

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  
ADRIEN CATELLA

# SCHÉMAS D'INTÉGRATION EN TEMPS EFFICACES POUR LA RÉOLUTION NUMÉRIQUE DES ÉQUATIONS DE MAXWELL INSTATIONNAIRES PAR DES MÉTHODES GALERKIN DISCONTINUES D'ORDRE ÉLEVÉ EN MAILLAGES NON-STRUCTURÉS

Thèse dirigée par Stéphane LANTERI  
Co-dirigée par Victorita DOLEAN  
soutenue le 12 Décembre 2008

Jury :

Mme. VICTORITA DOLEAN,	maître de conférence,	co-directeur
M. LUC GIRAUD,	professeur ,	rapporteur
M. STÉPHANE LANTERI,	directeur de recherche,	directeur
M. BONIFACE NKONGA,	professeur ,	rapporteur
Mme. MURIEL SESQUES,	ingénieur de recherche,	examineur
M. DAMIEN TROMEUR-DERVOUT,	professeur ,	examineur



# Table des matières

<b>Introduction générale</b>	<b>5</b>
L'électromagnétisme numérique et ses applications . . . . .	5
Les difficultés de la modélisation numérique . . . . .	5
Résolution numérique des équations de Maxwell . . . . .	6
Objectif et plan de la thèse . . . . .	7
<b>1 Cadres mathématique et numérique</b>	<b>11</b>
1.1 Equations de Maxwell . . . . .	11
1.1.1 Formes différentielles et lois constitutives des matériaux . . . . .	13
1.1.2 Symétrisation du système de Maxwell . . . . .	14
1.1.3 Discontinuité du champ électromagnétique et conditions aux limites . . . . .	16
1.1.4 Formulation vectorielle et hyperbolicité . . . . .	18
1.1.5 Redimensionnement du système de Maxwell . . . . .	20
1.1.6 Problèmes de diffraction d'onde . . . . .	21
1.2 Méthodes numériques en domaine temporel . . . . .	22
1.2.1 Méthodes FDTD . . . . .	22
1.2.2 Méthodes FETD . . . . .	24
1.2.3 Méthodes FVTD . . . . .	25
1.2.4 Méthodes DGTD . . . . .	26
1.2.5 Brève revue des schémas d'intégration en temps . . . . .	27
<b>2 Méthodes Galerkin discontinues en domaine temporel</b>	<b>39</b>
2.1 Introduction . . . . .	39
2.2 Discrétisation en espace . . . . .	39
2.2.1 Formulation faible . . . . .	41
2.2.2 Traitement numérique des conditions aux limites . . . . .	43
2.2.3 Calcul des intégrales de bord . . . . .	44
2.2.4 Equations semi-discrétisées . . . . .	45
2.3 Discrétisation en temps . . . . .	49
2.3.1 Préliminaires . . . . .	49
2.3.2 Schéma saute-mouton d'ordre 2 . . . . .	51
2.3.3 Schéma saute-mouton d'ordre 4 . . . . .	52
2.4 Résultats numériques . . . . .	52
2.4.1 Résultats en 2D . . . . .	52
2.4.2 Résultats en 3D . . . . .	61
2.5 Conclusion . . . . .	63

<b>3</b>	<b>Méthodes GDDT implicites</b>	<b>67</b>
3.1	Introduction	67
3.2	Méthode DGTD- $\mathbb{P}_k$ implicite du second ordre	68
3.2.1	Formulation	68
3.2.2	Propriétés	69
3.2.3	Résultats numériques en 2D	81
3.3	Méthode DGTD- $\mathbb{P}_k$ implicite du quatrième ordre	104
3.3.1	Formulation	105
3.3.2	Résultats numériques en 2D	107
3.4	Conclusion	108
<b>4</b>	<b>Résolution des systèmes implicites</b>	<b>113</b>
4.1	Introduction	113
4.2	Stratégies de résolution	114
4.2.1	Le solveur direct MUMPS	114
4.2.2	La méthode GMRES	114
4.2.3	Méthode de factorisation incomplète ILUT	116
4.3	Etude numérique	117
4.3.1	Méthode DGTD- $\mathbb{P}_1$	119
4.3.2	Méthode DGTD- $\mathbb{P}_2$	125
4.4	Conclusion	130
<b>5</b>	<b>Méthode GDDT hybride explicite/implicite</b>	<b>131</b>
5.1	Formulation	132
5.2	Résultats numériques en 2D	135
5.2.1	Mode propre dans une cavité	135
5.2.2	Diffraction d'une onde plane par un cylindre diélectrique	142
5.3	Conclusion	148
<b>6</b>	<b>Résultats numériques en 3D</b>	<b>151</b>
6.1	Préambule	151
6.2	Résultats numériques	153
6.2.1	Diffraction d'une onde plane par un jet d'affaire	153
6.2.2	Propagation dans une cavité sphérique ouverte	156
6.2.3	Diffraction d'une onde plane par une géométrie de missile	162
6.3	Conclusion	162
	<b>Conclusion générale et perspectives</b>	<b>167</b>
	<b>Références bibliographiques</b>	<b>171</b>



# 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 l'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 des besoins de planification du traitement par hyperthermie<sup>1</sup> de tumeurs cancéreuses [Lin et al., 2000]-[Siauve et al., 2003], soit pour évaluer les éventuels effets nocifs de l'exposition à ces champs [Bernardi et al., 2001]-[Clatz et al., 2000].

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

En dépit des avancées significatives de ces vingt dernières années, un grand nombre d'applications de l'électromagnétisme numérique reste hors de portée de la plupart des méthodes numériques pour la modélisation des équations de Maxwell [Reitich and Tamma, 2004].

---

<sup>1</sup>L'hyperthermie est un échauffement (local) des tissus biologiques.

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 :

- la plupart des phénomènes électromagnétiques à modéliser 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 longueurs d'onde mais peuvent dépasser la centaine, conduisant à des maillages qui peuvent contenir jusqu'à dix millions de mailles 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 être particulièrement bien parallélisable.

## 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 n'ait toutes les propriétés requises, le choix étant alors 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.

La présente étude porte sur la résolution numérique des équations de Maxwell en domaine temporel. Différentes familles de méthodes numériques ont été développées pour résoudre ces équations, faisant suite à la méthode la plus ancienne proposée en 1966 par K.S. Yee [Yee, 1966]. Nous les passons en revue plus loin dans ce manuscrit. Le point de départ de notre étude consiste en une méthode d'éléments finis discontinus, connue sous le nom de méthode Galerkin discontinue, initialement introduite dans [Piperno and Fezoui, 2003]-[Fezoui et al., 2005] pour la résolution numérique des équations de Maxwell en domaine temporel. Cette méthode est formulée sur des maillages triangulaires en 2D et tétraédriques

en 3D. Les méthodes Galerkin discontinues présentent plusieurs particularités pouvant être des avantages parmi lesquels :

- elles sont naturellement adaptées à la discrétisation de fonctions discontinues et à la prise en compte d'hétérogénéités (par exemple du milieu dans le cas d'un problème de propagation d'ondes),
- elles se prêtent bien à l'utilisation de maillages non-structurés pour la discrétisation de géométries complexes, bien évidemment cette particularité n'apporte rien sur une géométrie simple.
- elles autorisent un raffinement local non-conforme du maillage de calcul et un degré d'interpolation variable en espace (en d'autres termes, elles définissent un cadre idéal pour la mise au point de méthodes auto-adaptatives),
- elles sont naturellement parallélisables.

Le caractère discontinu de l'approximation impose d'avoir recours à une formulation faible locale c'est-à-dire dont le support d'intégration est l'élément (un triangle en 2D et un tétraèdre en 3D). Une intégration par parties fait alors apparaître un terme de bord dont le calcul conduit à l'introduction d'une fonction de flux numérique (similairement à ce qui est réalisé dans les méthodes volumes finis [Cioni et al., 1997]-[Piperno et al., 2002]). Une particularité de la méthode Galerkin discontinue introduite dans [Piperno and Fezoui, 2003]-[Fezoui et al., 2005] est d'utiliser une fonction de flux numérique centrée. Lorsque celle-ci est combinée à un schéma d'intégration en temps centré de type saute-mouton du second ordre, on montre que l'on obtient une méthode Galerkin discontinue non-dissipative et stable sous une condition de type CFL. Cependant, lorsque les simulations numériques mettent en jeu des maillages non-uniformes (localement raffinés), cette condition de stabilité impose un pas de temps très pénalisant, c'est-à-dire conduisant à des temps de calcul importants voire prohibitifs pour certains problèmes tridimensionnels (Chap. 2).

## Objectif et plan de la thèse

L'objectif général de notre étude a été de mettre au point des stratégies d'intégration en temps permettant d'améliorer l'efficacité de la méthode Galerkin discontinue introduite dans [Piperno and Fezoui, 2003]-[Fezoui et al., 2005].

Nous avons tout d'abord étudié des schémas d'intégration en temps globalement implicites. En premier lieu, nous avons considéré le schéma Crank-Nicolson, car très bien étudié dans la littérature, simple à mettre en œuvre et A-stable. La A-stabilité d'un schéma est son caractère inconditionnellement stable. Cela correspond parfaitement à nos attentes en termes de stabilité car il est possible ainsi d'utiliser un pas de temps arbitrairement large. Cependant, les problèmes de propagation considérés étant purement instationnaires ou périodiques, on conçoit bien que la précision du schéma définit implicitement une limite sur le pas de temps pour que la solution approchée soit proche de celle obtenue par le schéma explicite. Dans le cas du schéma Crank-Nicolson, la précision en temps est du second ordre et la méthode Galerkin discontinue implicite résultante est non-dissipative. Il convient donc d'étudier l'erreur de dispersion de cette méthode pour mieux appréhender le pas de temps maximal utilisable en pratique.

Dans un souci d'améliorer la précision de la méthode Galerkin discontinue basée sur le schéma Crank-Nicolson, nous avons ensuite évalué la possibilité d'utiliser un schéma d'intégration en temps implicite précis au quatrième ordre. Notre premier choix s'est tourné vers les schémas de la famille Adams-Moulton. En effet, il s'agit de la généralisation aux ordres supérieurs du schéma de Crank-Nicolson. Malheureusement, comme le démontre Dahlquist [Dahlquist, 1963], les méthodes implicites à pas multiples perdent leur A-stabilité pour un ordre strictement supérieur à deux. Nous avons dès lors préféré adopter une approche alternative qui permet d'augmenter l'ordre de précision du schéma de Crank-Nicolson tout en

préservant sa A-stabilité. Pour cela, l'approche considérée est basée sur une technique de défaut corrigé [Gustafsson and Hemmingsson-Frändén, 2002].

Contrairement à un schéma d'intégration en temps explicite, les méthodes Galerkin discontinues globalement implicites nécessitent la résolution d'un ou plusieurs systèmes linéaires globaux à chaque itération en temps. Pour des problèmes bidimensionnels de taille compatible avec une station de travail, une factorisation LU creuse peut-être appliquée à la matrice de ces systèmes linéaires dont les coefficients sont indépendants du temps (pour des matériaux non-dispersifs). Du coup, cette factorisation est réalisée une fois pour toutes hors de la boucle en temps et chaque résolution de système linéaire revient à l'inversion exacte de deux systèmes triangulaires (référéncée dans les méthodes directes comme la phase de descente/remontée). Ainsi, le temps de résolution des systèmes linéaires est amorti par le gain en pas de temps lorsque les simulations numériques mettent en jeu des maillages triangulaires fortement non-uniformes. Malheureusement, on n'obtient pas un comportement aussi favorable en trois dimensions d'espace car une factorisation LU globale est tout simplement trop consommatrice en place mémoire pour une utilisation sur station de travail. Pour améliorer cette situation, deux stratégies peuvent être envisagées :

- remplacer la méthode de résolution directe par une méthode itérative préconditionnée. La problématique sous-jacente concerne alors la technique de préconditionnement qui doit être à la fois peu gourmande en mémoire et efficace des points de vue de son effet sur la convergence de la méthode itérative et du surcoût en temps de calcul ;
- remplacer le schéma globalement implicite par un schéma hybride explicite/implicite (plus précisément, un schéma localement implicite). En effet, dans bon nombre de situations, les régions raffinées n'impliquent qu'un faible pourcentage des éléments du maillage de calcul. Idéalement, on souhaiterait appliquer le schéma d'intégration implicite aux éléments d'un sous-ensemble des éléments du maillage tout en préservant un schéma explicite pour les éléments dans le complémentaire de cet ensemble. Une telle méthode Galerkin discontinue hybride explicite/implicite a récemment été proposée et partiellement étudiée par Piperno [Piperno, 2006b]-[Piperno, 2006a]. Dans notre étude, nous avons repris cette méthode et l'avons mise en œuvre en deux et trois dimensions d'espace pour aboutir dans ce dernier cas à une méthodologie numérique particulièrement bien adaptée aux simulations numériques en maillages tétraédriques localement raffinés.

Le plan de la thèse est le suivant :

- le chapitre 1 est consacré à une présentation du système des équations de Maxwell qui modélise les phénomènes électromagnétiques, ainsi qu'à une revue rapide des principales techniques existantes pour la résolution numérique de ce système d'EDPs : méthode de différences finies en domaine temporel (FDTD), méthode d'éléments finis en domaine temporel (FETD), méthode de volumes finis en domaine temporel (FVTD) et méthode Galerkin discontinus en domaine temporel (DGTD).
- Dans le chapitre 2 nous présentons la méthode Galerkin discontinue à la base de notre étude [Piperno and Fezoui, 2003]-[Fezoui et al., 2005]. Nous détaillons dans un premier temps la discrétisation en espace en explicitant la formulation faible puis le traitement numérique des conditions aux limites. Dans [Piperno and Fezoui, 2003]-[Fezoui et al., 2005] l'intégration en temps des équations semi-discrétisées est réalisée au moyen d'un schéma saute-mouton précis au second ordre. Outre cette option, nous considérons aussi l'utilisation d'un schéma saute-mouton précis au quatrième ordre. De même, dans [Piperno and Fezoui, 2003]-[Fezoui et al., 2005], la mise en œuvre de la méthode Galerkin discontinue se limitait à des interpolations locales constantes (conduisant à la méthode de volumes finis [Remaki, 2000]-[Piperno et al., 2002]) et linéaires. Dans notre étude, le degré d'interpolation est arbitraire, l'interpolation locale s'appuyant sur des fonctions de base polynomiales nodales (interpolation de Lagrange  $\mathbb{P}_p$  [Ciarlet and Lions, 1991]). Nous concluons ce chapitre par une série de résultats numériques en deux et trois dimensions d'espace visant d'une part, à illustrer les possibilités de la

méthode Galerkin discontinue  $DGTD-\mathbb{P}_p$  et, d'autre part, à motiver notre étude sur l'amélioration de l'efficacité de cette méthode.

- Dans le chapitre 3, nous développons ce qui est la partie principale de ce travail de thèse à savoir la mise au point de méthodes  $DGTD-\mathbb{P}_p$  implicites. Nous considérons dans un premier temps une méthode  $DGTD-\mathbb{P}_p$  implicite basée sur le schéma d'intégration en temps de Crank-Nicolson. La méthode  $DGTD-\mathbb{P}_p$  résultante est précise au second ordre en temps, non-dissipative et inconditionnellement stable. Nous détaillons la formulation de cette méthode en trois dimensions d'espace puis nous étudions quelques unes de ses propriétés. Nous réalisons aussi une analyse de dispersion numérique en une dimension d'espace. La méthode est ensuite évaluée en deux dimensions d'espace et en maillages triangulaires, sur la base de différents cas tests de propagation en milieux homogène et hétérogène. Nous étudions ensuite la possibilité de construire une méthode  $DGTD-\mathbb{P}_p$  implicite précise au quatrième ordre en temps en exploitant une technique de défaut corrigé. Il s'agit plus ici d'une étude de faisabilité d'une telle approche, que du développement d'une méthode  $DGTD-\mathbb{P}_p$  implicite d'intérêt pratique. En effet, une étude numérique montre que la technique employée permet effectivement d'améliorer la précision de la méthode  $DGTD-\mathbb{P}_p$  implicite basée sur le schéma d'intégration en temps de Crank-Nicolson mais au détriment d'un coût par itération en temps prohibitif.
- Dans le chapitre 5 nous étudions la possibilité de restreindre l'utilisation d'un schéma d'intégration en temps implicite à un sous-ensemble des équations semi-discrétisées. On considère pour cela une méthode  $DGTD$  hybride explicite/implicite initialement proposée par Piperno [Piperno, 2006b]-[Piperno, 2006a]. Cette méthode combine un schéma saute-mouton du second ordre et un schéma de Crank-Nicolson du second ordre pour l'intégration en temps du système d'EDO résultant de la discrétisation des équations de Maxwell par une méthode Galerkin discontinue basée sur un schéma centré pour le calcul des flux numériques aux interfaces entre cellules voisines. Une étude numérique pour des problèmes de propagation bidimensionnels montre que cette méthode offre un bon compromis entre temps de calcul et occupation mémoire comparativement aux méthodes  $DGTD-\mathbb{P}_p$  globalement explicite ou implicite.
- Le chapitre 4 est consacré à une étude de la phase de résolution des systèmes linéaires obtenus à chaque itération en temps d'une méthode  $DGTD-\mathbb{P}_p$  implicite (ou hybride explicite/implicite). Nous comparons, dans le contexte de problèmes de propagation en deux dimensions d'espace, deux stratégies de résolution : une méthode directe (factorisation LU) adaptée au cas de matrices creuses [Amestoy et al., 2000] et une méthode itérative GMRES [Saad and Schultz, 1986] préconditionnée par une technique de factorisation incomplète ILUT [Saad, 1994].
- Dans le chapitre 6 nous présentons les résultats de quelques simulations numériques tridimensionnelles en utilisant les méthodes  $DGTD-\mathbb{P}_p$  globalement implicite et hybride explicite/implicite.
- Pour finir, la section 6.3 fait un bilan de cette étude et identifie quelques pistes pour des travaux futurs.



# Chapitre 1

## Cadres mathématique et numérique

Ce chapitre est consacré à une présentation du modèle mathématique qui permet de décrire les phénomènes électromagnétiques, ainsi qu'à une revue rapide des principales techniques existantes pour la résolution numérique du système d'EDPs associé. Pour plus de détails, on pourra se référer aux deux premiers tomes de l'excellent cours de Feynman [Feynman, 1963] et [Feynman, 1964] pour un point de vue physique de l'électromagnétisme, et au premier tome de la série *Analyse mathématique et calcul numérique pour les sciences et les techniques* de Dautray et Lions [Dautray and Lions, 1987] pour les aspects plus mathématiques. Dans la présentation qui suit, on s'est aussi inspiré de l'ouvrage de Nédélec [Nedelec, 1980]

La théorie classique de l'électromagnétisme repose sur les célèbres équations de Maxwell, proposées par leur auteur il y a plus d'un siècle [Maxwell, 1873]. Les équations de Maxwell ne sont pas démontrables de manière logique ou axiomatique. Au moment de leur publication, elles étaient seulement cohérentes avec les résultats connus, en particulier d'électrostatique et de magnétostatique, mais *extrapolaient* largement ces derniers. Elles furent accueillies avec réticence par beaucoup. Ce type de conflit n'a pu être résolu que par un verdict expérimental. Ce fut Hertz qui, finalement, força l'adoption de la théorie de Maxwell en mettant en évidence l'existence et la propagation d'ondes électromagnétiques à une vitesse égale ou voisine de la lumière (les radiations lumineuses n'étant d'ailleurs qu'un cas particulier des ondes électromagnétiques), phénomène qui avait été prévu par le physicien anglais. Rapidement cependant, le succès même de la théorie de Maxwell plongea la physique moderne dans la première de ses grandes crises. Il apparut en effet que la mécanique newtonienne classique était incompatible avec l'électromagnétisme. Cette incompatibilité se manifestait en particulier au niveau de l'existence d'une vitesse *absolue* (gardant la même valeur dans tous les repères) : la vitesse  $c$  de la lumière. La constante  $c$ , conséquence des équations de Maxwell expérimentalement confirmée, contredit les formules newtoniennes de composition des vitesses. Toutes ces difficultés furent finalement résolues grâce à Einstein et sa théorie de la relativité restreinte. Une nouvelle mécanique se mettait en place, compatible avec l'électromagnétisme de Maxwell, et dont la mécanique newtonienne représente une approximation quand les mobiles considérés gardent des vitesses petites devant  $c$ .

### 1.1 Equations de Maxwell

La définition des champs électromagnétiques résulte principalement des expériences sur les interactions mutuelles des charges électriques. L'idée est de supposer que chaque charge est entourée d'un champ électrique  $\mathbf{E}$ . Si la charge  $q$  est soumise à un champ  $\mathbf{E}$ , la force  $\mathbf{F}_E$  qu'elle subit est défini par  $\mathbf{F}_E = q\mathbf{E}$ . De plus, on peut observer que le déplacement d'une charge peut être affecté par une force proportionnelle à sa vitesse. Un autre champ a donc dû être défini, l'induction magnétique  $\mathbf{B}$ , de telle sorte que cette

force s'écrit  $\mathbf{F}_B = q\mathbf{v} \times \mathbf{B}$ . Une particule soumise à la fois aux forces  $\mathbf{F}_E$  et  $\mathbf{F}_B$  est ainsi dite présente dans une région où règne à la fois des champs électrique et magnétique. Ceux-ci appliquent une force globale sur la particule  $q$  de la forme  $\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$  qui est connue sous le nom de force de Lorentz. L'interdépendance de ces deux champs ne se révèle qu'au travers des équations de Maxwell que nous allons maintenant présenter.

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

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

**La loi de Gauss électrique.** Cette loi relie le flux de l'intensité 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  $\rho$  une densité de distribution continue de charges dans le volume  $V$ . La constante  $\varepsilon$  est la permittivité électrique du milieu. Une différence entre flux entrant et sortant à travers toute surface  $A$  indique la présence de sources à l'intérieur de cette surface. Cette loi montre qu'un champ électrique peut non seulement être généré par une variation de flux magnétique mais aussi par des charges électriques

**La loi de Gauss magnétique.** Malgré toutes les interdépendances qui existent entre champs électrique et magnétique, aucune charge magnétique analogue à la charge électrique n'a pu être mise en évidence et ce en dépit de nombreuses investigations. Les champs d'induction magnétique peuvent toutefois être décrits en termes de courants et l'équivalent d'un dipôle magnétique serait une boucle de courant infiniment petite. L'absence de monopôle magnétique entraîne la nullité du flux d'induction magnétique à travers toute surface  $A$  fermée :

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

**La loi de Maxwell-Ampère.** Elle est issue dans un premier temps de la loi d'Ampère qui relie la circulation de l'induction magnétique  $\mathbf{B}$  le long d'une courbe fermée  $C$  à l'intensité du courant  $i$  dans cette boucle définie par :

$$i = \iint_A \mathbf{j} \cdot d\mathbf{S}, \quad (1.4)$$

où  $\mathbf{j}$  est la densité de courant par unité de surface ( $A$  étant toujours une surface ouverte délimitée par  $C$ ). La principale contribution de Maxwell fut de corriger cette loi puisque le mouvement de charges n'est pas la seule source de champ magnétique. L'expression correcte de cette loi, connue sous le nom de Maxwell-Ampère, s'écrit :



$$\oint_C \mathbf{B} \cdot d\mathbf{l} = \mu \iint_A \left( \varepsilon \frac{\partial \mathbf{E}}{\partial t} + \mathbf{j} \right) \cdot d\mathbf{S}, \quad (1.5)$$

où la constante  $\mu$  est la perméabilité magnétique du milieu. Un champ électrique  $\mathbf{E}$  dépendant du temps est ainsi associé à un champ d'induction magnétique  $\mathbf{B}$  même en l'absence de courants sources ( $\mathbf{j}=0$ ).

**Système de Maxwell.** L'ensemble de ces expressions intégrales forment les équations de Maxwell qui se réduisent dans le cas de la propagation d'ondes électromagnétiques dans le vide à :

$$\left\{ \begin{array}{l} \oint_C \mathbf{E} \cdot d\mathbf{l} = - \iint_A \frac{\partial \mathbf{B}}{\partial t} \cdot d\mathbf{S}, \\ \oint_C \mathbf{B} \cdot d\mathbf{l} = \varepsilon_0 \mu_0 \iint_A \frac{\partial \mathbf{E}}{\partial t} \cdot d\mathbf{S}, \\ \oiint_A \mathbf{E} \cdot d\mathbf{S} = 0, \\ \oiint_A \mathbf{B} \cdot d\mathbf{S} = 0, \end{array} \right. \quad (1.6)$$

où  $\varepsilon_0$  et  $\mu_0$  sont la permittivité et la perméabilité du vide dont les valeurs sont données par :

$$\left\{ \begin{array}{l} \varepsilon_0 = 8,85 \times 10^{-12} F/m, \\ \mu_0 = 1,26 \times 10^{-6} H/m. \end{array} \right. \quad (1.7)$$

### 1.1.1 Formes différentielles et lois constitutives des matériaux

L'application du théorème de la divergence de Gauss et du théorème de Stokes permet d'écrire les équations de Maxwell (1.6) sous la forme d'un système d'équations aux dérivées partielles :

$$\left\{ \begin{array}{l} \varepsilon \mu \frac{\partial \mathbf{E}}{\partial t} - \text{rot}(\mathbf{B}) = -\mu \mathbf{j}, \\ \frac{\partial \mathbf{B}}{\partial t} + \text{rot}(\mathbf{E}) = 0, \\ \text{div}(\mathbf{E}) = \frac{\rho}{\varepsilon}, \\ \text{div}(\mathbf{B}) = 0. \end{array} \right. \quad (1.8)$$

Le champ  $\mathbf{E}(\mathbf{x}, t)$  (en  $V/m$ ) est une fonction vectorielle définie sur  $\mathbb{R}^3 \times \mathbb{R}$  et à valeurs dans  $\mathbb{R}^3$ . Le champ d'induction magnétique  $\mathbf{B}(\mathbf{x}, t)$  (en  $T$ ) et la densité de courant  $\mathbf{j}(\mathbf{x}, t)$  (en  $A/m^2$ ) sont également des fonctions vectorielles de  $\mathbb{R}^3$  définies sur les mêmes espaces que  $\mathbf{E}$ . Seule la densité de charge  $e$  (en  $C/m^3$ ) est une fonction scalaire à valeurs réelles. Jusqu'ici, la permittivité électrique  $\varepsilon$  et la perméabilité magnétique  $\mu$  du milieu ont été considérées comme constantes.

Pour des matériaux homogènes linéaires, isotropes, diélectriques, les champs  $\mathbf{D}$  et  $\mathbf{E}$  sont proportionnels :

$$\mathbf{D} = \varepsilon_0 \mathbf{E}.$$

De façon analogue, on peut introduire un vecteur de polarisation magnétique  $\mathbf{M}$  et définir une nouvelle quantité, le champ magnétique  $\mathbf{H}$  :

$$\mathbf{H} = \frac{\mathbf{B}}{\mu_0} - \mathbf{M}.$$

Pour des matériaux simples, les vecteurs  $\mathbf{B}$  et  $\mathbf{H}$  sont parallèles et proportionnels :

$$\mathbf{B} = \mu_0 \mathbf{H}.$$

Pour des matériaux hétérogènes isotropes, la permittivité et la perméabilité dépendent de la variable d'espace et sont alors des fonctions scalaires  $\varepsilon(\mathbf{x})$  et  $\mu(\mathbf{x})$ .

Outre les lois de comportement de base des matériaux, on ajoute en toute généralité au système de Maxwell d'autres lois constitutives. En particulier, dans les conducteurs à température constante, le flux de charge est conditionné par l'intensité du champ électrique via la loi d'Ohm  $\mathbf{j} = \sigma \mathbf{E}$  où  $\sigma$  est la conductivité du milieu en Siemens par mètres ( $S/m$ ). Il est facile de vérifier à partir de la loi d'Ohm et de la loi de conservation de la charge définie par :

$$\frac{\partial \rho}{\partial t} + \text{div}(\mathbf{j}) = 0, \quad (1.9)$$

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

$$\rho(x, t) = \rho(x, 0) \exp\left(\frac{-t\sigma}{\varepsilon}\right).$$

Autrement dit, le temps  $T = \frac{\varepsilon}{\sigma}$  est une constante de temps (temps de relaxation). C'est-à-dire le temps au bout duquel la densité de charge aura diminué de 36.8% de sa valeur initiale. Après un temps égal à  $3T$ , la densité ne vaudra plus 5% de sa valeur initiale. 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.*

### 1.1.2 Symétrisation du système de Maxwell

L'introduction du champ de déplacement électrique permet d'écrire les équations de Maxwell sous une forme qui possède des propriétés intéressantes d'un point de vue mathématique. La formulation la plus souvent rencontrée dans la littérature fait intervenir les quatre champs de vecteur 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. L'introduction d'une densité volumique de courant magnétique  $\mathbf{m}$  et d'une densité de charge magnétique  $\tilde{\rho}$  permet d'obtenir une forme totalement symétrique du système de Maxwell qui s'écrit alors :

$$\begin{cases} \frac{\partial \mathbf{D}}{\partial t} - \text{rot}(\mathbf{H}) = -\mathbf{j}, \\ \frac{\partial \mathbf{B}}{\partial t} + \text{rot}(\mathbf{E}) = -\mathbf{m}, \\ \text{div}(\mathbf{D}) = \rho, \\ \text{div}(\mathbf{B}) = \tilde{\rho}. \end{cases} \quad (1.10)$$

On ajoute également à ce système symétrisé une loi de conservation de la charge magnétique analogue à (1.9) :

$$\frac{\partial \tilde{\rho}}{\partial t} + \text{div}(\mathbf{m}) = 0. \quad (1.11)$$

Cette symétrisation n'a de sens que d'un point de vue mathématique puisqu'il n'existe pas réellement de courant et de source magnétique. Il est cependant licite de parler de courants magnétiques équivalents. Une boucle de courant électrique génère en effet un champ qui ressemble à celui qu'induirait un dipôle magnétique. On peut alors considérer qu'une boucle de courant infiniment petite est équivalente à un dipôle magnétique fictif. Une série de boucles de courants (un solénoïde, voir la Fig. 1.1 par exemple) génère également un champ similaire à celui induit par un courant magnétique fictif. On met alors en évidence les corrélations entre champs et sources électriques et champs et sources magnétiques. En effet, la symétrie des équations de Maxwell ainsi formulées implique que toute solution d'un problème électromagnétique correspond à celle d'un autre problème via un échange approprié entre grandeurs électriques et magnétiques. Cet échange qui laisse les équations de Maxwell (1.10) invariantes n'est en général pas unique. Un échange possible est donné par :

$$\mathbf{E} \rightarrow -\mathbf{H}, \quad \mathbf{H} \rightarrow \mathbf{E}, \quad \mathbf{j} \rightarrow -\mathbf{m}, \quad \mathbf{m} \rightarrow -\mathbf{j}, \quad \rho \rightarrow -\tilde{\rho}, \quad \tilde{\rho} \rightarrow \rho.$$

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

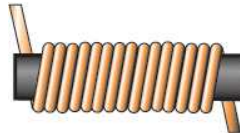


FIG. 1.1 – Un solénoïde

En l'absence d'une densité de charge magnétique et en utilisant la loi d'Ohm, nous pouvons ré-écrire (1.10) sous une forme plus exploitable, car n'utilisant plus les inductions électrique et magnétique :

$$\left\{ \begin{array}{l} \varepsilon \frac{\partial \mathbf{E}}{\partial t} - \text{rot}(\mathbf{H}) + \sigma \mathbf{E} = 0, \\ \mu \frac{\partial \mathbf{H}}{\partial t} + \text{rot}(\mathbf{E}) = 0, \\ \text{div}(\varepsilon \mathbf{E}) - \rho = 0, \\ \text{div}(\mu \mathbf{H}) = 0. \end{array} \right. \quad (1.12)$$

### 1.1.3 Discontinuité du champ électromagnétique et conditions aux limites

#### 1.1.3.1 Equations au sens des distributions

L'écriture classique des équations de Maxwell suppose implicitement que sources et champs ont une régularité qui assure un sens aux diverses dérivations. Or, dans les applications, cette régularité n'est pas toujours vérifiée. L'écriture des équations de Maxwell au sens des distributions permet alors de traiter ces singularités.

Soit une distribution vectorielle  $T_{\mathbf{A}}$  de  $\mathbb{R}^3$  associée au champ de vecteur  $\mathbf{A}$  dérivable au sens des fonctions vectorielles dans le complémentaire d'une surface de discontinuité  $S$  orientée et de normale  $\mathbf{n}$ . On rappelle que la divergence et le rotationnel d'une distribution vectorielle  $T_{\mathbf{A}}$  s'écrivent :

$$\left\{ \begin{array}{l} \text{div}(T_{\mathbf{A}}) = \text{div}(\mathbf{A}) + \mathbf{n} \cdot [\mathbf{A}]_S \delta_S, \\ \text{rot}(T_{\mathbf{A}}) = \text{rot}(\mathbf{A}) + \mathbf{n} \times [\mathbf{A}]_S \delta_S, \end{array} \right. \quad (1.13)$$

où  $\delta_S$  est la distribution de Dirac sur la surface  $S$  et  $[\mathbf{A}]_S$  le saut de  $\mathbf{A}$  à travers la surface  $S$ . Le terme  $\text{div}(\mathbf{A})$  représente l'opérateur divergence au sens des fonctions appliqué à  $\mathbf{A}$ , alors que  $\text{rot}(\mathbf{A})$  désigne quant à lui, l'opérateur rotationnel au sens des fonctions associé à  $\mathbf{A}$ . On considère maintenant les distributions vectorielles et scalaires associées aux grandeurs électromagnétiques intervenant dans les équations de Maxwell (1.10). On suppose que ces grandeurs sont dérivables dans le complémentaire de la surface de discontinuité  $S$ . De plus, la distribution scalaire  $T_\rho$  se décompose en la somme d'une densité de charge volumique  $\rho$  et d'une charge superficielle notée  $\rho_S$  ( $T_\rho = \rho + \rho_S \delta_S$ ). On peut en toute généralité y ajouter des charges ponctuelles mais nous ne les prendrons pas en compte ici. Quant aux distributions  $T_{\tilde{\rho}}$ ,  $T_{\tilde{\mathbf{j}}}$  et  $T_{\tilde{\mathbf{m}}}$ , elles se décomposent de façon analogue. On peut alors écrire, à partir de (1.13), les équations de Maxwell au sens des distributions :

$$\left\{ \begin{array}{l} \frac{\partial \mathbf{D}}{\partial t} - \text{rot}(\mathbf{H}) - \mathbf{n} \times [\mathbf{H}]_S \delta_S = -\mathbf{j} - \mathbf{j}_S \delta_S, \\ \frac{\partial \mathbf{B}}{\partial t} + \text{rot}(\mathbf{E}) + \mathbf{n} \times [\mathbf{E}]_S \delta_S = -\mathbf{m} - \mathbf{m}_S \delta_S, \\ \text{div}(\mathbf{D}) + \mathbf{n} \cdot [\mathbf{D}]_S \delta_S = \rho + \rho_S \delta_S, \\ \text{div}(\mathbf{B}) + \mathbf{n} \cdot [\mathbf{B}]_S \delta_S = \tilde{\rho} + \tilde{\rho}_S \delta_S. \end{array} \right. \quad (1.14)$$

On retrouve par identification les équations de Maxwell au sens des fonctions (1.10) d'une part, et les conditions aux limites sur la surface  $S$  d'autre part :

$$\begin{cases} \mathbf{n} \times [\mathbf{H}]_S = \mathbf{j}_S, \\ \mathbf{n} \times [\mathbf{E}]_S = -\mathbf{m}_S, \\ \mathbf{n} \cdot [\mathbf{D}]_S = \rho_S, \\ \mathbf{n} \cdot [\mathbf{B}]_S = \tilde{\rho}_S. \end{cases} \quad (1.15)$$

### 1.1.3.2 Conditions de transmission

Nous venons de décrire les conditions d'interface à travers une surface de discontinuité qui peut éventuellement porter des sources de courants (1.15). En l'absence de courants, les composantes tangentielles des champs électrique et magnétique sont continues à travers toute surface d'après les relations :

$$\begin{cases} \mathbf{n} \times [\mathbf{H}]_S = 0, \\ \mathbf{n} \times [\mathbf{E}]_S = 0. \end{cases} \quad (1.16)$$

On remarquera que les relations (1.16) sont toujours vérifiées y compris lorsque l'on est en présence de matériaux différents d'indices quelconques (i.e. quelles que soient les variations de la permittivité et de la perméabilité du milieu). En revanche, les deux autres relations obtenues à l'aide des lois de Gauss montrent que, même en l'absence de charges, les composantes normales des champs électrique et magnétique ne sont pas en général continues :

$$\begin{cases} \mathbf{n} \cdot [\mu\mathbf{H}]_S = 0, \\ \mathbf{n} \cdot [\varepsilon\mathbf{E}]_S = 0. \end{cases} \quad (1.17)$$

Ce type de relation s'applique également pour la définition des conditions aux limites, notamment pour les bords métalliques.

### 1.1.3.3 Conditions aux limites pour une frontière métallique

Nous assimilerons tout au long de cette étude une paroi métallique au modèle idéal et fictif du conducteur parfait. Nous considérons ici que les courants fictifs  $\mathbf{m}_S$  et les charges magnétiques fictives  $\tilde{\rho}_S$  sont nuls. Physiquement, il n'existe pas de courant surfacique  $\mathbf{j}_S$  à l'interface  $S$  d'un diélectrique et d'un métal. Le champ électromagnétique décroît très rapidement à l'intérieur du matériau. Ce phénomène est classiquement appelé effet de peau mais il n'est généralement pas pris en compte dans les modélisations numériques car le modèle de conducteur parfait est très simple à mettre en œuvre et assez précis pour un grand ensemble de dispositifs (voir la remarque 1). Nous considérons donc que le champ électromagnétique est nul à l'intérieur du métal. Soient  $\mathbf{E}_S$  et  $\mathbf{H}_S$  les champs électrique et magnétique sur la surface extérieure du métal, les relations de saut (1.15) deviennent :

$$\begin{cases} \mathbf{n} \times [\mathbf{H}]_S = \mathbf{j}_S, \\ \mathbf{n} \times [\mathbf{E}]_S = 0, \\ \mathbf{n} \cdot [\mathbf{E}]_S = \frac{\rho_S}{\varepsilon}, \\ \mathbf{n} \cdot [\mathbf{H}]_S = 0. \end{cases} \quad (1.18)$$

**Remarque 3** Les courants surfaciques  $\mathbf{j}_S$  et les charges surfaciques  $\rho_S$  ne sont pas des sources appliquées connues, mais leur présence est rendue nécessaire par le modèle de conducteur parfait.

On en déduit en particulier qu'à la surface d'un conducteur parfait le champ électrique est normal et le champ magnétique tangent. D'autre part, le champ magnétique peut se mesurer directement à partir de la densité de courants superficiels.

#### 1.1.3.4 Conditions aux limites pour une frontière absorbante

Les problèmes de propagation d'onde sont souvent posés en domaines non-bornés. Une des questions cruciales pour leur résolution numérique est de savoir borner artificiellement le domaine de calcul. Ce sujet est inépuisable et alimente depuis l'apparition des techniques de simulation numérique un nombre conséquent de recherches. Une approche possible consiste à imposer sur la frontière artificielle une condition limite qui évite autant que faire se peut la réflexion d'ondes dans le domaine de calcul. On parle alors de condition absorbante. Engquist et Majda [Engquist and Majda, 1977] ont proposé une famille de conditions absorbantes basée sur l'approximation de l'équation des ondes (on peut aussi se référer à Givoli [Givoli, 1991] pour la construction générale de ce type de condition). Ces conditions ont été proposées pour les ondes acoustiques et appliquées ensuite par Mur [Mur, 1981] aux problèmes électromagnétiques en 2D et 3D. Nous utilisons dans cette étude la condition de Silver-Müller qui s'écrit :

$$\mathbf{n} \times \mathbf{E} + \sqrt{\frac{\mu_0}{\varepsilon_0}} \mathbf{n} \times (\mathbf{n} \times \mathbf{H}) = \mathbf{n} \times \mathbf{E}^i + \sqrt{\frac{\mu_0}{\varepsilon_0}} \mathbf{n} \times (\mathbf{n} \times \mathbf{H}^i), \quad (1.19)$$

où  ${}^t(\mathbf{E}^i, \mathbf{H}^i)$  est un champ incident imposé. Nous utiliserons l'équation (1.19) dans la suite de la thèse lorsque nous ferons référence aux conditions absorbantes.

#### 1.1.4 Formulation vectorielle et hyperbolicité

Le système de Maxwell en fonction des quatre champs de vecteur  $\mathbf{E}$ ,  $\mathbf{H}$ ,  $\mathbf{D}$  et  $\mathbf{B}$  donné par les deux premières équations de (1.10) est dit écrit sous forme conservative. Ce type de formulation est très intéressant car de nombreux schémas conservatifs ayant prouvé leur efficacité et leur robustesse (dans le domaine de la mécanique des fluides notamment) ont été construits pour la résolution numérique de divers systèmes de loi de conservation. Afin de résoudre numériquement ce système, nous devons prendre en compte les lois constitutives des matériaux :

$$\mathbf{D} = \varepsilon(\mathbf{x})\mathbf{E} \quad \text{et} \quad \mathbf{B} = \mu(\mathbf{x})\mathbf{H}. \quad (1.20)$$

Les fonctions  $\varepsilon$  et  $\mu$  de la variable d'espace  $\mathbf{x}$  sont ici supposées scalaires mais peuvent être en toute généralité des tenseurs. Nous rappelons que ce type de loi de comportement est adapté aux matériaux linéaires, isotropes et hétérogènes. Dans la suite, nous avons souvent omis la dépendance en fonction de la variable d'espace de la permittivité et de la perméabilité afin d'alléger l'écriture.

On adjoint au système (1.10) les conditions aux limites introduites aux sous-sections 1.1.3.3 et 1.1.3.4, ainsi que les conditions initiales :

$$\mathbf{E}_0 = \mathbf{E}(\mathbf{x}, 0) \quad \text{et} \quad \mathbf{H}_0 = \mathbf{H}(\mathbf{x}, 0). \quad (1.21)$$

Notons que les lois de Gauss électrique et magnétique  $\text{div}(\mathbf{B}) = \tilde{\rho}$  et  $\text{div}(\mathbf{D}) = \rho$  sont redondantes dans le modèle continu (1.10) pour des données initiales (1.21) vérifiant ces contraintes et en considérant de plus vérifiées les équations de conservation de la charge (1.9) et (1.11).

On peut ainsi ne considérer que les deux premières équations de (1.10) qui s'écrivent alors pour la formulation conservative en fonction du champ électromagnétique  ${}^t(\mathbf{B}, \mathbf{D})$  :

$$\begin{cases} \frac{\partial \mathbf{B}}{\partial t} + \text{rot} \left( \frac{\mathbf{D}}{\varepsilon} \right) = -\mathbf{m}, \\ \frac{\partial \mathbf{D}}{\partial t} - \text{rot} \left( \frac{\mathbf{B}}{\mu} \right) = -\mathbf{j}. \end{cases} \quad (1.22)$$

Une écriture des équations de Maxwell (1.22) sous une forme vectorielle conduit à :

$$\mathbf{Q}_t + \mathbf{F}_1(\mathbf{Q})_x + \mathbf{F}_2(\mathbf{Q})_y + \mathbf{F}_3(\mathbf{Q})_z = -\mathbf{J}(\mathbf{Q}), \quad (1.23)$$

avec :

$$\mathbf{Q} = \begin{pmatrix} B_x \\ B_y \\ B_z \\ D_x \\ D_y \\ D_z \end{pmatrix}, \quad \mathbf{F}_1(\mathbf{Q}) = \begin{pmatrix} 0 \\ -D_z/\varepsilon \\ D_y/\varepsilon \\ 0 \\ B_z/\mu \\ -B_y/\mu \end{pmatrix}, \quad \mathbf{F}_2(\mathbf{Q}) = \begin{pmatrix} D_z/\varepsilon \\ 0 \\ -D_x/\varepsilon \\ -B_z/\mu \\ 0 \\ B_x/\mu \end{pmatrix},$$

$$\mathbf{F}_3(\mathbf{Q}) = \begin{pmatrix} -D_y/\varepsilon \\ D_x/\varepsilon \\ 0 \\ B_y/\mu \\ -B_x/\mu \\ 0 \end{pmatrix}, \quad \mathbf{J}(\mathbf{Q}) = \begin{pmatrix} m_x \\ m_y \\ m_z \\ j_x \\ j_y \\ j_z \end{pmatrix},$$

soit sous forme condensée :

$$\mathbf{Q}_t + \nabla \cdot \mathbb{F}(\mathbf{Q}) = -\mathbf{J}(\mathbf{Q}), \quad (1.24)$$

avec  $\mathbb{F}(\mathbf{Q}) = {}^t(\mathbf{F}_1(\mathbf{Q}), \mathbf{F}_2(\mathbf{Q}), \mathbf{F}_3(\mathbf{Q}))$ .

**Remarque 4** *En toute généralité, le courant  $\mathbf{J}$  peut être à la fois une donnée du problème (un courant appliqué connu) et une fonction du champ inconnu  $\mathbf{Q}$  lorsque l'on considère la loi d'Ohm par exemple. On écrit dans ce cas  $\mathbf{J}$  comme la somme du courant ohmique et du courant appliqué.*

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. Soit une combinaison linéaire des flux du système (1.23) de la forme :

$$\mathcal{F}(\mathbf{Q}, \mathbf{n}) = \eta \cdot \mathbb{F}(\mathbf{Q}),$$

où  $\mathbf{n} = {}^t(n_1, n_2, n_3)$  est un vecteur non nul de  $\mathbb{R}^3$ . La matrice jacobienne  $\mathcal{A}$  associée à  $\mathcal{F}$  est alors définie par :

$$\begin{cases} \mathcal{A}(\mathbf{n}) = \mathbf{n} \cdot \mathbb{F}'(\mathbf{Q}) = n_1 \mathcal{A}_1 + n_2 \mathcal{A}_2 + n_3 \mathcal{A}_3, \\ \mathcal{A}_i = \frac{\partial}{\partial \mathbf{Q}} \mathbf{F}_i(\mathbf{Q}) \text{ pour } i = \{1, 2, 3\}, \end{cases}$$

et s'écrit

$$\mathcal{A}(\mathbf{n}) = \begin{pmatrix} 0 & 0 & 0 & 0 & -n_3/\varepsilon & n_2/\varepsilon \\ 0 & 0 & 0 & n_3/\varepsilon & 0 & -n_1/\varepsilon \\ 0 & 0 & 0 & -n_2/\varepsilon & n_1/\varepsilon & 0 \\ 0 & n_3/\mu & -n_2/\mu & 0 & 0 & 0 \\ -n_3/\mu & 0 & n_1/\mu & 0 & 0 & 0 \\ n_2/\mu & -n_1/\mu & 0 & 0 & 0 & 0 \end{pmatrix}. \quad (1.25)$$

**Définition 1** Le système (1.23) de  $\mathbb{R}^6$  est dit hyperbolique si et seulement si la matrice jacobienne  $\mathcal{A}$  associée est diagonalisable dans  $\mathbb{R}$  pour tout vecteur  $\eta$  non nul de  $\mathbb{R}^3$  et pour tout vecteur  $\mathbf{Q}$  de  $\mathbb{R}^6$  (voir notamment [Serre, 2000]).

La matrice  $\mathcal{A}$  n'admet que des valeurs propres réelles et une base de  $\mathbb{R}^6$  constituée de vecteurs propres de  $\mathcal{A}$  peut être construite. Donc,  $\mathcal{A}$  est diagonalisable et le système de Maxwell est donc hyperbolique. En particulier, la matrice  $\mathcal{A}$  a trois valeurs propres réelles de multiplicité double qui sont données par :

$$\begin{cases} \lambda_1 = c\|\mathbf{n}\|, \\ \lambda_2 = -c\|\mathbf{n}\|, \\ \lambda_3 = 0, \end{cases} \quad (1.26)$$

où  $c = \frac{1}{\sqrt{\varepsilon\mu}}$  désigne la vitesse de la lumière dans le milieu de propagation.

### 1.1.5 Redimensionnement du système de Maxwell

Notons tout d'abord qu'il est classique d'exprimer la permittivité et la perméabilité des matériaux en fonction de celles du vide :

$$\varepsilon(\mathbf{x}) = \varepsilon_0\varepsilon_r(\mathbf{x}) \quad \text{et} \quad \mu(\mathbf{x}) = \mu_0\mu_r(\mathbf{x}).$$

On obtient alors des quantités relatives  $\varepsilon_r$  et  $\mu_r$  sans dimension. La vitesse relative de la lumière est maintenant notée  $c_r = \frac{1}{\sqrt{\varepsilon_r\mu_r}}$ . En particulier, on a  $\varepsilon_r = \mu_r = c_r = 1$  dans le vide.

Ensuite, en notant  $\tau = c_0t$  la nouvelle variable associée au temps, on redimensionne le système de Maxwell de telle sorte que les composantes du champ électromagnétique soient exprimées dans la même unité, plus précisément en volts par mètre (V/m). De plus, le système de Maxwell résultant est identique dans sa forme au système conservatif initial (1.23) puisqu'il s'écrit maintenant :

$$\mathbf{Q}_\tau + \mathbf{F}_1(\mathbf{Q})_x + \mathbf{F}_2(\mathbf{Q})_y + \mathbf{F}_3(\mathbf{Q})_z = -\mathbf{J}(\mathbf{Q}), \quad (1.27)$$



avec cette fois :

$$\left\{ \begin{array}{l} \mathbf{Q} = {}^t(c_0 \mathbf{B}, \mathbf{D}/\varepsilon_0) \\ \quad = {}^t(Q_1, Q_2, Q_3, Q_4, Q_5, Q_6), \\ \mathbf{F}_1(\mathbf{Q}) = {}^t(0, -Q_6/\varepsilon_r, Q_5/\varepsilon_r, 0, Q_3/\mu_r, -Q_2/\mu_r), \\ \mathbf{F}_2(\mathbf{Q}) = {}^t(Q_6/\varepsilon_r, 0, -Q_4/\varepsilon_r, -Q_3/\mu_r, 0, Q_1/\mu_r), \\ \mathbf{F}_3(\mathbf{Q}) = {}^t(-Q_5/\varepsilon_r, Q_4/\varepsilon_r, 0, Q_2/\mu_r, -Q_1/\mu_r, 0) \\ \mathbf{J}(\mathbf{Q}) = {}^t(J_1, J_2, J_3, J_4, J_5, J_6), \\ \quad = {}^t(m_x, m_y, m_z, j_x/\varepsilon_0 c_0, j_y/\varepsilon_0 c_0, j_z/\varepsilon_0 c_0). \end{array} \right.$$

D'autre part, toutes les relations obtenues pour l'hyperbolicité du système de Maxwell restent valables puisque le changement de variables est linéaire et les nouvelles valeurs propres du système sont alors exprimées en fonction de quantités relatives.

### 1.1.6 Problèmes de diffraction d'onde

Un problème classique en électromagnétisme numérique est la simulation de la diffraction d'une onde par un obstacle dont l'une des principales applications est le calcul de surface équivalente radar (SER). On considère un obstacle borné de  $\mathbb{R}^p$  ( $p = 2, 3$ ) de frontière  $\Sigma$ , parfaitement conducteur. Soit une onde électromagnétique incidente se propageant dans un milieu extérieur  $\Omega$  homogène non conducteur et arrivant sur l'obstacle. Cette onde incidente (indiquée  $i$ ) est une donnée du problème et est connue analytiquement. La présence de l'obstacle induit une perturbation de l'onde incidente qui se traduit par la formation d'une onde diffractée (indiquée  $d$ ). On cherche alors à déterminer le champ diffracté  ${}^t(\mathbf{E}^d, \mathbf{H}^d)$ , les équations de Maxwell (1.27) étant vérifiées par le champ total défini par :

$$\left\{ \begin{array}{l} \mathbf{E} = \mathbf{E}^d + \mathbf{E}^i, \\ \mathbf{H} = \mathbf{H}^d + \mathbf{H}^i. \end{array} \right.$$

Une écriture des équations de Maxwell en champ diffracté est généralement mieux adaptée à ce type de problème car le calcul d'une surface équivalente radar, par exemple, ne requiert que la connaissance du champ diffracté. Par ailleurs, les conditions aux limites absorbantes exprimées en termes de champ diffracté génèrent moins de réflexion parasites car celui-ci est généralement de plus faible amplitude que le champ total. La linéarité du système de Maxwell permet de ré-écrire les équations (1.27) en fonction du champ diffracté :

$$\mathbf{Q}_\tau^d + \mathbf{F}_1(\mathbf{Q}^d)_x + \mathbf{F}_2(\mathbf{Q}^d)_y + \mathbf{F}_3(\mathbf{Q}^d)_z = \mathbf{S}^i(\mathbf{Q}^i) - \mathbf{J}(\mathbf{Q}), \quad (1.28)$$

avec  $\mathbf{S}^i = -\mathbf{Q}_\tau^i - \mathbf{F}_1(\mathbf{Q}^i)_x - \mathbf{F}_2(\mathbf{Q}^i)_y - \mathbf{F}_3(\mathbf{Q}^i)_z$  où  $\mathbf{Q}^i$  est le vecteur associé au champ incident supposé connu. Les conditions aux limites sur un métal parfaitement conducteur s'écrivent alors :

$$\mathbf{n} \times \mathbf{E}^d = -\mathbf{n} \times \mathbf{E}^i,$$

alors que sur la frontière artificielle du domaine on a simplement :

$$\mathbf{n} \times \mathbf{E}^d + \sqrt{\frac{\mu_0}{\varepsilon_0}} \mathbf{n} \times (\mathbf{n} \times \mathbf{H}^d) = 0.$$

**Remarque 5** *L'onde incidente est toujours supposée être solution des équations de Maxwell dans le vide. Le second membre  $S^i$  issu des termes incidents est donc nul d'après nos hypothèses dans le vide.*

## 1.2 Méthodes numériques en domaine temporel

Plusieurs méthodes numériques permettent de résoudre le système des équations de Maxwell en domaine temporel :

- la méthode des Différences Finies en Domaine Temporel (FDTD),
- la méthode des Éléments Finis en Domaine Temporel (FETD),
- la méthode des Volumes Finis en Domaine Temporel (FVTD),
- la méthode Galerkin Discontinue en Domaine Temporel (DGTD),
- la méthode des équations intégrales en temps.

Nous rappelons brièvement ici l'histoire et les principales caractéristiques des méthodes FDTD, FETD, FVTD et DGTD.

### 1.2.1 Méthodes FDTD

La méthode des différences finies en domaine temporel est la plus simple conceptuellement et la plus aisée à mettre en oeuvre. Cette méthode a été introduite pour la première fois par Yee dans les années soixante [Yee, 1966]. Dans la formulation de base de la méthode FDTD, l'espace est divisé en parallélépipèdes rectangles identiques, et les composantes des champs électrique et magnétique sont calculées en un nombre fini de points. Des développements de Taylor permettent alors de discrétiser les dérivées partielles par rapport au temps et aux variables d'espace qui sont ici approchées par des différences centrées. Le système de Maxwell (avec  $\mathbf{J} = 0$ ) peut s'écrire sous la forme d'un système d'équations scalaires :

$$-\mu \frac{\partial H_x}{\partial t} = \frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z}, \quad (1.29a)$$

$$-\mu \frac{\partial H_y}{\partial t} = \frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x}, \quad (1.29b)$$

$$-\mu \frac{\partial H_z}{\partial t} = \frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y}, \quad (1.29c)$$

$$\varepsilon \frac{\partial E_x}{\partial t} = \frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z}, \quad (1.29d)$$

$$\varepsilon \frac{\partial E_y}{\partial t} = \frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x}, \quad (1.29e)$$

$$\varepsilon \frac{\partial E_z}{\partial t} = \frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y}. \quad (1.29f)$$

La méthode FDTD repose sur la localisation des composantes de  $\mathbf{E}$  et  $\mathbf{H}$  illustrée sur la Fig. 1.2. Un schéma aux différences finies centrées pour (1.29a) s'écrit :

$$\frac{B_x^{n+1/2} \left( i, j + \frac{1}{2}, k + \frac{1}{2} \right) - B_x^{n-1/2} \left( i, j + \frac{1}{2}, k + \frac{1}{2} \right)}{\Delta t} = \frac{E_y^n \left( i, j + \frac{1}{2}, k + 1 \right) - E_y^n \left( i, j + \frac{1}{2}, k \right)}{\Delta z} - \frac{E_z^n \left( i, j + 1, k + \frac{1}{2} \right) - E_z^n \left( i, j, k + \frac{1}{2} \right)}{\Delta y}$$

Les équations aux différences finies pour (1.29b)-(1.29c) sont construites de manière similaire. Pour l'équation (1.29d), nous avons :

$$\frac{D_x^n \left( i + \frac{1}{2}, j, k \right) - D_x^{n-1} \left( i + \frac{1}{2}, j, k \right)}{\Delta t} = \frac{H_z^{n-1/2} \left( i + \frac{1}{2}, j + \frac{1}{2}, k \right) - H_z^{n-1/2} \left( i + \frac{1}{2}, j - \frac{1}{2}, k \right)}{\Delta y} - \frac{H_y^{n-1/2} \left( i + \frac{1}{2}, j, k + \frac{1}{2} \right) - H_y^{n-1/2} \left( i + \frac{1}{2}, j, k - \frac{1}{2} \right)}{\Delta z}$$

Les équations aux différences finies pour (1.29e)-(1.29f) sont construites de manière similaire.

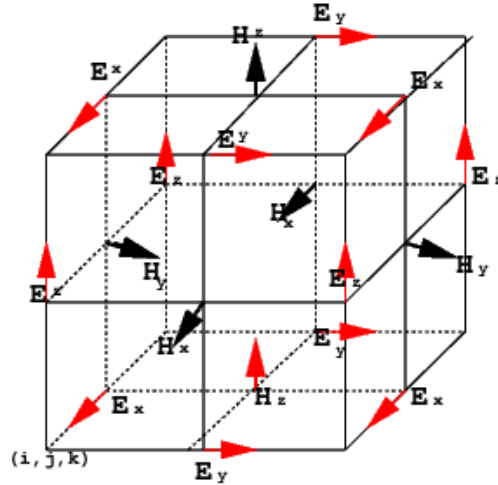


FIG. 1.2 – Cellule de Yee

La méthode FDTD résultante est précise au second ordre en espace et en temps, non-dissipative et stable sous la condition de CFL :

$$c\Delta t \leq \sqrt{(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2}$$

L'utilisation d'un maillage en parallélépipèdes rectangles est à la base de la simplicité de la méthode et de son efficacité. Néanmoins, c'est aussi la source de sa principale limitation. En effet, il est clair que ce type de discrétisation n'autorise pas une représentation fidèle de surfaces courbes (interfaces entre matériaux, surface parfaitement conductrice) telles que celles rencontrées dans la plupart des problèmes d'intérêt industriel. Des diffractions électromagnétiques apparaissent alors et dégradent la qualité de la solution. Un raffinement du maillage permet de diminuer ces perturbations mais au prix d'un accroissement important des ressources informatiques nécessaires au calcul, et du temps de calcul. Ainsi, dans le cas d'un raffinement global du maillage, si les pas de discrétisation en espace sont divisés par 2, alors le nombre de degrés de liberté pour le champ électromagnétique est multiplié en 3D par 8 ; de plus, pour vérifier le critère de stabilité, le pas de temps est lui aussi divisé par 2, et le temps de calcul en 3D se trouve alors multiplié par 16.

Plusieurs méthodes ont été proposées pour résoudre ce problème. La grille cartésienne peut être remplacé par un maillage en parallélépipèdes rectangles de tailles variables [Sheen et al., 1990], ou un maillage en hexaèdre à faces non parallèles [Lee et al., 1992]. Une autre approche consiste à conserver la structure du schéma de Yee et à lui apporter des modifications locales. La technique des intégrales de contour repose sur cette idée et a été développée pour la modélisation d'obstacles parfaitement conducteurs en deux dimensions [Jurgens et al., 1987] et trois dimensions d'espace [Jurgens and Taflove, 1993]. Enfin, des techniques de raffinement local du maillage donnent de bons résultats en 2D, mais présentent souvent des problèmes de stabilité en 3D [Ma et al., 1997]. Liu et Sarris [Liu and Sarris, 2005]-[Liu and Sarris, 2006] ont développé une méthode FDTD à fenêtre mobile (*Moving Window FDTD*). Cette méthode permet de prendre en compte la nature localisée des ondes électromagnétiques pendant leur propagation. Pour cela, les auteurs utilisent un maillage raffiné récursivement qui forme une structure d'arbre. Ainsi, lorsque l'on souhaite une meilleure précision, on utilise les sous-maillages les plus raffinés là où se trouve les pics d'énergie. Il est utile de noter qu'il existe des méthodes FDTD utilisant des éléments curvilignes permettant de mieux prendre en compte les géométries complexes. Malheureusement, l'utilisation de ce type d'élément se fait au détriment de la performance de la méthode.

### 1.2.2 Méthodes FETD

La méthode des éléments finis en domaine temporel repose sur la formulation variationnelle des équations de Maxwell du premier ordre en temps ou du second ordre en temps. Dans ce qui suit, nous considérons le cas de la formulation du second ordre en champ électrique :

$$\varepsilon \frac{\partial^2 \mathbf{E}}{\partial t^2} + \nabla \times \frac{1}{\mu} \nabla \times \mathbf{E} = 0.$$

Nous supposons aussi que la frontière du domaine de calcul est seulement de type métallique. Soit  $V$  l'espace fonctionnel naturel pour le champ électrique. La formulation variationnelle du problème consiste à définir le sous-espace  $V_0$  de  $V$  des champs électriques vérifiant la condition à la limite en posant :

$$V_0 = \{\mathbf{E} \in V, \mathbf{E} \times \mathbf{n}|_{\Gamma} = 0\},$$

et à résoudre le problème suivant :

$$\left\{ \begin{array}{l} \text{trouver } \mathbf{E} \text{ tel que } \forall \mathbf{E}^* \in V_0, \\ \iint_{\Omega} \varepsilon \mu \frac{\partial^2 \mathbf{E}}{\partial t^2} \cdot \mathbf{E}^* dV + \iint_{\Omega} \text{rot}(\mathbf{E}) \cdot \text{rot}(\mathbf{E}^*) dV = 0, \\ \mathbf{E}(t=0) = \mathbf{E}_0 \quad \text{et} \quad \frac{\partial \mathbf{E}}{\partial t}(t=0) = \mathbf{E}_1. \end{array} \right.$$

Ce problème est alors discrétisé en remplaçant l'espace  $V_0$  par un sous-espace vectoriel de dimension finie  $V_h$ . De plus, le schéma numérique n'est pas explicite en temps. Sans rentrer dans les détails, le schéma s'écrit :

$$M\Phi^{n+1} = \Phi^n, \quad (1.30)$$

où  $\Phi^n$  est le vecteur constitué des degrés de liberté du champ électrique, et  $M$  est une matrice bande réelle, symétrique, définie positive. A chaque pas de temps, le système linéaire (1.30) doit être résolu. Cette inversion ne pose aucun problème d'un point de vue mathématique, mais la très grande dimension de la matrice  $M$  entraîne des temps de calculs importants.

Des techniques de condensation de la matrice de masse permettent avec des approximations numériques de rendre diagonale la matrice  $M$ , ce qui évite l'inversion de cette matrice et fait ainsi gagner du temps de calcul et économiser les ressources informatiques. Malheureusement, ces techniques sont difficiles à mettre en œuvre pour les éléments finis utilisés dans les problèmes de diffraction électromagnétique [Elmkies and Joly, 1997a]-[Elmkies and Joly, 1997b]. Une autre voie d'étude pour diminuer le coût de l'inversion de la matrice  $M$  consiste à coupler la méthode des éléments finis temporels avec la méthode des différences finies [Rylander and Bondeson, 2002]. Il s'agit alors d'utiliser un maillage de type éléments finis confiné aux structures rayonnantes, et à appliquer la méthode FDTD pour compléter le volume de calcul et modéliser les conditions absorbantes.

Il est important de noter qu'il existe une autre méthode des éléments finis fréquemment utilisée. Elle a été introduite à la fin des années 70 par Nédélec [Nedelec, 1980]. Le plus petit ordre (approximation linéaire) en espace à été découvert par d'autres également comme Whitney [Whitney, 1957]. Appelée méthode des éléments finis d'arêtes, cette méthode évalue les champs non pas aux noeuds du maillage comme il est d'usage avec les éléments finis traditionnels mais le long des arêtes dans le cas d'une approximation linéaire.

Enfin, la stabilité des éléments finis en domaine temporel dépend du maillage, aussi bien par la taille des cellules hexaédriques ou tétraédriques que par leur forme. Le critère de stabilité est plus complexe que celui de la méthode FDTD et doit en toute rigueur être étudié pour chaque type de maillage utilisé. On peut citer pour les éléments finis traditionnels l'article de Jiao et Jin [Jiao and Jin, 2002], et [Butrylo et al., 2004] pour les éléments finis d'arêtes. Il existe néanmoins certaines techniques pour passer outre, ce qui permet d'obtenir des schémas inconditionnellement stables [Gedney and Navsariwala, 1995]. Le pas de temps reste toutefois limité par les caractéristiques physiques du problème considéré.

### 1.2.3 Méthodes FVTD

Les formulations de type volumes finis sont particulièrement bien adaptées à la discrétisation de systèmes d'équations aux dérivées partielles exprimant des lois de conservation. Le support de calcul est un volume de contrôle qui peut être un élément de base du maillage (cube, triangle, tétraèdre, etc.) ou une entité géométrique définie à partir d'un ensemble d'éléments attachés à un sommet du maillage. Dans le premier cas on obtient une formulation volume fini centrée élément et dans le second cas une formulation centrée sommet. Dans chaque cas, les inconnues du problème sont les valeurs moyennes sur le volume de contrôle des quantités physiques d'intérêt (les composantes des champs électrique et magnétique dans le cas présent). Dans une méthode volumes finis, les intégrales de volume sont converties en intégrales de surface en utilisant le théorème de la divergence. Le calcul numérique de ces intégrales correspond à l'évaluation d'un bilan de flux à la frontière d'un volume de contrôle. Une fonction de flux numérique est alors introduite qui doit vérifier certaines propriétés parmi lesquelles, la conservativité : le flux entrant dans un volume donné doit être identique au flux sortant du volume adjacent. Un des avantages des méthodes volumes finis réside dans le fait qu'elles se formulent aisément sur maillages non-structurés.

Le système de Maxwell peut s'écrire sous une forme conservative. Il est donc naturel d'adapter à la résolution des équations de Maxwell des schémas conservatifs ayant déjà prouvé leur efficacité (en mécanique des fluides par exemple). Il s'agit de plus d'un système hyperbolique. On peut utiliser cette propriété pour construire des schémas volumes finis décentrés (voir [Cioni, 1995]-[Cioni et al., 1997]-[Shankar et al., 1989] par exemple). Associés à une discrétisation temporelle de type Runge-Kutta, ces schémas sont très robustes. Ils ne sont toutefois pas satisfaisants car leur principal défaut est une diffusion numérique. Ainsi, si l'on simule l'évolution d'un mode propre dans une cavité métallique avec de tels schémas, la solution obtenue après un long temps de calcul est d'énergie nulle alors que l'énergie de la solution exacte est conservée au cours du temps. Signalons que des travaux ont été menés pour diminuer l'erreur de diffusion [Bonnet, 1997].

Un schéma volumes finis non diffusif conservant un équivalent discret de l'énergie électromagnétique a été proposé par Remaki [Remaki, 1999]-[Remaki, 2000]. Il repose sur une fonction de flux numérique centrée et une discrétisation temporelle de type saute-mouton. Tout comme le schéma de Yee, il est précis au second ordre en temps et en espace sur une grille régulière orthogonale. Il permet aussi de traiter des maillages tétraédrique de manière totalement explicite. L'atout majeur de ce schéma est de conserver une énergie discrète sur tout maillage (même non-conforme).

Un article de Fumeaux *et al.* [Fumeaux et al., 2006] expose les différents développements récents de cette méthode.

#### 1.2.4 Méthodes DGTD

Les méthodes Galerkin discontinues s'appuient sur une approximation des composantes du champ électromagnétique locale aux éléments du maillage sans imposer de continuité inter-éléments. D'un certain point de vue, il s'agit d'une méthode élément fini discontinue : l'approximation de la solution du problème est polynômiale par morceaux mais n'est pas globalement continue. Cette caractéristique rend la méthode très flexible vis-à-vis du choix de la représentation locale de la solution : le degré d'interpolation peut-être défini localement (i.e. par élément) et le maillage de calcul peut-être raffiné localement en autorisant la présence de nœuds flottants. En d'autres termes, la méthode se prête idéalement à la définition de stratégies de résolution adaptatives. La discontinuité globale de l'approximation impose une formulation faible locale (i.e. dont le support d'intégration est l'élément). Cette formulation faible fait apparaître un terme de bord qui requiert un traitement approprié du fait de la discontinuité inter-éléments de l'approximation. Ce traitement peut prendre la forme de l'ajout de termes de pénalisation ou de l'introduction d'une fonction de flux numérique suivant que l'on adopte un point de vue élément fini discontinu ou volume fini dans la formulation de la méthode Galerkin discontinue. En résumé, les méthodes Galerkin discontinues présentent les principales caractéristiques suivantes :

- elles sont naturellement adaptées à la discrétisation de fonctions discontinues et à la prise en compte de milieux de propagation hétérogènes.
- Elles permettent de traiter facilement des géométries complexes car de formulation aisée sur des maillages non-structurés.
- Elles sont parfaitement adaptées aux techniques de raffinement de maillage locales non-conformes.
- Le degré d'interpolation de l'approximation de la solution peut être choisi localement en fonction de la régularité de la solution exacte et de la précision désirée.
- Elles conduisent à des matrices de masse diagonales par blocs lorsqu'elles sont associées à des schémas d'intégration en temps explicites.
- Elles sont bien adaptées aux plate-formes de calcul parallèles.

Ce type de méthode a été introduit en 1973 par Reed et Hill [Reed and Hill, 1973] pour la résolution de l'équation de transport de neutrons. La pertinence de la méthode a ensuite été reconnue par Le-

saint et Raviart [Lesaint and Raviart, 1974], Johnson et Pitkaranta [Johnson and Pitkaranta, 1986], Richter [Richter, 1988] et Peterson [Peterson, 1991] qui, à partir de 1974, ont publié les premières analyses mathématiques dans le cas des équations scalaires hyperboliques et linéaires. Bien qu'étant apparues en 1973, les méthodes d'éléments finis discontinues n'ont évolué que récemment pour s'adapter à une large variété de problèmes (hyperboliques, parabolique, elliptique, linéaire ou non, etc). Pour plus de détails sur ces différents développements, on pourra se référer à l'ouvrage de Cockburn [Cockburn et al., 2000] et aux numéros spéciaux de journaux spécialisés [Cockburn and Shu, 2005]-[Dawson, 2006].

Dans le domaine de l'électromagnétisme, l'utilisation des méthodes Galerkin discontinues a commencé au milieu des années 90 avec les travaux de Helluy [Helluy, 1994] pour la résolution des équations de Maxwell dans le domaine harmonique et dans le domaine temporel, et avec les travaux de Remaki et Fezoui en 1998 [Remaki and Fezoui, 1998] pour la résolution des équations de Maxwell instationnaires.

### 1.2.5 Brève revue des schémas d'intégration en temps

Après avoir rappelé les principales méthodes numériques permettant de résoudre les équations de Maxwell en domaine temporel, nous discutons brièvement des schémas d'intégration en temps utilisés dans ce contexte, notamment pour mieux situer notre étude dont l'objectif est la prise en compte de schémas d'intégration en temps efficaces.

#### 1.2.5.1 Schémas explicites

Les schémas d'intégration en temps explicites sont les plus utilisés pour la résolution numérique des équations de Maxwell en domaine temporel. Il en existe principalement deux grandes classes :

- les schémas de type Verlet,
- les schémas de Runge-Kutta.

L'algorithme de Verlet est le schéma le plus utilisé en dynamique moléculaire. Il est la généralisation d'autres schémas connus auparavant comme par exemple le schéma saute-mouton. Feynman dès 1963 le décrit dans le chapitre 9 [Feynman, 1963] et 3 ans plus tard, Yee [Yee, 1966], le premier à avoir développé une méthode numérique pour les équations de Maxwell en domaine temporel, l'adopte pour sa méthode FDTD. Comme nous utilisons des schémas saute-mouton dans le cadre de cette thèse, nous reviendrons plus en détails sur ce schéma au chapitre 2.

Les schémas de Runge-Kutta reposent sur une approche multi-pas c'est-à-dire qu'une première estimation de la solution est utilisée pour calculer une seconde itération plus précise, et ainsi de suite. Dans [Cioni, 1995], on utilise le schéma suivant :

$$\begin{cases} \mathbf{Q}^{(0)} &= \mathbf{Q}^n, \\ \mathbf{Q}^{(k)} &= \mathbf{Q}^{(0)} - \frac{\Delta t}{(n_k + 1 - k)} \Psi \left( \mathbf{Q}^{(k-1)} \right) \quad \text{pour } k = 1, 2, \dots, n_k, \\ \mathbf{Q}^{n+1} &= \mathbf{Q}^{(n_k)}. \end{cases} \quad (1.31)$$

Ce schéma et certaines variantes sont généralement utilisés en combinaison avec des schémas décentrés pour le calcul des flux numériques entre cellules voisines dans les méthodes volumes finis [Cioni, 1995]-[Cioni et al., 1997] et Galerkin discontinues [Hesthaven and Warburton, 2002].

#### 1.2.5.2 Schémas implicites

Il existe des versions implicites des quatre familles de méthodes numériques discutées précédemment. Nous ne discutons ici que de variantes FDTD et FETD implicites.

**Méthodes FDTD implicites.** La méthode FDTD implicite, la plus populaire, repose sur l'utilisation d'un schéma ADI (*Alternate Direction Implicit*) en remplacement du schéma saute-mouton. Nous présentons ici la méthode FDTD-ADI étudiée par Zeng *et al.* [Zheng et al., 1999]-[Zheng et al., 2000] et Namiki [Namiki, 1999]-[Namiki, 2000]. Elle consiste en deux étapes :

– première étape :

$$\begin{aligned}
\frac{E_x|_{i,j,k}^{n+1/2} - E_x|_{i,j,k}^n}{\Delta t/2} &= \frac{1}{\varepsilon} \left( \frac{H_z|_{i,j+1,k}^{n+1/2} - H_z|_{i,j-1,k}^{n+1/2}}{2\Delta y} - \frac{H_y|_{i,j,k+1}^n - H_y|_{i,j,k-1}^n}{2\Delta z} \right) \\
\frac{E_y|_{i,j,k}^{n+1/2} - E_y|_{i,j,k}^n}{\Delta t/2} &= \frac{1}{\varepsilon} \left( \frac{H_x|_{i,j,k+1}^{n+1/2} - H_x|_{i,j,k-1}^{n+1/2}}{2\Delta z} - \frac{H_z|_{i+1,j,k}^n - H_z|_{i-1,j,k}^n}{2\Delta x} \right) \\
\frac{E_z|_{i,j,k}^{n+1/2} - E_z|_{i,j,k}^n}{\Delta t/2} &= \frac{1}{\varepsilon} \left( \frac{H_y|_{i+1,j,k}^{n+1/2} - H_y|_{i-1,j,k}^{n+1/2}}{2\Delta x} - \frac{H_x|_{i,j+1,k}^n - H_x|_{i,j-1,k}^n}{2\Delta y} \right) \\
\frac{H_x|_{i,j,k}^{n+1/2} - H_x|_{i,j,k}^n}{\Delta t/2} &= \frac{1}{\mu} \left( \frac{E_z|_{i,j,k+1}^{n+1/2} - E_z|_{i,j,k-1}^{n+1/2}}{2\Delta y} - \frac{E_y|_{i,j+1,k}^n - E_y|_{i,j-1,k}^n}{2\Delta z} \right) \\
\frac{H_y|_{i,j,k}^{n+1/2} - H_y|_{i,j,k}^n}{\Delta t/2} &= \frac{1}{\mu} \left( \frac{E_x|_{i+1,j,k}^{n+1/2} - E_x|_{i-1,j,k}^{n+1/2}}{2\Delta z} - \frac{E_z|_{i,j,k+1}^n - E_z|_{i,j,k-1}^n}{2\Delta x} \right) \\
\frac{H_z|_{i,j,k}^{n+1/2} - H_z|_{i,j,k}^n}{\Delta t/2} &= \frac{1}{\mu} \left( \frac{E_y|_{i,j+1,k}^{n+1/2} - E_y|_{i,j-1,k}^{n+1/2}}{2\Delta x} - \frac{E_x|_{i+1,j,k}^n - E_x|_{i-1,j,k}^n}{2\Delta y} \right)
\end{aligned} \tag{1.32}$$

– seconde étape :

$$\begin{aligned}
\frac{E_x|_{i,j,k}^{n+1} - E_x|_{i,j,k}^{n+1/2}}{\Delta t/2} &= \frac{1}{\varepsilon} \left( \frac{H_z|_{i,j+1,k}^{n+1/2} - H_z|_{i,j-1,k}^{n+1/2}}{2\Delta y} - \frac{H_y|_{i,j,k+1}^{n+1} - H_y|_{i,j,k-1}^{n+1}}{2\Delta z} \right) \\
\frac{E_y|_{i,j,k}^{n+1} - E_y|_{i,j,k}^{n+1/2}}{\Delta t/2} &= \frac{1}{\varepsilon} \left( \frac{H_x|_{i,j,k+1}^{n+1/2} - H_x|_{i,j,k-1}^{n+1/2}}{2\Delta z} - \frac{H_z|_{i+1,j,k}^{n+1} - H_z|_{i-1,j,k}^{n+1}}{2\Delta x} \right) \\
\frac{E_z|_{i,j,k}^{n+1} - E_z|_{i,j,k}^{n+1/2}}{\Delta t/2} &= \frac{1}{\varepsilon} \left( \frac{H_y|_{i+1,j,k}^{n+1/2} - H_y|_{i-1,j,k}^{n+1/2}}{2\Delta x} - \frac{H_x|_{i,j+1,k}^{n+1} - H_x|_{i,j-1,k}^{n+1}}{2\Delta y} \right) \\
\frac{H_x|_{i,j,k}^{n+1} - H_x|_{i,j,k}^{n+1/2}}{\Delta t/2} &= \frac{1}{\mu} \left( \frac{E_z|_{i,j,k+1}^{n+1/2} - E_z|_{i,j,k-1}^{n+1/2}}{2\Delta y} - \frac{E_y|_{i,j+1,k}^{n+1} - E_y|_{i,j-1,k}^{n+1}}{2\Delta z} \right) \\
\frac{H_y|_{i,j,k}^{n+1} - H_y|_{i,j,k}^{n+1/2}}{\Delta t/2} &= \frac{1}{\mu} \left( \frac{E_x|_{i+1,j,k}^{n+1/2} - E_x|_{i-1,j,k}^{n+1/2}}{2\Delta z} - \frac{E_z|_{i,j,k+1}^{n+1} - E_z|_{i,j,k-1}^{n+1}}{2\Delta x} \right) \\
\frac{H_z|_{i,j,k}^{n+1} - H_z|_{i,j,k}^{n+1/2}}{\Delta t/2} &= \frac{1}{\mu} \left( \frac{E_y|_{i,j+1,k}^{n+1/2} - E_y|_{i,j-1,k}^{n+1/2}}{2\Delta x} - \frac{E_x|_{i+1,j,k}^{n+1} - E_x|_{i-1,j,k}^{n+1}}{2\Delta y} \right)
\end{aligned} \tag{1.33}$$

Si on se limite à la première étape, on obtient un schéma précis à l'ordre un en temps ; en revanche le schéma devient d'ordre 2 en temps à l'issue de la seconde étape. On note par exemple que, dans la première équation de la première étape, dans le membre de droite, seule la composante de  $H_z$  est inconnue puisque calculée en  $n + \frac{1}{2}$  alors que la contribution de  $H_y$  est connue. Au contraire, dans la première équation de la seconde étape, c'est la composante de  $H_y$  qui est inconnue alors que  $H_z$  est cette fois connue en  $n + \frac{1}{2}$ .



A chaque étape, on peut se ramener à la résolution de systèmes linéaires à matrices tridiagonales pour  $E_x$ ,  $E_y$  et  $E_z$ . Par exemple, dans la première étape, on peut éliminer la composante  $H_z$  de la première et la dernière équations, et obtenir un système linéaire pour la composante  $E_x$ . Une fois les trois systèmes linéaires résolus pour les valeurs  $E_x$ ,  $E_y$  et  $E_z$ , les valeurs des composantes  $H_x$ ,  $H_y$  et  $H_z$  sont obtenues de façon explicite. La méthode FDTD-ADI est inconditionnellement stable. Il est important de garder en tête que la structure de cette méthode est intrinsèquement liée à l'utilisation de grilles cartésiennes.

Lee et Fornberg [Lee and Fornberg, 2004] décrivent l'application d'une technique de défaut corrigé pour améliorer l'ordre de précision en temps de la méthode FDTD-ADI. Chen *et al.* [Chen et al., 2000] présentent également une méthode de type FDTD-ADI appliquée au calcul de champs électromagnétiques créés par des circuits électroniques. Dans ce cas, la taille de l'objet à prendre en compte est bien plus petite que la longueur d'onde. Ce type de méthode est particulièrement intéressante dans ce contexte où la discrétisation de l'objet nécessite un grand nombre de points par longueur d'onde conduisant à la résolution d'un problème de taille conséquente. L'avantage de la méthode FDTD-ADI réside alors dans le fait qu'elle évite de traiter le problème dans son ensemble puisqu'elle traite chaque direction spatiale séparément.

Artuzi [Artuzi, 2004] propose une méthode de type FDTD originale et implicite, formulée sur des maillages tétraédriques. Cette méthode repose sur la distinction entre l'élément primal (formé des sommets 1, 2, 3, 4) et l'élément dual (formé des sommets  $\hat{1}$ ,  $\hat{2}$ ,  $\hat{3}$  et  $\hat{4}$ ) comme représentés sur la Fig. 1.3. La connection entre deux sommets définit une arête (ou segment) caractérisée par :

$$\vec{L}_{pq} = \vec{R}_p - \vec{R}_q,$$

où  $\vec{R}_p$  et  $\vec{R}_q$  représentent les vecteurs position des sommets  $p$  et  $q$ . On introduit une matrice de longueurs à partir des vecteurs  $\vec{L}_{pq}$  :

$$L = \begin{bmatrix} L_{12x} & L_{12y} & L_{12z} \\ L_{13x} & L_{13y} & L_{13z} \\ L_{14x} & L_{14y} & L_{14z} \\ L_{23x} & L_{23y} & L_{23z} \\ L_{24x} & L_{24y} & L_{24z} \\ L_{34x} & L_{34y} & L_{34z} \end{bmatrix}.$$

Trois arêtes forment une face triangulaire, dont le vecteur normal orienté vers l'extérieur de la cellule et de norme égale à la surface, est défini comme :

$$\vec{S}_s = \frac{(-1)^s}{2} \vec{L}_{pq} \times \vec{L}_{qr},$$

où  $p$ ,  $q$ ,  $r$  et  $s$  sont des permutations circulaires des sommets 1,2,3 et 4. On peut similairement à  $L$ , former une matrice  $S$  à partir de ces vecteurs :

$$S = \begin{bmatrix} S_{1x} & S_{1y} & S_{1z} \\ S_{2x} & S_{2y} & S_{2z} \\ S_{3x} & S_{3y} & S_{3z} \\ S_{4x} & S_{4y} & S_{4z} \end{bmatrix}.$$

L'élément dual est formé à partir du barycentre du tétraèdre référencé 0, et des barycentres de chaque face notés  $\hat{k}$  où  $k$  représente le sommet de l'élément primal n'appartenant pas à la face.

En reliant 0 avec le barycentre de chaque face, on définit de nouveaux segments et on déduit, de manière similaire à l'élément primal :

$$\vec{L}_s = \frac{\vec{L}_{ps} + \vec{L}_{qs} + \vec{L}_{rs}}{12},$$

qui permet d'obtenir une matrice de longueurs secondaire :

$$\hat{L} = \begin{bmatrix} L_{1x} & L_{1y} & L_{1z} \\ L_{2x} & L_{2y} & L_{2z} \\ L_{3x} & L_{3y} & L_{3z} \\ L_{4x} & L_{4y} & L_{4z} \end{bmatrix}.$$

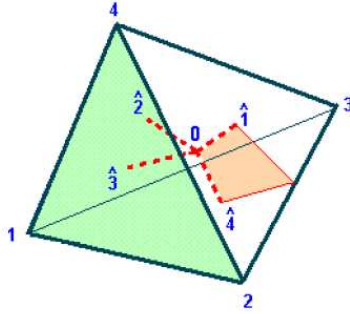


FIG. 1.3 – Eléments primal et dual dans la méthode FDTD en maillages tétraédriques (source : Artuzi [Artuzi, 2004])

Deux segments secondaires forment une face dont les indices correspondent aux sommets exclus de cette face :

$$\vec{S}_{pq} = \frac{\vec{S}_p - \vec{S}_q}{12}.$$

Ces vecteurs permettent de former la matrice :

$$\hat{S} = \begin{bmatrix} S_{12x} & S_{12y} & S_{12z} \\ S_{13x} & S_{13y} & S_{13z} \\ S_{14x} & S_{14y} & S_{14z} \\ S_{23x} & S_{23y} & S_{23z} \\ S_{24x} & S_{24y} & S_{24z} \\ S_{34x} & S_{34y} & S_{34z} \end{bmatrix}.$$

Le schéma est introduit à partir des équations de Maxwell sous forme intégrale :

$$\oint_L \mathbf{E}.dL = -\frac{\partial}{\partial t} \int_S \mathbf{B}.dS, \quad (1.34)$$

$$\oint_L \mathbf{H}.dL = \frac{\partial}{\partial t} \int_S \mathbf{D}.dS + \int_S \mathbf{j}.dS, \quad (1.35)$$

$$\oint_L \mathbf{D}.dS = \int_V \rho dV, \quad (1.36)$$

$$\oint_S \mathbf{B}.dS = 0. \quad (1.37)$$

Sous des hypothèses appropriées, on peut considérer que les équations (1.36) et (1.37) sont redondantes. En supposant constantes les composantes des champs sur un élément, la discrétisation de l'équation (1.34) sur les faces primales conduit à :

$$\Sigma L \mathbf{E} = -\frac{d}{dt} S \mathbf{B}, \quad (1.38)$$

avec :

$$\Sigma = \begin{bmatrix} 0 & 0 & 0 & -1 & 1 & -1 \\ 0 & 1 & -1 & 0 & 0 & 1 \\ -1 & 0 & 1 & 0 & -1 & 0 \\ 1 & -1 & 0 & 1 & 0 & 0 \end{bmatrix}.$$

La discrétisation de l'équation (1.35) sur les faces duales conduit à :

$${}^t \Sigma \hat{L} \mu \mathbf{B} = \frac{d}{dt} S \varepsilon \mathbf{E} + \hat{S} \sigma \mathbf{E} + \hat{S} \mathbf{j}. \quad (1.39)$$

Les équations (1.38) et (1.39) peuvent être combinées pour obtenir :

$$C \frac{d}{dt} \mathbf{V} + G \mathbf{V} + K \int_0^t \mathbf{V} dt = \mathbf{i}, \quad (1.40)$$

où :

$$\mathbf{i} = \hat{S} \mathbf{j} \quad \text{et} \quad \mathbf{V} = -L \mathbf{E},$$

et :

$$C = \hat{S} \varepsilon ({}^t L L)^{-1} {}^t L, \quad (1.41)$$

$$G = \hat{S} \sigma ({}^t L L)^{-1} {}^t L, \quad (1.42)$$

$$K = {}^t \Sigma \hat{L} \nu ({}^t S S)^{-1} {}^t S \Sigma, \quad (1.43)$$

sont respectivement les matrices de capacitance, conductance et reluctance de l'élément.

L'auteur utilise un schéma d'intégration en temps implicite de type Newmark. Appliqué à l'équation (1.40), ce schéma s'écrit :

$$C \frac{d^2 \mathbf{W}}{dt^2} + G \frac{d\mathbf{W}}{dt} + K \mathbf{W} = \mathbf{f},$$

qui est discrétisé par :

$$\begin{aligned} C \frac{\mathbf{W}^{n+1} - 2\mathbf{W}^n + \mathbf{W}^{n-1}}{\Delta t^2} + G \frac{\mathbf{W}^{n+1} - \mathbf{W}^{n-1}}{2\Delta t} \\ + K \frac{\mathbf{W}^{n+1} + 2\mathbf{W}^n + \mathbf{W}^{n-1}}{4} = \mathbf{f}^n, \end{aligned}$$

qui peut être reformulé comme un système de trois équations :

$$\begin{cases} \left( C + \frac{\Delta t}{2} G + \frac{\Delta t^2}{4} K \right) \mathbf{U}^{n+1} = \mathbf{f}^n - G \mathbf{V}^n - K \mathbf{W}^n, \\ \mathbf{V}^{n+1} = \mathbf{V}^n + \Delta t \mathbf{U}^{n+1}, \\ \mathbf{W}^{n+1} = \mathbf{W}^n + \Delta t \mathbf{V}^{n+1}. \end{cases}$$

Puisque  $C$ ,  $G$  et  $K$  sont des matrices symétriques réelles creuses, à coefficients diagonaux positifs, la résolution du système linéaire formulé pour l'obtention de  $\mathbf{U}^{n+1}$  est réalisée au moyen d'une méthode de gradient conjugué préconditionnée par une factorisation incomplète de Cholesky (ICCG). En effet, comme le précise les auteurs, une réduction trop importante du pas de temps dégrade la convergence de la méthode ICCG. On peut s'en rendre compte en regardant l'équation de  $\mathbf{U}^{n+1}$  : l'utilisation d'un pas de temps trop petit ne permet pas de prendre suffisamment en considération les termes  $\mathbf{G}$  et  $\mathbf{K}$  et tend donc à affaiblir la dominance diagonale de la matrice globale.

**Méthodes FETD implicites.** Il existe quelques travaux décrivant des méthodes FETD implicites. Nous considérons ici la méthode étudiée par Gedney et Navsariwala [Gedney and Navsariwala, 1995]. Les auteurs considèrent l'équation des ondes du second ordre :

$$\nabla \times \frac{1}{\mu_r} \nabla \times \mathbf{E} + \mu_0 \sigma \frac{\partial \mathbf{E}}{\partial t} + \frac{\varepsilon_r}{c_0^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} = -\mu_0 \frac{\partial \mathbf{J}}{\partial t}, \quad (1.44)$$

et la formulation faible associée :

$$\iiint_{\Omega} \left( \frac{1}{\mu_r} (\nabla \times \boldsymbol{\varphi}) \cdot (\nabla \times \mathbf{E}) + \mu_0 \sigma \boldsymbol{\varphi} \cdot \frac{\partial \mathbf{E}}{\partial t} + \frac{\varepsilon_r}{c_0^2} \boldsymbol{\varphi} \cdot \frac{\partial^2 \mathbf{E}}{\partial t^2} \right) d\Omega = - \iiint_{\Omega} \mu_0 \boldsymbol{\varphi} \cdot \frac{\partial \mathbf{J}}{\partial t} d\Omega. \quad (1.45)$$

Les champs de vecteurs sont développés en utilisant une forme des éléments d'arêtes de Whitney [Whitney, 1957]  $\mathbf{W}_j$  pondérés par des coefficients  $e_j$  qui sont des fonctions continues en temps, conduisant à l'EDO du second ordre suivante :

$$T_\varepsilon \frac{1}{c_0^2} \frac{d^2 e}{dt^2} + T_\sigma \frac{\eta_0}{c_0} \frac{de}{dt} + S e = -f. \quad (1.46)$$

Enfin, les auteurs utilisent pour l'intégration en temps un schéma implicite de type Newmark :

$$T_\varepsilon \frac{1}{(c_0 \Delta t)^2} (e^{n+1} - 2e^n + e^{n-1}) + T_\sigma \frac{\eta_0}{2c_0 \Delta t} (e^{n+1} - e^{n-1}) + S (\beta e^{n+1} + (1 - 2\beta) e^n + \beta e^{n-1}) = -f. \quad (1.47)$$

Une étude de stabilité permet d'établir que le schéma est inconditionnellement stable pour  $\beta \geq \frac{1}{4}$  et les simulations numériques réalisées par les auteurs montrent que pour  $\beta = \frac{1}{4}$ , l'erreur est minimisée.

### 1.2.5.3 Schémas hybrides explicites/implicites

Nous présentons dans cette section un état des lieux concernant l'hybridation de schémas en temps implicite et explicite dans le contexte de la résolution numérique des équations de Maxwell instationnaires. Dans notre étude, nous nous sommes particulièrement intéressés au schéma initialement proposé par Piperno [Piperno, 2006b] et [Piperno, 2006a]. L'étude de ce schéma fait l'objet du chapitre 5 de cette thèse.

Nous démarrons cette revue par le schéma étudié par Chaudhuri et Van [Chaudhuri and Van, 1999]. Les auteurs présentent dans cet article un schéma hybride pour simuler la propagation dans un guide d'ondes. Dans ce type de problème, il faut traiter des non-linéarités car les matériaux non-linéaires permettent une plus grande flexibilité dans la conception des guides d'ondes. L'idée à la base de la méthode développée par les auteurs est d'appliquer un schéma explicite au milieu linéaire et un schéma implicite au milieu non-linéaire, sur la base d'une discrétisation spatiale par un schéma différences finies.

Dans une structure de guide d'onde 2D, la propagation d'une onde polarisée TE est gouvernée par une équation scalaire :

$$\mu_0 \partial_t (\sigma E_y) + \frac{1}{c_0^2} \partial_{tt} (\varepsilon_r E_y) = (\partial_{xx} + \partial_{zz}) E_y, \quad (1.48)$$

où  $c_0$  désigne la vitesse de la lumière dans le vide,  $\sigma$  désigne la conductivité électrique et  $\varepsilon_r$  est la permittivité relative. Pour un milieu non-linéaire,  $\varepsilon_r$  dépend du champ électrique comme :

$$\varepsilon_r = \varepsilon_{r,L} + \alpha E_y^2.$$

Dans les régions où le milieu est linéaire, une forme discrétisée explicite de l'équation (1.48) est obtenue en utilisant une formule de dérivation en temps au premier ordre et une formulation de dérivation en espace du deuxième ordre. Pour un noeud explicite  $i$ , on obtient :

$$\begin{aligned} \mu_0 \sigma_i \left( \frac{E_i^{n+1} - E_i^n}{\Delta t} \right) + (\varepsilon_r)_i \frac{E_i^{n+1} - 2E_i^n + E_i^{n-1}}{(c\Delta t)^2} \\ = \frac{E_{i_L}^n - 2E_i^n + E_{i_R}^n}{(\Delta z)^2} + \frac{E_{i_A}^n - 2E_i^n + E_{i_B}^n}{(\Delta x)^2}, \end{aligned} \quad (1.49)$$

les indices  $i_L$ ,  $i_R$ ,  $i_A$  et  $i_B$  indiquant respectivement les voisins de gauche, droite, haut et bas du noeud  $i$ . Dans le milieu non-linéaire, les dérivées spatiales sont moyennées sur trois pas de temps  $n-1$ ,  $n$  et  $n+1$ . Il est utile de remarquer que dans l'équation (1.48), nous sommes en présence d'une dérivée en temps du premier ordre et une autre du second ordre. Ainsi, dans l'équation (1.49), seul un schéma en temps du premier ordre est appliqué.

Pour un noeud implicite  $j$ , on obtient alors :

$$\begin{aligned}
\mu_0 \left( \frac{(\sigma E)_j^{n+1} - (\sigma E)_j^n}{\Delta t} \right) &+ \left( \frac{(\varepsilon_r E)_j^{n+1} - 2(\varepsilon_r E)_j^n + (\varepsilon_r E)_j^{n-1}}{(c\Delta t)^2} \right) \\
&= a_1 \left( \frac{E_{j_i}^l - 2E_j^{n+1} + E_{j_R}^l}{(\Delta z)^2} + \frac{E_{j_A}^l - 2E_j^{n+1} + E_{j_B}^l}{(\Delta x)^2} \right) \\
&+ a_2 \left( \frac{E_{j_i}^n - 2E_j^n + E_{j_R}^n}{(\Delta z)^2} + \frac{E_{j_A}^n - 2E_j^n + E_{j_B}^n}{(\Delta x)^2} \right) \\
&+ a_3 \left( \frac{E_{j_i}^m - 2E_j^{n-1} + E_{j_R}^m}{(\Delta z)^2} + \frac{E_{j_A}^l - 2E_j^{n-1} + E_{j_B}^m}{(\Delta x)^2} \right), \tag{1.50}
\end{aligned}$$

où  $a_1$ ,  $a_2$  et  $a_3$  sont des coefficients de pondération tels que  $a_1 + a_2 + a_3 = 1$ . Si l'on pose  $a_2 = 1$  dans (1.50), on aboutit à une forme explicite non-linéaire et pour  $a_1 = 1$  on obtient un schéma totalement implicite dissipatif. Un schéma hybride présentant de bonnes propriétés de stabilité est obtenu pour  $a_1 = a_3 = 1/4$  et  $a_2 = 1/2$ .

Pour relier les zones explicites et implicites, les auteurs suivent une procédure où les indices de temps  $l$  et  $m$  des nœuds voisins dans (1.50) sont définis de la sorte : si un voisin du noeud  $j_\eta \in \{j_L, j_R, j_A, j_B\}$  est aussi un nœud implicite, alors  $l(j_\eta) = n + 1$  et  $m(j_\eta) = n - 1$ ; si par contre  $j_\eta$  est un nœud explicite alors  $l(j_\eta) = m(j_\eta) = n$ . Par ailleurs, une méthode de Newton-Raphson est utilisée pour la résolution du système discret associé à la partie implicite. Le schéma hybride résultant est conditionnellement stable et permet de modéliser de fortes non-linéarités que la méthode FDTD traditionnelle est incapable de traiter.

Contrairement au schéma précédent qui utilise un seul type de méthode de discrétisation, nous discutons maintenant de travaux combinant deux méthodes différentes.

Rylander et Bondeson [Rylander and Bondeson, 2002] proposent une méthode hybride FDTD-FETD qui permet d'allier l'efficacité de la méthode des différences finies à la méthode des éléments finis pour une bonne prise en compte de détails géométriques. Le schéma est explicite pour la partie FDTD alors qu'il est implicite pour la partie FETD. La méthode utilise des couches de tétraèdres pour la discrétisation des détails géométriques complexes associées à des hexaèdres (ou simplement des cubes) pour la partie explicite. Une jonction conforme entre ces deux maillages est réalisée par une couche de pyramides à cinq faces. Il faut souligner que le traitement de cette interface est particulier : les bords intérieurs du maillage structuré (en cubes) sont traités au moyen d'un schéma explicite en temps tandis que ceux des pyramides et tétraèdres utilisent par un schéma implicite en temps.

L'avantage majeur de la méthode hybride FDTD-FETD proposée par Rylander et Bondeson réside dans le fait que le pas de temps utilisé en pratique est celui lié à la stabilité de la méthode FDTD. En effet, la méthode des éléments finis étant appliquée sur des sous-maillages raffinés, l'utilisation du pas de temps calculé sur les cellules grossières où la méthode FDTD est appliquée est rendue possible par l'adoption d'un schéma d'intégration en temps implicite pour le système semi-discret associé aux régions traitées par la méthode des éléments finis.

Partant d'une équation du second ordre de la forme (1.44), le schéma global s'écrit :

$$S(\theta \mathbf{E}^{n+1} - (2\theta - 1)\mathbf{E}^n + \theta \mathbf{E}^{n-1}) + \frac{1}{\Delta t^2} M(\mathbf{E}^{n+1} - 2\mathbf{E}^n + \mathbf{E}^{n-1}) = 0, \tag{1.51}$$

où  $M$  est la matrice de masse et  $S$  la matrice de rigidité. L'algorithme est stable pour  $\theta \geq \frac{1}{4}$ . Dans le schéma proposé par Rylander et Bondeson,  $\theta$  est défini localement pour chaque élément. Ainsi (1.51) s'écrit encore comme :

$$\sum_{k=1}^K \left( S_k (\theta_k \mathbf{E}^{n+1} - (2\theta_k - 1) \mathbf{E}^n + \theta_k \mathbf{E}^{n-1}) + \frac{1}{\Delta t^2} M_k (\mathbf{E}^{n+1} - 2\mathbf{E}^n + \mathbf{E}^{n-1}) \right) = 0,$$

où  $\theta_k$  est choisi supérieur à  $\frac{1}{4}$  pour les éléments implicites et fixé à 0 pour les éléments du maillage structuré. Bien que les auteurs ne précisent aucun temps de calcul, pour une même précision et sur une géométrie complexe, leur méthode hybride est plus rapide qu'une méthode de FDTD.

Une approche hybride alternative est décrite par Ferrières *et al.* dans [Ferrieres et al., 2004] où les auteurs proposent de coupler des méthodes FDTD en maillages cartésiens et FVTD en maillages tétraédriques. L'hybridation est basée sur un recouvrement d'une ou plusieurs cellules sur lesquelles les composantes du champ électromagnétique sont évaluées par les deux méthodes FDTD et FVTD (voir la Fig. 1.4).

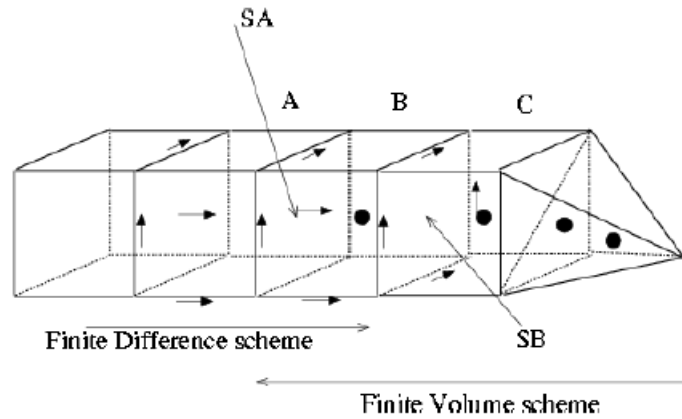


FIG. 1.4 – Maillage hybride avec recouvrement (source : Ferrières *et al.* [Ferrieres et al., 2004])

Dans ce qui suit, on note les champs associés à la méthode FDTD par un exposant  $DF$  alors que ceux associés à la méthode FVTD ont un exposant  $VF$ . Le champ électrique  $E^{DF}$  le long des arêtes de la surface  $S_B$  est calculé par interpolation du champ électrique  $E_i^{VF}$  pris au centre des cellules volumes finis qui se situent autour de l'arête considérée :

$$E^{DF} = \sum_{i=1}^4 \frac{E_i^{VF}}{4}.$$

Les champs à l'intérieur de la partie non-structurée du maillage sont évalués par la méthode FVTD. Les champs dans la cellule  $B$  nécessitent la connaissance des flux provenant de la face  $S_A$  qui dépendent des champs à l'intérieur de la cellule  $A$ . Ainsi, les auteurs interpolent chacune des quatre valeurs du champ électrique  $e_i^{DF}$  et les deux valeurs du champ magnétique  $h_i^{DF}$  données par la méthode FDTD au centre de la cellule  $A$  pour évaluer les valeurs des champs  $e^{VF}$  et  $h^{VF}$  et par conséquent, des flux souhaités :

$$e^{VF} = \frac{1}{4} \sum_{i=1}^4 e_i^{DF} \quad \text{et} \quad h^{VF} = \frac{1}{2} \sum_{i=1}^2 h_i^{DF}.$$

La méthode impose que les nœuds des deux parties des maillages coïncident. L'interface entre les cellules hexaédriques et tétraédriques est obtenue en insérant des cellules hexaédriques où une ou plusieurs faces sont définies par deux triangles. Ces cellules sont traitées par la méthode FVTD.

Etant donné que les champs électrique et magnétique sont calculés respectivement aux temps  $(n+1)\Delta t$  et  $(n+\frac{1}{2})\Delta t$  dans la méthode FDTD alors qu'ils sont calculés au temps  $(n+1)\Delta t$  dans la méthode FVTD, les auteurs proposent la stratégie suivante :

1. évaluer le champ électrique au temps  $n\Delta t$  sur la surface de la face  $S_b$  avec les valeurs calculées par la méthode FVTD,
2. calculer le champ magnétique au temps  $(n+\frac{1}{2})\Delta t$  avec la méthode FDTD dans la zone concernée,
3. calculer le champ électrique au temps  $(n+1)\Delta t$  avec la méthode FDTD dans la zone concernée,
4. interpoler le champ magnétique aux interfaces de la couche  $A$  au temps  $n\Delta t$  en utilisant les valeurs calculées à l'aide de la méthode FDTD aux temps  $(n-\frac{1}{2})\Delta t$  et  $(n+\frac{1}{2})\Delta t$ ,
5. interpoler le champ magnétique aux interfaces de la couche  $A$  au temps  $n\Delta t$  en utilisant les valeurs calculées à l'aide de la méthode FDTD au même instant,
6. calculer les champs électrique et magnétique au temps  $(n+1)\Delta t$  avec la méthode FVTD dans la zone concernée.

Une particularité de cette méthode est qu'elle utilise deux pas de temps,  $\Delta t$  et  $\Delta t_{loc}$ . Le pas de temps global  $\Delta t$  est utilisé par la méthode FDTD ainsi que sur les cellules les plus grossières du maillage où est appliqué la méthode FVTD. Le pas de temps local  $\Delta t_{loc}$  est quant à lui réservé aux petites cellules pour la méthode FVTD. En pratique, la procédure suivante est utilisée :

1. calculer les champs électrique et magnétique au temps  $(n+1)\Delta t$  pour les cellules relevant du pas de temps global,
2. pour chaque itération  $n_i$  de pas de temps local,
  - (a) pour les cellules relevant du pas de temps global voisines de cellules relevant du pas de temps local, calculer la valeur moyenne des champs électrique et magnétique à l'itération  $n_i\Delta t_{loc}$  en utilisant les valeurs aux temps  $n\Delta t$  et  $(n+1)\Delta t$ ,
  - (b) pour les cellules relevant du pas de temps local, calculer les valeurs des champs électrique et magnétique au temps  $(n_i+1)\Delta t_{loc}$ .

Les auteurs testent leur méthode hybride sur un problème de diffraction par une géométrie de cône plat. Les temps de calcul obtenus sont de 99 mn, 35 mn et 25 mn pour respectivement les méthodes FDTD, FVTD et FDTD/FVTD. La précision des résultats est moindre pour la méthode hybride tout en restant néanmoins très acceptable (voir [Ferrieres et al., 2004] pour plus de détails).

Nous concluons cette revue par une présentation de la méthode hybride proposée par Beilina et Grote [Beilina and Grote, 2004]. Il s'agit d'une méthode entièrement explicite, adaptative et hybride éléments finis/différences finies pour la résolution des équations de Maxwell en domaine temporel.



Le domaine de calcul est partitionné en deux parties  $\Omega_{FDM}$  et  $\Omega_{FEM}$  respectivement traitées par des méthodes FDTD et FETD. Dans le domaine  $\Omega_{FDM}$ , le schéma de Yee [Yee, 1966] est appliqué au système des équations de Maxwell du premier ordre. La méthode FETD est appliquée à une formulation du second ordre des équations de Maxwell du type (1.44). Dans le domaine  $\Omega_{FEM}$ , le système discret obtenu s'écrit :

$$M (\mathbf{E}^{n+1} - 2\mathbf{E}^n + \mathbf{E}^{n-1}) = -\tau^2 F^n + s\tau^2 C \mathbf{j}^n - \tau^2 K \mathbf{E}^n - s\tau^2 C \mathbf{E}^n,$$

avec  $M$  la matrice de masse,  $K$  la matrice de rigidité correspondant au terme du rotationnel,  $C$  la matrice de rigidité correspondant au terme de divergence et  $F$  le second membre correspond au courant imposé. Les blocs élémentaires des matrices s'écrivent :

$$\begin{aligned} M_{i,j}^e &= (\varepsilon \varphi_i, \varphi_j)_e, \\ K_{i,j}^e &= \left( \frac{1}{\mu} \nabla \times \varphi_i, \nabla \times \varphi_j \right)_e, \\ C_{i,j}^e &= \left( \frac{1}{\mu} \nabla \cdot \varphi_i, \nabla \cdot \varphi_j \right)_e. \end{aligned}$$

De façon assez classique, pour obtenir un schéma explicite, les auteurs remplacent  $M$  par la matrice de masse condensée  $M^L$  :

$$\begin{aligned} \mathbf{E}^{n+1} &= -\tau^2 (M^L)^{-1} F^n + 2\mathbf{E}^n - \tau^2 (M^L)^{-1} K \mathbf{E}^n \\ &= -s\tau^2 (M^L)^{-1} C \mathbf{E}^n + s\tau^2 (M^L)^{-1} C \mathbf{j}^n - \mathbf{E}^{n-1}. \end{aligned}$$

Les communications de données entre la méthode FETD sur la partie du maillage non-structuré, et la méthode FDTD sur la partie structurée s'appuient sur un recouvrement croisé du maillage d'une couche de deux cellules. On introduit les notations suivantes (voir la Fig. 1.5) pour une illustration en 1D) :

- $\omega_o$  est l'ensemble des nœuds 'o' intérieurs à  $\Omega_{FDM}$  qui sont liés à la frontière de  $\Omega_{FEM}$ .
- $\omega_x$  est l'ensemble des nœuds 'x' intérieurs à  $\Omega_{FEM}$  qui sont liés à la frontière de  $\Omega_{FDM}$ .
- $\omega_*$  est l'ensemble des nœuds '\*' intérieurs à  $\Omega_{FEM}$  qui ne sont pas contenus dans  $\Omega_{FDM}$ .
- $\omega_D$  est l'ensemble des nœuds 'D' intérieurs à  $\Omega_{FDM}$  qui ne sont pas contenus dans  $\Omega_{FEM}$ .

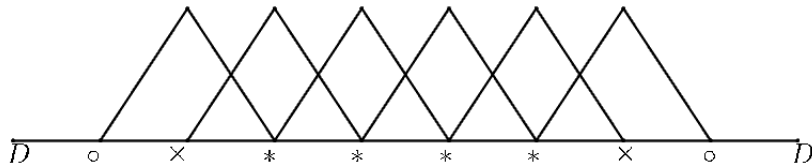


FIG. 1.5 – Couplage entre les méthodes FETD et FDTD en une dimension d'espace

Les nœuds de  $\omega_o$  et  $\omega_x$  sont stockés deux fois i.e. à la fois comme nœuds appartenant à  $\Omega_{FDM}$  et à  $\Omega_{FEM}$ . L'algorithme proposé pour l'avancement en temps du champ électrique est alors le suivant :

1. dans la partie structurée du maillage,  $\Omega_{FDM}$ , calculer  $\mathbf{H}^{n+\frac{1}{2}}$  avec la méthode FDTD, à partir de la valeur connue  $\mathbf{H}^{n-\frac{1}{2}}$ , et calculer  $\mathbf{E}^{n+1}$  avec  $\mathbf{E}^n$  et  $\mathbf{H}^{n+\frac{1}{2}}$

2. dans la partie non-structurée du maillage,  $\Omega_{FEM}$ , calculer  $\mathbf{E}^{n+1}$  en utilisant le schéma FETD
3. utiliser les valeurs du champ électrique  $\mathbf{E}_{FEM}^{n+1}$  aux nœuds  $\omega_{\times}$  comme une condition limite pour la méthode FDTD dans  $\Omega_{FDM}$ . Pour obtenir les valeurs de la composante  $E_{x,FDM}^{n+1}$  aux nœuds de  $\omega_{\times}$ , les auteurs utilisent l'approximation suivante :

$$E_{x,FDM}^{n+1}(p + \frac{1}{2}, q, r) = \frac{E_{x,FEM}^{n+1}(p + 1, q, r) + E_{x,FEM}^{n+1}(p, q, r)}{2}.$$

Toutes les autres composantes du champ électrique sont obtenues de manière similaire.

4. utiliser les valeurs du champ électrique aux nœuds de  $\omega_o$  comme condition limite pour la méthode FETD dans  $\Omega_{FEM}$ . L'approximation suivante est utilisée pour avoir les valeurs de la composante  $E_{x,FEM}^{n+1}$  aux nœuds de  $\omega_o$  :

$$E_{x,FEM}^{n+1}(p, q, r) = \frac{E_{x,FDM}^{n+1}(p + \frac{1}{2}, q, r) + E_{x,FDM}^{n+1}(p - \frac{1}{2}, q, r)}{2}.$$

Toutes les autres composantes du champ électrique sont obtenues de manière similaire.

Les auteurs proposent aussi une procédure d'adaptation du maillage basée sur un estimateur d'erreur a posteriori. Les expériences numériques visent à mettre en évidence le gain de précision avec l'adaptation de maillage.

## Chapitre 2

# Méthodes Galerkin discontinues en domaine temporel

### 2.1 Introduction

Nous présentons dans ce chapitre la méthode Galerkin discontinue qui est à la base de notre étude [Piperno and Fezoui, 2003]-[Fezoui et al., 2005]-[Bernacki et al., 2006a]. Nous détaillons dans un premier temps la discrétisation en espace en explicitant la formulation faible puis le traitement numérique des conditions aux limites. Dans [Piperno and Fezoui, 2003]-[Fezoui et al., 2005] l'intégration en temps des équations semi-discrétisées est réalisée au moyen d'un schéma saute-mouton précis au second ordre. Outre cette option, nous considérons aussi l'utilisation d'un schéma saute-mouton précis au quatrième ordre. De même, dans [Piperno and Fezoui, 2003]-[Fezoui et al., 2005], la mise en œuvre de la méthode Galerkin discontinue se limitait à des approximations par morceaux constantes (conduisant à la méthode de volumes finis [Remaki, 2000]-[Piperno et al., 2002]) et linéaires. Dans notre étude, le degré d'interpolation est arbitraire, l'interpolation locale s'appuyant sur des fonctions de base polynomiales nodales (interpolation de Lagrange  $\mathbb{P}_p$  [Ciarlet and Lions, 1991]). Nous concluons ce chapitre par une série de résultats numériques en deux et trois dimensions d'espace visant d'une part, à illustrer les possibilités de la méthode Galerkin discontinue DGTD- $\mathbb{P}_p$  et, d'autre part, à motiver notre étude sur l'amélioration de l'efficacité de cette méthode.

### 2.2 Discrétisation en espace

Soit  $\Omega$  un ouvert, borné et régulier de  $\mathbb{R}^3$  de frontière  $\Gamma = \Gamma^a \cup \Gamma^m$ . Nous partons ici des deux premières équations de (1.12) :

$$\begin{cases} \varepsilon \frac{\partial \mathbf{E}}{\partial t} - \text{rot}(\mathbf{H}) &= -z_0 \sigma \mathbf{E}, \\ \mu \frac{\partial \mathbf{H}}{\partial t} + \text{rot}(\mathbf{E}) &= 0, \end{cases} \quad (2.1)$$

où  $\mathbf{E} \equiv \mathbf{E}(\mathbf{x}, t)$  et  $\mathbf{H} \equiv \mathbf{H}(\mathbf{x}, t)$  désignent respectivement les champs électrique et magnétique;  $\varepsilon \equiv \varepsilon(\mathbf{x}, t)$ ,  $\mu \equiv \mu(\mathbf{x}, t)$  et  $\sigma \equiv \sigma(\mathbf{x}, t)$  sont respectivement les permittivité électrique, perméabilité magnétique et conductivité électrique. Les équations (2.1) ont été redimensionnées et  $\varepsilon$  et  $\mu$  définissent des quantités

relatives ( $z_0 = \sqrt{\frac{\mu_0}{\varepsilon_0}}$  est l'impédance du vide). On complète ce système par les conditions aux limites suivantes où  $\mathbf{n}$  désigne la normale sortante à  $\Gamma$  :

- des conditions métalliques sur  $\Gamma^m$  :  $\mathbf{n} \times \mathbf{E} = 0$ ,
- des conditions absorbantes sur  $\Gamma^a$  :  $\mathbf{n} \times \mathbf{E} + z\mathbf{n} \times (\mathbf{n} \times \mathbf{H}) = \mathbf{n} \times \mathbf{E}^\infty + z\mathbf{n} \times (\mathbf{n} \times \mathbf{H}^\infty)$ ,  
où  $z = \sqrt{\frac{\mu}{\varepsilon}}$  et  ${}^t(\mathbf{E}^\infty, \mathbf{H}^\infty)$  est un champ incident donné.

On peut aussi écrire le système (2.1) sous la forme suivante :

$$\begin{cases} \varepsilon \frac{\partial \mathbf{E}}{\partial t} + N_x \frac{\partial \mathbf{H}}{\partial x} + N_y \frac{\partial \mathbf{H}}{\partial y} + N_z \frac{\partial \mathbf{H}}{\partial z} = -z_0 \sigma \mathbf{E}, \\ \mu \frac{\partial \mathbf{H}}{\partial t} - N_x \frac{\partial \mathbf{E}}{\partial x} - N_y \frac{\partial \mathbf{E}}{\partial y} - N_z \frac{\partial \mathbf{E}}{\partial z} = 0, \end{cases} \quad (2.2)$$

avec :

$$\mathbf{E} = \begin{pmatrix} E_x \\ E_y \\ E_z \end{pmatrix} \quad \text{et} \quad \mathbf{H} = \begin{pmatrix} H_x \\ H_y \\ H_z \end{pmatrix},$$

et où les matrices (antisymétriques)  $N_x$ ,  $N_y$  et  $N_z$  sont données par :

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

Sous une forme condensée (pseudo-conservative), le système (2.2) se ré-écrit comme :

$$Q\mathbf{W}_t + \nabla \cdot F(\mathbf{W}) = \mathbf{J} \quad (2.3)$$

avec :

$$Q = \begin{bmatrix} \varepsilon I_{3 \times 3} & 0_{3 \times 3} \\ 0_{3 \times 3} & \mu I_{3 \times 3} \end{bmatrix} = \begin{bmatrix} Q_\varepsilon & 0_{3 \times 3} \\ 0_{3 \times 3} & Q_\mu \end{bmatrix},$$

$$\mathbf{W} = \begin{pmatrix} E_x \\ E_y \\ E_z \\ H_x \\ H_y \\ H_z \end{pmatrix} = \begin{pmatrix} \mathbf{E} \\ \mathbf{H} \end{pmatrix}, \quad \mathbf{J} = \begin{pmatrix} -z_0 \sigma E_x \\ -z_0 \sigma E_y \\ -z_0 \sigma E_z \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -z_0 \sigma \mathbf{E} \\ 0 \end{pmatrix}, \quad F(\mathbf{W}) = \begin{pmatrix} F_x(\mathbf{W}) \\ F_y(\mathbf{W}) \\ F_z(\mathbf{W}) \end{pmatrix},$$

où :

$$F_x(\mathbf{W}) = \begin{bmatrix} 0_{3 \times 3} & N_x \\ -N_x & 0_{3 \times 3} \end{bmatrix} \mathbf{W} = \begin{pmatrix} 0 \\ H_z \\ -H_y \\ 0 \\ -E_z \\ E_y \end{pmatrix}, \quad F_y(\mathbf{W}) = \begin{bmatrix} 0_{3 \times 3} & N_y \\ -N_y & 0_{3 \times 3} \end{bmatrix} \mathbf{W} = \begin{pmatrix} -H_z \\ 0 \\ H_x \\ E_z \\ 0 \\ -E_x \end{pmatrix},$$

$$F_z(\mathbf{W}) = \begin{bmatrix} 0_{3 \times 3} & N_z \\ -N_z & 0_{3 \times 3} \end{bmatrix} \mathbf{W} = \begin{pmatrix} H_y \\ -H_x \\ 0 \\ -E_y \\ E_x \\ 0 \end{pmatrix}.$$

### 2.2.1 Formulation faible

Le domaine  $\Omega$  est supposé discrétisé par une triangulation  $\mathcal{T}_h = \bigcup_{i=1}^N \tau_i$ , où les éléments  $\tau_i$  sont des triangles en 2D et des tétraèdres en 3D. On note  $\varphi$  une fonction test scalaire. On suppose  $\mathbf{J} = 0$  pour simplifier la présentation (la formulation obtenue s'étend aisément au cas  $\mathbf{J} \neq 0$ ). En multipliant (2.3) par  $\varphi$  et en intégrant sur  $\tau_i$ , nous obtenons :

$$\begin{aligned} \int_{\tau_i} \varphi (Q\mathbf{W}_t + \nabla \cdot F(\mathbf{W})) d\mathbf{x} &= 0 \\ \Leftrightarrow \int_{\tau_i} Q\mathbf{W}_t \varphi d\mathbf{x} + \int_{\tau_i} (\nabla \cdot F(\mathbf{W})) \varphi d\mathbf{x} &= 0. \end{aligned} \quad (2.4)$$

Une intégration par parties permet d'écrire :

$$\begin{aligned} \int_{\tau_i} Q\mathbf{W}_t \varphi d\mathbf{x} + \int_{\tau_i} (\nabla \cdot F(\mathbf{W})) \varphi d\mathbf{x} &= 0 \\ \Leftrightarrow \int_{\tau_i} Q\mathbf{W}_t \varphi d\mathbf{x} - \int_{\tau_i} \nabla \varphi \cdot F(\mathbf{W}) d\mathbf{x} + \int_{\partial\tau_i} (F(\mathbf{W}) \cdot \mathbf{n}) \varphi d\sigma &= 0. \end{aligned} \quad (2.5)$$

Soit  $\mathcal{P}_i = \mathbb{P}_{p_i}[\tau_i]$  l'espace des fonctions polynomiales de degré au plus  $p$  sur  $\tau_i$ . Les degrés de liberté locaux sont notés  $\mathbf{W}_{ij} \in \mathbb{R}^6$ . Soit  $\phi_i = (\varphi_{i1}, \varphi_{i2}, \dots, \varphi_{id_i})$  une base locale de  $\mathcal{P}_i$  et  $\mathbf{W}_i$  la projection  $L_2$ -orthogonale de  $\mathbf{W}$  sur  $\mathcal{P}_i$ , on a :

$$\mathbf{W}_i(\mathbf{x}) = \sum_{j=1}^{d_i} \mathbf{W}_{ij} \varphi_{ij}(\mathbf{x}). \quad (2.6)$$

En toute généralité, le degré d'interpolation  $p_i$  est défini localement pour chaque cellule (triangle ou tétraèdre) du maillage d'où l'introduction d'un nombre de degrés de liberté  $d_i$  local. Toutefois, dans cette étude, nous ne considérons en pratique que le cas  $p_i = p$  et  $d_i = d \forall \tau_i \in \mathcal{T}_h$ . Néanmoins, la formulation de la méthode DGTD est présentée dans le cadre général d'un degré d'interpolation  $p_i$  variable en espace.

En injectant  $\mathbf{W}_i$  dans (2.5) nous obtenons :

$$\begin{aligned} \int_{\tau_i} Q(\mathbf{W}_i)_t \varphi d\mathbf{x} - \int_{\tau_i} \nabla \varphi \cdot F(\mathbf{W})_i d\mathbf{x} + \int_{\partial\tau_i} (F(\mathbf{W}) \cdot \mathbf{n}) \varphi d\sigma &= 0, \\ \Leftrightarrow Q_i \int_{\tau_i} (\mathbf{W}_i)_t \varphi d\mathbf{x} - \int_{\tau_i} \nabla \varphi \cdot F(\mathbf{W})_i d\mathbf{x} + \sum_{j \in \mathcal{V}_i} \int_{a_{ij}} (F(\mathbf{W}) \cdot \mathbf{n}_{ij}) \varphi d\sigma &= 0, \end{aligned} \quad (2.7)$$

où :

$$F(\mathbf{W})_i \equiv F(\mathbf{W})|_{\tau_i} \quad \text{et} \quad Q_i = \begin{bmatrix} \varepsilon_i I_{3 \times 3} & 0_{3 \times 3} \\ 0_{3 \times 3} & \mu_i I_{3 \times 3} \end{bmatrix}.$$

Dans (2.7),  $a_{ij} = \tau_i \cap \tau_j$  représente un triangle en 3D et une arête en 2D,  $\mathcal{V}_i = \{j | \tau_i \cap \tau_j \neq \emptyset\}$  est l'ensemble des éléments voisins de  $\tau_i$  et  $a_{ij}$  est la face commune entre  $\tau_i$  et  $\tau_j$ ;  $\mathbf{n}_{ij}$  est le vecteur normal à l'interface  $a_{ij}$  dirigé de  $\tau_i$  vers  $\tau_j$ . Puisque la représentation locale de  $\mathbf{W}$  est discontinue d'un élément à un autre, un traitement particulier doit être introduit pour l'évaluation des intégrales de bord sur la face  $a_{ij}$ . Dans le contexte des méthodes volumes finis, on parle de *flux numérique*. Nous utilisons ici un flux numérique basé sur un schéma centré :

$$F(\mathbf{W})|_{a_{ij}} \approx \frac{F(\mathbf{W})_i + F(\mathbf{W})_j}{2}.$$

Par conséquent :

$$\begin{aligned} \sum_{j \in \mathcal{V}_i} \int_{a_{ij}} (F(\mathbf{W}) \cdot \mathbf{n}_{ij}) \varphi d\sigma &= \sum_{j \in \mathcal{V}_i} \int_{a_{ij}} \left( \frac{F(\mathbf{W})_i + F(\mathbf{W})_j}{2} \cdot \mathbf{n}_{ij} \right) \varphi d\sigma \\ &= \frac{1}{2} \sum_{j \in \mathcal{V}_i} \int_{a_{ij}} (F(\mathbf{W})_i \cdot \mathbf{n}_{ij}) \varphi d\sigma \\ &\quad + \frac{1}{2} \sum_{j \in \mathcal{V}_i} \int_{a_{ij}} (F(\mathbf{W})_j \cdot \mathbf{n}_{ij}) \varphi d\sigma \\ &= \frac{1}{2} \int_{\partial\tau_i} (F(\mathbf{W})_i \cdot \mathbf{n}_i) \varphi d\sigma \\ &\quad + \frac{1}{2} \sum_{j \in \mathcal{V}_i} \int_{a_{ij}} (F(\mathbf{W})_j \cdot \mathbf{n}_{ij}) \varphi d\sigma. \end{aligned} \tag{2.8}$$

En intégrant par parties nous obtenons :

$$\begin{aligned} \sum_{j \in \mathcal{V}_i} \int_{a_{ij}} (F(\mathbf{W}) \cdot \mathbf{n}_{ij}) \varphi d\sigma &= \frac{1}{2} \int_{\tau_i} ((\nabla \cdot F(\mathbf{W})_i) \varphi + \nabla \varphi \cdot F(\mathbf{W})_i) d\mathbf{x} \\ &\quad + \frac{1}{2} \sum_{j \in \mathcal{V}_i} \int_{a_{ij}} (F(\mathbf{W})_j \cdot \mathbf{n}_{ij}) \varphi d\sigma. \end{aligned} \tag{2.9}$$

Si nous introduisons l'expression (2.9) dans la dernière équation de (2.7), nous avons :

$$\begin{aligned} Q_i \int_{\tau_i} (\mathbf{W}_i)_t \varphi d\mathbf{x} - \int_{\tau_i} \nabla \varphi \cdot F(\mathbf{W})_i d\mathbf{x} &+ \frac{1}{2} \int_{\tau_i} ((\nabla \cdot F(\mathbf{W})_i) \varphi + \nabla \varphi \cdot F(\mathbf{W})_i) d\mathbf{x} \\ &+ \frac{1}{2} \sum_{j \in \mathcal{V}_i} \int_{a_{ij}} (F(\mathbf{W})_j \cdot \mathbf{n}_{ij}) \varphi d\sigma = 0, \end{aligned} \tag{2.10}$$

soit encore (en tenant compte de la linéarité de  $F(\mathbf{W})$ ) :

$$\begin{aligned} Q_i \int_{\tau_i} (\mathbf{W}_i)_t \varphi d\mathbf{x} &+ \frac{1}{2} \int_{\tau_i} ((\nabla \cdot F(\mathbf{W}_i)) \varphi - \nabla \varphi \cdot F(\mathbf{W}_i)) d\mathbf{x} \\ &+ \frac{1}{2} \sum_{j \in \mathcal{V}_i} \int_{a_{ij}} (F(\mathbf{W}_j) \cdot \mathbf{n}_{ij}) \varphi d\sigma = 0, \end{aligned} \tag{2.11}$$

qui peut être ré-écrite comme :

$$\begin{aligned}
Q_i \int_{\tau_i} (\mathbf{W}_i)_t \varphi d\mathbf{x} &+ \frac{1}{2} \int_{\tau_i} \left( \left( \sum_{k=1}^3 \partial_{x_k} F_{x_k}(\mathbf{W}_i) \right) \varphi - \sum_{k=1}^3 \partial_{x_k} \varphi F_{x_k}(\mathbf{W}_i) \right) d\mathbf{x} \\
&+ \frac{1}{2} \sum_{j \in \mathcal{V}_i} \int_{a_{ij}} \left( \sum_{k=1}^3 n_{ij}^{x_k} F_{x_k}(\mathbf{W}_j) \right) \varphi d\sigma = 0,
\end{aligned} \tag{2.12}$$

avec  $\mathbf{n}_{ij} = {}^t(n_{ij}^{x_1}, n_{ij}^{x_2}, n_{ij}^{x_3})$  et  $x_1 \equiv x$ ,  $x_2 \equiv y$  et  $x_3 \equiv z$ . Maintenant, si nous définissons :

$$\Psi^{x_k} = \begin{bmatrix} 0_{3 \times 3} & N_{x_k} \\ -N_{x_k} & 0_{3 \times 3} \end{bmatrix},$$

l'équation (2.12) devient :

$$\begin{aligned}
Q_i \int_{\tau_i} (\mathbf{W}_i)_t \varphi d\mathbf{x} &+ \frac{1}{2} \int_{\tau_i} \left( \left( \sum_{k=1}^3 \partial_{x_k} \Psi^{x_k} \mathbf{W}_i \right) \varphi - \sum_{k=1}^3 \partial_{x_k} \varphi \Psi^{x_k} \mathbf{W}_i \right) d\mathbf{x} \\
&+ \frac{1}{2} \sum_{j \in \mathcal{V}_i} \int_{a_{ij}} \left( \sum_{k=1}^3 n_{ij}^{x_k} \Psi^{x_k} \mathbf{W}_j \right) \varphi d\sigma = 0 \\
\Leftrightarrow Q_i \int_{\tau_i} (\mathbf{W}_i)_t \varphi d\mathbf{x} &+ \frac{1}{2} \int_{\tau_i} \left( \left( \sum_{k=1}^3 \partial_{x_k} \Psi^{x_k} \mathbf{W}_i \right) \varphi - \sum_{k=1}^3 \partial_{x_k} \varphi \Psi^{x_k} \mathbf{W}_i \right) d\mathbf{x} \\
&+ \frac{1}{2} \sum_{j \in \mathcal{V}_i} \int_{a_{ij}} M_{ij} \mathbf{W}_j \varphi d\sigma = 0,
\end{aligned} \tag{2.13}$$

où :

$$\begin{cases} M_{\mathbf{n}} = \sum_{k=1}^3 n^{x_k} \Psi^{x_k} = \begin{bmatrix} 0_{3 \times 3} & N_{\mathbf{n}} \\ -N_{\mathbf{n}} & 0_{3 \times 3} \end{bmatrix}, \\ M_{ij} = M_{\mathbf{n}_{ij}}, \\ N_{\mathbf{n}} = \sum_{k=1}^3 n^{x_k} N_{x_k}, \\ N_{ij} = N_{\mathbf{n}_{ij}}. \end{cases} \tag{2.14}$$

### 2.2.2 Traitement numérique des conditions aux limites

**Frontières métalliques** Nous notons  $\mathcal{F}_m$  l'ensemble des faces frontières métalliques du maillage  $\mathcal{T}_h$  et  $\mathcal{F}_m^i$  l'ensemble des faces métalliques du tétraèdre  $\tau_i$ . Si  $a_{ij} \in \mathcal{F}_m^i$  alors nous posons :

$$\mathbf{W}_j = \begin{bmatrix} -I_{3 \times 3} & 0_{3 \times 3} \\ 0_{3 \times 3} & I_{3 \times 3} \end{bmatrix} \mathbf{W}_i.$$

Alors,  $\mathbf{W}|_{a_{ij}} = \frac{\mathbf{W}_i + \mathbf{W}_j}{2}$  est tel que :

$$\mathbf{W}|_{a_{ij}} = \begin{bmatrix} 0_{3 \times 3} & 0_{3 \times 3} \\ 0_{3 \times 3} & I_{3 \times 3} \end{bmatrix} \mathbf{W}_i.$$

**Frontières absorbantes** Nous notons  $\mathcal{F}_a$  l'ensemble des faces frontières absorbantes du maillage  $\mathcal{T}_h$  et  $\mathcal{F}_a^i$  l'ensemble des faces absorbantes du tétraèdre  $\tau_i$ . Si  $a_{ij} \in \mathcal{F}_a^i$  alors nous posons :

$$\mathbf{W}_j = \begin{bmatrix} 0_{3 \times 3} & z_i N_{ij} \\ -z_i^{-1} N_{ij} & 0_{3 \times 3} \end{bmatrix} \mathbf{W}_i + \begin{bmatrix} I_{3 \times 3} & -z_i N_{ij} \\ z_i^{-1} N_{ij} & I_{3 \times 3} \end{bmatrix} \mathbf{W}_i^\infty,$$

où  $\mathbf{W}_i^\infty$  est la projection de  $\mathbf{W}^\infty$  sur  $\tau_i$ . Alors,  $\mathbf{W}|_{a_{ij}} = \frac{\mathbf{W}_i + \mathbf{W}_j}{2}$  est tel que :

$$\mathbf{W}|_{a_{ij}} = \frac{1}{2} \begin{bmatrix} I_{3 \times 3} & z_i N_{ij} \\ -z_i^{-1} N_{ij} & I_{3 \times 3} \end{bmatrix} \mathbf{W}_i + \frac{1}{2} \begin{bmatrix} I_{3 \times 3} & -z_i N_{ij} \\ z_i^{-1} N_{ij} & I_{3 \times 3} \end{bmatrix} \mathbf{W}_i^\infty.$$

### 2.2.3 Calcul des intégrales de bord

A partir de maintenant, nous désignons par  $\mathcal{F}^i$  l'ensemble des faces de  $\tau_i$  et on a  $\mathcal{F}^i = \mathcal{F}_d^i \cup \mathcal{F}_a^i \cup \mathcal{F}_m^i$  où  $\mathcal{F}_d^i$ ,  $\mathcal{F}_a^i$  et  $\mathcal{F}_m^i$  représentent respectivement l'ensemble des faces internes, absorbantes et métalliques du tétraèdre  $\tau_i$ . On a alors :

$$\begin{aligned} \sum_{j \in \mathcal{V}_i} \int_{a_{ij}} M_{ij} \mathbf{W}_j \varphi d\sigma &= \sum_{a_{ij} \in \mathcal{F}_d^i} \int_{a_{ij}} M_{ij} \mathbf{W}_j \varphi d\sigma \\ &+ \sum_{a_{ij} \in \mathcal{F}_a^i} \int_{a_{ij}} M_{ij} \mathbf{W}_j \varphi d\sigma \\ &+ \sum_{a_{ij} \in \mathcal{F}_m^i} \int_{a_{ij}} M_{ij} \mathbf{W}_j \varphi d\sigma. \end{aligned} \quad (2.15)$$

Le calcul des flux internes (i.e. des termes de bord pour  $a_{ij} \in \mathcal{F}_d^i$ ) est immédiat. Ensuite :

$$- \text{ pour } a_{ij} \in \mathcal{F}_a^i : \left\{ \begin{aligned} M_{ij} \mathbf{W}_j &= M_{ia} \begin{bmatrix} 0_{3 \times 3} & z_i N_{ia} \\ -z_i^{-1} N_{ia} & 0_{3 \times 3} \end{bmatrix} \mathbf{W}_j \\ &+ M_{ia} \begin{bmatrix} I_{3 \times 3} & -z_i N_{ij} \\ z_i^{-1} N_{ij} & I_{3 \times 3} \end{bmatrix} \mathbf{W}_i^\infty \\ &= \begin{bmatrix} -z_i^{-1} N_{ia}^2 & 0_{3 \times 3} \\ 0_{3 \times 3} & -z_i N_{ia}^2 \end{bmatrix} \mathbf{W}_i + \begin{bmatrix} z_i^{-1} N_{ia}^2 & N_{ia} \\ -N_{ia} & z_i N_{ia}^2 \end{bmatrix} \mathbf{W}_i^\infty \\ &\equiv P_{ia} \mathbf{W}_i + P_{i\infty} \mathbf{W}_i^\infty. \end{aligned} \right.$$

$$- \text{ pour } a_{ij} \in \mathcal{F}_m^i : M_{ij} \mathbf{W}_j = M_{im} \begin{bmatrix} -I_{3 \times 3} & 0_{3 \times 3} \\ 0_{3 \times 3} & I_{3 \times 3} \end{bmatrix} \mathbf{W}_i = \begin{bmatrix} 0_{3 \times 3} & N_{im} \\ N_{im} & 0_{3 \times 3} \end{bmatrix} \mathbf{W}_i \equiv P_{im} \mathbf{W}_i.$$



Dans les expressions ci-dessus,  $N_{ia}$  et  $N_{im}$  sont respectivement calculés à partir des vecteurs  $\mathbf{n}_{ia}$  et  $\mathbf{n}_{im}$  aux faces frontières correspondantes. A ce stade, une remarque s'impose concernant le traitement des termes de bord associés aux faces absorbantes. Le système des équations de Maxwell écrit sous forme conservative :

$$\begin{cases} \frac{\partial \mathbf{D}}{\partial t} - \text{rot} \left( \frac{\mathbf{B}}{\mu} \right) = 0, \\ \frac{\partial \mathbf{B}}{\partial t} + \text{rot} \left( \frac{\mathbf{D}}{\varepsilon} \right) = 0, \end{cases} \quad (2.16)$$

avec  $\mathbf{D} = \varepsilon \mathbf{E}$  et  $\mathbf{B} = \mu \mathbf{H}$  est hyperbolique. Par suite, la matrice :

$$\mathcal{M}_{\mathbf{n}} = M_{\mathbf{n}} Q^{-1} = \begin{bmatrix} 0_{3 \times 3} & \frac{N_{\mathbf{n}}}{\mu} \\ -\frac{N_{\mathbf{n}}}{\varepsilon} & 0_{3 \times 3} \end{bmatrix},$$

est diagonalisable et l'on montre que l'on a la décomposition suivante :

$$\begin{aligned} \mathcal{M}_{\mathbf{n}} &= M_{\mathbf{n}}^+ + M_{\mathbf{n}}^- \\ &= \frac{c}{2} \begin{bmatrix} -N_{\mathbf{n}}^2 & z^{-1} N_{\mathbf{n}} \\ -z N_{\mathbf{n}} & -N_{\mathbf{n}}^2 \end{bmatrix} + \frac{c}{2} \begin{bmatrix} N_{\mathbf{n}}^2 & z^{-1} N_{\mathbf{n}} \\ -z N_{\mathbf{n}} & N_{\mathbf{n}}^2 \end{bmatrix}. \end{aligned}$$

avec  $c = \frac{1}{\sqrt{\varepsilon \mu}}$ . On en déduit :

$$\begin{cases} M_{\mathbf{n}}^{\pm} &= M_{\mathbf{n}}^{\pm} Q^{-1}, \\ M_{\mathbf{n}}^+ &= \frac{1}{2} \begin{bmatrix} -z^{-1} N_{\mathbf{n}}^2 & N_{\mathbf{n}} \\ -N_{\mathbf{n}} & -z N_{\mathbf{n}}^2 \end{bmatrix}, \\ M_{\mathbf{n}}^- &= \frac{1}{2} \begin{bmatrix} z^{-1} N_{\mathbf{n}}^2 & N_{\mathbf{n}} \\ -N_{\mathbf{n}} & z N_{\mathbf{n}}^2 \end{bmatrix}. \end{cases}$$

On constate donc que  $P_{i\infty} \mathbf{W}_i^{\infty} = 2M_{\mathbf{n}_{i\infty}}^- \mathbf{W}_i^{\infty}$  est un terme qui traduit l'information entrante à la frontière artificielle  $\Gamma^a$  issue du champ incident  $\mathbf{W}_i^{\infty}$ .

#### 2.2.4 Equations semi-discrétisées

Commençons par introduire les notations suivantes :

$$\begin{cases} \Phi_i &= \int_{\tau_i} {}^t \phi_i \phi_i dx, \\ \Phi_i^{x_k} &= \int_{\tau_i} ({}^t \phi_i (\partial_{x_k} \phi_i) - {}^t (\partial_{x_k} \phi_i) \phi_i) dx, \\ \Phi_{ij} &= \int_{a_{ij}} {}^t \phi_i \phi_j d\sigma. \end{cases}$$

où :

- $\Phi_i$  est une matrice  $d_i \times d_i$  symétrique définie positive,
- $\Phi_i^{x_k}$  est une matrice  $d_i \times d_i$  antisymétrique,
- $\Phi_{ij}$  est a priori une matrice rectangulaire quelconque si  $d_i \neq d_j$  ( $\Phi_{ij}$  est une matrice  $d_i \times d_i$  symétrique positive si  $d_i = d_j$ ).

Nous allons maintenant proposer deux expressions du système d'équations semi-discrétisées associé à la formulation Galerkin discontinue présentée précédemment.

### 2.2.4.1 Formulation E

Cette formulation sera utilisée par la suite en combinaison avec un schéma d'intégration en temps explicite. Posons  $\mathbb{E}_i = (\mathbf{E}_{i1}, \mathbf{E}_{i2}, \dots, \mathbf{E}_{id_i})$ ,  $\mathbb{H}_i = (\mathbf{H}_{i1}, \mathbf{H}_{i2}, \dots, \mathbf{H}_{id_i})$  et  $\phi_i = (\varphi_{i1}, \varphi_{i2}, \dots, \varphi_{id_i})$  ( $\mathbb{E}_i$  et  $\mathbb{H}_i$  sont des matrices rectangulaires  $3 \times d_i$ ). Si dans la dernière équation de (2.13)  $\varphi$  est remplacée par  $\varphi_{ij}$  pour  $1 \leq j \leq d_i$  alors nous obtenons :

$$\forall \tau_i \in \mathcal{T}_h : \left\{ \begin{array}{l} 2Q_{\varepsilon,i}(\mathbb{E}_i)_t^t \Phi_i + \sum_{k=1}^3 N^{x_k} \mathbb{H}_i^t \Phi_i^{x_k} + \sum_{a_{ij} \in \mathcal{F}_d^i} N_{ij} \mathbb{H}_j^t \Phi_{ij} \\ \quad + \sum_{a_{ij} \in \mathcal{F}_m^i} N_{im} \mathbb{H}_i^t \Phi_{ij} - \sum_{a_{ij} \in \mathcal{F}_a^i} z_i^{-1} N_{ia}^2 \mathbb{E}_i^t \Phi_{ij} \\ = - \sum_{a_{ij} \in \mathcal{F}_a^i} (z_i^{-1} N_{ia}^2 \mathbb{E}_i^\infty + N_{ia} \mathbb{H}_i^\infty)^t \Phi_{ij}, \\ 2Q_{\mu,i}(\mathbb{H}_i)_t^t \Phi_i - \sum_{k=1}^3 N^{x_k} \mathbb{E}_i^t \Phi_i^{x_k} - \sum_{a_{ij} \in \mathcal{F}_d^i} N_{ij} \mathbb{E}_j^t \Phi_{ij} \\ \quad + \sum_{a_{ij} \in \mathcal{F}_m^i} N_{im} \mathbb{E}_i^t \Phi_{ij} - \sum_{a_{ij} \in \mathcal{F}_a^i} z_i N_{ia}^2 \mathbb{H}_i^t \Phi_{ij} \\ = - \sum_{a_{ij} \in \mathcal{F}_a^i} (-N_{ia} \mathbb{E}_i^\infty + z_i N_{ia}^2 \mathbb{H}_i^\infty)^t \Phi_{ij}. \end{array} \right. \quad (2.17)$$

Les inconnues étant habituellement sous forme de vecteurs, nous ré-écrivons les équations (2.17) en vectorisant les matrices  $\mathbb{E}_i$  et  $\mathbb{H}_i$ . Dans ce qui suit,  $\mathbf{E}_i$  et  $\mathbf{H}_i$  désignent maintenant des vecteurs  $3d_i \times 1$  des degrés de liberté locaux  $\mathbf{E}_{ik}$  et  $\mathbf{H}_{ik}$  pour  $k = 1, \dots, d_i$  associés à la cellule  $\tau_i$ . On obtient alors :

$$\forall \tau_i \in \mathcal{T}_h : \left\{ \begin{array}{l} 2\mathcal{X}_{\varepsilon,i} \frac{d\mathbf{E}_i}{dt} + \sum_{k=1}^3 \mathcal{X}_i^{x_k} \mathbf{H}_i + \sum_{a_{ij} \in \mathcal{F}_d^i} \mathcal{X}_{ij} \mathbf{H}_j \\ \quad + \sum_{a_{ij} \in \mathcal{F}_m^i} \mathcal{X}_{im} \mathbf{H}_i - \sum_{a_{ij} \in \mathcal{F}_a^i} \mathcal{X}_{ia}^E \mathbf{E}_i = - \sum_{a_{ij} \in \mathcal{F}_a^i} \mathcal{X}_{i\infty}^E \mathbf{W}_i^\infty, \\ 2\mathcal{X}_{\mu,i} \frac{d\mathbf{H}_i}{dt} - \sum_{k=1}^3 \mathcal{X}_i^{x_k} \mathbf{E}_i - \sum_{a_{ij} \in \mathcal{F}_d^i} \mathcal{X}_{ij} \mathbf{E}_j \\ \quad + \sum_{a_{ij} \in \mathcal{F}_m^i} \mathcal{X}_{im} \mathbf{E}_i - \sum_{a_{ij} \in \mathcal{F}_a^i} \mathcal{X}_{ia}^H \mathbf{H}_i = - \sum_{a_{ij} \in \mathcal{F}_a^i} \mathcal{X}_{i\infty}^H \mathbf{W}_i^\infty, \end{array} \right. \quad (2.18)$$

avec :

$$\begin{aligned}
\mathcal{X}_{\varepsilon,i} &= \begin{bmatrix} (\Phi_i)_{11}Q_{\varepsilon,i} & (\Phi_i)_{12}Q_{\varepsilon,i} & \cdots & (\Phi_i)_{1d_i}Q_{\varepsilon,i} \\ (\Phi_i)_{21}Q_{\varepsilon,i} & (\Phi_i)_{22}Q_{\varepsilon,i} & \cdots & (\Phi_i)_{2d_i}Q_{\varepsilon,i} \\ \vdots & \vdots & & \vdots \\ (\Phi_i)_{d_i1}Q_{\varepsilon,i} & (\Phi_i)_{d_i2}Q_{\varepsilon,i} & \cdots & (\Phi_i)_{d_id_i}Q_{\varepsilon,i} \end{bmatrix}, \\
\mathcal{X}_{\mu,i} &= \begin{bmatrix} (\Phi_i)_{11}Q_{\mu,i} & (\Phi_i)_{12}Q_{\mu,i} & \cdots & (\Phi_i)_{1d_i}Q_{\mu,i} \\ (\Phi_i)_{21}Q_{\mu,i} & (\Phi_i)_{22}Q_{\mu,i} & \cdots & (\Phi_i)_{2d_i}Q_{\mu,i} \\ \vdots & \vdots & & \vdots \\ (\Phi_i)_{d_i1}Q_{\mu,i} & (\Phi_i)_{d_i2}Q_{\mu,i} & \cdots & (\Phi_i)_{d_id_i}Q_{\mu,i} \end{bmatrix}, \\
\mathcal{X}_i^{x_k} &= \begin{bmatrix} (\Phi_i^{x_k})_{11}N^{x_k} & (\Phi_i^{x_k})_{12}N^{x_k} & \cdots & (\Phi_i^{x_k})_{1d_i}N^{x_k} \\ (\Phi_i^{x_k})_{21}N^{x_k} & (\Phi_i^{x_k})_{22}N^{x_k} & \cdots & (\Phi_i^{x_k})_{2d_i}N^{x_k} \\ \vdots & \vdots & & \vdots \\ (\Phi_i^{x_k})_{d_i1}N^{x_k} & (\Phi_i^{x_k})_{d_i2}N^{x_k} & \cdots & (\Phi_i^{x_k})_{d_id_i}N^{x_k} \end{bmatrix}, \\
\mathcal{X}_{ij} &= \begin{bmatrix} (\Phi_{ij})_{11}N_{ij} & (\Phi_{ij})_{12}N_{ij} & \cdots & (\Phi_{ij})_{1d_j}N_{ij} \\ (\Phi_{ij})_{21}N_{ij} & (\Phi_{ij})_{22}N_{ij} & \cdots & (\Phi_{ij})_{2d_j}N_{ij} \\ \vdots & \vdots & & \vdots \\ (\Phi_{ij})_{d_i1}N_{ij} & (\Phi_{ij})_{d_i2}N_{ij} & \cdots & (\Phi_{ij})_{d_id_j}N_{ij} \end{bmatrix}, \\
\mathcal{X}_{im} &= \begin{bmatrix} (\Phi_{ij})_{11}N_{im} & (\Phi_{ij})_{12}N_{im} & \cdots & (\Phi_{ij})_{1d_i}N_{im} \\ (\Phi_{ij})_{21}N_{im} & (\Phi_{ij})_{22}N_{im} & \cdots & (\Phi_{ij})_{2d_i}N_{im} \\ \vdots & \vdots & & \vdots \\ (\Phi_{ij})_{d_i1}N_{im} & (\Phi_{ij})_{d_i2}N_{im} & \cdots & (\Phi_{ij})_{d_id_i}N_{im} \end{bmatrix}, \\
\mathcal{X}_{ia}^E &= z_i^{-1} \begin{bmatrix} (\Phi_{ij})_{11}N_{ia}^2 & (\Phi_{ij})_{12}N_{ia}^2 & \cdots & (\Phi_{ij})_{1d_i}N_{ia}^2 \\ (\Phi_{ij})_{21}N_{ia}^2 & (\Phi_{ij})_{22}N_{ia}^2 & \cdots & (\Phi_{ij})_{2d_i}N_{ia}^2 \\ \vdots & \vdots & & \vdots \\ (\Phi_{ij})_{d_i1}N_{ia}^2 & (\Phi_{ij})_{d_i2}N_{ia}^2 & \cdots & (\Phi_{ij})_{d_id_i}N_{ia}^2 \end{bmatrix} \quad \text{et} \quad \mathcal{X}_{ia}^H = z_i^2 \mathcal{X}_{ia}^E,
\end{aligned}$$

les définitions des matrices  $\mathcal{X}_{i\infty}^E$  et  $\mathcal{X}_{i\infty}^H$  n'étant pas précisées ici. Si l'on suppose  $d_i$  identique pour toutes les cellules  $\tau_i$ , toutes ces matrices exceptées  $\mathcal{X}_{i\infty}^E$  et  $\mathcal{X}_{i\infty}^H$  sont de taille  $3d_i \times 3d_i$ . Les matrices  $\mathcal{X}_{i\infty}^E$  et  $\mathcal{X}_{i\infty}^H$  sont de taille  $3d_i \times 6d_i$ . De plus,  $\mathcal{X}_{\varepsilon,i}$  et  $\mathcal{X}_{\mu,i}$  sont des matrices symétriques définies positives,  $\mathcal{X}_i^{x_k}$  est une matrice symétrique (construite à partir de deux matrices antisymétriques);  $\mathcal{X}_{ij}$  et  $\mathcal{X}_{im}$  sont respectivement des matrices  $3d_i \times 3d_i$  et  $3d_i \times 3d_j$  antisymétriques.

#### 2.2.4.2 Formulation I

Cette formulation sera utilisée par la suite en combinaison avec un schéma d'intégration en temps implicite. Posons  $\mathbb{W}_i = (\mathbf{W}_{i1}, \mathbf{W}_{i2}, \dots, \mathbf{W}_{id_i})$  et  $\phi_i = (\varphi_{i1}, \varphi_{i2}, \dots, \varphi_{id_i})$  ( $\mathbb{W}_i$  est une matrice rectangulaire  $6 \times d_i$ ). Si dans la dernière équation de (2.13)  $\varphi$  est remplacée par  $\varphi_{ij}$  pour  $1 \leq j \leq d_i$  alors nous obtenons :

$$\begin{aligned}
\forall \tau_i \in \mathcal{T}_h : \quad & 2Q_i(\mathbb{W}_i)_t^t \Phi_i + \sum_{k=1}^3 \Psi^{x_k} \mathbb{W}_i^t \Phi_i^{x_k} + \sum_{a_{ij} \in \mathcal{F}_d^i} M_{ij} \mathbb{W}_j^t \Phi_{ij} \\
& + \sum_{a_{ij} \in \mathcal{F}_m^i} P_{im} \mathbb{W}_i^t \Phi_{ij} + \sum_{a_{ij} \in \mathcal{F}_a^i} P_{ia} \mathbb{W}_i^t \Phi_{ij} \\
& = - \sum_{a_{ij} \in \mathcal{F}_a^i} P_{i\infty} \mathbb{W}_i^{\infty t} \Phi_{ij}.
\end{aligned} \tag{2.19}$$

Comme nous l'avons fait pour la formulation E, nous ré-écrivons l'équation (2.19) en vectorisant la matrice  $\mathbb{W}_i$ . On obtient alors :

$$\begin{aligned}
\forall \tau_i \in \mathcal{T}_h : \quad & 2\mathcal{Y}_i \frac{d\mathbf{W}_i}{dt} + \sum_{k=1}^3 \mathcal{Y}_i^{x_k} \mathbf{W}_i + \sum_{a_{ij} \in \mathcal{F}_d^i} \mathcal{Y}_{ij} \mathbf{W}_j \\
& + \sum_{a_{ij} \in \mathcal{F}_m^i} \mathcal{Y}_{im} \mathbf{W}_i + \sum_{a_{ij} \in \mathcal{F}_a^i} \mathcal{Y}_{ia} \mathbf{W}_i = - \sum_{a_{ij} \in \mathcal{F}_a^i} \mathcal{Y}_{i\infty} \mathbf{W}_i^{\infty},
\end{aligned} \tag{2.20}$$

avec :

$$\begin{aligned}
\mathcal{Y}_i &= \begin{bmatrix} (\Phi_i)_{11} Q_i & (\Phi_i)_{12} Q_i & \cdots & (\Phi_i)_{1d_i} Q_i \\ (\Phi_i)_{21} Q_i & (\Phi_i)_{22} Q_i & \cdots & (\Phi_i)_{2d_i} Q_i \\ \vdots & \vdots & \ddots & \vdots \\ (\Phi_i)_{d_i 1} Q_i & (\Phi_i)_{d_i 2} Q_i & \cdots & (\Phi_i)_{d_i d_i} Q_i \end{bmatrix}, \\
\mathcal{Y}_i^{x_k} &= \begin{bmatrix} (\Phi_i^{x_k})_{11} \Psi^{x_k} & (\Phi_i^{x_k})_{12} \Psi^{x_k} & \cdots & (\Phi_i^{x_k})_{1d_i} \Psi^{x_k} \\ (\Phi_i^{x_k})_{21} \Psi^{x_k} & (\Phi_i^{x_k})_{22} \Psi^{x_k} & \cdots & (\Phi_i^{x_k})_{2d_i} \Psi^{x_k} \\ \vdots & \vdots & \ddots & \vdots \\ (\Phi_i^{x_k})_{d_i 1} \Psi^{x_k} & (\Phi_i^{x_k})_{d_i 2} \Psi^{x_k} & \cdots & (\Phi_i^{x_k})_{d_i d_i} \Psi^{x_k} \end{bmatrix}, \\
\mathcal{Y}_{ij} &= \begin{bmatrix} (\Phi_{ij})_{11} M_{ij} & (\Phi_{ij})_{12} M_{ij} & \cdots & (\Phi_{ij})_{1d_j} M_{ij} \\ (\Phi_{ij})_{21} M_{ij} & (\Phi_{ij})_{22} M_{ij} & \cdots & (\Phi_{ij})_{2d_j} M_{ij} \\ \vdots & \vdots & \ddots & \vdots \\ (\Phi_{ij})_{d_i 1} M_{ij} & (\Phi_{ij})_{d_i 2} M_{ij} & \cdots & (\Phi_{ij})_{d_i d_j} M_{ij} \end{bmatrix}, \\
\mathcal{Y}_{im} &= \begin{bmatrix} (\Phi_{ij})_{11} P_{im} & (\Phi_{ij})_{12} P_{im} & \cdots & (\Phi_{ij})_{1d_i} P_{im} \\ (\Phi_{ij})_{21} P_{im} & (\Phi_{ij})_{22} P_{im} & \cdots & (\Phi_{ij})_{2d_i} P_{im} \\ \vdots & \vdots & \ddots & \vdots \\ (\Phi_{ij})_{d_i 1} P_{im} & (\Phi_{ij})_{d_i 2} P_{im} & \cdots & (\Phi_{ij})_{d_i d_i} P_{im} \end{bmatrix}, \\
\mathcal{Y}_{ia} &= \begin{bmatrix} (\Phi_{ij})_{11} P_{ia} & (\Phi_{ij})_{12} P_{ia} & \cdots & (\Phi_{ij})_{1d_i} P_{ia} \\ (\Phi_{ij})_{21} P_{ia} & (\Phi_{ij})_{22} P_{ia} & \cdots & (\Phi_{ij})_{2d_i} P_{ia} \\ \vdots & \vdots & \ddots & \vdots \\ (\Phi_{ij})_{d_i 1} P_{ia} & (\Phi_{ij})_{d_i 2} P_{ia} & \cdots & (\Phi_{ij})_{d_i d_i} P_{ia} \end{bmatrix}.
\end{aligned}$$

## 2.3 Discrétisation en temps

Le choix de la discrétisation en temps est une étape cruciale pour l'efficacité globale de la méthode numérique. Deux types de schémas sont utilisés pour résoudre des problèmes transitoires :

- les schémas explicites. Le principal avantage de ces schémas est la facilité de mise en œuvre. En général, il est aussi relativement aisé de construire des schémas explicites d'ordre élevé. En revanche, la stabilité de ces schémas est contrainte par une condition sur le pas de temps qui peut s'avérer très restrictive lorsque le maillage est non-uniforme ou localement raffiné. Parmi ces schémas, les plus utilisés pour la résolution numérique des équations de Maxwell sont les schémas de type Runge-Kutta et le schéma saute-mouton utilisé dans la méthode FDTD.
- les schémas implicites. Le plus souvent, ces schémas sont inconditionnellement stables. Cependant, ils conduisent à la résolution d'un système linéaire à chaque pas de temps ce qui implique des surcoûts en termes de temps de calcul et d'occupation mémoire. Le choix ou la mise au point de la méthode de résolution de ce système linéaire est une étape importante qui conditionne notablement la pertinence d'un schéma implicite.

Dans ce chapitre nous nous concentrons sur l'utilisation de schémas explicites pour l'intégration en temps du système d'équations ordinaires (2.18). Nous considérons tout d'abord le schéma saute-mouton d'ordre 2 déjà adopté dans [Fezoui et al., 2005], puis un schéma saute-mouton d'ordre 4 [Young, 1999]. L'étude de méthodes DGTD basées sur des schémas d'intégration en temps implicites est l'objet du chapitre suivant.

### 2.3.1 Préliminaires

Pour simplifier l'écriture des schémas saute-mouton d'ordre 2 et d'ordre 4, on suppose que  $\mathcal{F}_a^i = \emptyset$ . Nous ré-écrivons les systèmes d'équations différentielles ordinaires (EDO) locaux (2.18) comme :

$$\forall \tau_i \in \mathcal{T}_h : \begin{cases} M_{\varepsilon,i} \frac{d\mathbf{E}_i}{dt} = -K_i \mathbf{H}_i - \sum_{a_{ij} \in \mathcal{F}_d^i} \mathcal{X}_{ij} \mathbf{H}_j - \sum_{a_{ij} \in \mathcal{F}_m^i} \mathcal{X}_{im} \mathbf{H}_i, \\ M_{\mu,i} \frac{d\mathbf{H}_i}{dt} = K_i \mathbf{E}_i + \sum_{a_{ij} \in \mathcal{F}_d^i} \mathcal{X}_{ij} \mathbf{E}_j - \sum_{a_{ij} \in \mathcal{F}_m^i} \mathcal{X}_{im} \mathbf{E}_i, \end{cases} \quad (2.21)$$

où  $M_{\varepsilon,i}$ ,  $M_{\mu,i}$  et  $K_i$  sont des matrices  $3d_i \times 3d_i$  symétriques, alors que  $\mathcal{X}_{ij}$  et  $\mathcal{X}_{im}$  sont respectivement des matrices  $3d_i \times 3d_i$  et  $3d_i \times 3d_j$  antisymétriques. On déduit de (2.21) deux systèmes d'EDO globaux :

$$\begin{cases} \mathbb{M}_\varepsilon \frac{d\mathbf{E}_h}{dt} = \mathbb{F} \mathbf{H}_h, \\ \mathbb{M}_\mu \frac{d\mathbf{H}_h}{dt} = \mathbb{G} \mathbf{E}_h, \end{cases} \quad (2.22)$$

avec :

$$\begin{cases} \mathbb{F} = -\mathbb{K} - \mathbb{A} - \mathbb{B}, \\ \mathbb{K} \text{ est une matrice symétrique,} \\ \mathbb{A} \text{ est une matrice symétrique,} \\ \mathbb{B} \text{ est une matrice antisymétrique,} \\ \mathbb{G} = -{}^t\mathbb{F}. \end{cases}$$

Les méthodes d'intégration en temps explicites à un pas comme le schéma saute-mouton impliquent le stockage d'une seule solution ancienne suivie par une étape de mise à jour. Par conséquent ces méthodes sont très efficaces du point de vue du temps de calcul et faciles à mettre en œuvre. On peut introduire l'intégrateur saute-mouton explicite d'ordre  $n$  comme une certaine approximation de la solution de l'EDO  $\dot{y}(t) = Ay(t)$  soit :

$$y(t) = e^{At}y(t_0), \quad (2.23)$$

avec  $y(t_0)$  la valeur initiale. L'équivalent en temps discret de (2.23) est alors :

$$y(n\Delta t) = e^{A\Delta t}y((n-1)\Delta t). \quad (2.24)$$

Nous pouvons ré-écrire (2.22) comme :

$$\frac{d}{dt} \begin{pmatrix} \mathbf{H}_h \\ \mathbf{E}_h \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & \mathbb{M}_\mu^{-1}\mathbb{G} \\ \mathbb{M}_\varepsilon^{-1}\mathbb{F} & 0 \end{pmatrix}}_{\mathbb{A}} \underbrace{\begin{pmatrix} \mathbf{H}_h \\ \mathbf{E}_h \end{pmatrix}}_{\mathbf{W}_h(t)}, \quad (2.25)$$

où la matrice  $\mathbb{A}$  ne dépend pas du temps. Cherchant une solution en temps discret de (2.25), on introduit une approximation en temps avec un pas de temps global. La solution en temps discret de l'EDO du premier ordre (2.25) est une version discrétisée de la solution exponentielle selon son équivalent scalaire (2.24) :

$$\mathbf{W}_h(n\Delta t) = \Phi(\Delta t)\mathbf{W}_h((n-1)\Delta t) \quad (2.26)$$

avec :

$$\Phi(\Delta t) = \sum_{i=0}^{\infty} \frac{\Delta t^i}{i!} \mathbb{A}^i := e^{\mathbb{A}\Delta t}. \quad (2.27)$$

Finalement, la solution de (2.25) s'écrit comme :

$$\begin{pmatrix} \mathbf{H}_h(n\Delta t) \\ \mathbf{E}_h(n\Delta t) \end{pmatrix} = \underbrace{\begin{pmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{pmatrix}}_{\Phi} \underbrace{\begin{pmatrix} \mathbf{H}_h((n-1)\Delta t) \\ \mathbf{E}_h((n-1)\Delta t) \end{pmatrix}}_{\mathbf{W}_h((n-1)\Delta t)}. \quad (2.28)$$

Introduisant une grille en temps décalée, comme c'est le cas pour les schémas saute-mouton, nous obtenons l'équation générale de mise à jour de ces schémas :

$$\mathbf{H}_h^{n+1} = [\Phi_{11}^2 - \Phi_{11}\Phi_{12}\Phi_{22}^{-1}\Phi_{21}]\mathbf{H}_h^{n-1} + [\Phi_{12} + \Phi_{11}\Phi_{12}\Phi_{22}^{-1}]\mathbf{E}_h^n, \quad (2.29a)$$

$$\mathbf{H}_h^{n+2} = [\Phi_{21} + \Phi_{22}\Phi_{21}\Phi_{11}^{-1}]\mathbf{H}_h^{n+1} + [\Phi_{22}^2 - \Phi_{22}\Phi_{21}\Phi_{11}^{-1}\Phi_{12}]\mathbf{E}_h^n. \quad (2.29b)$$

En l'absence de courant, nous avons la relation suivante :

$$[\Phi_{11}^2 - \Phi_{11}\Phi_{12}\Phi_{22}^{-1}\Phi_{21}] = [\Phi_{22}^2 - \Phi_{22}\Phi_{21}\Phi_{11}^{-1}\Phi_{12}] = \mathbf{I} + \mathcal{O}(\Delta t^{N+1}), \quad (2.30)$$

caractéristique des schémas saute-mouton.

### 2.3.2 Schéma saute-mouton d'ordre 2

Le schéma saute-mouton du second ordre s'écrit :

$$\begin{cases} \mathbf{H}_h^{n+1/2} &= \mathbf{H}_h^{n-1/2} + \Delta t \mathbb{M}_\mu^{-1} \mathbb{G} \mathbf{E}_h^n, \\ \mathbf{E}_h^{n+1} &= \mathbf{E}_h^n + \Delta t \mathbb{M}_\varepsilon^{-1} \mathbb{F} \mathbf{H}_h^{n+1/2}. \end{cases} \quad (2.31)$$

A partir de maintenant, on suppose  $P_i = p$  et donc  $d_i = d \forall \mathcal{T}_i \in \mathcal{T}_h$ . Nous rappelons ici certaines propriétés de la méthode Galerkin discontinue DGTD- $\mathbb{P}_p$  résultante (pour plus de détails, on consultera [Fezoui et al., 2005]). La méthode DGTD- $\mathbb{P}_p$  obtenue est non-dissipative car basée sur des schémas centrés pour l'intégration en temps et le calcul des flux numériques aux interfaces entre cellules voisines. Une conséquence est que l'énergie électromagnétique discrète définie par :

$$\begin{cases} \mathcal{E}^n &= \sum_{i=1}^N \mathcal{E}_i^n \\ \mathcal{E}_i^n &= \frac{1}{2} \int_{\mathcal{T}_i} \left( {}^t \mathbf{E}_i^n \varepsilon_i \mathbf{E}_i^n + {}^t \mathbf{H}_i^{n-\frac{1}{2}} \mu_i \mathbf{H}_i^{n+\frac{1}{2}} \right) = \frac{1}{2} \left( {}^t \mathbf{E}_i^n M_i^\varepsilon \mathbf{E}_i^n + {}^t \mathbf{H}_i^{n-\frac{1}{2}} M_i^\mu \mathbf{H}_i^{n+\frac{1}{2}} \right), \end{cases} \quad (2.32)$$

est exactement conservée lorsque  $\mathcal{F}_a = \emptyset$  (i.e. lorsque la frontière du domaine de calcul est de type métallique uniquement). Un critère de stabilité est alors obtenu en montrant que l'énergie (2.32) est effectivement une forme quadratique définie positive de ces arguments. Ainsi, sous l'hypothèse qu'il existe des constantes adimensionnées  $\alpha_i$  et  $\beta_{ik}$  ( $k \in \mathcal{V}_i$ ) telles que :

$$\forall \mathbf{X} \in \mathcal{P}_i : \begin{cases} \|\text{rot} \mathbf{X}\|_{\mathcal{T}_i} \leq \frac{\alpha_i P_i}{V_i} \|\mathbf{X}\|_{\mathcal{T}_i}, \\ \|\mathbf{X}\|_{a_{ik}}^2 \leq \frac{\beta_{ik} S_{ik}}{V_i} \|\mathbf{X}\|_{\mathcal{T}_i}^2, \end{cases} \quad (2.33)$$

où  $\|\mathbf{X}\|_{\mathcal{T}_i}$  et  $\|\mathbf{X}\|_{a_{ik}}$  représentent les normes  $L^2$  du champ de vecteur  $\mathbf{X}$  respectivement sur  $\mathcal{T}_i$  et sur la face  $a_{ik}$ , le schéma est  $L^2$ -stable si le pas de temps  $\Delta t$  est choisi tel que :

$$\forall i, \forall k \in \mathcal{V}_i, \quad c_i \Delta t \left( 2\alpha_i + \beta_{ik} \max \left( \sqrt{\frac{\mu_i}{\mu_k}}, \sqrt{\frac{\varepsilon_i}{\varepsilon_k}} \right) \right) < \frac{4V_i}{P_i}, \quad (2.34)$$

où  $P_i$  et  $V_i$  désignent respectivement le périmètre et le volume de  $\mathcal{T}_i$ . Les constantes  $\alpha_i$  et  $\beta_{ik}$  ne dépendent pas de la taille des cellules  $\mathcal{T}_i$  mais de leur géométrie et de la forme des fonctions de base de  $\varphi_{ij}$ .

Par ailleurs, l'étude de convergence de la méthode DGTD- $\mathbb{P}_p$  menée dans [Fezoui et al., 2005] conduit à une erreur d'ordre  $\mathcal{O}(Th^{\min(s,p)}) + \mathcal{O}(\Delta t^2)$  dans  $C^0([0, T]; L^2(\Omega))$ .

### 2.3.3 Schéma saute-mouton d'ordre 4

Le schéma saute-mouton du quatrième ordre adopté ici est issu d'une famille de schémas présentée dans [Young, 1999] et récemment étudiée dans [Spachmann et al., 2002]. En développant (2.27) à l'ordre 4, (2.29) peut être simplifiée sous la forme :

$$\begin{cases} \mathbf{H}_h^{n+1/2} &= \mathbf{H}_h^{n-1/2} + \Delta t \mathbf{M}_\mu^{-1} \mathbf{G} \left( 1 - \frac{\Delta t^2}{24} \mathbf{M}_\varepsilon^{-1} \mathbf{F} \mathbf{M}_\mu^{-1} \mathbf{G} \right) \mathbf{E}_h^n, \\ \mathbf{E}_h^{n+1} &= \mathbf{E}_h^n + \Delta t \mathbf{M}_\varepsilon^{-1} \mathbf{F} \left( 1 - \frac{\Delta t^2}{24} \mathbf{M}_\mu^{-1} \mathbf{G} \mathbf{M}_\varepsilon^{-1} \mathbf{F} \right) \mathbf{H}_h^{n+1/2}. \end{cases} \quad (2.35)$$

Les propriétés de stabilité et de convergence de la méthode DGTD- $\mathbb{P}_p$  couplée au schéma saute-mouton du quatrième ordre sont étudiées dans la thèse d'Hassan Fahs [Fahs, 2008]. Il y est notamment montré que le schéma saute-mouton du quatrième ordre permet l'utilisation de nombres CFL

## 2.4 Résultats numériques

Dans cette section, on présente les résultats d'expériences numériques en deux et trois dimensions d'espace permettant d'une part d'illustrer les possibilités des méthodes DGTD- $\mathbb{P}_p$  basées sur les schémas d'intégration en temps de type saute-mouton du second ordre et du quatrième ordre et, d'autre part, de motiver notre étude des chapitres suivants sur la mise au point de schémas d'intégration en temps efficaces pour les calculs en maillages localement raffinés. Toutes les simulations numériques sont réalisées sur une station de travail basée sur un processeur Intel Xeon 2.33 GHz avec 2 GB de mémoire RAM.

### 2.4.1 Résultats en 2D

#### 2.4.1.1 Mode propre dans une cavité carrée

Le premier cas test que nous considérons est la propagation d'un mode propre dans une cavité carrée ( $[0, 1] \times [0, 1]$ ) dont les parois sont parfaitement conductrices. Pour ce cas test, la solution analytique est donnée par :

$$\begin{cases} H_x &= -\frac{\pi}{\omega} \sin(\pi x) \cos(\pi y) \sin(\omega t), \\ H_y &= \frac{\pi}{\omega} \cos(\pi x) \sin(\pi y) \sin(\omega t), \\ E_z &= \sin(\pi x) \sin(\pi y) \cos(\omega t), \end{cases}$$

avec  $\omega = 2\pi F$  et  $F=212$  MHz. La simulation consiste à initialiser le champ électromagnétique à partir de la solution analytique ci-dessus en tenant compte d'un décalage de  $\frac{\Delta t}{2}$  entre les initialisations des composantes  $H_x$  et  $H_y$  d'une part, et  $E_z$  d'autre part.

Pour ce cas test, nous comparons les solutions numériques obtenues avec les deux schémas d'intégration en temps respectivement référencés comme LF2 et LF4 dans la suite. Nous utilisons des maillages uniformes et non-uniformes (voir la Fig. 2.1) composés respectivement de 200 et de 386 éléments. Pour les évaluations des taux de convergence numérique, on construit des séquences de maillages par raffinement global de maillages grossiers initiaux. La Fig. 2.2 visualise les lignes de contour de la composante  $E_z$  des solutions obtenues avec les méthodes DGTD- $\mathbb{P}_1$  et DGTD- $\mathbb{P}_2$  combinées au schéma LF2 sur le maillage non uniforme.



Les résultats obtenus avec des maillages uniformes sont montrés sur les Fig. 2.3 à Fig. 2.5 sous la forme, d'une part, des évolutions en temps de l'erreur  $L_2$  entre les solutions analytiques et numériques et de l'énergie discrète, et d'autre part des courbes de convergence numérique. Par ailleurs, on résume dans les Tab. 2.1 et 2.2 quelques éléments permettant d'évaluer les performances des différentes méthodes. On constate que l'on obtient bien les taux de convergence prédits par l'analyse [Fezoui et al., 2005]) et notamment que l'utilisation d'un schéma saute-mouton du quatrième ordre permet de mieux exploiter le gain en précision lié à l'interpolation spatiale d'ordre arbitrairement élevé. Le schéma LF4 permet aussi un gain en temps résultant des valeurs plus élevées du nombre de CFL, en dépit d'une complexité arithmétique accrue.

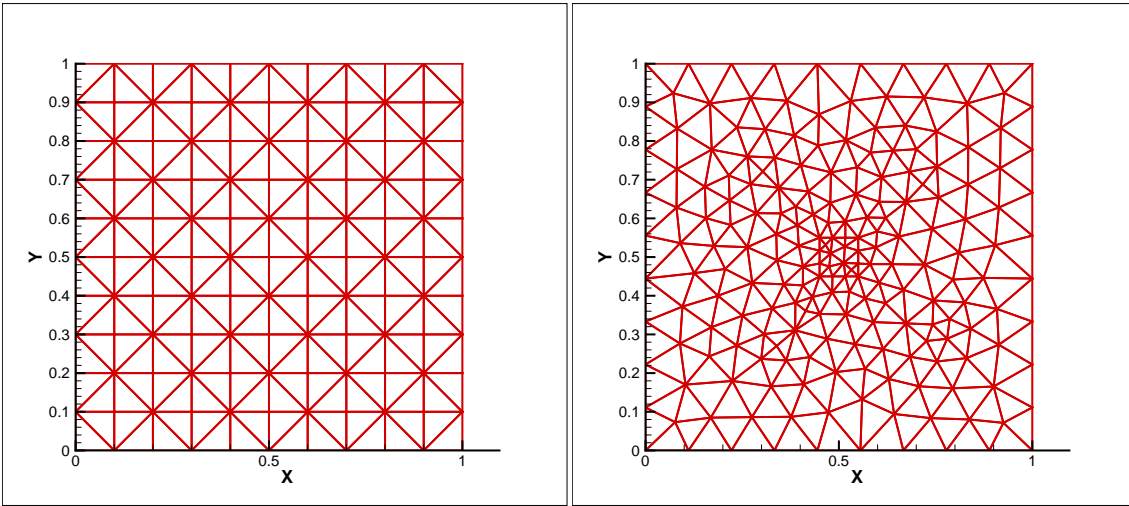


FIG. 2.1 – Mode propre dans un carré unité  
Maillages uniforme (gauche) et non-uniforme (droite)

Schéma en temps	Méthode	CFL	Erreur $L_2$ finale	Temps CPU
LF2	DGTD- $\mathbb{P}_1$	0.3	0.3387	0.6 sec
-	DGTD- $\mathbb{P}_2$	0.2	$0.1861 \cdot 10^{-1}$	2.3 sec
-	DGTD- $\mathbb{P}_3$	0.1	$0.4599 \cdot 10^{-2}$	10.2 sec
-	DGTD- $\mathbb{P}_4$	0.05	$0.1141 \cdot 10^{-2}$	40.5 sec
LF4	DGTD- $\mathbb{P}_1$	0.95	0.3116	0.3 sec
-	DGTD- $\mathbb{P}_2$	0.57	$0.1080 \cdot 10^{-2}$	1.6 sec
-	DGTD- $\mathbb{P}_3$	0.28	$0.1217 \cdot 10^{-3}$	7.6 sec
-	DGTD- $\mathbb{P}_4$	0.14	$0.5355 \cdot 10^{-5}$	30.9 sec

TAB. 2.1 – Mode propre dans un carré unité  
Maillage uniforme : # sommets = 121 et # triangles = 200

#### 2.4.1.2 Diffraction d'une onde plane par un profil d'aile

Le cas test considéré ici est la diffraction d'une onde plane de fréquence  $F=2$  GHz (i.e. de longueur d'onde  $\lambda = 0.15$  m) par un profil d'aile NACA0012 parfaitement conducteur. Le domaine de calcul est

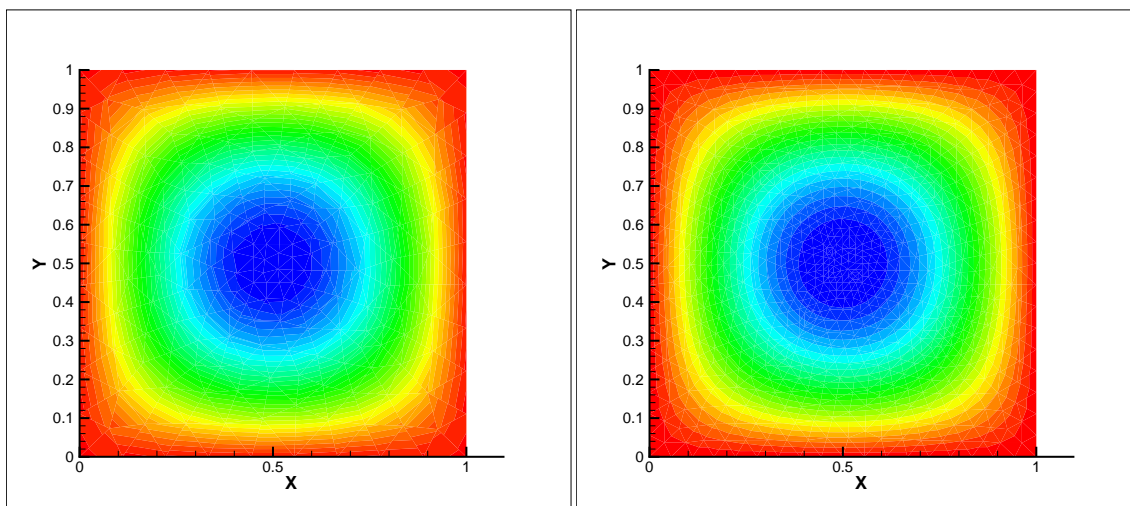


FIG. 2.2 – Mode propre dans un carré unité  
Lignes de contour de  $E_z$  - schéma saute-mouton LF2  
Gauche : méthode DGTD- $\mathbb{P}_1$  - Droite : méthode DGTD- $\mathbb{P}_2$

Schéma en temps	Méthode	Ordre de convergence
LF2	DGTD- $\mathbb{P}_1$	1.92
-	DGTD- $\mathbb{P}_2$	2.03
-	DGTD- $\mathbb{P}_3$	1.99
-	DGTD- $\mathbb{P}_4$	1.94
LF4	DGTD- $\mathbb{P}_1$	1.93
-	DGTD- $\mathbb{P}_2$	2.98
-	DGTD- $\mathbb{P}_3$	2.97
-	DGTD- $\mathbb{P}_4$	3.97

TAB. 2.2 – Mode propre dans un carré unité - Maillages uniformes  
Schémas saute-mouton du second ordre LF2 et du quatrième ordre LF4

Schéma en temps	Méthode	CFL	Erreur $L_2$ finale	Temps CPU
LF2	DGTD- $\mathbb{P}_1$	0.3	0.1340	5.0 sec
-	DGTD- $\mathbb{P}_2$	0.2	$0.1185 \cdot 10^{-2}$	19.0 sec
-	DGTD- $\mathbb{P}_3$	0.1	$0.2155 \cdot 10^{-3}$	84.6 sec
-	DGTD- $\mathbb{P}_4$	0.05	$0.5020 \cdot 10^{-4}$	333.0 sec
LF4	DGTD- $\mathbb{P}_1$	0.95	0.1325	2.5 sec
-	DGTD- $\mathbb{P}_2$	0.57	$0.7118 \cdot 10^{-3}$	13.0 sec
-	DGTD- $\mathbb{P}_3$	0.28	$0.3108 \cdot 10^{-4}$	63.1 sec
-	DGTD- $\mathbb{P}_4$	0.14	$0.5054 \cdot 10^{-5}$	261.9 sec

TAB. 2.3 – Mode propre dans un carré unité  
Maillage non-uniforme : # sommets = 202 et # triangles = 366

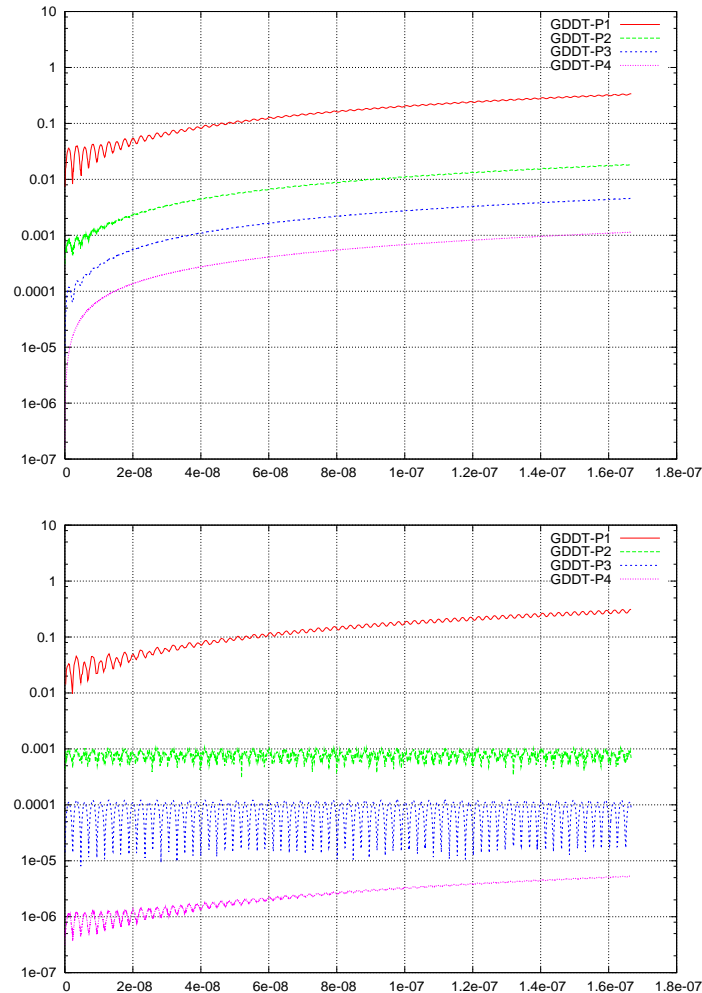


FIG. 2.3 – Mode propre dans un carré unité

Maillage uniforme : # sommets = 121 et # triangles = 200

Evolution temporelle de l'erreur  $L_2$  (abscisse : temps en sec - ordonnée : erreur  $L_2$  en échelle log)

Schémas saute-mouton du second ordre LF2 (haut) et du quatrième ordre LF4 (bas)

Schéma en temps	Méthode	Ordre de convergence
LF2	DGTD- $\mathbb{P}_1$	1.76
-	DGTD- $\mathbb{P}_2$	1.92
-	DGTD- $\mathbb{P}_3$	1.99
LF4	DGTD- $\mathbb{P}_1$	1.84
-	DGTD- $\mathbb{P}_2$	2.93
-	DGTD- $\mathbb{P}_3$	2.98

TAB. 2.4 – Mode propre dans un carré unité - Maillages non-uniformes

Schémas saute-mouton du second ordre LF2 et du quatrième ordre LF4

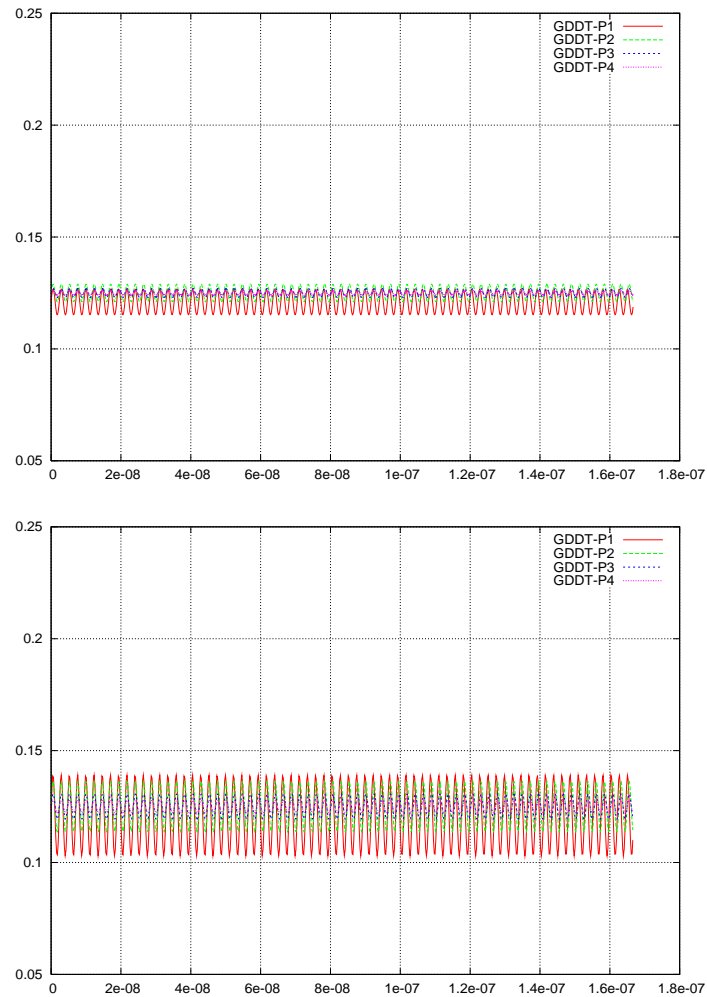


FIG. 2.4 – Mode propre dans un carré unité

Maillage uniforme :  $\#$  sommets = 121 et  $\#$  triangles = 200

Evolution temporelle de l'énergie discrète (abscisse : temps en sec - ordonnée : énergie discrète)

Schémas saute-mouton du second ordre LF2 (haut) et du quatrième ordre LF4 (bas)

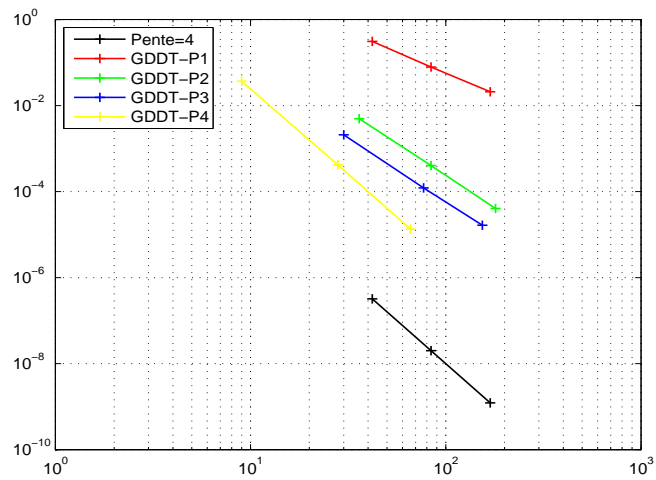
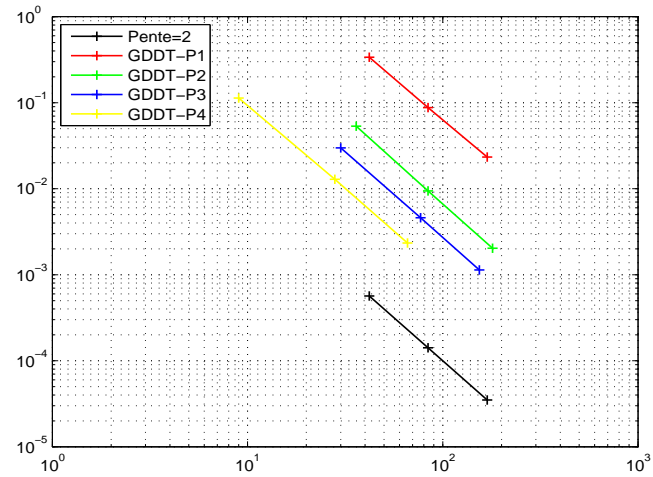


FIG. 2.5 – Mode propre dans un carré unité - Maillages uniformes  
 Convergence numérique (erreur  $L_2$  en fonction de  $\sqrt{\#DOF}$  en échelle log-log)  
 Schémas saute-mouton du second ordre LF2 (haut) et du quatrième ordre LF4 (bas)

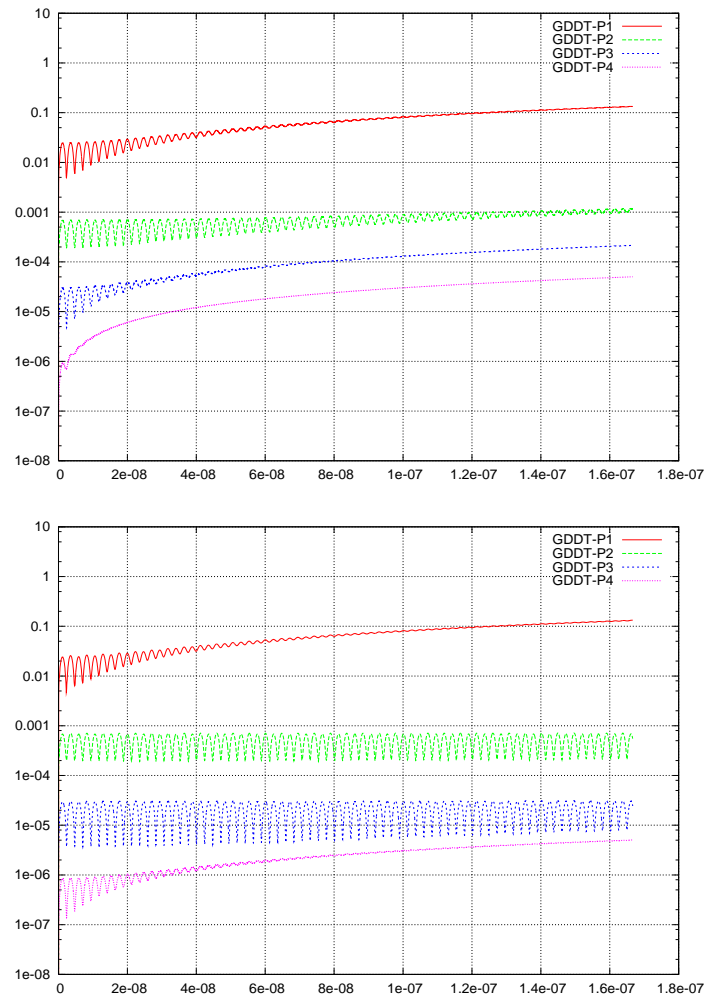


FIG. 2.6 – Mode propre dans un carré unité

Maillage non-uniforme :  $\#$  sommets = 202 et  $\#$  triangles = 366

Evolution temporelle de l'erreur  $L_2$  (abscisse : temps en sec - ordonnée : erreur  $L_2$  en échelle log)

Schémas saute-mouton du second ordre LF2 (haut) et du quatrième ordre LF4 (bas)

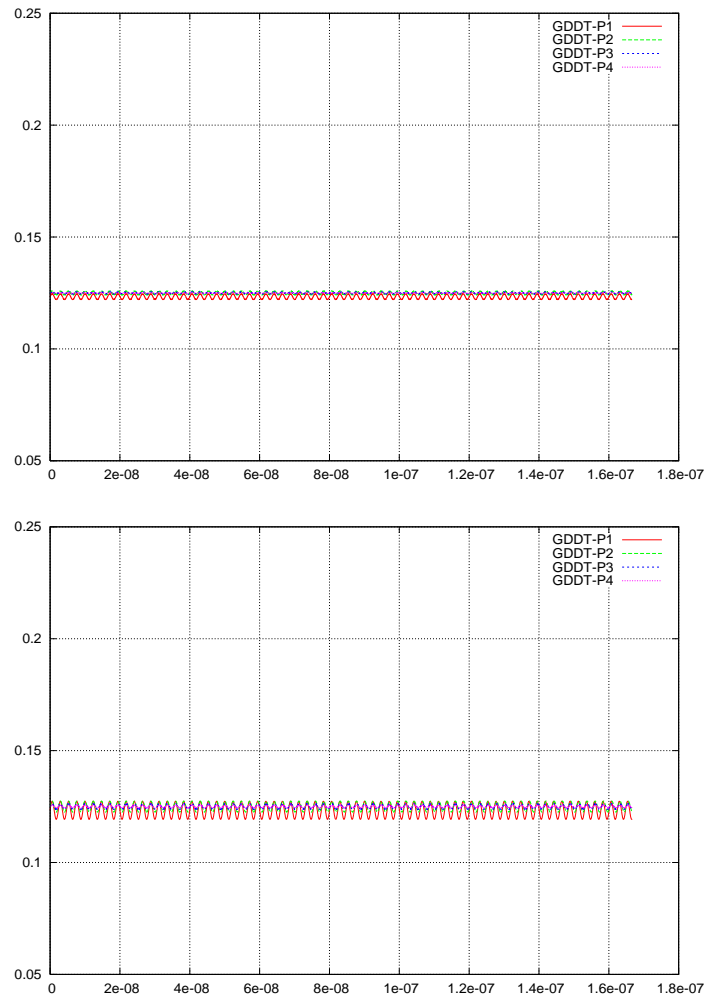


FIG. 2.7 – Mode propre dans un carré unité

Maillage non-uniforme : # sommets = 202 et # triangles = 366

Evolution temporelle de l'énergie discrète (abscisse : temps en sec - ordonnée : énergie discrète)

Schémas saute-mouton du second ordre LF2 (haut) et du quatrième ordre LF4 (bas)

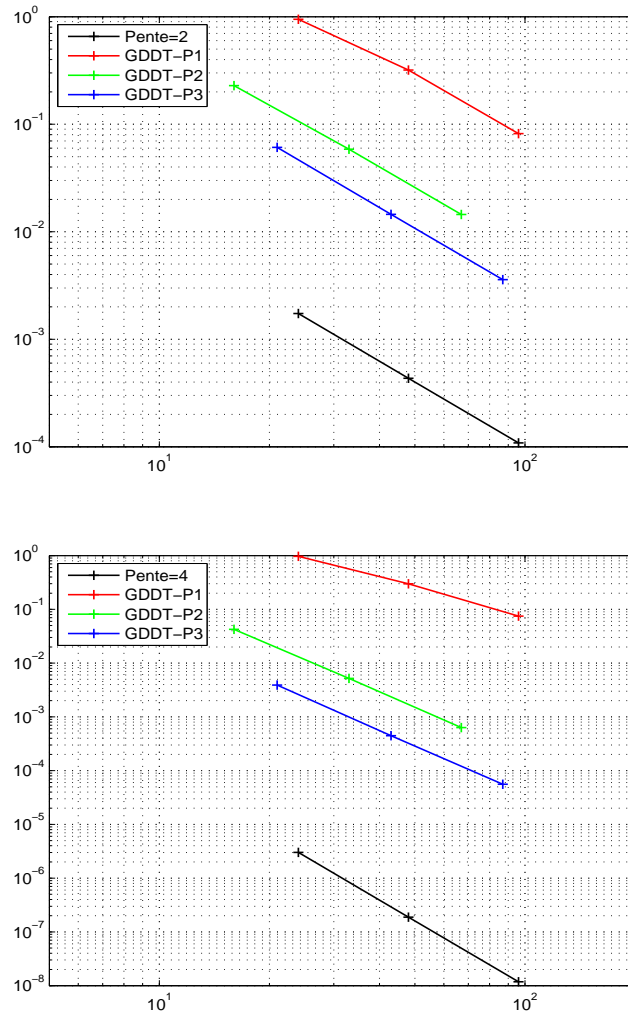


FIG. 2.8 – Mode propre dans un carré unité - Maillages non-uniformes  
 Schémas saute-mouton du second ordre LF2 (haut) et du quatrième ordre LF4 (bas)  
 Convergence numérique (erreur  $L_2$  en fonction de  $\sqrt{\#DOF}$  en échelle log-log)



délimité par un rectangle sur lequel on applique une condition absorbante de type Silver-Müller. Les simulations sont réalisées avec le schéma d'intégration en temps saute-mouton du second ordre et portent sur dix périodes de l'onde incidente. On utilise un maillage non-uniforme relativement grossier comportant 2727 sommets et 5152 triangles (voir la Fig. 2.9). Les longueurs minimale et maximale des arêtes de ce maillage sont respectivement égales à 0.0059364 m (ce qui représente  $\approx \lambda/25$  m) et 0.1795959 m (soit  $\approx 0.9\lambda$  m). Ainsi, à cette valeur de la fréquence, il est clair que le maillage est localement de trop faible résolution. On montre sur la Fig. 2.10 les lignes de contour de la composante  $E_z$  pour différentes méthodes DGTD- $\mathbb{P}_p$ . On peut nettement apprécier l'apport de l'interpolation d'ordre élevé pour ce problème de propagation haute fréquence.

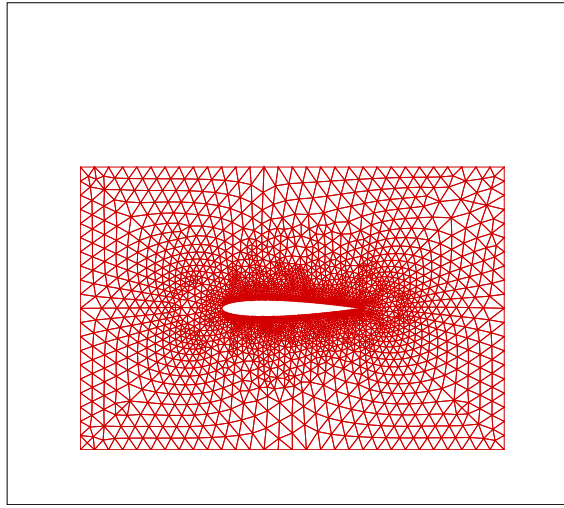


FIG. 2.9 – Diffraction d'une onde plane par un profil d'aile  
Maillage triangulaire : # sommets=2727 et # triangles=5152

### 2.4.2 Résultats en 3D

Le cas test présenté dans cette section a pour but de motiver la présente étude en mettant en avant l'impact de la non-uniformité du maillage sur le pas de temps maximal et le temps de calcul dans le contexte d'un problème de propagation d'onde tridimensionnel. Le problème considéré est la diffraction d'une onde plane de fréquence  $F=1$  GHz (i.e. de longueur d'onde  $\lambda = 0.3$  m) par une géométrie d'avion civil. Le domaine de calcul est délimité par un parallélépipède sur lequel on applique une condition absorbante de type Silver-Müller. Les simulations sont réalisées avec le schéma d'intégration en temps saute-mouton du second ordre et portent sur dix périodes de l'onde incidente. Le maillage tétraédrique utilisé comporte 153821 sommets et 883374 tétraèdres (voir la Fig. 2.11 pour des vues du maillage de la surface de l'avion). Les longueurs minimale et maximale des arêtes de ce maillage sont respectivement égales à  $6.01 \cdot 10^{-4}$  m (ce qui représente  $\approx \lambda/500$  m) et  $1.22 \cdot 10^{-1}$  m (soit  $\approx \lambda/2.5$  m). Les valeurs minimale et maximale du pas de temps (hors CFL) sont respectivement égales à  $\Delta t_{\min} = 0.24$  picosec et  $\Delta t_{\max} = 40.50$  picosec soit un rapport  $\Delta t_{\max}/\Delta t_{\min} \approx 168$ . On montre sur la Fig. 2.12 les lignes de contour de la composante  $E_z$  pour les méthodes DGTD- $\mathbb{P}_1$  et DGTD- $\mathbb{P}_2$  et sur la Fig. 2.13 on visualise l'évolution temporelle de  $E_z$  en un point du domaine pour un comparaison locale des deux méthodes. Les temps de calcul sont donnés dans la Tab. 2.5 pour des simulations réalisées avec un logiciel parallélisé pour des plateformes de

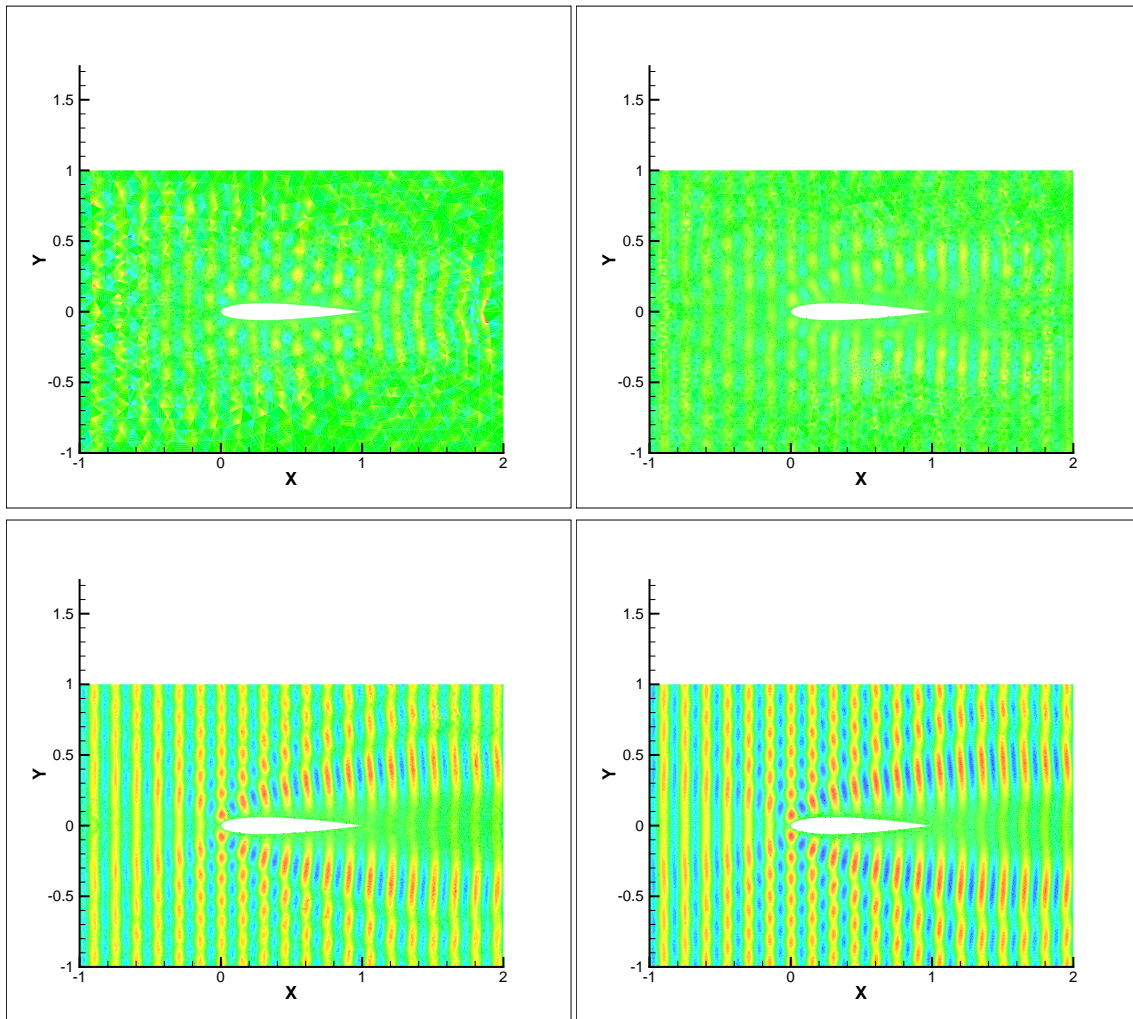


FIG. 2.10 – Diffraction d'une onde plane par un profil d'aile -  $F=2$  GHz

Lignes de contour de la composante  $E_z$

Haut gauche : méthode DGTD- $\mathbb{P}_1$  - Haut droite : méthode DGTD- $\mathbb{P}_2$

Bas gauche : méthode DGTD- $\mathbb{P}_3$  - Bas droite : méthode DGTD- $\mathbb{P}_4$

calcul à mémoire distribuée [Bernacki et al., 2006b]-[Bernacki et al., 2006a]. Compte tenu de l'adoption d'un schéma d'intégration en temps explicite, le pas de temps global utilisé pour les simulations est  $\Delta t = \Delta t_{\min}$ . Par suite, les temps de calcul sont ici pénalisés d'un facteur 168 par rapport à une situation idéale qui autoriserait l'utilisation d'un pas de temps  $\Delta t = \Delta t_{\max}$  et conduirait à des solutions numériques de précision comparable. Malheureusement, comme nous le verrons par la suite, ces deux objectifs sont plutôt antagonistes. Ainsi, un schéma d'intégration en temps globalement implicite inconditionnellement stable, précis au second ordre en temps et non-dissipatif comme le schéma de Crank-Nicolson, est caractérisé par une erreur de dispersion plus importante que celle du schéma saute-mouton du second ordre. Un compromis temps de calcul/précision est donc nécessaire si l'on souhaite envisager des simulations numériques 3D de grande taille.

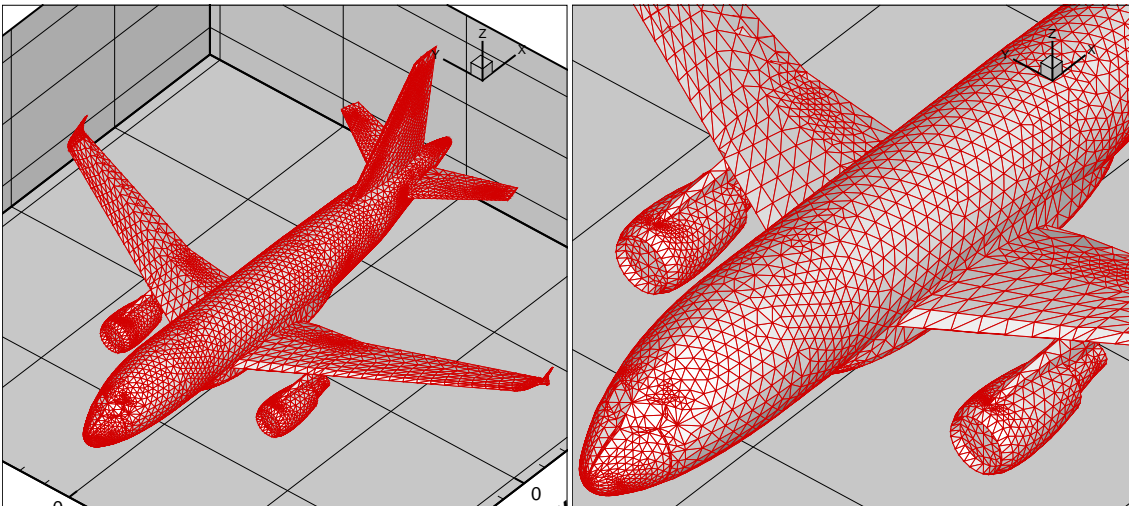


FIG. 2.11 – Diffraction d'une onde plane par un avion civil - Maillage de surface

Méthode	CFL- $\mathbb{P}_p$	Temps de calcul
DGTD- $\mathbb{P}_1$	0.3	2 h 26 mn
DGTD- $\mathbb{P}_2$	0.2	9 h 13 mn

TAB. 2.5 – Diffraction d'une onde plane par un avion civil -  $F=1$  GHz  
Cluster de 32 processeurs Intel Xeon 2.66 GHz et interconnexion Myrinet

## 2.5 Conclusion

Nous avons présenté dans ce chapitre la méthode Galerkin discontinue qui a servi de point de départ aux travaux décrits dans la suite de ce manuscrit. Initialement combinée à un schéma d'intégration en temps explicite saute-mouton du second ordre (LF2), cette méthode a été ici associée à un schéma de cette même famille précis au quatrième ordre (LF4). Cette combinaison permet de mieux exploiter l'ordre de précision arbitrairement élevé de la discrétisation en espace. En particulier, des expériences numériques sur la base d'un problème test modèle ont permis de montrer que la convergence numérique de la méthode LF4-DGTD- $\mathbb{P}_p$  est au moins d'ordre  $p$  pour  $p < 4$ . Par ailleurs, le schéma LF4 autorise un nombre CFL

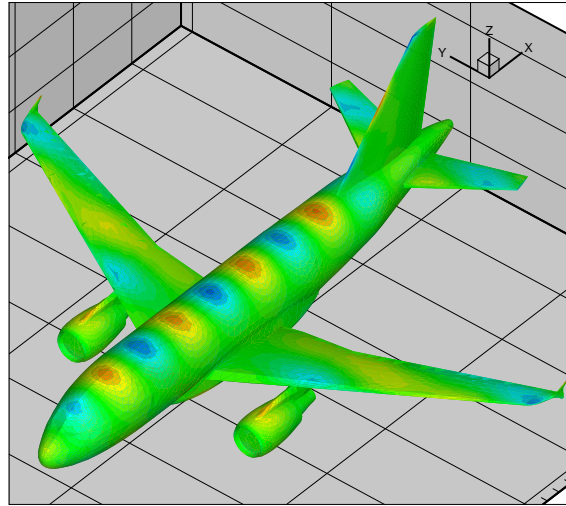


FIG. 2.12 – Diffraction d'une onde plane par un avion civil -  $F=1$  GHz  
Lignes de contour de la composante  $E_z$   
Méthode DGTD- $\mathbb{P}_1$

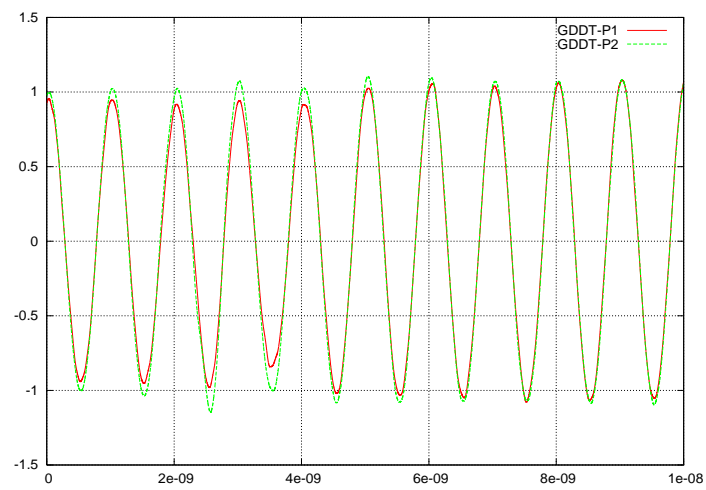


FIG. 2.13 – Diffraction d'une onde plane par un avion civil -  $F=1$  GHz  
Evolution temporelle de  $E_z$  en un point du domaine (abscisse : temps en sec - ordonnée :  $E_z$ )

2.85 plus élevé que celui du schéma LF2 et, en dépit d'une complexité arithmétique accrue, la méthode LF4-DGTD- $\mathbb{P}_p$  est moins coûteuse en temps de calcul que la méthode LF2-DGTD- $\mathbb{P}_p$ .



# Chapitre 3

## Méthodes GDDT implicites

### 3.1 Introduction

Dans le chapitre précédent, nous avons formulé deux méthodes DGTD- $\mathbb{P}_k$  explicites basées sur des schémas d'intégration en temps de type saute-mouton du second ordre et du quatrième ordre. Ces méthodes DGTD- $\mathbb{P}_k$  sont non-dissipatives et stables sous une condition de type CFL. Cependant, lorsque les simulations numériques mettent en jeu des maillages non-uniformes (localement raffinés), cette condition de stabilité impose un pas de temps très pénalisant, c'est-à-dire conduisant à des temps de calcul importants voire prohibitifs pour certains problèmes tridimensionnels. On distingue essentiellement deux stratégies pour palier ce problème :

- la première consiste à autoriser l'utilisation d'un pas de temps localement adapté à la taille des éléments du maillage. De telles stratégies sont étudiées dans [Fumeaux et al., 2004]-[Piperno, 2006b]-[Piperno, 2006a]-[Montseny et al., 2008].
- la seconde s'appuie sur l'adoption d'un schéma d'intégration implicite. C'est cette seconde stratégie que nous étudions dans ce chapitre.

Les schémas implicites sont connus pour leur large domaine de stabilité [Gear, 1971]. Néanmoins, les schémas implicites nécessitent la résolution d'un ou plusieurs systèmes linéaires à chaque itération en temps, induisant des surcoûts importants en termes de temps de calcul et d'occupation mémoire par rapport aux schémas explicites. Le choix de la méthode de résolution, directe (LU) ou itérative avec préconditionnement, joue un rôle primordial dans l'efficacité de la méthode numérique résultante. Un chapitre de cette thèse (chapitre 4) est entièrement consacré à une étude comparative de ces deux options de résolution dans un cadre bidimensionnel.

La littérature traitant de schémas implicites est assez réduite pour les équations de Maxwell. Namiki [Namiki, 1999, Namiki, 2000] a proposé une méthode DFDT implicite basée sur une technique de factorisation ADI (*Alternate Direction Implicit*) qui permet de se ramener à la résolution de systèmes linéaires tri-diagonaux. Nous nous sommes, pour notre part, intéressés au schéma Crank-Nicolson pour sa facilité de mise en œuvre et sa A-stabilité. Ce schéma a de plus fait l'objet de nombreux travaux et a été appliqué à de nombreux problèmes [Dubois et al., 2005]-[He and Sun, 2007]. Il est comparé à la méthode ADI pour la propagation d'onde électromagnétiques bidimensionnelles dans [Sun and Trueman, 2003] où il est démontré que la méthode DFDT implicite basée sur ce schéma est plus efficace que son homologue basée sur la technique ADI.

Ce chapitre se décompose en deux parties traitant de méthodes DGTD- $\mathbb{P}_k$  implicites respectivement du second ordre et du quatrième ordre en temps.

Dans un premier temps nous étudions une méthode DGTD- $\mathbb{P}_k$  implicite basée sur le schéma de Crank-Nicolson. La méthode DGTD- $\mathbb{P}_k$  résultante est précise au second ordre en temps, non-dissipative et inconditionnellement stable. Nous détaillons la formulation de cette méthode en trois dimensions d'espace puis nous étudions quelques unes de ses propriétés. Nous réalisons aussi une analyse de dispersion numérique en une dimension d'espace. La méthode est ensuite évaluée en deux dimensions d'espace et en maillages triangulaires, sur la base de différents cas tests de propagation en milieux homogène et hétérogène.

Nous étudions ensuite la possibilité de construire une méthode DGTD- $\mathbb{P}_k$  implicite précise au quatrième ordre en temps en exploitant une technique de défaut corrigé. Il s'agit plus ici d'une étude de faisabilité d'une telle approche, que du développement d'une méthode DGTD- $\mathbb{P}_k$  implicite d'intérêt pratique. En effet, une étude numérique montre que la technique employée permet effectivement d'améliorer la précision de la méthode DGTD- $\mathbb{P}_k$  implicite basée sur le schéma d'intégration en temps de Crank-Nicolson mais au détriment d'un coût par itération en temps prohibitif.

## 3.2 Méthode DGTD- $\mathbb{P}_k$ implicite du second ordre

Cette section est consacrée à l'étude d'une méthode DGTD implicite basée sur le schéma de Crank-Nicolson pour l'intégration en temps. La méthode DGTD résultante est non-dissipative (si  $\mathcal{F}_a = \emptyset$ ), inconditionnellement stable et précise au second ordre en temps du fait de l'utilisation du schéma de Crank-Nicolson.

### 3.2.1 Formulation

Partant de la formulation I (2.20),  $\forall \tau_i \in \mathcal{T}_h$  on obtient alors l'équation discrète suivante :

$$\begin{aligned}
2\mathcal{Y}_i \left( \frac{\mathbf{W}_i^{n+1} - \mathbf{W}_i^n}{\Delta t} \right) &+ \sum_{k=1}^3 \mathcal{Y}_i^{x_k} \left( \frac{\mathbf{W}_i^{n+1} + \mathbf{W}_i^n}{2} \right) &+ \sum_{a_{ij} \in \mathcal{F}_a^i} \mathcal{Y}_{ij} \left( \frac{\mathbf{W}_j^{n+1} + \mathbf{W}_j^n}{2} \right) \\
&+ \sum_{a_{ij} \in \mathcal{F}_m^i} \mathcal{Y}_{im} \left( \frac{\mathbf{W}_i^{n+1} + \mathbf{W}_i^n}{2} \right) &+ \sum_{a_{ij} \in \mathcal{F}_a^i} \mathcal{Y}_{ia} \left( \frac{\mathbf{W}_i^{n+1} + \mathbf{W}_i^n}{2} \right) \quad (3.1) \\
&= - \sum_{a_{ij} \in \mathcal{F}_a^i} \mathcal{Y}_{i\infty} \overline{\mathbf{W}}_i^\infty,
\end{aligned}$$

où  $\overline{\mathbf{W}}_i^\infty$  désigne une évaluation de  $\mathbf{W}_i^\infty$  en  $t^{n+\frac{1}{2}} = t^0 + (n + \frac{1}{2})\Delta t$ .



Posons  $\sigma = \frac{4}{\Delta t}$ . Alors  $\forall \tau_i \in \mathcal{T}_h$ , l'équation (3.1) devient :

$$\begin{aligned}
\sigma \mathcal{Y}_i \mathbf{W}_i^{n+1} &+ \sum_{k=1}^3 \mathcal{Y}_i^{x_k} \mathbf{W}_i^{n+1} &+ \sum_{a_{ij} \in \mathcal{F}_d^i} \mathcal{Y}_{ij} \mathbf{W}_j^{n+1} &+ \sum_{a_{ij} \in \mathcal{F}_m^i} \mathcal{Y}_{im} \mathbf{W}_i^{n+1} \\
&+ \sum_{a_{ij} \in \mathcal{F}_a^i} \mathcal{Y}_{ia} \mathbf{W}_i^{n+1} &= \\
\sigma \mathcal{Y}_i \mathbf{W}_i^n &- \sum_{k=1}^3 \mathcal{Y}_i^{x_k} \mathbf{W}_i^n &- \sum_{a_{ij} \in \mathcal{F}_d^i} \mathcal{Y}_{ij} \mathbf{W}_j^n &- \sum_{a_{ij} \in \mathcal{F}_m^i} \mathcal{Y}_{im} \mathbf{W}_i^n \\
&- \sum_{a_{ij} \in \mathcal{F}_a^i} \mathcal{Y}_{ia} \mathbf{W}_i^n &- 2 \sum_{a_{ij} \in \mathcal{F}_a^i} \mathcal{Y}_{i\infty} \overline{\mathbf{W}}_i^\infty.
\end{aligned} \tag{3.2}$$

Chaque itération en temps de la méthode DGTD- $\mathbb{P}_k$  implicite résultante nécessite donc la résolution d'un système linéaire :

$$\mathbb{P} \mathbf{W}^{n+1} = \mathbb{R} \mathbf{W}^n + b, \tag{3.3}$$

avec  $\forall \tau_i \in \mathcal{T}_h$  :

$$\left\{ \begin{array}{l}
\mathbb{P}_{ii} = \sigma \mathcal{Y}_i + \sum_{k=1}^3 \mathcal{Y}_i^{x_k} + \sum_{a_{ij} \in \mathcal{F}_m^i} \mathcal{Y}_{im} + \sum_{a_{ij} \in \mathcal{F}_a^i} \mathcal{Y}_{ia}, \\
\mathbb{R}_{ii} = \sigma \mathcal{Y}_i - \sum_{k=1}^3 \mathcal{Y}_i^{x_k} - \sum_{a_{ij} \in \mathcal{F}_m^i} \mathcal{Y}_{im} - \sum_{a_{ij} \in \mathcal{F}_a^i} \mathcal{Y}_{ia}, \\
\mathbb{P}_{ij} = \mathcal{Y}_{ij}, \text{ pour } a_{ij} \in \mathcal{F}_d^i, \\
\mathbb{R}_{ij} = -\mathbb{P}_{ij}, \\
b_i = -2 \sum_{a_{ij} \in \mathcal{F}_a^i} \mathcal{Y}_{i\infty} \overline{\mathbf{W}}_i^\infty.
\end{array} \right. \tag{3.4}$$

### 3.2.2 Propriétés

Nous étudions ici quelques propriétés de la méthode DGTD- $\mathbb{P}_k$  implicite du second ordre (3.3)-(3.4). On précise tout d'abord quelques propriétés de certaines des matrices élémentaires intervenant dans la définition de cette méthode.

1.  $N_{ij}$  est une matrice antisymétrique et  $N_{ij} = -N_{ji}$ .
2.  $N_{ij}^2$  est une matrice symétrique négative et admet les valeurs propres  $\{0, -1, -1\}$ .
3.  $P_{ia} = \begin{pmatrix} -z_i^{-1} N_{ia}^2 & 0_3 \\ 0_3 & -z_i N_{ia}^2 \end{pmatrix}$  est une matrice symétrique positive.

**Preuve** (de 3.)

Avec la symétrie de  $N_{ij}^2$ , on a :

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

De plus,  $\forall \mathbf{X} : {}^t\mathbf{X}P_{ia}\mathbf{X} = -z_i^{-1} {}^t\mathbf{X}_1 N_{ia}^2 \mathbf{X}_1 - z_i {}^t\mathbf{X}_2 N_{ia}^2 \mathbf{X}_2$  et  $N_{ia}^2$  est réelle symétrique négative, donc :

$${}^t\mathbf{X}_k N_{ia}^2 \mathbf{X} \leq 0 \text{ pour } k \in \{1; 2\} \Leftrightarrow {}^t\mathbf{X}_k P_{ia} \mathbf{X} \geq 0 \text{ (car } z_j > 0).$$

■

### 3.2.2.1 Inversibilité de la matrice implicite en 3D

On vérifie ici la solvabilité du système implicite (3.3)-(3.4).

**Proposition 1** *La matrice  $\mathbb{P}$  caractérisant le système implicite (3.3)-(3.4) est inversible.*

**Preuve** Pour étudier l'inversibilité de la matrice  $\mathbb{P}$ , nous décomposons celle-ci sous la forme  $\mathbb{D} + \mathbb{M}$  avec :

$$\mathbb{D}_{ii} = \sigma \mathcal{Y}_i + \sum_{a_{ij} \in \mathcal{F}_a^i} \mathcal{Y}_{ia},$$

et  $\mathbb{M} = \mathbb{P} - \mathbb{D}$ . Notons tout d'abord que  $\mathbb{D}$  est une matrice par bloc symétrique définie positive car la somme d'une matrice symétrique définie positive (i.e.  $\sigma \mathcal{Y}_i$ ) et d'une matrice positive (i.e.  $\sum_{a_{ij} \in \mathcal{F}_a^i} \mathcal{Y}_{ia}$ ).

D'autre part,  $\mathbb{M}$  peut se décomposer comme  $\mathbb{M} = \mathbb{M}_d + \mathbb{M}_m$  où  $\mathbb{M}_m = \sum_{a_{ij} \in \mathcal{F}_m^i} \mathcal{Y}_{im}$  est une matrice

diagonale par bloc antisymétrique, et  $\mathbb{M}_d$  est une matrice antisymétrique constituée des blocs  $\mathcal{Y}_{ij}$  sachant que  $\mathcal{Y}_{ij}$  est une matrice symétrique et que  $\mathcal{Y}_{ji} = -\mathcal{Y}_{ij}$ . Pour montrer l'inversibilité de  $\mathbb{P}$ , nous montrons que  $\mathbb{P}\mathbf{X} = 0 \Rightarrow \mathbf{X} = 0$  ou encore,  ${}^t\mathbf{X}\mathbb{P}\mathbf{X} = 0 \Rightarrow \mathbf{X} = 0$ . Cette dernière relation conduit à :

$${}^t\mathbf{X}(\mathbb{D} + \mathbb{M})\mathbf{X} = {}^t\mathbf{X}\mathbb{D}\mathbf{X} + {}^t\mathbf{X}\mathbb{M}\mathbf{X}.$$

Puisque  $\mathbb{M}$  est antisymétrique, on a  ${}^t\mathbf{X}\mathbb{M}\mathbf{X} = 0 \forall \mathbf{X}$ . Alors, nous devons juste vérifier que  ${}^t\mathbf{X}\mathbb{D}\mathbf{X} = 0 \Rightarrow \mathbf{X} = 0$  ce qui est justement le cas puisque  $\mathbb{D}$  est symétrique définie positive. ■

### 3.2.2.2 Stabilité par étude énergétique en 3D

Dans cette section, on suppose que le champ incident  $\mathbf{W}_i^\infty$  est nul et on étudie la stabilité de la méthode DGTD implicite du second ordre (3.3)-(3.4) en démontrant la conservation ou la décroissance d'une certaine énergie électromagnétique discrète.

**Proposition 2** *L'énergie électromagnétique discrète :*

$$\mathcal{E}^n = \frac{1}{2} \sum_{i=1}^N [\varepsilon_i {}^t\mathbf{E}_i^n \Phi_i \mathbf{E}_i^n + \mu_i {}^t\mathbf{H}_i^n \Phi_i \mathbf{H}_i^n]$$

ou sous forme plus condensée :

$$\mathcal{E}^n = \frac{1}{2} \sum_{i=1}^N [{}^t \mathbf{W}_i^n \mathcal{Y}_i \mathbf{W}_i^n],$$

est décroissante au cours du temps, c'est à dire que  $\mathcal{E}^{n+1} \leq \mathcal{E}^n$ . Si le domaine de calcul ne possède pas de frontières absorbantes (i.e. si  $\mathcal{F}_a = \emptyset$ ), cette énergie est exactement conservée.

**Preuve** Si nous multiplions l'équation (3.1) par  ${}^t \mathbf{W}_i^{n+\frac{1}{2}} = {}^t \left( \frac{\mathbf{W}_i^{n+1} + \mathbf{W}_i^n}{2} \right)$ , nous avons :

$$\begin{aligned} 2{}^t \mathbf{W}_i^{n+\frac{1}{2}} \mathcal{Y}_i \left( \frac{\mathbf{W}_i^{n+1} - \mathbf{W}_i^n}{\Delta t} \right) &+ \sum_{k=1}^3 {}^t \mathbf{W}_i^{n+\frac{1}{2}} \mathcal{Y}_i^{x_k} \mathbf{W}_i^{n+\frac{1}{2}} \\ &+ \sum_{a_{ij} \in \mathcal{F}_d^i} {}^t \mathbf{W}_i^{n+\frac{1}{2}} \mathcal{Y}_{ij} \mathbf{W}_j^{n+\frac{1}{2}} \\ &+ \sum_{a_{ij} \in \mathcal{F}_m^i} {}^t \mathbf{W}_i^{n+\frac{1}{2}} \mathcal{Y}_{im} \mathbf{W}_i^{n+\frac{1}{2}} \\ &+ \sum_{a_{ij} \in \mathcal{F}_a^i} {}^t \mathbf{W}_i^{n+\frac{1}{2}} \mathcal{Y}_{ia} \mathbf{W}_i^{n+\frac{1}{2}} = 0. \end{aligned} \quad (3.5)$$

Pour commencer, on a la relation :

$${}^t \mathbf{W}_i^{n+\frac{1}{2}} \left( \sum_{k=1}^3 \mathcal{Y}_i^{x_k} \right) \mathbf{W}_i^{n+\frac{1}{2}} = 0,$$

puisque la matrice  $\sum_{k=1}^3 \mathcal{Y}_i^{x_k}$  est antisymétrique. D'autre part :

$${}^t \mathbf{W}_j^{n+\frac{1}{2}} \mathcal{Y}_{ji} \mathbf{W}_i^{n+\frac{1}{2}} = -{}^t \mathbf{W}_j^{n+\frac{1}{2}} \mathcal{Y}_{ij} \mathbf{W}_i^{n+\frac{1}{2}} = -{}^t \mathbf{W}_i^{n+\frac{1}{2}} \mathcal{Y}_{ij} \mathbf{W}_j^{n+\frac{1}{2}},$$

en utilisant de nouveau la symétrie de  $\mathcal{Y}_{ij}$  et  $\mathcal{Y}_{ji} = -\mathcal{Y}_{ij}$  par construction. Par suite, lorsque l'on somme les équations (3.5) pour  $i = 1, \dots, N$ , les termes de flux internes s'éliminent deux à deux. Enfin, la contribution des frontières métalliques de l'équation (3.5) est nulle car  $\mathcal{Y}_{im}$  est antisymétrique. Donc, en sommant sur toutes les cellules, nous obtenons :

$$\mathcal{E}^{n+1} - \mathcal{E}^n = 0, \quad (3.6)$$

si  $\mathcal{F}_a = \emptyset$  et la méthode DGTD implicite du second ordre (3.3)-(3.4) est inconditionnellement stable. Lorsque  $\mathcal{F}_a \neq \emptyset$ ,  $\mathcal{Y}_{ia}$  étant une matrice positive,  ${}^t \mathbf{W}_i^{n+\frac{1}{2}} \mathcal{Y}_{ia} \mathbf{W}_i^{n+\frac{1}{2}}$  est positif et par conséquent

$$\mathcal{E}^{n+1} - \mathcal{E}^n \leq 0, \quad (3.7)$$

dont on déduit là encore la stabilité inconditionnelle de la méthode DGTD implicite du second ordre (3.3)-(3.4). ■

### 3.2.2.3 Dispersion numérique en 1D

La méthode DGTD implicite du second ordre (3.3)-(3.4) étant inconditionnellement stable, nous n'avons pas de contrainte sur le choix du pas de temps. Néanmoins, un pas de temps trop important peut augmenter l'erreur numérique de manière significative. Afin d'évaluer l'influence de ce paramètre, nous procédons ici à une analyse de l'erreur numérique. La méthode DGTD implicite considérée dans cette section étant non-dissipative, nous devons considérer l'erreur de dispersion. Plusieurs publications traitent de l'analyse de l'erreur numérique de méthodes DGTD comme [Sarmany et al., 2007] où les auteurs utilisent un schéma de type Runge-Kutta pour l'intégration en temps dans le cadre de la résolution numérique des équations de Maxwell 1D et 2D. Atkins et Hu réalisent une analyse de dispersion en 1D [Atkins and Hu, 2002a] et en 2D [Atkins and Hu, 2002b] sur un système d'équations hyperboliques. Toutes ces études ont pour point commun la résolution d'un problème aux valeurs propres pour déterminer l'évolution de l'erreur en fonction des pas de discrétisation en espace et en temps. Ainsworth [Ainsworth, 2004] propose une étude analytique pour obtenir des estimations des taux de convergence des termes d'erreur de dispersion et de dissipation d'une méthode Galerkin discontinue d'ordre élevé.

Nous allons effectuer l'analyse de dispersion dans le cas mono-dimensionnel pour des raisons de commodité. Nous décrivons tout d'abord le schéma DGTD 1D avant de détailler l'étude de dispersion à proprement dite.

**Equations de Maxwell 1D.** Afin d'obtenir une version mono-dimensionnelle des équations de Maxwell (1.12) nous considérons comme direction de propagation  $\mathbf{k} = (k, 0, 0)^t$  avec une polarisation du vecteur champ électrique telle que  $\mathbf{E} = (0, 0, E_z)^t$ . La polarisation du vecteur champ magnétique se déduit du produit vectoriel  $\mathbf{k} \times \mathbf{E}$  menant à  $\mathbf{H} = (0, H_y, 0)^t$ . De plus, nous supposons que  $E_z$  et  $H_y$  sont fonctions de  $x$  et  $t$ . Pour plus de clarté, nous allégeons les notations en remplaçant  $E_z$  par  $E$  et  $H_y$  par  $H$ . Les équations de Maxwell 1D sont alors données par :

$$\begin{cases} \varepsilon(x) \frac{\partial E}{\partial t} - \frac{\partial H}{\partial x} = 0, \\ \mu(x) \frac{\partial H}{\partial t} - \frac{\partial E}{\partial x} = 0. \end{cases} \quad (3.8)$$

Ce système peut être ré-écrit sous une forme pseudo-conservative :

$$QW_t + \partial_x F(W) = 0, \quad (3.9)$$

où  $W_t = \frac{\partial W}{\partial t}$ ,  $\partial_x F(W) = \frac{\partial F(W)}{\partial x}$  et :

$$Q = \begin{bmatrix} \varepsilon & 0 \\ 0 & \mu \end{bmatrix}, \quad W = \begin{pmatrix} E \\ H \end{pmatrix} \quad \text{et} \quad F(W) = \begin{pmatrix} -H \\ -E \end{pmatrix} = AW \quad \text{avec} \quad A = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}.$$

Les équations (3.9) sont résolues sur un intervalle  $[a, b]$ .

**Formulation faible et discrétisation.** L'intervalle  $[a, b]$  est discrétisé comme  $\bigcup_{i=1}^N C_i$  tel que  $C_i = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}]$  avec  $\Delta x_i = x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}}$  (voir la figure 3.1). Notons néanmoins que pour cette étude de dispersion nous simulons un domaine infini et donc, nous n'appliquons pas de conditions aux limites  $a$  et  $b$ , afin d'appliquer une analyse en ondes planes.

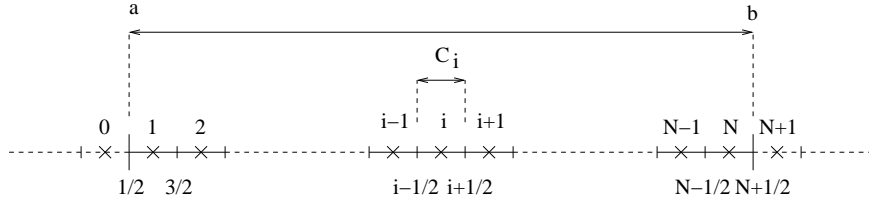


FIG. 3.1 – Discrétisation d'un intervalle 1D  $[a, b]$

En multipliant (3.9) par une fonction test  $\varphi$  et en intégrant sur la cellule  $C_i$ , nous obtenons :

$$\begin{aligned} & \int_{C_i} \varphi (QW_t + [F(W)]_x) dx = 0, \\ \Leftrightarrow & \int_{C_i} \varphi QW_t dx - \int_{C_i} \varphi_x F(W) dx + \varphi(x_{i+\frac{1}{2}})F(W(x_{i+\frac{1}{2}})) \\ & \quad - \varphi(x_{i-\frac{1}{2}})F(W(x_{i-\frac{1}{2}})) = 0. \end{aligned} \tag{3.10}$$

Nous introduisons les notations suivantes :

$$F_{i\pm\frac{1}{2}} \equiv F(W(x_{i\pm\frac{1}{2}})) \quad \text{et} \quad \varphi_i^\pm \equiv \varphi(x_{i\pm\frac{1}{2}}).$$

Avec ces notations, l'équation (3.10) devient :

$$\int_{C_i} \varphi Q W_t dx - \int_{C_i} \varphi_x F(W) dx + \varphi_i^+ F_{i+\frac{1}{2}} - \varphi_i^- F_{i-\frac{1}{2}} = 0. \quad (3.11)$$

Une formulation Galerkin discontinue est basée sur l'introduction d'un ensemble de fonctions de base locales  $(\varphi_{ij})_{1 \leq j \leq d_i}$  pour chaque cellule  $(C_i)_{1 \leq i \leq N}$  où  $d_i$  représente le nombre de degrés de liberté caractérisant l'approximation des variables d'état dans chaque cellule. La restriction du champ  $W(x, t)$  à  $C_i$  est ici approchée par une combinaison linéaire des fonctions de base locales polynomiales de degré au plus  $p_i$ , supposées linéairement indépendantes qui engendrent un espace noté  $\mathcal{P}_i = \mathbb{P}_{p_i}[C_i]$  :

$$W_i(x, t) \equiv W(x, t)|_{C_i} = \sum_{j=1}^{d_i} W_{ij}(t) \varphi_{ij}(x), \quad (3.12)$$

où  $W_i$  est la projection orthogonale de  $W$  sur  $\mathcal{P}_i$ . Comme au chapitre 2, on présente la formulation de la méthode DGTD- $\mathbb{P}_{p_i}$  puis, pour l'étude de dispersion, on ne considère que le cas  $p_i = p$  et  $d_i = d$ ,  $\forall C_i$ . Les degrés de liberté locaux sont notés  $W_{ij}$ . Nous notons  $\phi_i = (\varphi_{i1}, \varphi_{i2}, \dots, \varphi_{id_i})$  la base locale de  $\mathcal{P}_i$ . Il s'agit maintenant d'injecter  $W_i$  dans (3.11). Notre problème est alors de chercher  $W_i \in \mathcal{P}_i$  tel que :

$$\int_{C_i} \varphi Q_i(W_i)_t dx - \int_{C_i} \varphi_x F(W)_i dx + \varphi_i^+ F_{i+\frac{1}{2}} - \varphi_i^- F_{i-\frac{1}{2}} = 0, \quad \forall \varphi \in \mathcal{P}_i, \quad (3.13)$$

où :

$$Q_i = \begin{bmatrix} \varepsilon_i & 0 \\ 0 & \mu_i \end{bmatrix}, \quad \varepsilon_i = \varepsilon|_{C_i}, \quad \mu_i = \mu|_{C_i}.$$

Pour une cellule  $i$ , avec  $2 \leq i \leq N$ , on utilise un schéma centré pour le calcul des flux  $F_{i\pm\frac{1}{2}}$ , c'est-à-dire :

$$F_{i\pm\frac{1}{2}} = \frac{1}{2} \left( AW_{i\pm 1}^{i\pm\frac{1}{2}} + AW_i^{i\pm\frac{1}{2}} \right) \quad \text{avec} \quad W_i^{i\pm\frac{1}{2}} \equiv W_i(x_{i\pm\frac{1}{2}}),$$

ce qui conduit à :

$$\begin{aligned} \int_{C_i} \varphi Q_i(W_i)_t dx - \int_{C_i} \varphi_x AW_i dx &+ \frac{\varphi_i^+}{2} \left( AW_{i+1}^{i+\frac{1}{2}} + AW_i^{i+\frac{1}{2}} \right) \\ &- \frac{\varphi_i^-}{2} \left( AW_{i-1}^{i-\frac{1}{2}} + AW_i^{i-\frac{1}{2}} \right) = 0, \quad \forall \varphi \in \mathcal{P}_i, \end{aligned} \quad (3.14)$$

c'est-à-dire :

$$\begin{aligned} \int_{C_i} \varphi Q_i(W_i)_t dx - \int_{C_i} \varphi_x AW_i dx &+ \frac{\varphi_i^+}{2} AW_{i+1}^{i+\frac{1}{2}} \\ &+ \frac{1}{2} [\varphi AW_i]_{i-\frac{1}{2}}^{i+\frac{1}{2}} - \frac{\varphi_i^-}{2} AW_{i-1}^{i-\frac{1}{2}} = 0, \quad \forall \varphi \in \mathcal{P}_i. \end{aligned} \quad (3.15)$$

En utilisant une formule d'intégration par parties :

$$[\varphi AW_i]_{i-\frac{1}{2}}^{i+\frac{1}{2}} = \int_{C_i} \varphi_x AW_i dx + \int_{C_i} \varphi A(W_i)_x dx,$$

l'équation (3.15) devient :

$$\begin{aligned} \int_{C_i} \varphi Q_i(W_i)_t dx + \frac{1}{2} \int_{C_i} (\varphi A(W_i)_x - \varphi_x AW_i) dx \\ + \frac{\varphi_i^+}{2} AW_{i+1}^{i+\frac{1}{2}} - \frac{\varphi_i^-}{2} AW_{i-1}^{i-\frac{1}{2}} = 0, \quad \forall \varphi \in \mathcal{P}_i. \end{aligned} \quad (3.16)$$

**Étude du problème discret.** Nous désignons par  $\mathbb{W}_i = (W_{i1}, W_{i2}, \dots, W_{id_i})$  (ainsi  $\mathbb{W}_i$  est une matrice rectangulaire  $2 \times d_i$ ). La relation (3.12) peut se ré-écrire comme  $W_i = \mathbb{W}_i^t \phi_i$ . En utilisant cette expression dans (3.16), nous avons :

$$\begin{aligned} \int_{C_i} Q_i(\mathbb{W}_i)_t^t \phi_i \varphi dx + \frac{1}{2} \int_{C_i} (A \mathbb{W}_i^t (\phi_i)_x \varphi - A \mathbb{W}_i^t \phi_i \varphi_x) dx \\ + \frac{1}{2} A \mathbb{W}_{i+1}^t \phi_{i+1}(x_{i+\frac{1}{2}}) \varphi_i^+ \\ - \frac{1}{2} A \mathbb{W}_{i-1}^t \phi_{i-1}(x_{i-\frac{1}{2}}) \varphi_i^- = 0, \quad \forall \varphi \in \mathcal{P}_i. \end{aligned} \quad (3.17)$$

Nous introduisons les notations suivantes :

$$\left\{ \begin{array}{l} \Phi_i = \int_{C_i} {}^t \phi_i \phi_i dx = {}^t \Phi_i, \\ \Phi'_i = \int_{C_i} ({}^t \phi_i (\phi_i)_x - {}^t (\phi_i)_x \phi_i) dx, \\ \Phi_i^\pm = {}^t \phi_i^\pm \phi_i^\pm = {}^t \phi_i(x_{i\pm\frac{1}{2}}) \phi_i(x_{i\pm\frac{1}{2}}) = {}^t \Phi_i^\pm, \\ \Phi_{i\pm 1, i}^\pm = {}^t \phi_{i\pm 1}^\pm \phi_i^\pm = {}^t \phi_{i\pm 1}(x_{i\pm\frac{1}{2}}) \phi_i(x_{i\pm\frac{1}{2}}) = {}^t \Phi_{i, i\pm 1}^\pm. \end{array} \right. \quad (3.18)$$

Maintenant, si dans (3.17)  $\varphi$  est remplacé par toutes les fonctions de base  $\varphi_{ij}$  pour  $1 \leq j \leq d_i$ , on obtient :

$$2Q_i(\mathbb{W}_i)_t \Phi_i + A \mathbb{W}_i^t \Phi'_i + A \mathbb{W}_{i+1}^t \Phi_{i+1, i}^+ - A \mathbb{W}_{i-1}^t \Phi_{i-1, i}^- = 0. \quad (3.19)$$

Les inconnues étant traditionnellement sous forme de vecteurs, nous ré-écrivons le système d'équations (3.19) en vectorisant la matrice  $\mathbb{W}$  :

$$2\mathcal{Y}_i(\mathbf{W}_i)_t + \mathcal{Y}'_i \mathbf{W}_i + \mathcal{Y}_i^+ \mathbf{W}_{i+1} - \mathcal{Y}_i^- \mathbf{W}_{i-1} = 0, \quad (3.20)$$

avec :

$$\left\{ \begin{array}{l} \mathcal{Y}_i = \begin{bmatrix} (\Phi_i)_{11}Q_i & (\Phi_i)_{12}Q_i & \cdots & (\Phi_i)_{1d_i}Q_i \\ (\Phi_i)_{21}Q_i & (\Phi_i)_{22}Q_i & \cdots & (\Phi_i)_{2d_i}Q_i \\ \vdots & \vdots & & \vdots \\ (\Phi_i)_{d_i1}Q_i & (\Phi_i)_{d_i2}Q_i & \cdots & (\Phi_i)_{d_id_i}Q_i \end{bmatrix}, \\ \mathcal{Y}'_i = \begin{bmatrix} (\Phi'_i)_{11}A & (\Phi'_i)_{12}A & \cdots & (\Phi'_i)_{1d_i}A \\ (\Phi'_i)_{21}A & (\Phi'_i)_{22}A & \cdots & (\Phi'_i)_{2d_i}A \\ \vdots & \vdots & & \vdots \\ (\Phi'_i)_{d_i1}A & (\Phi'_i)_{d_i2}A & \cdots & (\Phi'_i)_{d_id_i}A \end{bmatrix}, \\ \mathcal{Y}_i^\pm = \begin{bmatrix} (\Phi_{i\pm 1,i}^\pm)_{11}A & (\Phi_{i\pm 1,i}^\pm)_{12}A & \cdots & (\Phi_{i\pm 1,i}^\pm)_{1d_j}A \\ (\Phi_{i\pm 1,i}^\pm)_{21}A & (\Phi_{i\pm 1,i}^\pm)_{22}A & \cdots & (\Phi_{i\pm 1,i}^\pm)_{2d_j}A \\ \vdots & \vdots & & \vdots \\ (\Phi_{i\pm 1,i}^\pm)_{d_i1}A & (\Phi_{i\pm 1,i}^\pm)_{d_i2}A & \cdots & (\Phi_{i\pm 1,i}^\pm)_{d_id_j}A \end{bmatrix}. \end{array} \right.$$

**Intégration en temps implicite.** Nous appliquons le schéma de Crank-Nicolson à l'équation (3.20), ce qui nous donne :

$$\begin{aligned} 2\mathcal{Y}_i \left( \frac{\mathbf{W}_i^{n+1} - \mathbf{W}_i^n}{\Delta t} \right) + \mathcal{Y}'_i \left( \frac{\mathbf{W}_i^{n+1} + \mathbf{W}_i^n}{2} \right) + \mathcal{Y}_i^+ \left( \frac{\mathbf{W}_{i+1}^{n+1} + \mathbf{W}_{i+1}^n}{2} \right) \\ - \mathcal{Y}_i^- \left( \frac{\mathbf{W}_{i-1}^{n+1} + \mathbf{W}_{i-1}^n}{2} \right) = 0. \end{aligned} \quad (3.21)$$

**Intégration en temps explicite.** Nous appliquons le schéma saute-mouton à l'équation (3.20), et l'on obtient :

$$2\mathcal{Y}_i \left( \frac{\mathbf{S}_i^{n+\frac{1}{2}} - \mathbf{S}_i^{n-\frac{1}{2}}}{\Delta t} \right) + \mathcal{Y}'_i \mathbf{S}_i^n + \mathcal{Y}_i^+ \mathbf{S}_{i+1}^n - \mathcal{Y}_i^- \mathbf{S}_{i-1}^n = 0, \quad (3.22)$$

avec :

$$\mathbf{S}_i^{n+\frac{1}{2}} = \begin{pmatrix} E_{i,1}^n \\ \vdots \\ E_{i,d_i}^n \\ H_{i,1}^{n+\frac{1}{2}} \\ \vdots \\ H_{i,d_i}^{n+\frac{1}{2}} \end{pmatrix}.$$



**Dispersion numérique pour la méthode DGTD implicite.** Nous nous plaçons dans le cas d'un milieu infini 1D. Nous partons du schéma (3.21) avec  $\varepsilon_i = \varepsilon$  et  $\mu_i = \mu$  ( $\varepsilon$  et  $\mu$  sont constants sur tout le domaine) et sur un maillage uniforme ( $\Delta x_i = \Delta x$ ). Nous supposons aussi que  $p_i = p$  et  $d_i = d \forall C_i$ . Nous cherchons la solution de (3.21) sous la forme d'une onde plane  $\mathbf{V}_i = \mathbf{G}_i e^{i\omega_n \Delta t + ik \Delta x}$  avec  $\mathbf{V}_i = (E_{i,1}, \dots, E_{i,d_i}, H_{i,1}, \dots, H_{i,d_i})$ ,  $\mathbf{i}^2 = -1$  et où  $\mathbf{G}_i$  est l'amplitude du vecteur. En injectant  $\mathbf{V}_i$  dans (3.21), nous obtenons le système suivant :

$$\begin{aligned} \left( 2\mathcal{Y}_i \left( \frac{e^{i\omega_h \Delta t} - 1}{\Delta t} \right) + \mathcal{Y}'_i \left( \frac{e^{i\omega_h \Delta t} + 1}{2} \right) \right) + \mathcal{Y}_i^+ \left( \frac{e^{i\omega_h \Delta t} + 1}{2} \right) e^{ik \Delta x} - \\ - \mathcal{Y}_i^- \left( \frac{e^{i\omega_h \Delta t} + 1}{2} \right) e^{-ik \Delta x} \mathbf{V}_i = 0, \end{aligned} \quad (3.23)$$

ou sous forme plus condensée :

$$\frac{4\mathbf{i}}{\Delta t} \tan \left( \frac{\omega_h \Delta t}{2} \right) \mathcal{Y}_i \mathbf{V}_i = (-\mathcal{Y}'_i - e^{ik \Delta x} \mathcal{Y}_i^+ + e^{-ik \Delta x} \mathcal{Y}_i^-) \mathbf{V}_i, \quad (3.24)$$

avec  $\omega_h$ , la valeur discrète de  $\omega$ . Soit  $\tilde{P}$  défini par :

$$\tilde{P} = \mathcal{Y}_i^{-1} (-\mathcal{Y}'_i - e^{ik \Delta x} \mathcal{Y}_i^+ + e^{-ik \Delta x} \mathcal{Y}_i^-).$$

En introduisant  $\tilde{P}$  dans l'équation (3.24), nous avons :

$$\frac{4\mathbf{i}}{\Delta t} \tan \left( \frac{\omega_h \Delta t}{2} \right) \mathbf{V}_i = \tilde{P} \mathbf{V}_i. \quad (3.25)$$

L'étape suivante consiste à décomposer  $\mathbf{V}_j$  dans une base de vecteurs propres de  $\tilde{P}$ . Par conséquent pour un vecteur propre  $\mathbf{X}_l$ , (3.25) devient :

$$\frac{4\mathbf{i}}{\Delta t} \tan \left( \frac{\omega_l \Delta t}{2} \right) = \lambda_l, \quad (3.26)$$

par suite :

$$\omega_l = \frac{2}{\Delta t} \arctan \left( \frac{\lambda_l \Delta t}{4\mathbf{i}} \right).$$

La relation de dispersion s'écrit comme suit :

$$q_{imp} = \frac{\omega_h}{kc} = \frac{2}{\bar{\omega}} \arctan \left( \frac{\lambda \Delta t}{4\mathbf{i}} \right), \quad (3.27)$$

avec  $\bar{\omega} = \omega \Delta t$ .

**Dispersion numérique pour la méthode DGTD explicite.** Nous appliquons  $\mathbf{V}_i$  à (3.22) et nous obtenons :

$$\left( 2\mathcal{Y}_i \left( \frac{e^{i\frac{\omega_h \Delta t}{2}} - e^{-i\frac{\omega_h \Delta t}{2}}}{\Delta t} \right) + \mathcal{Y}'_i + \mathcal{Y}_i^+ e^{ik \Delta x} - \mathcal{Y}_i^- e^{-ik \Delta x} \right) \mathbf{V}_i = 0, \quad (3.28)$$

ou sous forme plus condensée :

$$\frac{4\mathbf{i}}{\Delta t} \sin\left(\frac{\omega_h \Delta t}{2}\right) \mathcal{Y}_i \mathbf{V}_i = (-\mathcal{Y}'_i - e^{ik\Delta x} \mathcal{Y}_i^+ + e^{-ik\Delta x} \mathcal{Y}_i^-) \mathbf{V}_i. \quad (3.29)$$

En suivant la même méthode que pour le cas implicite, on obtient la relation de dispersion :

$$q_{exp} = \frac{\omega_h}{kc} = \frac{2}{\bar{\omega}} \arcsin\left(\frac{\lambda \Delta t}{4\mathbf{i}}\right). \quad (3.30)$$

**Étude numérique.** Les valeurs propres de  $\tilde{P}$ , notées  $\lambda$ , définies par l'équation (3.26) et nécessaires dans le calcul des relations de dispersion (3.27)-(3.30), sont obtenues à l'aide du logiciel de calcul symbolique ©Maple pour les méthodes DGTD- $\mathbb{P}_0$ , DGTD- $\mathbb{P}_1$  et DGTD- $\mathbb{P}_2$ . Nous donnons ci-dessous les valeurs propres de  $\tilde{P}$  pour des valeurs de  $K$  prises dans le voisinage de zéro :

– méthode DGTD- $\mathbb{P}_0$  :

$$\begin{cases} \frac{\lambda_1}{2} = 1 - \frac{1}{6}K^2 + \frac{1}{120}K^4 - \frac{1}{5040}K^6 + \mathcal{O}(K^8) \\ \frac{\lambda_2}{2} = -1 + \frac{1}{6}K^2 - \frac{1}{120}K^4 + \frac{1}{5040}K^6 + \mathcal{O}(K^8) \end{cases} \quad (3.31)$$

– méthode DGTD- $\mathbb{P}_1$  :

$$\begin{cases} \frac{\lambda_1}{2} = 1 + \frac{1}{48}K^2 - \frac{7}{15360}K^4 - \frac{599}{10321920}K^6 + \mathcal{O}(K^8) \\ \frac{\lambda_2}{2} = -1 - \frac{1}{48}K^2 + \frac{7}{15360}K^4 + \frac{599}{10321920}K^6 + \mathcal{O}(K^8) \end{cases} \quad (3.32)$$

– méthode DGTD- $\mathbb{P}_2$  :

$$\begin{cases} \frac{\lambda_1}{2} = 1 - \frac{1}{16800}K^6 - \frac{1}{162000}K^8 + \mathcal{O}(K^{10}) \\ \frac{\lambda_2}{2} = -1 + \frac{1}{16800}K^6 + \frac{1}{162000}K^8 + \mathcal{O}(K^{10}) \end{cases} \quad (3.33)$$

A partir des valeurs propres calculées à l'aide de ©Maple et de l'équation (3.27), nous obtenons pour la méthode DGTD- $\mathbb{P}_0$  :

$$\begin{aligned} q_{imp}(\alpha, K) &= 1 - \frac{1}{6} \left(1 + \frac{\alpha}{2}\right) K^2 \\ &+ \frac{1}{4} \left(\frac{1}{30} + \frac{1}{6}\alpha^2 + \frac{1}{20}\alpha^4\right) K^4 + \mathcal{O}(K^6), \end{aligned} \quad (3.34)$$

et avec l'équation (3.30), nous avons :

$$\begin{aligned} q_{exp}(\alpha, K) &= 1 + \frac{1}{24}(\alpha - 2)(\alpha + 2)K^2 \\ &+ \frac{1}{1920}(3\alpha + 2)(\alpha - 2)(\alpha + 2)(3\alpha - 2)K^4 + \mathcal{O}(K^6), \end{aligned} \quad (3.35)$$

où  $K = k\Delta x$  est le nombre d'onde numérique adimensionné et  $\alpha = \frac{c\Delta t}{\Delta x}$  le nombre CFL. Similairement, pour la méthode DGTD- $\mathbb{P}_1$  :

$$\begin{aligned} q_{imp}(\alpha, K) &= 1 - \frac{1}{24}K^2(2\alpha - 1)(2\alpha + 1) \\ &+ \frac{1}{23040}K^4(-11 - 320\alpha^2 + 736\alpha^4) + \mathcal{O}(K^6), \end{aligned} \quad (3.36)$$

et :

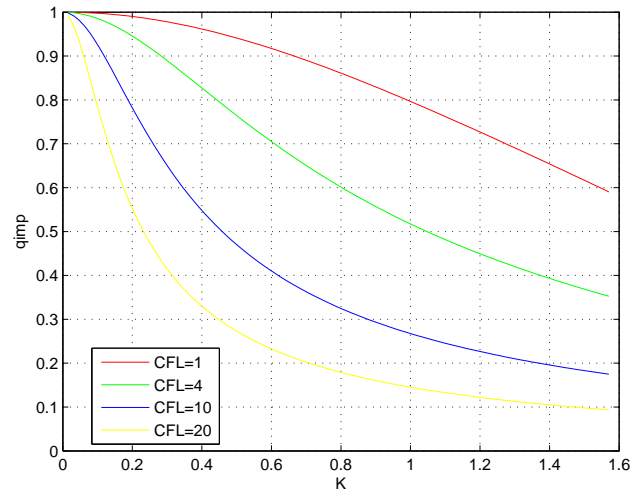
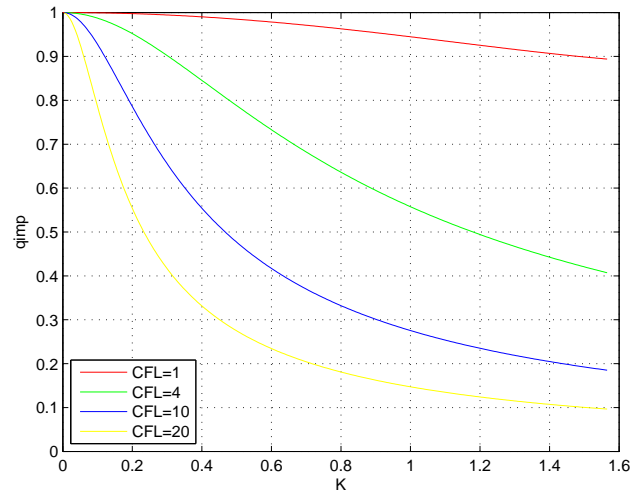
$$\begin{aligned} q_{exp}(\alpha, K) &= 1 + \frac{1}{24}K^2(2\alpha^2 + 1) \\ &+ \frac{1}{23040}K^4(4\alpha + 1)(4\alpha - 1)(16\alpha^2 + 11) + \mathcal{O}(K^6). \end{aligned} \quad (3.37)$$

Le logiciel ©Maple nous renvoyant pour la méthode DGTD- $\mathbb{P}_2$  un résultat sur plusieurs lignes, nous n'avons pas jugé nécessaire, dans un souci de simplicité, de détailler ici l'expression correspondante des quantités  $q_{imp}$  et  $q_{exp}$ .

**Analyse des résultats** Dans un premier temps, nous montrons sur les Fig. 3.2, 3.3 et 3.4 l'évolution de la relation de dispersion du schéma implicite ( $q_{imp}$ ) pour les ordres d'interpolation en espace 0, 1 et 2 en fonction du paramètre  $K$  pour différentes valeurs du nombre CFL (1, 4, 10 et 20). Nous considérons ici des valeurs de  $K$  comprises entre 0 et  $\frac{\pi}{2}$  pour des raisons de périodicité,  $q_{imp}$  étant défini à l'aide de la fonction arctan. Faire varier le paramètre  $K$  peut être vu comme une variation du nombre d'onde ou du pas d'espace, de par la définition du nombre d'onde numérique  $K = k\Delta x$ . Par ailleurs, on rappelle que pour le schéma explicite, la limite de stabilité de la méthode DGTD- $\mathbb{P}_k$  est obtenue avec un nombre CFL=1 pour  $p = 0$ , 0.3 pour  $p = 1$  et 0.2 pour  $p = 2$ . Le nombre CFL permet de relier la valeur du pas de temps au pas d'espace. Nous choisissons de n'utiliser que des valeurs du nombre CFL supérieures ou égales à la valeur maximale du nombre CFL autorisé pour la stabilité du schéma explicite de la méthode DGTD- $\mathbb{P}_0$ . On rappelle également que  $q_{imp} = \frac{\omega_h}{kc}$  où  $\omega_h$  désigne la pulsation discrète.

Nous constatons sur la Fig. 3.2, sans surprise, que l'erreur de dispersion croît lorsque l'on augmente d'une part le nombre CFL et d'autre part le nombre d'onde numérique. Lorsque l'on prend une valeur du CFL plus grande, le pas de temps augmente lui aussi et entraîne une dispersion numérique plus importante. Aussi, augmenter la longueur d'onde numérique revient à diminuer le nombre de points par longueur d'onde, c'est-à-dire à déraffiner le maillage, et l'erreur numérique croît. La dispersion numérique de la méthode DGTD- $\mathbb{P}_1$  implicite appelle les mêmes commentaires. Néanmoins, le niveau de dispersion numérique de cette méthode est notablement inférieur à celui caractérisant la méthode DGTD- $\mathbb{P}_0$  particulièrement pour CFL  $\approx 1$ . Pour les autres valeurs du nombre CFL, cette différence est moindre. On imagine qu'à partir d'une certaine valeur du nombre CFL, l'erreur numérique du schéma en temps est prépondérante par rapport à celle du schéma en espace. Enfin, les courbes pour la méthode DGTD- $\mathbb{P}_2$  implicite sont similaires à celles de la méthode DGTD- $\mathbb{P}_1$  implicite. Nous tirons les mêmes conclusions que précédemment, à savoir que l'erreur du schéma en temps est dans ce cas prédominante par rapport à celle du schéma en espace.

Afin de confirmer l'hypothèse précédente selon laquelle à partir d'une certaine valeur du nombre CFL l'erreur du schéma en temps est prépondérante par rapport à celle du schéma en espace, nous présentons sur la Fig. 3.5 l'évolution de  $q_{imp}$  pour les méthodes implicites DGTD- $\mathbb{P}_1$  et DGTD- $\mathbb{P}_2$  pour une valeur du nombre CFL=0.1. On observe ici que la dispersion de la méthode DGTD- $\mathbb{P}_1$  se dégrade alors que celle de la méthode DGTD- $\mathbb{P}_2$  reste minimale. Cette comparaison met bien en exergue que pour des petites valeurs

FIG. 3.2 – Dispersion de la méthode DGTD- $\mathbb{P}_0$  implicite pour différentes valeurs du nombre CFLFIG. 3.3 – Dispersion de la méthode DGTD- $\mathbb{P}_1$  implicite pour différentes valeurs du nombre CFL

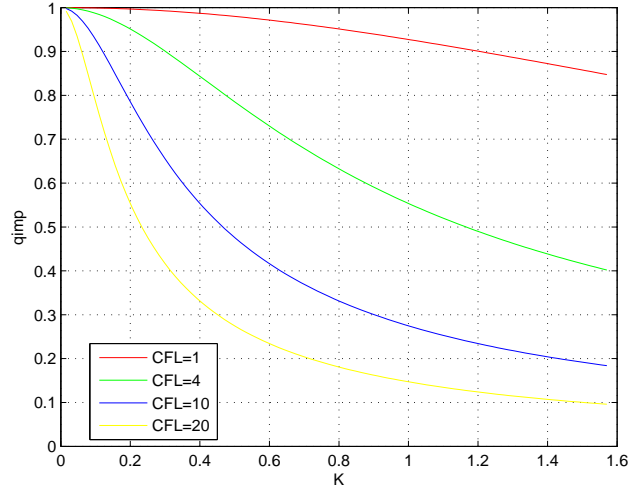


FIG. 3.4 – Dispersion de la méthode DGTD- $\mathbb{P}_2$  implicite pour différentes valeurs du nombre CFL

du nombre CFL (de l'ordre de la limite de stabilité des méthodes explicites), l'erreur du schéma en espace est plus importante que celle du schéma en temps et inversement.

Dans le but de valider cette étude, nous nous proposons maintenant de confronter les résultats précédents à des résultats de simulations numériques. On utilise pour cela un code de calcul 1D de propagation d'un pulse gaussien développé sous ©Matlab. Nous montrons sur les Fig. 3.6 et 3.7 l'évolution de l'erreur numérique en fonction de la longueur d'onde numérique  $K$  et pour différentes valeurs du nombre CFL. Nous voyons comme sur les courbes précédentes que l'erreur est plus faible pour des petites valeurs des nombres CFL et  $K$ . Par ailleurs, sur la Fig. 3.3, on note que la courbe correspondant à CFL=20 coupe la droite d'abscisse 0.2 à une certaine ordonnée. A cette même ordonnée, les courbes correspondant à CFL=4 et CFL=10 coupent respectivement les droites d'abscisse 0.4 et 1.0. Nous avons donc le même niveau de dispersion numérique pour ces différentes valeurs du nombre CFL à différentes valeurs de  $K$ . Il se trouve que cette similitude se retrouve sur la Fig. 3.6 pour une ordonnée de  $5.6 \cdot 10^{-3}$ , ce qui tend à confirmer notre étude théorique. Comme on a pu le constater auparavant, les méthodes DGTD- $\mathbb{P}_1$  et DGTD- $\mathbb{P}_2$  implicites ont des dispersions très proches. Nous retrouvons bien ce comportement sur les Fig. 3.4 et 3.7.

Enfin, nous visualisons sur les Fig. 3.8, 3.9 et Fig. 3.10 l'évolution de l'erreur de dispersion des méthodes DGTD- $\mathbb{P}_k$  explicites pour différentes valeurs du nombre CFL. Contrairement aux méthodes implicites, les nombres CFL utilisés avec les méthodes DGTD- $\mathbb{P}_k$  explicites sont plus faibles. On remarque sur la Fig. 3.8 que les courbes de dispersion sont très similaires. On remarque enfin sur la Fig. 3.10 que l'erreur de dispersion est moindre par rapport à la méthode DGTD- $\mathbb{P}_1$ . On peut donc en conclure que c'est le schéma en espace qui est limitatif pour ces ordres.

### 3.2.3 Résultats numériques en 2D

Nous présentons dans cette sous-section des résultats numériques visant à évaluer la précision et les performances de la méthode DGTD- $\mathbb{P}_k$  implicite basée sur le schéma de Crank-Nicolson. Cette étude est réalisée dans un cadre bidimensionnel. Nous considérons le cas des ondes transverses magnétiques  $\text{TM}_z$  pour lesquelles  $\mathbf{H} = {}^t(H_x, H_y, 0)$  et  $\mathbf{E} = {}^t(0, 0, E_z)$ . Seules les composantes non nulles des champs électrique et magnétique sont considérées. Ainsi, les équations de Maxwell en mode  $\text{TM}_z$  s'écrivent :

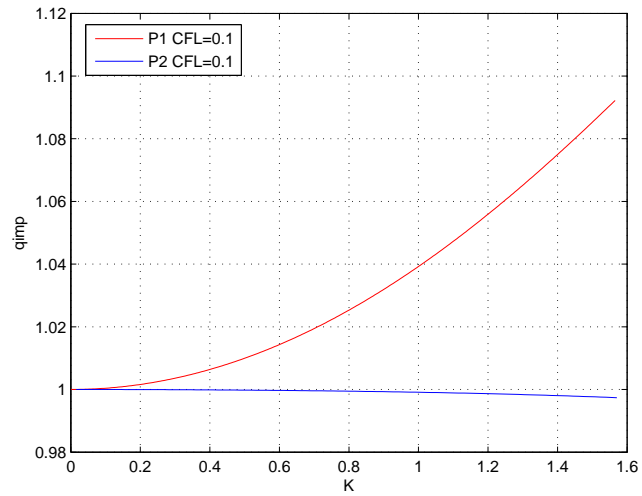


FIG. 3.5 – Dispersion des méthodes DGTD- $\mathbb{P}_1$  et DGTD- $\mathbb{P}_2$  implicites pour CFL=0.1

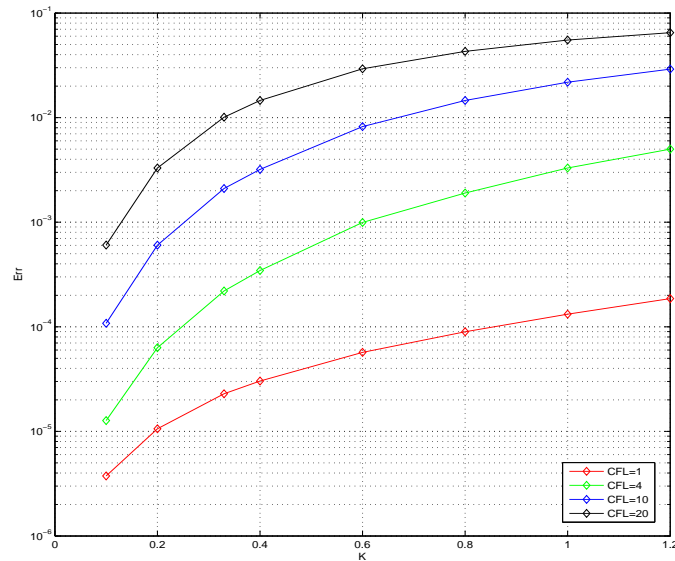


FIG. 3.6 – Erreur numérique de la méthode DGTD- $\mathbb{P}_1$  implicite

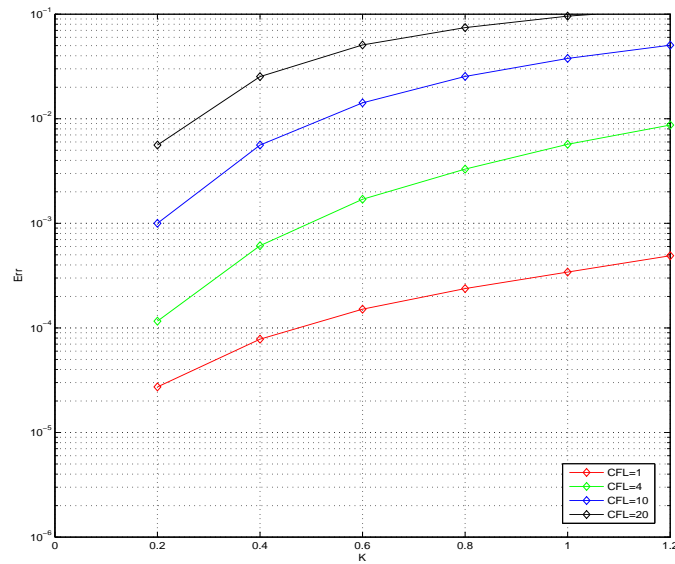


FIG. 3.7 – Erreur numérique de la méthode DGTD- $\mathbb{P}_2$  implicite

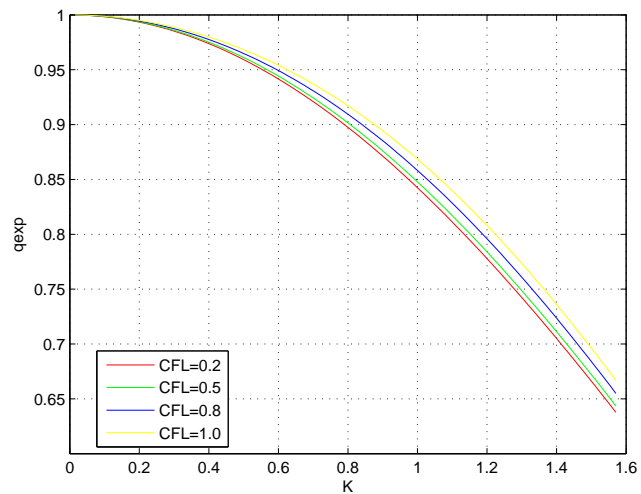
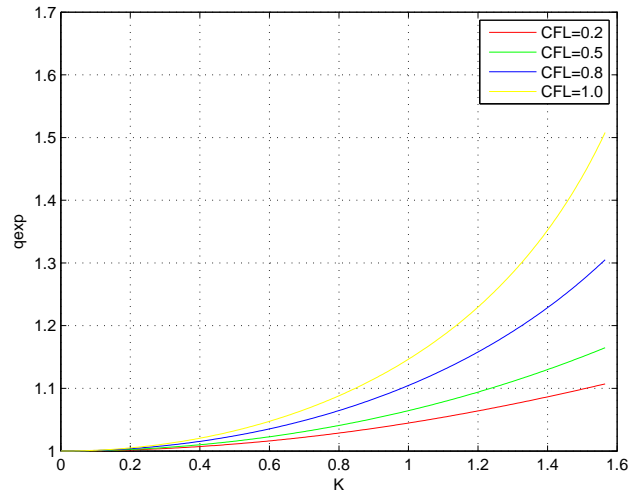
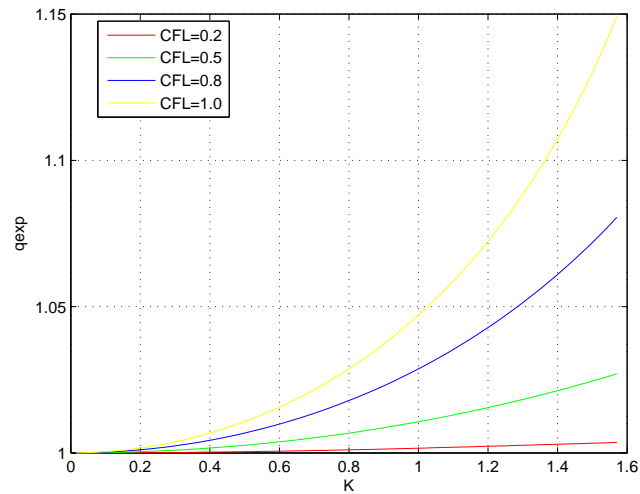


FIG. 3.8 – Dispersion de la méthode DGTD- $\mathbb{P}_0$  explicite pour différentes valeurs du nombre CFL

FIG. 3.9 – Dispersion de la méthode DGTD- $\mathbb{P}_1$  explicite pour différentes valeurs du nombre CFLFIG. 3.10 – Dispersion de la méthode DGTD- $\mathbb{P}_2$  explicite pour différentes valeurs du nombre CFL



$$\begin{cases} \epsilon \frac{\partial E_z}{\partial t} - \frac{\partial H_y}{\partial x} + \frac{\partial H_x}{\partial y} = -z_0 E_z, \\ \mu \frac{\partial H_x}{\partial t} + \frac{\partial E_z}{\partial y} = 0, \\ \mu \frac{\partial H_y}{\partial t} - \frac{\partial E_z}{\partial x} = 0, \end{cases} \quad (3.38)$$

que l'on peut aussi écrire sous la forme (3.9) avec :

$$Q = \begin{bmatrix} \epsilon & 0_{1 \times 2} \\ 0_{2 \times 1} & \mu I_{2 \times 2} \end{bmatrix}, \quad \mathbf{W} = \begin{pmatrix} E_z \\ H_x \\ H_y \end{pmatrix} \quad \text{et} \quad F(\mathbf{W}) = \begin{pmatrix} F_x(\mathbf{W}) \\ F_y(\mathbf{W}) \end{pmatrix},$$

$$F_x(\mathbf{W}) = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} \mathbf{W} = \begin{bmatrix} 0 & N_{e_x} \\ {}^t N_{e_x} & 0_{2 \times 2} \end{bmatrix} \mathbf{W} = \begin{pmatrix} -H_y \\ 0 \\ -E_z \end{pmatrix},$$

$$F_y(\mathbf{W}) = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \mathbf{W} = \begin{bmatrix} 0 & N_{e_y} \\ {}^t N_{e_y} & 0_{2 \times 2} \end{bmatrix} \mathbf{W} = \begin{pmatrix} H_x \\ E_z \\ 0 \end{pmatrix},$$

$$\text{où } {}^t N_n = \begin{pmatrix} n_y \\ -n_x \end{pmatrix}, \quad e_x = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{et} \quad e_y = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Nous considérons successivement trois cas tests. Le premier traite de la propagation d'un mode propre dans une cavité dont les parois sont parfaitement conductrices. Pour ce cas test, une solution analytique est connue et l'on peut donc évaluer l'erreur numérique commise avec la méthode DGTD- $\mathbb{P}_k$  implicite relativement à son homologue explicite. Les deux autres cas tests portent respectivement sur la diffraction d'une onde plane par un carré parfaitement conducteur et sur la diffraction d'une onde plane par un cylindre diélectrique. Les maillages utilisés pour ces simulations sont localement raffinés pour mettre en évidence le gain en temps de calcul apporté par la méthode DGTD- $\mathbb{P}_k$  implicite dans ce contexte. Enfin, comme nous l'évoquons en début de cette section, le choix de la méthode de résolution des systèmes linéaires obtenus est déterminant pour l'obtention d'une méthode DGTD- $\mathbb{P}_k$  implicite. Dans un cadre bidimensionnel, une méthode directe offre dans la plupart des cas le meilleur rapport vitesse d'exécution/occupation mémoire. Notre choix s'est porté sur le solveur direct creux MUMPS (MULTifrontal Massively Parallel sparse direct Solver) [Amestoy et al., 2000]. Notons enfin que dans le cas de milieux de propagation dont les caractéristiques électromagnétiques  $\epsilon$  et  $\mu$  ne dépendent pas du temps, la factorisation LU de la matrice  $\mathbb{P}$  caractérisant le système implicite (3.3)-(3.4) peut être effectuée une fois pour toutes avant la boucle en temps. Chaque étape de résolution consiste alors en l'inversion de deux systèmes triangulaires ce qui peut être réalisé efficacement.

### 3.2.3.1 Mode propre dans une cavité

Le premier cas test que nous considérons est la propagation d'un mode propre dans une cavité dont les parois sont parfaitement conductrices (voir aussi la sous-section 2.4.1.1). Ce cas test autorise une comparaison avec la solution exacte. Ici, nous verrons que nous ne pouvons tirer avantage du schéma implicite que nous proposons si le maillage utilisé est uniforme (ou quasi uniforme) alors que nous obtenons un gain en temps de calcul substantiel si le maillage utilisé est non-uniforme.

Pour ce faire, nous utilisons deux maillages triangulaires (voir la Fig. 3.11) :

- un maillage uniforme composé de 1681 sommets et de 3200 triangles. Le pas de temps adimensionné correspondant à un nombre  $\text{CFL-}\mathbb{P}_0 = 1$  est  $(\Delta t)_u = (\Delta t)_u / 3.10^8$  s. Pour une interpolation d'ordre  $p \geq 1$ , le pas de temps utilisé est donné par  $\text{CFL-}\mathbb{P}_p \times (\Delta t)_u$  où  $\text{CFL-}\mathbb{P}_k$  est le nombre CFL associé à la méthode  $\text{DGTD-}\mathbb{P}_k$ .
- Un maillage non-uniforme composé de 1400 sommets et de 2742 triangles. Le rapport de taille entre la plus grande et la plus petite arête du maillage est de 178. Dans ce cas, les valeurs minimale et maximale du pas de temps sont respectivement  $(\Delta t)_m = 4.34 \cdot 10^{-4}$  m et  $(\Delta t)_M = 7.06 \cdot 10^{-2}$  m. Le pas de temps utilisé pour les simulations est  $\text{CFL-}\mathbb{P}_p \times (\Delta t)_m$ .

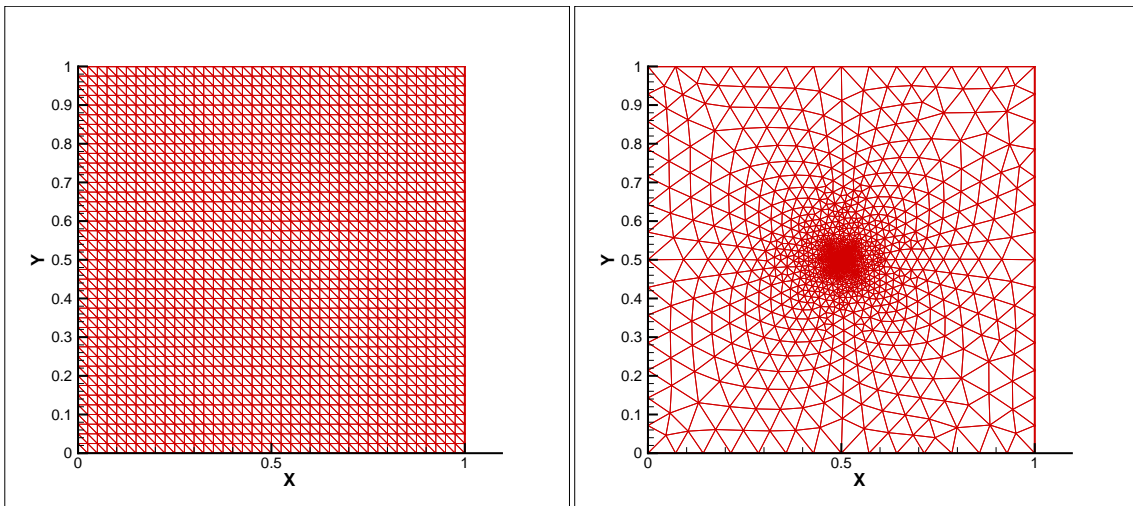


FIG. 3.11 – Mode propre dans un carré unité  
Maillages uniforme (gauche) et non-uniforme (droite) du carré unitaire

Pour la méthode  $\text{DGTD-}\mathbb{P}_k$  explicite,  $\text{CFL-}\mathbb{P}_p \leq 1$  est dictée par des considérations de stabilité tandis que pour la méthode  $\text{DGTD-}\mathbb{P}_k$  implicite, la valeur du nombre CFL peut être choisie arbitrairement grande. En pratique, le choix de cette valeur est guidé par des considérations de précision de la solution calculée numériquement. Les résultats obtenus sont montrés sur les Fig. 3.12 à 3.16 sous la forme de l'évolution en temps de différentes quantités : l'erreur  $L_2$  entre solutions exacte et approchée, l'énergie discrète et la composante  $E_z$  du champs électrique. Les simulations portent sur 10 périodes. Les temps de calcul sont répertoriés dans les Tab. 3.1 et 3.2 .

Tout d'abord, il est clair que la méthode  $\text{DGTD-}\mathbb{P}_k$  implicite du second ordre n'est pas un bon choix si le maillage utilisé est uniforme. En effet, la dispersion numérique introduite par le schéma de Crank-Nicolson dégrade notablement la précision globale du calcul. En revanche, si l'on utilise un maillage non-uniforme, la méthode  $\text{DGTD-}\mathbb{P}_k$  implicite devient alors intéressante malgré le surcoût engendré par la résolution du système linéaire à chaque itération en temps. Intéressons nous au comportement en temps de l'erreur  $L_2$  sur les Fig. 3.14 et 3.15 ; nous pouvons voir qu'une valeur du nombre  $\text{CFL}=12.0$  (respectivement 4.0) pour la méthode  $\text{DGTD-}\mathbb{P}_1$  (respectivement  $\text{DGTD-}\mathbb{P}_2$ ) conduit à une solution acceptable. Pour ces valeurs de CFL, le gain en temps de calcul global (CPU) entre les méthode implicite et explicite est respectivement de 3.3 et 2.2. On notera que ces temps de calcul incluent le temps de factorisation.

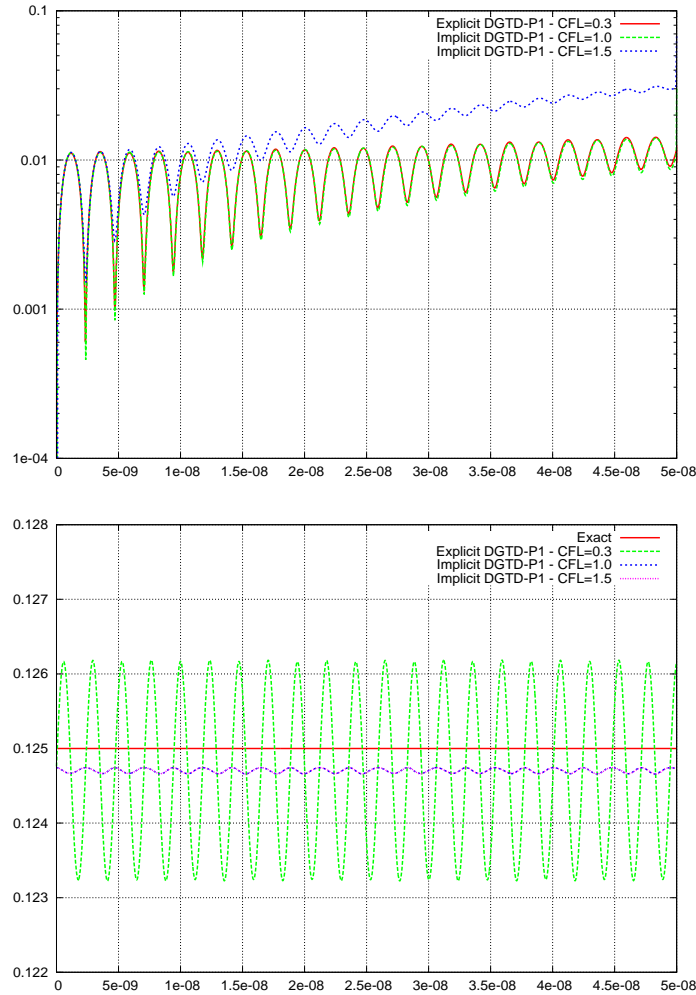


FIG. 3.12 – Mode propre dans un carré unité - Maillage uniforme  
 Méthode DGTD- $\mathbb{P}_1$  : évolution temporelle de l'erreur  $L_2$  (haut) et de l'énergie discrète (bas)  
 Abscisse : temps en sec - Ordonnée : erreur  $L_2$  (haut) énergie discrète (bas)

Intégration en temps	Méthode	CFL- $\mathbb{P}_k$	Temps CPU
Explicite	DGTD- $\mathbb{P}_1$	0.3	15 sec
Implicite	-	1.0	44 sec
-	-	1.5	30 sec

TAB. 3.1 – Mode propre dans un carré unité - Maillage uniforme  
 Temps CPU (station de travail basée sur un processeur AMD Opteron 2 GHz)

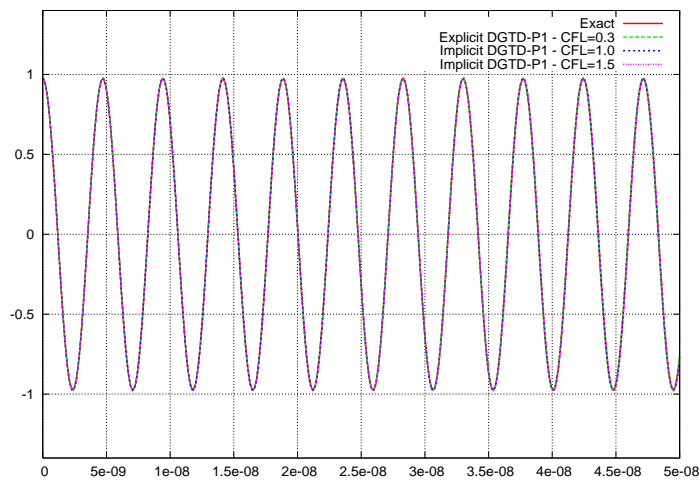


FIG. 3.13 – Mode propre dans un carré unité - Maillage uniforme  
Evolution temporelle de  $E_z$  : méthode DGTD- $\mathbb{P}_1$  (abscisse : temps en sec - ordonnée :  $E_z$ )

Intégration en temps	Méthode	CFL- $\mathbb{P}_k$	Temps CPU
Explicite	DGTD- $\mathbb{P}_1$	0.3	443 sec
Implicite	-	12.0	133 sec
-	-	24.0	67 sec
Explicite	DGTD- $\mathbb{P}_2$	0.2	2057 sec
Implicite	-	2.0	1923 sec
-	-	4.0	938 sec
-	-	6.0	620 sec

TAB. 3.2 – Mode propre dans un carré unité - Maillage non-uniforme  
Temps CPU (station de travail basée sur un processeur AMD Opteron 2 GHz)

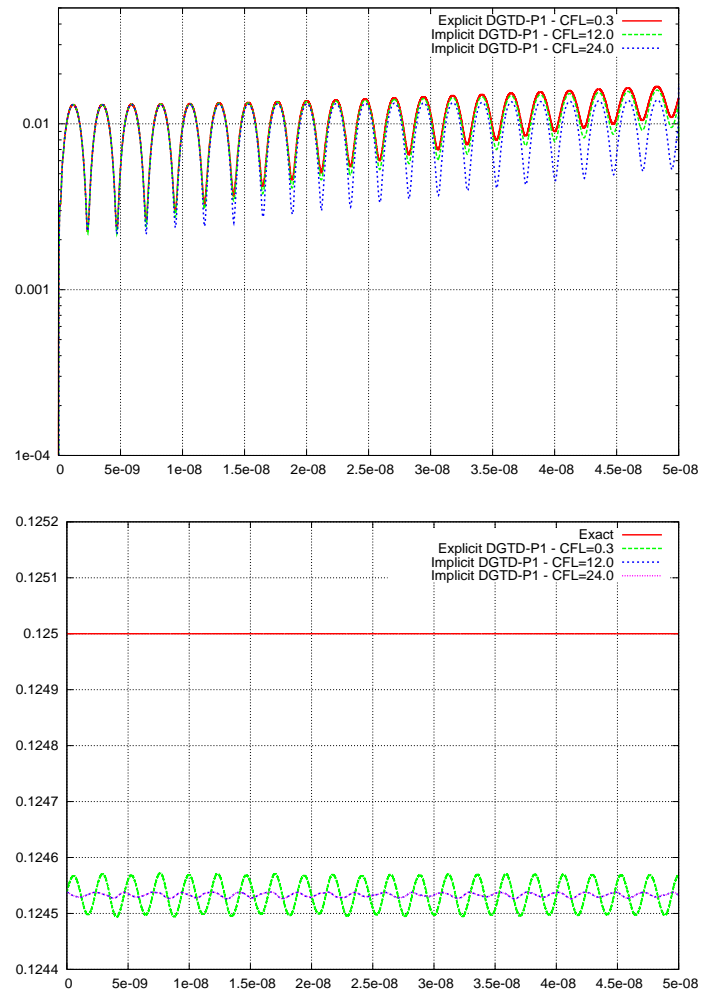


FIG. 3.14 – Mode propre dans un carré unité - Maillage non-uniforme  
 Méthode DGTD- $\mathbb{P}_1$  : évolution temporelle de l'erreur  $L_2$  (haut) et de l'énergie discrète (bas)  
 Abscisse : temps en sec - Ordonnée : erreur  $L_2$  (haut) énergie discrète (bas)

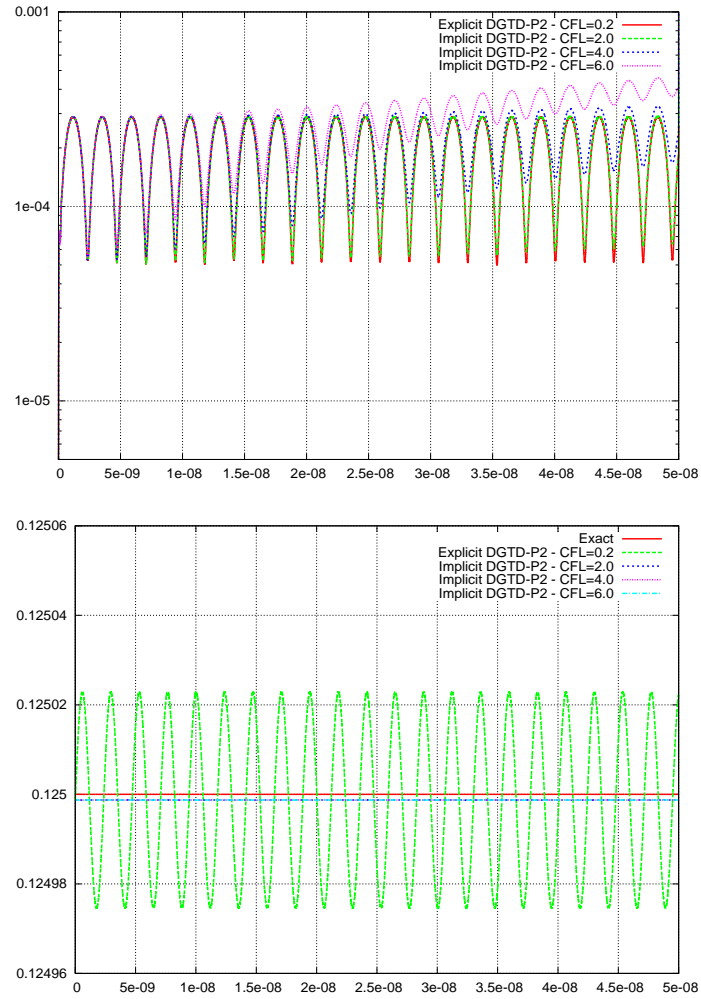


FIG. 3.15 – Mode propre dans un carré unité - Maillage non-uniforme  
 Méthode DGTD- $\mathbb{P}_2$  : évolution temporelle de l'erreur  $L_2$  (haut) et de l'énergie discrète (bas)  
 Abscisse : temps en sec - Ordonnée : erreur  $L_2$  (haut) énergie discrète (bas)

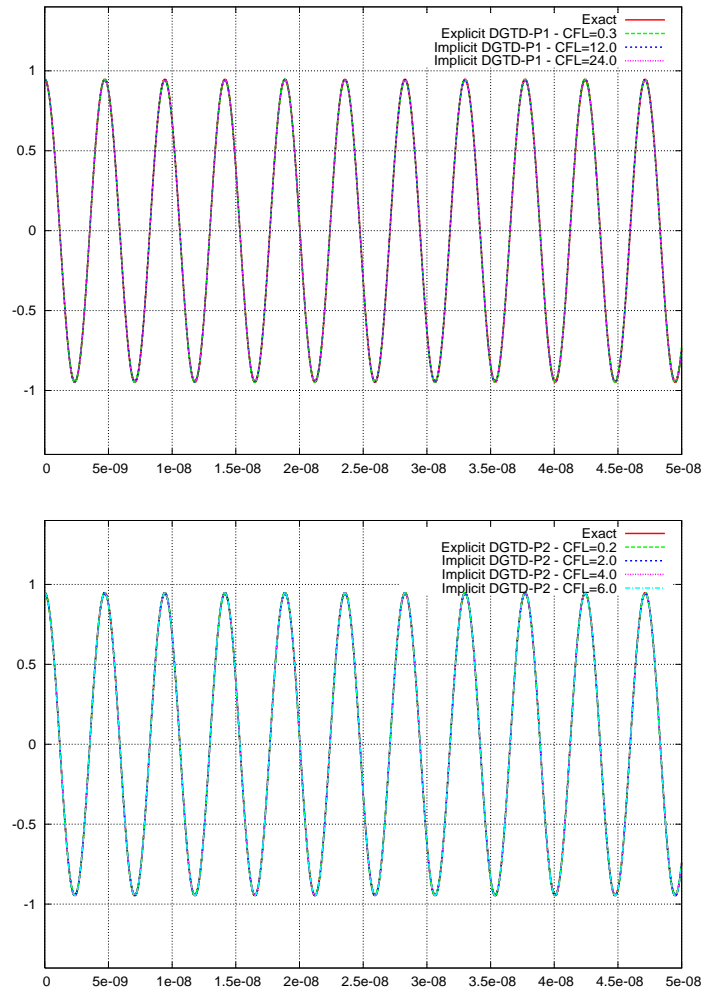


FIG. 3.16 – Mode propre dans un carré unité - Maillage non-uniforme  
 Evolution temporelle de  $E_z$  : Méthodes DGTD- $\mathbb{P}_1$  (haut) et DGTD- $\mathbb{P}_2$  (bas)  
 Abscisse : temps en sec - Ordonnée :  $E_z$

### 3.2.3.2 Diffraction d'une onde plane par un carré

Le second cas test que nous considérons est la diffraction d'une onde plane par un carré parfaitement conducteur de coté  $c = 0.25$  m. Un condition absorbante de type Silver-Müller est appliquée sur la frontière  $\Gamma_a$  qui est ici définie par un carré de coté  $c = 1.0$  m. Nous utilisons un maillage non-uniforme (voir la Fig.3.17) composé de 6018 sommets et de 10792 triangles. Le rapport de taille entre la plus grande et la plus petite arête du maillage est de 357. Dans ce cas, les valeurs minimale et maximale du pas de temps sont respectivement  $(\Delta t)_m = 2.86 \cdot 10^{-4}$  m et  $(\Delta t)_M = 9.86 \cdot 10^{-2}$  m. Comme précédemment, le pas de temps utilisé pour la simulation est  $\text{CFL} \cdot \mathbb{P}_p \times (\Delta t)_m$ . Les calculs ont été effectués pour trois fréquences différentes de l'onde plane incidente :  $F=300$  MHz,  $F=600$  MHz et  $F=900$  MHz. Dans chaque cas, les simulations se font sur 10 périodes et une transformée de Fourier discrète est appliquée sur les composantes du champ lors de la dernière période.

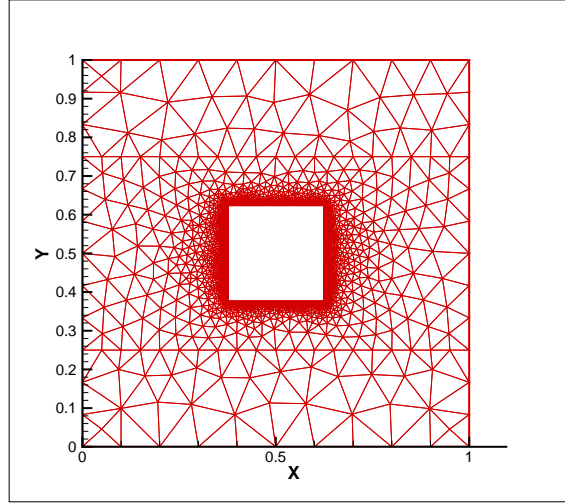


FIG. 3.17 – Diffraction d'une onde plane par un carré : maillage triangulaire

Les résultats obtenus sont présentés sur les Fig. 3.18 à 3.20 sous la forme de la transformée de Fourier discrète (DFT) de la coupe 1D suivant  $x$  pour  $y = 0.25$  m du champ  $E_z$ . Pour chaque configuration de calcul (méthode  $\text{DGTD-}\mathbb{P}_k$  et fréquence  $F$ ), nous montrons la distribution de  $\text{DFT}(E_z)$  pour la méthode  $\text{DGTD-}\mathbb{P}_k$  explicite considérée comme solution de référence, et les deux distributions de  $\text{DFT}(E_z)$  correspondant à la méthode  $\text{DGTD-}\mathbb{P}_k$  implicite. Les calculs sont effectués avec le nombre CFL maximal conduisant à une solution approchée qui se superpose à la solution de référence, et avec un nombre CFL légèrement plus élevé conduisant à une solution moins précise. On notera que ces distributions 1D sont tracées en termes des degrés de liberté globaux localisés sur la droite  $y = 0.25$  m (voir la Fig. 3.17) et la valeur associée de la  $\text{DFT}(E_z)$  est calculée comme la valeur moyenne des degrés de liberté locaux. Les temps de calculs sont récapitulés dans la Tab. 3.3.

Ces résultats motivent deux remarques principales :

- le nombre CFL maximal autorisé décroît quand la fréquence de l'onde plane incidente augmente. Ceci n'est pas surprenant car bien que la méthode  $\text{DGTD-}\mathbb{P}_k$  implicite soit inconditionnellement stable, le nombre CFL maximal est dicté par des considérations physiques.
- Pour un ordre d'interpolation donné, le gain en temps de calcul i.e. le ratio du temps de calcul de la méthode  $\text{DGTD-}\mathbb{P}_k$  explicite sur le temps de calcul de la méthode  $\text{DGTD-}\mathbb{P}_k$  implicite, décroît quand



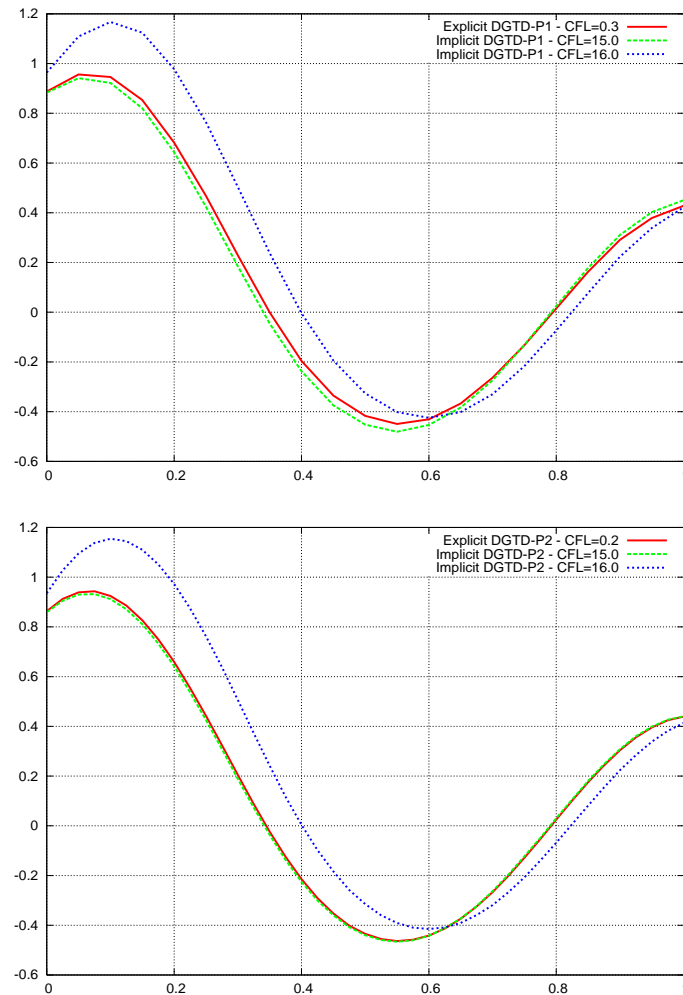


FIG. 3.18 – Diffraction d'une onde plane par un carré :  $F=300$  MHz  
Distribution 1D de la  $DFT(E_z)$ ,  $y = 0.75$  m  
Haut : DGTD- $\mathbb{P}_1$  méthode - Bas : DGTD- $\mathbb{P}_2$  méthode

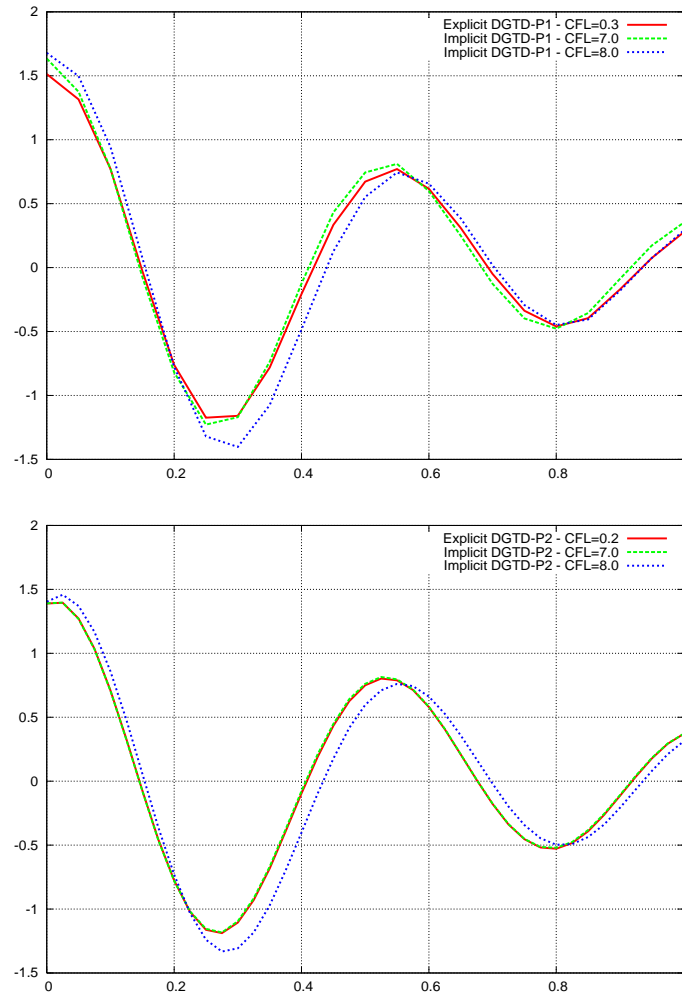


FIG. 3.19 – Diffraction d'une onde plane par un carré :  $F=600$  MHz  
 Distribution 1D de la  $DFT(E_z)$ ,  $y = 0.25$  m  
 Haut : DGTD- $\mathbb{P}_1$  méthode - Bas : DGTD- $\mathbb{P}_2$  méthode

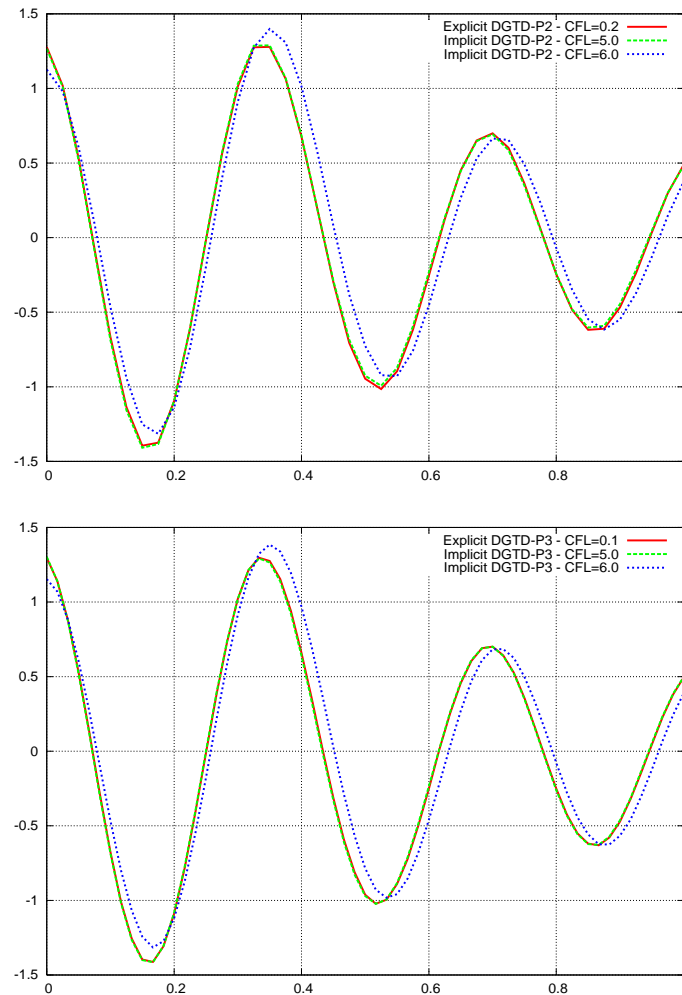


FIG. 3.20 – Diffraction d’une onde plane par un carré :  $F=900$  MHz  
 Distribution 1D de la  $DFT(E_z)$ ,  $y = 0.25$  m  
 Haut : DGTD- $\mathbb{P}_2$  méthode - Bas : DGTD- $\mathbb{P}_3$  méthode

Fréquence	Intégration en temps	Méthode	CFL- $\mathbb{P}_k$	Temps CPU
300 MHz	Explicite	DGTD- $\mathbb{P}_1$	0.3	1602 sec
-	Implicite	-	15.0	370 sec
-	Explicite	DGTD- $\mathbb{P}_2$	0.2	5677 sec
-	Implicite	-	15.0	762 sec
600 MHz	Explicite	DGTD- $\mathbb{P}_1$	0.3	758 sec
-	Implicite	-	7.0	383 sec
-	Explicite	DGTD- $\mathbb{P}_2$	0.2	3074 sec
-	Implicite	-	7.0	767 sec
900 MHz	Explicite	DGTD- $\mathbb{P}_2$	0.2	2191 sec
-	Implicite	-	5.0	746 sec
-	Explicite	DGTD- $\mathbb{P}_3$	0.1	8771 sec
-	Implicite	-	5.0	1591 sec

TAB. 3.3 – Diffraction d'une onde plane par un carré  
Temps CPU (station de travail basée sur un processeur AMD Opteron 2 GHz)

la fréquence augmente. Par exemple, pour  $p = 2$  ce ratio passe de 7.5 pour  $F=300$  Mhz à 3.0 pour  $F=900$  MHz. Cependant, pour une fréquence donnée, ce gain augmente avec l'ordre d'interpolation : pour  $F=900$  MHz, ce ratio est respectivement égal à 3.0 pour  $p = 2$  et 5.5 pour  $p = 3$ .

### 3.2.3.3 Diffraction d'une onde plane par un cylindre diélectrique

Le troisième cas test sélectionné est la diffraction d'une onde plane par un cylindre diélectrique, considérant ainsi un cas de propagation en milieu hétérogène. La configuration de ce problème est décrite sur la Fig. 3.21.

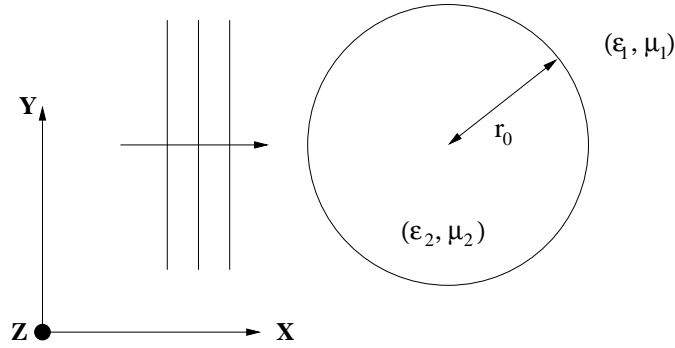


FIG. 3.21 – Diffraction d'une onde plane par un cylindre diélectrique

Nous supposons que le cylindre est éclairé par une onde plane monochromatique de la forme :

$$E_z^{\text{inc}} = \exp(-\mathbf{i}(k_1 x - \omega t)) \quad , \quad H_y^{\text{inc}} = -\exp(-\mathbf{i}(k_1 x - \omega t))$$

où  $k_1 = \omega \sqrt{\epsilon_1 \mu_1}$ . Dans ce cas, la solution exacte du problème de diffraction est donné par :

$$\begin{aligned}
E_z(x, y, t) &= E_z(r, \theta, t) \\
&= e^{i\omega t} \begin{cases} \sum_{n=-\infty}^{\infty} C_n^{\text{tot}} J_n(k_2 r) e^{in\theta}, & r \leq r_0, \\ \sum_{n=-\infty}^{\infty} (\mathbf{i}^{-n} J_n(k_1 r) + C_n^{\text{scat}} H_n^{(2)}(k_1 r)) e^{in\theta}, & r > r_0, \end{cases}
\end{aligned}$$

où  $J_n$  et  $H_n^{(2)}$  représentent respectivement la fonction de Bessel de première espèce d'ordre  $n$  et la fonction de Hankel de seconde espèce d'ordre  $n$ , et  $k_2 = \omega\sqrt{\epsilon_2\mu_2}$  est la constante de propagation pour un milieu homogène diélectrique sans perte. Enfin,  $(r, \theta) = (\sqrt{x^2 + y^2}, \arctan(y/x))$  représentent les coordonnées polaires. Les coefficients  $C_n^{\text{tot}}$  pour le champ total dans le cylindre diélectrique s'écrivent :

$$C_n^{\text{tot}} = \mathbf{i}^{-n} \frac{(k_1/\mu_1) J_n'(k_1 r_0) H_n^{(2)}(k_1 r_0) - (k_1/\mu_1) H_n^{(2)'}(k_1 r_0) J_n(k_1 r_0)}{(k_2/\mu_2) J_n'(k_2 r_0) H_n^{(2)}(k_1 r_0) - (k_1/\mu_1) H_n^{(2)'}(k_1 r_0) J_n(k_2 r_0)}.$$

Les coefficients correspondant au champ diffracté  $C_n^{\text{scat}}$  sont donnés par :

$$C_n^{\text{scat}} = \mathbf{i}^{-n} \frac{(k_1/\mu_1) J_n'(k_1 r_0) J_n(k_2 r_0) - (k_2/\mu_2) J_n'(k_2 r_0) J_n(k_1 r_0)}{(k_2/\mu_2) J_n'(k_2 r_0) H_n^{(2)}(k_1 r_0) - (k_1/\mu_1) H_n^{(2)'}(k_1 r_0) J_n(k_2 r_0)}.$$

Utilisant les équations de Maxwell, on en déduit les solutions des composantes du champ magnétique. La composante angulaire est donnée par :

$$H_\theta(r, \theta, t) = -e^{i\omega t} \begin{cases} \frac{-\mathbf{i}k_2}{\omega\mu_2} \sum_{n=-\infty}^{\infty} C_n^{\text{tot}} J_n'(k_2 r) e^{in\theta}, & r \leq r_0, \\ \frac{-\mathbf{i}k_1}{\omega\mu_1} \sum_{n=-\infty}^{\infty} (\mathbf{i}^{-n} J_n'(k_1 r) + C_n^{\text{scat}} H_n^{(2)'}(k_1 r)) e^{in\theta}, & r > r_0, \end{cases}$$

et la composante radiale par :

$$H_r(r, \theta, t) = -e^{i\omega t} \begin{cases} \frac{\mathbf{i}}{\omega\mu_2 r} \sum_{n=-\infty}^{\infty} \mathbf{i}n C_n^{\text{tot}} J_n(k_2 r) e^{in\theta}, & r \leq r_0, \\ \frac{\mathbf{i}}{\omega\mu_1 r} \sum_{n=-\infty}^{\infty} \mathbf{i}n (\mathbf{i}^{-n} J_n(k_1 r) + C_n^{\text{scat}} H_n^{(2)}(k_1 r)) e^{in\theta}, & r > r_0. \end{cases}$$

Dans ce qui suit, nous posons  $\mu_1 = \mu_2 = \epsilon_1 = 1$ , c'est à dire que le matériau est non-magnétique, et que l'extérieur du cylindre est occupé par le vide. La fréquence angulaire adimensionnée est fixée à  $\omega = 2\pi$  (ce qui correspond à une fréquence  $F=300$  MHz) et le domaine de calcul  $\Omega$  est choisi comme un cylindre de rayon unité centré en  $(0, 0)$ . La frontière du domaine  $\Gamma_a$  où sont appliquées les conditions de Silver-Müller du premier ordre est définie, quand à elle, par un cylindre de rayon  $r = 1.6$  m.

**Configuration C1** :  $r_0 = 0.6$  m et  $\epsilon_2 = 2.25$  Dans un premier temps, nous considérons une situation déjà traitée dans [Hesthaven et al., 1999]-[Cai and Deng, 2003]-[Shi et al., 2006] où le cylindre interne a un rayon  $r_0 = 0.6$  m et délimite un matériau ayant une permittivité relative  $\epsilon_2 = 2.25$ . Nous utilisons un maillage non-uniforme (voir la Fig. 3.22) composé de 4108 sommets et de 8054 triangles. Le rapport entre la plus grande et la plus petite arête de ce maillage est de 197. Dans ce cas, les valeurs minimales et maximales

du pas de temps sont respectivement données par  $(\Delta t)_m = 6.27 \cdot 10^{-4}$  m et  $(\Delta t)_M = 9.29 \cdot 10^{-2}$  m. Comme précédemment, le pas de temps utilisé pour la simulation est  $\text{CFL-}\mathbb{P}_p \times (\Delta t)_m$ .

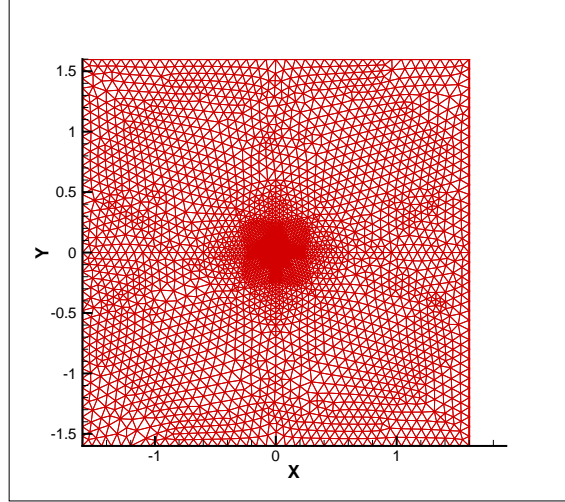


FIG. 3.22 – Diffraction d’une onde plane sur un diélectrique (C1) : maillage triangulaire

Les résultats sont présentés sur les Fig. 3.23 et 3.24 sous la forme, respectivement, des lignes de contour du champ  $E_z$  après 10 périodes et de la distribution 1D suivant l’axe  $x$  pour  $y = 0.0$  m de la transformée de Fourier de  $E_z$ . De plus, la Fig. 3.25 montre l’évolution en temps de l’erreur  $L_2$  entre la solution numérique et la solution exacte pour les méthodes DGTD- $\mathbb{P}_1$  et DGTD- $\mathbb{P}_2$  explicites et implicites. Les temps de calculs sont résumés dans la Tab. 3.4 tandis que la Tab. 3.5 liste les temps de factorisation et l’espace mémoire nécessaires au stockage des facteurs L et U. Pour cette configuration du problème, le gain apporté par la méthode implicite sur la méthode explicite est d’un facteur 5.4 pour  $p = 1$  et de 8.7 pour  $p = 2$ . Nous pouvons noter également que le temps de factorisation représente un faible pourcentage du temps de la simulation et ce pour les deux degrés d’interpolation.

Intégration en temps	Méthode	CFL- $\mathbb{P}_k$	Temps CPU
Explicite	DGTD- $\mathbb{P}_1$	0.3	542 sec
Implicite	-	21.0	102 sec
Explicite	DGTD- $\mathbb{P}_2$	0.2	1892 sec
Implicite	-	20.0	218 sec

TAB. 3.4 – Diffraction d’une onde plane par un cylindre diélectrique (C1)  
Temps CPU (station de travail basée sur un processeur AMD Opteron 2 GHz)

**Configuration C2 :**  $r_0 = 0.05$  m et  $\epsilon_2 = 10.0$ . Nous considérons maintenant le cas d’un cylindre intérieur de rayon  $r_0 = 0.05$  m et délimitant un matériau de permittivité relative  $\epsilon_2 = 10.0$ . Nous utilisons un maillage non-uniforme composé de 4228 sommets et de 8326 triangles. Le rapport entre la plus grande et la plus petite arête du maillage est de 5138. Dans ce cas, les valeurs minimales et maximales du pas de temps sont respectivement données par  $(\Delta t)_m = 2.6 \cdot 10^{-5}$  m et  $(\Delta t)_M = 1.22 \cdot 10^{-1}$  m. Comme

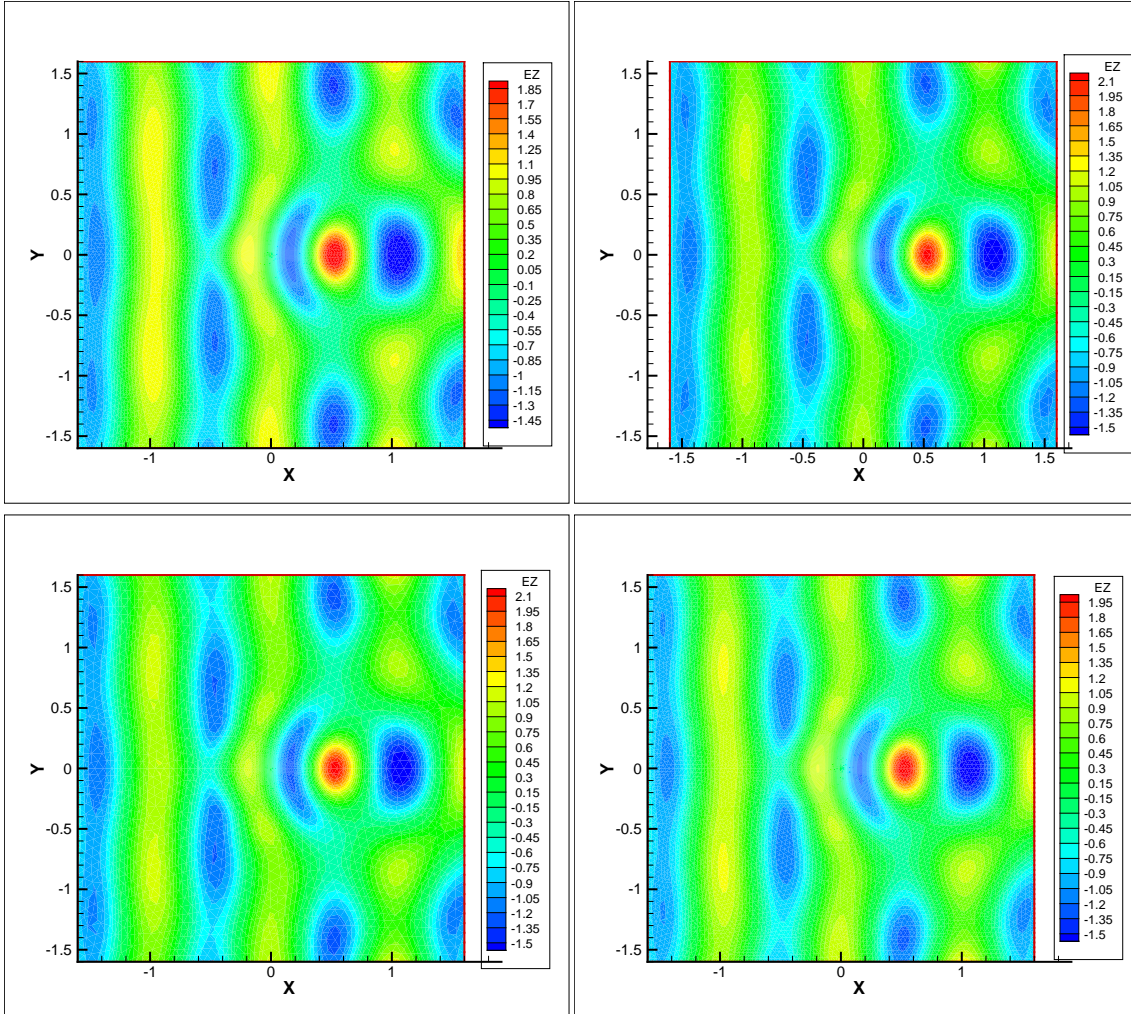


FIG. 3.23 – Diffraction d’une onde plane par un cylindre diélectrique (C1)

Lignes de contour  $E_z$  après 10 périodes

Haut gauche : solution analytique - Haut droite : méthode DGTD- $\mathbb{P}_1$  explicite, CFL=0.3

Bas gauche : méthode DGTD- $\mathbb{P}_1$  implicite, CFL=21.0

Bas droite : méthode DGTD- $\mathbb{P}_2$  implicite, CFL=20.0

Méthode	CFL- $\mathbb{P}_k$	Temps CPU	Mémoire RAM
DGTD- $\mathbb{P}_1$	21.0	3 sec	70 MB
DGTD- $\mathbb{P}_2$	20.0	6 sec	181 MB

TAB. 3.5 – Diffraction d’une onde plane par un cylindre diélectrique (C1)

Temps CPU (station de travail basée sur un processeur AMD Opteron 2 GHz)

Temps de factorisation et quantité de mémoire nécessaire pour le stockage des facteurs L et U

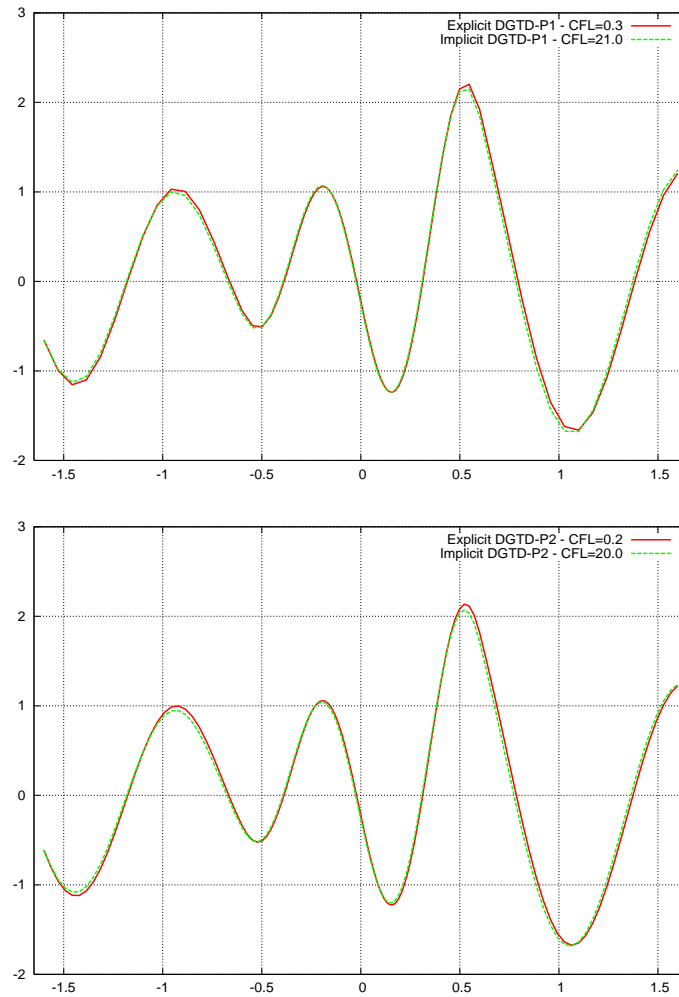


FIG. 3.24 – Diffraction d'une onde plane par un cylindre diélectrique (C1)

Distribution 1D de la DFT( $E_z$ ),  $y = 0.0$  mHaut : méthode DGTD- $\mathbb{P}_1$  - Bas : méthode DGTD- $\mathbb{P}_2$



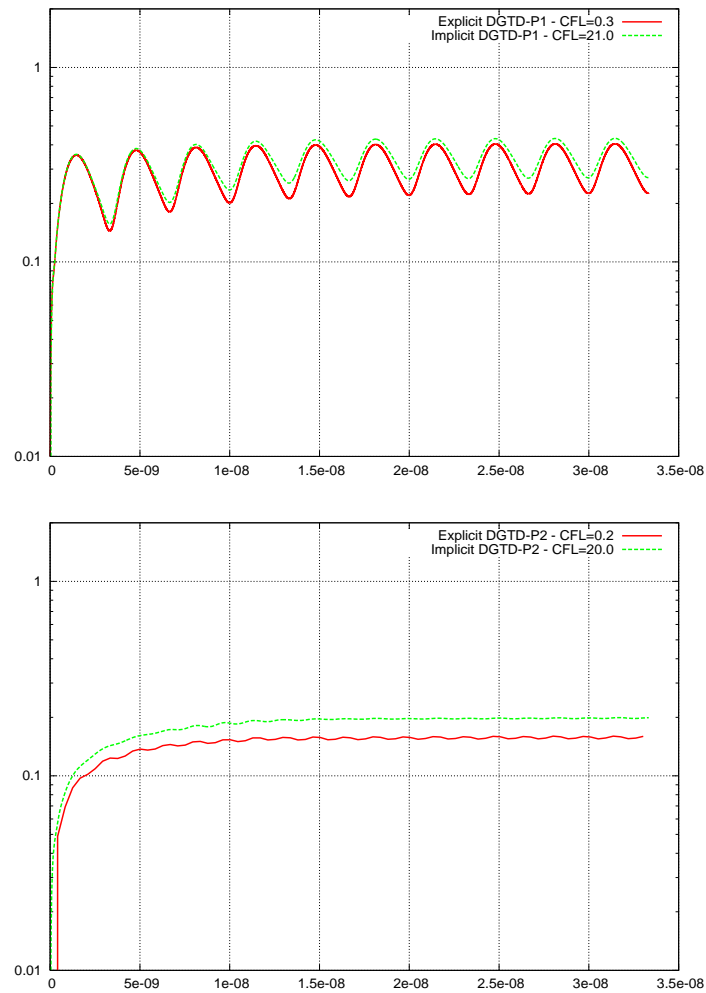


FIG. 3.25 – Diffraction d'une onde plane par un cylindre diélectrique (C1)  
 Evolution temporelle de l'erreur  $L_2$  (abscisse : temps en sec - ordonnée : erreur  $L_2$  en échelle log)  
 Haut : méthode DGTD- $\mathbb{P}_1$  - Bas : méthode DGTD- $\mathbb{P}_2$

précédemment, le pas de temps utilisé dans les simulations est  $\text{CFL-}\mathbb{P}_p \times (\Delta t)_m$ . Au premier abord, cette configuration de calcul peut sembler quelque peu artificielle. En réalité, elle modélise une situation que l'on rencontre très fréquemment dans les problèmes réels, tout spécialement dans les cas tridimensionnels, soit parce qu'un détail géométrique impose un raffinement local très important, ou comme le résultat d'un raffinement de maillage lié à un saut important des valeurs des caractéristiques électromagnétiques du milieu de propagation. Sur la Fig. 3.26, nous comparons l'évolution en temps de l'erreur  $L_2$  et la distribution 1D suivant l'axe  $x$  pour  $y = 0.0$  m de la transformée de Fourier discrète de  $E_z$  issus des méthodes DGTD- $\mathbb{P}_2$  explicite et implicite pour valeurs du nombre CFL de 0.2 et 200.0 respectivement. Les temps de calculs sont résumés dans la Tab. 3.6 tandis que la Tab. 3.7 donne le temps de factorisation et la mémoire nécessaire au stockage des facteurs L et U. Nous observons ici une très importante diminution du temps de calcul d'un facteur de 72. Ces derniers résultats motivent les remarques suivantes :

- lorsque, comme c'est le cas avec la configuration de calcul considérée ici, le maillage de calcul présente des zones fortement raffinées très localisées, une stratégie de résolution hybride explicite/implicite (i.e. une méthode DGTD localement implicite) devrait s'avérer très avantageuse à la fois en termes de gains en temps de calcul par rapport à une approche globalement explicite, et en termes d'occupation mémoire par rapport à une approche globalement implicite. En effet, dans un tel cas, la matrice implicite à factoriser ne porterait que sur un faible pourcentage du nombre total d'éléments du maillage ce qui est clairement un point positif important dans l'optique de la simulation numérique de problèmes tridimensionnels.
- Le gain en temps de calcul observé est pour une grande part dû à l'utilisation d'un solveur direct creux pour la résolution du système implicite (3.3)-(3.4). Il est néanmoins peu probable qu'une telle stratégie de résolution soit applicable au cas tridimensionnel pour des raisons évidentes liées à l'espace mémoire nécessaire pour le stockage des facteurs L et U. Une méthodologie numérique hybride explicite/implicite s'impose donc comme la direction à suivre pour étendre la méthode DGTD- $\mathbb{P}_k$  implicite étudiée dans cette section à la simulation numérique de problèmes tridimensionnels.

Intégration en temps	Méthode	CFL- $\mathbb{P}_k$	Temps CPU
Explicite	DGTD- $\mathbb{P}_2$	0.2	19970 sec
Implicite	-	200.0	275 sec

TAB. 3.6 – Diffraction d'une onde plane par un cylindre diélectrique (C2)  
Temps CPU (station de travail basée sur un processeur AMD Opteron 2 GHz)

Méthode	CFL- $\mathbb{P}_k$	Temps CPU	Mémoire RAM
DGTD- $\mathbb{P}_2$	200.0	6 sec	177 MB

TAB. 3.7 – Diffraction d'une onde plane par un cylindre diélectrique (C2)  
Temps CPU (station de travail basée sur un processeur AMD Opteron 2 GHz)  
Temps de factorisation et mémoire nécessaire au stockage des facteurs L et U

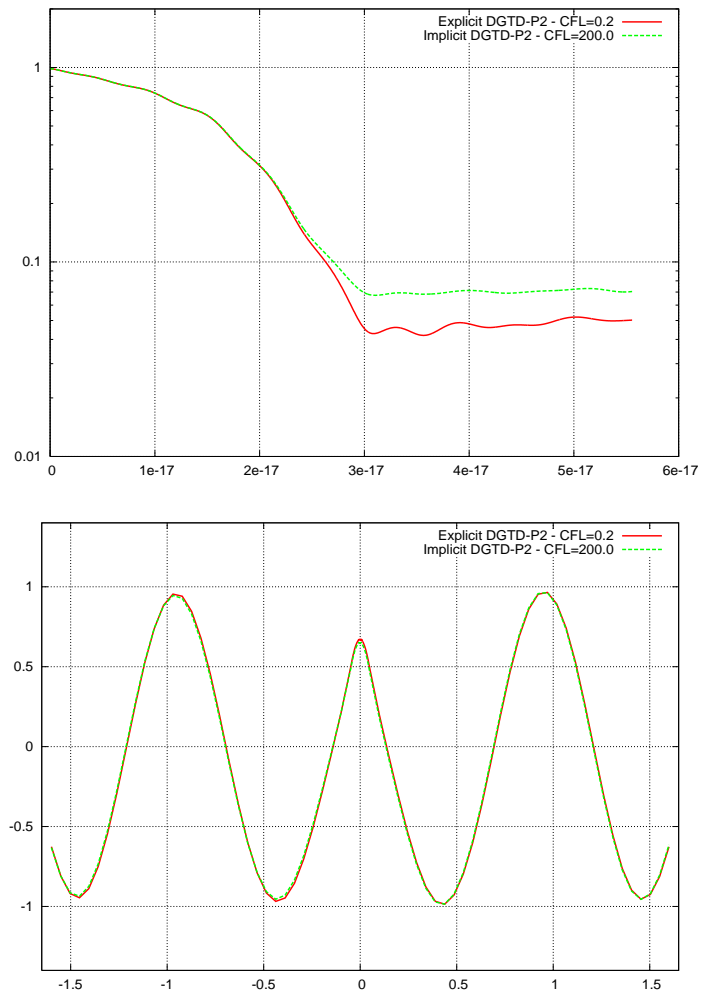


FIG. 3.26 – Diffraction d'une onde plane par un cylindre diélectrique (C2)

Haut : Evolution temporelle de l'erreur  $L_2$ Bas : Distribution 1D de la DFT( $E_z$ ),  $y = 0.0$  mMéthodes DGTD- $\mathbb{P}_2$  explicite et implicite

### 3.3 Méthode DGTD- $\mathbb{P}_k$ implicite du quatrième ordre

Les résultats numériques de la sous-section précédente ont montré que le schéma implicite Crank-Nicolson permet de diminuer le temps de calcul par rapport au schéma explicite de référence (schéma saute-mouton du second ordre) pour une erreur numérique de même ordre de grandeur. Néanmoins, on a aussi constaté que la stabilité inconditionnelle de la méthode DGTD- $\mathbb{P}_k$  implicite basée sur le schéma de Crank-Nicolson ne peut-être totalement exploitée, le nombre CFL (et donc le pas de temps) utilisé en pratique étant limité par des considérations de précision (voir par exemple les résultats du cas test de diffraction d'une onde plane par un carré, où le nombre CFL admissible décroît lorsque la fréquence de l'onde incidente augmente, pour un degré d'interpolation  $p$  donné). Nous étudions dans cette sous-section la possibilité de construire une méthode DGTD- $\mathbb{P}_k$  implicite précise au quatrième ordre afin d'autoriser l'utilisation de pas de temps sensiblement plus grands que ceux de la méthode DGTD- $\mathbb{P}_k$  implicite du second ordre.

Les schémas implicites les plus fréquemment utilisés sont les schémas d'Adams-Moulton dont le plus connu est Crank-Nicolson, ou les schémas aux différences rétrogrades (BDF pour *Backward differentiation Formulas*) [Curtiss and Hirschfelder, 1952] [Hairer et al., 1987, Hairer and Wanner, 1991]. Le schéma d'Euler rétrograde, qui est très utilisé, est le schéma d'ordre un de ces deux familles. Ces schémas sont tous d'ordre de précision peu élevé mais ils restent populaires principalement pour deux raisons : ils sont faciles à mettre en œuvre, et ils sont A-stables, c'est-à-dire que leur domaine de stabilité est dans le demi-plan gauche du plan complexe. Pour les schémas multipas linéaires A-stables, nous pouvons citer le résultat de Dahlquist [Dahlquist, 1963] qui montre que l'ordre de précision le plus élevé est 2. Pour des ordres de précision supérieurs, il faut s'orienter vers des schémas qui ont des régions de stabilité limitées, comme par exemple les schémas d'Adams-Moulton (qui sont la généralisation aux ordres plus élevés d'Euler rétrograde et de Crank-Nicolson). Cependant, par rapport à notre objectif, les régions de stabilité de ces schémas ne sont pas suffisamment grandes. Il y a une classe importante de schémas implicites de type Runge-Kutta qui n'ont pas ces restrictions, mais les schémas d'ordre élevé de cette famille sont difficiles à mettre en œuvre.

Les schémas basés sur l'extrapolation de Richardson sont faciles à implémenter. Pour un schéma basique d'ordre 2, nous pouvons obtenir un ordre de précision supérieur. Supposons que l'approximation d'une certaine quantité  $P$  peut se décomposer comme :

$$P = Q(h) + Ah^p + \mathcal{O}(h^{p+1}).$$

L'idée est donc d'éliminer le terme  $A$ . Pour ce faire, l'extrapolation de Richardson se base sur deux pas d'espace  $h_1$  et  $h_2$ . L'approximation de  $P$  devient :

$$P = \frac{h_2^p Q(h_1) - h_1^p Q(h_2)}{h_2^p - h_1^p} + \mathcal{O}(h^{p+1}).$$

La valeur de l'extrapolation de Richardson de  $P$  est alors :

$$P = \frac{h_2^p Q(h_1) - h_1^p Q(h_2)}{h_2^p - h_1^p}.$$

Ce processus peut alors être répété afin d'obtenir des ordres plus élevés. Le schéma est basé sur le principe que l'erreur est proportionnelle à une certaine puissance du pas de temps  $h$ .

Une autre approche, basée sur une amélioration itérative de la précision de l'approximation, a été proposée par Lindberg [Lindberg, 1980]. Ici, des opérateurs d'ordre élevé sont appliqués à l'erreur obtenue avec un schéma de bas ordre pour obtenir un schéma d'ordre globalement plus élevé. Cette approche est ainsi nommée *méthode du défaut corrigé*. Pour notre part, nous allons utiliser le principe du défaut corrigé

pour obtenir un schéma implicite d'ordre 4 à partir du schéma de Crank-Nicolson. Dans l'application de ce principe, on utilise la forme de l'erreur de troncature associée au schéma de Crank-Nicolson et non pas la forme exacte de l'erreur de la solution. Une approximation de cette erreur de troncature est introduite à chaque étape de l'algorithme, augmentant la précision de deux ordre à chaque pas de temps. Il est à noter que l'on utilise un seul et unique maillage pour une simulation donnée.

Le principe du défaut corrigé est bien connu pour les problèmes aux valeurs aux limites, tant pour les EDO que pour les EDP. On pourra notamment consulter les travaux de Pereyra [Pereyra, 1967a]-[Pereyra, 1967b]-[Pereyra, 1968]. Gustafsson et Hemmingsson-Frändén utilisent ce principe en espace dans le contexte d'une méthode de décomposition de domaine pour résoudre un problème de Poisson dans [Gustafsson and Hemmingsson-Frändén, 1999], et pour des problèmes dépendant du temps dans [Gustafsson and Hemmingsson-Frändén, 2000]-[Gustafsson and Hemmingsson-Frändén, 2002]. Une application du principe du défaut corrigé pour la construction de méthodes implicite d'ordre élevé pour les équations de Maxwell est proposée par Lee et Fornberg [Lee and Fornberg, 2004]

### 3.3.1 Formulation

Nous partons du système (2.3) que nous pouvons ré-écrire comme :

$$Q \frac{\partial \mathbf{W}}{\partial t} = A \mathbf{W}, \quad (3.39)$$

où  $A \mathbf{W}$  désigne formellement l'opération :

$$\begin{bmatrix} 0_{3 \times 3} & N_x \\ -N_x & 0_{3 \times 3} \end{bmatrix} \frac{\partial \mathbf{W}}{\partial x} + \begin{bmatrix} 0_{3 \times 3} & N_y \\ -N_y & 0_{3 \times 3} \end{bmatrix} \frac{\partial \mathbf{W}}{\partial y} + \begin{bmatrix} 0_{3 \times 3} & N_z \\ -N_z & 0_{3 \times 3} \end{bmatrix} \frac{\partial \mathbf{W}}{\partial z}.$$

En appliquant le schéma de Crank-Nicolson pour l'intégration en temps de (3.39) nous obtenons :

$$Q \frac{\mathbf{W}^{n+1} - \mathbf{W}^n}{\Delta t} = \frac{1}{2} A (\mathbf{W}^{n+1} + \mathbf{W}^n) + O(\Delta t^2), \quad (3.40)$$

soit encore :

$$\left( \frac{2}{\Delta t} Q - A \right) \mathbf{W}^{n+1} = \left( \frac{2}{\Delta t} Q + A \right) \mathbf{W}^n + O(\Delta t^2). \quad (3.41)$$

A partir de l'expression (3.41), nous pouvons déduire que l'erreur commise est :

$$\begin{aligned} \mathcal{E} &= \left[ \left( \frac{2}{\Delta t} Q - A \right) \mathbf{W}^{n+1} - \left( \frac{2}{\Delta t} Q + A \right) \mathbf{W}^n \right] \\ &= \left[ \frac{2}{\Delta t} Q (\mathbf{W}^{n+1} - \mathbf{W}^n) - A (\mathbf{W}^{n+1} + \mathbf{W}^n) \right]. \end{aligned} \quad (3.42)$$

En utilisant les développements de Taylor :

$$\begin{aligned} \mathbf{W}^{n+1} - \mathbf{W}^n &= \Delta t \mathbf{W}_t^{n+1/2} + \frac{(\Delta t)^3}{24} \mathbf{W}_{ttt}^{n+1/2} + \frac{(\Delta t)^5}{1920} \mathbf{W}_{tttt}^{n+1/2} + O(\Delta t^7), \\ \mathbf{W}^{n+1} + \mathbf{W}^n &= 2 \mathbf{W}^{n+1/2} + \frac{(\Delta t)^2}{4} \mathbf{W}_{tt}^{n+1/2} + \frac{(\Delta t)^4}{192} \mathbf{W}_{tttt}^{n+1/2} + O(\Delta t^6), \end{aligned} \quad (3.43)$$

nous obtenons le développement de l'erreur suivant :

$$\mathcal{E} = 2 \left( Q \mathbf{W}_t^{n+1/2} - A \mathbf{W}^{n+1/2} \right) + \left( -\frac{A}{4} \mathbf{W}_{tt}^{n+1/2} + \frac{Q}{12} \mathbf{W}_{ttt}^{n+1/2} \right) \Delta t^2 + \dots \quad (3.44)$$

Le premier terme du membre de droite de (3.44) disparaît par (3.39) alors que le second membre est à exploiter pour obtenir une méthode de défaut corrigé d'ordre 4. On développe maintenant les différentes composantes de ce terme d'erreur. On a :

$$\begin{aligned} -\frac{\Delta t^2}{4} A \mathbf{W}_{tt}^{n+1/2} &= -\frac{\Delta t^2}{4} A \frac{(\mathbf{W}_t^{n+1} - \mathbf{W}_t^n)}{\Delta t} \\ &= -\frac{\Delta t^2}{4} A \frac{(\mathbf{W}^{n+3/2} - \mathbf{W}^{n+1/2} - (\mathbf{W}^{n+1/2} - \mathbf{W}^{n-1/2}))}{\Delta t^2} \\ &= -\frac{\Delta t^2}{4} A \frac{(\mathbf{W}^{n+3/2} - 2\mathbf{W}^{n+1/2} + \mathbf{W}^{n-1/2})}{\Delta t^2} \\ &= -\frac{\Delta t^2}{4} A \frac{\left( \frac{1}{2} (\mathbf{W}^{n+2} + \mathbf{W}^{n+1}) - (\mathbf{W}^{n+1} + \mathbf{W}^n) + \frac{1}{2} (\mathbf{W}^n + \mathbf{W}^{n-1}) \right)}{\Delta t^2} \\ &= -\frac{1}{8} A (\mathbf{W}^{n+2} - \mathbf{W}^{n+1} - \mathbf{W}^n + \mathbf{W}^{n-1}). \end{aligned}$$

De même :

$$\begin{aligned} \frac{\Delta t^2}{12} Q \mathbf{W}_{ttt}^{n+1/2} &= \frac{\Delta t^2}{12} Q \frac{(\mathbf{W}_{tt}^{n+1} - \mathbf{W}_{tt}^n)}{\Delta t} \\ &= \frac{\Delta t^2}{12} Q \frac{(\mathbf{W}_t^{n+3/2} - \mathbf{W}_t^{n+1/2} - (\mathbf{W}_t^{n+1/2} - \mathbf{W}_t^{n-1/2}))}{\Delta t^2} \\ &= \frac{\Delta t^2}{12} Q \frac{(\mathbf{W}_t^{n+3/2} - 2\mathbf{W}_t^{n+1/2} + \mathbf{W}_t^{n-1/2})}{\Delta t^2} \quad (3.45) \\ &= \frac{\Delta t^2}{12} Q \frac{(\mathbf{W}^{n+2} - \mathbf{W}^{n+1} - 2(\mathbf{W}^{n+1} - \mathbf{W}^n) + \mathbf{W}^n - \mathbf{W}^{n-1})}{\Delta t^3} \\ &= \frac{1}{12\Delta t} Q (\mathbf{W}^{n+2} - 3\mathbf{W}^{n+1} + 3\mathbf{W}^n - \mathbf{W}^{n-1}). \end{aligned}$$

On en déduit donc que pour augmenter la précision de la méthode implicite initiale jusqu'à l'ordre 4 il nous faut prendre en compte le terme d'erreur :

$$\begin{aligned} \mathcal{E} &\equiv -\frac{1}{8} A (\mathbf{W}^{n+2} - \mathbf{W}^{n+1} - \mathbf{W}^n + \mathbf{W}^{n-1}) + \\ &\quad \frac{1}{12\Delta t} Q (\mathbf{W}^{n+2} - 3\mathbf{W}^{n+1} + 3\mathbf{W}^n - \mathbf{W}^{n-1}). \end{aligned} \quad (3.46)$$

Soit  $\mathbf{W}^{4,n}$  le champ évalué au temps  $t_n = n\Delta t$  solution des équations de Maxwell précisée à l'ordre 4 en temps. Ce champ résulte de la résolution du système :

$$\left(\frac{2}{\Delta t}Q - A\right) \mathbf{W}^{4,n+1} = \left(\frac{2}{\Delta t}Q + A\right) \mathbf{W}^{4,n} + \mathcal{E}, \quad (3.47)$$

où l'évaluation du terme d'erreur  $\mathcal{E}$  requiert la connaissance des champs  $\mathbf{W}^{4,n}$ ,  $\mathbf{W}^{2,n+2}$ ,  $\mathbf{W}^{2,n+1}$ ,  $\mathbf{W}^{2,n}$  et  $\mathbf{W}^{2,n-1}$ . Ici,  $\mathbf{W}^{2,n}$  désigne le champ précis au second ordre en temps. L'algorithme utilisé est alors le suivant :

1. Initialisations

–  $\mathbf{W}^{2,0}$  est donné. On calcule  $\mathbf{W}^{2,1}$  et on pose  $\mathbf{W}^{4,1} = \mathbf{W}^{2,1}$ . On calcule  $\mathbf{W}^{2,2}$ .

2. Boucle en temps

–  $\mathbf{W}^{2,n-1}$ ,  $\mathbf{W}^{2,n}$ ,  $\mathbf{W}^{2,n+1}$  et  $\mathbf{W}^{4,n}$  sont connus.

– Résolution du système :

$$\left(\frac{2}{\Delta t}Q - A\right) \mathbf{W}^{2,n+2} = \left(\frac{2}{\Delta t}Q + A\right) \mathbf{W}^{2,n+1}. \quad (3.48)$$

– Evaluation de l'erreur :

$$\begin{aligned} \mathcal{E} = & -\frac{1}{8}A(\mathbf{W}^{2,n+2} - \mathbf{W}^{2,n+1} - \mathbf{W}^{2,n} + \mathbf{W}^{2,n-1}) + \\ & \frac{1}{12\Delta t}Q(\mathbf{W}^{2,n+2} - 3\mathbf{W}^{2,n+1} + 3\mathbf{W}^{2,n} - \mathbf{W}^{2,n-1}). \end{aligned} \quad (3.49)$$

– Résolution du système :

$$\left(\frac{2}{\Delta t}Q - A\right) \mathbf{W}^{4,n+1} = \left(\frac{2}{\Delta t}Q + A\right) \mathbf{W}^{4,n} + \mathcal{E}. \quad (3.50)$$

La méthode de défaut corrigé implicite d'ordre 4 ci-dessus nécessite la résolution de deux systèmes linéaires au lieu d'un seul système pour la méthode implicite d'ordre 2 basée sur le schéma de Crank-Nicolson. Ces deux systèmes mettent en jeu la même matrice qui peut donc être factorisée une fois pour toutes si une méthode directe est adoptée pour la résolution de ces deux systèmes.

### 3.3.2 Résultats numériques en 2D

Dans cette section, nous présentons des résultats numériques afin de comparer les méthodes DGTD- $\mathbb{P}_k$  implicites du second et du quatrième ordre. Les schémas d'intégration en temps correspondants sont référencés comme CN2 et CN4. On montre aussi des comparaisons avec la méthode DGTD- $\mathbb{P}_k$  explicite basée sur le schéma LF2 (saute-mouton du second ordre). Pour cette étude numérique, nous ne considérons que le cas test de la propagation d'un mode propre dans une cavité dont les parois sont parfaitement conductrices (voir aussi la sous-section 2.4.1.1). On utilise trois maillages triangulaires non-uniformes dont les caractéristiques sont résumées dans la Tab. 3.8. Dans cette table,  $L_{\max}/L_{\min}$  désigne le rapport de taille entre la plus grande et la plus petite arête du maillage correspondant. Ces maillages sont donc modérément non-uniformes.

Les Fig. 3.27 à Fig. 3.29 comparent l'évolution en temps de l'erreur  $L_2$  pour des simulations réalisées avec le maillage M2 par les différentes méthodes. Par ailleurs, on résume dans la Tab. 3.9 quelques éléments

Maillage	# sommets	# triangles	$L_{\max}/L_{\min}$
M1	81	144	14.3
M2	309	584	14.5
M3	1148	2230	12.1

TAB. 3.8 – Mode propre dans un carré unité - Maillages non-uniformes

permettant d'évaluer les performances des différentes méthodes. Enfin, la Fig. 3.30 présente les courbes de convergence numérique, les taux de convergence correspondants étant résumés dans la Tab. 3.10. Pour cette étude de convergence numérique on a utilisé la séquence de maillages de la Tab. 3.8. Il est clair que le schéma CN4 permet d'améliorer sensiblement la précision globale de la méthode DGTD- $\mathbb{P}_k$  implicite dès lors que  $p \geq 3$ . On constate aussi que l'écart de précision entre les méthodes DGTD- $\mathbb{P}_k$  basées sur les schémas CN2 et CN4 est nettement plus notable pour  $p = 4$ . Des calculs sur des maillages fortement raffinés permettraient certainement de mieux évaluer l'apport du schéma CN4 en termes de précision, néanmoins le surcoût de calcul lié à la résolution de deux systèmes linéaires à chaque pas de temps risque de ne pas permettre un gain notable en temps de calcul. Un résultat préliminaire dans cette direction est montré sur la Fig. 3.31 pour des calculs réalisés avec la méthode DGTD- $\mathbb{P}_4$  combinée aux schémas LF2, CN2 et CN4, et le maillage non-uniforme déjà utilisé à la sous-section 3.2.3.1. Les temps de calcul associés sont rassemblés dans la Tab. 3.11.

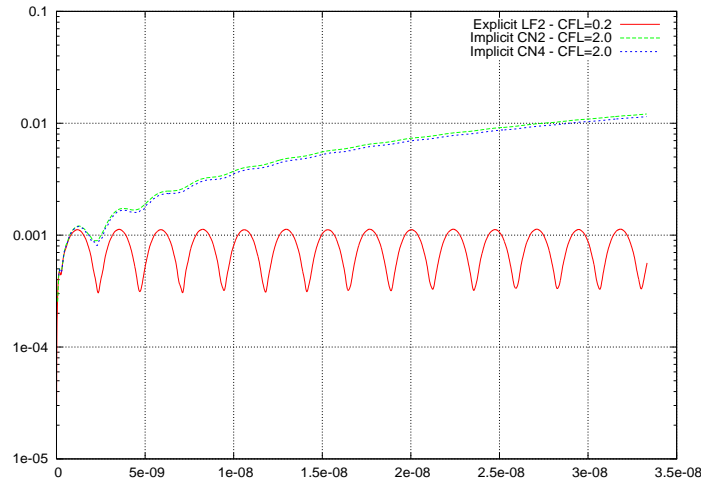


FIG. 3.27 – Mode propre dans un carré unité  
 Maillage non-uniforme : # sommets = 309 et # triangles = 584  
 Evolution temporelle de l'erreur  $L_2$  - Méthode DGTD- $\mathbb{P}_2$   
 Abscisse : temps en sec - Ordonnée : erreur  $L_2$  en échelle log

### 3.4 Conclusion

Nous avons proposé dans ce chapitre deux versions globalement implicites de la méthode DGTD- $\mathbb{P}_k$ . La méthode CN2-DGTD- $\mathbb{P}_k$  basée sur le schéma de Crank-Nicolson du second ordre a été étudiée



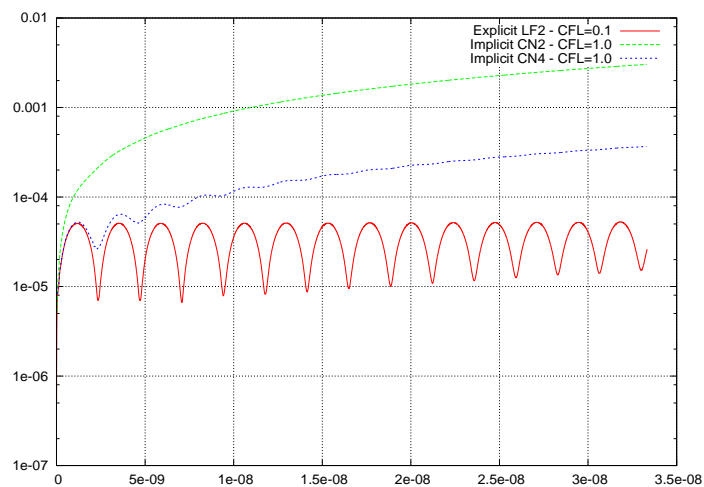


FIG. 3.28 – Mode propre dans un carré unité  
 Maillage non-uniforme : # sommets = 309 et # triangles = 584  
 Evolution temporelle de l'erreur  $L_2$  - Méthode DGTD- $\mathbb{P}_3$   
 Abscisse : temps en sec - Ordonnée : erreur  $L_2$  en échelle log

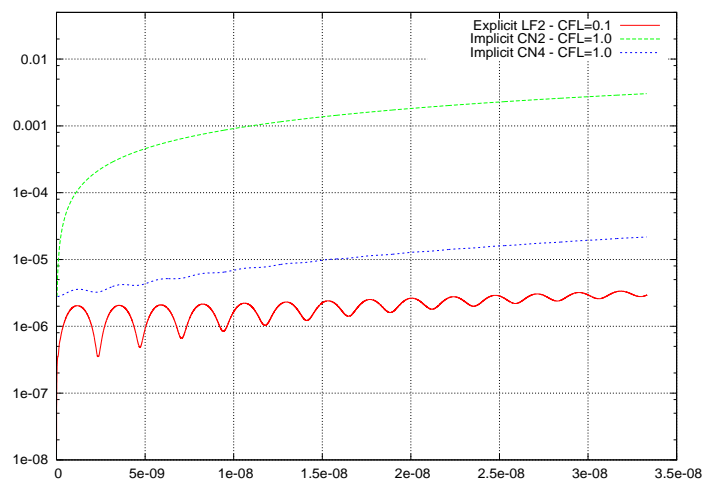


FIG. 3.29 – Mode propre dans un carré unité  
 Maillage non-uniforme : # sommets = 309 et # triangles = 584  
 Evolution temporelle de l'erreur  $L_2$  - Méthode DGTD- $\mathbb{P}_4$   
 Abscisse : temps en sec - Ordonnée : erreur  $L_2$  en échelle log

Schéma en temps	Méthode	CFL	Erreur $L_2$ finale	Temps CPU
LF2	DGTD- $\mathbb{P}_2$	0.2	$0.5640 \cdot 10^{-3}$	11.0 sec
-	DGTD- $\mathbb{P}_3$	0.1	$0.2611 \cdot 10^{-4}$	46.8 sec
-	DGTD- $\mathbb{P}_4$	0.05	$0.2946 \cdot 10^{-5}$	181.3 sec
CN2	DGTD- $\mathbb{P}_2$	2.0	$0.1147 \cdot 10^{-1}$	9.4 sec
-	DGTD- $\mathbb{P}_3$	1.0	$0.3034 \cdot 10^{-2}$	38.5 sec
-	DGTD- $\mathbb{P}_4$	1.0	$0.3034 \cdot 10^{-2}$	69.4 sec
CN4	DGTD- $\mathbb{P}_2$	2.0	$0.1209 \cdot 10^{-1}$	25.2 sec
-	DGTD- $\mathbb{P}_3$	1.0	$0.3680 \cdot 10^{-3}$	99.6 sec
-	DGTD- $\mathbb{P}_4$	1.0	$0.2174 \cdot 10^{-4}$	178.5 sec

TAB. 3.9 – Mode propre dans un carré unité - Maillage non-uniforme  
Maillage M2 : # sommets = 309 et # triangles = 584

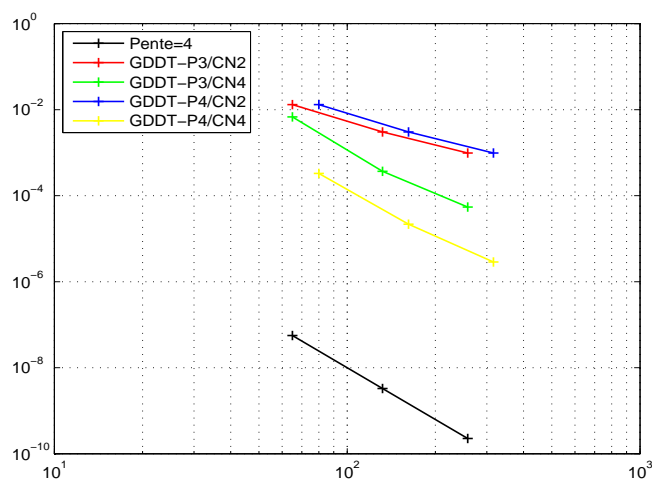


FIG. 3.30 – Mode propre dans un carré unité - Maillages non-uniformes  
Schémas Crank-Nicolson du second ordre CN2 et du quatrième ordre CN4  
Convergence numérique (erreur  $L_2$  en fonction de  $\sqrt{\#DOF}$  en échelle log-log)

Schéma en temps	Méthode	Ordre de convergence
CN2	DGTD- $\mathbb{P}_3$	1.88
-	DGTD- $\mathbb{P}_4$	1.89
CN4	DGTD- $\mathbb{P}_3$	3.50
-	DGTD- $\mathbb{P}_4$	3.45

TAB. 3.10 – Mode propre dans un carré unité - Maillages non-uniformes  
Schémas Crank-Nicolson du second ordre CN2 et du quatrième ordre CN4

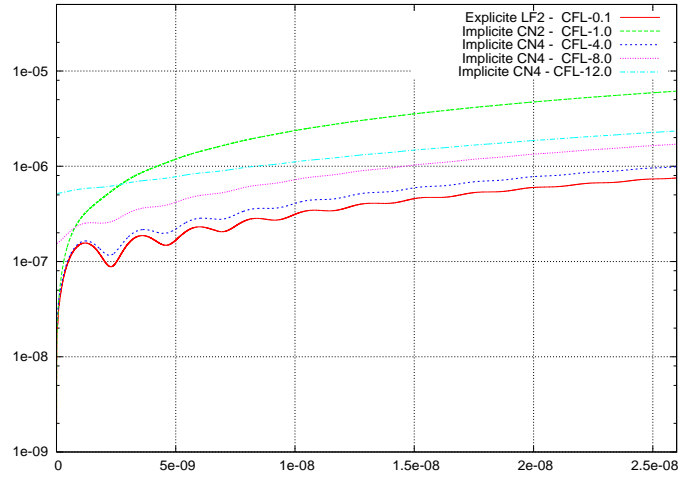


FIG. 3.31 – Mode propre dans un carré unité - Méthode DGTD- $\mathbb{P}_4$   
 Maillage non-uniforme : # sommets = 1400 et # triangles = 2742  
 Evolution temporelle de l'erreur  $L_2$  - DGTD- $\mathbb{P}_4$

Schéma en temps	CFL	Erreur $L_2$ finale	Temps CPU
LF2	0.1	$0.9690 \cdot 10^{-6}$	9225 sec
CN2	1.0	$0.6417 \cdot 10^{-5}$	7164 sec
CN4	4.0	$0.1264 \cdot 10^{-5}$	4485 sec
-	8.0	$0.2167 \cdot 10^{-5}$	2279 sec
-	12.0	$0.2987 \cdot 10^{-5}$	1497 sec

TAB. 3.11 – Mode propre dans un carré unité - Méthode DGTD- $\mathbb{P}_4$   
 Maillage non-uniforme : # sommets = 1400 et # triangles = 2742  
 Abscisse : temps en sec - Ordonnée : erreur  $L_2$  en échelle log

théoriquement (stabilité inconditionnelle, inversibilité du système implicite associé et dispersion numérique en 1D) et numériquement dans le contexte de la résolution des équations de Maxwell 2D. Des gains conséquents en termes de temps de calcul ont été obtenus pour des calculs en maillages non-uniformes. Néanmoins, la précision de cette méthode est limitée par la précision du second ordre du schéma de Crank-Nicolson qui, de plus, est caractérisé par une erreur de dispersion plus importante que celle du schéma explicite saute-mouton (LF2). Nous avons donc étudié la possibilité d'augmenter la précision en temps de la méthode DGTD- $\mathbb{P}_k$  implicite par l'intermédiaire d'une technique de défaut corrigé. La méthode CN4-DGTD- $\mathbb{P}_k$  résultante est effectivement plus précise mais le coût de calcul par itération en temps est doublé du fait de la résolution de deux systèmes linéaires au lieu d'un pour la méthode CN2-DGTD- $\mathbb{P}_k$ . Une exploitation efficace de la méthode CN4-DGTD- $\mathbb{P}_k$  semble plus envisageable dans le contexte d'une stratégie de résolution localement implicite telle que celle étudiée dans le chapitre suivant pour le couplage des schémas LF2 et CN2.

# Chapitre 4

## Résolution des systèmes implicites

### 4.1 Introduction

Les méthodes DGTD- $\mathbb{P}_k$  implicites étudiées au chapitre 3 nécessitent la résolution d'un ou plusieurs systèmes linéaires à chaque itération temporelle. Le choix de la méthode de résolution de ces systèmes linéaires est une question cruciale au regard des performances (numériques et computationnelles) de la méthode DGTD- $\mathbb{P}_k$  implicite. Ce chapitre est consacré à une étude numérique de la résolution des systèmes linéaires issus de la méthode DGTD- $\mathbb{P}_k$  implicite basée sur le schéma de Crank-Nicolson appliquée aux équations de Maxwell en deux dimensions d'espace. Dans ce contexte, on considère et compare ici deux stratégies de résolution : une méthode directe adaptée aux matrices creuses et une méthode itérative préconditionnée. Ces méthodes permettent de résoudre un système linéaire :

$$Ax = b, \tag{4.1}$$

où  $A \in \mathbb{R}^{n \times n}$  est une matrice générale,  $x \in \mathbb{R}^n$  et  $b \in \mathbb{R}^n$ . Dans le cas des méthodes DGTD- $\mathbb{P}_k$  implicites du chapitre 3, la matrice  $A$  est la somme de matrices symétriques et d'une matrice antisymétrique.  $A$  est donc une matrice quelconque. Toutefois, les coefficients de cette matrice ne dépendent pas de la variable temporelle, propriété que l'on exploite en pratique en réalisant certains traitements propres aux stratégies de résolution considérées, en amont de la boucle de calcul principale (itération temporelle).

La première stratégie de résolution est une méthode directe basée sur une factorisation LU. Elle permet de calculer la solution exacte du système (4.1). Performance et robustesse sont ses principaux avantages. En revanche, elle peut être très gourmande en ressource mémoire, particulièrement dans le cas tridimensionnel. La seconde stratégie permet, a priori, de pallier cette difficulté. Elle consiste à résoudre le système (4.1) suivant un processus itératif et il existe une grande variété de méthodes qui mettent en œuvre un tel processus. Les méthodes itératives conduisent à des solutions approchées. La plupart du temps, leur convergence est très sensible aux propriétés de la matrice sous-jacente. La convergence de certaines de ces méthodes peut être étudiée analytiquement permettant de définir les classes de matrices auxquelles elle s'appliquent idéalement. Quelques fois, la convergence d'une méthode itérative dépend du nombre de conditionnement de la matrice et dans presque tous les cas la prise en compte d'un préconditionneur est indispensable pour assurer la convergence. Un préconditionneur est une matrice  $M$  régulière qui approche  $A^{-1}$ . On est alors amené à résoudre le système linéaire transformé :

$$MAx = Mb. \tag{4.2}$$

Le système 4.2 est préconditionné à gauche, mais il peut l'être aussi à droite :

$$AMt = b, \quad (4.3)$$

voire simultanément à gauche et à droite. Une fois la solution  $t$  de (4.3) obtenue, la solution du système (4.1) est déduite de  $x = Mt$ . Nous utiliserons dans cette étude un préconditionnement à gauche pour améliorer l'efficacité de la résolution itérative.

Le plan de ce chapitre est le suivant : dans un premier temps (section 4.2), nous présentons brièvement les deux stratégies de résolution considérées dans cette étude à savoir, le solveur direct MUMPS et la méthode itérative GMRES associée au préconditionneur ILUT. Nous enchaînons par une étude comparative de ces deux stratégies de résolution (section 4.3) en termes d'efficacité numérique, précision et performances (temps de calcul et occupation mémoire). Nous concluons ce chapitre en résumant les principaux enseignements de cette étude numérique.

## 4.2 Stratégies de résolution

### 4.2.1 Le solveur direct MUMPS

Le solveur direct MUMPS (*MUltifrontal Massively Parallel sparse direct Solver*) [Amestoy et al., 1998]-[Amestoy et al., 2001] a été initialement développé durant le projet européen PARASOL (1996-1999) par le CERFACS, l'ENSEEIH-IRIT et le RAL. Depuis 1999, MUMPS est co-développé par différents organismes : le CERFACS, l'ENSEEIH-IRIT et l'INRIA. MUMPS est un solveur direct parallélisé pour matrices creuses à structure irrégulière. La matrice à factoriser peut être symétrique définie positive, symétrique ou anti-symétrique, à coefficients réels ou complexes. MUMPS est basé sur une technique multifrontale pour réaliser une factorisation LU ou LDLT de la matrice. La résolution se fait en trois phases : analyse symbolique, factorisation numérique et résolution. Pendant la phase d'analyse symbolique plusieurs prétraitements sont appliqués, comme la permutation des lignes et colonnes de la matrice, dans le but de minimiser le nombre de termes non-nuls introduits durant la phase de factorisation. Cette dernière est guidée par le résultat de la phase d'analyse symbolique. Dans cette phase, on peut aussi être amené à appliquer différentes techniques de pivotage numérique afin d'assurer autant que faire se peut la stabilité de la factorisation. MUMPS exploite à la fois le parallélisme dû au caractère creux des matrices et celui présent dans les noyaux de factorisation dense (BLAS). La phase de résolution applique des substitutions amont et aval sur les facteurs calculés. Le logiciel est écrit en Fortran 90 et une interface en C est disponible. La version parallèle de MUMPS nécessite la librairie MPI (Message Passing Interface) pour les étapes de communications entre processeurs d'un ordinateur parallèle, et utilise les bibliothèques BLACS et ScaLAPACK pour exploiter des noyaux d'algèbre linéaire dense parallélisés.

### 4.2.2 La méthode GMRES

La méthode GMRES (Generalized Minimum RESidual) a été proposée par Saad et Schultz en 1986 [Saad and Schultz, 1986] pour la résolution de systèmes linéaires à matrices non-hermitiennes. GMRES appartient à la classe des méthodes de Krylov [Saad, 2003]. On rappelle ici les principes de base et principales étapes du processus itératif associé (voir aussi l'ouvrage de Saad [Saad, 2003]).

Soit  $x_0$  un vecteur de dimension  $n$  qui représente l'initialisation de la résolution itérative du système linéaire (4.1) et  $r_0 = b - Ax_0$  le résidu associé. A l'étape  $j$ , l'algorithme GMRES construit une approximation de la solution de (4.1) de la forme :

$$x = x_0 + V_m y, \quad (4.4)$$

où  $y$  est un vecteur de dimension  $m$  et  $V_m = [q_1, \dots, q_m]$  est une base orthogonale de l'espace de Krylov de dimension  $m$  définie par :

$$\mathcal{K}(A, r_0, m) = \text{span}\{r_0, Ar_0, \dots, A^{m-1}r_0\}. \quad (4.5)$$

Le vecteur  $y$  est déterminé par la minimisation de la norme euclidienne du résidu  $r = b - Ax$  sur l'espace  $x_0 + \mathcal{K}(A, r_0, m)$ . La base  $V_m$  du sous-espace de Krylov  $\mathcal{K}(A, r_0, m)$  est obtenue grâce au procédé d'Arnoldi [Arnoldi, 1951]. La projection orthogonale de  $A$  sur  $\mathcal{K}(A, r_0, m)$  conduit à une matrice d'Hessenberg supérieure  $H_m = {}^*V_m A V_m$  d'ordre  $j$ . Le procédé d'Arnoldi satisfait la relation :

$$A V_m = V_m H_m + h_{m+1,m} v_{m+1}^t e_m, \quad (4.6)$$

où  $e_m$  est le  $m^{\text{ème}}$  vecteur de la base canonique. L'équation (4.6) peut être ré-écrite sous forme matricielle comme :

$$A V_m = V_{m+1} \tilde{H}_m, \quad (4.7)$$

où :

$$\tilde{H}_m = \begin{bmatrix} & & & H_m & \\ 0 & \dots & 0 & h_{m+1,m} & \end{bmatrix},$$

est une matrice  $(m+1) \times m$ . Soit  $v_1 = r_0/\beta$  où  $\beta = \|r_0\|_2$ . Le résidu  $r_m$  associé à la solution approchée  $x_m$ , définie par (4.4) satisfait :

$$\begin{aligned} r_m &= b - A x_m = b - A(x_0 + V_m y) \\ &= r_0 - A V_m y = r_0 - V_{m+1} \tilde{H}_m y \\ &= \beta v_1 - V_{m+1} \tilde{H}_m y \\ &= V_{m+1}(\beta e_1 - \tilde{H}_m y). \end{aligned}$$

Puisque  $V_{m+1}$  est une matrice dont les colonnes sont orthonormales, la norme du résidu  $\|r\|_2 = \|\beta e_1 - \tilde{H}_m y\|_2$  est minimisée lorsque  $y$  est solution du problème de moindres carrés :

$$\min_{y \in \mathbb{C}^m} \|\beta e_1 - \tilde{H}_m y\|_2. \quad (4.8)$$

Supposons que  $y_j$  soit solution de (4.8). Alors,  $x = x_0 + V_m y$  est une solution approchée de (4.1) pour laquelle le résidu est minimisé sur  $x_0 + \mathcal{K}(A, r_0, m)$ . La méthode GMRES tire son nom de la propriété de minimisation qui est l'élément clé permettant de diminuer la norme du résidu en itérant.

En arithmétique exacte, la méthode GMRES converge en exactement  $n$  itérations. Cependant, en pratique,  $n$  est potentiellement très grand et le stockage de la base orthonormale  $V_m$  peut devenir prohibitif. De plus, l'orthogonalisation de  $v_m$  avec ces prédécesseurs,  $v_1, \dots, v_{m-1}$  nécessite  $4nm$  opérations, ce qui pour des  $m$  grand, engendre des coûts de calculs particulièrement élevés. La variante GMRES( $m$ ) permet de pallier ces deux inconvénients. Pour un  $m$  fixé, la méthode GMRES( $m$ ) calcule une séquence de solutions

approchées  $x_j$ , jusqu'à ce que  $x_j$  soit acceptable ou que  $j = m$ . Si lorsque  $j = m$  la solution approchée n'est pas acceptable alors la méthode GMRES est redémarrée à partir de la dernière approximation calculée, c'est-à-dire avec  $x_0 = x_m$ , pour se conformer à la propriété de monotonicité qui veut que la norme du résidu décroisse même lorsque la méthode est réinitialisée. Le procédé est itératif jusqu'à ce qu'une approximation suffisamment précise soit trouvée.

Nous détaillons ci-dessous l'algorithme de base de GMRES. Pour plus de détails sur les différentes variantes on se référera à l'ouvrage de Saad [Saad, 2003].

**Algorithme 1** Méthode GMRES.

1. Calculer  $r_0 = b - Ax_0$ ,  $\beta := \|r_0\|_2$  et  $v_1 := r_0/\beta$
2. Définir la matrice  $\overline{H}_m = \{h_{ij}\}_{1 \leq i \leq m+1, 1 \leq j \leq m}$  de taille  $(m+1) \times m$  et poser  $\overline{H}_m = 0$
3. Pour  $j = 1, 2, \dots, m$  Faire
  - (a) Calculer  $w_j := Av_j$
  - (b) Pour  $i=1, \dots, j$  Faire :
    - i.  $h_{ij} = (w_j, v_i)$
    - ii.  $w_j := w_j - h_{ij}v_i$
  - (c) Fin de la boucle sur  $i$
  - (d) Calculer  $h_{j+1,j} = \|w_j\|_2$ . Si  $h_{j+1,j} = 0$ , poser  $m := j$  et aller à 5
  - (e) Calculer  $v_{j+1} = w_j/h_{j+1,j}$
4. Fin de la boucle sur  $j$
5. Calculer  $y_m$  qui minimise  $\|\beta e_1 - \tilde{H}_j y_j\|_2$  et poser  $x_m = x_0 + V_m y_m$

### 4.2.3 Méthode de factorisation incomplète ILUT

Pour améliorer la convergence de la méthode GMRES( $m$ ) nous avons choisi de lui associer la technique de préconditionnement ILUT proposée par Saad [Saad, 1994]. Il s'agit d'une méthode de factorisation LU incomplète qui utilise deux mécanismes de sélection des termes non-nuls lors de la construction des facteurs L et U incomplets. L'algorithme correspondant est décrit ci-dessous où  $a_{i*}$  désigne la  $i^{\text{ème}}$  ligne de  $A$  et  $w$  le vecteur ligne qui est utilisée pour stocker les combinaisons linéaires des lignes creuses lors du procédé d'élimination ( $w_k$  désigne le  $k^{\text{ème}}$  élément de cette ligne). Dans cet algorithme, règles suivantes sont utilisées :

- lors de l'opération (b)(ii.),  $w_k$  est mis à zéro s'il est inférieur en module à une tolérance relative  $\tau_i$  obtenue en multipliant  $\tau$  par la norme de la  $i^{\text{ème}}$  ligne.
- Lors de l'opération (d), une autre règle de rejet est appliquée. Tout d'abord, on annule tout élément de la ligne inférieur en module à  $\tau_i$ . Puis, on garde seulement les  $p$  plus grands éléments de la partie  $L$  de la ligne et les  $p$  plus grands éléments de la partie  $U$  de la ligne, en plus des éléments diagonaux qui sont toujours conservés.



**Algorithme 2** Factorisation incomplète ILUT( $\tau, p$ ).

1. Pour  $i = 1, \dots, n$  Faire
  - (a)  $w := a_{i*}$
  - (b) Pour  $k = 1, \dots, i - 1$  et quand  $w_k \neq 0$  Faire
    - i.  $w_k := w_k / a_{kk}$
    - ii. Appliquer une règle de rejet à  $w_k$
    - iii. Si  $w_k \neq 0$  alors  $w := w - w_k * u_{k*}$
  - (c) Fin de la boucle
  - (d) Appliquer une règle de rejet aux éléments de  $w$
  - (e)  $l_{ij} := w_j$  pour  $j = 1, \dots, i - 1$
  - (f)  $u_{ij} := w_j$  pour  $j = i, \dots, n$
  - (g)  $w := 0$
2. Fin de la boucle

### 4.3 Etude numérique

Les deux stratégies de résolution de systèmes linéaires discutées dans la section précédente sont ici évaluées numériquement dans le contexte de la méthode DGTD- $\mathbb{P}_k$  implicite basée sur le schéma de Crank-Nicolson introduite au chapitre 3, appliquée à la résolution des équations de Maxwell en deux dimensions d'espace. Le cas test sélectionné est celui qui traite de la propagation d'un mode propre dans une cavité dont les parois sont parfaitement conductrices (voir aussi la sous-section 2.4.1.1). Pour ce cas test, une solution analytique est connue et l'on peut donc évaluer l'influence des paramètres des stratégies de résolution considérées sur l'erreur numérique globale. Nous utilisons deux maillages triangulaires (voir la Fig. 4.1) dont les caractéristiques sont précisées dans le Tab. 4.3. On souhaite notamment évaluer l'impact de la non-uniformité du maillage sur les caractéristiques des factorisations LU et ILUT obtenues.

Toutes les simulations numériques sont réalisées sur une station de travail basée sur un processeur Intel Xeon 2.33 GHz avec 2 GB de mémoire RAM. Dans ce qui suit, sauf mention explicite, le terme «Erreur  $L_2$ » fait toujours référence à l'erreur commise entre la solution qui résulte de la méthode DGTD- $\mathbb{P}_k$  implicite et la solution analytique des équations de Maxwell 2D pour le cas test sélectionné. En particulier, dans les tables de résultats, il s'agit de l'erreur à la fin d'une simulation (i.e. au temps final). Par convention, les nombres CFL utilisés n'excèdent pas les valeurs qui conduisent à des solutions numériques comparables avec celles résultantes de la méthode DGTD- $\mathbb{P}_k$  explicite (voir les comparaisons menées dans le chapitre 3).

Rappelons les différents paramètres de résolution susceptibles de varier dans nos comparaisons :

- pour la méthode GMRES( $m$ )
  - $m$  : la taille de la base de Krylov,
  - le nombre maximum d'itérations,
  - $\varepsilon$  : le seuil de résolution linéaire (seuil de réduction du résidu initial i.e. tel que  $\frac{\|r_f\|}{\|r_0\|} < \varepsilon$ ).
- pour le préconditionneur ILUT( $\tau, p$ )
  - $p$  : le nombre maximum d'éléments non-nuls dans une ligne des facteurs L et U,

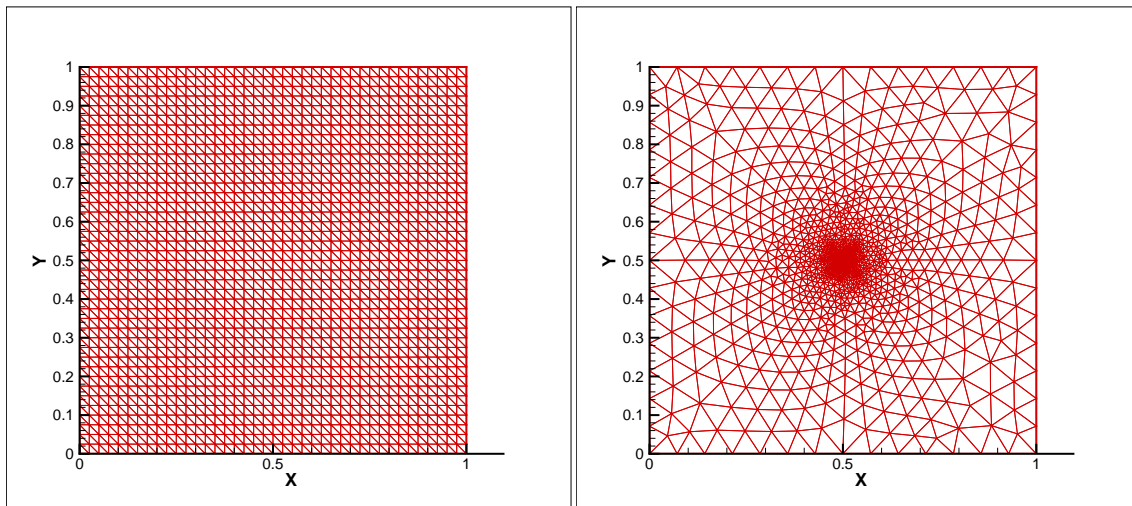


FIG. 4.1 – Maillages uniforme (gauche) et non uniforme (droite) du carré unitaire

	Maillage uniforme	Maillage non-uniforme
Nombre de sommets	1369	1400
Nombre de triangles	2592	2742
Aire minimale d'un triangle	$0.3858 \cdot 10^{-3} \text{ m}^2$	$0.1593 \cdot 10^{-6} \text{ m}^2$
Aire maximale d'un triangle	$0.3858 \cdot 10^{-3} \text{ m}^2$	$0.3004 \cdot 10^{-2} \text{ m}^2$
Rapport aire max/aire min	1	162

TAB. 4.1 – Caractéristiques des maillages de la Fig. 4.1

- $\tau$  : le seuil numérique de rejet d'un terme non-nul dans les facteurs L et U.

Clairement, les paramètres  $\tau$  et  $p$  vont conditionner le coût mémoire du préconditionneur ILUT mais également son efficacité. Malheureusement, ces deux objectifs sont souvent antagonistes. En règle générale, l'obtention d'un préconditionneur qui soit une approximation précise de  $A^{-1}$  nécessite des petites valeurs du seuil  $\tau$  et des grandes valeurs du seuil  $p$ .

Pour cette étude numérique nous ne considérons que les méthodes DGTD- $\mathbb{P}_1$  et DGTD- $\mathbb{P}_2$  implicites. Nous discutons plus en détail des résultats obtenus avec la DGTD- $\mathbb{P}_2$  implicite.

### 4.3.1 Méthode DGTD- $\mathbb{P}_1$

Les résultats pour les simulations numériques utilisant le maillage uniforme et la stratégie de résolution itérative GMRES/ILUT sont résumés dans la Tab. 4.2 et 4.3. Une interprétation graphique de ces résultats est donnée sur la Fig. 4.2. Les données des Tab. 4.4 et 4.5 permettent de comparer les deux stratégies de résolution. Dans ces tables,  $n_z$  désigne le nombre de termes non nuls dans les facteurs L et U complets ou incomplets. Cet indicateur permet une comparaison partielle des occupations mémoire des deux stratégies de résolution. En effet, on doit aussi tenir compte de l'espace mémoire nécessaire au stockage des vecteurs de la base de Krylov dans la méthode GMRES( $m$ ) qui est égal à  $3 \times d_i \times m$  où  $d_i$  représente le nombre de degrés de liberté caractérisant l'approximation des variables d'état dans chaque cellule. Néanmoins, pour la stratégie de résolution GMRES/ILUT, l'utilisation mémoire est largement dominée par l'espace requis pour le stockage des facteurs L et U incomplets.

On se propose aussi d'évaluer la précision de la factorisation ILUT en fonction des différents paramètres de résolution ainsi que du nombre CFL (i.e. du pas de temps). Pour cela, nous utilisons un mode opératoire en deux étapes :

1. on réalise le produit matrice vecteur  $b = Au$  où  $u$  est un vecteur composé de 1,
2. on résout le système  $MAx = Mb$ .

On calcule alors les erreurs  $L_2$  et  $L_\infty$  entre  $x$  et  $u$ . Les résultats obtenus sont résumés dans la Tab. 4.6.

Les résultats pour les simulations numériques utilisant le maillage non-uniforme et la stratégie de résolution itérative GMRES/ILUT sont résumés dans les Tab. 4.7 à 4.9. Une interprétation graphique de ces résultats est donnée sur les Fig. 4.3 et 4.4. Les données des Tab. 4.10 et 4.11 permettent de comparer les deux stratégies de résolution et la Tab. 4.12 permet d'évaluer la précision de la factorisation incomplète.

Taille de la base de Krylov ( $m$ )	Seuil de résolution ( $\varepsilon$ )	Erreur $L_2$	Temps CPU
15	$10^{-1}$	0.0738512	28 sec
15	$10^{-2}$	0.0738187	34 sec
15	$10^{-3}$	0.0737972	41 sec

TAB. 4.2 – Maillage uniforme, méthode DGTD- $\mathbb{P}_1$  : influence du paramètre  $\varepsilon$

Taille de la base de Krylov ( $m$ )	Nombre maximum de termes non-nuls par ligne	Tolérance $\tau$	Erreur $L_2$	Temps CPU
15	25	0.01	0.07490	35 sec
10	25	0.01	0.07990	35 sec
5	25	0.01	0.07490	35 sec
5	25	0.5	0.08511	198 sec
5	25	0.1	0.07472	42 sec
5	15	0.01	0.07531	44 sec
5	35	0.01	0.07368	38 sec
5	100	0.01	0.07426	36 sec

TAB. 4.3 – Maillage uniforme, méthode DGTD- $\mathbb{P}_1$  : influence de divers paramètres ( $\varepsilon = 10^{-1}$ )

Stratégie de résolution	Seuil de résolution ( $\varepsilon$ )	Erreur $L_2$	Temps CPU	$n_z$
MUMPS	-	0.0738512	16 sec	1399462
GMRES/ILUT	$10^{-1}$	0.0738512	28 sec	767287

TAB. 4.4 – Maillage uniforme, méthode DGTD- $\mathbb{P}_1$  : comparaison des stratégies de résolution

Paramètre	Valeur
Taille de la base Krylov ( $m$ )	15
Nombre maximum d'itérations	50
Seuil de résolution ( $\varepsilon$ )	$10^{-3}$
Nombre maximum de termes non-nuls ( $p$ )	25
Seuil de rejet numérique ( $\tau$ )	$10^{-1}$

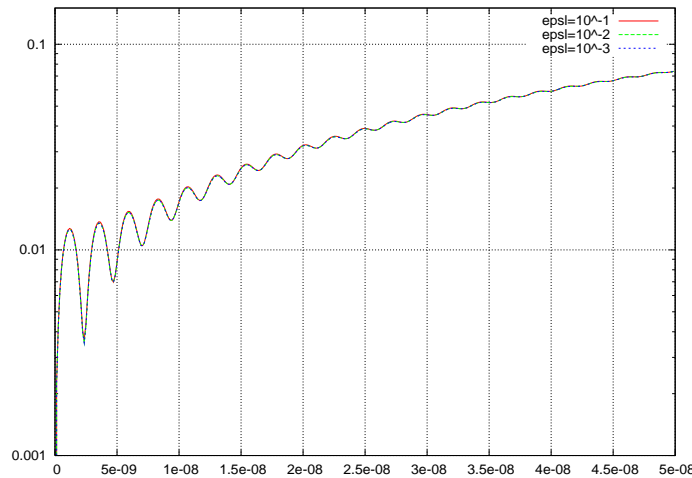
TAB. 4.5 – Maillage uniforme, méthode DGTD- $\mathbb{P}_1$  : paramètres des stratégies de résolution

FIG. 4.2 – Maillage uniforme, méthode DGTD- $\mathbb{P}_1$  : influence du paramètre  $\varepsilon$   
 Evolution temporelle de l'erreur  $L_2$  pour  $m$  fixé  
 Abscisse : temps en sec - Ordonnée : erreur  $L_2$  en échelle log

$p$	$\tau$	CFL	Nombre total de termes non nuls ( $n_z$ )	Taux de remplissage	Erreur $L_2$	Erreur $L_\infty$
25	$10^{-2}$	0.5	468899	0.0861	1.277137	0.028908
25	$10^{-2}$	1.0	637192	0.1170	4.734998	0.133020
50	$10^{-2}$	0.5	468899	0.0861	1.277137	0.028908
50	$10^{-2}$	1.0	659688	0.1211	4.385123	0.075635
75	$10^{-2}$	1.0	659688	0.1211	4.351230	0.075635
50	$5 \cdot 10^{-3}$	0.5	538868	0.0990	0.589043	0.013770
50	$5 \cdot 10^{-3}$	1.0	775584	0.1425	1.185052	0.021233
75	$5 \cdot 10^{-3}$	1.0	775584	0.1425	1.185052	0.021233
50	$10^{-3}$	0.5	685777	0.1260	0.114056	0.002730
50	$10^{-3}$	1.0	981358	0.1803	1.909762	0.035363
75	$10^{-3}$	0.5	1165410	0.2142	0.259764	0.006987
75	$10^{-3}$	1.0	1070193	0.1967	0.430781	0.009887
100	$10^{-3}$	1.0	1204511	0.2213	0.142912	0.004078
100	$10^{-4}$	1.0	1460751	0.2684	0.092920	0.002482
125	$10^{-4}$	1.0	1541404	0.2832	0.031929	0.000990
150	$10^{-4}$	1.0	1542354	0.2834	0.031461	0.000756

TAB. 4.6 – Maillage uniforme, méthode DGTD- $\mathbb{P}_1$  : précision de la factorisation ILUT

Taille de la base de Krylov ( $m$ )	Seuil de résolution ( $\varepsilon$ )	Erreur $L_2$	Temps CPU
5	$10^{-2}$	0.010726	519 sec
10	$10^{-2}$	0.010645	493 sec
15	$10^{-2}$	0.010742	480 sec
25	$10^{-2}$	0.010710	473sec
30	$10^{-2}$	0.010710	473sec
35	$10^{-2}$	0.010710	476 sec

TAB. 4.7 – Maillage non-uniforme, méthode DGTD- $\mathbb{P}_1$  : influence du paramètre  $m$ 

Taille de la base de Krylov ( $m$ )	Nombre maximum de termes non-nuls par ligne	Tolérance $\tau$	Erreur $L_2$	Temps CPU
15	25	$10^{-1}$	$\infty$	-
15	25	$10^{-2}$	0.06696	475 sec
15	25	$10^{-3}$	0.03069	465 sec
15	15	$10^{-2}$	$\infty$	-
15	35	$10^{-2}$	0.04741	384sec

TAB. 4.8 – Maillage non-uniforme, méthode DGTD- $\mathbb{P}_1$  : influence de divers paramètres ( $\varepsilon = 10^{-2}$ )

Taille de la base de Krylov ( $m$ )	Seuil de résolution ( $\varepsilon$ )	Erreur $L_2$	Temps CPU
15	$10^{-1}$	0.044047	396sec
15	$10^{-2}$	0.010742	480 sec
15	$10^{-3}$	0.010653	605 sec

TAB. 4.9 – Maillage non-uniforme, méthode DGTD- $\mathbb{P}_1$  : influence du paramètre  $\varepsilon$

Stratégie de résolution	Seuil de résolution ( $\varepsilon$ )	Erreur $L_2$	Temps CPU	$n_z$
MUMPS	-	0.010710	82 sec	1715396
GMRES/ILUT	$10^{-2}$	0.010710	473 sec	787299

TAB. 4.10 – Maillage non-uniforme, méthode DGTD- $\mathbb{P}_1$  : comparaison des stratégies de résolution

Paramètre	Valeur
Taille de la base Krylov ( $m$ )	25
Nombre maximum d'itérations	50
Seuil de résolution ( $\varepsilon$ )	$10^{-2}$
Nombre maximum de termes non-nuls ( $p$ )	25
Seuil de rejet numérique ( $\tau$ )	$10^{-3}$

TAB. 4.11 – Maillage non-uniforme, méthode DGTD- $\mathbb{P}_1$  : paramètres des stratégies de résolution

$p$	$\tau$	CFL	Nombre total de termes non nuls	Taux de remplissage	Erreur $L_2$	Erreur $L_\infty$
25	$10^{-2}$	24.0	787299	0.1293	71.924934	6.871774
25	$10^{-2}$	12.0	702681	0.1154	49.460821	4.451338
50	$10^{-2}$	24.0	1016477	0.1669	51.987159	4.696014
50	$10^{-2}$	12.0	867153	0.1424	33.699593	3.588880
75	$10^{-2}$	24.0	1093849	0.1796	48.080338	3.859168
100	$10^{-3}$	24.0	1684731	0.2766	36.573017	2.786219
200	$10^{-3}$	24.0	1903435	0.3125	22.183174	1.605774
750	$10^{-5}$	24.0	4008691	0.6582	1.078593	0.165045
1000	$10^{-5}$	24.0	4066727	0.6678	0.516973	0.062617
1000	$10^{-5}$	12.0	3100273	0.5091	0.054968	0.0063987
1200	$10^{-6}$	24.0	5248036	0.8617	0.161205	0.035302
1500	$10^{-6}$	24.0	5306657	0.8714	0.072543	0.015086

TAB. 4.12 – Maillage non-uniforme, méthode DGTD- $\mathbb{P}_1$  : précision de la factorisation ILUT

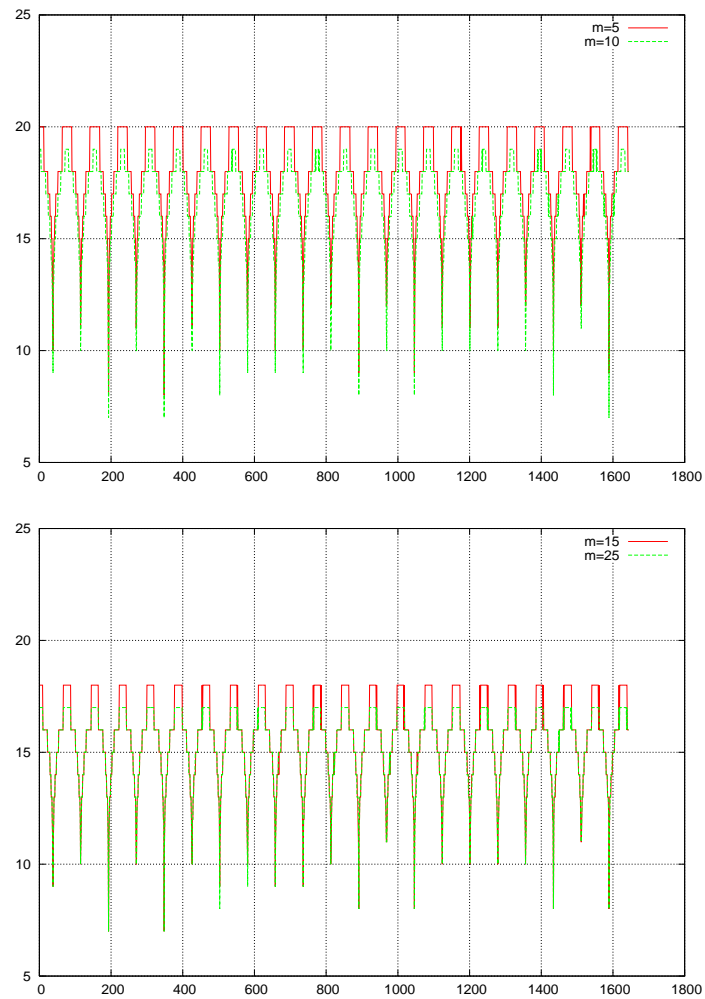


FIG. 4.3 – Maillage non-uniforme, méthode DGTD- $\mathbb{P}_1$  : influence du paramètre  $m$   
Evolution temporelle du nombre effectif d'itérations de GMRES/ILUT pour  $\varepsilon$  fixé

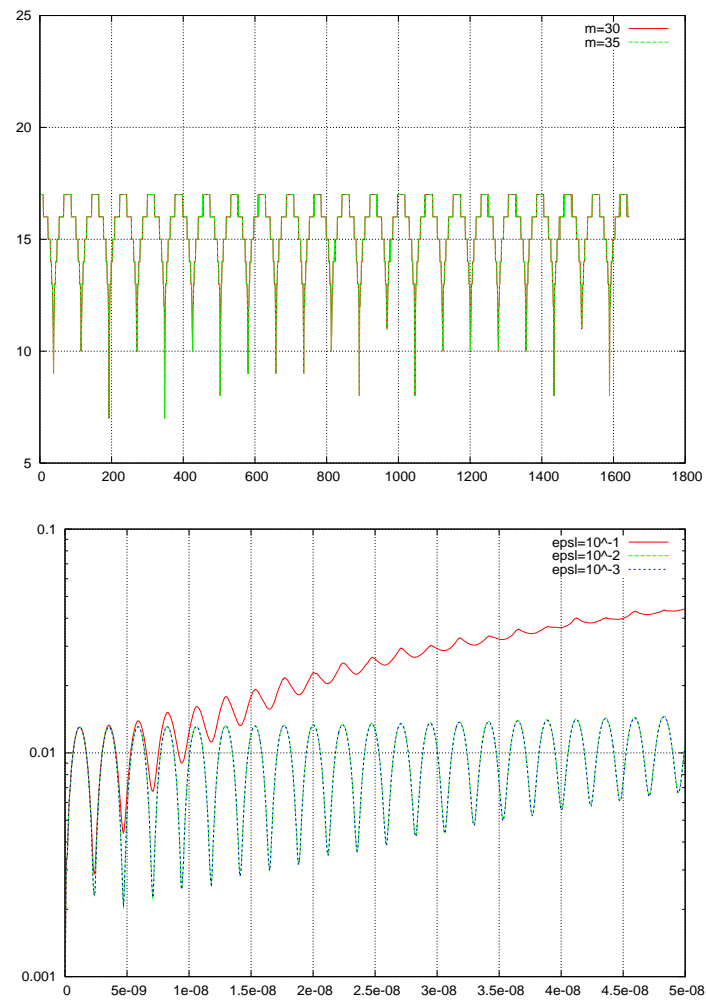


FIG. 4.4 – Maillage non-uniforme, méthode DGTD- $\mathbb{P}_1$  : influence des paramètres  $m$  et  $\epsilon$   
 Evolution temporelle du nombre effectif d'itérations de GMRES/ILUT pour  $\epsilon$  fixé (haut)  
 Evolution temporelle de l'erreur  $L_2$  pour  $m$  fixé (bas)



Dans le cas du maillage uniforme, les résultats de la Tab. 4.2 montrent que diminuer le seuil de résolution linéaire  $\varepsilon$  n'améliore pas sensiblement la précision de la solution mais augmente le temps de calcul. Pour les autres simulations numériques basées sur ce maillage, nous avons utilisé la valeur  $\varepsilon = 10^{-1}$ . Comme on peut le voir dans la Tab. 4.3, il est préférable d'utiliser une base de Krylov de petite taille et un nombre maximum de terme non-nuls par ligne des facteurs L et U incomplets suffisamment grand ( $\geq 25$ ). Même si nous ne l'avons pas indiqué dans cette table, l'espace mémoire requis reste modéré. Nous remarquons en revanche que dans le cas du maillage non-uniforme, il est préférable d'utiliser des valeurs plus élevées de la taille de la base de Krylov (voir la Tab. 4.7) et du nombre de termes non-nuls par ligne des facteurs L et U incomplets (voir la Tab. 4.8). De plus, la valeur  $\varepsilon = 10^{-2}$  du seuil de résolution linéaire semble offrir le meilleur compromis entre précision et temps de calcul (voir la Tab. 4.9). Enfin, le solveur direct MUMPS est nettement plus avantageux que la stratégie de résolution itérative GMRES/ILUT et ce particulièrement dans le cas du maillage non-uniforme (voir la Tab. 4.4 et 4.10).

Concernant la précision de la factorisation incomplète dans le cas du maillage uniforme, on peut remarquer sur la Tab. 4.6 que :

- l'utilisation d'un  $p$  trop petit amène une erreur très importante. En revanche, il est inutile d'augmenter  $p$  dès lors que tous les termes non nuls, autorisés par  $\tau$  sont stockés. On peut voir cet effet pour  $p = 25$  et  $p = 50$  avec  $\tau = 10^{-2}$  et CFL=0.5.
- Dans la plupart des cas, le remplissage des facteurs L et U incomplets est très sensible au CFL (mais cela ne semble pas être une règle générale).
- Choisir  $\tau = 10^{-4}$  n'est pas avantageux car le remplissage des facteurs L et U incomplets excède alors celui des facteurs L et U complets (voir la Tab. 4.4).

Dans le cas du maillage non-uniforme, les données de la Tab. 4.12 montrent que :

- contrairement au cas du maillage uniforme, l'erreur est beaucoup plus importante en maillage non-uniforme et ce même en diminuant la valeur du nombre CFL.
- Il faut utiliser des valeurs de  $\tau$  très faibles ( $10^{-5}$  ou  $10^{-6}$ ) et de  $p$  grandes (entre 1000 et 1500) pour avoir des erreur  $L_2$  de l'ordre de  $10^{-1}$ . Ce sont des valeurs bien supérieures à celles qu'on utilise en pratique.

### 4.3.2 Méthode DGTD- $\mathbb{P}_2$

Nous considérons tout d'abord des simulations numériques utilisant le maillage uniforme. La Tab. 4.13 résume un ensemble de résultats de simulations pour évaluer l'influence des paramètres de la stratégie de résolution GMRES/ILUT. Le seuil de résolution linéaire est fixé à  $\varepsilon = 10^{-3}$  et le nombre de CFL à 2. Ces résultats motivent les commentaires suivants :

- pour tous les réglages, nous obtenons une erreur  $L_2$  quasi-identique mais des temps de calcul assez différents. On a vérifié que toutes les configurations testées permettent d'atteindre le seuil de résolution linéaire fixé. Dès lors, la question restante est celle de la configuration de paramètres qui conduit au temps de simulation minimum.
- Concentrons nous sur la première partie de la Tab. 4.13 qui correspond à  $\tau = 10^{-3}$  :
  - nous remarquons qu'une taille de base de Krylov trop petite nuit au temps de calcul.
  - De même, un nombre maximum d'éléments non nuls par ligne des facteurs L et U incomplets trop faible fait croître également le temps de calcul. On remarque ici, néanmoins, qu'il est possible d'utiliser  $p = 25$  (comme pour la méthode DGTD- $\mathbb{P}_1$ ) et tout de même obtenir un préconditionneur qui garantisse la convergence de la stratégie GMRES/ILUT au seuil de résolution linéaire fixé.

Nous attribuons ce comportement à l'uniformité du maillage et nous verrons plus bas qu'il en est autrement lorsque le maillage est non-uniforme.

- Intéressons nous maintenant à la seconde partie de la Tab. 4.13 qui correspond à  $\tau = 10^{-2}$  :
- on remarque d'une part une diminution notable du temps de calcul qui s'explique bien-sûr par la moindre densité des facteurs L et U incomplets.
- Comme précédemment, on note qu'un nombre trop faible d'éléments non nuls autorisés dans les facteurs L et U incomplets ( $p = 25$ ) induit une augmentation du temps de calcul. Inversement, une valeur plus importante de ce paramètre ne semble plus avoir d'effet sur le temps de calcul passé un certain seuil ( $p = 100$ ).

Taille de la base de Krylov ( $m$ )	Nombre maximum d'itérations	Nombre maximum de termes non-nuls par ligne ( $p$ )	Tolérance ( $\tau$ )	Temps CPU	Erreur $L_2$
5	50	25	$10^{-3}$	240 sec	0.0839200
10	50	25	$10^{-3}$	226 sec	0.0839146
25	25	25	$10^{-3}$	216 sec	0.0839165
15	25	100	$10^{-3}$	207 sec	0.0839168
15	50	100	$10^{-3}$	207 sec	0.0839168
15	50	50	$10^{-3}$	207 sec	0.0839468
15	50	25	$10^{-3}$	216 sec	0.0839165
25	50	25	$10^{-3}$	216 sec	0.0839165
15	25	25	$10^{-2}$	215 sec	0.0839168
15	25	80	$10^{-2}$	161 sec	0.0839165
15	25	100	$10^{-2}$	145sec	0.0839165
15	25	200	$10^{-2}$	147 sec	0.0839165

TAB. 4.13 – Maillage uniforme, méthode DGTD- $\mathbb{P}_2$   
Performances de GMRES/ILUT pour un seuil de résolution linéaire fixé à  $\varepsilon = 10^{-3}$

Pour compléter ces résultats, on peut étudier la sensibilité des performances de la stratégie de résolution GMRES/ILUT au nombre CFL. La Tab. 4.14 rassemble des résultats obtenus pour des nombres CFL fixés à 0.1 et 1.0 et pour un seuil de résolution linéaire fixé à  $\varepsilon = 10^{-3}$ . L'évolution en temps de l'erreur  $L_2$  est montrée sur la Fig. 4.5. On note que les performances de la stratégie de résolution GMRES/ILUT semblent peu sensibles à la valeur du paramètre  $p$ . Néanmoins, on s'est aperçu qu'en faisant varier le paramètre qui contrôle le nombre maximum d'itérations de la stratégie GMRES/ILUT, on pouvait diminuer l'erreur ainsi que le temps de calcul comme illustré par les résultats de la Tab. 4.14.

Pour ce qui concerne les simulations avec le maillage non-uniforme, les données de la Tab. 4.16 montrent que la taille de la base de Krylov n'a pas d'impact sur la précision des calculs. Ce constat doit être tempéré par les valeurs prises par les autres paramètres et en particulier par les paramètres  $p$  et  $\tau$ . En effet, lorsque l'on diminue le nombre maximum d'itérations, la précision n'est pas altérée alors qu'une diminution de la valeur de  $p$  est quant à elle responsable d'une nette augmentation de l'erreur  $L_2$  finale.

Suivant le même mode opératoire que celui adopté pour la méthode DGTD- $\mathbb{P}_1$ , on évalue la précision de la factorisation ILUT en fonction des différents paramètres de résolution ainsi que du nombre CFL (i.e. du pas de temps). Les résultats correspondants sont rassemblés dans les Tab. 4.15 et 4.17 respectivement pour les calculs basés sur les maillages uniforme et non-uniforme. Ces résultats montrent que :

- le fait d'augmenter le degré d'interpolation se traduit par une augmentation du nombre de termes non nuls dans les facteurs L et U incomplets. On note aussi que les factorisations incomplètes obtenues sont moins précises. Comme on a pu le constater dans une moindre mesure pour la méthode DGTD- $\mathbb{P}_1$ , il faut utiliser des valeurs de  $p$  grandes et des valeurs de  $\tau$  petites pour obtenir des factorisations incomplètes acceptables.
- dans le cas d'un maillage non-uniforme, même l'utilisation de valeurs de  $p$  de plusieurs centaines et des valeurs de  $\tau$  de  $10^{-5}$  (plus de sept millions de termes stockés!) ne permettent pas d'obtenir une erreur  $L_2$  inférieure à 2.

On peut donc conclure de cette étude que la précision se dégrade lorsque le degré d'interpolation augmente et lorsque la non-uniformité du maillage est plus prononcée.

CFL	Taille de la base de Krylov	Nombre maximum de termes non-nuls par ligne ( $p$ )	Nombre maximum d'itérations	Temps CPU	Erreur $L_2$
0.1	15	100	25	1114 sec	0.0003602
0.1	15	50	25	1104 sec	0.0003602
1.0	15	50	25	216 sec	0.0211092
1.0	15	100	25	191 sec	0.0210921
1.0	15	100	50	191 sec	0.0210921

TAB. 4.14 – Maillage uniforme, méthode DGTD- $\mathbb{P}_2$   
Influence du nombre CFL sur les performances de GMRES/ILUT pour  $\tau = 10^{-3}$  et  $\varepsilon = 10^{-3}$

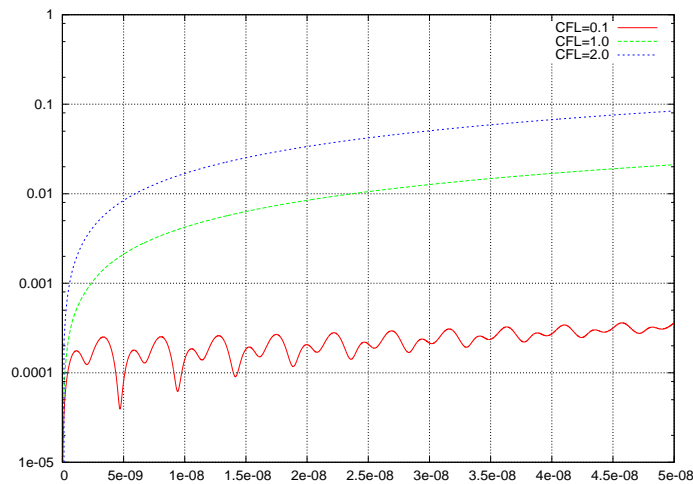


FIG. 4.5 – Maillage uniforme, méthode DGTD- $\mathbb{P}_2$  : influence du nombre de CFL  
Evolution temporelle de l'erreur  $L_2$  pour  $\tau = 10^{-3}$  et  $\varepsilon = 10^{-3}$   
Abscisse : temps en sec - Ordonnée : erreur  $L_2$  en échelle log

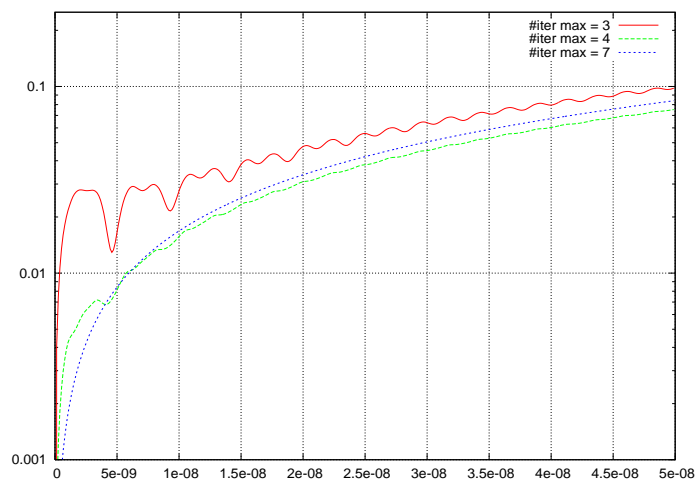


FIG. 4.6 – Maillage uniforme, méthode DGTD- $\mathbb{P}_2$   
 Influence du nombre maximum d'itérations de GMRES/ILUT  
 Evolution temporelle de l'erreur  $L_2$  pour CFL=2,  $\tau = 10^{-3}$  et  $\varepsilon = 10^{-3}$   
 Abscisse : temps en sec - Ordonnée : erreur  $L_2$  en échelle log

$p$	$\tau$	CFL	Nombre total de termes non nuls ( $n_z$ )	Taux de remplissage	Erreur $L_2$	Erreur $L_\infty$
50	$10^{-2}$	1.0	2440663	0.1122	19.488182	0.522237
50	$10^{-2}$	0.5	1967758	0.0904	9.069281	0.175176
100	$10^{-2}$	1.0	2616920	0.1202	16.577810	0.413245
100	$10^{-2}$	0.5	1974973	0.0907	9.005359	0.167772
100	$5 \cdot 10^{-3}$	1.0	3062521	0.1407	4.560378	0.286915
200	$5 \cdot 10^{-3}$	1.0	3148772	0.1447	3.977966	0.152040
300	$10^{-3}$	1.0	4371662	0.2008	2.195897	0.057392
450	$10^{-4}$	1.0	6281401	0.2886	0.122005	0.004260
600	$10^{-5}$	1.0	8222266	0.3777	0.019207	0.000378

TAB. 4.15 – Maillage uniforme, méthode DGTD- $\mathbb{P}_2$  : précision de la factorisation ILUT

Taille de la base de Krylov ( $m$ )	Nombre maximum d'itérations	Nombre maximum de termes non-nuls par ligne ( $p$ )	Tolérance ( $\tau$ )	Erreur $L_2$	Temps CPU
10	100	100	$10^{-2}$	0.01320	2730 sec
15	100	100	$10^{-2}$	0.01320	2755 sec
20	100	100	$10^{-2}$	0.01320	2752 sec
25	100	100	$10^{-2}$	0.01320	2759 sec
10	125	100	$10^{-2}$	0.01320	2745 sec
10	75	100	$10^{-2}$	0.01320	2741 sec
10	50	100	$10^{-2}$	0.01320	2733 sec
10	25	100	$10^{-2}$	0.01320	2730 sec
10	25	100	$10^{-3}$	0.01598	2822 sec
10	25	100	$5 \cdot 10^{-2}$	0.00153	3045 sec
10	100	125	$10^{-2}$	0.01338	2756 sec
10	100	75	$10^{-2}$	0.04280	3037 sec
10	100	50	$10^{-2}$	0.04280	3037 sec

TAB. 4.16 – Maillage non-uniforme, méthode DGTD- $\mathbb{P}_2$  : influence de divers paramètres ( $\varepsilon = 10^{-2}$ )

$p$	$\tau$	CFL	Nombre total de termes non nuls ( $n_z$ )	Taux de remplissage	Erreur $L_2$	Erreur $L_\infty$
100	$10^{-2}$	4	2357237	0.0968	31.537152	2.074252
100	$10^{-2}$	2	1827796	0.0750	14.235360	1.083303
100	$10^{-3}$	4	3503988	0.1438	22.772591	3.200968
100	$10^{-3}$	2	2829841	0.1162	7.829283	1.217322
200	$10^{-4}$	4	5139582	0.2110	8.027295	0.749083
300	$10^{-4}$	4	5449059	0.2237	4.057447	1.201992
400	$10^{-5}$	4	7001564	0.2874	2.020208	0.617489

TAB. 4.17 – Maillage non-uniforme, méthode DGTD- $\mathbb{P}_2$  : précision de la factorisation ILUT

## 4.4 Conclusion

Dans ce chapitre, nous avons étudié numériquement l'influence sur la précision des calculs et les temps de calcul, des différents paramètres de la stratégie itérative GMRES( $m$ )/ILUT( $\tau, p$ ) appliquée à la résolution des systèmes linéaires issus des méthodes DGTD- $\mathbb{P}_k$  implicites étudiées au chapitre 3. Les principaux enseignements de cette étude sont les suivants :

- la taille de la base de Krylov doit être suffisamment grande (de l'ordre de 10) dans la plupart des cas considérés pour obtenir une convergence de la méthode GMRES. Afin d'avoir le meilleur compromis temps de calcul/précision, il est préférable d'utiliser une taille de la base de Krylov de l'ordre de 15.
- Le seuil de résolution linéaire doit être fixé à  $10^{-2}$  ou  $10^{-3}$ . Des valeurs  $\varepsilon < 10^{-2}$  conduisent à une erreur  $L_2$  globale trop importante alors que des valeurs  $\varepsilon > 10^{-3}$  ne permettent pas d'améliorer la précision des simulations au détriment d'un temps de calcul qui augmente inutilement.
- Concernant les paramètres du préconditionneur ILUT, le nombre maximum de termes non nuls dans une ligne des facteurs incomplets L et U est fortement corrélé à la valeur de  $\tau$  comme on a pu le voir dans l'étude de la précision de la factorisation. D'une façon générale, le choix des valeurs des paramètres  $p$  et  $\tau$  dépend du degré d'interpolation dans la méthode DGTD- $\mathbb{P}_k$  et du degré de non-uniformité du maillage.

## Chapitre 5

# Méthode GDDT hybride explicite/implicite

Dans le chapitre 3, nous avons étudié des méthodes DGTD- $\mathbb{P}_k$  globalement implicites afin de pallier à la réduction inévitable du pas de temps pour des raisons de stabilité des méthodes DGTD- $\mathbb{P}_k$  explicites en maillages fortement non-uniformes (localement raffinés). Même si l'utilisation d'un schéma d'intégration en temps implicite permet une diminution notable du temps de calcul pour un niveau d'erreur numérique acceptable, il ne faut cependant pas perdre de vue que la résolution d'un système implicite global reste très coûteuse. En trois dimensions d'espace ou même pour des problèmes bidimensionnels de taille conséquente, un solveur direct adapté aux matrices creuses pour la résolution des systèmes implicites à chaque itération en temps n'est simplement pas viable et il faut envisager une approche différente. La solution classiquement adoptée dans ce contexte consiste à remplacer le solveur direct par une méthode de résolution itérative (méthode de Krylov) convenablement préconditionnée. Cette option a été étudiée numériquement au chapitre 4 où on a considéré une stratégie de résolution itérative combinant la méthode GMRES à une technique de factorisation incomplète ILUT.

Une stratégie alternative jusqu'ici peu considérée pour la résolution numérique des équations de Maxwell en domaine temporel (voir aussi la sous-section 1.2.5.3 du chapitre 1) consiste à restreindre l'application d'un schéma d'intégration en temps implicite à un sous-ensemble des équations semi-discrétisées et à préserver un schéma explicite pour le reste. Les inconnues concernées peuvent par exemple être associées aux éléments du maillage localisés dans les zones les plus raffinées. Autrement dit, on réalise un partitionnement des éléments en deux ensembles,  $\mathcal{S}_e$  et  $\mathcal{S}_i$ , à partir d'un critère approprié, et on formule ensuite une méthode DGTD- $\mathbb{P}_k$  hybride explicite/implicite ou localement implicite. En pratique, on peut penser utiliser un pas de temps global qui assurera la stabilité de la partie explicite  $\mathcal{S}_e$  et qui sera donc plus grand que celui garantissant la stabilité de la méthode DGTD- $\mathbb{P}_k$  explicite appliquée au maillage complet. De plus, si la partie implicite  $\mathcal{S}_i$  est réduite comme cela est souvent le cas lorsque le maillage est localement raffiné, le système implicite à résoudre sera de petite taille et une méthode de résolution directe sera parfaitement applicable.

Dans ce chapitre, nous nous proposons d'étudier une telle stratégie, plus précisément la méthode DGTD- $\mathbb{P}_k$  hybride explicite/implicite initialement proposée par Piperno [Piperno, 2006b]-[Piperno, 2006a]. Cette méthode combine un schéma saute-mouton du second ordre et un schéma de Crank-Nicolson du second ordre pour l'intégration en temps du système d'EDO résultant de la discrétisation des équations de Maxwell par une méthode Galerkin discontinue basée sur un schéma centré pour le calcul des flux numériques aux interfaces entre cellules voisines. La méthode résultante n'a jusqu'ici pas fait l'objet d'une mise en œuvre

pour la résolution numérique du système d'EDP qui nous intéresse et sa stabilité n'a été que partiellement étudiée dans [Piperno, 2006b]. Notre objectif premier dans ce chapitre est de démontrer numériquement les bénéfices de cette approche lorsque les simulations mettent en jeu des maillages triangulaires non-uniformes. Dans la section 5.1, nous détaillons la formulation de ce schéma hybride et dans la section 5.2, nous présentons les résultats d'expériences numériques.

## 5.1 Formulation

Dans [Piperno, 2006b] et [Piperno, 2006a], Piperno propose une méthode DGTD- $\mathbb{P}_k$  hybride explicite/implicite précise au second ordre en temps. La formulation de cette méthode s'appuie sur un partitionnement des éléments du maillage en deux ensembles,  $\mathcal{S}_e$  et  $\mathcal{S}_i$ . Elle combine un schéma saute-mouton pour l'avancement en temps de la partie explicite  $\mathcal{S}_e$  et un schéma de Crank-Nicolson pour l'avancement en temps de la partie implicite  $\mathcal{S}_i$ . Nous rappelons ici les principes de base de cette méthode DGTD- $\mathbb{P}_k$  hybride explicite/implicite.

On suppose que  $\mathcal{F}_a^i = \emptyset$ . On note  $d_t = \sum_i d_i$ . Nous partons de la formulation (2.21) de la section 2.3 du chapitre 2 :

$$\forall \tau_i \in \mathcal{T}_h : \begin{cases} M_{\varepsilon,i} \frac{d\mathbf{E}_i}{dt} = -K_i \mathbf{H}_i - \sum_{a_{ij} \in \mathcal{F}_d^i} \mathcal{X}_{ij} \mathbf{H}_j - \sum_{a_{ij} \in \mathcal{F}_m^i} \mathcal{X}_{im} \mathbf{H}_i, \\ M_{\mu,i} \frac{d\mathbf{H}_i}{dt} = K_i \mathbf{E}_i + \sum_{a_{ij} \in \mathcal{F}_d^i} \mathcal{X}_{ij} \mathbf{E}_j - \sum_{a_{ij} \in \mathcal{F}_m^i} \mathcal{X}_{im} \mathbf{E}_i, \end{cases} \quad (5.1)$$

où  $M_{\varepsilon,i}$ ,  $M_{\mu,i}$  et  $K_i$  sont des matrices  $3d_i \times 3d_i$  symétriques, alors que  $\mathcal{X}_{ij}$  et  $\mathcal{X}_{im}$  sont des matrices  $3d_i \times 3d_i$  antisymétriques. En regroupant toutes les inconnues du champ électrique (respectivement magnétique) au sein d'un même vecteur colonne  $\mathbf{E}_h$  (respectivement  $\mathbf{H}_h$ ) de taille  $d_t$ , on déduit de (5.1) deux systèmes d'EDO globaux :

$$\begin{cases} \mathbb{M}_\varepsilon \frac{d\mathbf{E}_h}{dt} = -\mathbb{K} \mathbf{H}_h - \mathbb{A} \mathbf{H}_h - \mathbb{B} \mathbf{H}_h, \\ \mathbb{M}_\mu \frac{d\mathbf{H}_h}{dt} = \mathbb{K} \mathbf{E}_h + \mathbb{A} \mathbf{E}_h - \mathbb{B} \mathbf{E}_h, \end{cases} \quad (5.2)$$

où :

1.  $\mathbb{M}_\varepsilon$ ,  $\mathbb{M}_\mu$  et  $\mathbb{K}$  sont des matrices diagonales par blocs où les blocs sont respectivement égaux à  $M_{\varepsilon,i}$ ,  $M_{\mu,i}$  et  $K_i$ . On en déduit que les matrices  $\mathbb{M}_\varepsilon$  et  $\mathbb{M}_\mu$  sont symétriques définies positives alors que la matrice  $\mathbb{K}$  est symétrique.
2.  $\mathbb{A}$  est une matrice  $d \times d$  creuse où les blocs non nuls sont égaux à  $\mathcal{X}_{ij}$  quand  $a_{ik}$  est une face interne.  $\mathbb{A}$  est une matrice symétrique.
3.  $\mathbb{B}$  est une matrice  $d \times d$  diagonale par blocs où les blocs non nuls sont égaux à  $\mathcal{X}_{im}$  quand  $a_{ij}$  est une face métallique.  $\mathbb{B}$  est une matrice antisymétrique.



Par suite, si  $\mathbb{S} = \mathbb{K} - \mathbb{A} - \mathbb{B}$ , les systèmes globaux (5.2) s'écrivent aussi comme :

$$\begin{cases} \mathbb{M}_\varepsilon \frac{d\mathbf{E}_h}{dt} = \mathbb{S}\mathbf{H}_h, \\ \mathbb{M}_\mu \frac{d\mathbf{H}_h}{dt} = -{}^t\mathbb{S}\mathbf{E}_h. \end{cases} \quad (5.3)$$

Le schéma saute-mouton du second ordre pour l'intégration en temps de (5.3) s'écrit :

$$\begin{cases} \mathbb{M}_\varepsilon \frac{\mathbf{E}_h^{n+1} - \mathbf{E}_h^n}{\Delta t} = \mathbb{S}\mathbf{H}_h^{n+\frac{1}{2}}, \\ \mathbb{M}_\mu \frac{\mathbf{H}_h^{n+\frac{3}{2}} - \mathbf{H}_h^{n+\frac{1}{2}}}{\Delta t} = -{}^t\mathbb{S}\mathbf{E}_h^{n+1}. \end{cases} \quad (5.4)$$

Piperno [Piperno, 2006b]-[Piperno, 2006a] base son schéma hybride explicite/implicite sur une forme équivalente du schéma saute-mouton connue comme la méthode Verlet :

$$\begin{cases} \mathbb{M}_\mu \frac{\mathbf{H}_h^{n+\frac{1}{2}} - \mathbf{H}_h^n}{\Delta t/2} = -{}^t\mathbb{S}\mathbf{E}_h^n, \\ \mathbb{M}_\varepsilon \frac{\mathbf{E}_h^{n+1} - \mathbf{E}_h^n}{\Delta t} = \mathbb{S}\mathbf{H}_h^{n+\frac{1}{2}}, \\ \mathbb{M}_\mu \frac{\mathbf{H}_h^{n+1} - \mathbf{H}_h^{n+\frac{1}{2}}}{\Delta t/2} = -{}^t\mathbb{S}\mathbf{E}_h^{n+1}. \end{cases} \quad (5.5)$$

L'avantage de ce schéma est de permettre le calcul des composantes des champs  $\mathbf{E}_h$  et  $\mathbf{H}_h$  aux mêmes instants. Supposons maintenant que l'on a réalisé un partitionnement des éléments du maillage en deux ensembles  $\mathcal{S}_e$  et  $\mathcal{S}_i$ . On déduit de ce partitionnement un découpage des champs  $\mathbf{E}_h$  et  $\mathbf{H}_h$  comme :

$$\mathbf{E}_h = \begin{pmatrix} \mathbf{E}_h^e \\ \mathbf{E}_h^i \end{pmatrix} \quad \text{et} \quad \mathbf{H}_h = \begin{pmatrix} \mathbf{H}_h^e \\ \mathbf{H}_h^i \end{pmatrix},$$

et des matrices  $\mathbb{M}_\varepsilon$ ,  $\mathbb{M}_\mu$ ,  $\mathbb{K}$  et  $\mathbb{B}$  :

$$\mathbb{M}_\varepsilon = \begin{pmatrix} \mathbb{M}_\varepsilon^e & \mathbb{O} \\ \mathbb{O} & \mathbb{M}_\varepsilon^i \end{pmatrix}, \quad \mathbb{M}_\mu = \begin{pmatrix} \mathbb{M}_\mu^e & \mathbb{O} \\ \mathbb{O} & \mathbb{M}_\mu^i \end{pmatrix}, \quad \mathbb{K} = \begin{pmatrix} \mathbb{K}^e & \mathbb{O} \\ \mathbb{O} & \mathbb{K}^i \end{pmatrix} \quad \text{et} \quad \mathbb{B} = \begin{pmatrix} \mathbb{B}^e & \mathbb{O} \\ \mathbb{O} & \mathbb{B}^i \end{pmatrix}.$$

La matrice  $\mathbb{A}$  qui correspond aux flux calculés aux interfaces entre éléments voisins se décompose comme :

$$\mathbb{A} = \begin{pmatrix} \mathbb{A}^{ee} & \mathbb{A}^{ei} \\ \mathbb{A}^{ie} & \mathbb{A}^{ii} \end{pmatrix},$$

où  $\mathbb{A}^{ee}$  et  $\mathbb{A}^{ii}$  sont des matrices symétriques et  $\mathbb{A}^{ei} = {}^t\mathbb{A}^{ie}$ . Finalement, en définissant les deux matrices symétriques  $\mathbb{S}^e = -\mathbb{K}^e - \mathbb{A}^{ee} - \mathbb{B}^e$  et  $\mathbb{S}^i = -\mathbb{K}^i - \mathbb{A}^{ii} - \mathbb{B}^i$ , le système d'EDO (5.3) est séparé en deux sous-systèmes :

$$\begin{cases} \mathbb{M}_\varepsilon^e \frac{d\mathbf{E}_h^e}{dt} = \mathbb{S}^e \mathbf{H}_h^e - \mathbb{A}^{ei} \mathbf{H}_h^i, \\ \mathbb{M}_\mu^e \frac{d\mathbf{H}_h^e}{dt} = -{}^t \mathbb{S} \mathbf{E}_h^e + \mathbb{A}^{ei} \mathbf{E}_h^i, \end{cases} \quad (5.6)$$

$$\begin{cases} \mathbb{M}_\varepsilon^i \frac{d\mathbf{E}_h^i}{dt} = \mathbb{S}^i \mathbf{H}_h^i - \mathbb{A}^{ie} \mathbf{H}_h^e, \\ \mathbb{M}_\mu^i \frac{d\mathbf{H}_h^i}{dt} = -{}^t \mathbb{S} \mathbf{E}_h^i + \mathbb{A}^{ie} \mathbf{E}_h^e. \end{cases}$$

L'algorithme hybride explicite-implicite proposé par Piperno [Piperno, 2006b]-[Piperno, 2006a] pour avancer de  $t^n = n\Delta t$  à  $t^{n+1} = (n+1)\Delta t$  est alors le suivant :

1. on avance de  $\Delta t/2$  le domaine explicite avec un schéma d'Euler progressif,
2. on avance de  $\Delta t$  le domaine implicite avec le schéma de Crank-Nicolson,
3. on avance de  $\Delta t/2$  le domaine explicite avec un schéma d'Euler rétrograde.

soit encore, pour le système d'EDO (5.6) :

$$\begin{cases} \mathbb{M}_\mu^e \left( \frac{\mathbf{H}_h^{e,n+\frac{1}{2}} - \mathbf{H}_h^{e,n}}{\Delta t/2} \right) = -{}^t \mathbb{S}^e \mathbf{E}_h^{e,n} + \mathbb{A}^{ei} \mathbf{E}_h^{i,n}, \\ \mathbb{M}_\varepsilon^e \left( \frac{\mathbf{E}_h^{e,n+\frac{1}{2}} - \mathbf{E}_h^{e,n}}{\Delta t/2} \right) = \mathbb{S}^e \mathbf{H}_h^{e,n+\frac{1}{2}} - \mathbb{A}^{ei} \mathbf{H}_h^{i,n}, \\ \mathbb{M}_\varepsilon^i \left( \frac{\mathbf{E}_h^{i,n+1} - \mathbf{E}_h^{i,n}}{\Delta t} \right) = \mathbb{S}^i \left( \frac{\mathbf{H}_h^{i,n+1} + \mathbf{H}_h^{i,n}}{2} \right) - \mathbb{A}^{ie} \mathbf{H}_h^{e,n+\frac{1}{2}}, \\ \mathbb{M}_\mu^i \left( \frac{\mathbf{H}_h^{i,n+1} - \mathbf{H}_h^{i,n}}{\Delta t} \right) = -{}^t \mathbb{S}^i \left( \frac{\mathbf{E}_h^{i,n+1} + \mathbf{E}_h^{i,n}}{2} \right) + \mathbb{A}^{ie} \mathbf{E}_h^{e,n+\frac{1}{2}}, \\ \mathbb{M}_\varepsilon^e \left( \frac{\mathbf{E}_h^{e,n+1} - \mathbf{E}_h^{e,n+\frac{1}{2}}}{\Delta t/2} \right) = \mathbb{S}^e \mathbf{H}_h^{e,n+\frac{1}{2}} - \mathbb{A}^{ei} \mathbf{H}_h^{i,n+1}, \\ \mathbb{M}_\mu^e \left( \frac{\mathbf{H}_h^{e,n+1} - \mathbf{H}_h^{e,n+\frac{1}{2}}}{\Delta t/2} \right) = -{}^t \mathbb{S} \mathbf{E}_h^{e,n+1} + \mathbb{A}^{ei} \mathbf{E}_h^{i,n+1}. \end{cases} \quad (5.7)$$

La démonstration rigoureuse de la stabilité de cet algorithme reste un problème ouvert. Dans [Piperno, 2006b] l'auteur montre qu'une certaine énergie électromagnétique discrète est conservée mais il ne précise pas la condition de stabilité associée.

Par ailleurs, on remarque que lors de la première étape de (5.7), on avance la partie explicite du champ magnétique tout comme on le fait lors de la dernière étape de (5.7). Par suite, on peut regrouper la dernière étape d'une itération avec la première de l'itération suivante. Cela a pour conséquence de réduire le coût de l'algorithme.

## 5.2 Résultats numériques en 2D

Dans cette section, nous présentons des résultats numériques visant à évaluer cette méthode DGTD- $\mathbb{P}_k$  hybride par rapport aux méthodes DGTD- $\mathbb{P}_k$  globalement explicite ou implicite. Les deux cas tests considérés sont la propagation d'un mode propre dans une cavité carrée parfaitement conductrice et la diffraction d'une onde plane par un cylindre diélectrique. Ces cas tests ont aussi été considérés dans le chapitre 3. Différentes configurations d'ensembles  $\mathcal{S}_e$  et  $\mathcal{S}_i$  sont testés afin d'évaluer au mieux les gains apportés par ce schéma.

Dans les expériences numériques, les triangles du maillage sont étiquetés « explicite » ou « implicite » à l'aide d'un paramètre noté  $l_{\max}$ . La procédure de sélection adoptée est triviale : si l'aire de  $\tau_k$  est  $\geq l_{\max}$  alors  $\tau_k \in \mathcal{S}_e$  sinon  $\tau_k \in \mathcal{S}_i$ . Nous mentionnons dans la suite le rapport de mailles explicites  $R_e$  :

$$R_e = \frac{\max_{k \in \mathcal{S}_e} \text{aire}(\tau_k)}{\min_{k \in \mathcal{S}_e} \text{aire}(\tau_k)}.$$

Cette quantité est un bon indicateur permettant de paramétrer au mieux  $l_{\max}$ . En effet, si  $R_e$  est par exemple de l'ordre de 100, il existe encore des cellules cent fois plus petites que d'autres dans le sous-ensemble des cellules explicites  $\mathcal{S}_e$ . Il est donc possible d'utiliser une valeur de  $l_{\max}$  plus grande pour traiter plus de cellules de manière implicite et ainsi augmenter le pas de temps global utilisé en pratique. Nous définissons le nombre CFL implicite par :

$$\text{CFL}_i = \text{CFL}_e \times \frac{\min_{k \in \mathcal{S}_e} \text{aire}(\tau_k)}{\min_{k \in \mathcal{S}_i} \text{aire}(\tau_k)}.$$

où  $\text{CFL}_e$  est le nombre CFL admissible pour le sous-ensemble  $\mathcal{S}_e$ .

Toutes les simulations numériques sont réalisées sur une station de travail basée sur un processeur Intel Xeon 3 GHz avec 2 GB de mémoire RAM. Dans ce qui suit, sauf mention explicite, le terme « Erreur  $L_2$  » fait toujours référence à l'erreur commise entre la solution qui résulte de la méthode DGTD- $\mathbb{P}_k$  implicite et la solution analytique des équations de Maxwell 2D pour le cas test sélectionné. En particulier, dans les tables de résultats, il s'agit de l'erreur à la fin d'une simulation (i.e. au temps final).

### 5.2.1 Mode propre dans une cavité

Le premier cas test que nous traitons est la propagation d'un mode propre dans une cavité dont les parois sont parfaitement conductrices (voir aussi la sous-section 2.4.1.1). Nous utilisons le maillage triangulaire non-uniforme déjà considéré dans la sous-section 3.2.3.1 du chapitre 3. Pour rappel, ce maillage est composé de 1400 sommets et de 2742 triangles. On visualise sur la Fig. 5.1 la distributions des aires des triangles du maillage. A l'aide de cette courbe, on peut déterminer une valeur de  $l_{\max}$  permettant de définir les sous-ensembles  $\mathcal{S}_e$  et  $\mathcal{S}_i$ . On constate que pour ce maillage, la distribution des triangles suivant leur aire est relativement homogène. Les maillages considérés dans la suite cas tests n'auront pas cette répartition.

#### 5.2.1.1 Méthode DGTD- $\mathbb{P}_1$

Avant d'évaluer le gain en temps de calcul résultant de l'utilisation du schéma hybride nous précisons dans la Tab. 5.1 les résultats obtenus avec la méthode DGTD- $\mathbb{P}_1$  basée sur les schémas explicite saute-mouton du second ordre (pour la valeur maximale autorisée du nombre CFL garantissant la stabilité) et

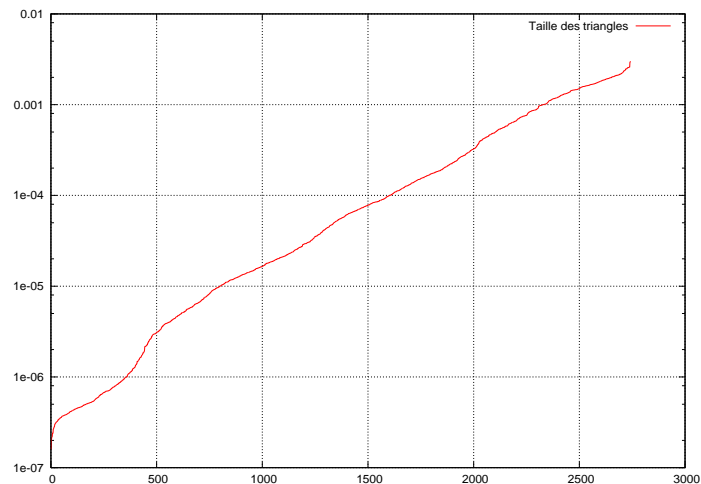


FIG. 5.1 – Mode propre dans une cavité carrée  
Distribution des aires des triangles (abscisse : index du triangle - ordonnée : aire en échelle log)

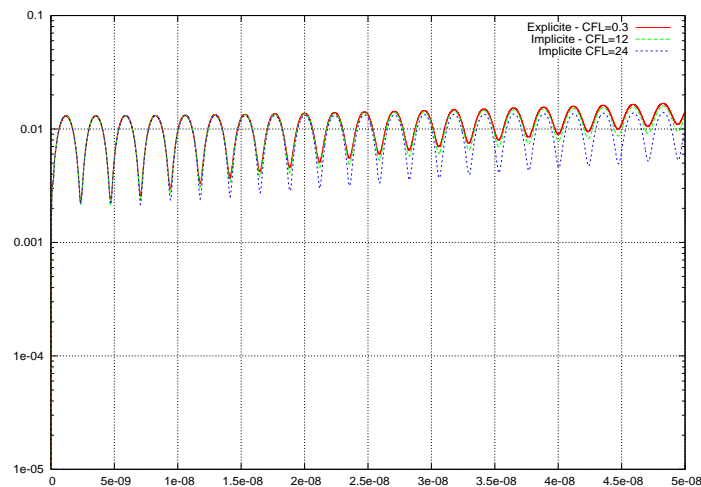


FIG. 5.2 – Mode propre dans une cavité carrée - Méthodes DGTD- $\mathbb{P}_1$  explicite et implicite  
Evolution temporelle de l'erreur  $L_2$  (abscisse : temps en sec - ordonnée : erreur  $L_2$  en échelle log)

Intégration en temps	CFL	Mémoire RAM	Erreur $L_2$ finale	Temps CPU
Explicite	0.3	5 MB	$0.17 \cdot 10^{-1}$	572 sec
Implicite	12	65 MB	$0.16 \cdot 10^{-1}$	143 sec
-	24	65 MB	$0.14 \cdot 10^{-1}$	71 sec

TAB. 5.1 – Mode propre dans une cavité carrée  
Méthode DGTD- $\mathbb{P}_1$  : schémas en temps globalement explicite ou implicite

implicite de Crank-Nicolson du second ordre (pour deux valeurs du nombre CFL). L'évolution temporelle de l'erreur  $L_2$  pour les différents schémas en temps est montrée sur la Fig. 5.2.

Nous allons maintenant nous intéresser au schéma hybride et voir s'il permet de concilier le meilleur des deux méthodes précédentes à savoir un temps de calcul relativement faible et une utilisation mémoire modérée. Nous résumons dans les Tab. 5.2 et 5.3 les performances de la méthode DGTD- $\mathbb{P}_1$  hybride explicite/implicite en fonction du paramètre  $l_{\max}$  (voir aussi la Fig. 5.3 pour une illustration de l'évolution temporelle de l'erreur  $L_2$  pour certaines des configurations de ces tables). Différentes valeurs de ce paramètre sont considérées et pour certaines d'entre elles le nombre CFL implicite (i.e.  $CFL_i$ ) est contraint à la valeur utilisée pour la méthode DGTD- $\mathbb{P}_1$  globalement implicite. Pour ce faire, il suffit de diminuer le nombre CFL appliqué aux triangles du sous-ensemble  $\mathcal{S}_e$  (i.e.  $CFL_e$ ). Cela permet ainsi d'évaluer l'impact du pas de temps sur les performances en temps de calcul et précision de la méthode DGTD- $\mathbb{P}_1$  hybride sans modifier le coût mémoire car les nombres d'éléments traités dans les sous-ensembles  $\mathcal{S}_e$  et  $\mathcal{S}_i$  ne changent pas. Dans toutes les tables qui suivent, la valeur du nombre CFL implicite indiquée entre parenthèses est celle qui correspond au paramètre  $l_{\max}$  et la valeur indiquée à la gauche de cette dernière est le nombre CFL implicite utilisé en pratique (i.e. le nombre CFL implicite contraint).

$l_{\max}$	Nombre de triangles implicites	Nombre de triangles explicites	Nombre d'arêtes mixtes	$R_e$
$10^{-5}$	799	1943	97	22.16
$3 \cdot 10^{-5}$	1211	1531	99	13.00
$5 \cdot 10^{-5}$	1339	1403	89	9.69
$10^{-4}$	1601	1141	117	6.64
$3 \cdot 10^{-4}$	1977	765	81	3.86
$3 \cdot 10^{-4}$	1977	765	81	3.86
$5 \cdot 10^{-4}$	2102	640	54	2.58
$6 \cdot 10^{-4}$	2162	580	48	2.26
$7 \cdot 10^{-4}$	2213	529	49	2.11
$8 \cdot 10^{-4}$	2255	487	49	1.96
$8 \cdot 10^{-4}$	2255	487	49	1.96
$10^{-3}$	2327	415	51	1.79
$10^{-3}$	2327	415	51	1.79
$10^{-3}$	2327	415	51	1.79

TAB. 5.2 – Mode propre dans une cavité carrée  
Méthode DGTD- $\mathbb{P}_1$  : schéma en temps hybride explicite/implicite

Nous déduisons des Tab. 5.2 et 5.3 plusieurs constatations :

- pour  $l_{\max}$  relativement petit (de l'ordre de  $10^{-5}$ ), les méthodes DGTD- $\mathbb{P}_1$  explicite et hybride explicite/implicite conduisent au même niveau d'erreur  $L_2$  finale. Par contre, l'occupation mémoire augmente de 5 MB à 22 MB et le temps de calcul diminue de 572 sec à 267 sec. Comparativement à la méthode DGTD- $\mathbb{P}_1$  implicite, même si le temps de calcul demeure supérieur, la quantité de mémoire utilisée par la méthode DGTD- $\mathbb{P}_1$  hybride explicite/implicite est bien inférieure.
- Si l'on peut se permettre d'utiliser plus de mémoire (choix d'un  $l_{\max}$  plus grand), on peut réduire de manière significative le temps de calcul par rapport à la méthode DGTD- $\mathbb{P}_1$  explicite. Par exemple, pour  $l_{\max}$  de l'ordre de  $10^{-4}$ , nous conservons la précision de la méthode DGTD- $\mathbb{P}_1$  explicite tout en réduisant le temps de calcul de 572 sec à 138 sec.

$l_{\max}$	$CFL_i$	Mémoire RAM	Erreur $L_2$ finale	Temps CPU
$10^{-5}$	2.20	22 MB	$0.17 \cdot 10^{-1}$	267 sec
$3 \cdot 10^{-5}$	3.79	30 MB	$0.17 \cdot 10^{-1}$	215 sec
$5 \cdot 10^{-5}$	5.63	33 MB	$0.17 \cdot 10^{-1}$	175 sec
$10^{-4}$	7.35	39 MB	$0.17 \cdot 10^{-1}$	138 sec
$3 \cdot 10^{-4}$	12.63	42 MB	$0.19 \cdot 10^{-1}$	99 sec
$3 \cdot 10^{-4}$	12 (12.63)	42 MB	$0.20 \cdot 10^{-1}$	103 sec
$5 \cdot 10^{-4}$	12 (18.93)	50 MB	$0.20 \cdot 10^{-1}$	110 sec
$6 \cdot 10^{-4}$	21 (21.5)	52 MB	$0.31 \cdot 10^{-1}$	64 sec
$7 \cdot 10^{-4}$	23.02	53 MB	$0.35 \cdot 10^{-1}$	60 sec
$8 \cdot 10^{-4}$	24.85	54 MB	$0.41 \cdot 10^{-1}$	56 sec
$8 \cdot 10^{-4}$	24 (24.85)	54 MB	$0.40 \cdot 10^{-1}$	58 sec
$10^{-3}$	27.2	55 MB	$0.51 \cdot 10^{-1}$	53 sec
$10^{-3}$	24 (27.2)	55 MB	$0.42 \cdot 10^{-1}$	60 sec
$10^{-3}$	12 (27.2)	55 MB	$0.22 \cdot 10^{-1}$	123 sec

TAB. 5.3 – Mode propre dans une cavité carrée  
Méthode DGTD- $\mathbb{P}_1$  : schéma en temps hybride explicite/implicite

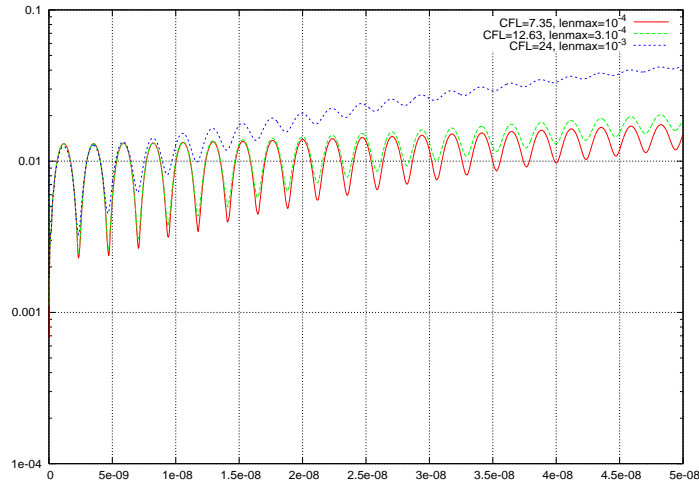


FIG. 5.3 – Mode propre dans une cavité carrée - Méthode DGTD- $\mathbb{P}_1$  hybride explicite/implicite  
Evolution temporelle de l'erreur  $L_2$  (abscisse : temps en sec - ordonnée : erreur  $L_2$  en échelle log)

- A partir d'un certain nombre CFL implicite (aux alentours de 20), l'erreur  $L_2$  augmente. On remarque que ce n'est pas lié au pourcentage de triangles traités avec le schéma implicite mais au pas de temps de référence utilisé pour les simulations. Ainsi si on contraint le nombre CFL implicite, on observe bien une réduction de l'erreur  $L_2$  finale.

### 5.2.1.2 Méthode DGTD- $\mathbb{P}_2$

Comme pour la méthode DGTD- $\mathbb{P}_1$ , nous commençons par rappeler dans la Tab. 5.4 les résultats obtenus avec la méthode DGTD- $\mathbb{P}_2$  basée sur les schémas explicite saute-mouton du second ordre (pour la valeur maximale autorisée du nombre CFL garantissant la stabilité) et implicite de Crank-Nicolson du second ordre (pour deux valeurs du nombre CFL). L'évolution temporelle de l'erreur  $L_2$  pour les différents schémas en temps est montrée sur la Fig. 5.4. Dans la suite, nous considérons la solution numérique obtenue pour CFL=4 comme résultat de référence pour la méthode DGTD- $\mathbb{P}_2$  globalement implicite.

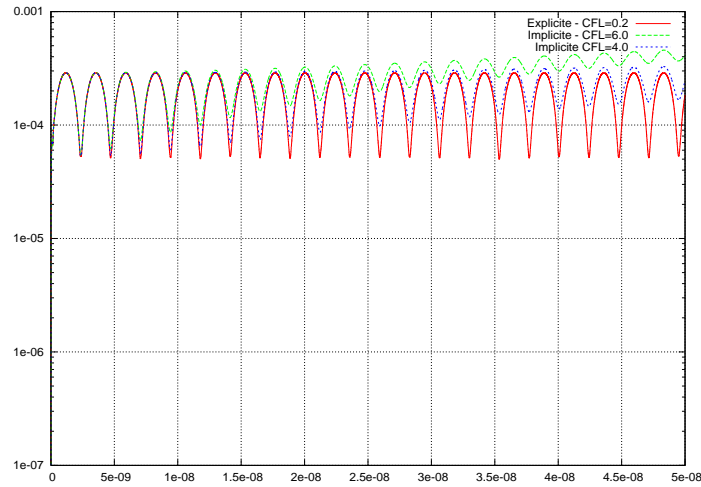


FIG. 5.4 – Mode propre dans une cavité carrée - Méthodes DGTD- $\mathbb{P}_2$  explicite et implicite  
Evolution temporelle de l'erreur  $L_2$  (abscisse : temps en sec - ordonnée : erreur  $L_2$  en échelle log)

Intégration en temps	CFL	Mémoire RAM	Erreur $L_2$ finale	Temps CPU
Explicite	0.2	5 MB	$0.28 \cdot 10^{-3}$	1801 sec
Implicite	4	102 MB	$0.32 \cdot 10^{-3}$	979 sec
-	6	102 MB	$0.88 \cdot 10^{-3}$	655 sec

TAB. 5.4 – Mode propre dans une cavité carrée  
Méthode DGTD- $\mathbb{P}_2$  : schémas en temps globalement explicite ou implicite

Nous résumons dans les Tab. 5.5 et 5.6 les performances de la méthode DGTD- $\mathbb{P}_2$  hybride explicite/implicite en fonction du paramètre  $l_{\max}$  (voir aussi la Fig. 5.5 pour une illustration de l'évolution temporelle de l'erreur  $L_2$  pour certaines des configurations de ces tables).

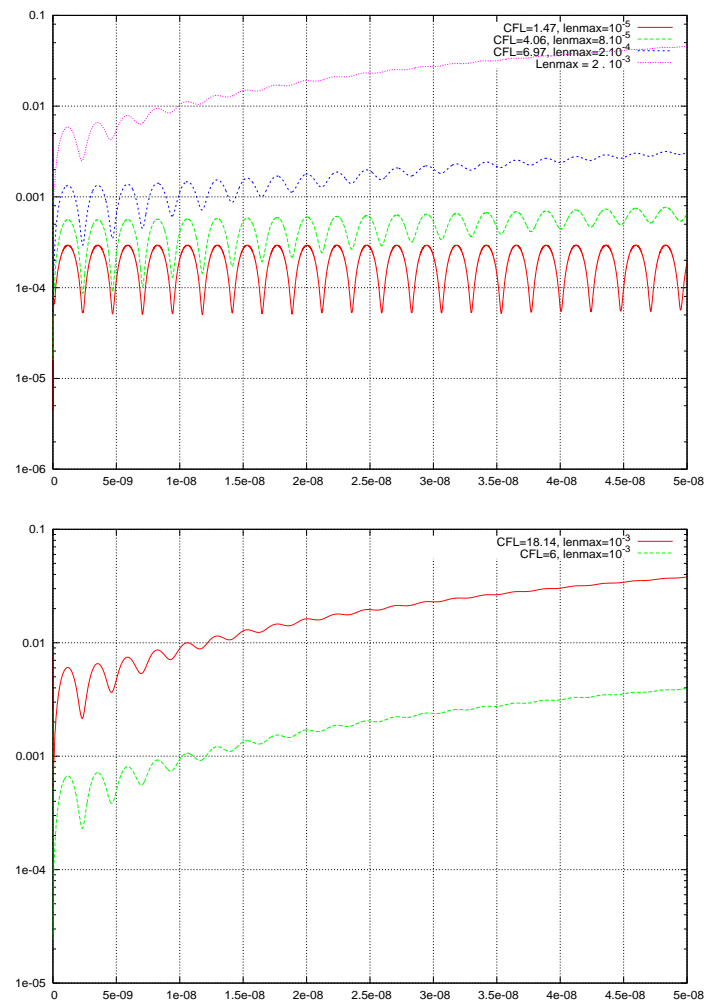


FIG. 5.5 – Mode propre dans une cavité carrée - Méthode DGTD- $\mathbb{P}_2$  hybride explicite/implicite  
Evolution temporelle de l'erreur  $L_2$  (abscisse : temps en sec - ordonnée : erreur  $L_2$  en échelle log)



$l_{\max}$	Nombre de triangles implicites	Nombre de triangles explicites	Nombre d'arêtes mixtes	$R_e$
$10^{-5}$	799	1943	97	22.16
$5 \cdot 10^{-5}$	1339	1403	89	9.69
$6 \cdot 10^{-5}$	1393	1349	111	8.45
$7 \cdot 10^{-5}$	1454	1288	128	8.28
$8 \cdot 10^{-5}$	1509	1233	127	8.01
$9 \cdot 10^{-5}$	1566	1176	120	6.99
$9 \cdot 10^{-5}$	1566	1176	120	6.99
$10^{-4}$	1601	1141	117	6.64
$1.5 \cdot 10^{-4}$	1741	1001	111	5.28
$2 \cdot 10^{-4}$	1859	883	111	4.66
$2 \cdot 10^{-4}$	1859	883	111	4.66
$3 \cdot 10^{-4}$	1977	765	81	3.86
$3 \cdot 10^{-4}$	1977	765	81	3.86
$5 \cdot 10^{-4}$	2102	640	54	2.57
$5 \cdot 10^{-4}$	2102	640	54	2.57
$10^{-3}$	2327	415	51	1.79
$10^{-3}$	2327	415	51	1.79
$2 \cdot 10^{-3}$	2653	89	75	1.37

TAB. 5.5 – Mode propre dans une cavité carrée  
Méthode DGTD- $\mathbb{P}_2$  : schéma en temps hybride explicite/implicite

$l_{\max}$	$CFL_i$	Mémoire RAM	Erreur $L_2$ finale	Temps CPU
$10^{-5}$	1.47	31 MB	$0.71 \cdot 10^{-3}$	916 sec
$5 \cdot 10^{-5}$	3.35	50 MB	$0.16 \cdot 10^{-2}$	592 sec
$6 \cdot 10^{-5}$	3.84	52 MB	$0.18 \cdot 10^{-2}$	535 sec
$7 \cdot 10^{-5}$	3.93	54 MB	$0.19 \cdot 10^{-2}$	547 sec
$8 \cdot 10^{-5}$	4.06	57 MB	$0.20 \cdot 10^{-2}$	551 sec
$9 \cdot 10^{-5}$	4.64	58 MB	$0.22 \cdot 10^{-2}$	511 sec
$9 \cdot 10^{-5}$	4 (4.64)	58 MB	$0.19 \cdot 10^{-2}$	570 sec
$10^{-4}$	4.89	59 MB	$0.23 \cdot 10^{-2}$	479 sec
$1.5 \cdot 10^{-4}$	6 (6.16)	66 MB	$0.28 \cdot 10^{-2}$	425 sec
$2 \cdot 10^{-4}$	6.97	68 MB	$0.33 \cdot 10^{-2}$	383 sec
$2 \cdot 10^{-4}$	6 (6.97)	68 MB	$0.28 \cdot 10^{-2}$	445 sec
$3 \cdot 10^{-4}$	8.41	73 MB	$0.46 \cdot 10^{-2}$	337 sec
$3 \cdot 10^{-4}$	6 (8.41)	73 MB	$0.28 \cdot 10^{-2}$	474 sec
$5 \cdot 10^{-4}$	12.62	78 MB	$0.12 \cdot 10^{-1}$	238 sec
$5 \cdot 10^{-4}$	6 (12.62)	78 MB	$0.28 \cdot 10^{-2}$	501 sec
$10^{-3}$	18.14	87 MB	$0.38 \cdot 10^{-1}$	186 sec
$10^{-3}$	6 (18.14)	87 MB	$0.39 \cdot 10^{-2}$	564 sec
$2 \cdot 10^{-3}$	23.78	98 MB	$0.45 \cdot 10^{-1}$	160 sec

TAB. 5.6 – Mode propre dans une cavité carrée  
Méthode DGTD- $\mathbb{P}_2$  : schéma en temps hybride explicite/implicite

Nous déduisons des Tab. 5.5 et 5.6 les commentaires suivants :

- lorsque l'on utilise un ordre d'interpolation plus élevé, ici l'ordre 2, on note que l'on n'atteint pas le même niveau de précision que celui caractérisant les méthodes DGTD- $\mathbb{P}_2$  explicite et implicite (voir la Tab. 5.4). Plus précisément, pour ces méthodes, l'erreur  $L_2$  finale est de l'ordre de  $0.3 \cdot 10^{-3}$  alors que pour la méthode DGTD- $\mathbb{P}_2$  hybride explicite/implicite l'erreur  $L_2$  la plus faible est de l'ordre  $0.7 \cdot 10^{-3}$ .
- Le nombre CFL qui assure la stabilité de la partie explicite diminue lorsque l'ordre d'interpolation augmente. Ainsi, le temps CPU de la méthode DGTD- $\mathbb{P}_k$  hybride explicite/implicite augmente avec le degré d'interpolation  $p$ .
- Pour  $l_{\max}$  de l'ordre de  $8 \cdot 10^{-5}$ , nous utilisons moitié moins de mémoire que la méthode DGTD- $\mathbb{P}_2$  implicite et le temps de calcul est réduit de 70% par rapport à la méthode DGTD- $\mathbb{P}_2$  explicite. De même, pour un  $l_{\max}$  de l'ordre de  $1.5 \cdot 10^{-4}$ , on réduit le temps de calcul à 425 sec tout en n'utilisant que 66 MB de mémoire.
- Lorsque l'on contraint la valeur du nombre CFL implicite pour  $l_{\max}$  donné, l'erreur  $L_2$  finale est équivalente à celle obtenue pour une valeur de  $l_{\max}$  induisant comme nombre CFL implicite la valeur plafond fixée. Par exemple, pour  $l_{\max}$  de l'ordre de  $1.5 \cdot 10^{-4}$ , l'erreur  $L_2$  est de  $0.28 \cdot 10^{-2}$  pour un nombre CFL implicite de 6. Pour  $l_{\max}$  de l'ordre de  $5 \cdot 10^{-4}$ , le nombre de CFL implicite est 12.63 et l'erreur  $L_2$  finale est de  $0.12 \cdot 10^{-1}$  tandis que si l'on diminue le nombre CFL implicite à 6, l'erreur  $L_2$  finale vaut  $0.28 \cdot 10^{-2}$ . Ce résultat reste toutefois à tempérer pour des valeurs plus grandes de  $l_{\max}$  puisque l'on observe que pour  $l_{\max} = 10^{-3}$ , si la diminution du nombre CFL implicite permet de réduire l'erreur  $L_2$  finale, celle-ci n'atteint pas la valeur obtenue pour  $l_{\max} = 3 \cdot 10^{-4}$  (et CFL=12.63).

En résumé, cette série de résultats montre que la DGTD- $\mathbb{P}_k$  hybride explicite/implicite permet de réduire le temps de calcul et l'occupation mémoire pour le cas test et pour le maillage considérés ici. Nous remarquons par ailleurs que le nombre d'arêtes mixtes ne joue en rien sur la précision du schéma. Cela peut paraître étonnant puisque c'est sur celles-ci que ce font les échanges de flux implicites/explicites.

## 5.2.2 Diffraction d'une onde plane par un cylindre diélectrique

Nous nous proposons ici d'évaluer les performances de la méthode DGTD- $\mathbb{P}_k$  hybride explicite/implicite pour le cas test de la diffraction d'une onde plane par un cylindre diélectrique. Nous considérons les configurations de calcul étudiées dans la sous-section 3.2.3.3 du chapitre 3.

### 5.2.2.1 Configuration C1

Pour cette configuration de calcul, le maillage triangulaire non-uniforme utilisé comporte 4108 sommets et 8054 triangles. On visualise sur la Fig. 5.6 la distributions des aires des triangles du maillage. A partir de cette courbe, on peut facilement déterminer une valeur de  $l_{\max}$  permettant de séparer les triangles en deux sous-ensembles  $\mathcal{S}_e$  et  $\mathcal{S}_i$ . La courbe présente un plateau qui débute à partir d'une aire de triangle légèrement supérieure à  $10^{-3}$ . En choisissant une valeur de  $l_{\max}$  proche de cette valeur, nous pouvons séparer les triangles en deux groupes, similaires en nombre. Nous verrons par la suite que ce choix est pertinent.

**Méthode DGTD- $\mathbb{P}_1$**  Les résultats obtenus avec la méthode DGTD- $\mathbb{P}_1$  basée sur les schémas explicite saute-mouton du second ordre et implicite de Crank-Nicolson du second ordre sont données dans la Tab. 5.7. L'évolution temporelle de l'erreur  $L_2$  pour les différents schémas en temps est montrée sur la Fig. 5.7.

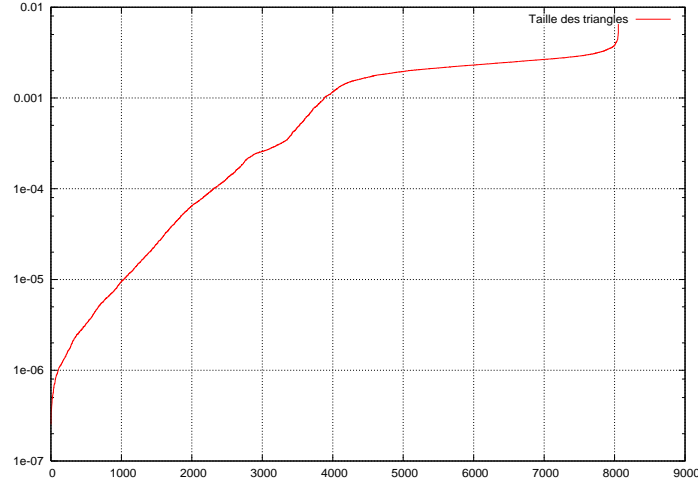


FIG. 5.6 – Diffraction d'une onde plane par un cylindre diélectrique (C1)  
Distribution des aires des triangles (abscisse : index du triangle - ordonnée : aire en échelle log)

Intégration en temps	CFL	Mémoire RAM	Erreur $L_2$ finale	Temps CPU
Explicite	0.3	9 MB	0.40	327 sec
Implicite	20	187 MB	0.43	55 sec

TAB. 5.7 – Diffraction d'une onde plane par un cylindre diélectrique (C1)  
Méthode DGTD- $\mathbb{P}_1$  : schémas en temps globalement explicite ou implicite

Nous résumons dans les Tab. 5.8 et Tab. 5.9 les performances de la méthode DGTD- $\mathbb{P}_1$  hybride explicite/implicite en fonction du paramètre  $l_{\max}$  (voir aussi la Fig. 5.8 pour une illustration de l'évolution temporelle de l'erreur  $L_2$  pour certaines des configurations de ces tables). Nous pouvons faire plusieurs remarques sur les résultats des Tab. 5.8 et 5.9 :

- pour  $l_{\max}$  proche de  $10^{-3}$ , les sous-ensembles  $\mathcal{S}_e$  et  $\mathcal{S}_i$  sont quasiment de même taille. Ce choix donne de bon résultats car il permet de diviser par deux l'occupation mémoire et obtenir un temps de calcul inférieur à la méthode DGTD- $\mathbb{P}_1$  globalement implicite.
- Choisir  $l_{\max}$  légèrement plus grand ( $2.5 \cdot 10^{-3}$ ) conduit à traiter une majeure partie des triangles avec le schéma d'intégration en temps implicite. Dans ce cas, on ne tire pas profit de l'hybridation des deux schémas d'intégration en temps.
- Autrement, choisir une valeur plus petite de  $l_{\max}$  ne permet pas de diminuer suffisamment le temps de calcul mais se traduit par une diminution notable de l'occupation mémoire.

Nous montrons sur la Fig. 5.9 l'évolution temporelle de la composante  $E_z$  du champ électrique en un point du maillage pour la méthode DGTD- $\mathbb{P}_1$  explicite avec un nombre CFL=0.3, la méthode DGTD- $\mathbb{P}_1$  implicite avec un nombre CFL=20 et de la méthode DGTD- $\mathbb{P}_1$  hybride explicite/implicite lorsque  $l_{\max} = 10^{-3}$  (et CFL=14.64). On peut remarquer la parfaite superposition des trois courbes.

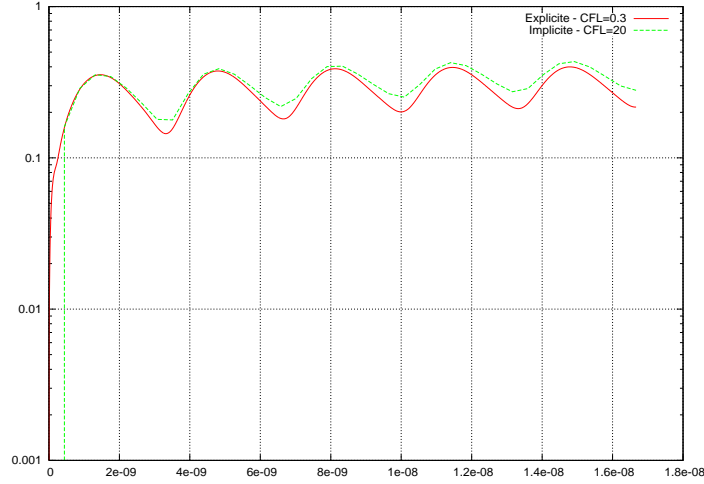


FIG. 5.7 – Diffraction d’une onde plane par un cylindre diélectrique (C1)  
Méthode DGTD- $\mathbb{P}_1$  - Evolution temporelle de l’erreur  $L_2$   
Abscisse : temps en sec - Ordonnée : erreur  $L_2$  en échelle log

$l_{\max}$	Nombre de triangles implicites	Nombre de triangles explicites	Nombre d’arêtes mixtes	$R_e$
$10^{-5}$	1032	7022	86	26.51
$10^{-4}$	2309	5745	121	8.50
$10^{-3}$	3880	4174	158	3.03
$2.5 \cdot 10^{-3}$	6572	1482	716	1.93
$2.5 \cdot 10^{-3}$	6572	1482	716	1.93
$3 \cdot 10^{-3}$	7628	426	434	1.59

TAB. 5.8 – Diffraction d’une onde plane par un cylindre diélectrique (C1)  
Méthode DGTD- $\mathbb{P}_1$  : schéma en temps hybride explicite/implicite

$l_{\max}$	$CFL_i$	Mémoire RAM	Erreur $L_2$ finale	Temps CPU
$10^{-5}$	1.62	30 MB	0.40	146 sec
$10^{-4}$	5.22	59 MB	0.41	78 sec
$10^{-3}$	14.64	95 MB	0.44	42 sec
$2.5 \cdot 10^{-3}$	23	150 MB	0.49	39 sec
$2.5 \cdot 10^{-3}$	21 (23)	150 MB	0.47	43 sec
$3 \cdot 10^{-3}$	28	174 MB	0.45	37 sec

TAB. 5.9 – Diffraction d’une onde plane par un cylindre diélectrique (C1)  
Méthode DGTD- $\mathbb{P}_1$  : schéma en temps hybride explicite/implicite

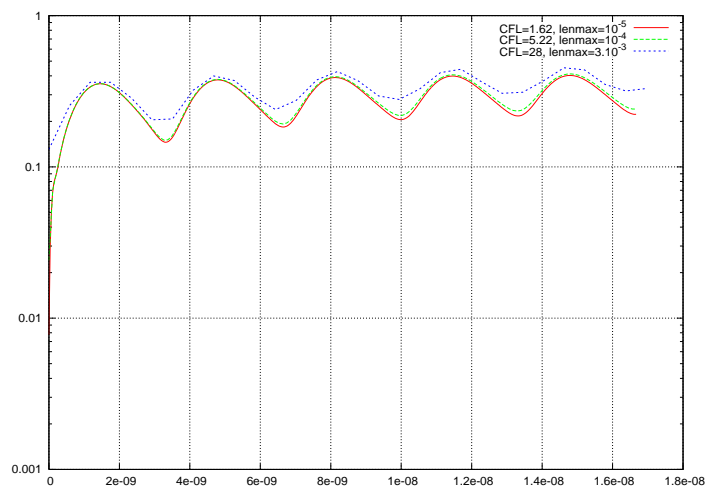


FIG. 5.8 – Diffraction d'une onde plane par un cylindre diélectrique (C1)  
Méthode DGTD- $\mathbb{P}_1$  hybride explicite/implicite  
Evolution temporelle de l'erreur  $L_2$  (abscisse : temps en sec - ordonnée : erreur  $L_2$  en échelle log)

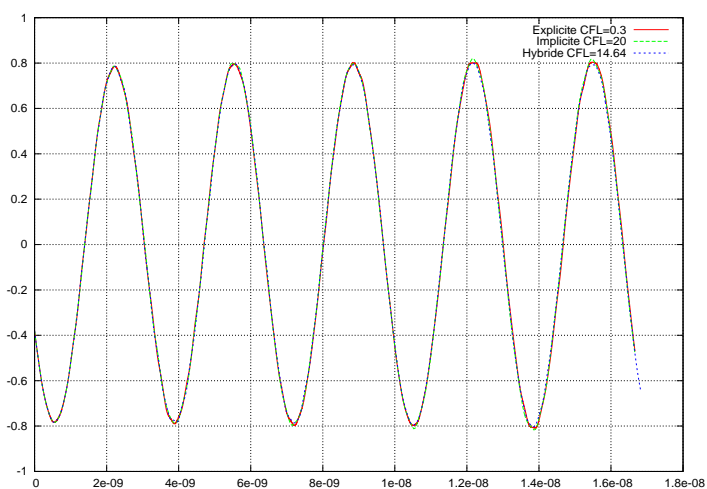


FIG. 5.9 – Diffraction d'une onde plane par un cylindre diélectrique (C1)  
Evolution temporelle de la composante  $E_z$  (abscisse : temps en sec - ordonnée :  $E_z$ )

**Méthode DGTD- $\mathbb{P}_2$**  Les résultats obtenus avec la méthode DGTD- $\mathbb{P}_2$  basée sur les schémas explicite saute-mouton du second ordre et implicite de Crank-Nicolson du second ordre sont données dans la Tab. 5.10. On note que pour le maillage considéré, la méthode DGTD- $\mathbb{P}_2$  globalement implicite devient relativement gourmande en occupation mémoire.

Intégration en temps	CFL	Mémoire RAM	Erreur $L_2$ finale	Temps CPU
Explicite	0.2	9	0.16	761 sec
Implicite	20	303	0.19	136 sec

TAB. 5.10 – Diffraction d’une onde plane par un cylindre diélectrique (C1)  
Méthode DGTD- $\mathbb{P}_2$  : schémas en temps globalement explicite ou implicite

Nous résumons dans les Tab. 5.11 et 5.12 les performances de la méthode DGTD- $\mathbb{P}_2$  hybride explicite/implicite en fonction du paramètre  $l_{\max}$ . Nous remarquons que le choix de la valeur  $10^{-3}$  pour  $l_{\max}$  permet de concilier au mieux le temps de calcul (identique à celui de la méthode implicite) avec l’occupation mémoire pour une erreur  $L_2$  finale du même ordre. L’évolution temporelle de l’erreur  $L_2$  pour les différents schémas en temps est montrée sur la Fig. 5.10.

$l_{\max}$	Nombre de triangles implicites	Nombre de triangles explicites	Nombre d’arêtes mixtes	$R_e$
$10^{-4}$	2309	5745	121	8.50
$10^{-3}$	3880	4174	158	3.03
$3 \cdot 10^{-3}$	7628	426	434	1.59

TAB. 5.11 – Diffraction d’une onde plane par un cylindre diélectrique (C1)  
Méthode DGTD- $\mathbb{P}_2$  : schéma en temps hybride explicite/implicite

$l_{\max}$	CFL <sub><i>i</i></sub>	Mémoire RAM	Erreur $L_2$ finale	Temps CPU
$10^{-4}$	3.48	89 MB	0.16	255 sec
$10^{-3}$	9.76	150 MB	0.18	139 sec
$3 \cdot 10^{-3}$	18.66	274 MB	0.21	135 sec

TAB. 5.12 – Diffraction d’une onde plane par un cylindre diélectrique (C1)  
Méthode DGTD- $\mathbb{P}_2$  : schéma en temps hybride explicite/implicite

### 5.2.2.2 Configuration C2

Nous concluons cette série d’expériences numériques en présentant les résultats obtenus pour la seconde configuration de calcul considérée pour ce cas test dans la sous-section 3.2.3.3 du chapitre 3. Nous utilisons un maillage non-uniforme composé de 4228 sommets et de 8326 triangles. On visualise sur la Fig. 5.11 la distributions des aires des triangles du maillage. Cette courbe présente deux plateaux. Le premier débute à partir d’une aire de triangle légèrement supérieure à  $10^{-6}$  tandis que le deuxième commence pour une aire de triangle proche de  $10^{-3}$ . Nous allons donc choisir ces deux valeurs pour  $l_{\max}$ . La première permettra de traiter environ 10% des éléments de manière implicite alors qu’avec la seconde, ce pourcentage grimpe à 60%.

Nous résumons dans les Tab. 5.13 à 5.15 les performances des méthodes DGTD- $\mathbb{P}_1$  et DGTD- $\mathbb{P}_2$  globalement implicites et hybrides explicites/implicites. L’évolution temporelle de l’erreur  $L_2$  pour les différentes

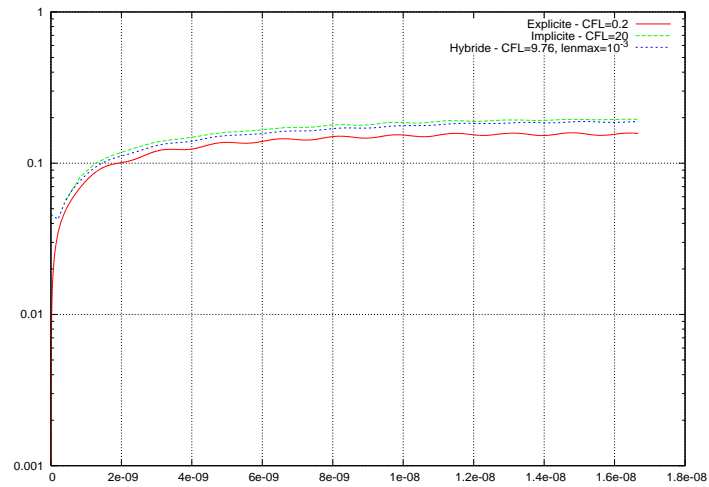


FIG. 5.10 – Diffraction d'une onde plane par un cylindre diélectrique (C1)  
Méthode DGTD- $\mathbb{P}_2$  - Evolution temporelle de l'erreur  $L_2$   
Abscisse : temps en sec - Ordonnée : erreur  $L_2$  en échelle log

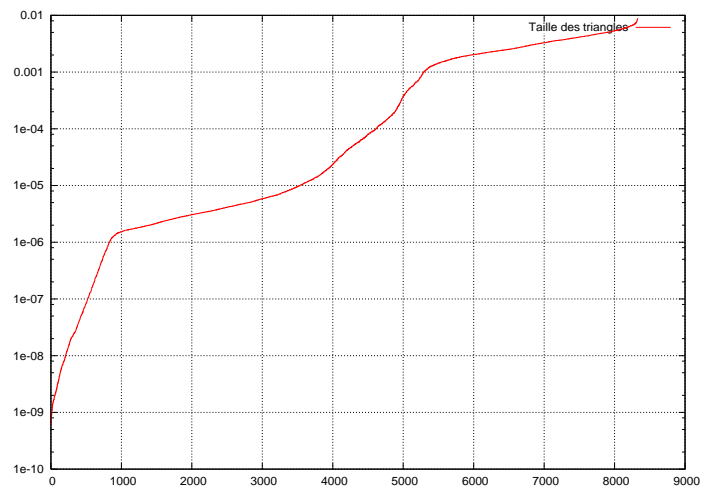


FIG. 5.11 – Diffraction d'une onde plane par un cylindre diélectrique (C2)  
Distribution des aires des triangles (abscisse : index du triangle - ordonnée : aire en échelle log)

méthodes est montrée sur la Fig. 5.12. Il est clair que le choix  $l_{\max} = 10^{-3}$  permet de réduire les temps de calcul et l'occupation mémoire pour les deux degrés d'interpolation considérés ici, alors que pour  $l_{\max} = 1.5 \cdot 10^{-6}$  on note seulement une réduction notable de la quantité de mémoire utilisée. Pour  $l_{\max} = 10^{-3}$ , les temps de calcul sont même nettement inférieurs à ceux des méthodes DGTD- $\mathbb{P}_1$  et DGTD- $\mathbb{P}_2$  globalement implicites. En effet, pour cette configuration de calcul, le même pas de temps de référence ( $\Delta t = 17.5$  picosec) est utilisé pour les méthodes globalement implicite et hybride explicite/implicite, mais les coûts par itération temporelle des méthodes DGTD- $\mathbb{P}_1$  et DGTD- $\mathbb{P}_2$  globalement implicites n'autorisent pas une réduction significative des temps de calcul.

Intégration en temps	$l_{\max}$	Nombre de triangles implicites	Nombre de triangles explicites	Nombre d'arêtes mixtes	$R_e$
Implicite	-	8326	0	0	-
Hybride	$1.5 \cdot 10^{-6}$	981	7345	209	131
-	$10^{-3}$	5292	3034	110	4.56

TAB. 5.13 – Diffraction d'une onde plane par un cylindre diélectrique  
Méthodes DGTD- $\mathbb{P}_1$  et DGTD- $\mathbb{P}_2$  : schémas en temps implicite et hybride explicite/implicite (C2)

Intégration en temps	$l_{\max}$	$CFL_i$	Mémoire RAM	Erreur $L_2$ finale	Temps CPU
Implicite	-	200	379 MB	0.47	150 sec
Hybride	$1.5 \cdot 10^{-6}$	10.59	55 MB	0.45	798 sec
-	$10^{-3}$	200 (304)	245 MB	0.47	98 sec

TAB. 5.14 – Diffraction d'une onde plane par un cylindre diélectrique (C2)  
Méthode DGTD- $\mathbb{P}_1$  : schémas en temps implicite et hybride explicite/implicite  
Temps de référence méthode DGTD- $\mathbb{P}_1$  explicite ( $CFL=0.3$ ) : 13438 sec

Intégration en temps	$l_{\max}$	$CFL_i$	Mémoire RAM	Erreur $L_2$ finale	Temps CPU
Implicite	-	200	507 MB	0.47	402 sec
Hybride	$1.5 \cdot 10^{-6}$	7.06	68 MB	0.45	2108 sec
-	$10^{-3}$	200	327 MB	0.47	256 sec

TAB. 5.15 – Diffraction d'une onde plane par un cylindre diélectrique (C2)  
Méthode DGTD- $\mathbb{P}_2$  : schémas en temps implicite et hybride explicite/implicite  
Temps de référence méthode DGTD- $\mathbb{P}_2$  explicite ( $CFL=0.2$ ) : 21456 sec

### 5.3 Conclusion

Dans ce chapitre, nous avons étudié numériquement une méthode DGTD- $\mathbb{P}_k$  hybride explicite/implicite initialement proposée par Piperno [Piperno, 2006b]-[Piperno, 2006a]. Les équations différentielles ordinaires résultant de la semi-discrétisation du système de Maxwell par une formulation Galerkin discontinue sont partitionnées en deux sous-ensembles. En pratique, ces sous-ensembles correspondent à une séparation des éléments du maillage sur la base d'un critère géométrique ici basé sur l'aire d'un triangle. Un schéma d'intégration en temps explicite saute-mouton est appliqué aux EDO associées aux éléments dont l'aire est supérieure à un seuil fixé, le reste des EDO étant traité par un schéma implicite de Crank-Nicolson. La



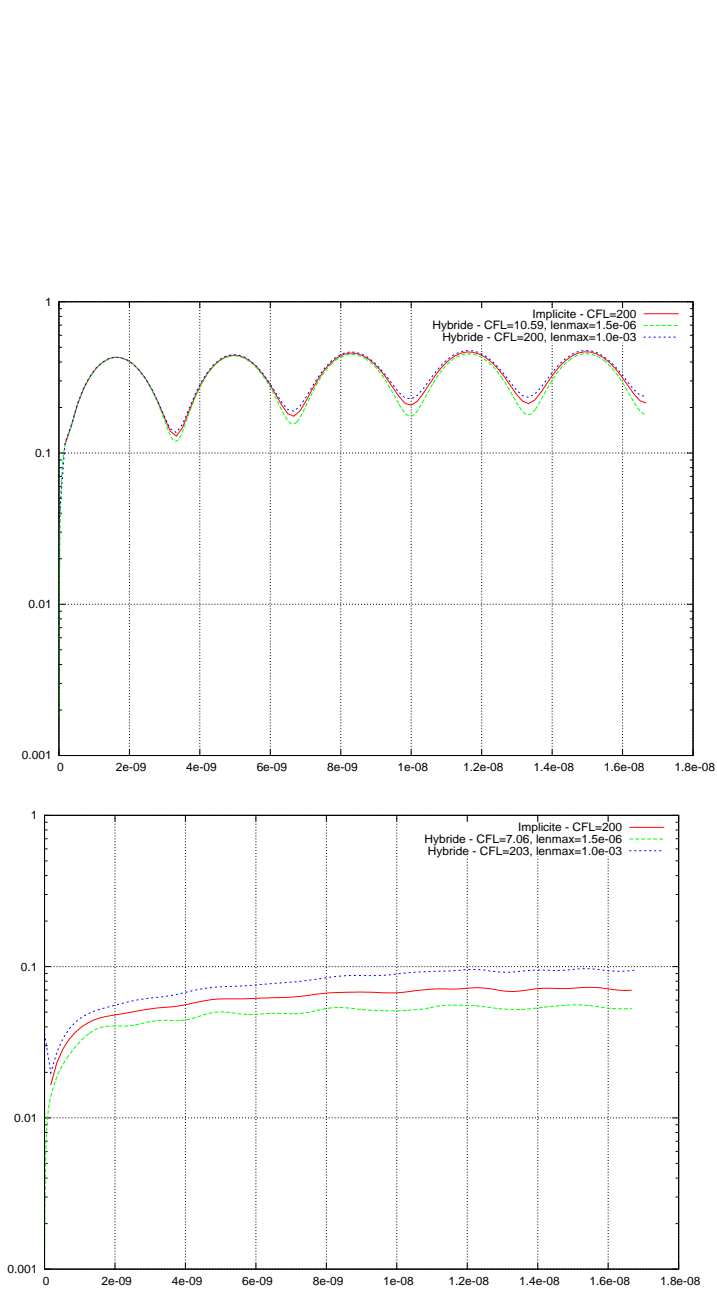


FIG. 5.12 – Diffraction d’une onde plane par un cylindre diélectrique (C2)  
Méthodes DGTD- $\mathbb{P}_1$  (haut) et DGTD- $\mathbb{P}_2$  (bas) - Evolution temporelle de l’erreur  $L_2$

méthode  $DGTD-\mathbb{P}_k$  résultante est d'ordre arbitrairement élevé en espace et du second ordre en temps. La stabilité et la convergence de cette méthode sont des questions ouvertes qui n'ont volontairement pas été abordées dans notre étude. On peut néanmoins penser que cette méthode est stable sous une condition de type CFL qui contraint le pas de temps à la valeur maximale admissible garantissant la stabilité des EDO traitées par le schéma explicite. Cette hypothèse est à la base de la mise en œuvre de la méthode  $DGTD-\mathbb{P}_k$  hybride qui a été réalisée dans cette étude et en pratique la méthode s'est toujours avérée stable.

Plusieurs expériences numériques ont permis d'étudier en détail les performances en termes de précision, temps de calcul et occupation mémoire de la méthode  $DGTD-\mathbb{P}_k$  hybride explicite/implicite pour  $p = 1, 2$  et pour des simulations mettant en jeu des maillages non-uniformes. Dans tous les cas, cette méthode offre un bon compromis entre temps de calcul et occupation mémoire comparativement aux méthodes  $DGTD-\mathbb{P}_k$  globalement explicite ou implicite. Elle devrait être particulièrement avantageuse pour les simulations tridimensionnelles où la méthode  $DGTD-\mathbb{P}_k$  globalement implicite basée sur un solveur direct pour la résolution des systèmes linéaires à chaque itération temporelle est bien trop gourmande en ressource mémoire pour être une méthodologie numérique viable.

# Chapitre 6

## Résultats numériques en 3D

### 6.1 Préambule

Dans ce chapitre, nous présentons quelques résultats d'expériences numériques portant sur des problèmes de propagation d'ondes en trois dimensions d'espace. L'objectif est avant tout d'évaluer les performances, en termes de temps de calcul et de consommation mémoire, de la méthode GDDT- $\mathbb{P}_p$  hybride explicite/implicite étudiée et validée numériquement dans un cadre bidimensionnel au chapitre 5. En effet, une extension directe au cas 3D de la méthode GDDT- $\mathbb{P}_p$  globalement implicite basée sur le schéma d'intégration en temps de Crank-Nicolson du second ordre d'une part, et sur la résolution du système implicite à chaque itération temporelle par une méthode directe (factorisation LU adaptée aux matrices creuses) d'autre part, n'est pas une stratégie viable. L'élément limitatif est bien évidemment le coût de la factorisation LU en 3D.

Pour illustrer ce dernier point, on a considéré le cas test de propagation d'un mode propre dans une cavité cubique unitaire dont les parois sont parfaitement conductrices (voir aussi les sous-sections 2.4.1.1 du chapitre 2 et 3.2.3.1 du chapitre 3). On utilise un maillage tétraédrique non-uniforme comprenant 3815 sommets et 19540 tétraèdres. Les longueurs minimale et maximale des arêtes de ce maillage sont respectivement égales à  $1.44 \cdot 10^{-3}$  m et  $1.6061 \cdot 10^{-1}$  m. Les valeurs minimale et maximale du pas de temps (hors CFL) sont respectivement égales à  $\Delta t_{\min} = 0.84$  picosec et  $\Delta t_{\max} = 107.1$  picosec soit un rapport  $\Delta t_{\max}/\Delta t_{\min} \approx 127$ . Sur la Fig. 6.1 on montre l'évolution temporelle de l'erreur  $L_2$  et de la composante  $E_z$  en un point du domaine (zoom sur la dernière période, la durée totale de la simulation étant de 80 nanosec) associées aux solutions numériques fournies par les méthodes GDDT- $\mathbb{P}_1$  globalement explicite et implicite. Dans ce dernier cas, on a considéré deux valeurs du nombre CFL. Les résultats de performance sont résumés dans la Tab. 6.1. On notera que l'on a mis en œuvre ici un code de calcul parallélisé (le solveur direct creux MUMPS est adapté aux architectures de calcul parallèles à mémoire distribuée [Amestoy et al., 2000]). Il est clair que le gain en temps de calcul lié à l'utilisation du schéma d'intégration en temps implicite est pour une bonne part absorbé par le surcoût de calcul de la factorisation. De plus, la consommation mémoire est excessive même si le maillage sous-jacent est relativement grossier. Les problèmes 3D dimensionnants mettent en jeu des maillages comportant plusieurs centaines de millier voire quelques millions de sommets. Il est donc peu probable qu'une stratégie globalement implicite de ce type passera à l'échelle.

A partir de maintenant, nous nous concentrons donc sur la stratégie de résolution hybride explicite/implicite. Celle-ci a fait l'objet d'une mise en œuvre en mode séquentiel. Comme en 2D, le solveur direct MUMPS est appliqué au système linéaire formé à partir des tétraèdres traités en implicite.

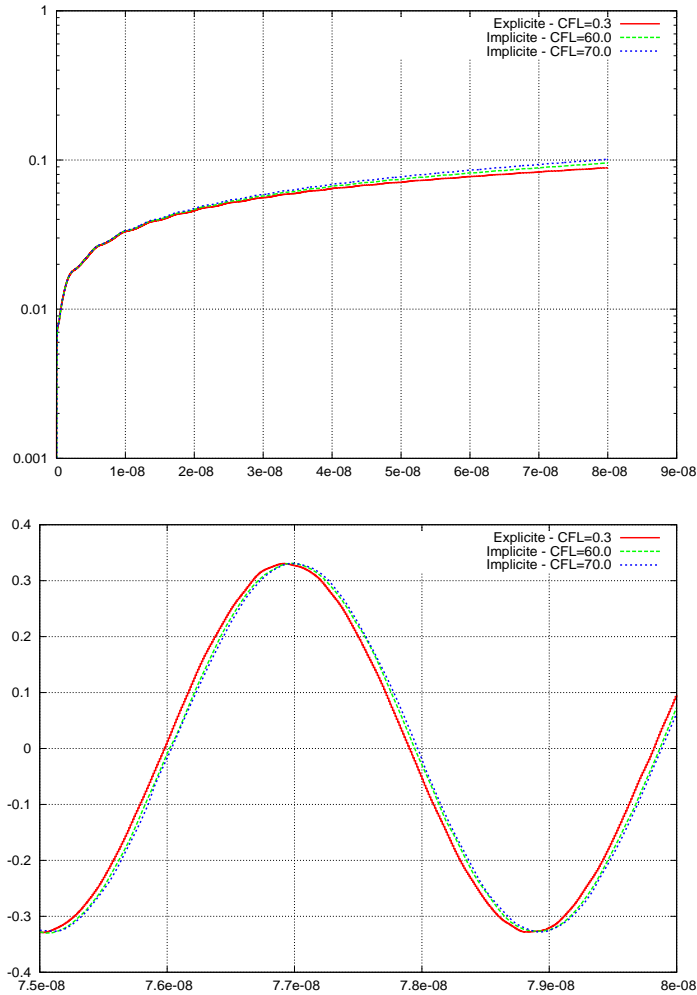


FIG. 6.1 – Mode propre dans un cube unité - Méthode GDDT- $\mathbb{P}_1$   
 Evolution temporelle de l'erreur  $L_2$  (haut) et de la composante  $E_z$  en un point du domaine (bas)  
 Abscisse : temps en sec - Ordonnée : erreur  $L_2$  (haut) énergie discrète (bas)

Intégration en temps	CFL	Mémoire RAM (LU)	Temps CPU (LU)	Temps CPU (total)
Explicite	0.3	-	-	2670 sec
Implicite	60.0	1600 MB	1484 sec	1312 sec
-	70.0	1600 MB	1052 sec	1049 sec

TAB. 6.1 – Mode propre dans un cube unité - Méthode GDDT- $\mathbb{P}_1$   
 Temps CPU (cluster d'AMD Opteron 2.2 GHz et interconnexion Myrinet)

Par ailleurs, le critère géométrique utilisé pour la construction des sous-ensembles  $\mathcal{S}_e$  et  $\mathcal{S}_i$  est défini par :

$$\mathcal{C}(\tau_i) = 4 \min_{j \in \mathcal{V}_i} \frac{V_i V_j}{P_i P_j} \quad (6.1)$$

où  $V_i$  et  $P_i$  désignent respectivement le volume du tétraèdre  $\tau_i$  et la surface i.e. la somme des aires des faces du tétraèdre  $\tau_i$ . On rappelle que  $\mathcal{V}_i = \{j | \tau_i \cap \tau_j \neq \emptyset\}$  est l'ensemble des éléments voisins de  $\tau_i$ . Cette expression intervient dans la formule de calcul du pas de temps global  $\Delta t$  garantissant la stabilité de la méthode GDDT- $\mathbb{P}_0$ . Pour les méthode GDDT- $\mathbb{P}_p$ ,  $p \geq 1$ , le pas de temps utilisé est  $\text{CFL}_p \times \Delta t_r$  où  $\text{CFL}_p < 1$ .

Nous considérons ci-dessous trois cas tests et dans chaque cas, on discute principalement des aspects relatifs aux performances des stratégies de résolution globalement explicite et hybride explicite/implicite.

## 6.2 Résultats numériques

### 6.2.1 Diffraction d'une onde plane par un jet d'affaire

Le problème considéré ici est la diffraction d'une onde plane de fréquence  $F=200$  MHz (i.e. de longueur d'onde  $\lambda = 1.5$  m) par une géométrie de jet d'affaire. Le domaine de calcul est délimité par un parallélépipède sur lequel on applique une condition absorbante de type Silver-Müller. Le maillage tétraédrique utilisé comporte 360495 sommets et 2024924 tétraèdres (voir la Fig. 6.2 pour des vues du maillage de la surface de l'avion). Les longueurs minimale et maximale des arêtes de ce maillage sont respectivement égales à  $9.166 \cdot 10^{-3}$  m (ce qui représente  $\approx \lambda/163$  m) et  $6.83161 \cdot 10^{-1}$  m (soit  $\approx \lambda/2.2$  m). Les valeurs minimale et maximale du pas de temps (hors CFL) sont respectivement égales à  $\Delta t_{\min} = 7.7$  picosec et  $\Delta t_{\max} = 444.3$  picosec soit un rapport  $\Delta t_{\max}/\Delta t_{\min} \approx 58$ . La distribution du critère  $\mathcal{C}(\tau_i)$  pour les éléments  $\tau_i$  de ce maillage est visualisée sur la Fig. 6.3 (la figure inférieure est un zoom sur les plus petites valeurs de ce critère). Le temps de simulation est fixé à 5 périodes de l'onde incidente.

Pour la construction des sous-ensembles  $\mathcal{S}_e$  et  $\mathcal{S}_i$ , on a sélectionné trois valeurs  $l_{\max}$  du critère  $\mathcal{C}(\tau_i)$  pour lesquelles on obtient les configurations résumées dans la Tab. 6.2. On montre sur la Fig. 6.4 les lignes de contour de la composante  $E_z$  et du module de  $\mathbf{E}$  sur la surface de l'avion, associées à la solution numérique fournie par la méthode GDDT- $\mathbb{P}_1$  hybride explicite/implicite. La Fig. 6.5 compare les solutions numériques fournies par les méthodes GDDT- $\mathbb{P}_1$  globalement explicite et hybride explicite/implicite sur la base de l'évolution temporelle de la composante  $E_z$  en un point donné du maillage. Les résultats de performance sont rassemblés dans la Tab. 6.3. On note que le temps de construction de la factorisation LU de la matrice associée à la partie implicite  $\mathcal{S}_i$  est très largement négligeable devant le temps de simulation total. Pour  $l_{\max} = 0.02$ , le gain en temps de calcul total entre les deux méthodes est de l'ordre de 6.

$l_{\max}$	$ \mathcal{S}_e $	$ \mathcal{S}_i $	$\text{CFL}_i$
0.0125	2024320	604	1.81
0.0175	2022464	2460	2.53
0.02	2018543	6381	2.90

TAB. 6.2 – Diffraction d'une onde plane par un jet d'affaire - F=200 MHz  
Caractéristiques des sous-ensembles  $\mathcal{S}_e$  et  $\mathcal{S}_i$

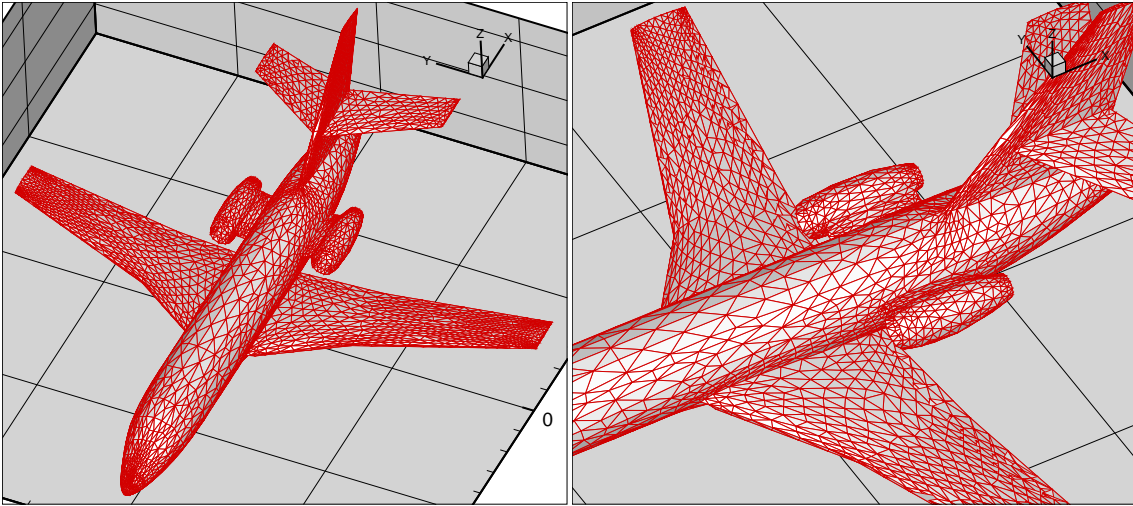


FIG. 6.2 – Diffraction d'une onde plane par un jet d'affaire -  $F=200$  MHz  
Maillage triangulaire de surface

$l_{\max}$	Mémoire RAM (LU)	Temps CPU (LU)	Temps CPU (total)
0.0125 m	12 MB	0.3 sec	6 h 39 mn
0.0175 m	48 MB	1.5 sec	4 h 44 mn
0.02 m	117 MB	4.2 sec	4 h 08 mn

TAB. 6.3 – Diffraction d'une onde plane par un jet d'affaire -  $F=200$  MHz  
Temps CPU (station de travail basée sur un processeur Intel Xeon 2.33 GHz)  
Méthode GDDT hybride explicite/implicite  
Temps CPU de référence (méthode GDDT- $\mathbb{P}_1$  explicite) : 25 h 3 mn

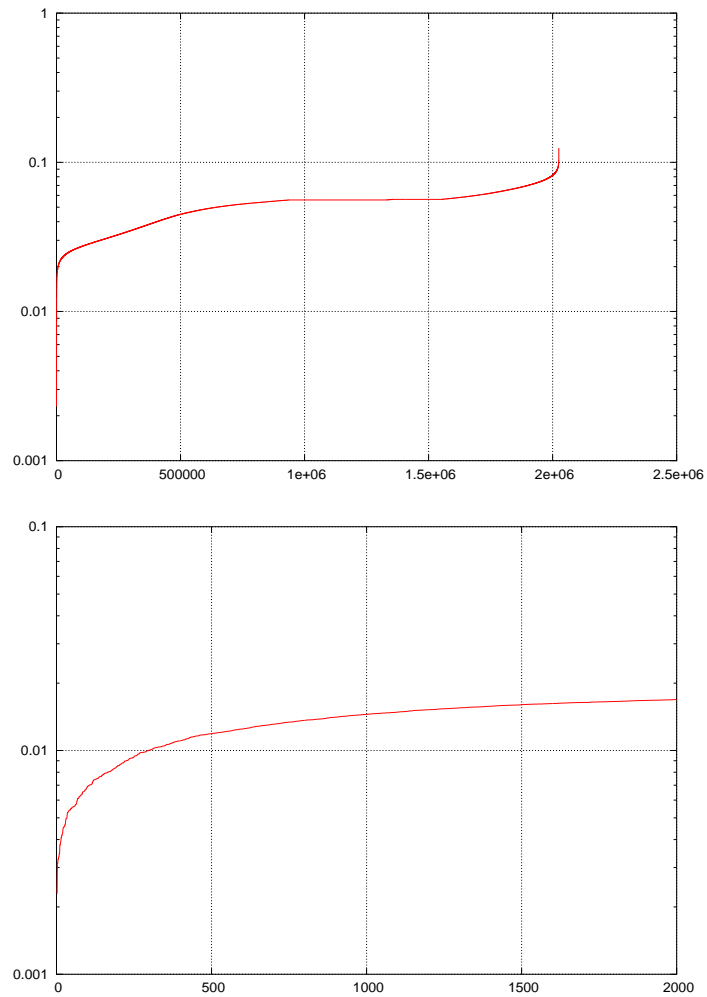


FIG. 6.3 – Diffraction d'une onde plane par un jet d'affaire -  $F=200$  MHz  
Distribution du critère géométrique pour la définition de  $S_e$  et  $S_i$   
Abscisse : indice de tétraèdre - Ordonnée : valeur du critère

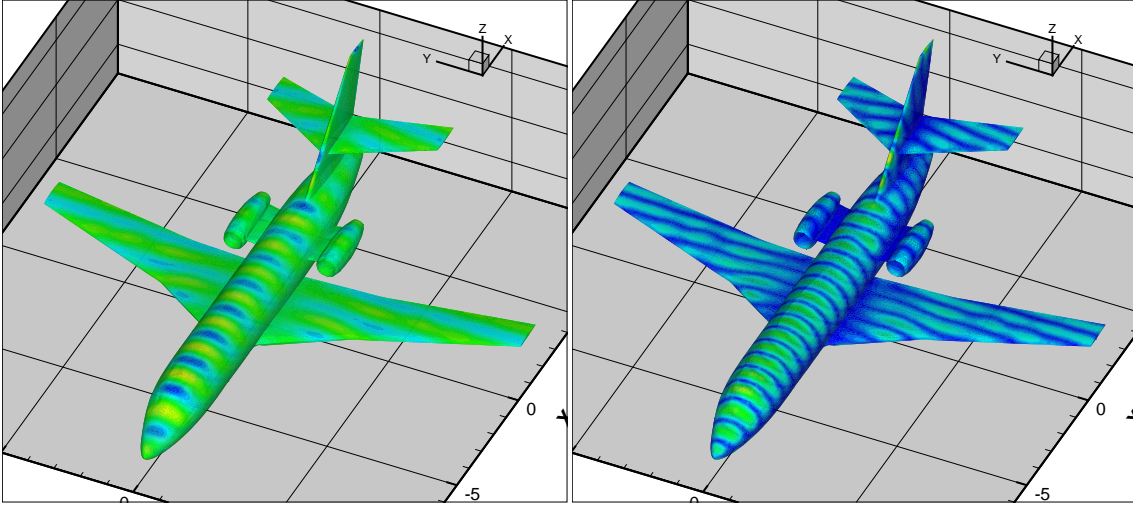


FIG. 6.4 – Diffraction d'une onde plane par un jet d'affaire -  $F=200$  MHz  
Méthode GDDT- $\mathbb{P}_1$  - Lignes de contour de  $E_z$  (gauche) et de  $|\mathbf{E}|$  sur la surface de l'avion

### 6.2.2 Propagation dans une cavité sphérique ouverte

Nous considérons maintenant un problème de propagation d'une onde plane dans un domaine parallélépipédique incluant une cavité sphérique comportant en l'un de ses pôles une ouverture circulaire. La cavité sphérique a un rayon extérieur  $R=0.2$  m et une épaisseur  $e=5 \cdot 10^{-3}$  m. Le rayon de l'ouverture circulaire est  $r=2.5 \cdot 10^{-2}$  m. Un maillage tétraédrique est construit comme l'assemblage d'une partie uniforme et d'une partie non-uniforme raffinée au voisinage de l'ouverture circulaire. Différentes vues du maillage et de la géométrie sont montrées sur la Fig. 6.6. Une onde plane incidente de fréquence  $F=1$  GHz (i.e. de longueur d'onde  $\lambda = 0.3$  m) est partiellement diffractée par la sphère et se propage aussi à l'intérieur de la cavité. Le maillage tétraédrique utilisé comporte 56484 sommets et 301116 tétraèdres. Les longueurs minimale et maximale des arêtes de ce maillage sont respectivement égales à  $5.04 \cdot 10^{-4}$  m (ce qui représente  $\approx \lambda/595$  m) et  $5.0943 \cdot 10^{-2}$  m (soit  $\approx \lambda/5.9$  m). Les valeurs minimale et maximale du pas de temps (hors CFL) sont respectivement égales à  $\Delta t_{\min} = 1.01$  picosec et  $\Delta t_{\max} = 34.57$  picosec soit un rapport  $\Delta t_{\max}/\Delta t_{\min} \approx 34$ . La distribution du critère  $\mathcal{C}(\tau_i)$  pour les éléments  $\tau_i$  de ce maillage est visualisée sur la Fig. 6.7 (la figure inférieure est un zoom sur les plus petites valeurs de ce critère). Le domaine de calcul est délimité par un parallélépipède sur lequel on applique une condition absorbante de type Silver-Müller. La cavité sphérique est supposée parfaitement conductrice. Le temps de simulation est fixé à 5 périodes de l'onde incidente.

Pour la construction des sous-ensembles  $\mathcal{S}_e$  et  $\mathcal{S}_i$ , on a sélectionné la valeur  $l_{\max} = 2.5 \cdot 10^{-3}$  m et dans ce cas on obtient  $|\mathcal{S}_e| = 300526$  et  $|\mathcal{S}_i| = 590$  (soit environ 0.2% des tétraèdres sont traités en implicite). On montre sur les Fig. 6.8 et Fig. 6.9 les lignes de contour de la composante  $E_x$  sur la surface de la sphère externe et dans un plan de coupe, associées aux solutions numériques fournies par les méthodes GDDT- $\mathbb{P}_1$  et GDDT- $\mathbb{P}_2$  hybrides explicites/implicites. La Fig. 6.10 compare les solutions numériques fournies par les méthodes GDDT- $\mathbb{P}_1$  et GDDT- $\mathbb{P}_2$  globalement explicites et hybrides explicites/implicites sur la base de l'évolution temporelle de la composante  $E_x$  en un point donné du maillage. Les résultats de performance sont rassemblés dans la Tab. 6.4. Le gain en temps de calcul total entre les deux méthodes est respectivement de l'ordre de 5.8 et de 5.2 pour les interpolations  $\mathbb{P}_1$  et  $\mathbb{P}_2$ .



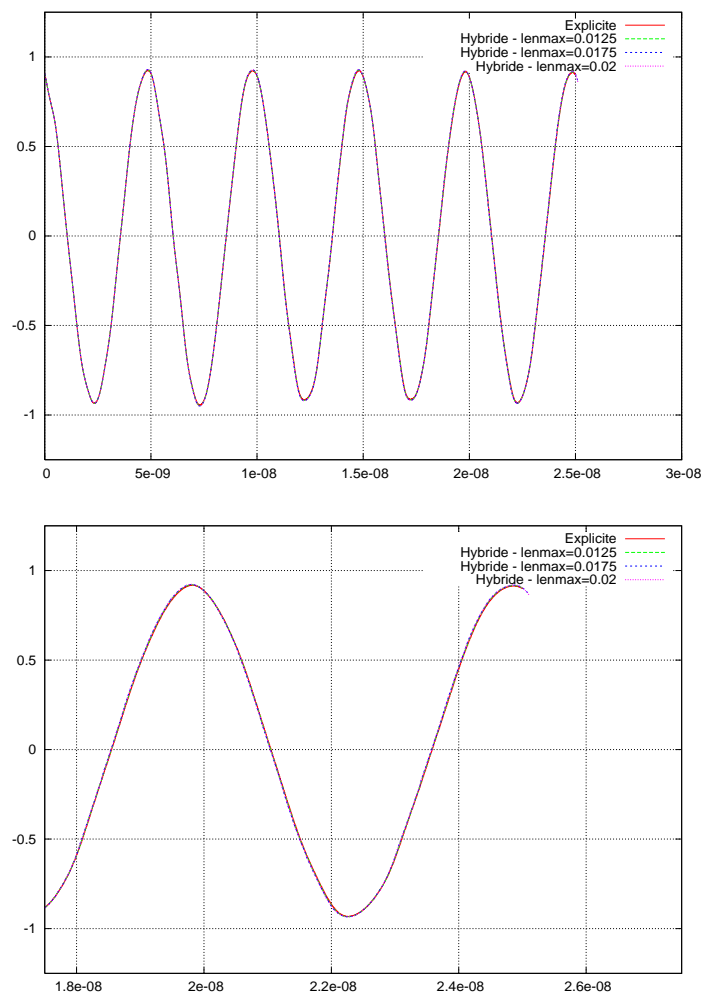


FIG. 6.5 – Diffraction d'une onde plane par un jet d'affaire -  $F=200$  MHz  
Evolution temporelle de  $E_z$  en un point du domaine (abscisse : temps en sec - ordonnée :  $E_z$ )

Méthode	Intégration en temps	Mémoire RAM (LU)	Temps CPU (LU)	Temps CPU (total)
GDDT- $\mathbb{P}_1$	Explicite	-	-	4 h 38 mn
-	Hybride explicite/implicite	66 MB	1.5 sec	48 mn
GDDT- $\mathbb{P}_2$	Explicite	-	-	30 h 39 mn
-	Hybride explicite/implicite	278 MB	9.5 sec	5 h 49 mn

TAB. 6.4 – Diffraction d'une onde plane par une cavité sphérique ouverte -  $F=1$  GHz  
Temps CPU (station de travail basée sur un processeur Intel Xeon 2.33 GHz)

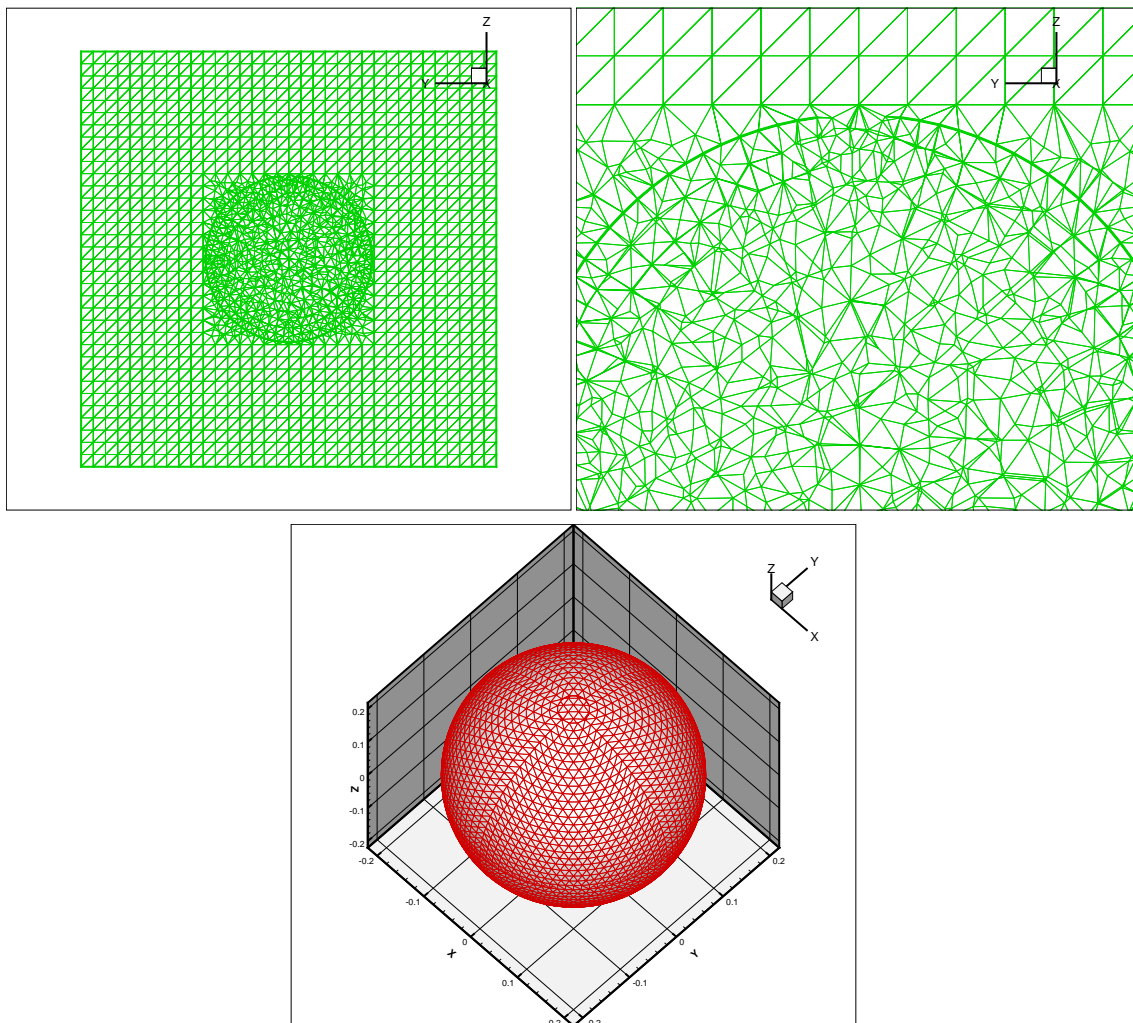


FIG. 6.6 – Diffraction d'une onde plane par une cavité sphérique ouverte -  $F=1$  GHz  
Coupe du maillage volumique (haut) et maillage triangulaire de surface (bas)

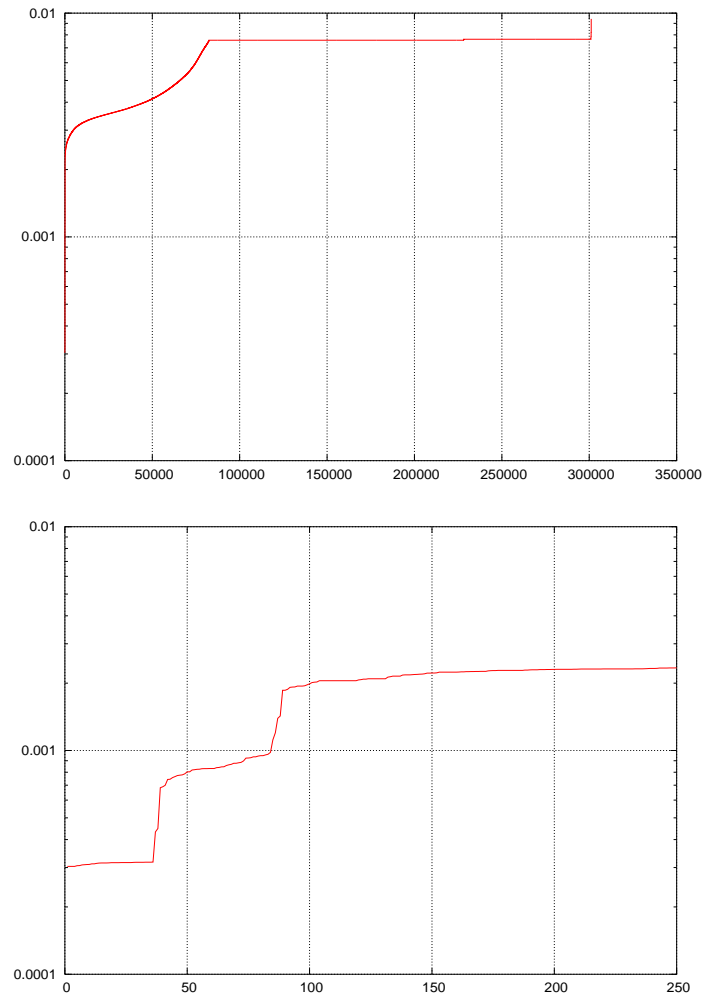


FIG. 6.7 – Diffraction d'une onde plane par une cavité sphérique ouverte -  $F=1$  GHz  
Distribution du critère géométrique pour la définition de  $S_e$  et  $S_i$   
Abscisse : indice de tétraèdre - Ordonnée : valeur du critère

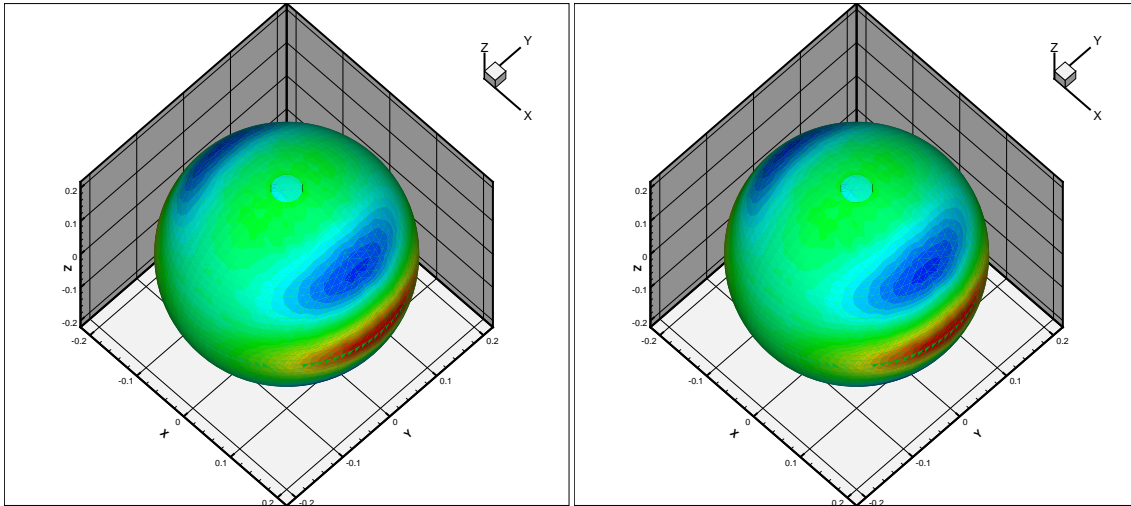


FIG. 6.8 – Diffraction d'une onde plane par une cavité sphérique ouverte -  $F=1$  GHz  
Lignes de contour de  $E_x$  sur la surface de la sphère externe : méthodes GDDT- $\mathbb{P}_1$  (gauche) et GDDT- $\mathbb{P}_2$  (droite)

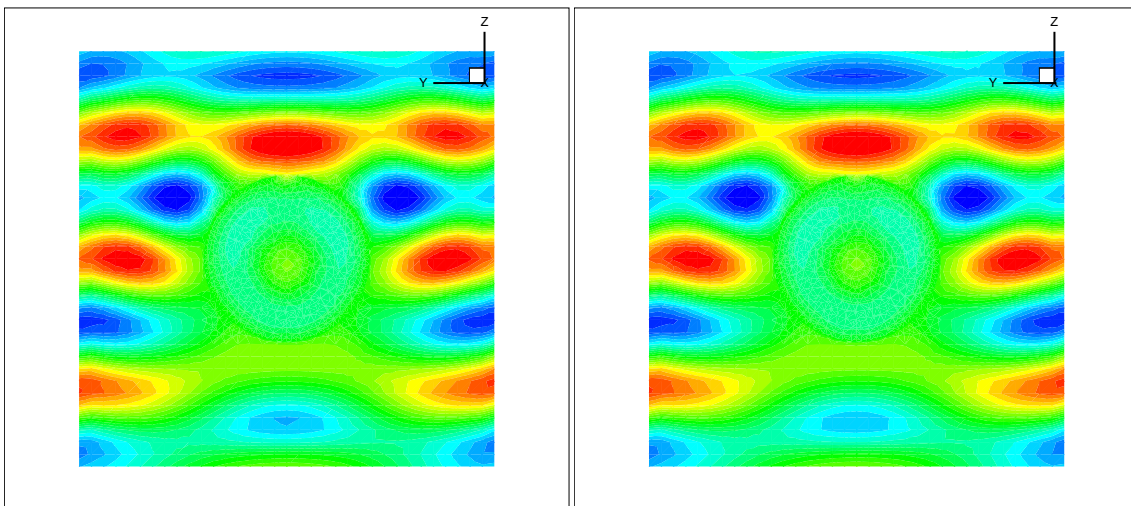


FIG. 6.9 – Diffraction d'une onde plane par une cavité sphérique ouverte -  $F=1$  GHz  
Lignes de contour de  $E_x$  dans un plan de coupe : méthodes GDDT- $\mathbb{P}_1$  (gauche) et GDDT- $\mathbb{P}_2$  (droite)

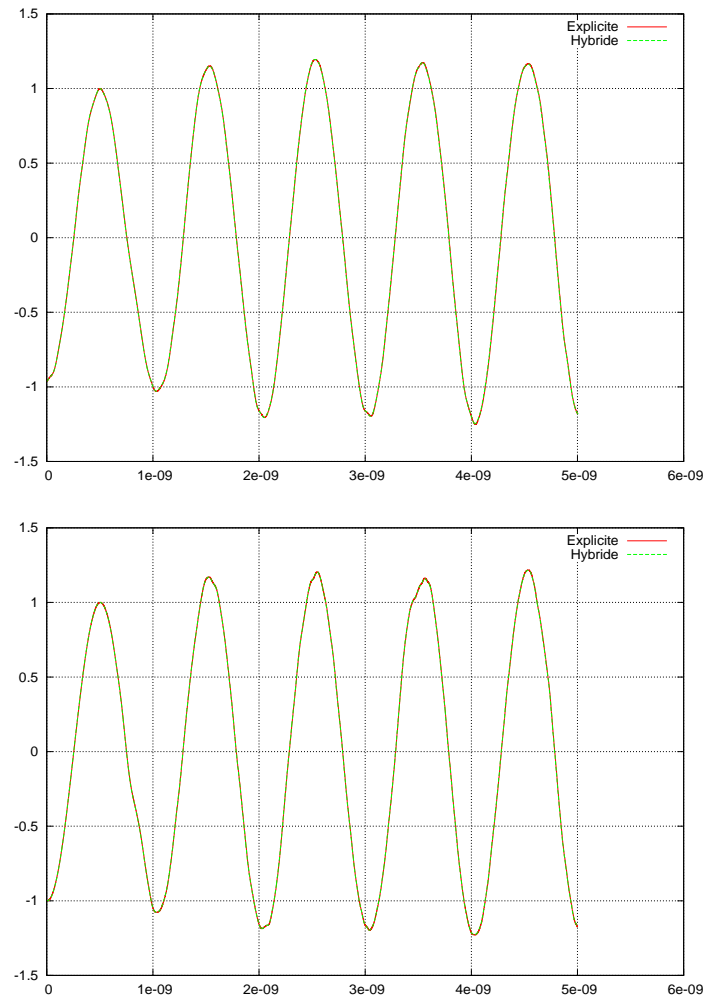


FIG. 6.10 – Diffraction d'une onde plane par une cavité sphérique ouverte -  $F=1$  GHz  
Evolution temporelle de  $E_x$  en un point du domaine (abscisse : temps en sec - ordonnée :  $E_x$ )  
Méthodes GDDT- $\mathbb{P}_1$  (haut) et GDDT- $\mathbb{P}_2$  (bas)

### 6.2.3 Diffraction d'une onde plane par une géométrie de missile

On traite ici d'un problème de propagation mettant en jeu un maillage construit comme l'assemblage d'une partie uniforme et d'une partie non-uniforme raffinée. La partie non-uniforme contient une géométrie de missile. Différentes vues du maillage et de la géométrie sont montrées sur la Fig. 6.11. Afin de mettre en avant le gain potentiel en temps de calcul dans une situation de non-uniformité extrême, on a artificiellement raffiné le maillage volumique au voisinage de la tête du missile. Le problème considéré est la diffraction d'une onde plane de fréquence  $F=300$  MHz (i.e. de longueur d'onde  $\lambda = 1.0$  m) par la géométrie de missile. Le maillage tétraédrique utilisé comporte 39660 sommets et 205485 tétraèdres. Les longueurs minimale et maximale des arêtes de ce maillage sont respectivement égales à  $8.24 \cdot 10^{-4}$  m (ce qui représente  $\approx \lambda/1213$  m) et  $1.79579 \cdot 10^{-1}$  m (soit  $\approx \lambda/5.7$  m). Les valeurs minimale et maximale du pas de temps (hors CFL) sont respectivement égales à  $\Delta t_{\min} = 0.86$  picosec et  $\Delta t_{\max} = 164.7$  picosec soit un rapport  $\Delta t_{\max}/\Delta t_{\min} \approx 192$ . La distribution du critère  $\mathcal{C}(\tau_i)$  pour les éléments  $\tau_i$  de ce maillage est visualisée sur la Fig. 6.12 (la figure inférieure est un zoom sur les plus petites valeurs de ce critère). Le domaine de calcul est délimité par un parallélogramme sur lequel on applique une condition absorbante de type Silver-Müller. Le missile est supposé parfaitement conducteur. Le temps de simulation est fixé à 5 périodes de l'onde incidente.

Pour la construction des sous-ensembles  $\mathcal{S}_e$  et  $\mathcal{S}_i$ , on a sélectionné la valeur  $l_{\max} = 0.01$  m et dans ce cas on obtient  $|\mathcal{S}_e| = 204677$  et  $|\mathcal{S}_i| = 808$  (soit environ 0.4% des tétraèdres sont traités en implicite). On montre sur les Fig. 6.13 et Fig. 6.14 les lignes de contour des composantes  $E_y$  sur la surface du missile et  $E_x$  dans un plan de coupe, associées aux solutions numériques fournies par les méthodes GDDT- $\mathbb{P}_1$  et GDDT- $\mathbb{P}_2$  hybrides explicites/implicites. La Fig. 6.15 compare les solutions numériques fournies par les méthodes GDDT- $\mathbb{P}_1$  et GDDT- $\mathbb{P}_2$  globalement explicites et hybrides explicites/implicites sur la base de l'évolution temporelle de la composante  $E_x$  en un point donné du maillage. Les résultats de performance sont rassemblés dans la Tab. 6.5. Le gain en temps de calcul total entre les deux méthodes est respectivement de l'ordre de 30 et de 28 pour les interpolations  $\mathbb{P}_1$  et  $\mathbb{P}_2$ .

Méthode	Intégration en temps	Mémoire RAM (LU)	Temps CPU (LU)	Temps CPU (total)
GDDT- $\mathbb{P}_1$	Explicite	-	-	12 h 14 mn
-	Hybride explicite/implicite	66 MB	1.5 sec	27 mn
GDDT- $\mathbb{P}_2$	Explicite	-	-	83 h 40 mn
-	Hybride explicite/implicite	278 MB	9.5 sec	2 h 58 mn

TAB. 6.5 – Diffraction d'une onde plane par une géométrie de missile -  $F=300$  MHz  
Temps CPU (station de travail basée sur un processeur Intel Xeon 2.33 GHz)

## 6.3 Conclusion

Les résultats numériques présentés dans ce chapitre ont permis de montrer l'intérêt d'une stratégie de résolution hybride explicite/implicite pour des simulations 3D mettant en jeu des maillages tétraédriques non-uniformes. Le gain en temps de calcul entre les méthodes GDDT- $\mathbb{P}_p$  globalement explicite et hybride explicite/implicite est de l'ordre de 5 lorsque les maillages sont tels que le rapport  $\Delta t_{\max}/\Delta t_{\min}$  est de l'ordre de quelques dizaines. Pour un cas de calcul reposant sur un maillage localement raffiné au voisinage d'une singularité géométrique, un gain en temps de calcul de l'ordre de 30 a été obtenu. Le surcoût en consommation mémoire du fait du stockage des facteurs L et U reste acceptable lorsque le maillage

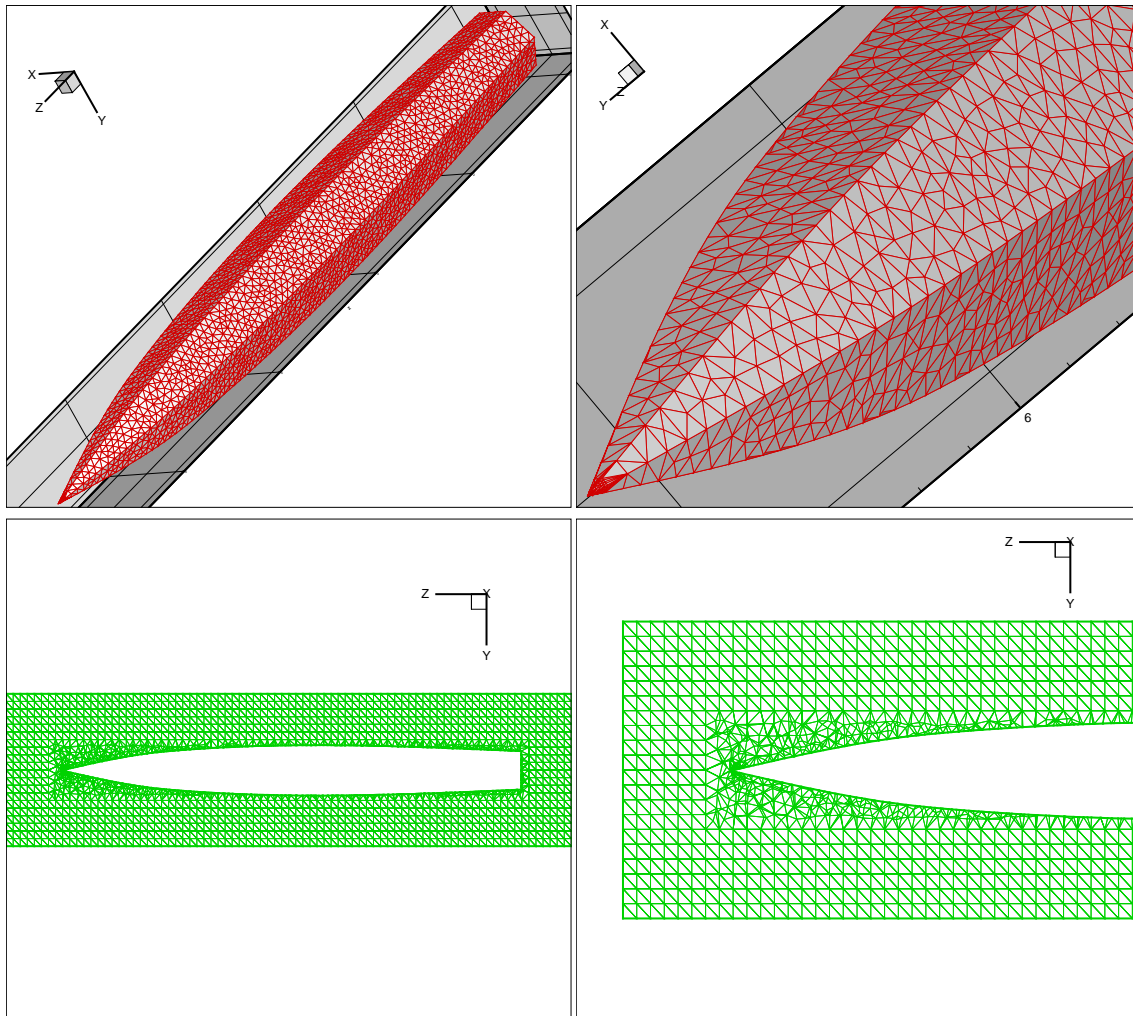


FIG. 6.11 – Diffraction d'une onde plane par une géométrie de missile -  $F=300$  MHz  
Coupe du maillage volumique (haut) et maillage triangulaire de surface (bas)

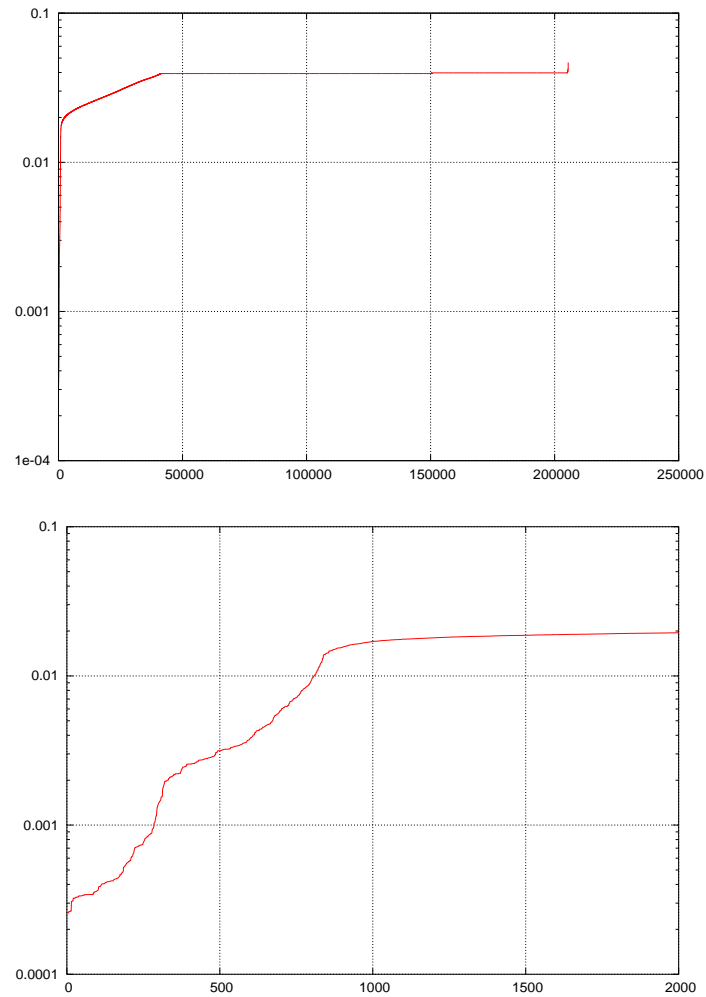


FIG. 6.12 – Diffraction d'une onde plane par une géométrie de missile -  $F=300$  MHz  
Distribution du critère géométrique pour la définition de  $\mathcal{S}_e$  et  $\mathcal{S}_i$   
Abscisse : indice de tétraèdre - Ordonnée : valeur du critère



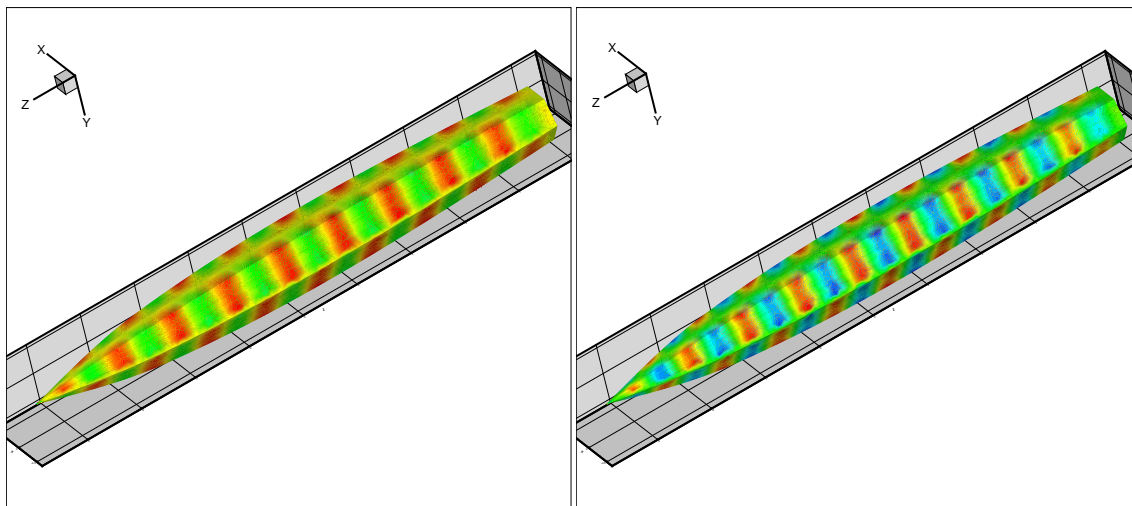


FIG. 6.13 – Diffraction d'une onde plane par une géométrie de missile -  $F=300$  MHz  
Lignes de contour de  $E_y$  sur la surface du missile : méthodes GDDT- $\mathbb{P}_1$  (gauche) et GDDT- $\mathbb{P}_2$  (droite)

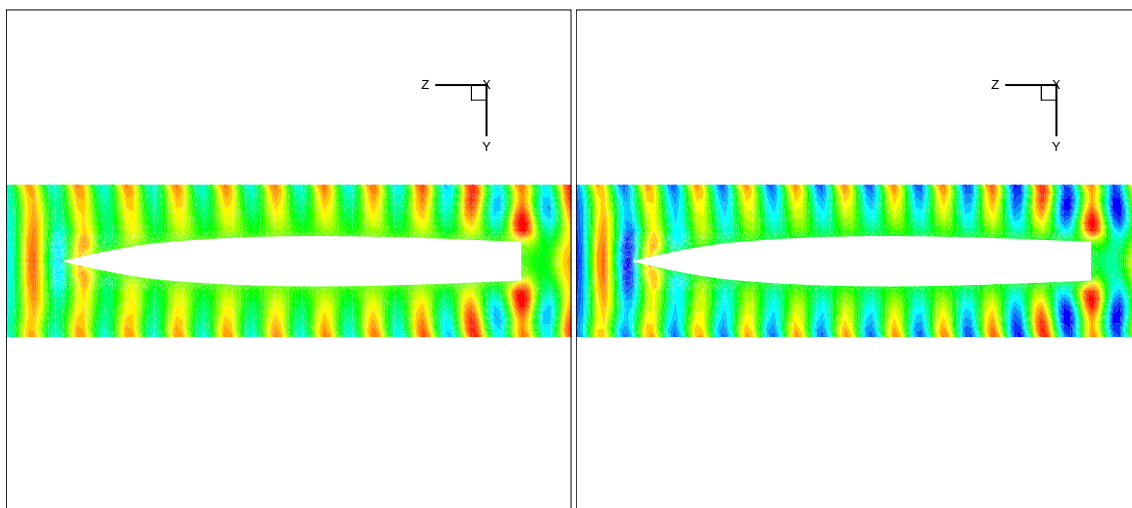


FIG. 6.14 – Diffraction d'une onde plane par une géométrie de missile -  $F=300$  MHz  
Lignes de contour de  $E_x$  dans un plan de coupe : méthodes GDDT- $\mathbb{P}_1$  (gauche) et GDDT- $\mathbb{P}_2$  (droite)

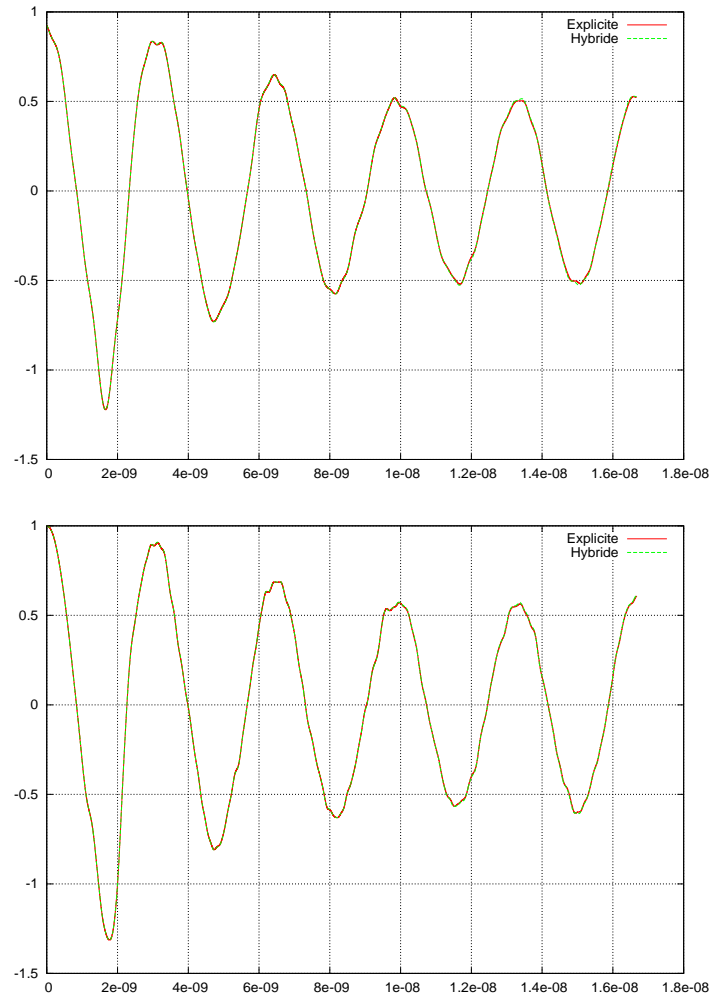


FIG. 6.15 – Diffraction d’une onde plane par une géométrie de missile -  $F=300$  MHz  
Evolution temporelle de  $E_x$  en un point du domaine (abscisse : temps en sec - ordonnée :  $E_x$ )  
Méthodes GDDT- $\mathbb{P}_1$  (haut) et GDDT- $\mathbb{P}_2$  (bas)

est localement raffiné et le nombre d'éléments traités en implicite représente moins d'1% du nombre total d'éléments. On s'attend à plus de flexibilité dans l'exploitation de la méthode GDDT- $\mathbb{P}_p$  hybride explicite/implicite avec sa parallélisation ou/et l'utilisation d'un solveur itératif préconditionné pour la résolution du système implicite.



# Conclusion générale et perspectives

Rappelons que l'objectif principal de cette thèse était de mettre au point des stratégies d'intégration en temps permettant d'améliorer l'efficacité d'une méthode Galerkin discontinue  $\text{GDDT-}\mathbb{P}_k$  arbitrairement d'ordre élevé en espace initialement introduite dans [Piperno and Fezoui, 2003]-[Fezoui et al., 2005], et reposant sur l'utilisation de maillages en simplexes (triangles en 2D et tétraèdres en 3D), pour la résolution numérique des équations de Maxwell en domaine temporel. Il s'agissait notamment de traiter des situations où le maillage de calcul est non-uniforme (localement raffiné) qui, lorsque les équations semi-discrétisées sont intégrées en temps par un schéma explicite, conduisent à des temps de calcul importants voire prohibitifs pour certains problèmes tridimensionnels du fait de la condition de stabilité imposée sur le pas de temps.

Nous avons tout d'abord considéré des schémas d'intégration en temps globalement implicites en remplacement du schéma explicite saute-mouton utilisé dans la méthode Galerkin discontinue de départ (méthode  $\text{LF2-GDDT-}\mathbb{P}_k$  dans la terminologie adoptée dans ce manuscrit). En premier lieu, nous avons considéré le schéma de Crank-Nicolson déjà très utilisé par ailleurs mais pas à notre connaissance dans le contexte considéré ici. La méthode  $\text{CN2-GDDT-}\mathbb{P}_k$  résultante a été étudiée théoriquement (stabilité inconditionnelle, inversibilité du système implicite associé et dispersion numérique en 1D) et numériquement dans le contexte de la résolution des équations de Maxwell 2D. Dans ce cadre 2D, le système linéaire associé au schéma implicite est résolu au moyen d'un solveur direct (factorisation LU) optimisé pour les matrices creuses. Des gains conséquents en termes de temps de calcul ont été obtenus pour des calculs en maillages non-uniformes. Cependant, les problèmes de propagation considérés étant purement instationnaires ou périodiques, on conçoit bien que la précision du schéma définit implicitement une limite sur le pas de temps pour que l'erreur d'approximation reste acceptable. Dans le cas du schéma Crank-Nicolson, la précision en temps est du second ordre et la méthode Galerkin discontinue implicite résultante est non-dissipative mais dispersive et la dispersion numérique est plus importante que celle du schéma explicite saute-mouton (LF2). Nous avons donc étudié la possibilité d'augmenter la précision en temps de la méthode  $\text{CN2-GDDT-}\mathbb{P}_k$  par l'intermédiaire d'une technique de défaut corrigé. La méthode  $\text{CN4-GDDT-}\mathbb{P}_k$  résultante est effectivement plus précise mais le coût de calcul par itération en temps est doublé du fait de la résolution de deux systèmes linéaires au lieu d'un pour la méthode  $\text{CN2-GDDT-}\mathbb{P}_k$ , si bien que l'exploitation efficace de cette méthode ne semble possible que pour des maillages présentant des taux de non-uniformité importants et pour des degrés d'interpolation en espace élevés. Une piste plus prometteuse est discutée plus bas.

Même si l'utilisation d'un schéma d'intégration en temps implicite permet une diminution notable du temps de calcul pour un niveau d'erreur numérique identique à celui du schéma d'intégration en temps explicite, la résolution d'un système implicite global reste très coûteuse. En trois dimensions d'espace ou même pour des problèmes bidimensionnels de taille conséquente, un solveur direct adapté aux matrices creuses pour la résolution des systèmes implicites à chaque itération en temps n'est simplement pas viable. Plusieurs stratégies peuvent être envisagées pour améliorer cette situation. Dans cette thèse, nous nous sommes penchés sur la possibilité de combiner des schémas d'intégration en temps explicite et implicite

ou, en d'autres termes, de concevoir une méthodologie numérique localement implicite. Une telle stratégie a récemment été proposée dans [Piperno, 2006b]-[Piperno, 2006a]. Jusqu'ici, elle n'avait pas fait l'objet d'une étude concrète dans le contexte de la résolution numérique des équations de Maxwell en domaine temporel. Nous en avons donc proposé une mise en œuvre dans laquelle les équations différentielles ordinaires résultant de la semi-discrétisation du système de Maxwell par une formulation Galerkin discontinue sont partitionnées en deux sous-ensembles. En pratique, ces sous-ensembles correspondent à une séparation des éléments du maillage sur la base d'un critère géométrique. La méthode LF2CN2-GDDT- $\mathbb{P}_k$  résultante est d'ordre arbitrairement élevé en espace et du second ordre en temps. Plusieurs expériences numériques en 2D ont permis d'étudier en détail les performances en termes de précision, temps de calcul et occupation mémoire de la méthode GDDT- $\mathbb{P}_k$  hybride explicite/implicite pour  $p = 1, 2$  et pour des simulations mettant en jeu des maillages non-uniformes. Dans tous les cas, cette méthode offre un bon compromis entre temps de calcul et occupation mémoire comparativement aux méthodes GDDT- $\mathbb{P}_k$  globalement explicite ou implicite.

Sur la base des résultats de l'étude numérique menée en 2D, il était important de démontrer les possibilités de la méthode hybride LF2CN2-GDDT- $\mathbb{P}_k$  pour la simulation numérique de problèmes de propagation tridimensionnels. Une implémentation en mode séquentiel a donc été réalisée où le système linéaire associé aux éléments traités par le schéma implicite est résolu au moyen d'un solveur direct. Cette implémentation a permis d'obtenir des résultats encourageants avec des gains en termes de temps de calcul entre les méthodes LF2-GDDT- $\mathbb{P}_k$  et LF2CN2-GDDT- $\mathbb{P}_k$  au minimum de l'ordre de 5. De plus, le surcoût en consommation mémoire du fait du stockage des facteurs L et U reste acceptable lorsque le maillage est localement raffiné et le nombre d'éléments traités en implicite représente moins d'1% du nombre total d'éléments.

On peut envisager deux perspectives principales aux travaux réalisés dans cette thèse :

- la méthodologie numérique hybride explicite/implicite étudiée ici mériterait d'être consolidée et étendue pour permettre la simulation numérique efficace de problèmes de propagation d'ondes tridimensionnels réalistes.

Tout d'abord, il serait souhaitable d'étudier théoriquement la stabilité et la convergence de la méthode LF2CN2-GDDT- $\mathbb{P}_k$ . La stabilité a été partiellement analysée via une approche énergétique par Piperno dans [Piperno, 2006a] mais l'auteur n'a pas obtenu une expression explicite du pas de temps admissible. L'étude de convergence *a priori* pourrait s'inspirer de la démarche adoptée dans [Fezoui et al., 2005].

Par ailleurs, le critère nécessaire à la construction des sous-ensembles d'éléments  $S_e$  et  $S_i$  a été défini de façon empirique. Une approche plus robuste, et aussi plus ambitieuse, pourrait s'appuyer sur un critère qui combine un indicateur d'erreur (de dispersion numérique dans le cas présent) pour un maillage donné, à une évaluation dynamique de la capacité mémoire disponible (surtout dans l'optique de l'utilisation d'une factorisation LU pour la résolution du système implicite).

Dans cette thèse, nous avons aussi proposé et évalué un schéma d'intégration en temps implicite de Crank-Nicolson précis au quatrième ordre. Les expériences numériques réalisées en 2D ont bien montré une amélioration de la précision et de la convergence de la méthode GDDT- $\mathbb{P}_k$  associée mais l'augmentation du coût de calcul ne permet pas d'imaginer une exploitation de cette méthode en 3D. En revanche, il serait intéressant d'étudier la possibilité de mettre au point une méthode GDDT- $\mathbb{P}_k$  hybride explicite/implicite combinant les schémas LF4 et CN4.

Enfin, d'un point de vue plus informatique, une mise œuvre de la méthode LF2CN2-GDDT- $\mathbb{P}_k$  adaptée aux plate-formes de calcul parallèles est une étape incontournable. La difficulté principale dans cette optique concerne la stratégie de partitionnement du maillage qui doit tenir compte de l'hétérogénéité des calculs et des communications liée à l'approche hybride explicite/implicite.

- Les méthodes GDDT- $\mathbb{P}_k$  globalement implicite et hybride explicite/implicite nécessitent la résolution d'un ou plusieurs systèmes linéaires à chaque itération temporelle. Dans nos travaux, nous avons

principalement adopté un solveur direct pour la résolution de ces systèmes. Nous avons aussi réalisé une étude numérique en 2D visant à évaluer une stratégie de résolution itérative basée sur l'association de la méthode GMRES et de la technique de préconditionnement ILUT. Pour les problèmes 3D, deux directions éventuellement complémentaires méritent d'être étudiées.

En premier lieu, on pourrait chercher à concevoir un algorithme de résolution itératif parallèle par décomposition de domaine. Pour la résolution des équations de Maxwell en régime harmonique discrétisées par une méthode Galerkin discontinue, des algorithmes de Schwarz ont été proposés dans [Dolean et al., 2007]-[Dolean et al., 2008] qui pourraient servir de base à une telle étude.

Ensuite, en vue de mettre au point une stratégie de résolution itérative efficace et exploitable en 3D, basée sur une méthode de Krylov comme GMRES, la recherche d'un préconditionneur adapté aux caractéristiques (mathématiques et structurales) des systèmes linéaires en question s'avèrera indispensable.





# Bibliographie

- [Ainsworth, 2004] Ainsworth, M. (2004). Dispersive and dissipative behaviour of high order discontinuous Galerkin finite element methods. *J. Comput. Phys.*, 198(1) :106–130.
- [Amestoy et al., 2000] Amestoy, P., Duff, I., and L'Excellent, J.-Y. (2000). Multifrontal parallel distributed symmetric and unsymmetric solvers. *Comput. Methods Appl. Mech. Eng.*, 184 :501–520.
- [Amestoy et al., 2001] Amestoy, P. R., Duff, I. S., Koster, J., and L'Excellent., J. Y. (2001). A fully asynchronous multifrontal solver using distributed dynamic scheduling. *SIAM J. Mat. Anal. and Appl.*, 23 :15–41.
- [Amestoy et al., 1998] Amestoy, P. R., Duff, I. S., and L'Excellent., J. Y. (1998). Multifrontal parallel distributed symmetric and unsymmetric solvers. *Comput. Methods in Appl. Mech. Eng.*, 184 :501–520.
- [Arnoldi, 1951] Arnoldi, W. (1951). The principle of minimized iterations in the solution of the matrix eigenvalue problem. *Quarterly of applied mathematics*, 9 :17–29.
- [Artuzi, 2004] Artuzi, W. (2004). An unconditionally stable FDTD method using tetrahedral cells. MO-MAG.
- [Atkins and Hu, 2002a] Atkins, H. and Hu, F. (2002a). Eigensolution analysis of the discontinuous Galerkin method with non uniform grids I. one space dimension. *J. Comput. Phys.*, 182 :516–545.
- [Atkins and Hu, 2002b] Atkins, H. and Hu, F. (2002b). Two dimensional wave analysis of the discontinuous Galerkin method with non-uniform grids and boundary conditions. *AIAA*, 2002-2514.
- [Beilina and Grote, 2004] Beilina, L. and Grote, M. J. (2004). Adaptive hybrid finite element/difference method for Maxwell's equations. Technical report, University of Basel, Switzerland.
- [Bernacki et al., 2006a] Bernacki, M., Fezoui, L., Lanteri, S., and Piperno, S. (2006a). Parallel unstructured mesh solvers for heterogeneous wave propagation problems. *Appl. Math. Model.*, 30(8) :744–763.
- [Bernacki et al., 2006b] Bernacki, M., Lanteri, S., and Piperno, S. (2006b). 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, 1997] Bonnet, F. (1997). *Méthodes de résolution efficaces pour le système de Maxwell instationnaire*. PhD thesis, Université de Nice-Sophia Antipolis.
- [Butrylo et al., 2004] Butrylo, B., Vollaire, C., and Nicolas, L. (2004). stability and fidelity of the finite element time doamin method with distorted mesh. *IEEE Trans. on Magn*, 40(2) :1424–1427.
- [Cai and Deng, 2003] Cai, W. and Deng, S. (2003). An upwinding embedded boundary method for Maxwell's equations in media with material interfaces : 2D case. *J. Comput. Phys.*, 190(1) :159–183.

- [Chaudhuri and Van, 1999] Chaudhuri, S. and Van, V. (1999). A hybrid implicit-explicit FDTD scheme for nonlinear optical waveguide modeling. *IEEE Trans. Microwave Theory Tech.*, 47(5) :540–545.
- [Chen et al., 2000] Chen, C.-P., Lee, T.-W., Murugesan, N., and Hagness, S. (2000). Generalized FDTD-ADI : an unconditionally stable full-wave Maxwell's equations solver for VLSI interconnect modeling. In *2000 IEEE/ACM International Conference on Computer-Aided Design*, pages 156–164.
- [Ciarlet and Lions, 1991] Ciarlet, P. and Lions, J., editors (1991). *Finite element methods*, volume 2 of *Handbook of Numerical Analysis*. North-Holland.
- [Cioni, 1995] Cioni, J. (1995). *Résolution numérique des équations de Maxwell instationnaires par une méthode Volumes Finis*. PhD thesis, Université de Nice Sophia Antipolis.
- [Cioni et al., 1997] Cioni, J., Fezoui, L., Anne, L., and Poupaud, F. (1997). A parallel FVTD Maxwell solver using 3D unstructured meshes. In *13th Annual Review of Progress in Applied Computational Electromagnetics (PIERS)*, pages 359–365.
- [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.
- [Cockburn et al., 2000] Cockburn, B., Karniadakis, G., and Shu, C., editors (2000). *Discontinuous Galerkin methods. Theory, computation and applications*, volume 11 of *Lecture Notes in Computational Science and Engineering*. Springer-Verlag.
- [Cockburn and Shu, 2005] Cockburn, B. and Shu, C., editors (2005). *Special issue on discontinuous Galerkin methods*, volume 22-23 of *J. Sci. Comput.* Springer.
- [Curtiss and Hirschfelder, 1952] Curtiss, C. F. and Hirschfelder, J. O. (1952). Integration of stiff equation. *US Nat Acad. Sci.*, 38(3) :235–243.
- [Dahlquist, 1963] Dahlquist, G. (1963). A special stability problem for linear multistep methods. *BIT*, 3(1) :27–43.
- [Dautray and Lions, 1987] Dautray, R. and Lions, J. (1987). *Analyse mathématique et calcul numérique pour les sciences et les techniques*, volume 1. MASSON.
- [Dawson, 2006] Dawson, C., editor (2006). *Special issue on discontinuous Galerkin methods*, volume 195 of *Comput. Meth. App. Mech. Engng.* Elsevier.
- [Dolean et al., 2007] Dolean, V., Lanteri, S., and Perrussel, R. (2007). A domain decomposition method for solving the three-dimensional time-harmonic Maxwell equations discretized by discontinuous Galerkin methods. *J. Comput. Phys.*, 227(3) :2044–2072.
- [Dolean et al., 2008] Dolean, V., Lanteri, S., and Perrussel, R. (2008). Optimized Schwarz algorithms for solving time-harmonic Maxwell's equations discretized by a discontinuous Galerkin method. *IEEE. Trans. Magn.*, 44(6) :954–957.
- [Dubois et al., 2005] Dubois, T., Jauberteau, F., Temam, R., and Tribbia, J. (2005). Multilevel schemes for the shallow water equations. *J. Comput. Phys.*, 207(2) :660–694.
- [Elmkies and Joly, 1997a] Elmkies, A. and Joly, P. (1997a). Eléments finis d'arete et condensation de masse pour les équations de Maxwell : le cas 2D. *C.R. Acad. Sci. Paris*, 324(11) :1287–1293.
- [Elmkies and Joly, 1997b] Elmkies, A. and Joly, P. (1997b). Eléments finis d'arete et condensation de masse pour les équations de Maxwell : le cas de dimension 3. *C.R. Acad. Sci. Paris*, 325(11) :1217–1222.
- [Engquist and Majda, 1977] Engquist, B. and Majda, A. (1977). Absorbing boundary conditions for the numerical simulation of waves. *Math. Comp.*, 31 :629–651.

- [Fahs, 2008] Fahs, H. (2008). *Méthodes de type Galerkin discontinu d'ordre élevé pour la résolution numérique des équations de Maxwell instantanées sur des maillages simplexes non-conformes*. PhD thesis, Université de Nice-Sophia Antipolis.
- [Ferrieres et al., 2004] Ferrieres, X., Parmantier, J., Bertuol, S., and Ruddle, A. (2004). Application of a hybrid finite difference/finite volume method to solve an automotive EMC problem. *IEEE Trans. Electromagnetic Compatibility*, 46(4) :624–634.
- [Feynman, 1963] Feynman, R. (1963). *Lectures on physics : mechanics, radiations and heat*, volume 1. Addison-Wesley.
- [Feynman, 1964] Feynman, R. (1964). *Lectures on physics : electromagnetism and matter*, volume 2. Addison-Wesley.
- [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 : Math. Model. Num. Anal.*, 39(6) :1149–1176.
- [Fumeaux et al., 2006] Fumeaux, C., Baumann, D., Bonnet, P., and Vahldieck, R. (2006). Developments of finite volume techniques for electromagnetic modeling on unstructured meshes. *17th International Zurich Symposium on Electromagnetic Compatibility*.
- [Fumeaux et al., 2004] Fumeaux, C., Baumann, D., Leuchtmann, P., and Vahldieck, R. (2004). A generalized local time-step scheme for efficient FDTD simulations in strongly inhomogeneous meshes. *IEEE Trans. Microwave Theory Tech.*, 52(3) :1067–1076.
- [Gear, 1971] Gear, C. W. (1971). *Numerical initial value problems in ordinary differential equations*. Prentice Hall.
- [Gedney and Navsariwala, 1995] Gedney, S. D. and Navsariwala, U. (1995). An unconditionally stable finite element time domain solution of the vector wave equation. *IEEE Microwave Guided Wave Letters*, 5 :332–334.
- [Givoli, 1991] Givoli, D. (1991). Non-reflecting boundary conditions. *J. Comput. Phys.*, 94 :1–29.
- [Gustafsson and Hemmingsson-Frändén, 1999] Gustafsson, B. and Hemmingsson-Frändén, L. (1999). A fast domain decomposition high order Poisson solver. *J. Scient. Comput.*, 14 :223–243.
- [Gustafsson and Hemmingsson-Frändén, 2000] Gustafsson, B. and Hemmingsson-Frändén, L. (2000). Implicit high order difference methods and domain decomposition. *Appl. Numer. Math.*, 33 :493–500.
- [Gustafsson and Hemmingsson-Frändén, 2002] Gustafsson, B. and Hemmingsson-Frändén, L. (2002). Deferred correction in space and time. *J. Scient. Comput.*, 17(1-4) :541–550.
- [Hairer et al., 1987] Hairer, E., Norsett, S. P., and Wanner, G., editors (1987). *Solving ordinary differential equations I - Non-stiff problems*, volume 8 of *Computational Mathematics*. Springer-Verlag.
- [Hairer and Wanner, 1991] Hairer, E. and Wanner, G., editors (1991). *Solving ordinary differential equations II - Stiff and differential-algebraic problems*, volume 8 of *Computational Mathematics*. Springer-Verlag.
- [He and Sun, 2007] He, Y. and Sun, W. (2007). Stability and convergence of the Crank-Nicolson/Adams-Bashforth scheme for the time-dependent Navier-Stokes equations. *SIAM J. Numer. Anal.*, 45(2) :837–869.
- [Helluy, 1994] Helluy, P. (1994). *Résolution numérique des équations de Maxwell harmonique par une méthode d'éléments finis discontinus*. PhD thesis, Ecole Nationale Supérieure de l'Aéronautique et de l'Espace.
- [Hesthaven et al., 1999] Hesthaven, J., Dinesen, P. G., and Lynov, J. P. (1999). Spectral collocation time-domain modeling of diffractive optical elements. *J. Comput. Phys.*, 155(2) :287–306.

- [Hesthaven and Warburton, 2002] Hesthaven, J. S. and Warburton, T. (2002). Nodal high-order methods on unstructured grids. I. Time-domain solution of Maxwell's equations. *J. Comput. Phys.*, 181 :186–221.
- [Jiao and Jin, 2002] Jiao, D. and Jin, J.-M. (2002). A general approach for the stability analysis of the time domain finite element method for electromagnetic simulations. *IEEE Trans. Antennas and Propagation*, 50(11) :1624–1632.
- [Johnson and Pitkaranta, 1986] Johnson, C. and Pitkaranta, J. (1986). An analysis of the discontinuous Galerkin method for a scalar hyperbolic equation. *Math. Comp.*, 46 :1–26.
- [Jurgens and Taflove, 1993] Jurgens, T. and Taflove, A. (1993). Three-dimensional contour FD-TD modeling of scattering from single and multiple bodies. *IEEE Trans. Antennas and Propagat.*, 41 :1703–1708.
- [Jurgens et al., 1987] Jurgens, T., Taflove, A., Umashankar, K., and Moore, T. (1987). Finite-Difference Time-Domain modeling of curved surfaces. *IEEE Trans. Antennas and Propagat.*, 35 :1248–1257.
- [Lee and Fornberg, 2004] Lee, J. and Fornberg, B. (2004). Some unconditionally stable time stepping methods for the 3d Maxwell's equations. *J. Comput. and Appl. Math.*, 166 :497–523.
- [Lee et al., 1992] Lee, J.-F., Palendech, R., and Mittra, R. (1992). Modeling three-dimensional discontinuities in waveguides using nonorthogonal FD-TD algorithm. *IEEE Trans. Microwave Theory Tech.*, 40 :346–352.
- [Lesaint and Raviart, 1974] Lesaint, P. and Raviart, P. (1974). *On a finite element method for solving the neutron transport equation*, pages 89–145. Academic Press.
- [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 interstitial arrays of sleeved-slot microwave antennas for hyperthermia treatment of brain tumors. *IEEE Trans. Microwave Theory Tech.*, 48(11) :2191–2198.
- [Lindberg, 1980] Lindberg, B. (1980). Error estimation and iterative improvement for discretization algorithms. *BIT*, 20(4) :486–500.
- [Liu and Sarris, 2005] Liu, Y. and Sarris, C. D. (2005). Amr-fdtd : a dynamically adaptive mesh refinement scheme for the finite-difference time-domain technique. *IEEE AP-S International Symposium on Antennas and Propagation*, pages 134–137.
- [Liu and Sarris, 2006] Liu, Y. and Sarris, C. D. (2006). Modeling of dielectric waveguide structures with a multilevel dynamically adaptative mesh refinement (AMR)-FDTD method. *IEEE AP-S International Symposium on Antennas and Propagation*, pages 3841–3844.
- [Ma et al., 1997] Ma, K., Li, M., Drewniak, L., T.H, H., and Doren, T. V. (1997). Comparison of FDTD algorithms for subcellular modeling of slots in shelding enclosures. *IEEE Trans. Electromag. Compat.*, 39 :147–155.
- [Maxwell, 1873] Maxwell, J. (1873). *A treatise on electricity and magnetism*. Clarendon Press, Oxford.
- [Montseny et al., 2008] Montseny, E., Pernet, S., Ferrières, X., and Cohen, G. (2008). Dissipative terms and local time-stepping improvements in a spatial high order Discontinuous Galerkin scheme for the time-domain Maxwell's equations. *J. Comput. Phys.*, 227(14) :6795–6820.
- [Mur, 1981] Mur, G. (1981). Absorbing boundary conditions for the finite-difference approximation of the time domain electromagnetic field equations. *IEEE Trans. Electromag. Compat.*, 37(23).
- [Namiki, 1999] Namiki, T. (1999). A new FDTD algorithm based on alternating-direction implicit method. *IEEE Trans. Microwave Theory Tech.*, 47(10) :2003–2007.
- [Namiki, 2000] Namiki, T. (2000). 3D ADI-FDTD method unconditionnally stable time-domain algorithm for solving full vector Maxwell's equations. *IEEE Trans. Microwave Theory Tech.*, 48(10) :1743–1748.

- [Nedelec, 1980] Nedelec, J. C. (1980). Mixed finite elements in  $\mathbb{R}^3$ . *Numerische Mathematik*, 35(3) :315–341.
- [Pereyra, 1967a] Pereyra, V. (1967a). Accelerating the convergence of discretization algorithms. *SIAM J. Numer. Anal.*, 4(4) :508–533.
- [Pereyra, 1967b] Pereyra, V. (1967b). Iterated deferred corrections for nonlinear operator equations. *J. Numer. Math.*, 10(4) :316–323.
- [Pereyra, 1968] Pereyra, V. (1968). Iterated deferred corrections for nonlinear boundary value problems. *J. Numer. Math.*, 11(2) :111–125.
- [Peterson, 1991] Peterson, T. (1991). A note on the convergence of the discontinuous Galerkin method for a scalar hyperbolic equation. *SIAM J. Numer. Anal.*, 28 :133–140.
- [Piperno, 2006a] Piperno, S. (2006a). DGTD methods using modal basis functions and symplectic local time-stepping : application to wave propagation problems. *Europ. J. of Comp. Mechanics*, 15(6) :643–670.
- [Piperno, 2006b] Piperno, S. (2006b). Symplectic local time-stepping in non-dissipative DGTD methods applied to wave propagation problems. *ESAIM : M2AN*, 40(5) :815–841.
- [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.
- [Reed and Hill, 1973] Reed, W. and Hill, T. (1973). Triangular mesh methods for the neutron transport equation. Technical Report LA-UR-73-479, Los Alamos National Laboratory, Los Alamos, New Mexico.
- [Reitich and Tamma, 2004] Reitich, F. and Tamma, K. K. (2004). State-of-the-art, trends and directions in Computational Electromagnetics. *CMES Comput. Model. Eng. Sci.*, 5 :287–294.
- [Remaki, 1999] Remaki, M. (1999). *Méthodes numériques pour les équations de Maxwell instationnaires en milieux hétérogènes*. PhD thesis, Ecole Nationale des Ponts et Chaussées.
- [Remaki, 2000] Remaki, M. (2000). A new finite volume scheme for solving Maxwell's system. *COMPEL*, 19 :913–931.
- [Remaki and Fezoui, 1998] Remaki, M. and Fezoui, L. (1998). Une méthode de Galerkin discontinu pour la résolution des équations de Maxwell en milieu hétérogène. Technical Report 3501, INRIA.
- [Richter, 1988] Richter, G. R. (1988). An optimal order error estimate for the discontinuous Galerkin method. *Math. Comp.*, 50 :75–88.
- [Rylander and Bondeson, 2002] Rylander, T. and Bondeson, A. (2002). Stability of explicit-implicit hybrid time-stepping schemes for Maxwell's equations. *J. Comput. Phys.*, 179 :426–438.
- [Saad, 1994] Saad, Y. (1994). ILUT : dual threshold incomplete ILU factorization for solving non-symmetric linear systems. *Numer. Lin. Alg. with Appl.*, 1 :387–402.
- [Saad, 2003] Saad, Y. (2003). *Iterative methods for sparse linear systems*. SIAM, 2nd edition.
- [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.
- [Sarmany et al., 2007] Sarmany, D., Botchev, M. A., and van der Vegt, J. (2007). Dispersion and dissipation error in high-order Runge-Kutta discontinuous Galerkin discretisations of the Maxwell equations. *J. Scient. Comput.*, 33(1).

- [Serre, 2000] Serre, D. (1999-2000). *Systems of conservation laws*, volume I and II. Cambridge University Press.
- [Shankar et al., 1989] Shankar, V., Hall, W., and Mohammadian, A. (1989). A time-domain differential solver for electromagnetic problems. In *IEE*, volume 77, pages 709–721.
- [Sheen et al., 1990] Sheen, D., Ali, S., Abouzahra, M., and Kong, J. (1990). Application of the three dimensional Finite-Difference Time-Domain method to the analysis of a planar microstrip circuits. *IEEE Trans. Microwave Theory Tech.*, 38 :849–856.
- [Shi et al., 2006] Shi, Y., Li, L., and Liang, C.-H. (2006). Two dimensional multidomain pseudospectral time-domain algorithm based on alternating-direction implicit method. *IEEE Trans. Antennas and Propagat.*, 54 :1207–1214.
- [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.
- [Spachmann et al., 2002] Spachmann, H., Schuhmann, R., and Weiland, T. (2002). Higher order explicit time integration schemes for Maxwell's equations. *Int. J. Numer. Model. : Electronic Networks, Devices and Fields*, 15(4) :419–437.
- [Sun and Trueman, 2003] Sun, C. and Trueman, C. (2003). Unconditionally stable Crank-Nicolson scheme for solving two-dimensional Maxwell's equations. *Electronics Letters*, 39(7) :595–597.
- [Whitney, 1957] Whitney, H. (1957). *Geometric integration Theory*. Princeton University Press, Princeton.
- [Yee, 1966] Yee, K. (1966). Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media. *IEEE Trans. Antennas and Propagat.*, AP-16 :302–307.
- [Young, 1999] Young, J. (1999). The design of high order, leap-frog integrators for Maxwell's equations. *IEEE Antennas and Propagation Society International Symposium*, 1 :176–179.
- [Zheng et al., 1999] Zheng, F., Chen, Z., and Zhang, J. (1999). A finite-difference time-domain method to develop without the Courant stability conditions. *IEEE Microwave and Guided Wave Letters*, 9 :441–443.
- [Zheng et al., 2000] Zheng, F., Chen, Z., and Zhang, J. (2000). Toward the development of a three-dimensional unconditionally stable finite-difference time-domain method. *IEEE Trans. Microwave Theory Technol.*, 48 :1550–1558.



**schémas d'intégration en temps efficaces pour la résolution numérique  
des équations de Maxwell instationnaires  
par des méthodes Galerkin discontinues d'ordre élevé en maillage non-structurés**

L'objectif général de cette étude est le développement et l'évaluation de schémas en temps efficaces pour des méthodes de type Galerkin discontinu (GD) en maillages tétraédriques non-structurés pour la résolution numérique des équations de Maxwell en domaine temporel. Dans la première partie de cette thèse, nous rappelons les équations de Maxwell et faisons une rapide revue des principales méthodes numériques utilisées pour résoudre ce système. Dans la seconde partie de cette thèse, nous présentons la méthode Galerkin discontinue basée sur des approximations centrées d'ordre générique. Dans ce chapitre, nous nous intéresserons qu'aux schémas en temps explicite. Nous détaillerons dans le troisième chapitre la partie principale de ce travail de thèse, c'est-à-dire les schémas implicites en temps, plus particulièrement, le schéma implicite, très étudié dans la littérature, de Crank-Nicolson et dans un second temps un schéma implicite d'ordre 4 obtenu à l'aide de la technique du défaut corrigé. Nous réalisons une étude comparative de deux solveurs (direct et itératif) pour la résolution du système linéaire au chapitre 4. Pour des questions d'espace mémoire, nous nous intéressons au chapitre 5 à appliquer le schéma implicite à un sous ensemble du domaine de calcul. Pour cela, nous utilisons un schéma hybride explicite/implicite. Au chapitre 6, nous présentons les résultats 3D obtenus avec cette méthode. Les problèmes considérés ont plusieurs millions d'inconnues.

**Mots clefs :** Equations de Maxwell - Eléments finis - Galerkin discontinu - Ordre élevé - Maillage tétraédrique - schéma implicite - schéma hybride en temps - Calcul parallèle.

**Discontinuous Galerkin methods for the numerical resolution  
of the 3D frequency domain Maxwell equations**

This general objective of this study is the development and assesment of efficient time integration scheme for Discontinuous Galerkin time domain (DGTD) method on unstructured tetrahedral meshes for numerical resolution of Maxwell equations. In first part of this thesis, we remind Maxwell's equations ans summarize main numerical methods used to solve this system. In the second part, we present the Discontinuous Galerkin method based on centered approximations for generic order. In this chapter, we focuse to time explicit scheme. We detailed, in third chapter, the main part of this work, in other words time implicit scheme, especially the Crank-Nicolson scheme, which is most studied in scientific litterature and in a second time a scheme of order 4 obtained by the defect correction technique. We realized a comparative study of both solvers (iterative and direct) to solve the linear system in chapter 4. For a memory space consideration , we apply the implicit scheme on a subdomain only. To do this, we use a hybrid explicit/implicit scheme. On chapter 6, we present the results 3D obtained with this method. Problems considered has several millions unknowns.

**Key words :** Maxwell equations - Finite elements - Discontinuous Galerkin - Tetrahedral mesh - Implicit scheme - hybrid scheme - High order - Parallel computing.