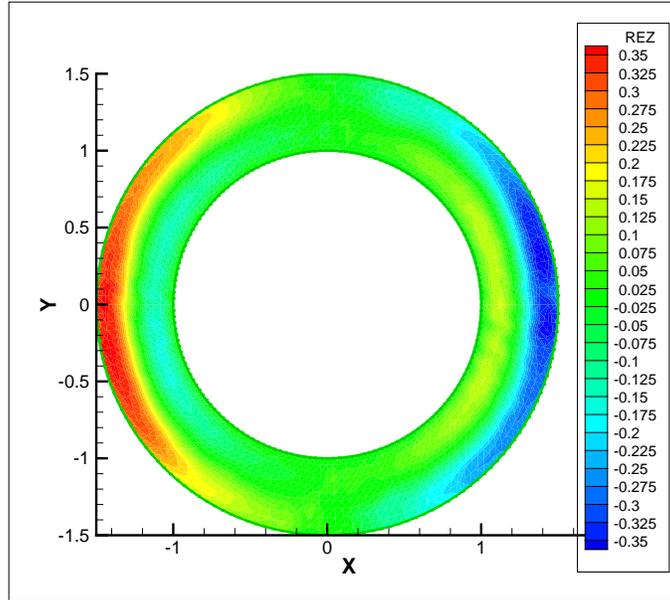


FIG. 1.23 – Equations de Maxwell 3D : diffraction d’une onde plane par une sphère parfaitement conductrice.

Maillage M1, méthode GD-P0, $F = 600$ MHz.
Partie réelle de E_z : lignes de contour dans le plan $Z = 0$.

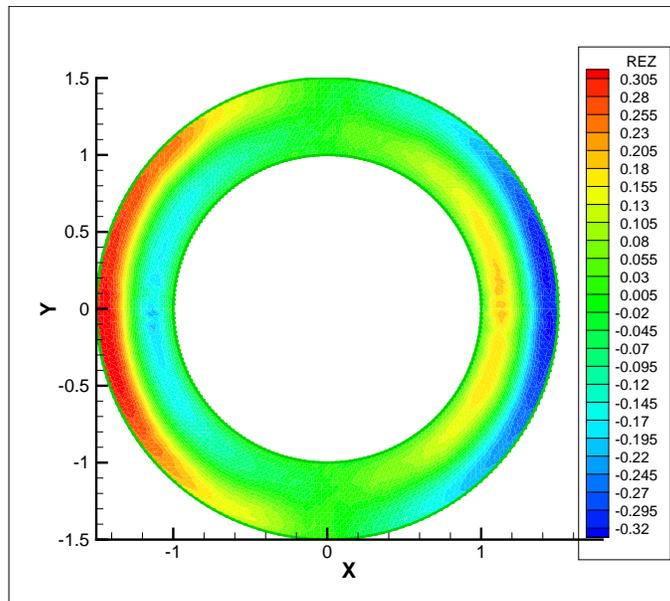


FIG. 1.24 – Equations de Maxwell 3D : diffraction d’une onde plane par une sphère parfaitement conductrice.

Maillage M2, méthode GD-P0, $F = 600$ MHz.
Partie réelle de E_z : lignes de contour dans le plan $Z = 0$.

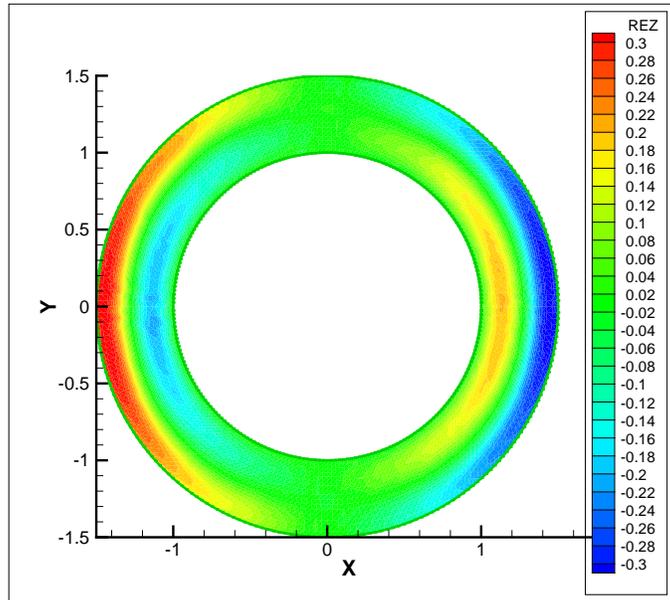


FIG. 1.25 – Equations de Maxwell 3D : diffraction d’une onde plane par une sphère parfaitement conductrice.

Maillage M3, méthode GD-P0, $F = 600$ MHz.
Partie réelle de E_z : lignes de contour dans le plan $Z = 0$.

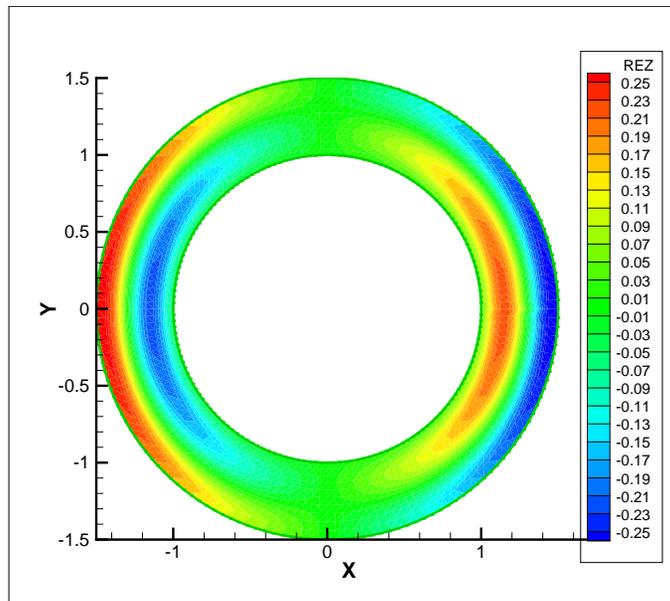


FIG. 1.26 – Equations de Maxwell 3D : diffraction d’une onde plane par une sphère parfaitement conductrice.

Maillage M1, méthode GD-P1, $F = 600$ MHz.
Partie réelle de E_z : lignes de contour dans le plan $Z = 0$.

4.1.3 Diffraction d'une onde plane par un cube

Cette fois, on étudie la diffraction d'une onde plane par un cube de côté $C = 1/3$ m placé au centre d'un cube unitaire sur les faces duquel on applique une condition absorbante de type Silver-Müller. Le cube interne est supposé parfaitement conducteur. La fréquence est fixée à 900 MHz et l'onde plane est définie par :

$$\mathbf{k} = (k_x, 0, 0)^t, \quad \mathbf{E} = (0, E_y, 0)^t \quad \text{et} \quad \mathbf{H} = (0, 0, H_z)^t$$

Dans le vide, la longueur d'onde correspondant à cette fréquence est $\lambda = 1/3$ m. On utilise un maillage tétraédrique uniforme dont les caractéristiques sont résumées dans la table 4.2 (voir aussi la figure 1.27 pour une vue du maillage dans le plan $Z = 0.5$). Les solutions fournies par les méthodes GD-P0 et GD-P1 sont ici comparées avec celles obtenues avec la méthode GDDT-P1 décrite dans [Piperno and Fezoui, 2003] pour la résolution des équations de Maxwell en domaine temporel. Dans la méthode GDDT-P1, les équations de Maxwell sont discrétisées par une formulation de type Galerkin discontinu essentiellement identique à celle adoptée dans cette étude si ce n'est qu'un schéma saute-mouton est utilisé pour l'intégration en temps des équations semi-discrétisées en espace.

Maillage	# sommets	# tétraèdres	L_{\min} (en m)	L_{\max} (en m)	L_{avg} (en m)
M1	129,276	725,424	0.020000	0.034641	0.025528

TAB. 4.2 – Equations de Maxwell 3D : diffraction d'une onde plane par un cube. Caractéristiques des maillages tétraédriques, $F = 900$ MHz.

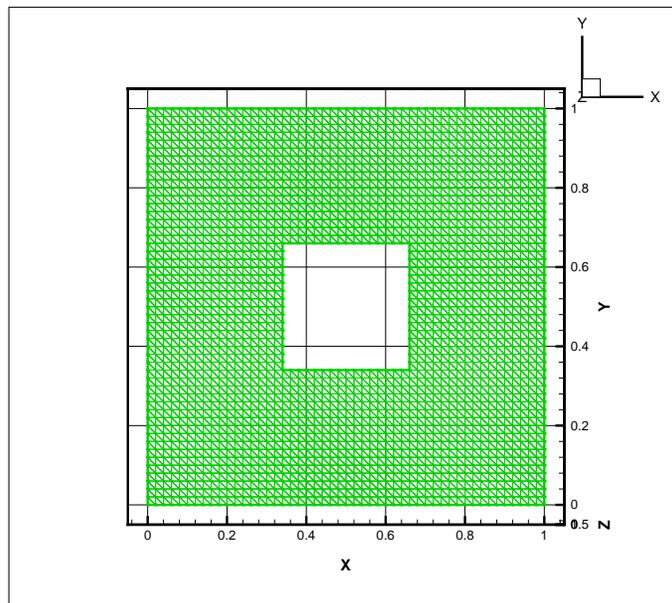


FIG. 1.27 – Equations de Maxwell 3D : diffraction d'une onde plane par un cube. Maillage dans le plan $Z = 0.5$.

4.1.3.1 Propagation dans le vide

Dans un premier temps, on suppose que le milieu environnant le cube interne est le vide. Les solutions obtenues sont montrées sous la forme des lignes de contour de la partie réelle de E_x et de E_y respectivement sur les figures 1.28 et 1.31 pour ce qui concerne la méthode GDDT-P1, les figures 1.29 et 1.32 pour ce qui concerne la méthode GD-P0 et enfin, les figures 1.30 et 1.33 pour ce qui concerne la méthode GD-P1. Pour permettre une comparaison plus fine des différentes solutions, on a représenté sur les figures 1.34 à 1.36 des coupes 1D de la partie réelle des composantes E_x , E_y et E_z (évolution de E_x , E_y et E_z en fonction de X pour $Y = Z = 0.7$). On constate un bon accord entre les solutions fournies par les méthodes GDDT-P1 et GD-P1, la méthode GD-P0 nécessitant probablement l'utilisation d'un pas de discrétisation plus fin.

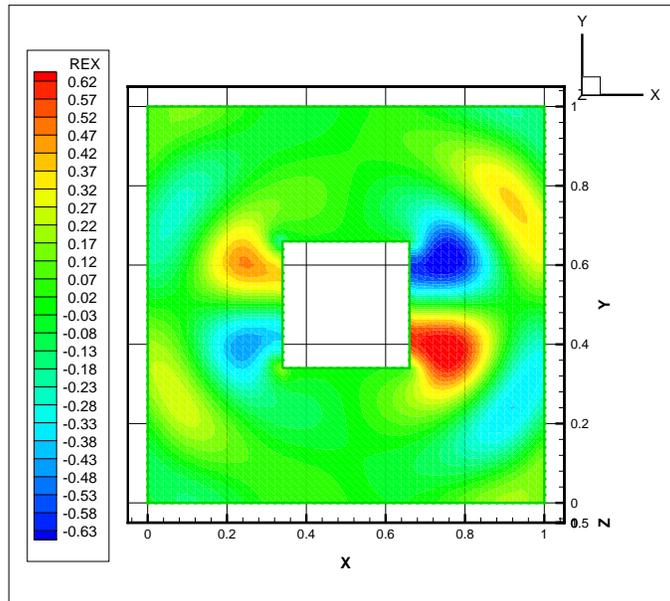


FIG. 1.28 – Equations de Maxwell 3D : diffraction d'une onde plane par un cube.
Propagation dans le vide, $F = 900$ MHz, méthode GDDT-P1.
Partie réelle de E_x : lignes de contour dans le plan $Z = 0.5$.

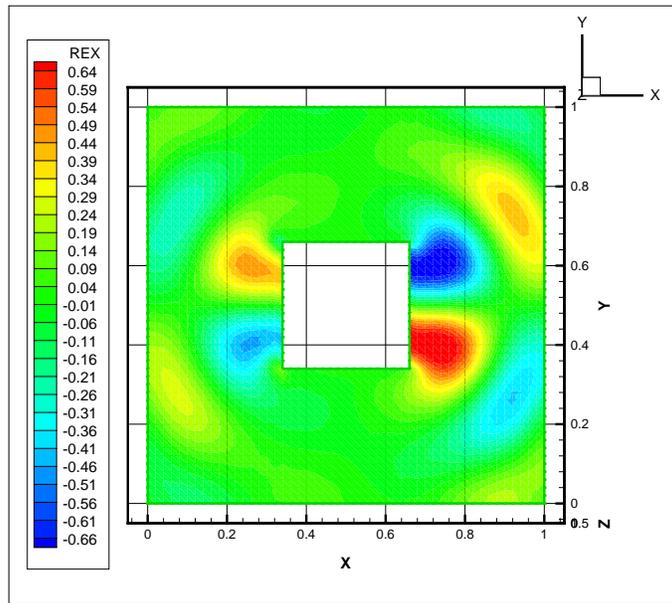


FIG. 1.29 – Equations de Maxwell 3D : diffraction d’une onde plane par un cube.
 Propagation dans le vide, $F = 900$ MHz, méthode GD-P0.
 Partie réelle de E_x : lignes de contour dans le plan $Z = 0.5$.

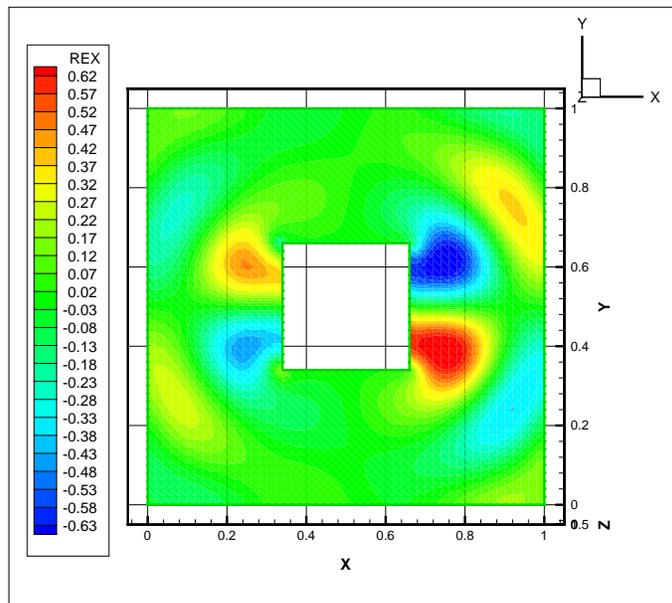


FIG. 1.30 – Equations de Maxwell 3D : diffraction d’une onde plane par un cube.
 Propagation dans le vide, $F = 900$ MHz, méthode GD-P1.
 Partie réelle de E_x : lignes de contour dans le plan $Z = 0.5$.

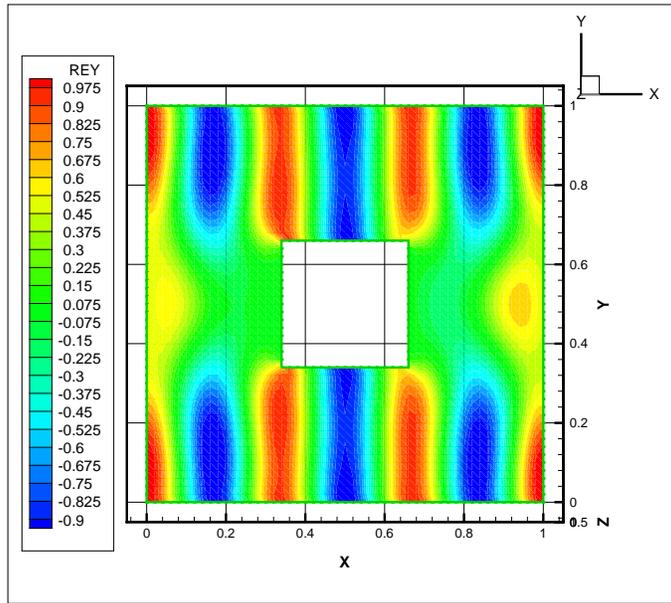


FIG. 1.31 – Equations de Maxwell 3D : diffraction d’une onde plane par un cube.
 Propagation dans le vide, $F = 900$ MHz, méthode GDDT-P1.
 Partie réelle de E_y : lignes de contour dans le plan $Z = 0.5$.

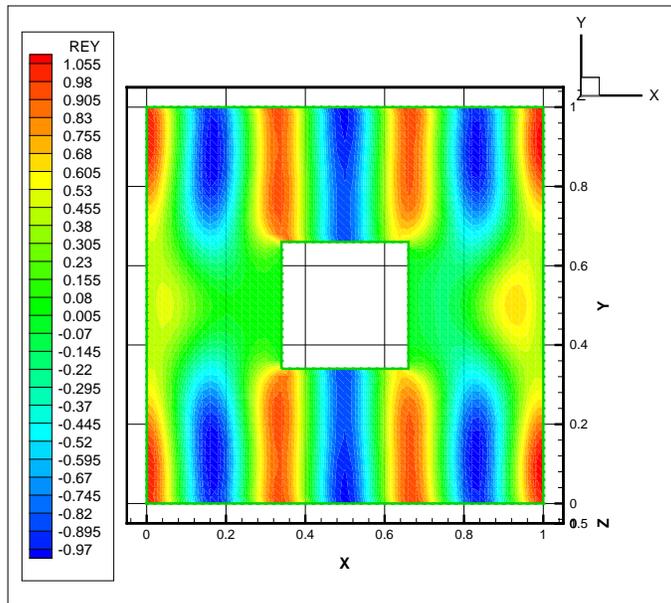


FIG. 1.32 – Equations de Maxwell 3D : diffraction d’une onde plane par un cube.
 Propagation dans le vide, $F = 900$ MHz, méthode GD-P0.
 Partie réelle de E_y : lignes de contour dans le plan $Z = 0.5$.

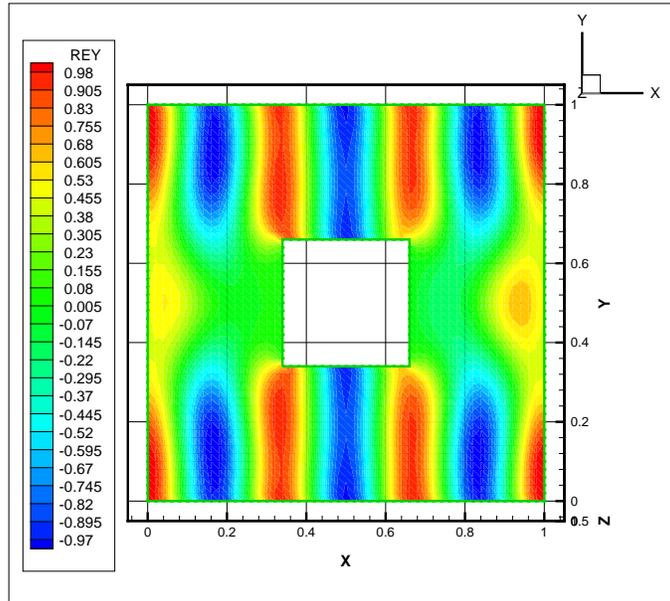


FIG. 1.33 – Equations de Maxwell 3D : diffraction d’une onde plane par un cube.
 Propagation dans le vide, $F = 900$ MHz, méthode GD-P1.
 Partie réelle de E_y : lignes de contour dans le plan $Z = 0.5$.

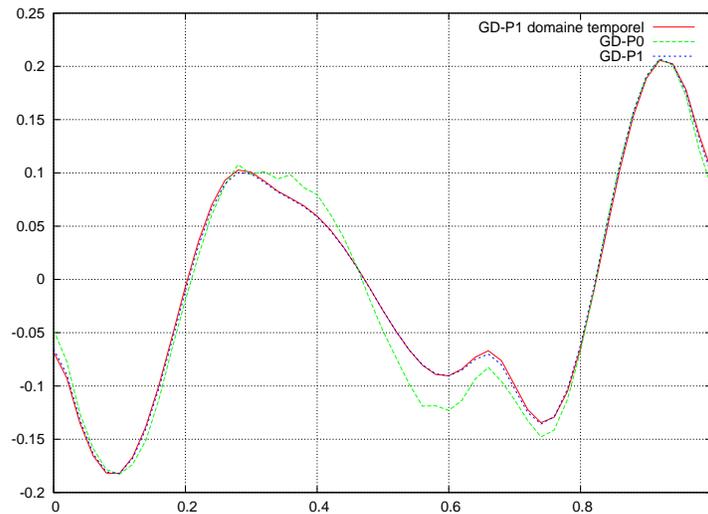


FIG. 1.34 – Equations de Maxwell 3D : diffraction d’une onde plane par un cube.
 Propagation dans le vide, $F = 900$ MHz.
 Partie réelle de E_x : coupe 1D en $Y = Z = 0.7$.

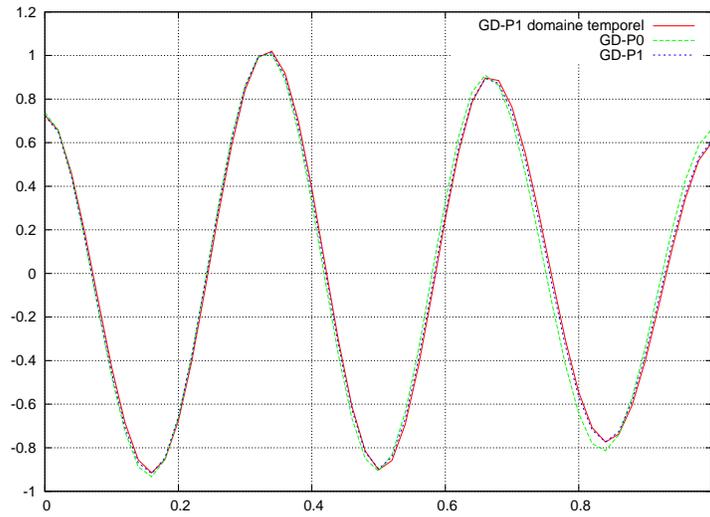


FIG. 1.35 – Equations de Maxwell 3D : diffraction d’une onde plane par un cube.
 Propagation dans le vide, $F = 900$ MHz.
 Partie réelle de E_y : coupe 1D en $Y = Z = 0.7$.

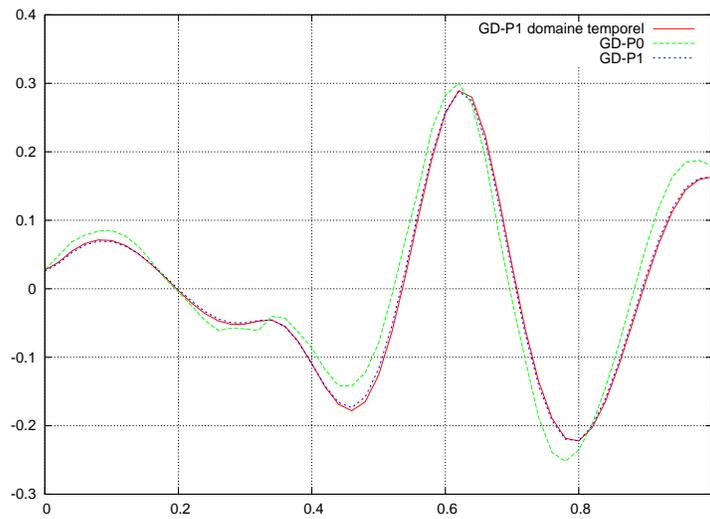


FIG. 1.36 – Equations de Maxwell 3D : diffraction d’une onde plane par un cube.
 Propagation dans le vide, $F = 900$ MHz.
 Partie réelle de E_z : coupe 1D en $Y = Z = 0.7$.

4.1.3.2 Cas d'un matériau diélectrique non-conducteur

On suppose maintenant que le cube interne est entouré d'une épaisseur d'un matériau de permittivité relative $\epsilon_r = 4$ (voir la figure 1.37). Dans cette zone, la longueur d'onde correspondant à la fréquence $F = 900$ MHz et à cette valeur de la permittivité est $\lambda = 1/6$ m. Comme dans le cas précédent, les solutions obtenues sont montrées sous la forme des lignes de contour de la partie réelle de E_x et E_y respectivement sur les figures 1.38 et 1.41 pour ce qui concerne la méthode GDDT-P1, les figures 1.39 et 1.42 pour ce qui concerne la méthode GD-P0 et enfin, les figures 1.40 et 1.43 pour ce qui concerne la méthode GD-P1. Ces figures sont complétées par les coupes 1D en $Y = Z = 0.7$ des composantes E_x , E_y et E_z représentées sur les figures 1.44 à 1.46. On constate que la solution fournie par la méthode GD-P0 est nettement différente des solutions résultant des méthodes GDDT-P1 et GD-P1. Ce résultat n'est pas surprenant compte tenu du fait que dans la zone du matériau le pas d'espace utilisé correspond à environ 5 points par longueur d'onde (au lieu de 10 dans le vide) ce qui est clairement insuffisant.

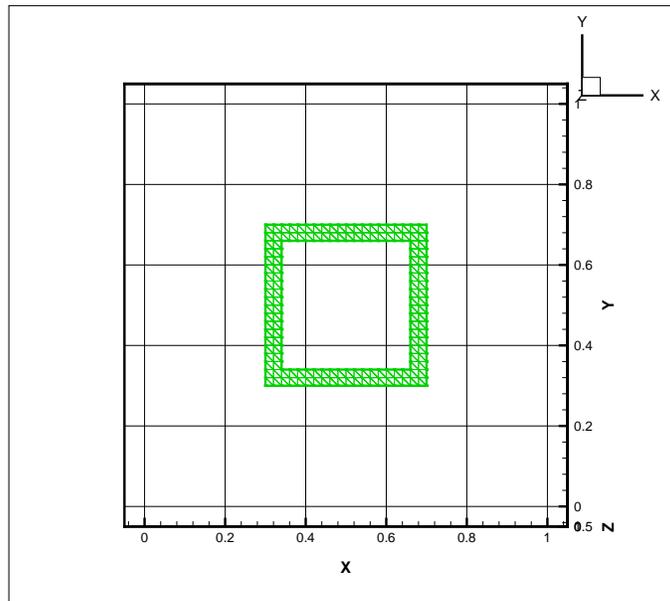


FIG. 1.37 – Equations de Maxwell 3D : diffraction d'une onde plane par un cube.
Maillage dans le plan $Z = 0.5$ restreint à la zone de matériau.

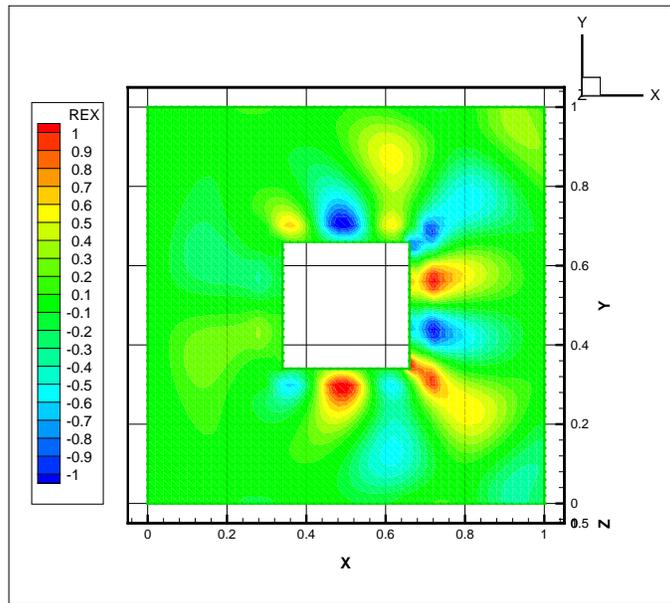


FIG. 1.38 – Equations de Maxwell 3D : diffraction d’une onde plane par un cube.
 Matériau diélectrique non-conducteur, $F = 900$ MHz, méthode GDDT-P1.
 Partie réelle de E_x : lignes de contour dans le plan $Z = 0.5$.

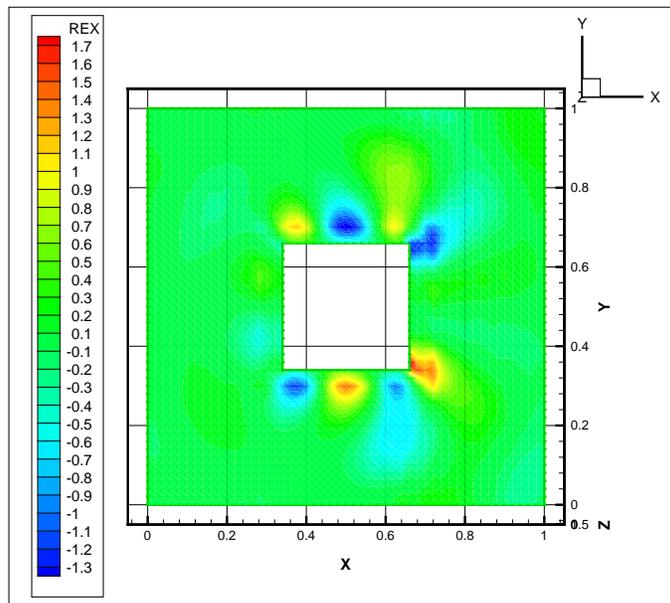


FIG. 1.39 – Equations de Maxwell 3D : diffraction d’une onde plane par un cube.
 Matériau diélectrique non-conducteur, $F = 900$ MHz, méthode GD-P0.
 Partie réelle de E_x : lignes de contour dans le plan $Z = 0.5$.

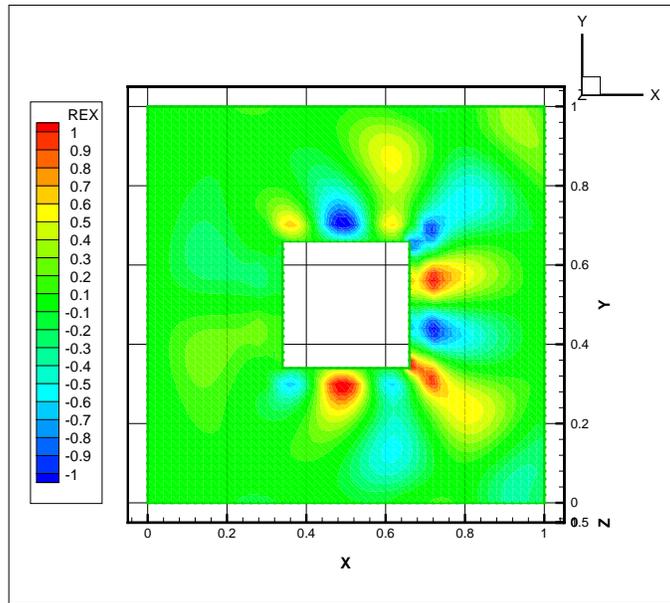


FIG. 1.40 – Equations de Maxwell 3D : diffraction d’une onde plane par un cube.
 Matériau diélectrique non-conducteur, $F = 900$ MHz, méthode GD-P1.
 Partie réelle de E_x : lignes de contour dans le plan $Z = 0.5$.

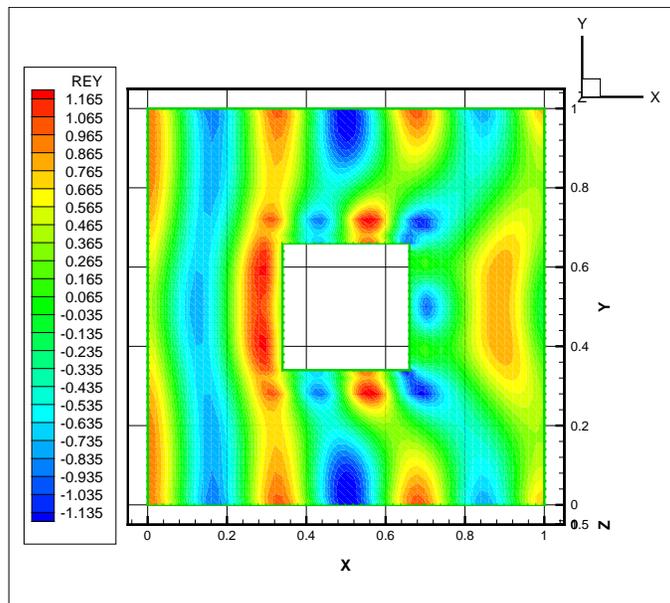


FIG. 1.41 – Equations de Maxwell 3D : diffraction d’une onde plane par un cube.
 Matériau diélectrique non-conducteur, $F = 900$ MHz, méthode GDDT-P1.
 Partie réelle de E_y : lignes de contour dans le plan $Z = 0.5$.

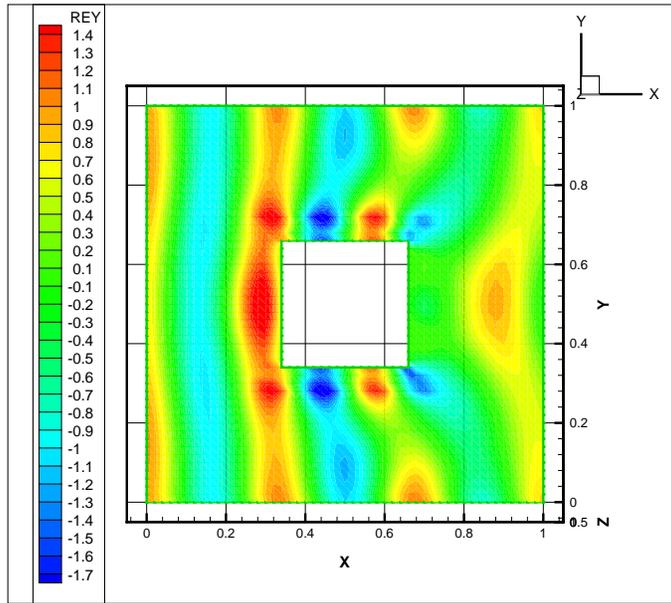


FIG. 1.42 – Equations de Maxwell 3D : diffraction d’une onde plane par un cube.
 Matériau diélectrique non-conducteur, $F = 900$ MHz, méthode GD-P0.
 Partie réelle de E_y : lignes de contour dans le plan $Z = 0.5$.

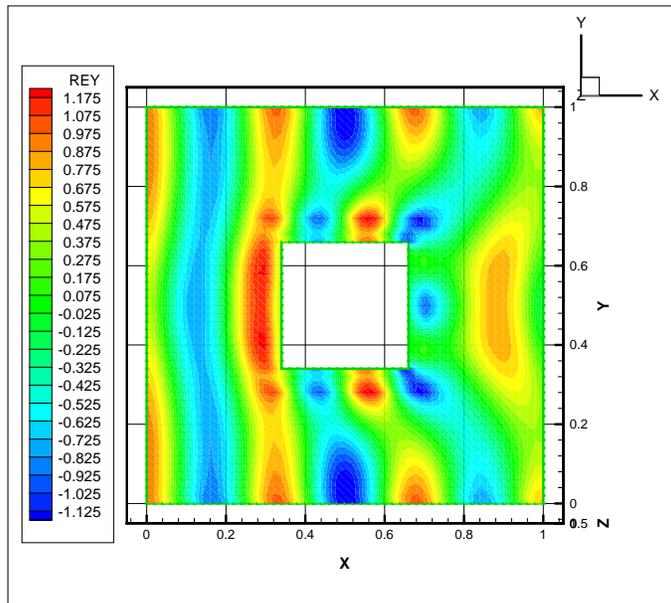


FIG. 1.43 – Equations de Maxwell 3D : diffraction d’une onde plane par un cube.
 Matériau diélectrique non-conducteur, $F = 900$ MHz, méthode GD-P1.
 Partie réelle de E_y : lignes de contour dans le plan $Z = 0.5$.

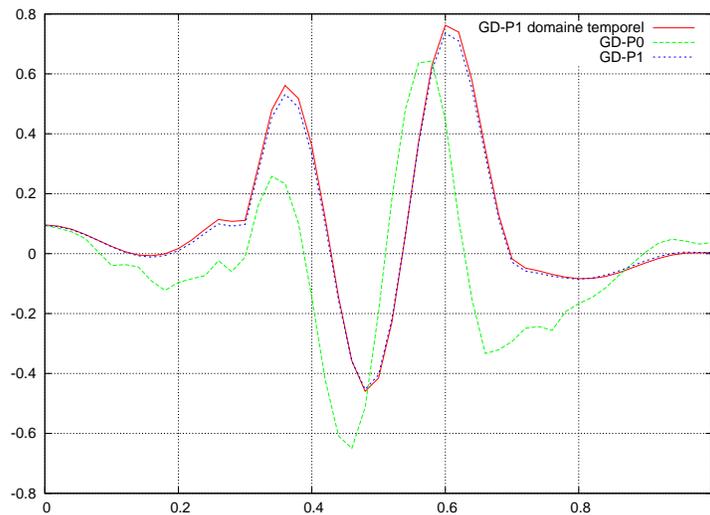


FIG. 1.44 – Equations de Maxwell 3D : diffraction d’une onde plane par un cube.
 Matériau diélectrique non-conducteur, $F = 900$ MHz.
 Partie réelle de E_x : coupe 1D en $Y = Z = 0.7$.

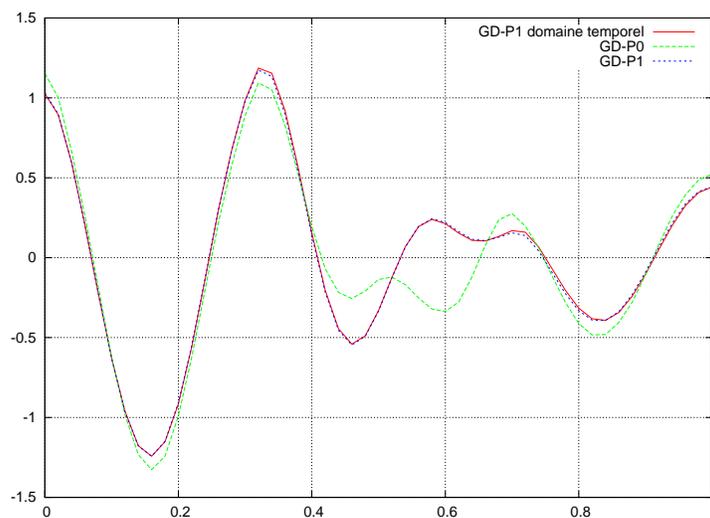


FIG. 1.45 – Equations de Maxwell 3D : diffraction d’une onde plane par un cube.
 Matériau diélectrique non-conducteur, $F = 900$ MHz.
 Partie réelle de E_y : coupe 1D en $Y = Z = 0.7$.

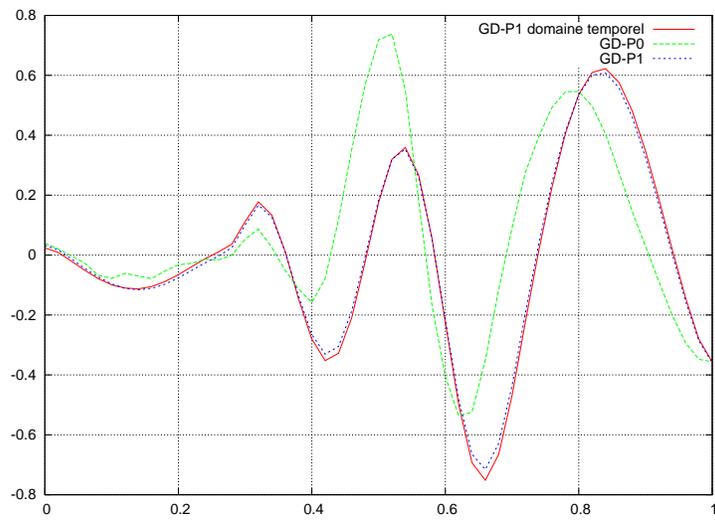


FIG. 1.46 – Equations de Maxwell 3D : diffraction d’une onde plane par un cube.
 Matériau diélectrique non-conducteur, $F = 900$ MHz, méthode GD-P1.
 Partie réelle de E_z : coupe 1D en $Y = Z = 0.7$.

4.1.3.3 Cas d'un matériau diélectrique conducteur

Le matériau est maintenant supposé conducteur de conductivité $\sigma = 1 \text{ S.m}^{-1}$ alors que la permittivité relative est toujours égale à $\varepsilon_r = 4$. Les solutions obtenues sont montrées sous la forme des lignes de contour de la partie réelle de E_x et E_y respectivement sur les figures 1.47 et 1.50 pour ce qui concerne la méthode GDDT-P1, les figures 1.48 et 1.51 pour ce qui concerne la méthode GD-P0 et enfin, les figures 1.49 et 1.52 pour ce qui concerne la méthode GD-P1. Ces figures sont complétées par les coupes 1D en $Y = Z = 0.7$ des composantes E_x , E_y et E_z représentées sur les figures 1.53 à 1.55. Les remarques faites sur les résultats obtenus dans le cas d'un matériau non-conducteur s'appliquent encore ici.

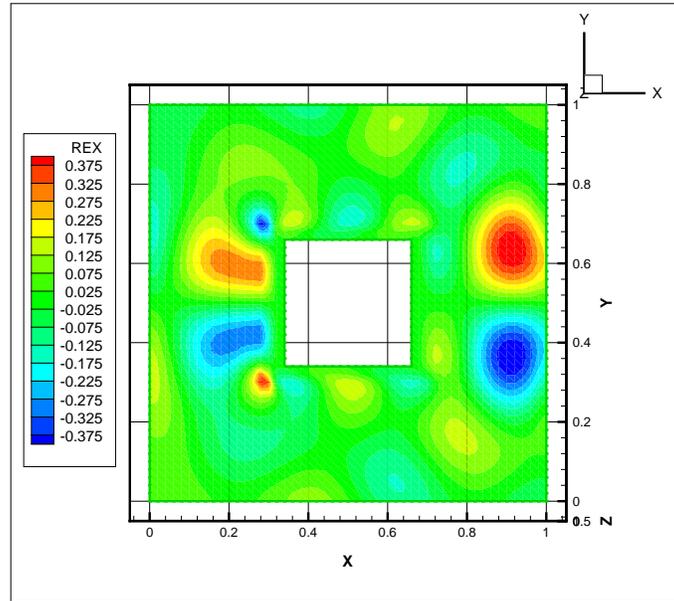


FIG. 1.47 – Equations de Maxwell 3D : diffraction d'une onde plane par un cube.
Matériau diélectrique conducteur, $F = 900 \text{ MHz}$, méthode GDDT-P1.
Partie réelle de E_x : lignes de contour dans le plan $Z = 0.5$.

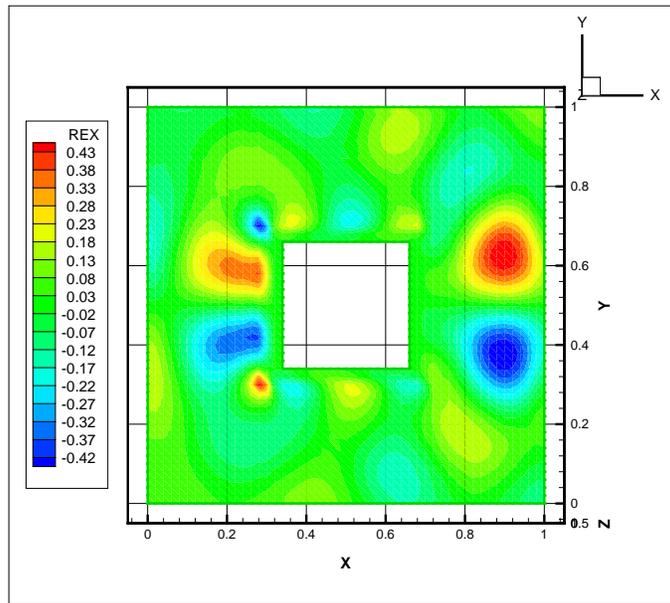


FIG. 1.48 – Equations de Maxwell 3D : diffraction d’une onde plane par un cube.
 Matériau diélectrique conducteur, $F = 900$ MHz, méthode GD-P0.
 Partie réelle de E_x : lignes de contour dans le plan $Z = 0.5$.

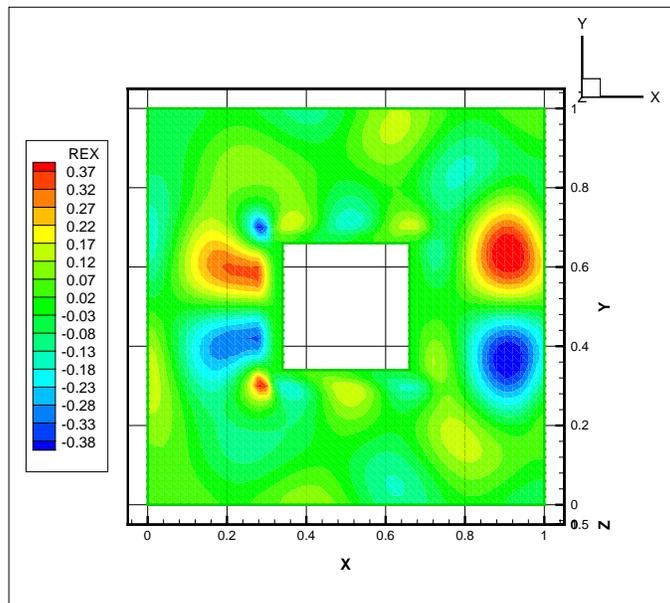


FIG. 1.49 – Equations de Maxwell 3D : diffraction d’une onde plane par un cube.
 Matériau diélectrique conducteur, $F = 900$ MHz, méthode GD-P1.
 Partie réelle de E_x : lignes de contour dans le plan $Z = 0.5$.

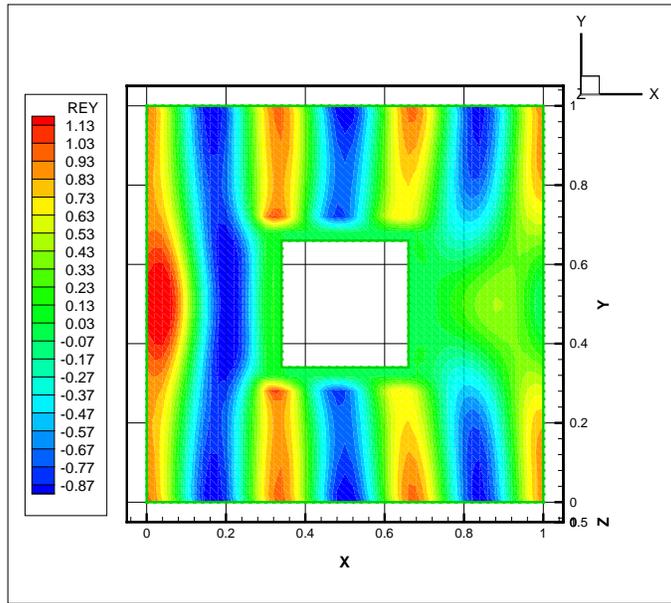


FIG. 1.50 – Equations de Maxwell 3D : diffraction d’une onde plane par un cube.
 Matériau diélectrique conducteur, $F = 900$ MHz, méthode GDDT-P1.
 Partie réelle de E_y : lignes de contour dans le plan $Z = 0.5$.

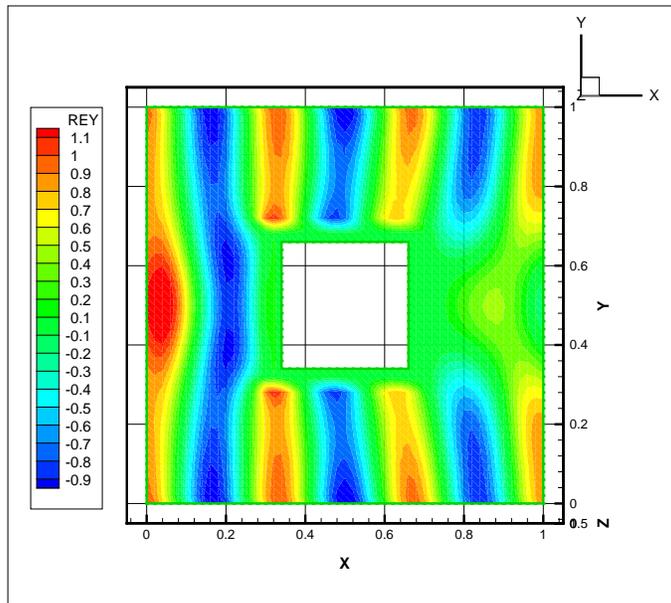


FIG. 1.51 – Equations de Maxwell 3D : diffraction d’une onde plane par un cube.
 Matériau diélectrique conducteur, $F = 900$ MHz, méthode GD-P0.
 Partie réelle de E_y : lignes de contour dans le plan $Z = 0.5$.

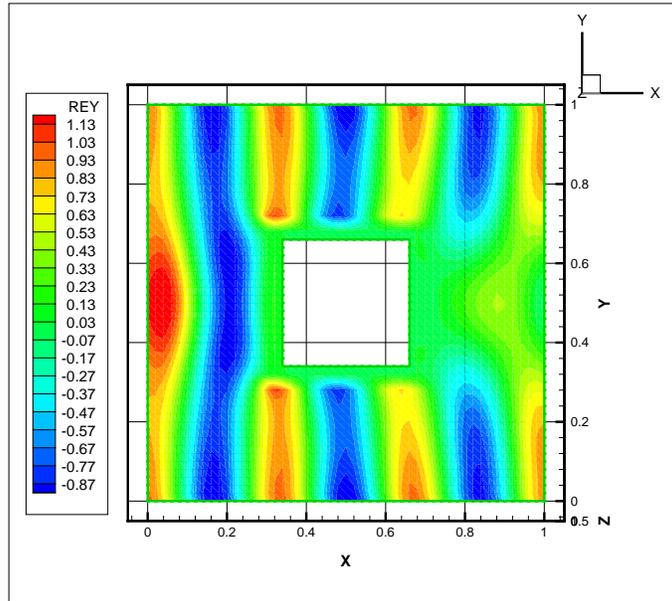


FIG. 1.52 – Equations de Maxwell 3D : diffraction d’une onde plane par un cube.
 Matériau diélectrique conducteur, $F = 900$ MHz, méthode GD-P1.
 Partie réelle de E_y : lignes de contour dans le plan $Z = 0.5$.

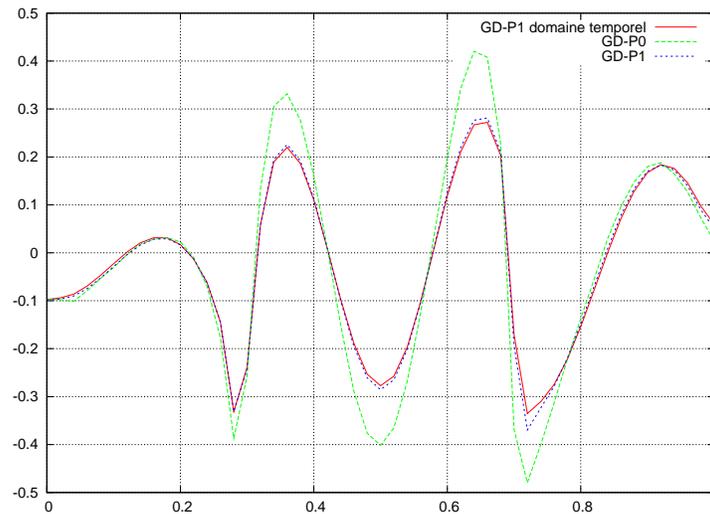


FIG. 1.53 – Equations de Maxwell 3D : diffraction d’une onde plane par un cube.
 Matériau diélectrique conducteur, $F = 900$ MHz.
 Partie réelle de E_x : coupe 1D en $Y = Z = 0.7$.

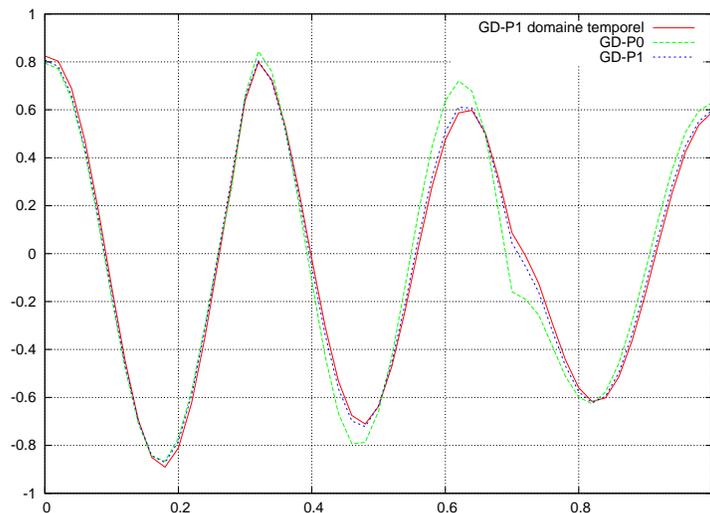


FIG. 1.54 – Equations de Maxwell 3D : diffraction d’une onde plane par un cube.
 Matériau diélectrique conducteur, $F = 900$ MHz.
 Partie réelle de E_y : coupe 1D en $Y = Z = 0.7$.

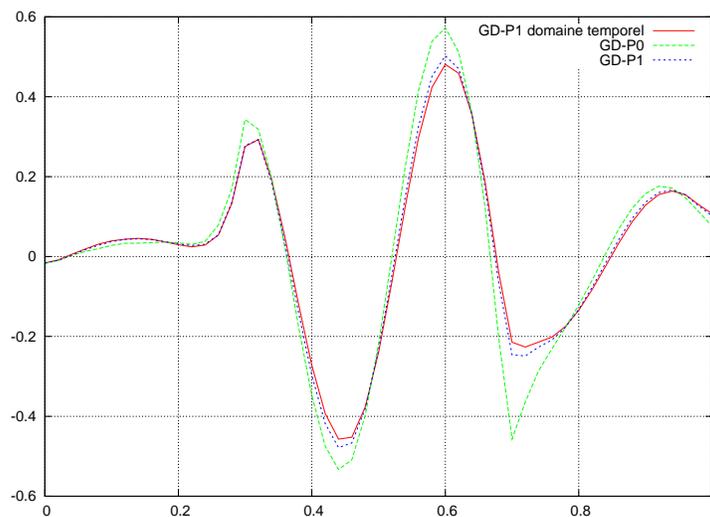


FIG. 1.55 – Equations de Maxwell 3D : diffraction d’une onde plane par un cube.
 Matériau diélectrique conducteur, $F = 900$ MHz, méthode GD-P1.
 Partie réelle de E_z : coupe 1D en $Y = Z = 0.7$.

4.2 Résolution numérique en mode séquentiel

Cette section est consacrée à la résolution des systèmes linéaires issus des formulations de type Galerkin discontinu GD-P0 et GD-P1 en maillages tétraédriques étudiées au chapitre 2. Les matrices de ces systèmes linéaires sont creuses, à structure irrégulière, non-symétriques et non-définies. De plus, l'utilisation d'une fonction de flux numériques centrés aux interfaces entre éléments voisins ne favorise pas la propriété de dominance diagonale des matrices en question. On peut donc s'attendre à devoir traiter des matrices mal conditionnées. Deux options principales sont classiquement envisageables dans ce contexte : une méthode de résolution directe basée sur une factorisation LU ou une méthode itérative préconditionnée. Nous considérons ici ces deux options sous la forme des stratégies de résolution suivantes :

- une méthode de résolution directe LU adaptée aux matrices creuses et aux plateformes de calcul parallèles. Cette dernière caractéristique sera évaluée dans la section qui suit. Il s'agit d'une méthode multifrontale et plus précisément du solveur MUMPS¹[Amestoy et al., 2000], développé par des chercheurs du CERFACS, de l'ENSEEIH-IRIT et de l'INRIA Rhône-Alpes. Cette méthode se décompose en deux phases : une phase d'analyse et la factorisation numérique à proprement dite. La phase d'analyse inclue une étape de renumérotation qui a pour but de minimiser le remplissage des facteurs durant la factorisation numérique, et une étape d'analyse symbolique qui implique la construction d'un arbre d'élimination. Cet arbre constitue le graphe de dépendance des calculs et est exploité pour définir les tâches qui peuvent être exécutées en parallèle. D'autre part, la méthode multifrontale permet de se ramener pour chaque nœud de l'arbre à des factorisations partielles sur des matrices denses. Cette approche permet une bonne localisation des données et une bonne utilisation des caches.
- une méthode itérative de type Krylov. Il existe plusieurs méthodes de ce type adaptées aux matrices non-symétriques [Saad, 1996]. Dans cette étude, nous avons sélectionné la méthode BiCGstab(ℓ)[Sleijpen and Fokkema, 1993] applicable aux matrices à spectre complexe. Cette méthode combine les avantages des méthodes BiCG (bi-gradient conjugué) et GMRES(ℓ)[Saad and Schultz, 1986]. Dans le cas $\ell = 1$, cette méthode est équivalent à la méthode BiCGstab[van der Vorst, 1992]. L'occupation mémoire interne requise par la méthode BiCGstab(ℓ) est de $2\ell + 5$ vecteurs de dimension le nombre d'inconnues du problème contre $\ell + 3$ pour la méthode GMRES(ℓ). La méthode BiCGstab(ℓ) s'emploie donc en général avec $\ell \leq 10$. Nous reviendrons plus loin sur la question du préconditionnement de cette méthode.

Le cas test considéré pour l'évaluation de ces deux stratégies est celui de la propagation d'une onde plane dans le vide (cf. sous-section 4.1.1). Les calculs sont réalisés sur un ordinateur équipé d'un noeu biprocesseur AMD Opteron/2 GHz et de 2 GB de mémoire RAM. Sauf mention contraire, tous les calculs sont réalisés en double précision (64 bits).

4.2.1 Résolution itérative non-préconditionnée

Dans un premier temps, nous évaluons l'influence du paramètre ℓ de la méthode BiCGstab(ℓ) utilisée sans préconditionneur. La fréquence de l'onde incidente est tout d'abord fixée à 300 MHz et le maillage utilisé est caractérisé par $N_x = 21$ et $N_y = N_z = 11$. Les résultats concernant l'évolution de la norme euclidienne du résidu sont montrés sur les figures 2.56 (méthode d'approximation GD-P0) et 2.57 (méthode de d'approximation GD-P1). Les temps de calcul pour une réduction de la norme du résidu

¹MULTifrontal Massively Parallel sparse direct Solver

initial d'un facteur 10^6 sont résumés dans la table 4.3. Comme on pouvait s'y attendre, on constate qu'augmenter la valeur du paramètre ℓ permet d'améliorer la vitesse de convergence de la méthode de résolution $\text{BiCGstab}(\ell)$. Néanmoins, on constate aussi que la réduction du nombre d'itérations pour converger bien que notable ne se traduit pas forcément par une diminution du temps de calcul. Ceci est effectivement le cas pour des systèmes de taille suffisamment importante, augmentation de taille traduite ici par le passage de la méthode d'approximation GD-P0 à la méthode GD-P1. Pour le maillage considéré ici (12,000 éléments), les nombres d'inconnues sont respectivement égaux à 72,000 (méthode GD-P0, 6 degrés de liberté par élément) et 288,000 (méthode GD-P1, 24 degrés de liberté par élément).

Le paramètre ℓ est maintenant fixé à 6 et on évalue l'influence de la fréquence de l'onde incidente sur la convergence de la méthode de résolution $\text{BiCGstab}(6)$. On se limite à l'application de la méthode d'approximation GD-P0 sur un maillage tel que $N_x = 41$ et $N_y = N_z = 11$. Les résultats concernant l'évolution de la norme euclidienne du résidu sont montrés sur la figure 2.58 tandis que les temps de calcul sont résumés dans la table 4.4. L'augmentation de la fréquence se traduit par une amélioration significative de la vitesse de convergence et une diminution du temps de calcul. Ceci s'explique simplement par le fait que l'augmentation de la fréquence (et donc de la pulsation ω) se traduit par une augmentation de la dominance diagonale (voir les expressions (3.79) de la sous-section 2.3.1.2).

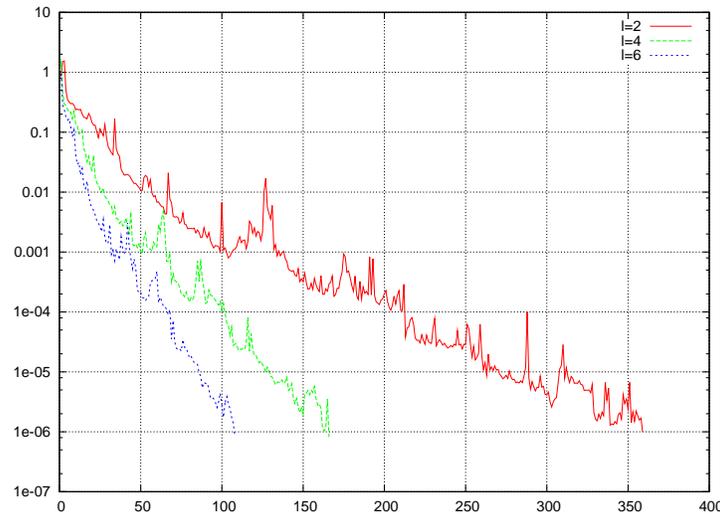


FIG. 2.56 – Equations de Maxwell 3D : propagation d'une onde plane dans le vide.

$N_x = 21$, $N_y = N_z = 11$, $F = 300$ MHz, méthode d'approximation GD-P0.

Influence du paramètre ℓ de la méthode de résolution $\text{BiCGstab}(\ell)$ non-préconditionnée.

Abscisse : nombre d'itérations - Ordonnée : résidu normalisé (échelle logarithmique).

4.2.2 Résolution itérative préconditionnée

Nous étudions maintenant l'apport d'une technique de préconditionnement sur la convergence de la méthode $\text{BiCGstab}(\ell)$. Notre choix s'est porté sur la méthode $\text{ILUT}(\tau, p)$ [Saad, 1994]. Dans cette méthode, le remplissage des facteurs L et U est guidé par deux paramètres : p est un entier qui désigne le nombre maximal de termes dans une ligne de L ou U ; τ est un paramètre réel qui définit le seuil d'élimination des termes dans les facteurs L et U (voir [Saad, 1994] pour plus de détails). Généralement,

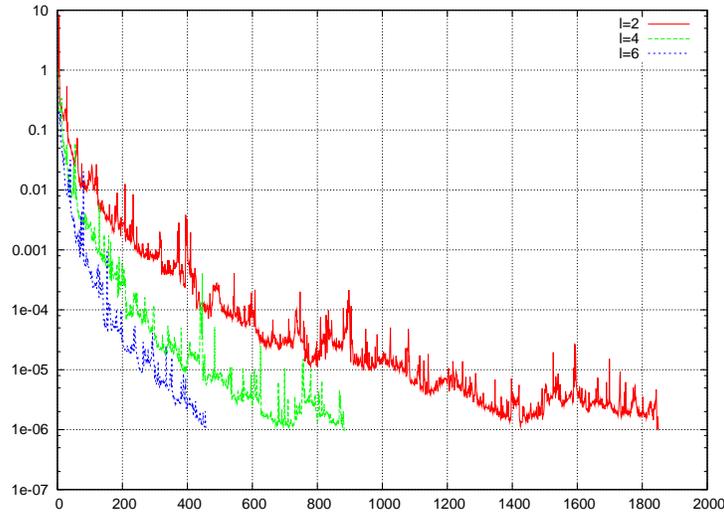


FIG. 2.57 – Equations de Maxwell 3D : propagation d’une onde plane dans le vide.
 $N_x = 21$, $N_y = N_z = 11$, $F = 300$ MHz, méthode d’approximation GD-P1.
 Influence du paramètre ℓ de la méthode de résolution BiCGstab(ℓ) non-préconditionnée.
 Abscisse : nombre d’itérations - Ordonnée : résidu normalisé (échelle logarithmique).

Méthode d’approximation	ℓ	# iter	CPU
GD-P0	2	359	57.0 sec
-	4	166	58.0 sec
-	6	108	55.0 sec
GD-P1	2	1850	1770.0 sec
-	4	883	1728.0 sec
-	6	457	1308.0 sec

TAB. 4.3 – Equations de Maxwell 3D : propagation d’une onde plane dans le vide.
 $N_x = 21$, $N_y = N_z = 11$, $F = 300$ MHz.
 Influence du paramètre ℓ de la méthode BiCGstab(ℓ) non-préconditionnée.
 Nombre d’itérations et temps de calcul.

Fréquence de l’onde incidente	# iter	CPU
300 MHz	494	527.0 sec
600 MHz	385	397.0 sec
900 MHz	265	305.0 sec

TAB. 4.4 – Equations de Maxwell 3D : propagation d’une onde plane dans le vide.
 $N_x = 41$, $N_y = N_z = 11$, méthode d’approximation GD-P0.
 Méthode de résolution BiCGstab(6) non-préconditionnée.
 Influence de la fréquence : nombre d’itérations et temps de calcul.

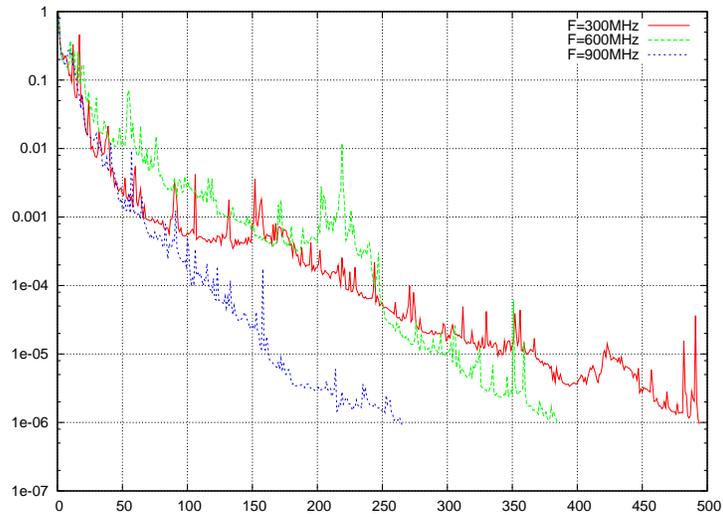


FIG. 2.58 – Equations de Maxwell 3D : propagation d’une onde plane dans le vide.
 $N_x = 41$, $N_y = N_z = 11$, méthode d’approximation GD-P0.
Méthode de résolution BiCGstab(6) non-préconditionnée.
Influence de la fréquence F .

Abscisse : nombre d’itérations - Ordonnée : résidu normalisé (échelle logarithmique).

pour des matrices qui se prêtent bien à ce type de factorisation, p est compris entre 10 et 50 et τ entre 10^{-3} et 10^{-2} . Nous avons tout d’abord constaté que le préconditionneur $ILUT(\tau, p)$ n’est utilisable dans le cas présent (i.e. conduit à une factorisation LU qui est une bonne approximation de l’inverse exact de la matrice du système linéaire à résoudre) que pour des grandes valeurs du paramètre p . Ainsi, pour un maillage tel que $N_x = 21$ et $N_y = N_z = 11$, une valeur de p inférieure à 600 conduit à une divergence de la méthode BiCGstab(6). Sur la figure 2.59 nous avons représenté les courbes de convergence obtenues lorsque $\tau = 10^{-3}$ et pour différentes valeurs du paramètre p . La valeur $p = 750$ conduit à une convergence en 3 itérations et le choix $p = 1000$ n’apporte aucun bénéfice comme en témoigne les données de la table 4.5. Dans les tables de cette sous-section, $nz_{|L|U}$ désigne le nombre de termes non nuls total des facteurs L et U. Le maillage utilisé comporte 12,000 tétraèdres et le nombre d’inconnues du système linéaire est donc égal à 72,000 dans le cas de la méthode d’approximation GD-P0. La matrice de ce système comporte 88,808 termes non nuls. L’influence du paramètre τ à p fixé est quant à elle moins notable (voir la figure 2.60 et la table 4.6). Compte tenu des valeurs du paramètre p requises par ces expériences numériques et des taux de remplissage des facteurs L et U résultants, nous pouvons faire les deux remarques suivantes :

- il ne paraît pas raisonnable d’envisager l’utilisation de ce type de préconditionneur pour la matrice du système linéaire associé à la formulation GD-P1,
- il semble opportun d’évaluer l’applicabilité d’une méthode de résolution directe comme une alternative robuste aux stratégies étudiées jusqu’ici. C’est ce que nous réalisons dans la sous-section suivante.

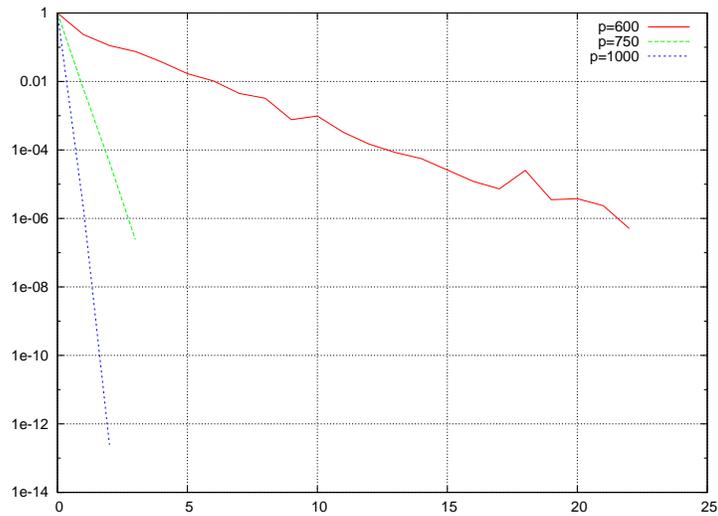


FIG. 2.59 – Equations de Maxwell 3D : propagation d’une onde plane dans le vide.
 $N_x = 21$, $N_y = N_z = 11$, $F = 300$ MHz, méthode d’approximation GD-P0.
 Méthode de résolution BiCGstab(6) préconditionnée.
 Influence du paramètre p de $\text{ILUT}(\tau, p)$ avec $\tau = 10^{-3}$.
 Abscisse : nombre d’itérations - Ordonnée : résidu normalisé (échelle logarithmique).

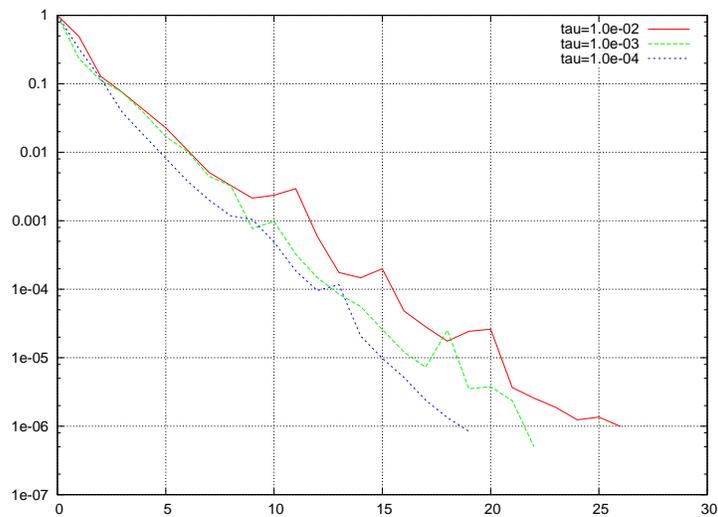


FIG. 2.60 – Equations de Maxwell 3D : propagation d’une onde plane dans le vide.
 $N_x = 21$, $N_y = N_z = 11$, $F = 300$ MHz, méthode d’approximation GD-P0.
 Méthode de résolution BiCGstab(6) préconditionnée.
 Influence du paramètre τ de $\text{ILUT}(\tau, p)$ avec $p = 600$.
 Abscisse : nombre d’itérations - Ordonnée : résidu normalisé (échelle logarithmique).

p	# iter	CPU factorisation	CPU résolution	nz_{ilu}
600	22	163.0 sec	84.0 sec	15,183,142
750	3	184.0 sec	13.0 sec	17,443,632
1000	2	210.0 sec	10.0 sec	20,061,178

TAB. 4.5 – Equations de Maxwell 3D : propagation d’une onde plane dans le vide.
 $N_x = 21$, $N_y = N_z = 11$, $F = 300$ MHz, méthode d’approximation GD-P0.
Méthode de résolution BiCGstab(6) préconditionnée.
Influence du paramètre p de ILUT(τ, p) avec $\tau = 10^{-3}$.
Nombre d’itérations et temps de calcul.

τ	# iter	CPU factorisation	CPU résolution	nz_{ilu}
10^{-2}	26	138.0 sec	99.0 sec	14,832,752
10^{-3}	22	163.0 sec	84.0 sec	15,183,142
10^{-4}	19	160.0 sec	73.0 sec	15,203,740

TAB. 4.6 – Equations de Maxwell 3D : propagation d’une onde plane dans le vide.
 $N_x = 21$, $N_y = N_z = 11$, $F = 300$ MHz, méthode d’approximation GD-P0.
Méthode de résolution BiCGstab(6) préconditionnée.
Influence du paramètre τ de ILUT(τ, p) avec $p = 600$.
Nombre d’itérations et temps de calcul.

4.2.3 Résolution directe

Le solveur MUMPS est utilisé pour résoudre le système linéaire résultant de la méthode d’approximation GD-P0 pour une fréquence $F = 300$ MHz. La programmation du solveur MUMPS s’appuie sur les procédures BLAS pour lesquelles il existe différentes implémentations optimisées pour un type de processeur donné. Nous utilisons la librairie GotoBLAS[Goto and van de Geijn, 2006] spécifique au processeur AMD Opteron. Dans ce cas de stratégie de résolution, il est intéressant de se pencher sur la question de l’occupation mémoire requise pour la construction et le stockage des facteurs L et U. Dans les tables de cette sous-section, N_{dl} désigne le nombre de degrés de liberté du problème calculé comme $6 \times \#$ éléments du maillage. La table 4.7 résume les résultats obtenus pour différents maillages. Ces résultats mettent en évidence le principal inconvénient de ce type de solveur à savoir, l’occupation mémoire. On constate aussi que les temps de résolution sont quasiment négligeables devant les temps de factorisation. Une méthode de résolution directe s’avèrera donc profitable si on peut amortir le coût de factorisation dans un contexte de résolution multiple à factorisation fixée comme démontré dans la section 4.3.

Nous évaluons maintenant une stratégie alternative qui consiste à utiliser une factorisation MUMPS en simple précision (32 bits) comme préconditionneur de la méthode BiCGstab(ℓ) avec $\ell = 1$ dans le cadre d’un calcul global en double précision (64 bits). Bien que cette approche soit utilisée ici sans justification théorique, nous retiendrons qu’il existe actuellement une certaine dynamique sur une problématique très proche [Kurzak and Dongarra, 2006]-[Langou et al., 2006] motivée par le constat que les performances en simple précision des architectures des processeurs modernes sont bien supérieures à celles atteignables en double précision. Ainsi, dans [Kurzak and Dongarra, 2006]-[Langou et al., 2006], les auteurs présentent une approche relativement générale, basée sur une technique de raffinement itératif,

pour obtenir une précision 64 bits à partir d'un calcul initial en précision 32 bits dans le cadre de la résolution de systèmes linéaires à matrice dense. Dans notre cas, nous considérons que la factorisation LU simple précision est une approximation de la factorisation exacte et de ce fait conduit à une approximation de l'inverse de la matrice initiale. Les résultats obtenus sont rassemblés dans la table 4.8. On constate une diminution notable de l'occupation mémoire mais aussi du temps de factorisation au détriment d'une augmentation sensible du coût de résolution.

N_x	N_{dl}	Mémoire	CPU factorisation	CPU résolution
21	72,000	385 MB	13.0 sec	0.3 sec
31	108,000	626 MB	22.0 sec	0.5 sec
41	144,000	878 MB	33.0 sec	0.5 sec
51	180,000	1117 MB	45.0 sec	0.7 sec

TAB. 4.7 – Equations de Maxwell 3D : propagation d'une onde plane dans le vide.
 $N_y = N_z = 11$, méthode d'approximation d'approximation GD-P0.
Méthode de résolution directe à l'aide solveur MUMPS.
Occupation mémoire et temps de calcul.

N_x	N_{dl}	Mémoire	CPU factorisation	CPU résolution
21	72,000	204 MB	8.0 sec	1.0 sec
31	108,000	328 MB	14.0 sec	1.5 sec
41	144,000	462 MB	21.0 sec	2.0 sec
51	180,000	591 MB	28.0 sec	2.5 sec

TAB. 4.8 – Equations de Maxwell 3D : propagation d'une onde plane dans le vide.
 $N_y = N_z = 11$, méthode d'approximation GD-P0.
Méthode de résolution itérative BiCGstab(1) préconditionnée par MUMPS.
Occupation mémoire et temps de calcul.

4.3 Résolution numérique en mode parallèle

Dans cette section, nous évaluons différentes stratégies de résolution parallèle appliquée à certains des cas de calcul décrits dans la section 4.1. Nous considérons plus particulièrement les stratégies de résolution suivantes :

- la méthode BiCGstab(ℓ)[Sleijpen and Fokkema, 1993] utilisée comme solveur global, sans technique de préconditionnement et avec $\ell = 6$,
- la méthode BiCGstab(ℓ) utilisée comme solveur global avec $\ell = 1$, préconditionnée par une factorisation MUMPS[Amestoy et al., 2000] en simple précision,
- la méthode itérative (2.51) associée à l'algorithme de Schwarz (2.7) basé sur des conditions d'interface naturelles,
- la méthode BiCGstab(ℓ) appliquée au système interface (2.49) avec $\ell = 6$.

Pour ce qui concerne la résolution des systèmes locaux, deux options sont utilisées :

- la méthode BiCGstab(ℓ) avec $\ell = 6$, préconditionnée par une technique ILUT(τ, p)[Saad, 1994] avec $\tau = 10^{-3}$ et $p = 1000$,
- la méthode BiCGstab(ℓ) avec $\ell = 6$, préconditionnée par une factorisation MUMPS en simple précision.

Dans les tables et figures qui suivent nous adoptons les conventions de notation précisées dans la table 4.9. Par ailleurs, 'CPU' désigne le maximum du temps CPU relevé sur chaque processus alors que 'REEL' est le temps écoulé (identique pour tous les processus). La quantité '%CPU' est le rapport des quantités 'CPU' et 'REEL' et est utilisée ici comme une métrique d'évaluation de l'efficacité parallèle. Les expériences numériques sont réalisées sur un cluster de nœuds biprocesseurs AMD Opteron/2 GHz interconnectés par un commutateur Gigabit Ethernet.

Nom	Méthode de résolution
GLOB	Méthode BiCGstab(6) (solveur global)
GLOBD	BiCGstab(1)+MUMPS (solveur global)
DDMINT1	Résolution itérative du système interface et solveur local basé sur BiCGstab(6)+ILUT($10^{-3}, 1000$)
DDMINT2	Méthode BiCGstab(6) appliquée au système interface et solveur local basé sur BiCGstab(6)+ILUT($10^{-3}, 1000$)
DDMINT3	Méthode BiCGstab(6) appliquée au système interface et solveur local basé sur BiCGstab(1)+MUMPS

TAB. 4.9 – Les différentes méthodes de résolution considérées.

4.3.1 Propagation d'une onde plane dans le vide

On considère de nouveau le cas test de la propagation d'une onde plane dans le vide (cf. sous-section 4.1.1). Dans un premier temps, on utilise un maillage défini par $N_x = 41$ et $N_y = N_z = 15$ (# sommets = 9,225 - # tétraèdres = 47,040 - $N_{dl} = 282,240$ pour la méthode d'approximation GD-P0). La fréquence est fixée à 600 MHz. Le seuil de résolution des systèmes linéaires (système global, systèmes locaux et système interface) est fixé à $\epsilon = 10^{-6}$ (i.e. on considère que la convergence est atteinte lorsque

la norme euclidienne du résidu initial est réduite d'un facteur 10^6). Sur la figure 3.61, on a représenté les évolutions de la norme euclidienne du résidu pour les différentes stratégies de résolution. La méthode GLOB converge en 383 itérations. Le nombre maximal d'itérations de la méthode DDMINT1 a été fixé à 100 et il s'avère que le seuil de résolution fixé n'est pas atteint à ce nombre d'itérations. Ce comportement est néanmoins conforme avec le résultat de l'analyse de convergence menée dans la sous-section 3.2.2.2 du chapitre 3. L'accélération de la résolution du système interface par la méthode BiCGstab(ℓ) est clairement démontrée (convergence en 5 itérations pour les stratégies DDMINT2 et DDMINT3). Les temps de résolution correspondants sont résumés dans la table 4.10 (ces temps n'incluent pas les temps de construction des factorisations LU et ILU). Dans cette table, le temps de calcul de la méthode DDMINT1 est celui nécessaire à la réalisation de 100 itérations. Par ailleurs, dans la table 4.11, on précise les occupations mémoire des stratégies de résolution qui font appel au solveur direct MUMPS et les temps de construction des factorisations LU associées. En résumé cette première série de résultats montre que la stratégie GLOB basée sur la factorisation globale (parallèle) MUMPS est de loin la plus performante suivie par la stratégie DDMINT3. Néanmoins, en termes de temps de construction des factorisations et d'occupation mémoire associées, la stratégie de résolution par décomposition de domaine est nettement plus avantageuse. A partir de maintenant on se concentre sur les stratégies de résolution GLOB et DDMINT3.

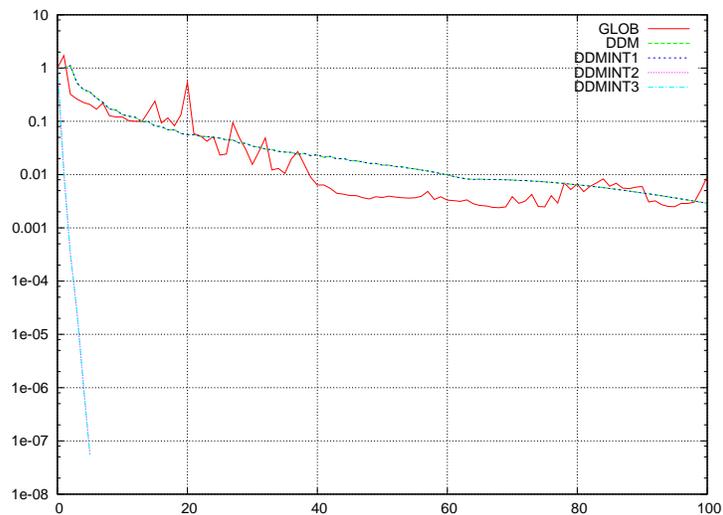


FIG. 3.61 – Equations de Maxwell 3D : propagation d'une onde plane dans le vide.

$N_x = 41$, $N_y = N_z = 15$, $F = 600$ MHz, méthode d'approximation GD-P0.

Convergence des différentes méthodes de résolution.

Abscisse : nombre d'itérations - Ordonnée : résidu normalisé (échelle logarithmique).

On considère ensuite un maillage défini par $N_x = 81$ et $N_y = N_z = 15$ (# sommets = 18,225 - # tétraèdres = 94,080 - $N_{dl} = 564,480$ pour la méthode d'approximation GD-P0). La fréquence est toujours fixée à 600 MHz. Les résultats obtenus sont rassemblés dans les tables 4.12 et 4.13 dans lesquelles on a fait apparaître l'accélération parallèle $S(N_{procs})$ évaluée à partir des temps écoulés. Il s'agit ici d'une accélération relative au temps obtenu pour le plus petit nombre de sous-domaines utilisé pour un cas de calcul donné. On constate que la méthode GLOB ne tire pas avantage de l'augmentation du nombre de processus au contraire de la stratégie DDMINT3 qui se singularise par des accélérations

Stratégie de résolution	# iter	CPU	REEL	%CPU
GLOBD	1	2.0 sec	3.0 sec	66.0%
GLOB	383	227.0 sec	239.0 sec	95.0%
DDM	> 100	1070.0 sec	1074.0 sec	99.0%
DDMINT1	> 100	997.0 sec	1025.0 sec	97.0%
DDMINT2	5	698.0 sec	702.0 sec	99.0%
DDMINT3	5	59.0 sec	63.0 sec	94.0%

TAB. 4.10 – Equations de Maxwell 3D : propagation d’une onde plane dans le vide.
 $N_x = 41$, $N_y = N_z = 15$, $F = 600$ MHz, méthode d’approximation GD-P0.
Performances des différentes stratégies de résolution (phase de résolution).
Cluster AMD Opteron/2 GHz, Gigabit Ethernet, N_{procs} .

Stratégie de résolution	RAM (min/max)	CPU	REEL
GLOBD	407 MB/452 MB	25.0 sec	29.0 sec
DDMINT3	164 MB/164 MB	7.0 sec	7.0 sec

TAB. 4.11 – Equations de Maxwell 3D : propagation d’une onde plane dans le vide.
 $N_x = 41$, $N_y = N_z = 15$, $F = 600$ MHz, méthode d’approximation GD-P0.
Performances des différentes stratégies de résolution (phase de factorisation).
Cluster AMD Opteron/2 GHz, Gigabit Ethernet, N_{procs} .

superlinéaires dans la phase de construction des factorisations locales.

Stratégie de résolution	N_{procs}	# iter	CPU	REEL	$S(N_{\text{procs}})$
GLOBD	4	1	3.2 sec	5.5 sec	1.00
-	8	1	2.1 sec	5.8 sec	0.95
-	16	1	1.6 sec	5.1 sec	1.10
DDMINT3	4	6	157.0 sec	159.0 sec	1.00
-	8	6	68.0 sec	69.0 sec	2.30
-	16	7	38.0 sec	40.0 sec	3.95

TAB. 4.12 – Equations de Maxwell 3D : propagation d’une onde plane dans le vide.
 $N_x = 81$, $N_y = N_z = 15$, $F = 600$ MHz, méthode d’approximation GD-P0.
Performances des différentes stratégies de résolution (phase de résolution).
Cluster AMD Opteron/2 GHz, Gigabit Ethernet.

Si on utilise maintenant un maillage défini par $N_x = 81$ et $N_y = N_z = 31$ (# sommets = 77,841 - # tétraèdres = 432,000 - $N_{\text{procs}} = 2,592,000$ pour la méthode d’approximation GD-P0), il n’est pas possible de réaliser un calcul avec la stratégie GLOBD pour un nombre modéré de processeurs. Ainsi l’occupation mémoire minimale requise atteint 2478 MB pour $N_{\text{procs}} = 8$ pour une capacité maximale disponible de 2048 MB par nœud biprocesseur dans le cas du cluster à notre disposition.

Enfin, on complète les résultats précédents en considérant la méthode d’approximation GD-P1. Les calculs ont été réalisés avec la stratégie de résolution DDMINT3 exclusivement. Les résultats sont résumés dans les tables 4.14 et 4.15 pour un maillage défini par $N_x = 41$ et $N_y = N_z = 15$ (# sommets

Stratégie de résolution	N_{procs}	RAM (min/max)	CPU	REEL	$S(N_{\text{procs}})$
GLOBD	4	789 MB/831 MB	64.0 sec	74.0 sec	1.00
-	8	458 MB/727 MB	39.0 sec	65.0 sec	1.15
-	16	286 MB/371 MB	31.0 sec	50.0 sec	1.50
DDMINT3	4	363 MB/363 MB	17.0 sec	18.0 sec	1.00
-	8	132 MB/132 MB	4.5 sec	5.5 sec	3.25
-	16	64 MB/ 64 MB	2.3 sec	2.7 sec	6.65

TAB. 4.13 – Equations de Maxwell 3D : propagation d’une onde plane dans le vide.

$N_x = 81$, $N_y = N_z = 15$, $F = 600$ MHz, méthode d’approximation GD-P0.

Performances des différentes stratégies de résolution (phase de factorisation).

Cluster AMD Opteron/2 GHz, Gigabit Ethernet.

= 9,225 - # tétraèdres = 47,040 - $N_{\text{dl}} = 1,128,960$), et dans les tables 4.16 et 4.17 pour un maillage défini par $N_x = 81$ et $N_y = N_z = 15$ (# sommets = 18,225 - # tétraèdres = 94,080 - $N_{\text{dl}} = 2,257,920$ pour la méthode GD-P1). On note une certaine stabilité du nombre d’itérations à convergence de la méthode BiCGstab(ℓ) lorsque la taille du problème et le nombre de processus augmentent. Ce constat doit cependant être modéré par le fait que l’on a considéré une faible variation du nombre d’inconnues et de processeurs.

Stratégie de résolution	N_{procs}	# iter	CPU	REEL	$S(N_{\text{procs}})$
DDMINT3	8	7	265.0 sec	273.0 sec	1.00
-	16	7	124.0 sec	128.0 sec	2.15

TAB. 4.14 – Equations de Maxwell 3D : propagation d’une onde plane dans le vide.

$N_x = 41$, $N_y = N_z = 15$, $F = 600$ MHz, méthode d’approximation GD-P1.

Performances de la stratégie de résolution DDMINT3 (phase de résolution).

Cluster AMD Opteron/2 GHz, Gigabit Ethernet.

Stratégie de résolution	N_{procs}	RAM (min/max)	CPU	REEL	$S(N_{\text{procs}})$
DDMINT3	8	1184 MB/1184 MB	112.0 sec	114.0 sec	1.00
-	16	553 MB/ 553 MB	39.0 sec	40.0 sec	2.85

TAB. 4.15 – Equations de Maxwell 3D : propagation d’une onde plane dans le vide.

$N_x = 41$, $N_y = N_z = 15$, $F = 600$ MHz, méthode d’approximation GD-P1.

Performances de la stratégie de résolution DDMINT3 (phase de factorisation).

Cluster AMD Opteron/2 GHz, Gigabit Ethernet.

4.3.2 Diffraction par une sphère parfaitement conductrice

Nous considérons ici le cas de calcul étudié dans la sous-section 4.1.2. Précisons tout de suite que l’essentiel des résultats obtenus l’ont été avec la méthode d’approximation GD-P0. En effet, le cluster à notre disposition comporte 64 processeurs et un total de 64 GB de mémoire RAM. L’application de la stratégie de résolution DDMINT3 n’a pas été possible pour la méthode d’approximation GD-P1 même

Stratégie de résolution	N_{procs}	# iter	CPU	REEL	$S(N_{\text{procs}})$
DDMINT3	16	8	459.0 sec	491.0 sec	1.00
-	32	8	209.0 sec	231.0 sec	2.15

TAB. 4.16 – Equations de Maxwell 3D : propagation d’une onde plane dans le vide.
 $N_x = 81$, $N_y = N_z = 15$, $F = 600$ MHz, méthode d’approximation GD-P1.
Performances de la stratégie de résolution DDMINT3 (phase de résolution).
Cluster AMD Opteron/2 GHz, Gigabit Ethernet.

Stratégie de résolution	N_{procs}	RAM (min/max)	CPU	REEL	$S(N_{\text{procs}})$
DDMINT3	16	1184 MB/1184 MB	112.0 sec	118.0 sec	1.00
-	32	553 MB/ 553 MB	39.0 sec	40.0 sec	2.95

TAB. 4.17 – Equations de Maxwell 3D : propagation d’une onde plane dans le vide.
 $N_x = 81$, $N_y = N_z = 15$, $F = 600$ MHz, méthode d’approximation GD-P1.
Performances de la stratégie de résolution DDMINT3 (phase de factorisation).
Cluster AMD Opteron/2 GHz, Gigabit Ethernet.

pour le maillage le plus grossier (cf. la table 4.1) du fait de l’occupation mémoire requise pour construire et stocker les facteurs L et U par sous-domaine. Les figures 3.62 et 3.63 montrent respectivement les évolutions de la norme euclidienne du résidu pour les stratégies de résolution GLOB et DDMINT3 pour chacun des maillages de la table 4.1 et lorsque l’on utilise la méthode d’approximation GD-P0. Les maillages sont partitionnés en utilisant l’outil ParMeTiS qui repose sur un algorithme multiniveaux pour le partitionnement de graphes irréguliers[Karypis and Kumar, 1999]. Les résultats de performance sont rassemblés dans les tables 4.18 à 4.21. Plusieurs commentaires s’imposent :

- la stratégie de résolution DDMINT3 est particulièrement à son avantage lorsque le pas de discrétisation spatial diminue et que le nombre de sous-domaines augmente. Ainsi, le rapport entre les temps de calcul des deux stratégies passe de 3.5 (maillage M1 pour $N_{\text{procs}} = 16$) à 6.5 (maillage M3 pour $N_{\text{procs}} = 64$).
- en dépit de l’utilisation d’un outil de partitionnement performant, le défaut de répartition de la charge de calcul, même minime (pour le maillage M3, le nombre d’éléments d’un sous-maillage est compris entre 4520 et 5010), se fait particulièrement ressentir pour la stratégie DDMINT3 à travers les factorisations LU locales comme en témoigne les données de la table 4.21.
- un avantage incontestable de la stratégie de résolution DDMINT3 est son efficacité parallèle évaluée ici par le taux '%CPU' qui, pour nos expériences numériques, est toujours supérieure à 89%.

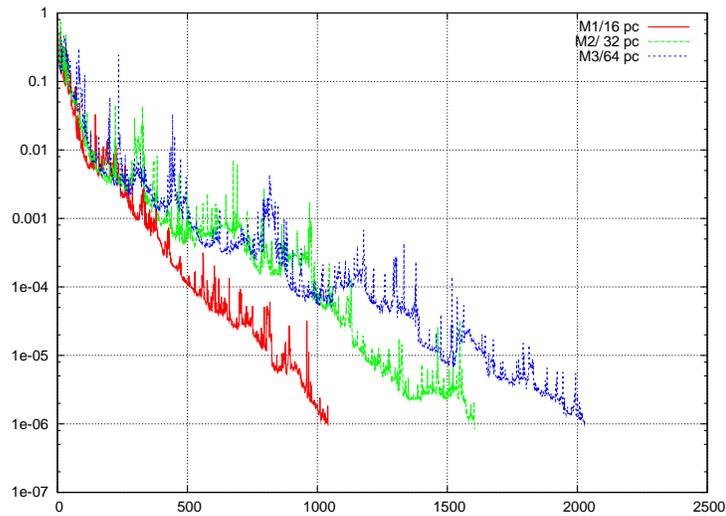


FIG. 3.62 – Equations de Maxwell 3D : diffraction d’une onde plane par une sphère parfaitement conductrice.

$F = 600$ MHz, méthode d’approximation GD-P0.

Convergence de la stratégie de résolution GLOB.

Abscisse : nombre d’itérations - Ordonnée : résidu normalisé (échelle logarithmique).

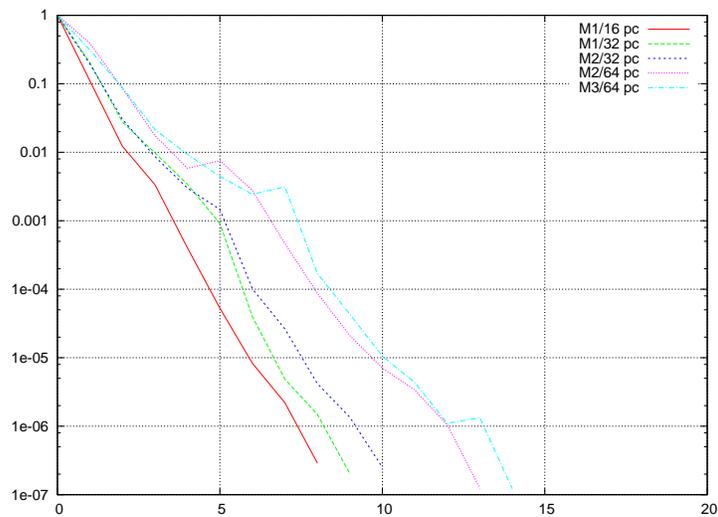


FIG. 3.63 – Equations de Maxwell 3D : diffraction d’une onde plane par une sphère parfaitement conductrice.

$F = 600$ MHz, méthode d’approximation GD-P0.

Convergence de la stratégie de résolution DDMINT3.

Abscisse : nombre d’itérations - Ordonnée : résidu normalisé (échelle logarithmique).

Maillage	N_{procs}	# iter	CPU (min/max)	REEL	% CPU
M1	16	1036	1098.0 sec/1200.0 sec	1383.0 sec	87.0%
-	32	1042	577.0 sec/ 636.0 sec	831.0 sec	76.5%
M2	32	1605	1846.0 sec/1978.0 sec	2409.0 sec	82.0%
-	64	1614	986.0 sec/1101.0 sec	1458.0 sec	75.5%
M3	64	2031	1940.0 sec/2142.0 sec	2919.0 sec	73.5%

TAB. 4.18 – Equations de Maxwell 3D : diffraction d’une onde plane par une sphère parfaitement conductrice.

$F = 600$ MHz, méthode d’approximation GD-P0.
Performances de la stratégie de résolution GLOB.
Cluster AMD Opteron/2 GHz, Gigabit Ethernet.

Maillage	N_{procs}	# iter	CPU (min/max)	REEL	% CPU
M1	64	5832	7733.0 sec/8409.0 sec	11725.0 sec	72.0%

TAB. 4.19 – Equations de Maxwell 3D : diffraction d’une onde plane par une sphère parfaitement conductrice.

$F = 600$ MHz, méthode d’approximation GD-P1.
Performances de la stratégie de résolution GLOB.
Cluster AMD Opteron/2 GHz, Gigabit Ethernet.

Maillage	N_{procs}	# iter	CPU (min/max)	REEL	% CPU
M1	16	8	211.0 sec/390.0 sec	398.0 sec	98.0%
-	32	9	97.0 sec/129.0 sec	143.0 sec	90.0%
M2	32	10	260.0 sec/477.0 sec	481.0 sec	99.0%
-	64	13	131.0 sec/191.0 sec	215.0 sec	89.0%
M3	64	14	259.0 sec/413.0 sec	452.0 sec	91.5%

TAB. 4.20 – Equations de Maxwell 3D : diffraction d’une onde plane par une sphère parfaitement conductrice.

$F = 600$ MHz, méthode d’approximation GD-P0.
Performances de la stratégie de résolution DDMINT3 (phase de résolution).
Cluster AMD Opteron/2 GHz, Gigabit Ethernet.

Maillage	N_{procs}	RAM (min/max)	CPU (min/max)	REEL
M1	16	725 MB/1129 MB	55.0 sec/121.0 sec	125.0 sec
-	32	269 MB/ 404 MB	11.0 sec/ 26.0 sec	27.0 sec
M2	32	650 MB/1318 MB	43.0 sec/144.0 sec	146.0 sec
-	64	188 MB/ 462 MB	7.0 sec/ 34.0 sec	36.0 sec
M3	64	384 MB/1001 MB	18.0 sec/100.0 sec	102.0 sec

TAB. 4.21 – Equations de Maxwell 3D : diffraction d’une onde plane par une sphère parfaitement conductrice.

$F = 600$ MHz, méthode d’approximation GD-P0.

Performances de la stratégie de résolution DDMINT3 (phase de factorisation).

Cluster AMD Opteron/2 GHz, Gigabit Ethernet.

Conclusion générale

Conclusion

Ce manuscrit est consacré à l'étude et la mise en œuvre de méthodes de type Galerkin discontinu en maillages non-structurés pour la résolution numérique des équations de Maxwell en régime harmonique. Nous nous sommes ici limités à des méthodes basées sur des fonctions de base polynomiales d'ordre 0 (méthode volumes finis) et 1 (méthode Galerkin discontinu linéaire). Un des principaux objectifs de cette thèse était d'adapter des méthodes de ce type précédemment développées dans le projet CAIMAN de l'INRIA Sophia Antipolis, pour la résolution numérique des équations de Maxwell en domaine temporel, à la résolution des équations de Maxwell en régime harmonique. Un autre objectif était de mettre au point des stratégies performantes pour la résolution des systèmes algébriques résultants. Ainsi, nous avons pu bâtir un socle théorique et numérique de base qui permet dès lors de connaître les directions à prendre pour améliorer ces travaux.

Nous avons dans un premier temps étudié le problème discret issu des équations de Maxwell harmoniques perturbées dont la discrétisation repose sur une formulation Galerkin discontinu d'ordre quelconque. Nous avons montré le caractère bien posé du problème dans le cas d'un ouvert polyédrique muni de conditions absorbantes de type Silver-Müller (aproximation d'ordre un de la condition de radiation exacte). La dissipativité du schéma induite par la perturbation a permis d'aboutir à une forme bilinéaire coercive et de conclure quant à l'existence et l'unicité de la solution.

Nous avons ensuite étudié les propriétés élémentaires de l'opérateur matriciel résultant de la formulation du schéma Galerkin discontinu d'ordre m en une et trois dimensions d'espace. Ces propriétés ont permis en l'occurrence de montrer l'existence et l'unicité de la solution dans le cas 1D pour des fonctions nodales \mathbb{P}_0 et \mathbb{P}_1 en raisonnant directement sur l'opérateur matriciel provenant de la discrétisation. Les tests numériques préliminaires 1D ont par la suite donné une première impression des forces et des faiblesses de ces schémas. En effet, bien que le schéma basé sur l'approximation locale \mathbb{P}_0 soit peu coûteux en terme de place mémoire, il n'est pas adapté aux maillages fortement non-uniformes et génère des oscillations parasites. De ce point de vue, le schéma \mathbb{P}_1 est plus robuste. Il semble également que ce schéma tienne mieux compte des fortes hétérogénéités du milieu de propagation. Notons aussi que l'étude de la dispersion numérique a montré que le schéma fréquentiel est plus dispersif que le schéma temporel, et que le schéma GD-P1 fréquentiel est lui-même bien moins dispersif que le schéma fréquentiel GD-P0.

En ce qui concerne les stratégies de résolution des systèmes algébriques résultants des formulations Galerkin discontinu GD-P0 et GD-P1 en maillages tétraédriques, et en vue d'aboutir à des algorithmes utilisables sur des architectures de calcul parallèles, nous avons opté pour des méthodes de décomposition de domaine. Nous nous sommes attardés à une famille de méthodes s'appuyant sur la formulation d'un algorithme de Schwarz additif avec ou sans recouvrement, et pour deux types de conditions d'interface :

des conditions absorbantes approchées et des conditions optimisées. Ces méthodes ont été appliquées aux équations de Maxwell 2D en régime harmonique discrétisées par la méthode Galerkin discontinu GD-P0 sur un domaine canonique homogène en maillage quadrangulaire décomposé en deux sous-domaines. Nous avons étudié ces méthodes en terme de vitesse de convergence, tant sur le plan théorique en calculant le taux de convergence relatif à chaque type de condition d'interface par une analyse de Fourier, que sur le plan numérique en simulant de nombreux cas tests basés sur la propagation d'une onde plane dans le vide. Nous avons ainsi constaté le gain obtenu en terme de vitesse de convergence de la méthode utilisant les conditions optimisées par rapport à la méthode utilisant les conditions classiques, ainsi que la robustesse de la méthode qui dépend très faiblement de la finesse du maillage. Néanmoins, le gain escompté par le troisième jeu de conditions d'interface optimisées ne fut pas significatif en raison du manque de précision de l'algorithme de minimisation numérique. Bien qu'ayant constaté un gain important entre l'algorithme muni des conditions classiques et l'algorithme muni des conditions optimisées, on peut légitimement se demander si le gain attendu par le deuxième jeu de conditions optimisées n'est pas négligeable par rapport au premier.

Enfin, nous avons validé les méthodes GD-P0 et GD-P1 en trois dimensions d'espace sur des cas tests allant de la propagation d'une onde plane dans le vide, jusqu'à la diffraction par un cube ou une sphère métallique. Les observations concernant les résultats 3D sont en accord avec celles qui ont été faites en 1D : la méthode GD-P0 est précise du moment que le maillage n'est pas trop non-uniforme, même si de meilleurs résultats ont été obtenus avec le schéma GD-P1 sur un maillage grossier et davantage non-uniforme. Les comparaisons entre les différentes stratégies de résolution parallèle développées ont montré la supériorité de l'approche par décomposition de domaine basée dans le cas présent sur des conditions d'interface classiques. En particulier, cette stratégie de résolution est celle qui tire profit le plus nettement de l'augmentation du nombre de processeurs comparativement à une approche globale. Le seul bémol notable concerne le défaut de répartition de charge entre les sous-domaines lié à l'utilisation de factorisations LU pour le préconditionnement de systèmes locaux. A ce sujet, le préconditionnement local effectué grâce au solveur direct MUMPS en simple précision a montré des performances inattendues et très intéressantes. Ainsi, aussi bien dans le cas d'un préconditionnement local dans la méthode de décomposition de domaine que dans celui d'un préconditionnement global du solveur itératif BiCGstab(ℓ), cette approche a montré de bien meilleures performances que les stratégies utilisant un préconditionnement par une méthode ILU.

Perspectives

A l'heure actuelle, il est clair que les méthodes que nous avons développées à ce stade pour traiter le problème de la diffraction d'ondes ne sont pas suffisamment matures pour rivaliser avec les méthodes intégrales accélérées par les méthodes multipôles, ou même les méthodes d'éléments finis d'arête de Nédélec en ce qui concerne les formulations volumiques. En revanche, certains points forts des méthodes de type Galerkin discontinu permettront lors de prochains travaux de concurrencer sérieusement ces méthodes de référence pour les raisons suivantes :

- le caractère discontinu de la méthode facilite l'implémentation en maillage non-conforme et s'adapte aux très fortes hétérogénéités de matériaux,
- le caractère local facilite l'adaptativité pour traiter des zones de maillages par des fonctions de base dont l'ordre varie spatialement,
- le caractère local induit aussi un haut degré de parallélisme dans les stratégies de résolution parallèle exploitant la décomposition de domaine.

Concernant des perspectives à très court terme qui seraient directement issues du prolongement des travaux réalisés dans cette thèse, on propose de traiter certains obstacles auxquels nous nous sommes heurtés ainsi qu'une amélioration des autres points.

Plus précisément, le premier problème qu'il faut résoudre est de montrer le caractère bien posé du problème discret pour une perturbation nulle. La coercivité de la forme bilinéaire étant perdue, ceci rend son traitement beaucoup plus délicat. Une solution à explorer consisterait à utiliser le principe d'absorption limite ainsi que des méthodes provenant de la théorie du contrôle. Cette démonstration permettrait par la même occasion de résoudre le problème rencontré dans la formulation du schéma en 3D concernant l'inversibilité de l'opérateur matriciel.

Ensuite, pour ce qui est du schéma à proprement parler, il serait peut être judicieux d'utiliser des fonctions de base différentes telles que les fonctions de base hiérarchiques qui faciliteraient la mise en œuvre de stratégies de variation locale de l'ordre d'interpolation. De même, il serait intéressant d'implémenter la méthode en utilisant un schéma à flux décentrés et de comparer ces deux type d'approximation des flux afin de conclure sur l'option la plus appropriée en régime harmonique. Il est clair que le caractère non-diffusif du schéma à flux centrés a beau se révéler indispensable pour des simulations en domaine temporel, ce caractère offrirait en régime harmonique des matrices bien mieux conditionnées et faciliterait donc la résolution du système linéaire. En revanche, il est d'actualité d'utiliser dans ce cas des méthodes d'ordre élevé dans la mesure où les approximations GD d'ordre 0 ont déjà été étudiées et ont montré leurs limites.

Quant aux méthodes de décomposition de domaine, la première extension qui s'impose est de réaliser pour un ordre d'interpolation $m \geq 1$, en géométrie quelconque et en milieu hétérogène, une étude similaire à celle effectuée pour un schéma GD-P0 afin de voir si les résultats obtenus sont aussi encourageants. Par ailleurs, l'utilisation de conditions d'interface optimisées fournirait à l'ordre m un algorithme en sous-domaines non-recouvrants où la formulation du complément de Schur reposerait non plus sur les cellules mais sur les degrés de liberté communs à chaque interface. Il serait également utile de procéder à une optimisation numérique plus poussée du deuxième jeu de conditions optimisées en cherchant un algorithme de minimisation numérique performant pour résoudre les problèmes de type min-max dont la solution n'est pas un point-selle.

Enfin, précisons que le solveur MUMPS a montré en tant que préconditionneur des performances qui méritent d'approfondir cette voie, et serait une alternative intéressante à l'utilisation de préconditionneurs approchés. Cela n'exclue pas pour autant de chercher des préconditionneurs approchés moins gourmands en mémoire.

Bibliographie

- [Alonso-Rodriguez and Gerardo-Giorda, 2004] Alonso-Rodriguez, A. and Gerardo-Giorda, L. (2004). New non-overlapping domain decomposition methods for the harmonic Maxwell system. Technical Report UTM 667, University of Trento.
- [Amestoy et al., 2000] Amestoy, P., Duff, I., and L'Excellent, J.-Y. (2000). Multifrontal parallel distributed symmetric and unsymmetric solvers. *Comput. Meth. App. Mech. Engng.*, 184 :501–520.
- [Arnold, 1982] Arnold, D. (1982). An interior penalty finite element method with discontinuous elements. *SIAM J. Numer. Anal.*, 19 :742–760.
- [Arnold et al., 2002] Arnold, D., Brezzi, F., Cockburn, B., and Marini, L. (2002). Unified analysis of discontinuous Galerkin methods for elliptic problems. *SIAM J. Numer. Anal.*, 39(5) :1749–1779.
- [Assous et al., 1999] Assous, F., Jr., P. C., Raviart, P., and Sonnendrucker, E. (1999). Characterization of the singular part of the solution of Maxwell's equations in a polyhedral domain. *Math. Meth. Appl. Sci.*, 22 :485–499.
- [Babuska and Ihlenburg, 1995] Babuska, I. and Ihlenburg, F. (1995). Finite element solution to the Helmholtz equation with high wavenumber. *Comput. Math. Appl.*, 30(9) :9–37.
- [Bernacki et al., 2006] Bernacki, M., Lanteri, S., and Piperno, S. (2006). Time-domain parallel simulation of heterogeneous wave propagation on unstructured grids using explicit non-diffusive, discontinuous Galerkin methods. *J. Comp. Acoustics*, 14(1) :57–81.
- [Bernardi et al., 2001] Bernardi, P., Cavagnaro, M., Pisa, S., and Piuze, E. (2001). Power absorption and temperature elevations induced in the human head by a dual-band monopole-helix antenna phone. *IEEE Trans. Microwave Theory Tech.*, 49(12) :2539–2546.
- [Bonnet, 1998] Bonnet, P. (1998). Résolution des équations de maxwell instationnaires et harmoniques par une technique de volumes finis. Application à des problèmes de compatibilité électromagnétique. Technical Report NT 1998-8, ONERA.
- [Bourgat et al., 1994] Bourgat, J. F., Tallec, P. L., Perthame, B., and Qiu, Y. (1994). Coupling Boltzmann and Euler equations without overlapping. In Quarteroni, A., Périaux, J., Kuznetsov, Y., and Widlund, O., editors, *6th International Conference on Domain Decomposition Methods*, pages 377–398. AMS.
- [Cai and Widlund, 1992] Cai, X.-C. and Widlund, O. (1992). Domain decomposition algorithms for indefinite elliptic problems. *SIAM J. Sci. Statist. Comput.*, 13 :243–258.
- [Champaney et al., 1997] Champaney, L., Cognard, J.-Y., Dureisseix, D., and Ladevèze, P. (1997). Large scale applications on parallel computers of a mixed domain decomposition method. *Comput. Mech.*, (19) :253–263.

- [Clatz et al., 2000] Clatz, O., Lanteri, S., Oudot, S., Pons, J.-P., Piperno, S., Scarella, G., and Wiart, J. (2000). Modélisation numérique réaliste de l'exposition des tissus de la tête à un champ électromagnétique issu d'un téléphone mobile. In *13ème Colloque International et Exposition sur la Compatibilité Electromagnétique*, pages 377–397, Saint Malo, France.
- [Clerc, 1998] Clerc, S. (1998). Non-overlapping Schwarz method for systems of first order equations. *Cont. Math.*, 218 :408–416.
- [Cockburn and Shu, 1989] Cockburn, B. and Shu, C. (1989). TVD Runge-Kutta projection discontinuous Galerkin finite element method for conservatoin laws II : general framework. *Math. Comp.*, 52 :411–435.
- [Cohen and Monk, 1998] Cohen, G. and Monk, P. (1998). Gauss point mass lumping schemes for Maxwell's equations. *Numer. Meth. Part. Diff. Eqns.*, 14 :63–68.
- [Darrigrand, 2002] Darrigrand, E. (2002). *Couplage méthodes multipôles - discrétisation microlocale pour les équations intégrales de l'électromagnétisme*. Thèse de doctorat en mathématiques appliquées, Université de Bordeaux 1.
- [Dautray and Lions, 1985] Dautray, R. and Lions, J. (1985). *Analyse mathématique et calcul numérique pour les sciences et les techniques*, volume 1. Masson.
- [de La Bourdonnaye et al., 1998] de La Bourdonnaye, A., Farhat, C., Macedo, A., Magoulès, F., and Roux, F.-X. (1998). A non-overlapping domain decomposition method for exterior Helmholtz problems. In *10th International Conference on Domain Decomposition Methods*, pages 42–66, Boulder, Colorado, USA. AMS.
- [Demkowicz and Xue, 2005] Demkowicz, L. and Xue, D. (2005). Control in geometry induced in hp finite element simulations. *Numer. Anal. and Model.*, 2(3) :283–300.
- [Després, 1991] Després, B. (1991). *Méthodes de décomposition de domaine pour les problèmes de propagation d'ondes en régimes harmoniques*. Thèse de doctorat en mathématiques, Paris IX.
- [Després et al., 1992] Després, B., Joly, P., and Roberts, J. E. (1992). A domain decomposition for the harmonic Maxwell equation. *Iter. Meth. Lin. Algbr.*, pages 475–484.
- [Dolean, 2001] Dolean, V. (2001). *Algorithmes par décomposition de domaine et accélération multigrille pour le calcul d'écoulements compressibles*. Thèse de doctorat en mathématiques, Université de Nice-Sophia Antipolis.
- [Dolean et al., 2006] Dolean, V., Gander, M., and Gerardo-Giorda, L. (2006). On the relation between the Schwarz algorithms applied to systems of PDE and scalar problems. In preparation.
- [Dolean et al., 2004] Dolean, V., Lanteri, S., and Nataf, F. (2004). Convergence analysis of additive Schwarz for the Euler equations. *Appl. Numer. Math.*, 49 :153–186.
- [Duruflé, 2006] Duruflé, M. (2006). *Intégration numérique et éléments finis d'ordre élevé appliqués aux équations de Maxwell en régime harmonique*. Thèse de doctorat en mathématiques Appliquées, Université Paris Dauphine.
- [Farhat and Roux, 1991] Farhat, C. and Roux, F.-X. (1991). A method of finite element tearing and interconnecting and its parallel solution algorithm. *Int. J. Numer. Meth. in Engrg.*, 32 :1205–1227.
- [Fatone et al., 1999] Fatone, L., Gervasio, P., and Quarteroni, A. (1999). Multimodels for incompressible flows. Technical Report 09.99, EPFL.
- [Fezoui et al., 2005] Fezoui, L., Lanteri, S., Lohrengel, S., and Piperno, S. (2005). Convergence and stability of a discontinuous Galerkin time-domain method for the 3D heterogeneous Maxwell equations on unstructured meshes. *ESAIM : Modél. Math. Anal. Numér.*, 39(6) :1149–1176.

- [Gander, 2006] Gander, M. (2006). Optimized Schwarz methods. *SIAM J. Numer. Anal.*, 44(2) :699–731.
- [Gander et al., 2002] Gander, M. J., Magoulès, F., and Nataf, F. (2002). Optimized Schwarz methods without overlap for the Helmholtz equation. *SIAM J. Sci. Stat. Comput.*, 24 :38–60.
- [Gerdes and Ilhenburg, 1999] Gerdes, K. and Ilhenburg, F. (1999). On the pollution effect in FE solutions of the 3D Helmholtz equation with large wavenumber. *Comput. Methods Appl. Mech. Engrg.*, 156 :155–172.
- [Goto and van de Geijn, 2006] Goto, K. and van de Geijn, R. (2006). Anatomy of high-performance matrix multiplication. *ACM Transactions on Mathematical Software*. soumis.
- [Helluy, 1994] Helluy, P. (1994). *Résolution numérique des équations de Maxwell harmoniques par une méthode d'éléments finis discontinus*. Thèse de doctorat en mathématiques appliquées, Ecole Nationale Supérieure de l'Aéronautique.
- [Helluy and Dayma, 1994] Helluy, P. and Dayma, S. (1994). Convergence d'une approximation discontinus des systèmes du premier ordre. *C. R. Acad. Sci. Paris*, 319 :1331–1335.
- [Helluy et al., 1994] Helluy, P., Mazet, P., and Klotz, P. (1994). Sur une approximation en domaine non borné des équations de Maxwell instationnaires : comportement asymptotique. *La Recherche Aérospatiale*, 5 :365–377.
- [Hénon et al., 2002] Hénon, P., Ramet, P., and Roman, J. (2002). PaStiX : a high-performance parallel direct solver for sparse symmetric definite systems. *Parallel Comput.*, 28 :301–321.
- [Hesthaven and Warburton, 1999] Hesthaven, J. and Warburton, T. (1999). Nodal high-order methods for symmetric hyperbolic equations. *SIAM J. Numer. Anal.*, 36 :935–952.
- [Hesthaven and Warburton, 2004a] Hesthaven, J. and Warburton, T. (2004a). Discontinuous Galerkin methods for the time-domain Maxwell's equations : an introduction. *ACES Newsletter*, 19 :10–29.
- [Hesthaven and Warburton, 2004b] Hesthaven, J. and Warburton, T. (2004b). High-order nodal Discontinuous Galerkin methods for the Maxwell eigenvalue problem. *Phil. Trans. Roy. Soc. London A*, 362 :493–524.
- [Hesthaven and Warburton, 2006] Hesthaven, J. and Warburton, T. (2006). The role of the penalty in the local discontinuous Galerkin method for Maxwell's eigenvalue problem. *Comput. Meth. App. Mech. Engrg.*, 195(25–28) :3205–3223.
- [Hill and Reed, 1973] Hill, T. and Reed, W. (1973). Triangular mesh methods for the neutron transport equation. Technical Report LA-UR-73-479, Los Alamos Scientific Laboratory.
- [Hiptmair, 2001] Hiptmair, R. (2001). Higher order Whitney forms. *Prog.In Electr. Res. (PIER)*, 32 :271–299.
- [Houston et al., 2004] Houston, P., Perugia, I., and Schötzau, D. (2004). Mixed discontinuous Galerkin approximation of the Maxwell operator. *SIAM J. Numer. Anal.*, 42 :434–459.
- [Ilhenburg, 1998] Ilhenburg, F., editor (1998). *Finite element analysis of acoustic scattering*. Springer-Verlag, Berlin.
- [J. Douglas Jr. and Dupont, 1976] J. Douglas Jr. and Dupont, T. (1976). Interior penalty procedures for elliptic and parabolic Galerkin methods. In *Lecture notes in Physics*, volume 58. Springer-Verlag Berlin.
- [Karypis and Kumar, 1999] Karypis, G. and Kumar, V. (1999). Parallel multilevel k-way partition scheme for irregular graphs. *SIAM Review*, 41(2) :278–300.

- [Kurzak and Dongarra, 2006] Kurzak, J. and Dongarra, J. (2006). Implementation of the mixed-precision in solving systems of linear equations on the CELL processor. Technical Report UT-CS-06-580, University of Tennessee.
- [Langou et al., 2006] Langou, J., Langou, J., Luszczek, P., Kurzak, J., Buttari, A., and Dongarra, J. (2006). Exploiting the performance of 32 bit floating point arithmetic in obtaining 64 bit accuracy. Technical Report UT-CS-06-574, University of Tennessee.
- [Lanteri, 1996] Lanteri, S. (1996). Parallel solutions of compressible flows using overlapping and non-overlapping mesh partitioning strategies. *Parallel Comput.*, 22 :943–968.
- [Le Tallec, 1994] Le Tallec, P. (1994). Domain decomposition methods in computational mechanics. In Oden, J., editor, *Computational Mechanics Advances*, volume 1, pages 121–220. North-Holland.
- [Lee et al., 2005] Lee, S.-C., Vouvakis, M., and Lee, J.-F. (2005). A non-overlapping domain decomposition method with non-matching grids for modeling large finite antenna arrays. *J. Comput. Phys.*, 203 :1–21.
- [Lin et al., 2000] Lin, J., Hirai, S., Chiang, C.-L., Hsu, W.-L., Su, J.-L., and Y.-J.Wang (2000). Computer simulation and experimental studies of SAR distributions of sleeved-slot microwave antennas for hyperthermia treatment of brain tumors. *IEEE Trans. Microwave Theory Tech.*, 48(11) :2191–2198.
- [Lions, 1988] Lions, P. (1988). On the Schwarz alternating method i. In Glowinski, R., Golub, G., Meurant, G., and Périaux, J., editors, *1st International Symposium on Domain Decomposition Methods for Partial Differential Equations*, pages 1–42. SIAM.
- [Lions, 1989] Lions, P. (1989). On the Schwarz alternating method ii. In Chan, T., Glowinski, R., Périaux, J., and Widlund, O., editors, *2nd International Symposium on Domain Decomposition Methods for Partial Differential Equations*, pages 47–70. SIAM.
- [Lions, 1990] Lions, P. (1990). On the Schwarz alternating method iii : a variant for nonoverlapping subdomains. In Chan, T., Glowinski, R., Périaux, J., and Widlund, O.-B., editors, *3rd International Symposium on Domain Decomposition Methods for Partial Differential Equations*, pages 202–223. SIAM.
- [Magoulès et al., 2004] Magoulès, F., Roux, F., and Salmon, S. (2004). Optimal discrete transmission conditions for a nonoverlapping domain decomposition method for the Helmholtz equation. *SIAM J. Sci. Comput.*, 25(5) :1497–1515.
- [Maxwell, 1987] Maxwell, J. (1985-1987). *Traité d'électricité et de magnétisme*, volume 1 et 2. Gauthier-Villars.
- [Monk, 2003] Monk, P., editor (2003). *Finite element methods for Maxwell's equations*. Oxford University Press.
- [Monk and Richter, 2005] Monk, P. and Richter, G. R. (2005). A discontinuous Galerkin method for linear symmetric hyperbolic systems in inhomogeneous media. *J. of Scient. Comput.*, 22–23 :443–477.
- [Nataf et al., 1994] Nataf, F., Rogier, F., and de Sturler, E. (1994). Optimal interface conditions for domain decomposition methods. Technical Report RI301, Centre de Mathématiques Appliquées, Ecole Polytechnique.
- [Nedelec, 1980] Nedelec, J. (1980). Mixed finite elements in \mathbb{R}^3 . *Numer. Math.*, 35 :315–341.
- [Nedelec, 1986] Nedelec, J. (1986). A new family of mixed finite elements in \mathbb{R}^3 . *Numer. Math.*, 50 :97–112.

- [Paglieri et al., 1997] Paglieri, L., Scheinine, A., Formaggia, L., and Quarteroni, A. (1997). Parallel conjugate gradient with Schwarz preconditioner applied to fluid dynamics problems. In *Parallel Computational Fluid Dynamics, Algorithms and Results using Advanced Computer (PCFD'96)*, pages 21–30.
- [Perugia and Schötzau, 2003] Perugia, I. and Schötzau, D. (2003). The hp -local discontinuous Galerkin method for low-frequency time-harmonic Maxwell equations. *Math. Comp.*, 72 :1179–1214.
- [Perugia et al., 2002] Perugia, I., Schötzau, D., and Monk, P. (2002). Stabilized interior penalty method for low-frequency time-harmonic Maxwell equations. *Comput. Meth. Appl. Mech. Engrg.*, 191 :4675–4697.
- [Pierson, 2000] Pierson, K. (2000). *A family of domain decomposition methods for the massively parallel solution of computational mechanics problems*. PhD thesis, University of Colorado at Boulder.
- [Piperno, 2000] Piperno, S. (2000). l^2 -stability of the upwind first order finite volume scheme for the Maxwell equations in two and three dimensions on arbitrary unstructured meshes. *RAIRO : Modél. Math. Anal. Numér.*, 34 :139–158.
- [Piperno and Fezoui, 2003] Piperno, S. and Fezoui, L. (2003). A centered discontinuous Galerkin finite volume scheme for the 3D heterogeneous Maxwell equations on unstructured meshes. Technical Report 4733, INRIA.
- [Piperno et al., 2002] Piperno, S., Remaki, M., and Fezoui, L. (2002). A nondiffusive finite volume scheme for the three-dimensional Maxwell's equations on unstructured meshes. *SIAM J. Num. Anal.*, 39(6) :2089–2108.
- [Quarteroni and Valli, 1991] Quarteroni, A. and Valli, A. (1991). *Theory and application of Steklov-Poincaré operators for boundary value problems*, pages 179–203. Applied and Industrial Mathematics. Kluwer Academic Publishers.
- [Quarteroni and Valli, 1999] Quarteroni, A. and Valli, A. (1999). *Domain decomposition methods for partial differential equations*. Oxford University Press.
- [Rauch, 1978] Rauch, J. (1978). Asymptotic behaviour of solutions to hyperbolic partial differential equations with zero speeds. *Comm. Pure Appl. Math.*, 31 :431–480.
- [Remaki, 2000] Remaki, M. (2000). A new finite volume scheme for solving Maxwell's system. *COMPEL*, 19 :913–931.
- [Saad, 1994] Saad, Y. (1994). ILUT : dual threshold incomplete ILU factorization for solving non-symmetric linear systems. *Numer. Lin. Alg.a with Appl.*, 1 :387–402.
- [Saad, 1996] Saad, Y. (1996). *Iterative methods for sparse linear systems*. PWS.
- [Saad and Schultz, 1986] Saad, Y. and Schultz, H. (1986). GMRES : Generalized minimal residual algorithm for solving non-symmetric linear systems. *SIAM J. Sci. Stat. Comput.*, 7 :856–869.
- [Schwarz, 1870] Schwarz, H. (1870). Über einen grenzübergang durch alternierendes verfahren. *Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich*, 15 :272–286.
- [Siauve et al., 2003] Siauve, N., Nicolas, L., Vollaire, C., and Marchal, C. (2003). 3D modeling of electromagnetic fields in local hyperthermia. *Eur. Phys. J. AP*, 21 :243–250.
- [Sleijpen and Fokkema, 1993] Sleijpen, G. and Fokkema, D. (1993). BiCGstab(ℓ) for linear equations involving unsymmetric matrices with complex spectrum. *Electron. Trans. Numer. Anal.*, 1 :11–32 (electronic only).

- [Smith, 1992] Smith, B. (1992). An optimal domain decomposition preconditioner for the finite element solution of linear elasticity problems. *SIAM J. Sci. Stat. Comput.*, 13 :364–378.
- [Smith et al., 1996] Smith, B., Bjorstad, P., and Gropp, W. (1996). *Domain decomposition and parallel multilevel methods for elliptic partial differential equations*. Cambridge University Press.
- [Steger and Warming, 1981] Steger, J. and Warming, R. (1981). Flux vector splitting for the inviscid gas dynamic with applications to finite difference methods. *J. of Comp. Phys.*, 40 :263–293.
- [Sylvand, 2002] Sylvand, G. (2002). *La méthode multipôle rapide en électromagnétisme : performances, parallélisation, applications*. Thèse de doctorat en mathématiques appliquées, Ecole Nationale des Ponts et Chaussées.
- [van der Vorst, 1992] van der Vorst, H. (1992). Bi-CGStab : a fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear systems. *SIAM J. Sci. Stat. Comput.*, 13 :631–644.
- [Wheeler, 1978] Wheeler, M. (1978). An elliptic collocation-finite element method with interior penalties. *SIAM J. Numer. Anal.*, 15 :152–161.
- [Yee, 1966] Yee, K. (1966). Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media. *IEEE Trans. Antennas and Propagation*, AP-16 :302–307.

Méthodes de type Galerkin discontinu pour la résolution numérique des équations de Maxwell 3D en régime harmonique

L'objectif général de cette étude est le développement et l'évaluation de méthodes de type Galerkin discontinu (GD) en maillages tétraédriques non-structurés pour la résolution numérique des équations de Maxwell en formulation du premier ordre et en régime harmonique. Dans la première partie de cette thèse, nous formulons et analysons des méthodes Galerkin discontinu basées sur des approximations centrées d'ordre 0 (méthode de volumes finis ou GD-P0) et d'ordre 1 (méthode de type Galerkin discontinu linéaire ou GD-P1). La seconde partie est consacrée à l'étude de méthodes de décomposition de domaine pour la résolution des systèmes algébriques issus de la discrétisation par des méthodes GD des équations de Maxwell en régime harmonique. On considère tout d'abord le système continu et on analyse la convergence d'algorithmes de Schwarz sans ou avec recouvrement basés sur des conditions d'interface naturelles. Ces conditions consistent à imposer aux interfaces les variables caractéristiques associées aux ondes entrantes dans un domaine. On s'intéresse ensuite à la convergence de ces algorithmes dans le cas discret sur la base de la méthode d'approximation volumes finis (méthode GD-P0) formulée sur un maillage quadrangulaire. On étudie enfin des conditions d'interface optimisées ayant pour but d'accélérer la convergence de l'algorithme de Schwarz sans recouvrement. Des tests préliminaires en 2D permettent de montrer clairement les gains résultant de l'utilisation de ces conditions. La troisième partie de la thèse est dédiée à l'évaluation numérique des méthodes d'approximation GD-P0 et GD-P1 en maillages tétraédriques. On considère pour cela une série de cas tests de complexité croissante portant sur des problèmes de diffraction en milieux homogènes et hétérogènes. En particulier, on évalue en détail les performances parallèles d'un algorithme de Schwarz avec recouvrement basé sur des conditions d'interface naturelles. On présente notamment les résultats de calculs portant sur plusieurs millions d'inconnues.

Mots clefs : Equations de Maxwell - Eléments finis - Galerkin discontinu - Maillage tétraédrique - Décomposition de domaine - Algorithmes de Schwarz - Calcul parallèle.

Discontinuous Galerkin methods for the numerical resolution of the 3D frequency domain Maxwell equations

The general objective of this study is the development and the evaluation of Discontinuous Galerkin (DG) methods on unstructured tetrahedral meshes for the numerical resolution of the first order system of 3D Maxwell equations in the frequency domain. In the first part of this thesis, we formulate and analyze centered DG methods based on a P0 local approximation (i.e. finite volume or DG-P0 method) and a P1 local approximation (i.e. linear discontinuous Galerkin or DG-P1 method). The second part is devoted to the design of domain decomposition methods for the solution of the algebraic systems associated to DG methods for the discretization of the time-harmonic Maxwell equations. We first consider the system of Maxwell equations in the continuous case and study the convergence of overlapping and non-overlapping Schwarz algorithms based on a first order (natural) interface condition that corresponds to a Dirichlet condition for characteristic variables associated to incoming waves. We then conduct a convergence analysis in the discrete case corresponding to the finite volume formulation (DG-P0 method) on a quadrilateral mesh. Finally, we study optimized interface conditions in order to accelerate the convergence of the non-overlapping Schwarz algorithm. Preliminary tests in 2D illustrate the performance gains resulting from the use of optimized interface conditions. The third part of the thesis is concerned with a numerical evaluation of the DG-P0 et DG-P1 formulations on tetrahedral meshes. We make use of a series of test cases of increasing complexity dealing with diffraction problems in homogeneous and heterogeneous media. We conduct a detailed analysis of the parallel performances of an overlapping Schwarz algorithm based on a natural interface condition. We present results of numerical simulations involving several million unknowns.

Key words : Maxwell equations - Finite elements - Discontinuous Galerkin - Tetrahedral mesh - Domain decomposition - Schwarz algorithm - Parallel computing.