Contribution à la résolution des équations de Maxwell dans les structures périodiques par la méthode des éléments finis

Romain Garnier

Université de Toulouse

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par *l'Université Toulouse III - Paul Sabatier*
Discipline ou spécialité : *Micro-ondes électromagnétisme et optoélectroniques*

Présentée et soutenue par *Romain Garnier*
Le *30 janvier 2013*

Titre : *Contribution à la résolution des équations de Maxwell dans les structures périodiques par la méthode des éléments finis*

JURY

André Barka, Ingénieur de recherche, ONERA, Toulouse	*Directeur de thèse*
Olivier Pascal, professeur, UPS/Laplace, Toulouse	*Co-directeur de thèse*
Paul Soudais, Ingénieur, Dassault aviation, St Cloud	*Examinateur*
Xavier Begaud, professeur, TELECOM ParisTech, Paris	*Rapporteur*
Alain Bachelot, professeur, Université Bordeaux-1, Bordeaux	*Rapporteur*

Ecole doctorale : *Génie Electrique, Electronique, Télécommunications*
Unité de recherche : *ONERA département électromagnétique et radar*

Remerciements

Je tiens tout d'abord à adresser mes remerciements à mes directeurs de thèse qui ont imaginé et proposé ce sujet.

Merci à André Barka qui a su proposer des directions de recherche innovantes et qui a largement contribué à l'élaboration du code décrit dans ma thèse autant par son expertise pratique que par ses réflexions théoriques.

Merci à Olivier Pascal qui a contribué à la mise en forme et à la justesse des raisonnements que nous avons élaborés et utilisés.

Mes remerciements s'adressent ensuite à tous les membres du jury qui ont pris le temps de juger mon travail et sans qui, je n'aurais jamais pu finir cette thèse.

Merci à mes rapporteurs qui ont enrichi ma réflexion et mon travail de leurs remarques pertinentes.

Merci à Xavier Begaud pour son analyse éclairée sur l'utilisation et les applications possibles des outils développés.

Merci à Alain Bachelot de m'avoir recommandé pour cette thèse. J'ai eu le privilège de suivre ses excellents cours d'électromagnétisme à l'université de Bordeaux et j'ai bénéficié de son analyse élaborée sur la partie mathématique numérique.

Merci à Paul Soudais pour son apport sur les méthodes numériques existantes, notamment pour son analyse des équations intégrales.

Merci aussi à Ronan Perrussel du laboratoire Laplace et Xavier Ferrières de l'ONERA pour leur aide et leurs conseils sur les parties théoriques abordées dans les perspectives.

Je tiens également à remercier tous mes collègues de l'ONERA qui m'ont apporté des points de vue et des réflexions enrichissantes pour mon travail.

Mes remerciements s'adressent enfin à toute ma famille dont le soutien m'a été précieux.

"The greatest enemy of knowledge is not ignorance, it is the illusion of knowledge"

Stephen Hawking.

Résumé

En électromagnétisme les structures périodiques suscitent un grand intérêt. Ces structures agissent ainsi comme des filtres fréquentiels et permettent la fabrication de méta-matériaux, composites et artificiels. Elles présentent des propriétés électromagnétiques inédites pour les matériaux naturels telles que des bandes interdites. On a ainsi pu fabriquer de nouveaux dispositifs permettant de guider, de focaliser ou de stopper la propagation. C'est par exemple utile pour éviter le couplage entre différents éléments rayonnants notamment via la caractérisation des ondes de surface qui se propagent à l'interface entre l'air et la structure périodique. Ce travail de thèse s'inscrit dans ce contexte et propose une description de la méthode des éléments finis dédiée à la caractérisation des structures périodiques. La modélisation numérique aboutit à des problèmes de valeurs propres de grandes tailles. Elle implique la résolution de systèmes linéaires composés de matrices creuses. Une méthode est abordée pour résoudre ce type de problème, en optimisant et combinant différents algorithmes.
Avant d'aborder les différents aspects de la méthode développée, nous établissons une liste exhaustive de l'ensemble des méthodes qui existent en énonçant leurs avantages et leurs inconvénients. Nous constatons notamment que la méthode des éléments finis permet de traiter un large éventail de structures périodiques en trois dimensions sans limitation sur leur forme géométrique. Nous présentons alors les différentes formulations de cette méthode. Ensuite les aspects algorithmiques de la méthode sont détaillés. Nous montrons notamment qu'une analyse des paramètres de résolution permet de préciser les interprétations physiques des résultats obtenus. Finalement nous présentons les performances de notre outil sur des cas d'applications issus de la littérature et nous abordons la caractérisation des ondes de surface. Pour cela, l'étude d'un réseau d'antennes patchs insérées dans des cavités métalliques est conduite.
Notons pour conclure que les études conduites au cours de cette thèse ont abouti à la production d'un code utilisable dans un environnement de calcul initialement présent à l'ONERA.

Mots clés Structures périodiques, solveur mode propre, modes de surface, méthode des éléments finis, problèmes aux valeurs propres, bande interdite électromagnétique, matrices creuses, couplage mutuel.

Abstract

Electromagnetic periodic structures are of great interest. These structures act as frequency filters and allow the manufacturing of meta-materials which appear to be composite and artificial. They exhibit electromagnetic properties that are unusual to natural materials such as band gaps. This allows new devices to guide, focus or stop the propagation. This is for example useful to avoid coupling between various radiating elements via the characterization of the surface waves which propagate at the interface between air and the periodic structure. This thesis provides a description of the finite element method dedicated to the characterization of periodic structures. Numerical modelling results in eigenvalue problems of large sizes. It involves solving linear systems compounds of sparse matrices. A method is therefore discussed for solving this type of problem, optimizing and combining different algorithms. Before discussing the different aspects of the developed method, we establish an exhaustive list of all the existing methods by stating their advantages and drawbacks. We note that the finite element method can handle a wide range of periodic structures in three dimensions without limitation on their shape. We present different formulations of this method. Then the algorithmic aspects of the method are detailed. We show that an analysis of the resolution settings can impact the physical interpretations of the results. Finally we show the performance of our tool on classical validation results from the bibliography and we discuss the characterization of surface waves. Therefore, the study of a patch antenna array included in metal cavities is conducted. To conclude we can say that the studies conducted in this thesis have resulted in the production of a code used in an environment calculation initially present at ONERA.

Key words Periodic structures, eigenmode solver, surface modes, Finite Element Method, Eigenvalue problems, electromagnetic band gap, sparse matrices, mutual coupling.

ACRONYMES

EBG : Electromagnetic Band Gap (Bande interdite électromagnétique).

BIE : Bande Interdite Électromagnétique.

FSS : Frequency Selective Surface (Surface sélective en fréquence).

SER : Surface équivalente radar.

PEC : Perfect Electric Conductor (Électrique parfait conducteur).

PMC : Perfect Magnetic Conductor (Magnétique parfait conducteur).

PBC : Periodic Boundary Conditions (Conditions limites périodiques).

PML : Perfectly Matched Layer (Couche absorbante parfaite).

TEM : Transverse Electric and Magnetic.

PRS : Partial Reflector Surface (Surface partiellement réflective)

FDTD : Finite-Difference Time-Domain (Différence finie dans le domaine temporelle)

MDOA : Minimum Degree Ordering Algorithm (Algorithme du degré minimum d'ordonnancement)

YSM : Yale sparse matrix (Format de Yale des matrices creuses)

FMM : Fourier Modal Method.

RCWA : Rigorous Coupled-Wave Analysis (Analyse rigoureuse d'ondes couplées)

LU : Lower, Upper (descente, remonté).

IRAM : Implicitly Restarted Arnoldi Method (redémarrage implicite de la méthode d'Arnoldi).

GMRES : Generalized minimal residual method (méthode généralisée du résidu minimum).

MKL : Math Kernel Library http ://software.intel.com/en-us/intel-mkl/

NOTATIONS

λ : Longueur d'onde exprimée en mètre.

C : Vitesse de la lumière dans le vide exprimée en mètres par seconde.

k_0 : Nombre d'onde dans le vide en radians par mètre.

ω : Pulsation en radians par seconde.

ϵ_0 : Permittivité du vide.

ϵ_r : Permittivité relative du matériau.

ϵ : Permittivité effective du matériau.

μ_0 : Perméabilité du vide.

μ_r : Perméabilité relative du matériau.

μ : Perméabilité effective du matériau.

ρ : Densité de charges électriques.

σ_e : Conductivité électrique.

$\mathbf{J}$: Vecteur densité de courant.

$\mathbf{r}$: Vecteur position. $\mathbf{r} = (x, y, z) \in \mathbb{R}^3$.

$\mathbf{D}$: Vecteur d'induction électrique.

$\mathbf{B}$: Vecteur d'induction magnétique.

$\mathbf{E}$: Vecteur champ électrique.

$\mathbf{H}$: Vecteur champ magnétique.

$\mathcal{E}$: Vecteur amplitude du champ électrique $\mathbf{E}$.

$\boldsymbol{\nu}$: Normale extérieure.

$\boldsymbol{\nabla}$: Opérateur différentiel Nabla.

$\boldsymbol{\nabla}.$ ou div : Divergence.

$\times$: Produit vectoriel, ou produit cartésien.

$\boldsymbol{\nabla}\times$: Rotationnel.

$\triangle$: Opérateur Laplacien scalaire.

$\triangle$: Opérateur Laplacien vectoriel.

$\square$: Opérateur vectoriel de l'équation des ondes (défini équation (4.II) section II.1.1).

$\otimes$: Produit tensoriel.

m_e : Masse d'un électron.

F_ν : Fréquence de collision.

j : Nombre complexe, $j^2 = -1$.

DL : Degrés de liberté.

$\mathbf{W}_p$: Fonction de base de Nedelec d'ordre 0 d'indice p.

$\underline{\mathbf{W}}$: $(\mathbf{W}_1, \ldots, \mathbf{W}_{DL})$.

Ω : Domaine ouvert ($\bar{\Omega}$ fermé de Ω).

Ω_h : Triangulation de Ω.

$\mathcal{H}$: Espace solution.

$\mathcal{H}_h$: Espace de dimension finie dense dans $\mathcal{H}$.

h : $\max_{k\in\Omega_h} h_k$ où h_k est le diamètre du tétraèdre K.

$L^2(\Omega)$: Espace des fonctions de carré sommable dans Ω.

$H^1(\Omega)$: $\left\{u \in L^2(\Omega); \forall i = 1...n, \frac{\partial u}{\partial x_i} \in L^2(\Omega)\right\}$.

$(H^1(\Omega))^3$: Fonctions vectorielles à valeurs dans $\mathbb{R}^3(\Omega)$ tel que chaque composante appartient à $H^1(\Omega)$.

$H_0^1(\Omega)$: Sous-espace vectoriel de $H^1(\Omega)$ obtenu par la complétion des fonctions $C^\infty(\Omega)$ à support compact dans Ω.

$<,>$: Produit scalaire hermitien.

$\|.\|_\infty$: Norme infinie d'un vecteur $\mathbf{v} \in \mathbb{C}^n$: $\|\mathbf{v}\|_\infty = \max(|v_1|, \ldots, |v_n|)$.

H^* : Transposée conjuguée de H.

H^T : Transposée de H.

δ : Distribution de Dirac, aussi appelée par abus de langage fonction δ de Dirac.

$\mathbf{Ar}_p$: Arête de numéro p.

■ : Fin de démonstration.

LOGICIELS UTILISÉS

Gid : Mailleur avec interface graphique.

Tecplot : Outil de visualisation 3d, mis en œuvre pour la cartographie des champs.

Gnuplot : Traceur de courbe à partir de fichiers textes organisés.

Gimp : Dessin et conversion d'images.

Latex : Outil de traitement de texte.

Word : Outil de traitement de texte.

Power Point : Outil de traitement de texte pour les présentations.

PPTM : Outil de post traitement de Maillage [1].

Table des matières

. . . . . .

I Introduction générale

I.1 Contexte de l'étude

Nous allons, dans cette partie, introduire le contexte par le biais d'exemples d'applications connues. Auparavant, nous pouvons nous appuyer sur les travaux récents de Juslan Lo[2], effectués au laboratoire Laplace pour situer l'historique et le contexte propres au métamatériaux pour l'électromagnétisme. "Historiquement, nous trouvons la trace d'utilisation des structures périodiques ou répétitives pour contrôler la propagation d'une onde aussi loin qu'au XIXème siècle avec les travaux de Lord Rayleigh en 1887[3]. Lorsque nous alternons une succession de couches diélectriques transparentes d'indices de réfraction différents, une lumière incidente à la surface de ces multicouches peut se trouver complètement réfléchie par des phénomènes d'interférences successives. Ce phénomène peut s'interpréter de la manière suivante : sur l'interface de chaque couche, la lumière est partiellement réfléchie, si l'espacement entre chaque couche est périodique, les réflexions successives finissent par annuler la propagation de l'onde lumineuse au sein de ces couches. Le miroir de Bragg mis au point par William Lawrence Bragg est un parfait exemple de l'exploitation de ce phénomène en une dimension. Le concept du contrôle de l'onde par des structures périodiques comme nous venons de le décrire émerge donc du domaine photonique et l'on parlait alors de matériaux à bande interdite photonique (BIP), ou de cristaux photoniques. Le principe s'est ensuite rapidement étendu aux longueurs d'ondes supérieures, donnant lieu aux cristaux électromagnétiques, qui sont plus couramment appelés les matériaux à bande interdite électromagnétique (BIE). En effet, les équations de Maxwell obéissent à une loi d'échelle, et les mêmes propriétés peuvent par conséquent être observées quelle que soit la longueur d'onde [4]. En fait, ces matériaux BIE sont classés sous une famille de matériaux plus large, dénommés métamatériaux. Ces métamatériaux doivent leurs propriétés intéressantes plus à leur structuration interne qu'à leur composition chimique. Mises à part les structures BIE, les autres types de métamatériaux sont classés selon la propriété exotique qui est mise en avant [5] :

- Les matériaux dits à main gauche (LHM) sont des métamatériaux qui possèdent un indice de réfraction négatif où la loi de Snell-Descartes est inversée [6, 7].
- Les surfaces à haute impédance sont des métamatériaux qui annulent les ondes de surface car l'impédance de surface de la structure est plus importante que l'impédance en espace libre [8].
- Les métamatériaux phononiques traitent des structures périodiques pour les ondes acoustiques, qui plus récemment encore, trouvent une application pour les ondes sismiques dans les travaux de Farhat et al. [9]. Il s'agit alors de créer une "cape d'invisibilité" pour protéger une bâtisse contre les ondes sismiques. Pour ces différents métamatériaux, nous distinguons dans la plupart des cas le régime d'homogénéisation où la structuration est de dimension très inférieure à la longueur d'onde (régime "méta") et où au contraire, elle est du même ordre de grandeur que la longueur d'onde. Les structures BIE se classifient plutôt dans le dernier cas. Dans le domaine des micro-

ondes et des hyperfréquences, les propriétés exotiques d'une structure à bande interdite électromagnétique (l'anisotropie, la réfraction négative, la bande interdite) reçoivent des attentions particulièrement importantes pour les applications dans les dispositifs de télécommunications (antennes, téléphones mobiles, GPS haute précision etc[10])." Extrait de de la thèse de Juslan Lo "Étude de la reconfigurabilité d'une structure à bande interdite électromagnétique (BIE) métallique par plasmas de décharge" pages 6-7[2].

Nous allons maintenant énoncer des exemples d'applications qui illustrent certaines possibilités des structures périodiques.

I.2 État de l'art non exhaustif pour les structures périodiques

Cette partie traite d'abord du cas général des surfaces sélectives en fréquence (FSS) dans la section I.2.1. Ensuite des exemples d'applications pour la conception d'antennes sont abordés dans les sections I.2.2, I.2.3 et I.2.4 et la section I.2.5 aborde un cas original d'application pratique au blindage électromagnétique.

I.2.1 Surfaces sélectives en fréquence (FSS) [13]

À partir des années 1960, pour des applications militaires liées à la conception d'antennes, les surfaces sélectives en fréquence ont fait l'objet de diverses études [11, 12]. L'une des propriétés très utilisée par les FSS est le filtrage fréquentiel. Après passage de l'onde électromagnétique à travers la FSS, certaines fréquences sont transmises tandis que d'autres sont réfléchies. Les filtres sont utilisés pour atténuer les signaux indésirables et limiter les phénomènes d'interférence. Les FSS sont en général constituées d'éléments métalliques et diélectriques arrangés en réseaux périodiques planaires (cf Fig 1.I). La réaction d'une FSS face à une excitation est déterminée par la géométrie des motifs périodisés, par leurs propriétés de dispersion et par le pas du réseau. Les propriétés de dispersion des FSS sont généralement établies lorsque l'on considère des surfaces infinies.

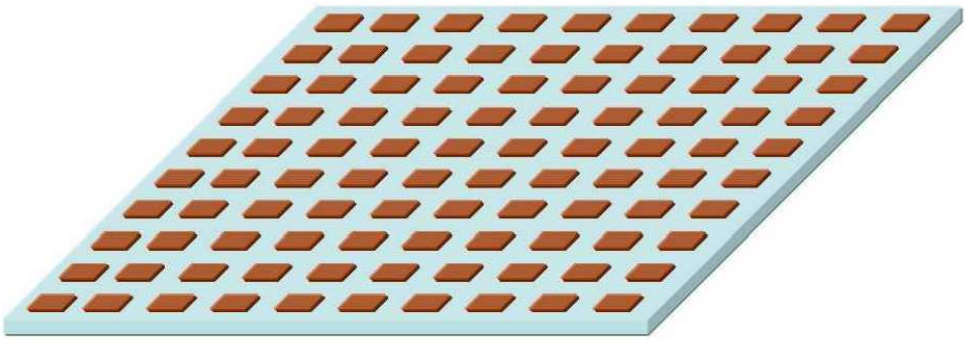


Fig 1.I Réseau bidimensionnel périodique de patchs métalliques ([13] page 5).

Si on excite la FSS par une onde plane de longueur d'onde λ, on peut configurer certains éléments pour qu'ils réfléchissent l'onde incidente avec un déphasage. Les phénomènes de résonance apparaissent quand la taille effective des éléments du réseau est un multiple de la longueur d'onde d'excitation [14, 15]. De plus, on sait que l'on peut agir sur ces déphasages par la modification des tailles et des formes des éléments. Ainsi, on peut par exemple arranger les éléments Fig 1.I pour que la somme des ondes réfléchies sur chacun des patchs forme une onde cohérente. C'est le principe du réflecteur de Marconi et Franklin[16]. Ils ont d'abord contribuer en 1919 à l'élaboration d'un réflecteur parabolique qui utilise des sections de fil demi-longueur d'onde.

En plus des différentes caractéristiques déjà énoncées, les propriétés de résonance d'une FSS dépendent de l'angle d'incidence. Trouver des solutions pour réduire cette dépendance

a constitué un enjeu majeur dans beaucoup d'applications qui utilisent les FSS notamment pour contrôler les caractéristiques d'une antenne.

Les FSS peuvent aussi permettre de rendre des objets non détectables par des radars. Elles peuvent notamment être insérées dans les Radômes qui sont principalement conçus pour protéger une antenne des intempéries. Dans ce cas, la FSS sera utilisée par exemple comme un filtre passe-bande qui réduit la surface équivalente radar (SER) d'une antenne en dehors de sa bande de fréquence de fonctionnement.

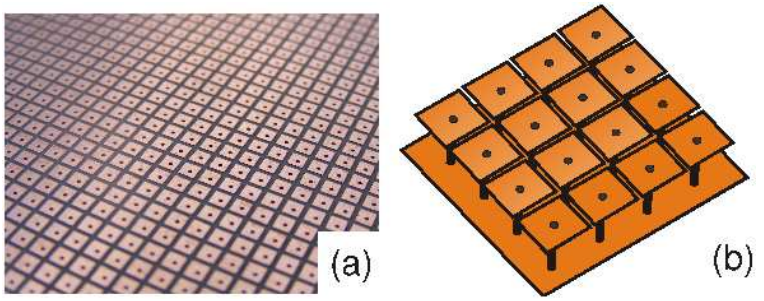
Fig 2.I Radômes rigides au Centre des Opérations de Cryptologie, Misawa, Japon [18].

Cette dernière application nous renvoie à la notion de furtivité qui doit permettre d'opérer à l'insu des radars d'un ennemi potentiel ce qui a toujours constitué un objectif de la technologie militaire (voir Fig 3.I). Dans la même idée que pour le radôme le but est de restreindre la capacité de détection, en recouvrant les objets à cacher par des FSS.

Fig 3.I Avion F-117 furtif[19].

I.2.2 Surfaces HIS

Les surfaces HIS[20] sont utilisées pour limiter les ondes de surfaces. En effet elles peuvent être insérées entre deux antennes afin de limiter les effets de couplage [21].



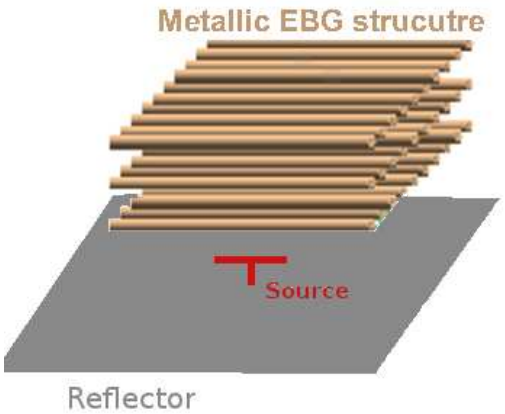
Taille des cellules = $\ll \lambda$

Fig 4.I Surface haute impédance (H.I.S)[22].

La taille des motifs périodisés est très petite devant la longueur d'onde d'étude de ces matériaux (de l'ordre de $\lambda/100$). Dans ce régime, la lumière ne voit pas la structuration, de telle sorte que l'on peut considérer le milieu comme un milieu effectif, avec un indice de réfraction moyen. Autrement dit, on homogénéise le milieu périodique. Mais attention, ce type de modélisation simplifie la réalité physique des surfaces à haute impédance qui peuvent entrer en résonance alors qu'elles sont éclairées par des longueurs d'ondes d'observation très grandes devant la taille des éléments périodisés [20].

I.2.3 Antennes EBG [23]

Les antennes EBG sont composées, comme on le voit sur le dessin Fig 5.I, d'un réflecteur métallique devant lequel est installé un élément rayonnant comme par exemple un dipôle. Au-dessus du dipôle on dispose un panneau partiellement réflecteur (PRS) sur lequel sont dessinés des motifs périodiques. L'avantage principal de cette antenne est qu'elle contribue à un accroissement de la directivité. L'inconvénient est que sa largeur de bande est étroite. Une des solutions pour augmenter la largeur de bande est d'augmenter le nombre de PRS [23].



Taille des cellules = $\simeq \frac{\lambda}{2}$

Fig 5.I Antenne EBG[24].

Pour ce type de dispositif la taille des motifs périodisés est de l'ordre de la longueur

d'onde d'étude. Ainsi on ne peut pas homogénéiser ce type de structure.

I.2.4 Réseaux de tiges diélectriques

Les réseaux périodiques de tiges diélectriques sont des structures à Bandes interdites électromagnétiques (BIE).

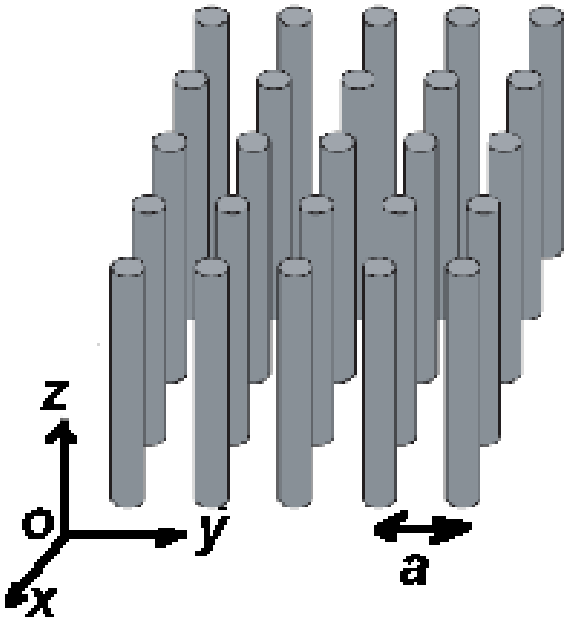


Fig 6.I Réseaux de tiges diélectriques de pas a.

La permittivité et le rayon des tiges vont influencer les propriétés de bande interdite qui s'avèrent particulièrement utiles pour piéger ou guider l'énergie, par le biais de cavités (défaut ponctuel) ou de guides d'ondes (défaut étendu, ou linéaire). On peut ainsi construire des dispositifs rayonnants en exploitant les modes de surface qui se propagent à l'interface entre l'air et les tiges de surface qui ont un rayon plus petit que le rayon des tiges du réseau [25, 26] :

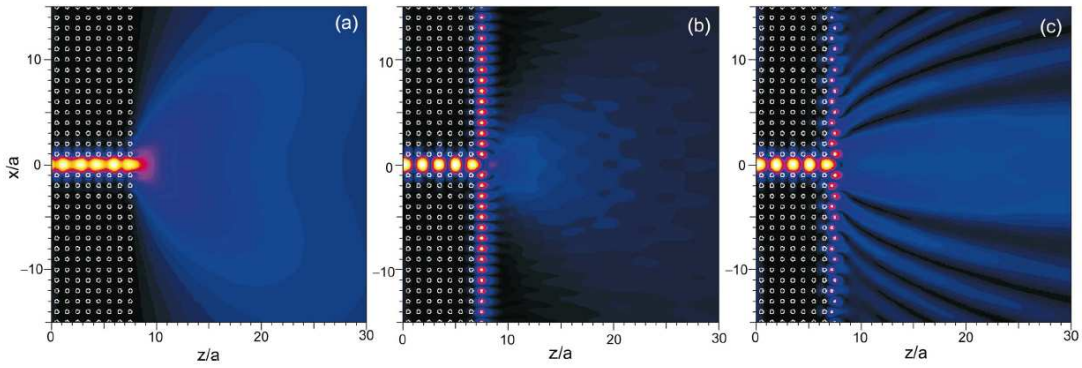


Fig 7.I Modification des tiges de surface [25].

Fig 7.I (a) un trou dans le réseau de tiges permet d'abord de guider l'énergie. Ensuite

Fig 7.I (b) les tiges de surface ont un rayon deux fois plus petit que les autres tiges, ce qui permet de faire apparaître de l'énergie à l'interface entre l'air et le dispositif. Fig 7.I (c) le décalage des tiges de rayon inférieur permet de rendre le mode de surface radiatif. Ce type de dispositif a fait l'objet d'une étude approfondie dans la thèse de Stéphane Varault [26]. Il permet aussi de créer des structures plus complexes telles que des coupleurs, des multiplexeurs, ou même des interféromètres. Évidemment un bon nombre de technologies ont déjà été réalisées pour accomplir ces fonctions de manière efficace. Cependant, l'utilisation des matériaux à bande interdite électromagnétique permet de réaliser plusieurs de ces fonctions avec un même support. Cela peut constituer un avantage dans le cadre de caractérisations expérimentales.

I.2.5 Isolant aux ondes Wifi

L'institut Polytechnique de Grenoble[27] a mis au point un papier peint qui filtre les ondes électromagnétiques des téléphones portables et systèmes wifi. Ce type de dispositif peut être utilisé pour limiter l'exposition aux ondes de certaines parties d'une habitation. Ce papier peint peut aussi permettre de sécuriser sa connexion wifi au piratage externe. Il est composé de motifs métalliques qui sont imprimés sur le papier peint. Les propriétés de filtrage sont dues à la taille et à la forme des motifs périodisés.

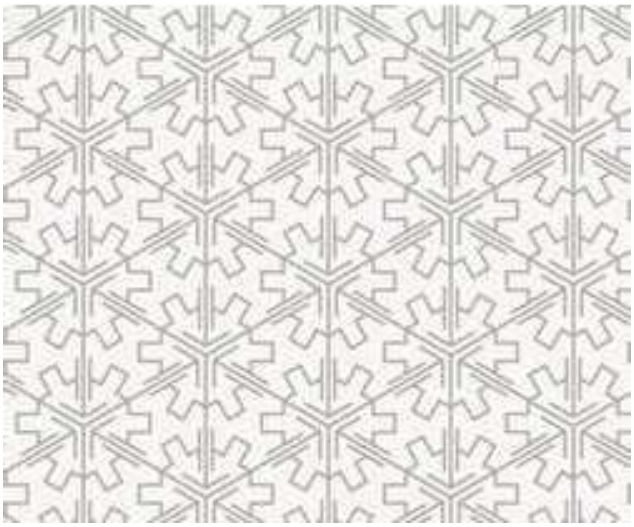


Fig 8.I Papier électromagnétiquement isolant composé de motifs périodiques[27].

Certains points restent à éclaircir au sujet de ce dispositif. Par exemple nous ne savons pas dans quelle mesure le milieu sur lequel est posé le papier influence ses propriétés. Si le mur est humide par exemple. Les tailles différentes des motifs imbriqués semblent nous indiquer qu'il y a plusieurs échelles correspondant sans doute à plusieurs fréquences absorbées. Malheureusement ces suppositions ne peuvent pas être vérifiées puisqu'il n'y a pas de publication scientifique à ce sujet à notre connaissance.

I.3 Problématique

Au cours de cette étude, nous allons nous intéresser à la mise en œuvre d'un modèle numérique dédié à la caractérisation des structures à bande interdite électromagnétique (BIE). Ces structures sont constituées d'un arrangement périodique de divers matériaux en trois dimensions. Leur propriété la plus utilisée est l'existence d'une bande de fréquence dans laquelle la structure est électromagnétiquement opaque. De plus, d'autres propriétés peu communes pour les matériaux plus classiques (anisotropie, indice de réfraction négatif) offrent la possibilité de contrôler la propagation des ondes électromagnétiques au sein d'une structure BIE de façon plus sophistiquée. On s'intéresse notamment à la propagation des ondes de surface à l'interface entre l'air et une structure périodique. Beaucoup d'outils de simulation existent mais ils peuvent comporter des facteurs limitant.

- La difficulté à traiter des structures anisotropes et à géométrie complexe.
- La difficulté à pouvoir interpréter les résultats obtenus à cause de l'inaccessibilité à des indicateurs qui garantissent la qualité de la résolution.
- La difficulté à traiter des structures dispersives.

On peut classer les méthodes existantes en deux familles :

- Les méthodes analytiques (cf section II.3) dont la modélisation découle principalement de considérations physiques.
- Les méthodes numériques (cf sections II.4 II.5 et III) dont la modélisation découle principalement de découpages temporels ou spatiaux.

Quelle que soit la méthode utilisée, la précision des résultats numériques est conditionnée par le nombre d'inconnues et l'organisation de l'espace mémoire. Ces facteurs auront bien sûr une influence sur le temps de calcul et la mémoire requise.

I.4 Objectifs

L'objectif de la thèse a été de développer un outil de résolution par éléments finis en maîtrisant tous les paramètres d'entrée et de sortie afin de caractériser la propagation modale dans une structure périodique infinie avec le moins de facteurs limitants possibles. On doit pouvoir ainsi traiter des structures périodiques infinies hétérogènes, anisotropes, sans contraintes géométriques. Notre outil doit aussi permettre dans un second temps de caractériser les solutions pouvant se propager à l'interface air-structure qu'on appelle les "modes de surface".

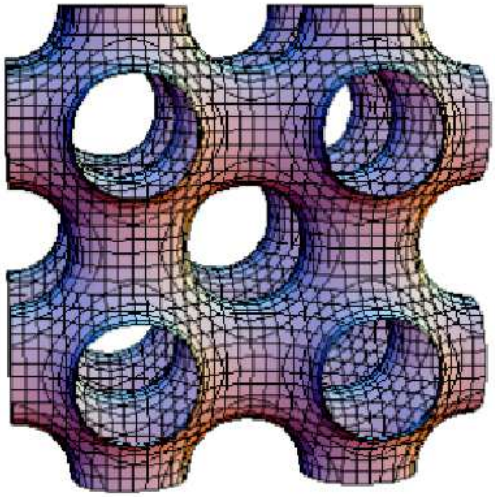


Fig 9.I Exemple d'une structure périodique à forme complexe[28].

Durant la thèse, nous nous sommes intéressés aux paramètres de résolution qui nous donnent des renseignements indispensables à l'interprétation des résultats. Nous avons envisagé la possibilité de traiter des matériaux périodiques dispersifs en injectant un modèle type Debye ou Drude[29, 30] dans la formulation. Nous avons aussi cherché à optimiser les algorithmes de stockage et de résolution des systèmes linéaires afin de minimiser le temps de calcul et l'espace mémoire utilisé. Les études conduites au cours de cette thèse ont abouti à la production d'un code utilisable dans l'environnement de calcul FACTOPO initialement développé à l'ONERA[31]. Pour atteindre les objectifs définis il a fallu modéliser correctement le problème éléments finis en développant un outil d'abord capable de calculer des modes propres dans des cavités métalliques[32]. Nous avons pu observer l'existence de valeurs propres nulles que nous avons éliminées grâce à la transformation *Shift and Invert* [33]. Parmi ces valeurs propres nulles il existe des solutions parasites définies dans la section IV.5.1 que nous avons éliminées dans les problèmes de cavités métalliques en développant un algorithme automatique décrit dans [34] et dans la section IV.5.2. Ensuite la possibilité de traiter les structures périodiques a été permise via l'introduction de conditions aux limites de type Floquet [35]. L'algorithme d'élimination automatique de solutions parasites utilisé initialement pour les problèmes de cavités métalliques n'est pas toujours possible quand on introduit ces conditions limites[36]. Nous nous sommes alors contentés d'éliminer toutes les

valeurs propres nulles en développant l'algorithme décrit dans la section IV.4. En parallèle avec cette approche nous avons optimisé le stockage des matrices creuses abordé dans la section IV.6 et la résolution des systèmes matriciels en combinant les librairies informatiques ARPACK[33] et PARDISO [37].

I.5 Plan de l'étude

Dans une première partie nous allons présenter le contexte de l'étude et les bases théoriques indispensables pour la compréhension du manuscrit. Dans une deuxième partie nous présentons les méthodes déjà existantes de caractérisation des structures EBG en précisant leurs atouts et leurs points faibles. Ensuite nous parlons des bases théoriques de la méthode des éléments finis en expliquant quelles transformations sont nécessaires pour traiter les structures périodiques. Dans une quatrième partie nous présentons les algorithmes de calcul et les réglages nécessaires pour résoudre le problème et interpréter les solutions. Dans une cinquième partie nous présentons des résultats de validation ainsi que de nouvelles interprétations possibles pour la caractérisation des modes de surface. Enfin nous concluons et donnons des perspectives à l'ensemble de ce travail.

I.6 Conclusion

Dans la section I.1 nous avons présenté les propriétés et les possibilités d'utilisation des structures EBG. Ensuite section I.2 nous avons présenté les applications principales qui justifient l'intérêt de la méthode que nous développons. Pour modéliser la propagation dans ces structures il est nécessaire de bien comprendre la théorie de Floquet que nous allons aborder dans le chapitre suivant. Cette théorie est complétée par des outils de représentation issus de l'étude des cristaux photoniques. À partir de la section II.3 nous allons lister les qualités et les défauts de toutes les méthodes existantes afin de les comparer entre elles et nous présentons la méthode que nous développons par la suite.

. . . . . .

II État de l'art des méthodes d'analyse des structures périodiques

II.1 Théorie de floquet

II.1.1 Introduction

La théorie de Floquet est la théorie essentielle utilisée pour la modélisation de la propagation dans une structure périodique dans une ou plusieurs directions. Dans la première partie II.1.2 nous présentons tout d'abord son principe en traitant le cas simple de la propagation dans des structures périodiques ayant leurs directions de périodicités confondues avec les axes du repère cartésien. Ensuite partie II.1.5 nous traitons le cas où une direction de périodicité n'est pas confondue avec un des axes du repère cartésien. Finalement, dans la section II.1.7 nous énonçons des considérations sur les cristaux photoniques afin d'introduire la notion de diagramme de bandes dans la section II.1.8. Toutes les grandeurs sont définies dans $\mathbb{R}^3 \times \mathbb{R}^t$. On note par $\mathbf{E}(\mathbf{r}, t)$ et $\mathbf{H}(\mathbf{r}, t)$ les champs solutions à valeurs dans $\mathbb{C}^3$. On suppose par la suite que les champs sont des solutions physiquement acceptables des équations de Maxwell. De tels champs doivent posséder un énergie électromagnétique localement bornée, ce qui entraîne qu'ils doivent être de carré localement sommable. On se place dans un domaine infiniment périodique et ouvert de $\mathbb{R}^3$ noté Ω. Dans Ω, les équations de Maxwell peuvent alors s'écrire[38] :

$$\begin{cases} \text{Maxwell Faraday} \quad \boldsymbol{\nabla} \times \mathbf{E} + \frac{\partial}{\partial t}(\mu \mathbf{H}) = \mathbf{0} \\ \text{Maxwell Gauss} \quad \nabla.(\epsilon \mathbf{E}) = \rho, \ \rho \in \mathbb{R} \\ \text{Conservation du flux magnétique} \quad \nabla.(\mu \mathbf{H}) = 0 \\ \text{Maxwell Ampère} \quad \boldsymbol{\nabla} \times \mathbf{H} - \frac{\partial}{\partial t}(\epsilon \mathbf{E}) - \mathbf{J} = \mathbf{0}, \ \mathbf{J} \in \mathbb{C}^3 \end{cases}$$

Dans le cas général, Ω est un domaine hétérogène et linéaire comportant des matériaux à pertes et ϵ, μ sont des matrices hermitiennes. Des définitions plus précises de ϵ et μ sont données en annexe A.2. Dans la suite, on pose $\epsilon^{-1} = \frac{1}{\epsilon}$ (respectivement $\mu^{-1} = \frac{1}{\mu}$), parce que dans la majorité des cas traités, ϵ et μ sont des nombres complexes. Si on veut garder le plus de généralité possible, il suffit de remplacer $\frac{1}{\epsilon}$ (respectivement $\frac{1}{\mu}$) par la matrice $[\epsilon]^{-1}$ (respectivement $[\mu]^{-1}$) dans tout ce qui suit. On appelle $\Omega_{obs} \in \Omega$ la région de l'espace de laquelle est observée la périodicité. On suppose Ω_{obs} est un ouvert de Ω qui ne comporte pas de charge ni de courant donc $\rho = 0$ et $\mathbf{J} = \mathbf{0}$. On suppose aussi que dans Ω_{obs}, $\epsilon_r = \epsilon/\epsilon_0$ et $\mu_r = \mu/\mu_0$ sont constants et réels. On peut alors écrire :

$$\boldsymbol{\nabla} \times \mathbf{H}(\mathbf{r}, t) - \epsilon_0 \epsilon_r \frac{\partial}{\partial t} \mathbf{E}(\mathbf{r}, t) = \mathbf{0}, \quad \mathbf{r} \in \Omega_{obs}. \tag{1.II}$$

$$\frac{\boldsymbol{\nabla} \times \mathbf{E}(\mathbf{r}, t)}{\mu_r} + \mu_0 \frac{\partial}{\partial t} \mathbf{H}(\mathbf{r}, t) = \mathbf{0}, \quad \mathbf{r} \in \Omega_{obs}. \tag{2.II}$$

Les équations (1.II) et (2.II) sont à prendre au sens des distributions [39]. Autrement dit, elles incluent les conditions aux limites sur les interfaces entre les matériaux, en l'occurence la continuité des composantes tangentielles du champ électrique (cf annexe A.1). En combinant (1.II) et (2.II) et en posant $C = \frac{1}{\sqrt{\epsilon_0 \mu_0}}$ on obtient alors :

$$\triangle \mathbf{E} + \frac{\epsilon_r \mu_r}{C^2} \frac{\partial^2}{\partial^2 t} \mathbf{E} = \mathbf{0} \Rightarrow \square \mathbf{E} = \mathbf{0}, \quad \mathbf{r} \in \Omega_{obs} \tag{3.II}$$

Où $\square$ est l'opérateur défini par :

$$\square = \triangle + \frac{\epsilon_r \mu_r}{C^2} \frac{\partial^2}{\partial^2 t} \tag{4.II}$$

Pour une amplitude complexe $\boldsymbol{\mathcal{E}}_i donne$, on définit le champ d'une onde plane incidente :

$$\mathbf{E}_i(x, y, z, t) = \boldsymbol{\mathcal{E}}_i e^{j(-\omega t - \mathbf{k}_i . \mathbf{r})}, \ \mathbf{k}_i = (k_{ix}, k_{iy}, k_{iz}), \ \mathbf{r} = (x, y, z). \tag{5.II}$$

$\mathbf{E}_i(x, y, z, t)$ est une solution triviale de (3.II) lorsque $\|\mathbf{k}_i\|^2 = \frac{\omega^2}{C^2} \epsilon_r \mu_r$.

II.1.2 Périodicité selon un axe

On veut par exemple déterminer les solutions de propagations le long d'une structure périodique selon l'axe (Ox) de période d. Prenons le cas d'un réseau de fentes métalliques. On considère une onde plane [40] incidente dans le plan du vecteur d'ondes :

$$\mathbf{k}_i = (k_i \sin(\theta), -k_i \cos(\theta), 0)$$

On suppose que la projection d'une onde plane dans le plan $z = 0$ fait un angle θ avec la normale et génère une ou plusieurs ondes. Comme dans la relation (5.II) la composante du champ électrique $(E_i)_z$ est donnée par :

$$(E_i)_z(x, y, z, t) = \mathcal{E}_{iz} e^{j(-\omega t - \vec{k}_i . \vec{r})} = \mathcal{E}_i e^{-j\omega t} . e^{j(k_i y . cos(\theta) - k_i x . sin(\theta))}.$$

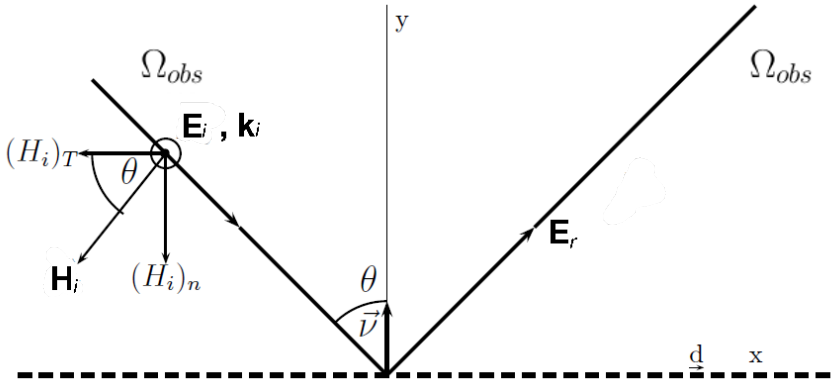


Fig 1.II Réseau périodique de fentes métalliques selon l'axe Ox.

On suppose que le domaine d'observation Ω_{obs} est le demi espace supérieur contenu dans

l'air. On peut alors écrire [41] de façon abusive que la solution représentant le champ total $\mathbf{E}(x,y,z,t) = \mathbf{E}_i(x,y,z,t) + \mathbf{E}_r(x,y,z,t)$ (à condition qu'elle existe) vérifie :

$$\mathbf{E}(x,y,z,t) = \Box^{-1}(\mathbf{E}_i(x,y,z,t)), \quad (x,y,z) = \mathbf{r} \in \Omega_{obs}$$

Ou l'opérateur $\Box$ est l'opérateur vectoriel de l'équation des ondes défini par (4.II). Autrement dit la solution $\mathbf{E}(x,y,z,t)$ du problème électromagnétique dans Ω_{obs} est l'image réciproque de l'onde incidente $\mathbf{E}_i(x,y,z,t)$ par l'opérateur $\Box$. Comme cet opérateur est linéaire et invariant par translation de longueur d selon (Ox) on peut écrire :

$$\mathbf{E}(x+d,y,zt) = \Box^{-1}(\mathbf{E}_i(x+d,y,z,t)) = e^{-k_i d \sin(\theta)} \Box^{-1}(\mathbf{E}_i(x,y,z,t))$$

Ceci revient en fait à dire que l'opérateur des ondes est invariant par translation mais la condition initiale change puisque l'onde incidente est décalée. En définissant $\mathbf{E}_0(x,y,z,t) = \mathbf{E}(x,y,z,t)e^{jk_i.x\sin(\theta)}$ on peut facilement vérifier que $\mathbf{E}_0$ est périodique en x de période d. On peut donc écrire $\mathbf{E}_0$ sous forme d'une série de Fourier :

$$\mathbf{E}_0(x,y,z,t) = \sum_{n=-\infty}^{+\infty} \hat{\mathbf{E}}_n(y,z,t) e^{j\left(\frac{2n\pi}{d}\right)x}$$

avec

$$\hat{\mathbf{E}}_n(y,z,t) = \frac{1}{d}\int_0^d \mathbf{E}_0(x,y,z,t) e^{j\left(-\frac{2n\pi}{d}\right)x} dx$$

Et donc

$$\mathbf{E}(x,y,z,t) = \sum_{n=-\infty}^{+\infty} \hat{\mathbf{E}}_n(y,z,t) e^{j(G_n - k_i\sin(\theta))x} \text{ avec } G_n = \frac{2n\pi}{d}(x,y,z) = \mathbf{r} \in \Omega_{obs}. \qquad (6.\text{II})$$

Remarque II.1. Soit f une fonction intégrable à valeur dans $\mathbb{R}$.
Dans la suite de la thèse, toutes les fonctions sont intégrables et si la variable temps n'est pas précisée, c'est que l'on considère les transformées de Fourier des grandeurs étudiées.
La transformée de Fourier que nous utilisons pour passer du domaine temporel au domaine fréquentiel s'écrit :

$$\mathcal{F}(f) : \omega \mapsto \hat{f}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(t)\, \mathrm{e}^{-j\omega t}\, \mathrm{d}t.$$

Et par conséquent :

$$\frac{\partial}{\partial t}\hat{f}(\omega) = -j\omega \hat{f}(\omega)$$

II.1.3 Exemple d'une structure bi-périodique simple

On peut étendre ce résultat à une structure périodique dans deux directions et invariante dans la direction z. Comme le problème est périodique l'étude se limite à une cellule de base.

Tous les points $\tilde{\mathbf{P}}$ dans le domaine périodique infini ont un point $\mathbf{P}$ correspondant dans le domaine rouge et si l'on connaît la valeur du champ $\underline{\mathbf{E}}$ dans ce domaine, toutes les solutions dans le domaine périodique peuvent être reproduites en utilisant les conditions de Floquet.

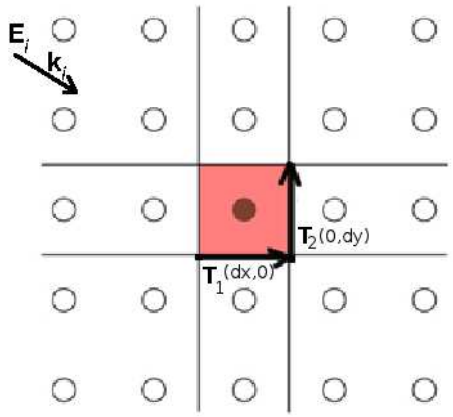


Fig 2.II En rouge cellule unitaire d'une structure périodique dans deux directions.
$\mathbf{T}_1 = (dx, 0),\ \mathbf{T}_2 = (0, dy),\ \mathbf{k}_i = (k_{ix}, k_{iy}, 0)$.

Plus précisément :

$$\exists (m,n) \in \mathbb{Z}^2,\ \tilde{\mathbf{P}} = \mathbf{P} + m\mathbf{T}_1 + n\mathbf{T}_2$$
$$\Rightarrow \mathbf{E}(\tilde{\mathbf{P}}) = \underline{\mathbf{E}}(\mathbf{P}) e^{-j\left(m\mathbf{T}_1 + n\mathbf{T}_2\right).\mathbf{k}_i}. \tag{7.II}$$

La fonction $\mathbf{E}_0 = e^{+j\left(\mathbf{k}_i.\mathbf{r}\right)}\mathbf{E}$ est périodique en x et y et le champs $\mathbf{E}$ s'écrit comme une série pseudo-périodique [41] :

$$\mathbf{E}(x,y,z) = \sum_{n=-\infty}^{+\infty} \sum_{m=-\infty}^{+\infty} \hat{\mathbf{E}}_{nm}(z) e^{j\left((\frac{2\pi n}{d_x} - k_{ix})x + (\frac{2\pi m}{d_y} - k_{iy})y\right)} \tag{8.II}$$

$$\hat{\mathbf{E}}_{nm}(z) = \frac{1}{d_x d_y} \int_0^{d_x} \int_0^{d_y} \mathbf{E}_0(x,y,z) e^{-j\left((\frac{2\pi n}{d_x})x + (\frac{2\pi m}{d_y})y\right)} dxdy$$

Définition. Dans la suite du document, on appelle "cellule de base" la cellule de référence sur laquelle tous les calculs sont effectués. Grâce aux relations de déphasage (7.II), les valeurs des champs obtenus sur la cellule de base peuvent être étendus sur l'ensemble de la structure périodique infinie.

II.1.4 Floquet 3d

On peut étendre ce résultat à la périodicité dans les trois directions de l'espace. En effet les systèmes auxquels nous nous intéressons sont périodiques dans deux voire trois directions. Supposons que le domaine est périodique dans trois directions d'espace (x,y,z) de périodes respectives (d_1, d_2, d_3). En considérant une onde incidente de vecteur d'onde $\mathbf{k}_i = (k_{ix}, k_{iy}, k_{iz})$

et le champ total solution $\mathbf{E}$, la fonction $\mathbf{E}_0(x,y,z)$ définie par

$$\mathbf{E}_0(x,y,z) = \mathbf{E}(x,y,z)e^{jk_{ix}x}.e^{jk_{iy}y}.e^{jk_{iz}z}$$

est développable en série de Fourier 3D soit :

$$\mathbf{E}_0(x,y,z) = \sum_{n=-\infty}^{+\infty}\sum_{m=-\infty}^{+\infty}\sum_{p=-\infty}^{+\infty}\hat{\mathbf{E}}_{nmp}e^{j\mathbf{G}_{nmp}.\mathbf{r}}, \quad \mathbf{G}_{nmp} = (\frac{2\pi n}{d_1}, \frac{2\pi m}{d_2}, \frac{2\pi p}{d_3}), \ \mathbf{r} = (x,y,z)$$

$$\hat{\mathbf{E}}_{nmp} = \frac{1}{d_1 d_2 d_3}\int_0^{d_1}\int_0^{d_2}\int_0^{d_3}\mathbf{E}_0(x,y,z)e^{-j\mathbf{G}_{nmp}.\mathbf{r}}dxdydz$$

Donc le champ total solution $\mathbf{E}$ est développable en série pseudo-périodique :

$$\mathbf{E}(x,y,z) = \sum_{n=-\infty}^{+\infty}\sum_{m=-\infty}^{+\infty}\sum_{p=-\infty}^{+\infty}\hat{\mathbf{E}}_{nmp}e^{j(\mathbf{G}_{nmp}-\mathbf{k}_i).\mathbf{r}} \tag{9.II}$$

II.1.5 Périodicité selon une droite vectorielle

On peut étendre le cas à des périodicités inclinées selon une ou plusieurs droites vectorielles qui ne sont pas alignées avec les axes du repère cartésien. Soit la droite vectorielle

$$\Delta = Vect\left(\sin(\theta)\cos(\phi), \sin(\theta)\sin(\phi), \cos(\theta)\right) = Vect\left(\tan(\theta)\cos(\phi), \tan(\theta)\sin(\phi), 1\right)$$

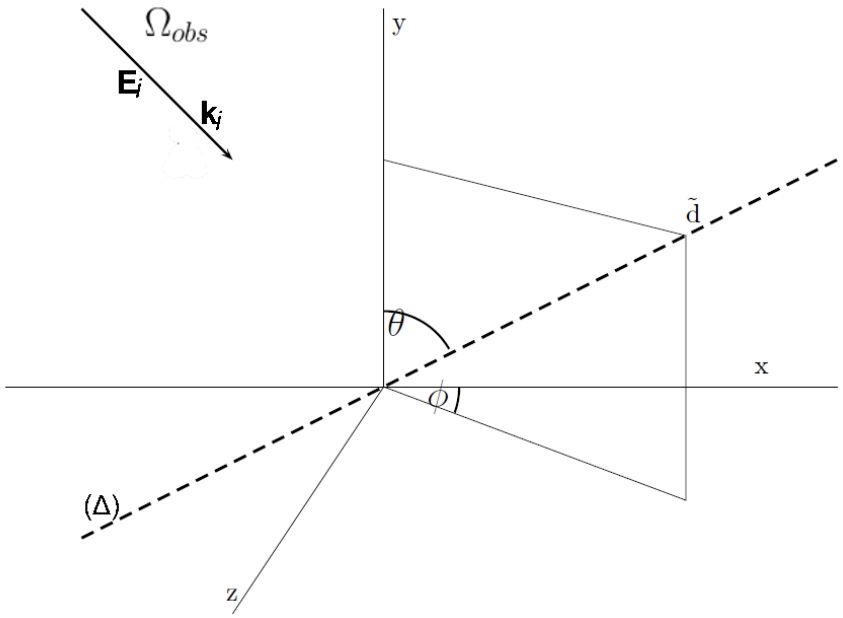


Fig 3.II Périodicité selon une droite vectorielle.

Comme dans la section précédente II.1.2 l'opérateur des ondes ($\square$) est linéaire et invariant par translation de longueur $\tilde{d}$ dans Ω_{obs} selon la droite vectorielle et on peut écrire en notant

$d = \tilde{d}\cos\theta$:

$$\mathbf{E}\left(x + n\tilde{d}\sin(\theta)\cos(\phi), y + n\tilde{d}.\sin(\theta)\sin(\phi), z + n\tilde{d}\cos(\theta)\right) =$$

$$\mathbf{E}\left(x + nd\tan(\theta)\cos(\phi), y + nd.\tan(\theta)\sin(\phi), z + nd\right) =$$

$$e^{-j(\mathbf{k}_i.(nd\tan(\theta)\cos(\phi), nd.\tan(\theta)\sin(\phi), nd))}\square^{-1}\left(\mathbf{E}_i(x,y,z)\right) =$$

$$\mathbf{E}(x,y,z,t)e^{-j(\mathbf{k}_i.(nd\tan(\theta)\cos(\phi), nd.\tan(\theta)\sin(\phi), nd))}, \quad (x,y,z) = \mathbf{r} \in \Omega_{obs}$$

Donc en posant :

$$\mathbf{E}_0(x,y,z) = \mathbf{E}(x,y,z)e^{j(\mathbf{k}_i.\vec{r})}$$

Alors $\mathbf{E}_0$ vérifie :

$$\mathbf{E}_0\left(x + n\tilde{d}\sin(\theta)\cos(\phi), y + n\tilde{d}.\sin(\theta)\sin(\phi), z + n\tilde{d}\cos(\theta)\right) =$$

$$\mathbf{E}_0\left(x + nd\tan(\theta)\cos(\phi), y + nd.\tan(\theta)\sin(\phi), z + nd\right) = \mathbf{E}_0(x,y,z)$$

II.1.5.1 Développement en série de Fourier

Avec un exemple de la littérature[42] on va expliciter les conditions de Floquet grâce au développement en série de Fourier selon deux axes non perpendiculaires.

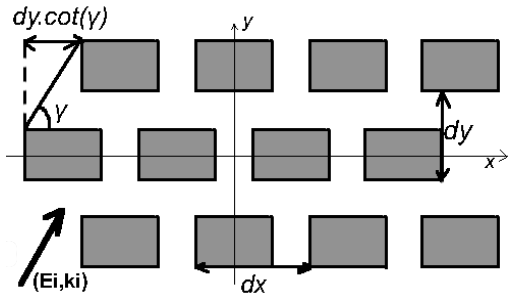


Fig 4.II Cellules périodiques le long de deux axes non perpendiculaires.

Fig 4.II on constate clairement qu'une onde qui arrive sous n'importe qu'elle incidence translatée de

$$(m.dx \pm n.\cot(\gamma)dy)\,\mathbf{e}_x + n.dy\mathbf{e}_y,$$

voit un problème identique au problème non translaté. On a alors :

$$\mathbf{E}(x + n.dy\cot(\gamma) + mdx, y + n.dy, z) = \mathbf{E}(x,y,z)e^{-j(\mathbf{k}_i.(n.dy\cot(\gamma)+mdx, n.dy, 0))}$$

La fonction $\mathbf{E}_0$ définie par $\mathbf{E}_0(x,y,z) = \mathbf{E}(x,y,z)e^{+j(k_{ix}x+k_{iy}y)}$ est périodique dans la direction Ox et la direction $y = \tan(\gamma x)$. $\mathbf{E}_0$ est donc développable en série de Fourier sur la surface Γ_r ($z = h$ fixé dans la cellule de base) selon ces deux directions et sa restriction à l'ensemble Γ_r peut s'écrire :

$$\mathbf{E}_0(x,y,h)\mathbf{1}_{\Gamma_r} = \mathbf{F}(x-\cot(\gamma)y,y)\mathbf{1}_{\Gamma_r},\quad \mathbf{E}_0(x+\cot(\gamma)y,y,h)\mathbf{1}_{\Gamma_r} = \mathbf{F}(x,y)\mathbf{1}_{\Gamma_r}$$

La fonction $\mathbf{F}$ est périodique en x et en y de période dx et dy.
En posant $\alpha = \cot(\gamma),\ u = \frac{2\pi.m}{dx},\ v = \frac{2\pi.n}{dy},\ (n,m) \in \mathbb{Z}^2$ sa transformée de Fourier est :

$$\hat{\mathbf{F}}(u,v) = \frac{1}{dxdy}\int_0^{dy}\int_0^{dx} \mathbf{E}_0(x+\alpha y,y,h)e^{-j(ux+vy)}dxdy$$

En effectuant d'abord l'intégration par rapport à x

$$\hat{\mathbf{F}}(u,v) = \frac{1}{dxdy}\int_0^{dy} e^{-jvy}\left[\int_0^{dx} \mathbf{E}_0(x+\alpha y,y,h)e^{-jux}dx\right]dy$$

On peut faire le changement de variable $x+\alpha y = s$

$$\hat{\mathbf{F}}(u,v) = \frac{1}{dxdy}\int_0^{dy} e^{-jvy}\left[\int_{\alpha y}^{dx+\alpha y} \mathbf{E}_0(s,y,h)e^{-jus+ju\alpha y}ds\right]dy$$

L'intégrale en fonction de la variable s est la transformée de Fourier d'une ligne horizontale de l'image $\mathbf{E}_0(s,y,h)$. On lui fait subir un déphasage $e^{ju\alpha y}$ correspondant à une translation de longueur α.

$$\hat{\mathbf{F}}(u,v) = \frac{1}{dxdy}\int_0^{dy} e^{-jvy+ju\alpha y}\left[\int_{\alpha y}^{dx+\alpha y} \mathbf{E}_0(s,y,h)e^{-jus}ds\right]dy$$

$\hat{\mathbf{F}}(u,v)$ est la transformée de Fourier de $\mathbf{E}_0(s,y,h)$ calculée pour u et $(v-\alpha m)$

$$\hat{\mathbf{F}}(u,v) = \hat{\mathbf{E}}_0(u,v-u\alpha,h)$$

$$\mathbf{F}(x,y) = \sum_{n=-\infty}^{+\infty}\sum_{m=-\infty}^{+\infty} \hat{\mathbf{F}}_{nm}(h)e^{j\left(\frac{2\pi.m}{dx}x+\frac{2\pi.n}{dy}y\right)}$$

avec

$$\hat{\mathbf{F}}_{nm}(h) = \frac{1}{dxdy}\int_0^{dy} e^{-j\frac{2\pi.n}{dy}y+j\frac{2\pi.m}{dx}\alpha y}\left[\int_{\alpha y}^{dx+\alpha y} \mathbf{E}_0(s,y,h)e^{-j\frac{2\pi.m}{dx}s}ds\right]dy$$

On en déduit l'expression du champs $\mathbf{E}$ sur Γ_r

$$\mathbf{E}(x,y,h) = \mathbf{F}(x-\cot(\gamma)y,y)e^{-j(\mathbf{k}_i.(x,y,0))} = \sum_{n=-\infty}^{+\infty}\sum_{m=-\infty}^{+\infty} \hat{\mathbf{F}}_{nm}(h)e^{j\left((\frac{2\pi.m}{dx}-k_{ix})x+(\frac{2\pi.n}{dy}-\frac{2\pi.m\cot(\gamma)}{dx}-k_{iy})y\right)}. \tag{10.II}$$

Plus généralement si une fonction est périodique selon trois axes non alignés avec les axes

du repère cartésien. Le champ $\mathbf{E}$ peut s'écrire :

$$\mathbf{E}(x,y,z) = \sum_{p=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} \sum_{m=-\infty}^{+\infty} \hat{\mathbf{E}}_{nmp} e^{j\left((G_{x_{nmp}}-k_{ix})x+(G_{y_{nmp}}-k_{iy})y+(G_{z_{nmp}}-k_{iz})z\right)} \tag{11.II}$$

Où les valeurs de $(G_{x_{nmp}}, G_{y_{nmp}}, G_{z_{nmp}})$ dépendent des angles entre les axes et des distances à franchir dans les directions de périodicité pour passer d'une maille à une autre. Plus précisément les vecteurs $(G_{x_{nmp}}, G_{y_{nmp}}, G_{z_{nmp}})$, $(n,m,p) \in \mathbb{Z}^3$ sont combinaisons linéaires des vecteurs de base du réseau réciproque de Bravais[43] que nous allons définir sections II.1.7.2 et II.1.7.3.

II.1.6 Conclusion

Dans cette partie nous avons présenté quelles sont les conséquences de la théorie de Floquet sur la modélisation des structures périodiques. Nous allons maintenant introduire le vocabulaire classique initialement utilisé pour traiter les cristaux photoniques qui sont des structures périodiques particulières. Ce vocabulaire aura une utilité pour la compréhension précise du diagramme de bandes qui est défini dans la section II.1.8.

II.1.7 Considérations sur les cristaux photoniques

II.1.7.1 Introduction

Les définitions utilisées pour les cristaux photoniques sont transposables à toutes les structures périodiques c'est pourquoi nous allons brièvement énoncer ces définitions et expliquer comment nous allons les utiliser.

II.1.7.2 Réseau de Bravais et maille primitive

Dans le domaine de la cristallographie, on définit un réseau de Bravais [43] comme une distribution périodique de nœuds dans l'espace. Les nœuds sont des sommets de la maille primitive qui est le motif géométrique le plus petit permettant de reconstruire le réseau cristallin infini par translations successives dans les directions de périodicité.

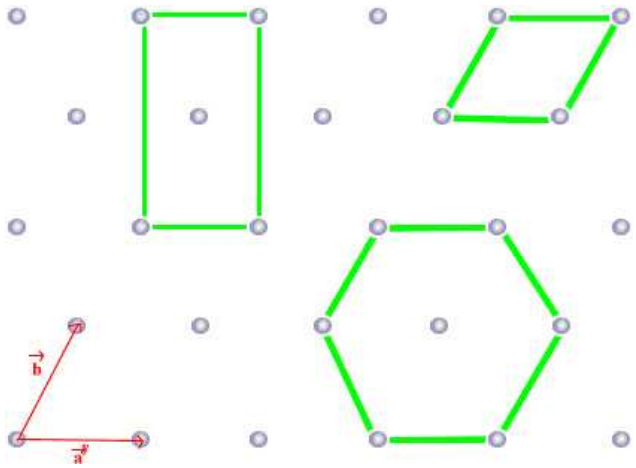


Fig 5.II $(\vec{a}, \vec{b})$ vecteurs de bases du réseau de Bravais[44] d'un réseau de tiges décalées.

II.1.7.3 Réseau réciproque

Le réseau réciproque d'un réseau de Bravais est défini par l'ensemble des vecteurs $\mathbf{G}$ tels que $e^{j\mathbf{G}\cdot\mathbf{R}} = 1$ pour tous les vecteurs position $\mathbf{R}$ du réseau de Bravais. Le réseau réciproque est un réseau de Bravais dont le réseau réciproque est le réseau de Bravais de départ.

Cette définition est directement liée à la périodicité des éléments du réseau de Bravais. En effet soit $\rho(\mathbf{r})$ une grandeur invariante par translation d'un vecteur $\mathbf{R}$ du réseau de Bravais on a :

$$\rho(\mathbf{r} + \mathbf{R}) = \rho(\mathbf{r}).$$

Cette fonction périodique peut être décomposée en série de Fourier, soit

$$\rho(\mathbf{r}) = \sum_G \rho_G e^{(j\mathbf{G}.\mathbf{r})} \text{ avec } \rho_G = \frac{1}{|Scell|} \int_{Scell} \rho(\mathbf{r}) e^{(j\mathbf{G}.\mathbf{r})} dr$$

Où $|Scell|$ est la surface de la cellule élémentaire. L'espace réciproque est défini par les vecteurs du réseau réciproque.

II.1.7.4 Maille de Wigner-Seitz

Une maille primitive peut aussi être obtenue en traçant les lignes qui relient un nœud donné à ses voisins, puis les plans médians de ces segments. La zone délimitée de cette manière s'appelle la maille primitive de Wigner-Seitz.

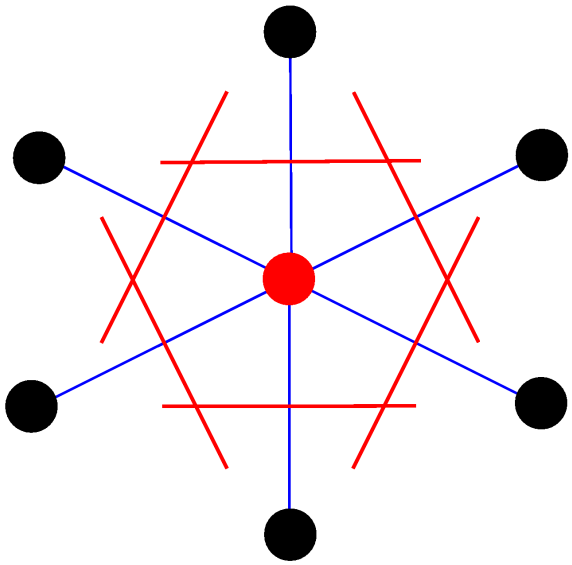


Fig 6.II Maille de Wigner-Seitz[45] d'un réseau de tiges décalées.

La première zone de Brillouin est définie de manière unique par la méthode permettant de construire la maille de Wigner-Seitz. La zone de Brillouin peut être encore réduite pour des raisons de symétrie en zone de Brillouin irréductible.

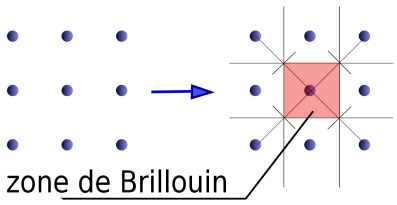


Fig 7.II Zone de Brillouin[46].

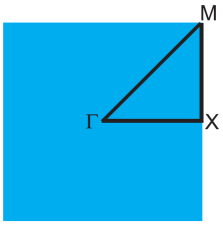


Fig 8.II ΓXM : Zone de Brillouin irréductible correspondante.

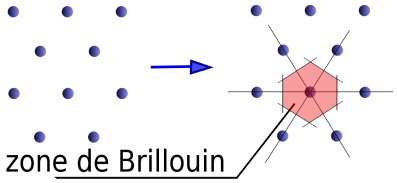


Fig 9.II Zone de Brillouin.

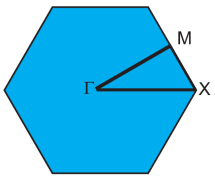


Fig 10.II ΓXM : Zone de Brillouin irréductible correspondante.

Dans nos calculs nous ferons en sorte que la cellule de base de la structure périodique soit incluse dans un rectangle ou un parallélépipède (ou sur une ligne pour un réseau 1d). Par exemple pour traiter la structure figure Fig 7.II la cellule de base et la zone de Brillouin seront identiques. Mais pour traiter la structure Fig 9.II on choisira la cellule de base suivante :

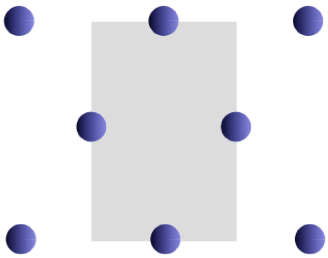


Fig 11.II Cellule de base du réseau de tiges décalé Fig 9.II en gris.

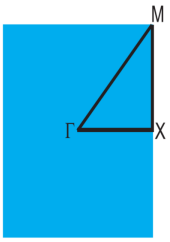


Fig 12.II ΓXM : Zone de "balayage" correspondante.

Comme la première zone Fig 11.II n'est plus la zone de Brillouin, on appellera la deuxième zone Fig 12.II : zone de "balayage". La zone de Brillouin irréductible ou de "balayage" représente pour nos expériences des directions incidentes définies par le vecteur d'onde de l'onde plane incidente. Pour une étude complète, la zone de balayage doit comporter suffisamment de directions incidentes pour connaître le comportement de l'ensemble de la structure face à une excitation.

II.1.7.5 Conclusion

Nous venons d'exposer brièvement quelques notions indispensables sur les cristaux photoniques. Elles sont transposables à l'étude des structures périodiques en général. Notamment nous avons vu que l'étude de la structure périodique se limite à une zone de Brillouin irréductible ou de "balayage". Ainsi cette zone est l'espace permettant de tracer le diagramme de bandes que nous allons définir dans la section II.1.8 suivante.

II.1.8 Diagramme de bandes

II.1.8.1 Introduction

Le diagramme de bandes est un outil qui permet de connaître la bande interdite de la structure périodique. Autrement dit on peut identifier les directions "aveugles" pour lesquelles il n'y a pas de propagation possible dans la structure. L'autre utilité du diagramme de bande qui est étudiée section V.4 est de pouvoir y lire des modes guidés ou de surface grâce à une signature spécifique.

II.1.8.2 Présentation

Pour une direction de propagation définie par le vecteur d'onde incident, on reporte un nombre donné de modes propres qui peuvent se propager dans la structure. Les modes propres sont des solutions fréquentielles des équations de Maxwell, ils peuvent être exprimés dans plusieurs unités : Hertz, rad/m... Ce diagramme s'appelle diagramme de bandes ou diagramme de dispersion. Dans une structure périodique en 1d, 2d et 3d il est utile de connaître le diagramme de dispersion dans plusieurs directions de propagation. Pour des raisons de périodicité dans le matériau on se limite à l'étude de la cellule de base du motif périodisé (cf section précèdente II.1.7).

II.1.8.3 Exemple de plaques périodiques infinies selon l'axe Ox .

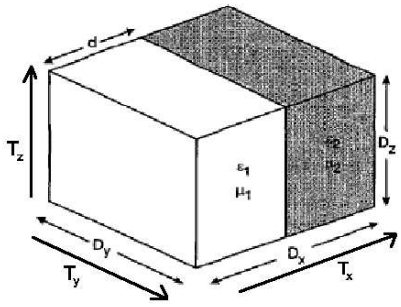


Fig 13.II La cellule de base est le cube de hauteur $D_X = D_Y = D_Z = 0.2m,\ d = 0.1m.$
Les caractéristiques des Deux matériaux sont $\epsilon_1 = \mu_1 = \mu_2 = 1$ et $\epsilon_2 = 9$.

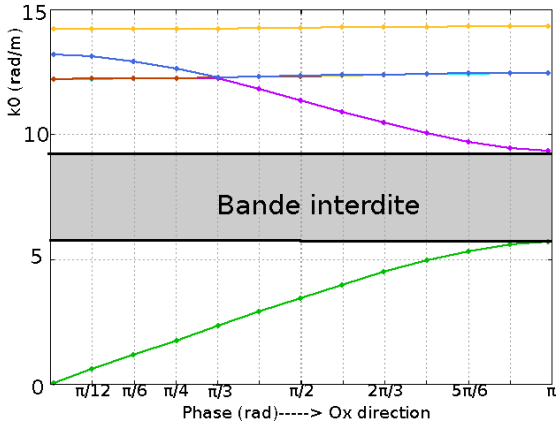


Fig 14.II Diagramme de bandes correspondant.

Soit $\mathbf{k}_i = (k_{ix}, k_{iy}, k_{iz})$ le vecteur d'onde du champ incident $\mathbf{E}_i$. La structure Fig 13.II est périodique dans la direction Ox et est invariante donc périodique dans les directions Oy,Oz. Par analogie avec la formule (9.II) on peut écrire le champ total comme :

$$\mathbf{E}(x,y,z) = \sum_{n=-\infty}^{+\infty} \sum_{m=-\infty}^{+\infty} \sum_{p=-\infty}^{+\infty} \hat{\mathbf{E}}_{nmp} e^{j(\mathbf{G}_{nmp} - \mathbf{k}_i).\mathbf{r}} \text{ avec } \mathbf{G}_{nmp} = (\frac{2n\pi}{D_X}, \frac{2m\pi}{D_Y}, \frac{2p\pi}{D_Z}),$$

Les déphasages selon les trois axes du repère cartésien $\phi_x,\ \phi_y,\ \phi_z$ sont exprimés en radian et

sont respectivement égaux à :

$$\phi_x = \mathbf{k}_i.\mathbf{T}_{\mathrm{x}} = k_{ix}.\mathrm{D_X},\ \phi_y = \mathbf{k}_i.\mathbf{T}_{\mathrm{y}} = k_{iy}.\mathrm{D_Y},\ \phi_z = \mathbf{k}_i.\mathbf{T}_{\mathrm{z}} = k_{iz}.\mathrm{D_Z}$$

Où $\mathbf{T}_{\mathrm{x}},\ \mathbf{T}_{\mathrm{y}},\ \mathbf{T}_{\mathrm{z}}$ sont les vecteurs de translation entre les faces en regard dans les directions de périodicité (cf Fig 13.II). Dans le diagramme Fig 14.II sont représentées les solutions en rad/m en fonction de la phase ϕ_x . Comme la structure est infinie dans les directions Oz et Oy, les valeurs obtenues seront invariantes par variation de $\mathbf{k}_i$ dans ces directions et donc k_{iy} et k_{iz} sont fixés à zéro. D'autre part pour éviter une situation déjà rencontrée à cause de la périodicité, k_{ix} ne doit varier que dans l'intervalle $[-\frac{\pi}{\mathrm{D_X}}, \frac{\pi}{\mathrm{D_X}}]$. Si k_{ix} appartient à l'intervalle $[-\frac{\pi}{\mathrm{D_X}}, \frac{\pi}{\mathrm{D_X}}]$ on peut facilement comprendre qu'on obtiendra des solutions symétriques par rapport à 0 et donc l'intervalle d'étude peut encore être réduit à $[0, \frac{\pi}{\mathrm{D_X}}]$ ce qui revient à faire varier la phase $\phi_x = \mathbf{k}_i.\mathbf{T}_x$ de 0 à π. Sur certains diagrammes les phases seront indiquées alors que sur d'autre c'est la norme et la direction du vecteur $\mathbf{k}_i$. On observe une bande interdite entre six et neuf rad/m. Pour plus de précision on pourra se reporter à la section V où les zones de balayage sont détaillées pour chaque cas traité.

II.1.8.4 Conclusion

Après avoir brièvement rappelé les fondements de la théorie de Floquet et les outils qui en découlent, nous allons aborder les différentes méthodes permettant de traiter les problèmes périodiques.

II.2 Différentes façon d'aborder la propagation dans les structures périodiques

Nous allons voir dans la suite du document qu'il existe deux façons différentes de connaître les propriétés de bandes interdites d'une structure périodique :

1. La première consiste à chercher directement les modes propres de la cellule de base périodisée grâce aux conditions limites de Floquet (cf section II.1).
2. La deuxième consiste à introduire une source explicite et à regarder les coefficients de transmission et de réflexion pour une fréquence incidente donnée.

Remarque II.2.

1. La première méthode suppose que la cellule de base est "fermée" par des conditions limites sur toutes ses faces extérieures. Les conditions de fermeture peuvent être des conditions périodiques, des conditions métalliques ou des conditions absorbantes (cf section III.3). Dans ce cas il n'y a pas de source explicite extérieure, bien que les conditions de Floquet supposent l'existence implicite d'une onde plane incidente (cf section II.1.8). La formulation de cette méthode aboutit à un problème de calcul de valeurs propres et donc les solutions modales sont multiples à chaque calcul.
2. La deuxième méthode introduit une source de type onde plane, ainsi cette technique nécessite l'existence d'un demi-espace ouvert. Cette méthode ne peut pas gérer les problèmes périodiques dans les trois directions de l'espace. En revanche, elle s'applique très bien pour les milieux mono ou bi-périodiques. La formulation de cette méthode aboutit à un problème de résolution de système linéaire avec second membre conduisant à un champ électrique solution pour une fréquence donnée. Une mise en œuvre de celle ci pour des simulations de SER d'antennes réseaux est proposée dans [47] et [48] et pour les calculs de rayonnement d'antenne réseau dans [42].

Nous allons dans la suite présenter les méthodes une à une pour finir sur la méthode que nous développons dans le chapitre III. Nous expliquons aussi comment choisir une méthode en fonction de la structure traitée. Pour toutes les méthodes, nous abordons quand c'est possible les deux types de modélisation que l'on vient de présenter :

1. Avec source explicite conduisant à la résolution d'un système linéaire avec second membre.
2. Sans source explicite ce qui aboutit à la résolution d'un problème de calcul de valeurs propres.

II.3 Méthodes analytiques modales

II.3.1 Introduction

Les méthodes analytiques modales utilisent la décomposition en séries de Floquet (cf section II.1) de la solution du problème périodique recherchée. Il existe (cf remarque II.2) deux façons de traiter le problème.

1. La méthode des ondes planes [49]. Les conditions aux limites de la cellule de base sont périodiques dans les trois directions de l'espace et il n'y a pas de source explicite.
2. La méthode RCWA[52, 51] pour Rigorous Coupled-Wave Analysis. Cette méthode nécessite l'introduction d'une source explicite.

II.3.2 Méthode des ondes planes

On suppose que la structure est périodique selon trois directions non nécessairement confondues avec les axes du repère cartésien. On peut alors, après changements successifs de repère, écrire le champ $\mathbf{E}$ comme une série pseudo-périodique. Si on se place dans le cas le plus général où les directions de périodicité ne sont pas confondues avec les trois axes du repère cartésien, alors d'après les notations définies équation (11.II) :

$$\mathbf{E}(x,y,z) = \sum_{p=-\infty}^{+\infty}\sum_{n=-\infty}^{+\infty}\sum_{m=-\infty}^{+\infty} \hat{\mathbf{E}}_{nmp} e^{j\left((G_{x_{nmp}}-k_{ix})x+(G_{y_{nmp}}-k_{iy})y+(G_{z_{nmp}}-k_{iz})z\right)}$$

On en déduit les expressions de $\mathbf{E}$ et $\mathbf{H}$:

$$\mathbf{E} = \sum_{p=-\infty}^{+\infty}\sum_{n=-\infty}^{+\infty}\sum_{m=-\infty}^{+\infty} \hat{\mathbf{E}}_{nmp} e^{j\left((\mathbf{G}_{nmp}-\mathbf{k}_i).\mathbf{r}\right)}, \quad \mathbf{H} = \sum_{p=-\infty}^{+\infty}\sum_{n=-\infty}^{+\infty}\sum_{m=-\infty}^{+\infty} \hat{\mathbf{H}}_{nmp} e^{j\left((\mathbf{G}_{nmp}-\mathbf{k}_i).\mathbf{r}\right)}$$

Si μ et ϵ désignent respectivement la perméabilité et la permittivité effective des milieux étudiés alors les équations de Maxwell harmoniques en champs $\mathbf{E}$ et $\mathbf{H}$ peuvent s'écrire[38] :

$$\boldsymbol{\nabla} \times \mathbf{E} - j\omega\mu\mathbf{H} = \mathbf{0}, \quad \boldsymbol{\nabla} \times \mathbf{H} = -j\omega\epsilon\mathbf{E} + \mathbf{J} \tag{12.II}$$

On considère qu'il n'y a pas de charges ni de courant dans le milieu ($\mathbf{J} = \mathbf{0}$ et $\rho = 0$). Et comme :

$$\boldsymbol{\nabla} \times \hat{\mathbf{E}}_{nmp} e^{j\left((\mathbf{G}_{nmp}-\mathbf{k}_i).\mathbf{r}\right)} = j(\mathbf{G}_{nmp}-\mathbf{k}_i) \times \hat{\mathbf{E}}_{nmp} e^{j\left((\mathbf{G}_{nmp}-\mathbf{k}_i).\mathbf{r}\right)}.$$

On en déduit que :

$$\sum_{p=-\infty}^{+\infty}\sum_{n=-\infty}^{+\infty}\sum_{m=-\infty}^{+\infty}(\mathbf{G}_{nmp}-\mathbf{k}_i)\times\hat{\mathbf{E}}_{nmp}e^{j\left((\mathbf{G}_{nmp}-\mathbf{k}_i).\mathbf{r}\right)}+j\omega\mu\hat{\mathbf{H}}_{nmp}e^{j\left((\mathbf{G}_{nmp}-\mathbf{k}_i).\mathbf{r}\right)}=\mathbf{0}$$

$$\Rightarrow j(\mathbf{G}_{nmp}-\mathbf{k}_i)\times\hat{\mathbf{E}}_{nmp}=+j\omega\mu\hat{\mathbf{H}}_{nmp},\ j(\mathbf{G}_{nmp}-\mathbf{k}_i)\times\hat{\mathbf{H}}_{nmp}=-j\omega\epsilon\hat{\mathbf{E}}_{nmp}\ \ \forall(n,m,p)\in\mathbb{Z} \tag{13.II}$$

$$\Rightarrow(\mathbf{G}_{nmp}-\mathbf{k}_i).\hat{\mathbf{E}}_{nmp}=(\mathbf{G}_{nmp}-\mathbf{k}_i).\hat{\mathbf{H}}_{nmp}=\mathbf{0}\ \ \forall(n,m,p)\in\mathbb{Z} \tag{14.II}$$

$$\Rightarrow\nabla.\mathbf{E}=\nabla.\mathbf{H}=\mathbf{0}. \tag{15.II}$$

Ceci est équivalent à dire que la solution s'écrit comme étant la superposition d'ondes planes [40]. À cette écriture un peu lourde on préfère dans la littérature [49] une écriture plus condensé en remplaçant $\mathbf{G}_{nmp}$ par un vecteur noté $\mathbf{G}$ appartenant au réseau réciproque (cf section II.1.7.3) noté $R.R$ les quantités s'écrivent alors :

$$\tilde{\mathbf{E}}(\tilde{\mathbf{r}})=\mathbf{E}(\mathbf{r})=\sum_{G\in R.R}\hat{\mathbf{E}}_G e^{j\left((\mathbf{G}-\mathbf{k}_i).\mathbf{r}\right)},\qquad\mathbf{H}=\sum_{\mathbf{G}\in R.R}\hat{\mathbf{H}}_G e^{j\left((\mathbf{G}-\mathbf{k}_i).\mathbf{r}\right)}.$$

Et donc les relations (13.II) et (14.II) deviennent :

$$\boldsymbol{\nabla}\times\hat{\mathbf{E}}_G e^{j\left((\mathbf{G}-\mathbf{k}_i).\mathbf{r}\right)}=(\mathbf{G}-\mathbf{k}_i)\times\hat{\mathbf{H}}_G e^{j\left((\mathbf{G}-\mathbf{k}_i).\mathbf{r}\right)}$$

et

$$(\mathbf{G}-\mathbf{k}_i).\hat{\mathbf{E}}_G=(\mathbf{G}-\mathbf{k}_i).\hat{\mathbf{H}}_G=\mathbf{0}\ \ \forall\mathbf{G}\in R.R. \tag{16.II}$$

II.3.2.1 Résolution en champ E

On suppose que la perméabilité relative $\mu_r=1$ et que la permittivité relative ϵ_r sont constantes par morceaux et périodiques selon trois directions données par les vecteurs de base du réseau de Bravais (cf section II.1.7.2). Ainsi ϵ_r s'écrit sous forme d'une série de Fourier :

$$\epsilon_r(\mathbf{r})=\sum_{\mathbf{G}\in R.R}\hat{\epsilon}_G e^{j\left(\mathbf{G}.\mathbf{r}\right)},$$

On en déduit l'expression :

$$\epsilon_r\mathbf{E}=\sum_{G',G''\in R.R}\hat{\epsilon}_{G'}\hat{\mathbf{E}}_{G''}e^{j\left((\mathbf{G}'+\mathbf{G}''-\mathbf{k}_i).\mathbf{r}\right)}.$$

Soit $k_0=\omega\sqrt{\epsilon_0\mu_0}$. En transformant l'équations (12.II), l'équation de Maxwell en champ $\mathbf{E}$ pour un milieu non dispersif s'écrit [38] :

$$\boldsymbol{\nabla}\times(\boldsymbol{\nabla}\times\mathbf{E})=k_0^2\epsilon_r\mathbf{E}$$

et comme

$$\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \hat{\mathbf{E}}_G e^{j\left((\mathbf{G}-\mathbf{k}_i).\mathbf{r}\right)}) = (\mathbf{G}-\mathbf{k}_i) \times \left((\mathbf{G}-\mathbf{k}_i) \times \hat{\mathbf{E}}_G e^{j\left((\mathbf{G}-\mathbf{k}_i).\mathbf{r}\right)}\right),$$

on en déduit en reportant dans l'équation de Maxwell en champ $\mathbf{E}$ et en identifiant les coefficients de Fourier de la fonction nulle que :

$$(\mathbf{G}-\mathbf{k}_i) \times \left((\mathbf{G}-\mathbf{k}_i) \times \hat{\mathbf{E}}_G\right) + k_0^2 \sum_{\mathbf{G}'+\mathbf{G}''=\mathbf{G}} \hat{\epsilon}_{G'} \hat{\mathbf{E}}_{G''} = \mathbf{0}.$$

En remarquant que $\mathbf{G}' = \mathbf{G} - \mathbf{G}''$ on peut écrire :

$$(\mathbf{G}-\mathbf{k}_i) \times \left((\mathbf{G}-\mathbf{k}_i) \times \hat{\mathbf{E}}_G\right) + k_0^2 \sum_{\mathbf{G}'' \in R.R} \hat{\epsilon}_{G-G''} \hat{\mathbf{E}}_{G''} = \mathbf{0} \quad \forall \mathbf{G} \in R.R.$$

En appliquant la formule du double produit vectoriel $(\vec{u} \wedge (\vec{v} \wedge \vec{w}) = (\vec{u} \cdot \vec{w})\vec{v} - (\vec{u} \cdot \vec{v})\vec{w}$) on obtient alors :

$$((\mathbf{G}-\mathbf{k}_i).\hat{\mathbf{E}}_G)(\mathbf{G}-\mathbf{k}_i) - \|\mathbf{G}-\mathbf{k}_i\|^2 \hat{\mathbf{E}}_G + k_0^2 \sum_{\mathbf{G}'' \in R.R} \hat{\epsilon}_{G-G''} \hat{\mathbf{E}}_{G''} = \mathbf{0} \quad \forall \mathbf{G} \in R.R.$$

Mais l'équation se simplifie d'après la relation (16.II). Ce qui donne :

$$\|\mathbf{G}-\mathbf{k}_i\|^2 \hat{\mathbf{E}}_G = k_0^2 \sum_{\mathbf{G}'' \in R.R} \hat{\epsilon}_{G-G''} \hat{\mathbf{E}}_{G''} = \mathbf{0} \quad \forall \mathbf{G} \in R.R.$$

Ce système infini d'équations aboutit à un système matriciel infini dont les inconnues sont les composantes des vecteurs $\hat{\mathbf{E}}_G$. Si on tronque ce système en restreignant le réseau réciproque à N éléments, on obtient un système matriciel du type :

$$A\mathbf{X} = \frac{1}{k_0^2} M\mathbf{X} \tag{17.II}$$

Où $\mathbf{X}$ est le vecteur $\hat{\mathbf{E}}_{G_1}, \ldots, \hat{\mathbf{E}}_{G_N}$, A est une matrice de dimension $3N \times 3N$ avec des coefficients égaux trois par trois sur une même ligne :

$$A_{ij} = \hat{\epsilon_r}(\mathbf{G_i} - \mathbf{G_j})$$

Et M est la matrice diagonale $3N \times 3N$ composée des matrices par bloc :

$$\|\mathbf{G}-\mathbf{k}_i\|^2 I_{3\times 3}, \quad \mathbf{G} \in R.R$$

Remarque II.3. Comme nous venons de le voir le nombre d'inconnues est égal à $3 \times N$ où N désigne le nombres d'ondes planes de la série de Floquet. En théorie ce système doit être infini, mais en le tronquant on commet évidemment des erreurs numériques. Plus le nombre d'ondes planes est grand plus on a de chance de se rapprocher de la solution réelle. D'autre

part, le problème (17.II) revient à calculer les valeurs propres de la matrice AM^{-1} et donc nécessite l'inversion de la matrice M.

II.3.2.2 Résolution en champs H L'équation en champs **H** s'écrit

$$\boldsymbol{\nabla} \times (\frac{\boldsymbol{\nabla} \times \mathbf{H}}{\epsilon_r}) = k_0^2 \mu_r \mathbf{H}$$

Où μ_r est la perméabilité relative du milieu étudié. On utilise cette fois-ci la transformée de la fonction inverse de ϵ_r :

$$\frac{1}{\epsilon_r} = \sum_{\mathbf{G} \in R.R} \widehat{\frac{1}{\epsilon_G}} e^{j(\mathbf{G}.\mathbf{r})}$$

Ensuite le principe est le même que pour l'équation du champ électrique **E** que l'on vient d'établir. On pourra regarder la démonstration détaillée en [49, 26].

Pour les structures canoniques comme les tiges diélectriques cylindriques, la permittivité relative ϵ_r à une valeur constante différente de un dans le cylindre. On peut alors facilement calculer la transformée de de Fourier à l'aide des fonctions de Bessel [49, 26].

Mais pour une cellule de base avec une géométrie plus complexe(non canonique) il faut calculer la transformer de Fourier des coefficients ϵ_r dont le terme :

$$\hat{\epsilon}_G = \frac{1}{|Scell|} \int_{Scell} \epsilon_r(\mathbf{r}) e^{j\mathbf{G}.\mathbf{r}} d\mathbf{r}$$

doit être calculé grâce aux formules numériques d'approximation d'intégrales en découpant en pixels la forme "non canonique". En d'autres termes, on calcule la transformée de Fourier discrète de la permittivité.

Fig 15.II Exemple de géométrie non canonique : coupe 2d d'un réseau de tiges en forme d'étoile.

Remarque II.4. Si on essaie de traiter des structures périodiques à forme non canonique, on accumule alors les erreurs dues à la troncature du nombre d'ondes planes et les erreurs sur le calcul de l'intégrale de la transformée de Fourier de la permittivité relative. En effet lorsqu'il

y a de forts contrastes d'indices entre les matériaux qui constituent la structure modélisée, il faut garder beaucoup de termes dans la décomposition de Fourier de la permittivité (ce qui revient à faire sa transformée de Fourier discrète sur un grand nombre de termes) car il faut prendre en compte correctement les fortes discontinuités aux interfaces entre matériaux. Il est alors difficile de savoir quel est le nombre minimum de termes à garder dans la série de Fourier pour avoir une précision acceptable. D'autre part, garder un grand nombre de termes dans la série de Fourier peut être très handicapant en termes de ressources mémoire et de temps de calcul.

Remarque II.5. Il est possible de traiter des matériaux dispersifs en faisant dépendre la permittivité de la pulsation (cf section VI.2) mais la taille du problème de calcul de modes propres est multiplié par deux au minimum.

II.3.3 Présentation de la méthode RCWA [52, 51]

La méthode RCWA [50, 52, 53] (RCWA, pour Rigorous Coupled Wave Analysis) est un algorithme qui permet de calculer des coefficients de réflexion et de transmission de structures bi-périodiques. Dans le même genre, la méthode FMM [54] (FMM, pour Fourier Modal Method) est une adaptation de la méthode RCWA quand les deux vecteurs de périodicité ne sont pas orthogonaux [54]. Ces méthodes utilisent un découpage spatial dans la direction orthogonale au réseau. Fig 16.II on considère une structure bi-périodique en (x, y) de périodes respectives T_1 et T_2. Le nombre de couches du découpage spatial est noté Nb_C. Dans chacune de ces couches, la permittivité est bi-périodique et supposée invariante en z.

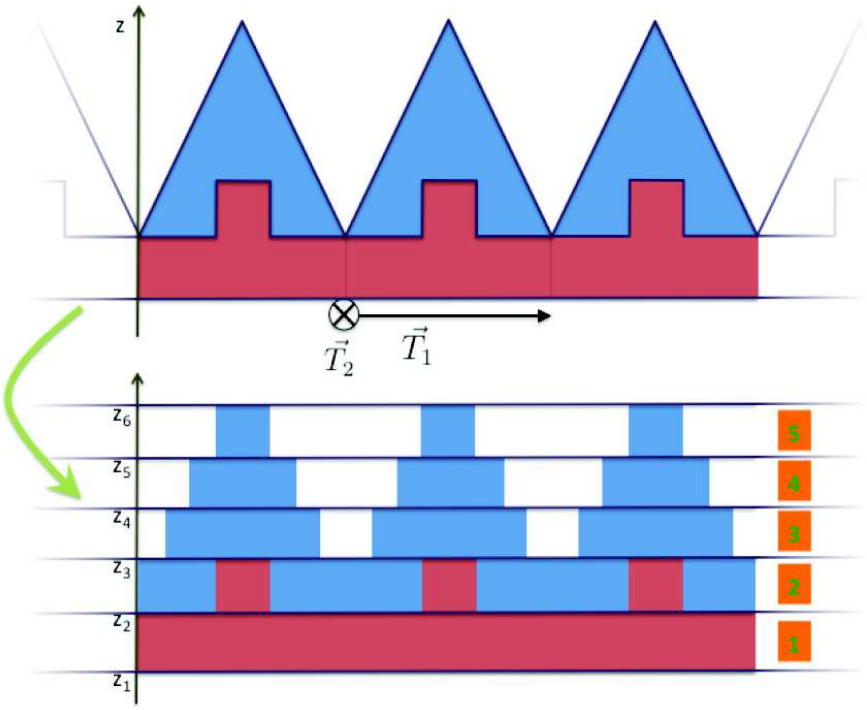


Fig 16.II Exemple simplifié de modélisation d'une structure bipériodique par la méthode RCWA. Il y a trois matériaux différents (un par couleur) et cinq couches ($Nb_C = 5$) ([51] page 37).

Pour $1 \leq p \leq Nb_C,\ (x,y) \in \mathbb{R}^2,\ z \in [z_p, z_{p+1}[$, on note :

$$\epsilon(x,y,z) = \epsilon(x,y,z_p), = \epsilon_p(x,y)$$

La source est une onde plane localisée au bord de la dernière couche Nb_C (z_6 sur la figure Fig 16.II). Son vecteur d'onde incident est noté $\mathbf{k}_i = (k_{ix}, k_{iy}, k_{iz})$. Dans chaque couche p, on écrit les champs $\mathbf{E}$ et $\mathbf{H}$ sous forme de séries pseudo-périodiques comme dans l'équation (8.II) :

$$\mathbf{E}(x,y,z) = \sum_{n=-\infty}^{+\infty} \sum_{m=-\infty}^{+\infty} \hat{\mathbf{E}}_{nm}(z) e^{j\left(\alpha_n x + \beta_m y\right)},\ \mathbf{H}(x,y,z) = \sum_{n=-\infty}^{+\infty} \sum_{m=-\infty}^{+\infty} \hat{\mathbf{H}}_{nm}(z) e^{j\left(\alpha_n x + \beta_m y\right)} \tag{18.II}$$

avec

$$\alpha_n = \frac{2\pi n}{T_1} - k_{ix},\ \beta_m = \frac{2\pi m}{T_2} - k_{iy}$$

On écrit aussi la décomposition en série de Fourier de la permittivité ϵ_p [51] :

$$\epsilon_p(x,y) = \sum_{n=-\infty}^{+\infty} \sum_{m=-\infty}^{+\infty} \hat{\epsilon}_{p,nm} e^{j\left(\frac{2\pi n}{T_1}x + \frac{2\pi m}{T_2}y\right)} \tag{19.II}$$

Les expressions (18.II) et (19.II) sont ensuite injectées dans les équations de Maxwell harmoniques que l'on va écrire dans chaque couche $p \in [1, \ldots Nb_C]$ en imposant des conditions de continuité aux interfaces. On obtient alors un système différentiel du premier ordre en z [51] :

$$\frac{d}{dz}\begin{pmatrix} \vdots \\ \hat{\mathbf{E}}_{nm}(z) \\ \vdots \\ \hat{\mathbf{H}}_{nm}(z) \\ \vdots \end{pmatrix} = \mathcal{M}_p \begin{pmatrix} \vdots \\ \hat{\mathbf{E}}_{nm}(z) \\ \vdots \\ \hat{\mathbf{H}}_{nm}(z) \\ \vdots \end{pmatrix} \tag{20.II}$$

Les coefficients de la permittivité $\hat{\epsilon}_{p,nm}$ interviennent dans l'expression de la matrice $\mathcal{M}_p$. En posant :

$$\mathbf{U}(z) = \begin{pmatrix} \vdots \\ \hat{\mathbf{E}}_{nm}(z) \\ \vdots \\ \hat{\mathbf{H}}_{nm}(z) \\ \vdots \end{pmatrix} \text{ l'équation (20.II) s'écrit } \frac{d}{dz}\mathbf{U}(z) = \mathcal{M}_p \mathbf{U}(z) \tag{21.II}$$

Afin de pouvoir intégrer cette équation différentielle, il faut tronquer le système et pour

$z \in [z_p, z_{p+1}[$, la solution de (20.II) est :

$$\mathbf{U}(z) = exp\{(z - z_p)\mathcal{M}_p\}\mathbf{U}(z_p)$$

La continuité du champ aux interfaces entre les couches nous permet d'écrire :

$$\mathbf{U}(z_{p+1}) = exp\{(z_{p+1} - z_p)\mathcal{M}_p\}\mathbf{U}(z_p).$$

On peut alors relier les champs de part et d'autre de la couche p. Ainsi les composantes du champ de la dernière couche $\mathbf{U}(z_{Nb_C})$ seront reliées aux composantes du champ de la première couche $\mathbf{U}(z_1)$ [52] :

$$\mathbf{U}(z_{Nb_C}) = \prod_{p=1}^{Nb_C} exp\{(z_{p+1} - z_p)\mathcal{M}_p\}\mathbf{U}(z_1). \tag{22.II}$$

Cette expression se calcule en diagonalisant ou trigonalisant (si elles ne sont pas diagonalisables) les matrices $\mathcal{M}_p,\ p \in [1, \dots Nb_C]$. La première couche et la dernière couche sont les milieux d'observation. Ces milieux sont supposés homogènes et uniformes et donc les champs $\mathbf{E}$ et $\mathbf{H}$ y vérifient les équations de Helmoltz. Autrement dit pour $z \in [z_{Nb_C-1}, z_{Nb_C}[\cup[z_1, z_2[$ on peut écrire [51] :

$$\triangle\Big(\sum_{n=-\infty}^{+\infty}\sum_{m=-\infty}^{+\infty}\hat{\mathbf{E}}_{nm}(z)e^{j\left(\alpha_n x+\beta_m y\right)}\Big) + k_{obs}^2\Big(\sum_{n=-\infty}^{+\infty}\sum_{m=-\infty}^{+\infty}\hat{\mathbf{E}}_{nm}(z)e^{j\left(\alpha_n x+\beta_m y\right)}\Big) = \mathbf{0} \tag{23.II}$$

Où $k_{obs}^2 = \omega^2\epsilon\mu$ avec ϵ et μ la permittivité et la perméabilité des milieux d'observations. Comme l'opérateur $\triangle$ est linéaire, on déduit que chaque coefficient $\hat{\mathbf{E}}_{nm}(z)$ de la série de (23.II) vérifie une équation de Helmoltz dont les solutions s'écrivent soit comme un terme incident ou "entrant" noté $\hat{\mathbf{E}}_{i_{nm}}(z)$, soit comme un terme réfléchi ou "sortant" noté $\hat{\mathbf{E}}_{r_{nm}}(z)$ [51] :

$$\hat{\mathbf{E}}_{i_{nm}}(z) = \hat{\mathbf{E}}_{i_{nm}}e^{-j\gamma_{nm}z},\ \hat{\mathbf{E}}_{r_{nm}}(z) = \hat{\mathbf{E}}_{r_{nm}}e^{+j\gamma_{nm}z} \text{ avec } \gamma_{nm} = \sqrt{k_{obs}^2 - \alpha_n^2 - \beta_m^2}.$$

Cette écriture nous permet de donner une autre formulation du champ $\mathbf{E}$ dans les milieux d'observation :

$$\mathbf{E}(x,y,z) = \sum_{n=-\infty}^{+\infty}\sum_{m=-\infty}^{+\infty}\hat{\mathbf{E}}_{i_{nm}}e^{j\left(\alpha_n x+\beta_m y-\gamma_{nm}z\right)} + \sum_{n=-\infty}^{+\infty}\sum_{m=-\infty}^{+\infty}\hat{\mathbf{E}}_{r_{nm}}e^{j\left(\alpha_n x+\beta_m y+\gamma_{nm}z\right)}. \tag{24.II}$$

Remarque II.6. L'écriture équation (24.II) est générale pour un champ $\mathbf{E}$ contenu dans un milieu homogène duquel on peut observer une structure bi-périodique. Cette écriture est utilisée quand on veut exprimer une condition limite de radiation sur la limite supérieur ou inférieur d'un domaine bi-périodique (cf section VI.2.2.4)

II.3.3.1 Exemple d'écriture des conditions limites

Considérons par exemple que les deux milieux d'observation la couche 1 et la couche Nb_C soient de l'air, et que la source $\mathbf{k}_i = (k_{ix}, k_{iy}, k_{iz})$ soit dans la couche Nb_C. Soit $\mathbf{E}_1$ et $\mathbf{E}_{Nb_C}$ les champs dans la couche 1 et dans la couche Nb_C. Leur expression est donnée par (24.II) et leur nombre d'onde d'observation est $k_{obs} = k_0 = \omega^2\epsilon_0\mu_0$. Si on appelle $\mathbf{W}_{Nb_C}$ et $\mathbf{W}_1$ les coordonnées $(\ldots, \hat{\mathbf{E}}_{i_{nm}}, \ldots \hat{\mathbf{E}}_{r_{nm}}, \ldots)_l,\ l = 1, Nb_C$ de $\mathbf{E}_{Nb_C}$ et $\mathbf{E}_1$ alors on peut écrire [52] :

$$\mathbf{W}_{Nb_C} = \begin{pmatrix} \mathbf{1} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \hat{\mathbf{E}}^{Nb_C}_{r_{11}} \\ \vdots \\ \hat{\mathbf{E}}^{Nb_C}_{r_{nm}} \\ \vdots \end{pmatrix},\ \mathbf{W}_1 = \begin{pmatrix} \hat{\mathbf{E}}^{1}_{i_{11}} \\ \vdots \\ \hat{\mathbf{E}}^{1}_{i_{nm}} \\ \vdots \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \vdots \end{pmatrix} \tag{25.II}$$

En effet pour $\mathbf{E}_{Nb_C}$ le champ incident est connu et c'est une onde plane d'amplitude choisie égale à $\mathbf{1}$ et pour $\mathbf{E}_1$ il n'y a pas de réflexion si on considère que toute l'énergie est dissipée dans l'air.

En effectuant la transformation nécessaire pour passer des coordonnées $\mathbf{U}_{Nb_C}$ (respectivement $\mathbf{U}_1$) aux coordonnées $\mathbf{W}_{Nb_C}$ (respectivement $\mathbf{W}_1$) défines en (21.II) et (25.II), on peut alors retrouver les coefficients de réflexion $(\hat{\mathbf{E}}^{Nb_C}_{r_{11}}, \ldots, \hat{\mathbf{E}}^{Nb_C}_{r_{nm}}, \ldots)$ et d'incidence $(\hat{\mathbf{E}}^{1}_{i_{11}}, \ldots, \hat{\mathbf{E}}^{1}_{i_{nm}}, \ldots)$ par la résolution d'un système linéaire construit à partir de la relation (11.III).

II.3.4 Intérêts des méthodes modales

- Ces méthodes sont relativement faciles à coder sur des géométries canoniques.
- Quand on traite des géométries simples la solution converge pour un nombre d'inconnues relativement peu élevé. Dans ce cas, bien que les matrices soient pleines (cf remarque II.3), les ressources informatiques utilisées pour un calcul sont donc réduites par rapport aux autres méthodes existantes.
- Les matériaux peuvent être dispersifs : leurs paramètres peuvent dépendre de la fréquence de calcul [29, 30].

II.3.5 Aspects limitant de la méthode

- Quand on traite des géométries non canoniques la transformée de Fourier discrète de la permittivité nécessite un découpage spatial suffisamment fin pour prendre en compte les détails des zones discrétisées.
- Quand on traite des géométries qui comprennent des détails de taille caractéristique très petite devant le pas du réseau, le nombre de termes de la série de Fourier doit

être suffisamment grand pour assurer la convergence de la méthode. Dans ce cas, les ressources informatiques nécessaires à un calcul deviennent pénalisantes (cf remarque II.4).

– Quand les matériaux sont dispersifs : la taille du problème de calcul de modes propres est multipliée [30] comme nous le verrons par la suite.

II.4 Méthode FDTD (Finite-Difference Time-Domain)

II.4.1 Principe de la méthode

C'est une méthode de calcul de différences finies, qui permet de résoudre les équations de Maxwell dans le domaine temporel. La formulation temporelle des équations de Maxwell peut s'écrire de la façon suivante[55] :

$$\begin{cases} \boldsymbol{\nabla} \times \mathbf{E}(\mathbf{r},t) = -\mu_0 \dfrac{\partial(\mathbf{H})}{\partial t} \quad \text{pour} \quad (\mathbf{r},t) \in \mathbb{R}^3 \times \mathbb{R}^t \\ \boldsymbol{\nabla} \times \mathbf{H}(\mathbf{r},t) = \frac{\partial}{\partial t}\big[\mathbf{D}_0(\mathbf{r},t) + \mathbf{P}_d(\mathbf{r},t)\big] \end{cases}$$

Où $\mathbf{E}(\mathbf{r},t)$, $\mathbf{H}(\mathbf{r},t)$ et μ_0 dénotent le champ électrique, le champ magnétique la perméabilité relative du vide et $\mathbf{P}_d(\mathbf{r},t) = \mathcal{P}\delta(\mathbf{r}-\mathbf{r}_0)e^{-i\omega t}$ représente le moment dipolaire oscillant. $\mathcal{P}$ et $\mathbf{r}_0$ sont l'amplitude et la position du moment dipolaire et δ la fonction de Dirac. $\mathbf{D}_0(\mathbf{r},t)$ représente le déplacement électrique causé par la structure diélectrique des éléments périodiques. Le vecteur densité de courant $\mathbf{J}(\mathbf{r},t)$ est implicitement contenu dans l'expression de $\mathbf{D}_0$ puisqu'il peut s'écrire en fonction du champ $\mathbf{E}(\mathbf{r},t)$ grâce à la loi d'Ohm (cf annexe A.5). Elle est généralement égale au produit de convolution entre le champ électrique et la fonction $[\Phi(\mathbf{r},t)]$ représentant la réponse temporelle du diélectrique de permittivité $[\epsilon(\mathbf{r},\omega)]$:

$$\mathbf{D}_0(\mathbf{r},t) = \epsilon_0 \int_{-\infty}^{+\infty} \Phi(\mathbf{r},t-t')\mathbf{E}(\mathbf{r},t')dt' \ , \ [\Phi(\mathbf{r},t)] = \frac{1}{2\pi}\int_{-\infty}^{+\infty}[\epsilon(\mathbf{r},\omega)]e^{-i\omega t}d\omega$$

Pour plus de détails sur la représentation temporelle de la permittivité on peut se reporter à l'annexe A.2.3.

Remarque II.7. On peut facilement envisager de traiter les milieux dispersifs grâce aux modèles explicites type Drude (cf section VI.2.1.3 et [56]) ou Debye (cf annexe A.2.5 et [57]) en transformant l'expression de la permittivité en fonction de la fréquence et en appliquant la transformée de Fourier[58]. De plus, nous pouvons faire face à d'autres cas dispersifs aussi longtemps que la présumée dépendance en fréquence de la constante diélectrique satisfait le principe causalité ou la relation de Kramers-Kronig [55] (cf annexe A.2.2).

De façon classique, on peut par exemple approcher les dérivées partielles figurant dans les équations de Maxwell par les différences finies centrées :

$$\frac{\partial f(x,y,z,t)}{\partial x} = \frac{f(x+\delta x/2,y,z,t) - f(x-\delta x/2,y,z,t)}{\delta x} + o(\delta x^2)$$

Où δx représente le pas de discrétisation suivant la direction x. Si on appelle δt le pas de temps, chaque composante spatiale du champ à l'instant $t+\delta t$ est mise à jour à partir de la composante calculée à partir de l'instant t plus précisément nous utilisons le schéma de Yee [59].

Remarque II.8. Contrairement aux autres méthodes présentées et comme le schéma de Yee est explicite, il n'y a pas ici de processus d'inversions de matrices, mais une mise à jour des

valeurs des composantes en chaque nœuds et à chaque instant. Notons que les schémas aux différences finies implicites nécessitent en général une inversion [60].

Les composantes électriques se calculent aux points de la cellule de Yee que l'on appelle les nœuds électriques. Ces nœuds se situent au milieu des arêtes tandis que les composantes magnétiques se calculent aux centres des faces aussi appelés nœuds magnétiques.

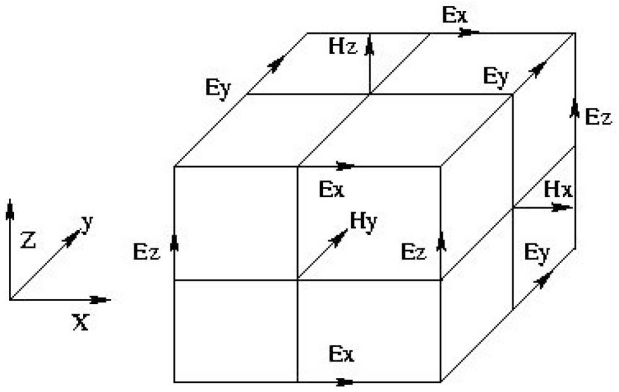


Fig 17.II Cellule de YEE[59] .

Cette répartition des composantes permet au schéma de Yee de respecter la continuité des composantes tangentielles électriques et normales magnétiques à l'interface de deux cellules. La grille FDTD est définie par les champs en chaque point du domaine discrétisé sur l'ensemble des pas de temps réalisés. Voici un exemple de mise à jour d'une coordonnée d'un point de la grille FDTD [58, 59, 61] :

$$H_y^{n+1/2}(i,j,k) = H_y^{n-1/2}(i,j,k) + \frac{\delta t}{\mu_0}\Big[\frac{E_z^n(i+1,j,k) - E_z^n(i,j,k)}{\delta x} - \frac{E_x^n(i,j,k+1) - E_x^n(i,j,k)}{\delta z}\Big]$$

II.4.2 Stabilité numérique

Pour éviter les instabilités numériques, le schéma doit vérifier une condition du type [58, 61] :

$$\delta t \leq \frac{1}{v_{max} \cdot \sqrt{\frac{1}{\delta x^2} + \frac{1}{\delta y^2} + \frac{1}{\delta z^2}}}$$

Où v_{max} est la vitesse maximum de propagation dans le milieu et $\delta x,\ \delta y,\ \delta z$ représentent les pas de discrétisation suivant les directions $x, y\ et\ z$.

II.4.3 Les problèmes périodiques

L'adaptation de la FDTD aux structures périodiques a fait l'objet d'une étude approfondie dans la thèse de Belkhir Abderrahmane [61]. Tout d'abord, on applique les conditions de Floquet (cf section III.3.5) sur les interfaces du domaine de calcul dans le cas classique du réseau de tiges de pas a. Soit une onde plane incidente de vecteur d'onde $\mathbf{k}_i = (k_{ix}, k_{iy}, k_{iz})$, les conditions de Floquet dans le plan $z = 0$ s'écrivent [61] :

$$\mathbf{E}(x=a,y,z=0,t) = \mathbf{E}(x=0,y,z=0,t)e^{jk_{ix}.a}$$
$$\mathbf{E}(x,y=a,z=0,t) = \mathbf{E}(x,y=0,z=0,t)e^{jk_{iy}.a}$$
$$\mathbf{H}(x=a,y,z=0,t) = \mathbf{H}(x=0,y,z=0,t)e^{jk_{ix}.a}$$
$$\mathbf{H}(x,y=a,z=0,t) = \mathbf{H}(x,y=0,z=0,t)e^{jk_{iy}.a}$$

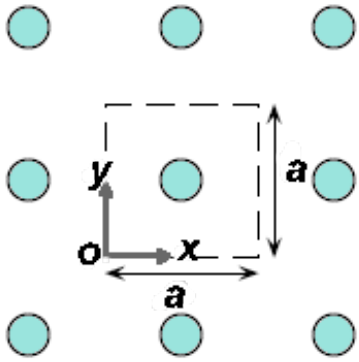


Fig 18.II Cellule de base du réseau carré de tiges.

Ensuite on injecte une série de Floquet à l'instant $t = 0$ et on laisse le signal se propager pendant une durée suffisamment grande pour exciter les premières fréquences propres. La série est construite à partir des vecteurs du réseau réciproque du réseau de Bravais et peut s'écrire (cf section II.3.2) :

$$\sum_{G \in R.R} \hat{\mathbf{E}}_G e^{j\left((\mathbf{G}-\mathbf{k}_i).\mathbf{r}\right)}$$

Après avoir injecté le signal, on calcule l'évolution de la densité volumique d'énergie en fonction de la fréquence grâce la transformée de Fourier du signal temporel. Le spectre obtenu fait apparaitre des pics d'énergie qui correspondent aux fréquences propres de la structure étudiée.

Remarque II.9. Souvent pour que les pics soient stables, un grand nombre d'itérations temporelles sont nécessaires (»10000) [61] et ceci pour chaque valeur de $\mathbf{k}_i$ et donc pour chaque point du diagramme de bandes. On peut alors imaginer que sur des structures maillées finement le temps de calcul pour réaliser un diagramme de bandes puisse devenir interminable. De plus, il faut pouvoir stocker toutes les valeurs des champs aux nœuds à chaque pas de temps ce qui peut nécessiter l'utilisation d'un espace mémoire considérable.

II.4.4 Intérêts

- Les structures sont faciles à mailler puisque le maillage est cubique et le codage de la méthode est relativement simple.
- La connaissance des propriétés temporelles des matériaux suffit au codage de la FDTD qui peut ainsi traiter des problèmes non linéaires et dispersifs (cf remarque II.7).
- Il n'y a pas ici de processus d'inversion de matrice (cf remarque II.8).

II.4.5 Limitations

- Le temps de calcul et l'espace mémoire qui sont utilisés par la grille FDTD peuvent devenir très important notamment parce qu'ils sont liés par une condition CFL nécessaire pour assurer la stabilité du schéma (cf remarques II.9).
- Les cellules de Yee ne doivent pas avoir de rapports de tailles importants pour assurer

la stabilité du schéma [58]. La taille des cubes est donc calibrée par rapport aux zones ou le maillage doit être raffiné.

– À chaque pas de temps, la mise à jour des conditions aux limites périodiques introduisent des approximations, qui peuvent avoir une influence sur la précision du calcul.

II.5 Méthode des équations intégrales

II.5.1 Introduction

Il n'y a pas, à notre connaissance, de méthode de résolution type "modes propres" qui utilise les équations intégrales pour traiter les structures périodiques. En revanche, il existe une méthode d'équations intégrales pour les structures bipériodiques avec l'introduction d'une source explicite (cf remarque II.2). Les méthodes qui font appel aux équations intégrales sont aussi appelées Méthode des moments (MoM, [62]) ou méthode des éléments de frontière (BEM, pour Boundary Element Method). Ces méthodes peuvent être adaptées aux calculs de champs sur des structures tridimensionnelles et bipériodiques. L'adaptation de ces méthodes à ce type de structures a notamment fait l'objet d'une étude approfondie dans la thèse de Samuel Nosal [51]. Les inconnues de ces équations sont les courants magnétiques et surfaciques calculés aux interfaces de matériaux homogènes (i.e dioptres). Une fois ces courants calculés, on peut en déduire la valeur du champ en tout point de l'espace tridimensionnel.

II.5.2 Principe de la méthode [51]

On considère un objet constitué d'un domaine diélectrique homogène Ω_2, de surface extérieure Γ_2 ayant une frontière commune avec un domaine Ω_1, de surface extérieure Γ_1. Une onde plane incidente $(\mathbf{E}^{inc}, \mathbf{H}^{inc})$ est introduite dans le domaine Ω_1.

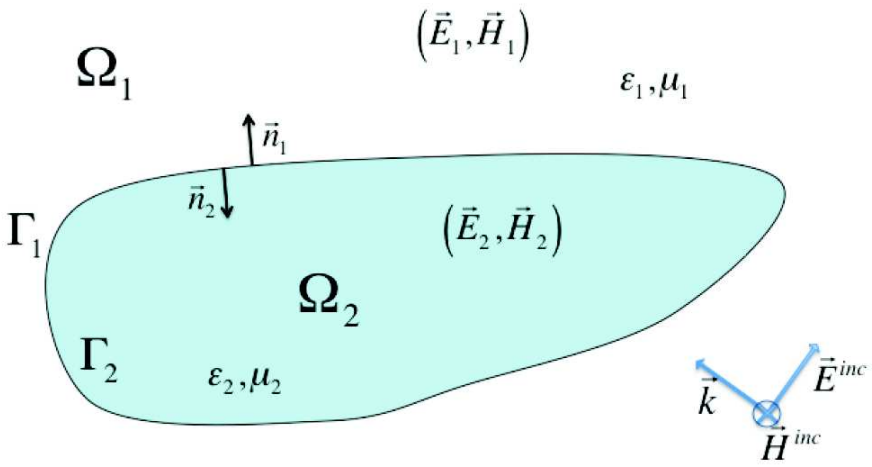


Fig 19.II Diffraction d'une onde incidence par un objet tridimensionnel : les conventions adoptées ([51] page 49).

On exprime le champ électrique dans l'espace à partir les courants magnétiques et surfaciques calculés sur la surface extérieure de chaque domaine à l'aide des équations intégrales

EFIE (Electric Field Integral Equation) [63, 51] :

$$\begin{aligned}
&\frac{1}{2}\mathbf{E}_1(\mathbf{r}) = \mathbf{E}^{inc}(\mathbf{r}) - \frac{1}{4\pi}\int_{\Gamma_1}\left\{j\omega\mu_1\mathbf{J}_1(\mathbf{r}') + \frac{j}{\omega\epsilon_1}\boldsymbol{\nabla}\boldsymbol{\nabla}'.\mathbf{J}_1(\mathbf{r}')\right\}G_1(\mathbf{r}-\mathbf{r}')ds' \\
&+\frac{1}{4\pi}\int_{\Gamma_1}\mathbf{M}_1(\mathbf{r}')\times\boldsymbol{\nabla}G_1(\mathbf{r}-\mathbf{r}')ds' \text{ pour } \mathbf{r}\in\Gamma_1 \\
&\frac{1}{2}\mathbf{E}_2(\mathbf{r}) = -\frac{1}{4\pi}\int_{\Gamma_2}\left\{j\omega\mu_1\mathbf{J}_2(\mathbf{r}') + \frac{j}{\omega\epsilon_1}\boldsymbol{\nabla}\boldsymbol{\nabla}'.\mathbf{J}_2(\mathbf{r}')\right\}G_2(\mathbf{r}-\mathbf{r}')ds' \\
&+\frac{1}{4\pi}\int_{\Gamma_2}\mathbf{M}_2(\mathbf{r}')\times\boldsymbol{\nabla}G_2(\mathbf{r}-\mathbf{r}')ds' \text{ pour } \mathbf{r}\in\Gamma_2
\end{aligned} \tag{26.II}$$

Les inconnues sont les grandeurs $\mathbf{M}_i = \mathbf{E}_i\times\mathbf{n}_i,\ i=1,2$ et $\mathbf{J}_i = \mathbf{n}_i\times\mathbf{H}_i,\ i=1,2$ représentant respectivement les densités de courants magnétiques et les densités de courants électriques. $G_i(\mathbf{r}-\mathbf{r}'),\ i=1,2$ est la fonction de Green qui s'écrit de la façon suivante dans le cas tridimensionnel [64, 51] :

$$G_i(\mathbf{r}) = \frac{e^{jk_i|\mathbf{r}|}}{|\mathbf{r}|},\ i=1,2 \text{ et } k_i = \omega\sqrt{\epsilon_i\mu_i}\ i=1,2.$$

La figure Fig 20.II représente un exemple de réseau bipériodique de périodes $\mathbf{T}_1$ et $\mathbf{T}_2$:

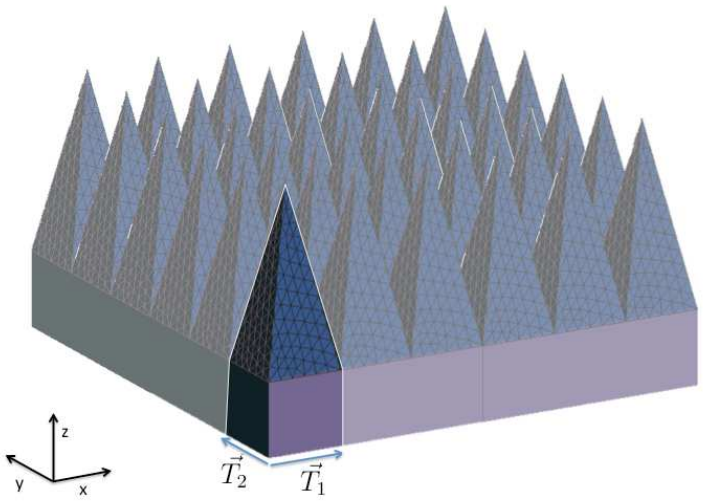


Fig 20.II Exemple de structure bi-périodique : revêtement de chambre anéchoïque dont la cellule de périodicité est constituée d'une pyramide de mousse en matériau absorbant ([51] page 7).

Écrivons l'équation intégrale pour $\mathbf{E}_1$ sur $\Gamma = \cup_{n,m}\Gamma_{n,m}$ sur la surface extérieure du réseau bi-périodique représentée Fig 20.II, $\Gamma_{n,m}$ étant la restriction de la surface extérieure à la cellule de base $\Gamma_{n,m}$. Le réseau est éclairé par une onde incidente pseudo-périodique [51] :

$$\begin{aligned}
&\frac{1}{2}\mathbf{E}_1(\mathbf{r}) = \mathbf{E}^{inc}(\mathbf{r}) - \sum_{n,m}\frac{1}{4\pi}\int_{\Gamma_{n,m}}\left\{j\omega\mu_1\mathbf{J}_1(\mathbf{r}') + \frac{j}{\omega\epsilon_1}\boldsymbol{\nabla}\boldsymbol{\nabla}'.\mathbf{J}_1(\mathbf{r}')\right\}G_1(\mathbf{r}-\mathbf{r}')ds' \\
&+\sum_{n,m}\frac{1}{4\pi}\int_{\Gamma_{n,m}}\mathbf{M}_1(\mathbf{r}')\times\boldsymbol{\nabla}G_1(\mathbf{r}-\mathbf{r}')ds'
\end{aligned} \tag{27.II}$$

Par un changement de variable on se ramène à une somme d'équations intégrales sur la cellule de base $\Gamma_{0,0}$ [51] :

$$\begin{aligned}&\frac{1}{2}\mathbf{E}_1(\mathbf{r}) = \mathbf{E}^{inc}(\mathbf{r})\\&-\sum_{n,m}\frac{1}{4\pi}\int_{\Gamma_{0,0}}\left\{j\omega\mu_1\mathbf{J}_1(\mathbf{r}^{'})+\frac{j}{\omega\epsilon_1}\boldsymbol{\nabla}\boldsymbol{\nabla}^{'}.\mathbf{J}_1(\mathbf{r}^{'}+n\mathbf{T}_1+m\mathbf{T}_2)\right\}G_1(\mathbf{r}-\mathbf{r}^{'}+n\mathbf{T}_1+m\mathbf{T}_2)ds^{'}\\&+\sum_{n,m}\frac{1}{4\pi}\int_{\Gamma_{0,0}}\mathbf{M}_1(\mathbf{r}^{'}+n\mathbf{T}_1+m\mathbf{T}_2)\times\boldsymbol{\nabla}G_1(\mathbf{r}-\mathbf{r}^{'}+n\mathbf{T}_1+m\mathbf{T}_2)ds^{'}\end{aligned} \tag{28.II}$$

La théorie de Floquet (cf section II.1) nous assure que les courants $\mathbf{J}_1$ et $\mathbf{M}_1$ sont pseudo périodiques. On a donc [51] :

$$\mathbf{J}_1(\mathbf{r}^{'}+n\mathbf{T}_1+m\mathbf{T}_2)=\mathbf{J}_1(\mathbf{r}^{'})e^{-j(n\mathbf{k}_i.\mathbf{T}_1+m\mathbf{k}_i.\mathbf{T}_2)},\ \mathbf{M}_1(\mathbf{r}^{'}+n\mathbf{T}_1+m\mathbf{T}_2)=\mathbf{M}_1(\mathbf{r}^{'})e^{-j(n\mathbf{k}_i.\mathbf{T}_1+m\mathbf{k}_i.\mathbf{T}_2)}$$

L'équation intégrale devient alors :

$$\begin{aligned}&\frac{1}{2}\mathbf{E}_1(\mathbf{r}) = \mathbf{E}^{inc}(\mathbf{r})-\frac{1}{4\pi}\int_{\Gamma_{0,0}}\left\{j\omega\mu_1\mathbf{J}_1(\mathbf{r}^{'})+\frac{j}{\omega\epsilon_1}\boldsymbol{\nabla}\boldsymbol{\nabla}^{'}.\mathbf{J}_1(\mathbf{r}^{'})\right\}\tilde{G}_1(\mathbf{r}-\mathbf{r}^{'})ds^{'}\\&+\frac{1}{4\pi}\int_{\Gamma_{0,0}}\mathbf{M}_1(\mathbf{r}^{'})\times\boldsymbol{\nabla}\tilde{G}_1(\mathbf{r}-\mathbf{r}^{'})ds^{'}\end{aligned} \quad [51] \tag{29.II}$$

Où $\tilde{G}_1$ est appelée fonction de Green 3D-bipériodique [51, 64]. Son expression est [51] :

$$\tilde{G}_1(\mathbf{r})=\sum_n\sum_m\frac{e^{-jk_1r_{nm}}}{r_{nm}}e^{-j(n\mathbf{k}_i.\mathbf{T}_1+m\mathbf{k}_i.\mathbf{T}_2)},\ \text{avec } r_{nm}=\sqrt{(n\mathbf{T}_1+x)^2+(m\mathbf{T}_2+y)^2+z^2}, \tag{30.II}$$

et $\mathbf{k}_i$ le vecteur d'onde de l'onde plane incidente. Ensuite, une fois la formulation écrite, on applique la méthode des moments en écrivant les courants $\mathbf{M}_i$ et $\mathbf{J}_i$ comme combinaisons linéaires des fonctions de base du type RWG (de Rao-Wilton-Glisson)[65], pour finalement obtenir le système linéaire à résoudre. On calcule donc le champ sur la surface pour une fréquence incidente donnée.

Remarque II.10. On peut combiner cette méthode avec des équations intégrales volumiques [66, 51] mais on ne pourra pas traiter de domaines constitués d'un matériau anisotrope.

II.5.3 Intérêts

- Le maillage surfacique demande beaucoup moins de degrés de liberté par rapport à un maillage volumique utilisé dans les autres méthodes numériques (FEM, FDTD).
- Les calculs s'effectuant aux interfaces des domaines, il n'est pas nécessaire d'introduire de conditions aux limites absorbantes pour modéliser l'atténuation des sources extérieures.
- L'expression de la périodicité se reporte uniquement sur la formulation et ne nécessite pas l'introduction de conditions aux limites supplémentaires.

– Cette méthode est très performante pour le calcul de la diffraction par des surfaces sélectives en fréquence sans épaisseur placées dans des empilements de couches diélectriques, notamment les radômes et l'exemple du méta-papier isolant (cf section I.2.5).
– Les matériaux peuvent être dispersifs : leurs paramètres peuvent dépendre de la fréquence de calcul.

II.5.4 Limitations

– La résolution des équations intégrales sur la structure périodique infinie nécessite le calcul de la fonction de Green bipériodique qui s'écrit sous la forme d'une série double et lentement convergente [51].
– La matrice du système linéaire est pleine et donc prend beaucoup de temps à inverser.
– On est limité à certains matériaux : linéaires homogènes isotropes et PEC.
– À cause de à la formulation des équations intégrales (26.II), on peut avoir des problèmes pour traiter les fréquences proches de zéro.

II.6 Conclusion

Dans cette section, nous avons sommairement présenté la plupart des méthodes existantes permettant de connaître les propriétés de bandes interdites d'une structure périodique. Nous avons également abordé leurs qualités et leurs défauts respectifs. Dans nos travaux, nous décidons de nous focaliser sur la méthode des éléments finis pour ses qualités évoquées en III.4, son aspect innovant et parce qu'elle utilise un environnement qui s'inscrit dans une suite de travaux de recherche [67, 31]. Dans la section suivante, nous allons donc traiter les différents aspects théoriques de sa mise en œuvre.

III La méthode des éléments finis

III.1 Introduction

La méthode des éléments finis, initialement utilisée pour résoudre numériquement des équations aux dérivées partielles, est bien connue[68] et permet, entre autre, de résoudre les équations de Maxwell fréquentielles. Notre démarche consiste à modifier la formulation classique des éléments finis par l'ajout de conditions aux limites particulières que nous allons détailler. Nous proposons notamment une description originale des conditions limites *Maître-esclave* dans la section III.3.5 et une nouvelle adaptation des conditions limites absorbantes pour les problèmes de modes propres dans la section III.3.3.

Les propriétés de bande interdite des structures périodiques peuvent être obtenues de plusieurs façons avec la méthode des éléments finis.

1. La première méthode consiste à résoudre le problème périodique en calculant le champ solution à chaque fréquence donnée.
2. La deuxième méthode calcule les modes propres solutions dans la cellule de base du réseau périodisé à l'aide des conditions de Floquet (cf section II.1).

La première méthode nécessite l'existence d'un demi-espace ouvert (cf remarque II.2) et aboutit à la résolution d'un système linéaire construit à partir de la formulation variationnelle des équations de Maxwell (cf section III.2). Dans cette thèse nous avons choisi de développer la deuxième méthode appliquée dans un premier temps à des problèmes de cavités métalliques puis étendue à des réseaux périodiques infinis. Cette méthode a déjà fait l'objet d'une étude dans [35] mais tous les algorithmes utilisés n'y sont pas clairement détaillés. Or seule une connaissance précise de l'ensemble des paramètres de résolution nous permet une interprétation exacte des résultats obtenus.

III.2 Formulation variationnelle

III.2.1 Introduction

La formulation variationnelle est le point de départ de la méthode de Galerkin[69] qui permet de calculer les solutions des équations de Maxwell dans le domaine fréquentiel. On a choisi une formulation générale qui s'inspire de [32]. On se place dans un domaine Ω hétérogène linéaire et on suppose que ϵ et μ sont des matrices hermitiennes définies positives. Les problèmes liés à l'existence des solutions ne sont pas abordés dans cette section puisqu'ils dépendent des conditions aux limites imposées aux bords du domaine $\partial\Omega$ qui ne sont pas précisées dans cette section afin de garder une vision globale.

Définition. On note $H(curl, \Omega)$ l'espace défini par :

$$H(curl, \Omega) = \{\mathbf{u} \in L^2(\Omega)^3,\ \boldsymbol{\nabla} \times \mathbf{u}\ \in L^2(\Omega)^3\}.$$

C'est un espace de Hilbert pour la norme

$$\sqrt{||\mathbf{u}||_{L^2(\Omega)^3} + ||\boldsymbol{\nabla} \times \mathbf{u}||_{L^2(\Omega)^3}}$$

et les solutions au sens des distributions[39] appartiennent à un sous espace de $H(curl, \Omega)$ noté $\mathcal{H}$ qui dépend des conditions limites imposées sur $\partial\Omega$.

Pour une meilleure visibilité on rappelle les équations de Maxwell en régime harmonique[38] :

$$\begin{cases} \text{Maxwell Faraday} \quad \boldsymbol{\nabla} \times \mathbf{E} - j\omega\mu\mathbf{H} = \mathbf{0} \\ \text{Maxwell Gauss} \quad \nabla.(\epsilon\mathbf{E}) = \rho \\ \text{Conservation du flux magnétique} \quad \nabla.(\mu\mathbf{H}) = 0 \\ \text{Maxwell Ampère} \quad \boldsymbol{\nabla} \times \mathbf{H} = -j\omega\epsilon\mathbf{E} + \mathbf{J} \end{cases} \tag{1.III}$$

On considère dans un premier temps qu'il n'y a pas de charge circulant dans le milieu ($\mathbf{J} = \mathbf{0}$). Le cas d'existence d'un vecteur de densité de courant non nul est traité section VI.2.1.2. En combinant Maxwell Ampère et Maxwell Faraday on obtient :

$$\boldsymbol{\nabla} \times (\frac{\boldsymbol{\nabla} \times \mathbf{E}}{\mu}) = \omega^2\epsilon\mathbf{E}, \quad \boldsymbol{\nabla} \times (\frac{\boldsymbol{\nabla} \times \mathbf{H}}{\epsilon}) = \omega^2\mu\mathbf{H}.$$

On multiplie ensuite par une fonction test $\boldsymbol{\varphi} \in \mathcal{H}$ et on intègre sur un domaine Ω de frontière $\partial\Omega$ qui a une normale sortante $\boldsymbol{\nu}$:

$$\int_\Omega \boldsymbol{\nabla} \times (\frac{\boldsymbol{\nabla} \times \mathbf{E}}{\mu}).\boldsymbol{\varphi} = \int_\Omega \omega^2\epsilon\mathbf{E}.\boldsymbol{\varphi}, \quad \int_\Omega \boldsymbol{\nabla} \times (\frac{\boldsymbol{\nabla} \times \mathbf{H}}{\epsilon}).\boldsymbol{\varphi} = \int_\Omega \omega^2\mu\mathbf{H}.\boldsymbol{\varphi} \tag{2.III}$$

Remarque III.1. Les fonctions $\boldsymbol{\varphi}$ et $\boldsymbol{E}$ sont à valeurs dans $\mathbb{R}^3$ mais les transformations que nous aborderons dans le cadre des structures périodiques section III.3.5 introduirons des déphasages complexes dans la formulation (2.III). Dans ce cas l'espace $\mathcal{H}$ sera muni d'une

structure hermitienne et la formulation (2.III) s'écriera :

$$\int_\Omega \boldsymbol{\nabla} \times (\frac{\boldsymbol{\nabla} \times \mathbf{E}}{\mu}).\boldsymbol{\varphi}^* = \int_\Omega \omega^2 \epsilon \mathbf{E}.\boldsymbol{\varphi}^*, \quad \int_\Omega \boldsymbol{\nabla} \times (\frac{\boldsymbol{\nabla} \times \mathbf{H}}{\epsilon}).\boldsymbol{\varphi}^* = \int_\Omega \omega^2 \mu \mathbf{H}.\boldsymbol{\varphi}^* \tag{3.III}$$

Où $\boldsymbol{\varphi}^*$ représente le complexe conjugué de $\boldsymbol{\varphi}$. Cette formulation sera rappelée quand les déphasages complexes seront introduit section III.3.5. Pour le moment les fonctions $\boldsymbol{\varphi}$ et $\boldsymbol{E}$ restent à valeur dans $\mathbb{R}^3$, mais pour ne pas perdre en généralité nous gardons la notation complexe.

Sachant que $\nabla.(\mathbf{A} \times \mathbf{B}) = (\boldsymbol{\nabla} \times \mathbf{A}).\mathbf{B} - (\boldsymbol{\nabla} \times \mathbf{B}).\mathbf{A}$ et que $\int_\Omega \nabla.\mathbf{A} = \int_{\partial\Omega} \mathbf{A}.\boldsymbol{\nu}$, on obtient :

$$\int_\Omega \boldsymbol{\nabla} \times \boldsymbol{\varphi}^*.(\frac{\boldsymbol{\nabla} \times \mathbf{E}}{\mu}) = \int_\Omega \omega^2 \epsilon \mathbf{E}.\boldsymbol{\varphi}^* - \int_{\partial\Omega} \big((\frac{\boldsymbol{\nabla} \times \mathbf{E}}{\mu}) \times \boldsymbol{\varphi}^*\big).\boldsymbol{\nu}$$

et

$$\int_\Omega \boldsymbol{\nabla} \times \boldsymbol{\varphi}^*.(\frac{\boldsymbol{\nabla} \times \mathbf{H}}{\epsilon}) = \int_\Omega \omega^2 \mu \mathbf{H}.\boldsymbol{\varphi}^* - \int_{\partial\Omega} \big((\frac{\boldsymbol{\nabla} \times \mathbf{H}}{\epsilon}) \times \boldsymbol{\varphi}^*\big).\boldsymbol{\nu}$$

Comme $\mathbf{A}.(\mathbf{B} \times \mathbf{C}) = \mathbf{C}.(\mathbf{A} \times \mathbf{B})$ on obtient :

$$\int_\Omega \boldsymbol{\nabla} \times \boldsymbol{\varphi}^*.(\frac{\boldsymbol{\nabla} \times \mathbf{E}}{\mu}) = \int_\Omega \omega^2 \epsilon \mathbf{E}.\boldsymbol{\varphi}^* - \int_{\partial\Omega} \big(\boldsymbol{\nu} \times (\frac{\boldsymbol{\nabla} \times \mathbf{E}}{\mu})\big).\boldsymbol{\varphi}^*$$

et

$$\int_\Omega \boldsymbol{\nabla} \times \boldsymbol{\varphi}^*.(\frac{\boldsymbol{\nabla} \times \mathbf{H}}{\epsilon}) = \int_\Omega \omega^2 \mu \mathbf{H}.\boldsymbol{\varphi}^* - \int_{\partial\Omega} \big(\boldsymbol{\nu} \times (\frac{\boldsymbol{\nabla} \times \mathbf{H}}{\epsilon})\big).\boldsymbol{\varphi}^*$$

En utilisant les notations $\epsilon_r = \frac{\epsilon}{\epsilon_0}$, la perméabilité magnétique relative $\mu_r = \frac{\mu}{\mu_0}$, le nombre d'ondes $k_0 = \frac{\omega}{C} = \omega\sqrt{\varepsilon_0 \mu_0}$ et l'impédence d'onde caractéristique du vide $Z_0 = \sqrt{\frac{\mu_0}{\epsilon_0}}$, les équations peuvent alors s'écrire[32] :

$$\int_\Omega \boldsymbol{\nabla} \times \boldsymbol{\varphi}^*.\frac{\boldsymbol{\nabla} \times \mathbf{E}}{\mu_r} = k_0^2 \int_\Omega \epsilon_r \mathbf{E}.\boldsymbol{\varphi}^* - \int_{\partial\Omega} \big(\boldsymbol{\nu} \times (\frac{\boldsymbol{\nabla} \times \mathbf{E}}{\mu_r})\big).\boldsymbol{\varphi}^* \tag{4.III}$$

et

$$\int_\Omega \boldsymbol{\nabla} \times \boldsymbol{\varphi}^*.\frac{\boldsymbol{\nabla} \times \mathbf{H}}{\epsilon_r} = k_0^2 \int_\Omega \mu_r \mathbf{H}.\boldsymbol{\varphi}^* - \int_{\partial\Omega} \big(\boldsymbol{\nu} \times (\frac{\boldsymbol{\nabla} \times \mathbf{H}}{\epsilon_r})\big).\boldsymbol{\varphi}^* \tag{5.III}$$

Remarque III.2. En remplaçant les termes de bord à l'aide des équations de Maxwell, on a une formulation variationnelle équivalente aux formulations précédentes (4.III) et (5.III). En champs $\mathbf{H}$:

$$\int_\Omega \boldsymbol{\nabla} \times \boldsymbol{\varphi}^*.\frac{\boldsymbol{\nabla} \times \mathbf{H}}{\epsilon_r} = k_0^2 \int_\Omega \mu_r \mathbf{H}.\boldsymbol{\varphi}^* + j\frac{k_0}{Z_0} \int_{\partial\Omega} (\boldsymbol{\nu} \times \mathbf{E}).\boldsymbol{\varphi}^*$$

Et en champs $\mathbf{E}$[42] :

$$\int_\Omega \boldsymbol{\nabla} \times \boldsymbol{\varphi}^*.\frac{\boldsymbol{\nabla} \times \mathbf{E}}{\mu_r} = k_0^2 \int_\Omega \epsilon_r \mathbf{E}.\boldsymbol{\varphi}^* - jk_0 Z_0 \int_{\partial\Omega} (\boldsymbol{\nu} \times \mathbf{H}).\boldsymbol{\varphi}^* \tag{6.III}$$

Ainsi le problème variationnel pour une résolution d'inconnue $\mathbf{E}$ devient :

$$\begin{cases} \forall \boldsymbol{\varphi} \in \mathcal{H}, \quad a(\mathbf{E}, \boldsymbol{\varphi}) = k_0^2 * m(\mathbf{E}, \boldsymbol{\varphi}) \\ \text{avec } a(\mathbf{E}, \boldsymbol{\varphi}) = \displaystyle\int_\Omega \boldsymbol{\nabla} \times \boldsymbol{\varphi}^* . \frac{\boldsymbol{\nabla} \times \mathbf{E}}{\mu_r} + \int_{\partial\Omega} \left(\boldsymbol{\nu} \times (\frac{\boldsymbol{\nabla} \times \mathbf{E}}{\mu_r})\right) . \boldsymbol{\varphi}^*, \\ \text{et } m(\mathbf{E}, \boldsymbol{\varphi}) = \displaystyle\int_\Omega \epsilon_r \mathbf{E} . \boldsymbol{\varphi}^* \end{cases} \tag{7.III}$$

Remarque III.3. Dans les problèmes de modes propres que nous traitons, les conditions aux limites sont soit des conditions de type métal parfait (cf sectionIII.3.2), soit des conditions périodiques (cf section 1.V). Ces deux types de conditions aux limites annulent les termes d'intégration de bords sur $\partial\Omega$ dans l'équation (7.III) (cf remarque III.11).

Pour la résolution en champ $\mathbf{H}$ il suffit d'inverser la permittivité relative ϵ_r et la perméabilité relative μ_r dans le problème variationnel (7.III).

Dans la suite de la thèse la formulation variationnelle utilisée est celle définie par l'équation (7.III) d'inconnue $\mathbf{E}$.

III.2.2 Assemblage des matrices via la méthode de Galerkin

III.2.2.1 Introduction

Dans cette section nous indiquons comment se construisent les matrices du problème éléments finis avec la méthode de Galerkin à partir de la formulation variationnelle établie équation (7.III) pour les milieux non dispersifs.

III.2.2.2 Construction du problème variationnel approché

Nous présentons ici une approche euristique de la construction des matrices éléments finis issues du problème variationnel défini en (7.III). Un algorithme général de construction des matrices est ici donné comme repère de tous les problèmes que nous allons aborder. Les solutions sont cherchées dans un espace de dimension finie $\mathcal{H}_h$ qui est dense dans l'espace $\mathcal{H}$ quand le pas du maillage h tend vers 0. Dans Ω, le champ solution $\mathbf{E} \in \mathcal{H}$ est approché par $\tilde{\mathbf{E}} \in \mathcal{H}_h$ qui est une combinaison linéaire de fonctions de base d'arêtes de Nedelec[70] d'ordre 0 (cf annexe A.3) $(\mathbf{W}_1, \ldots, \mathbf{W}_p, \ldots \mathbf{W}_{DL}) = \underline{\mathbf{W}}$ où DL désigne le nombre d'inconnues. La discrétisation de $\bar{\Omega}$ est nomé Ω_h. Ainsi si on note par e_p l'inconnue représentant la circulation de $\tilde{\mathbf{E}}$ sur chaque arête $\mathbf{Ar}_p$, on a :

$$\mathbf{E} \simeq \tilde{\mathbf{E}} = \sum_{p=1}^{DL} e_p \mathbf{W}_p = (e_1, \ldots, e_{DL})^T \underline{\mathbf{W}}, \;\; e_p = \int_{\mathbf{Ar}_p} \tilde{\mathbf{E}} \tag{8.III}$$

Le problème variationnel approché dans $\mathcal{H}_h$ s'écrit alors :
Trouver $e_p, \;\; p \in [1 \ldots DL]$ tel que

$$a(\tilde{\mathbf{E}}, \mathbf{W}_q) = k_0^2 * m(\tilde{\mathbf{E}}, \mathbf{W}_q), \quad q = 1 \dots DL$$
$$\Rightarrow \sum_{p=1}^{DL} e_p * a(\mathbf{W}_p, \mathbf{W}_q) = k_0^2 \sum_{p=1}^{DL} e_k * m(\mathbf{W}_p, \mathbf{W}_q), \quad q = 1 \dots DL. \tag{9.III}$$

où $a(,)$ et $m(,)$ sont les formes bilinéaires définies équation (7.III).
En posant $\mathbf{X} = (e_1, \dots e_{DL})$ on obtient :

$$A\mathbf{X} = k_0^2 M\mathbf{X}$$
$$A = (A)_{qp} = a(\mathbf{W}_p, \mathbf{W}_q)_{q,p\in[1\dots DL]}, \quad M = (M)_{qp} = m(\mathbf{W}_p, \mathbf{W}_q)_{q,p\in[1\dots DL]}. \tag{10.III}$$

Les modes propres sont les solutions $(k_0^2, \mathbf{X})$ d problème (10.III). Lorsqu'il y a une source explicite introduite sur $\partial\Omega$, on peut, pour une valeur de k_0^2 connue, résoudre le problème (10.III) qui s'écrit alors comme le système linéaire :

$$(A - k_0^2 M)\mathbf{X} = \mathbf{E}^{inc} \tag{11.III}$$

Où $\mathbf{E}^{inc}$ est un vecteur qui contient les termes du champ incidents. Plus de détails sont donnés sur la construction de $\mathbf{E}^{inc}$, section VI.2.2. Les matrices A et M y sont modifiées via l'ajout de conditions aux limites que nous allons introduire dans la section III.3.

Remarque III.4. Le choix des fonctions Nedelec d'ordre 0 est fait pour des raisons de facilité d'implémentation. Mais on pourrait très bien monter en ordre en choisissant des éléments de Nedelec d'ordre 1[71]. Ce qui aurait pour conséquence de passer à 20 fonctions de base par tétraèdre et d'améliorer la précision par rapport à l'ordre 0 pour un même maillage donné. Cependant, comme le nombre de fonction de base est plus grand pour l'ordre 1, le temps d'assemblage des matrices élémentaires (cf annexe A.4) est aussi augmenté.

Remarque III.5. Les termes de bords :

$$\int_{\partial\Omega} (\boldsymbol{\nu} \times (\frac{\boldsymbol{\nabla} \times \mathbf{W}_p}{\mu_r})).\mathbf{W}_q$$

Sont nuls si l'intersection des supports de $\mathbf{W}_p$ et de $\mathbf{W}_q$ n'appartient pas à $\partial\Omega$. De plus l'intersection des supports de $\mathbf{W}_p$ et de $\mathbf{W}_q$ doit être contenue dans un même triangle. Pour les calculs à l'intérieur du domaine Ω, les termes d'intégrales surfaciques vont disparaître. En effet lorsqu'on considère deux tétraèdres collés l'un à l'autre les deux normales respectives de la face commune ont des sens opposés et comme la composante tangentielle des fonctions de base est continue à l'interface, les éléments surfaciques calculés s'annulent deux à deux.

Remarque III.6. Lorsque ϵ_r et μ_r^{-1} sont des matrices hermitiennes définies positives la matrice assemblée M est toujours définie positive. En effet, d'après la décomposition de Schur[72] $\epsilon_r = UDU^*$, avec D une matrice diagonale réelle définie positive et U une matrice unitaire.

On pose alors :

$$\sqrt{\epsilon_r} = U\sqrt{D},$$

et on définit le produit scalaire :

$$< \mathbf{W}_p, \mathbf{W}_q >= (\epsilon_r \mathbf{W}_p)^* \mathbf{W}_q = (\sqrt{\epsilon_r}\mathbf{W}_p)^*(\sqrt{\epsilon_r}\mathbf{W}_q)$$

Ensuite on pose $\underline{\mathbf{W}} = (\mathbf{W}_1, \dots \mathbf{W}_{DL})$, on peut définir le produit tensoriel :

$$(\underline{\mathbf{W}}\tilde{\otimes}\underline{\mathbf{W}})_{pq} =< \mathbf{W}_p, \mathbf{W}_q >= (\sqrt{\epsilon_r}\mathbf{W}_p)^*(\sqrt{\epsilon_r}\mathbf{W}_q).$$

Un élément de $\mathcal{H}_h$ s'écrit alors $\mathbf{X}^T\underline{\mathbf{W}}$ et on a :

$$\mathbf{X}^* M \mathbf{X} = \mathbf{X}^*(\int_\Omega \underline{\mathbf{W}}\tilde{\otimes}\underline{\mathbf{W}})\mathbf{X} = \int_\Omega (\sqrt{\epsilon_\mathbf{r}}\underline{\mathbf{W}}^T\mathbf{X})^*(\sqrt{\epsilon_\mathbf{r}}\underline{\mathbf{W}}^T\mathbf{X}) > 0$$

S'il n'y a pas de terme de bord la matrice A est semi-définie positive parce qu'elle peut s'écrire à l'aide d'un produit tensoriel $\boldsymbol{\nabla} \times \underline{\mathbf{W}}\tilde{\otimes}\boldsymbol{\nabla} \times \underline{\mathbf{W}}$ (défini avec μ^{-1}), mais son noyau est non nul. En effet la décomposition discrète de Helmholtz[73, 74] nous assure qu'il existe un vecteur $\mathbf{U} \in (H^1(\Omega))^3$ dérivant d'un potentiel vecteur $\mathbf{V} \in H(curl, \Omega)$ et un potentiel scalaire $\phi \in H^1(\Omega)$ tel que :

$$\tilde{\mathbf{E}} = \mathbf{U} + \boldsymbol{\nabla}\phi, \text{ tel que } \boldsymbol{\nabla}.\mathbf{U} = 0 \text{ et } \mathbf{U} = \boldsymbol{\nabla} \times \mathbf{V}.$$

Le noyau de A contient alors les champs $\tilde{\mathbf{E}} = \mathbf{X}^T\underline{\mathbf{W}}$ qui ont une composante rotationnelle nulle (i.e $\mathbf{U} \equiv \mathbf{0}$). Cependant, comme M est inversible, la matrice $A - \sigma M$ l'est aussi, mais elle est rarement définie positive lorsque σ est positif réel.

Le calcul numérique des matrices élémentaires et des fonctions de base est résumé en annexe A.4.

III.2.3 Conclusion

Nous avons présenté dans cette section comment se modélisent les problèmes éléments finis quand les milieux traités ne sont pas dispersifs. Nous allons aborder dans la section suivante III.3 les différentes conditions limites associées à la formulation (10.III). Des méthodes pour traiter les milieux dispersifs sont abordées en perspective dans les sections VI.2 et VI.2.2.

III.3 Conditions aux limites pour la résolution mode propre

III.3.1 Introduction

Les conditions limites ont des conséquences importantes sur la formulation variationnelle (7.III) et sur l'assemblage des matrices (cf section III.2.2). Nous commençons d'abord par énoncer des conditions aux limites classiques section III.3.2 pour ensuite développer des conditions limites plus spécifiques et adaptées à nos problèmes sections III.3.3, III.3.5. Dans les perspectives, section VI.2.2 les conditions aux limites avec présence d'une source explicite sont détaillées.

III.3.2 Conducteur parfait

La condition de conducteur parfait est traduite par le fait que la composante tangentielle du champs électrique est nulle sur le bord domaine. Ceci s'explique parce que le champ dans le milieu métallique est nul et qu'il y a continuité du champ tangent à l'interface (cf annexe A.1). Nous allons voir que cette propriété peut aussi s'expliquer grâce à la loi d'Ohm (cf annexe A.5) :

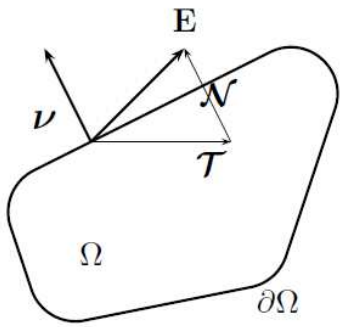


Fig 1.III Décomposition du champs selon les composantes normale et tangentielle.

En adoptant les conventions définies Fig 1.III et en appliquant la loi d'ohm (cf Annexe A.5) sur la surface extérieur $\partial\Omega$, on a alors :

$$J|_{\partial\Omega} = \sigma_e \mathbf{E}|_{\partial\Omega}$$

En décomposant le champ $\mathbf{E}$ selon sa composante normale et sa composante tangentielle, on en déduit :

$$\mathbf{E}|_{\partial\Omega} = E_T \boldsymbol{\mathcal{T}} + E_N \boldsymbol{\mathcal{N}} \Rightarrow E_T = \mathbf{E}|_{\partial\Omega}.\boldsymbol{\mathcal{T}} = \frac{\mathbf{J}|_{\partial\Omega}.\boldsymbol{\mathcal{T}}}{\sigma_e}$$

Par définition d'un conducteur parfait, σ_e est supposée infinie et $\mathbf{J}|_{\partial\Omega}.\boldsymbol{\mathcal{T}}$ fini, donc $E_T = 0$. En considérant le vecteur $\boldsymbol{\nu}$ normal à la surface on remarque que $\boldsymbol{\nu} = \mathbf{N}$ et donc la relation de conducteur parfait peut aussi s'écrire :

$$\mathbf{E} \times \boldsymbol{\nu} = E_T \boldsymbol{\mathcal{T}} \times \boldsymbol{\nu} = \mathbf{0}$$

D'après l'équation (8.III) le champ approché $\tilde{\mathbf{E}} \in \mathcal{H}_h$ est une combinaison linéaire de fonctions de base d'arêtes de Nedelec[70] d'ordre 0 $\mathbf{W}_p,\ p \in [1, \dots DL]$ où DL désigne le nombre d'inconnues et e_p représente la circulation de $\tilde{\mathbf{E}}$ sur chaque arête $\mathbf{Ar}_p$:

$$\mathbf{E} \simeq \tilde{\mathbf{E}} = \sum_{p=1}^{DL} e_p \mathbf{W}_p,\ e_p = \int_{\mathbf{Ar}_p} \tilde{\mathbf{E}}$$

Remarque III.7. **Pour les conditions limites du type conducteur parfait, les inconnues sont nulles sur les arêtes situées sur le matériaux PEC**. En effet soit $\mathbf{a}_i$ et $\mathbf{a}_j$ les sommets de l'arête $\mathbf{Ar}_p$ située sur une face PEC. Le chemin de $\mathbf{a}_i$ à $\mathbf{a}_j$ est paramétré par $\gamma(t) = t\mathbf{a}_j + (1-t)\mathbf{a}_i$ et l'intégrale le long de ce chemin du champs $\tilde{\mathbf{E}}$ s'écrit :

$$\int_{\mathbf{Ar}_p} \tilde{\mathbf{E}} = \int_0^1 \tilde{\mathbf{E}}(\gamma(t)).\overrightarrow{\mathbf{a}_i\mathbf{a}_j}\ dt = e_p = 0.$$

III.3.3 Couches absorbantes (PML) pour le solveur mode propre

III.3.3.1 Présentation et formulation

Les zones absorbantes parfaitement adaptées appelées PML sont utilisées essentiellement pour limiter la réflection des ondes solutions au bord du domaine de calcul et ainsi simuler un espace libre. Il existe une autre modélisation possible des conditions PML qui est abordée dans la section suivante III.3.4, mais nous avons choisi la formulation proposée dans [75] parce qu'elle est plus simple à mettre en œuvre et que le réglage des paramètres est plus intuitif. L'idée est de multiplier la perméabilité et la permittivité par un tenseur γ défini en (12.III) et permettant de simuler l'absorption d'une onde plane dans le milieu PML :

$$\gamma = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix} \tag{12.III}$$

On considère alors un milieu 2d (Oxz) et une zone PML absorbante dans la direction Oz Fig 2.III. Pour annuler la première réflection entre l'air et le milieu PML alors les coefficients devront vérifier [75] :

$$a = b = \frac{1}{c} \tag{13.III}$$

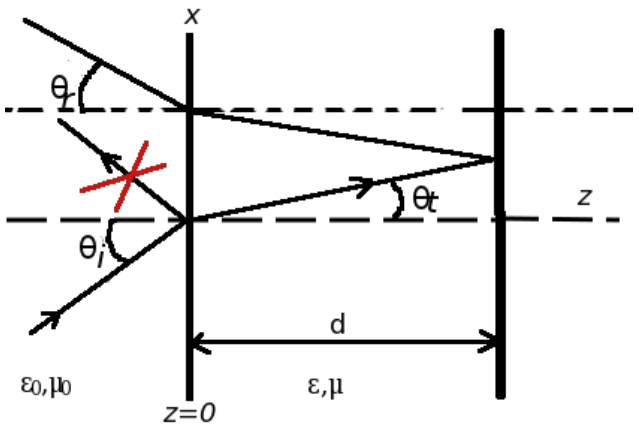


Fig 2.III Parcours d'une onde dans un milieu absorbant. La relation (13.III) annule la première réflexion.

Fig 2.III, on considère une onde plane qui se propage dans le milieu absorbant. Le champ incident a pour expression :

$$\mathbf{E}_i(x,z) = \mathcal{E} e^{-jk_0(\cos(\theta_i)z + \sin(\theta_i)x)}$$

Si on pose $a = \alpha - j\beta$, l'onde plane transmise dans le milieu PML a pour expression [75] :

$$\mathbf{E}_t(x,z) = \mathcal{E} e^{-k_0\beta\cos(\theta_t)z} e^{-jk_0(\sin(\theta_t)x + \alpha\cos(\theta_t z))} \quad \text{avec } k_0 = \omega\sqrt{\epsilon_0\mu_0} \tag{14.III}$$

L'absorption est donc contrôlée par le coefficient β. Pour que la zone PML ne perturbe pas le champ solution dans l'air, il est nécessaire que β ne soit pas choisi trop grand. Ainsi la variation d'indice entre l'air et la PML ne sera pas importante et on peut considérer Fig 2.III que $\theta_i = \theta_t = \theta_r = \theta$. Par conséquent d'après l'équation (14.III), le coefficient $|R(\theta)|$ de réflexion l'onde après le passage de l'onde dans la zone PML est égal à :

$$|R(\theta)| \simeq |Q(\theta)| e^{-2\beta d k_0 \cos(\theta)}. \tag{15.III}$$

$|Q(\theta)|$ étant le coefficient de réflexion de la surface qui termine la PML. Si cette surface est du métal on peut alors prendre $|Q(\theta)| = 1$. Dans ce cas nous pouvons alors déterminer la constante β qui contrôle l'absorption dans l'équation (15.III) afin d'obtenir un coefficient de réflexion inférieur à 1% :

$$-2\beta d k_0 \cos(\theta) < \ln(10^{-2}) \Rightarrow \beta > \frac{\ln(10)}{k_0 d \cos(\theta)}. \tag{16.III}$$

Il apparaît donc que le coefficient α dépend des directions d'ondes incidentes et qu'il est indéterminé pour les ondes rasantes $\theta = \pi/2$ et quand la zone PML est trop petite (d trop petit).

III.3.3.2 Cas particuliers pour le solveur mode propre

L'onde incidente absorbée à une longueur d'onde et une fréquence respectivement égales

à :

$$\lambda_i = \frac{2\pi}{k_0}, \quad F_i = \frac{Ck_0}{2\pi}$$

Lors d'un calcul de modes propres, on obtient plusieurs fréquences solutions. On doit alors choisir un coefficient β(16.III) de sorte que toutes solutions soient absorbées. Or d'après (16.III) pour que l'absorption soit suffisante il faut :

$$\beta > \frac{\ln(10)}{k_0 d\cos(\theta)} = \frac{C\ln(10)}{2\pi F_i d\cos(\theta)}$$

Soit F_{im} la fréquence minimum absorbée et k_{im} le nombre d'onde correspondant. Si on pose :

$$\beta = \frac{C\ln(10)}{2\pi F_{im} d\cos(\theta)} = \frac{\ln(10)}{k_{im} d\cos(\theta)} \tag{17.III}$$

Alors toutes les fréquences supérieurs à F_{im} seront absorbées. Pour que l'absorption soit suffisamment efficace et que le coefficient β ne soit pas choisi trop grand pour éviter une absorption trop efficace qui pourrait perturber l'ensemble des solutions, la longueur de la PML d doit être plus grande qu'une longueur notée d_{im} définie de la façon suivante :

$$d \geq d_{im} = \frac{\lambda_{im}}{4} = \frac{C}{4F_{im}} = \frac{2\pi}{4k_{im}} \tag{18.III}$$

Il est aussi possible d'appliquer ce type de PML pour des problèmes à pulsation fixe mais il existe une formulation plus adaptée qui consiste à complexifier les variables dans les zones PML [76]. On en énonce rapidement le principe dans la section suivante III.3.4.

III.3.4 Autre formulation des PML : complexification des coordonnées

Une façon classique de coder les conditions PML a été mise en œuvre dans [76]. Si une onde plane se propage suivant l'axe Ox son expression est :

$$\mathbf{E}(x,t) = \mathcal{E}e^{j(-\omega t - k_{ix}x)}.$$

On veut que l'onde s'atténue dans la direction Ox. Pour ce faire on prolonge analytiquement la variable d'espace x dans l'espace complexe :

$$\tilde{x} = x - jf(x).$$

L'expression de l'onde plane dans cet espace devient alors :

$$\mathbf{E}(\tilde{x},t) = \mathcal{E}e^{j(-\omega t - k_i x) - k_{ix} f(x)}. \tag{19.III}$$

On voit bien apparaître une atténuation selon l'axe des x à condition que la fonction f soit croissante. On veut que cette atténuation ne dépende pas de l'onde incidente. On va donc

choisir

$$f(x) = \frac{1}{k_{ix}} g_x(x).$$

Où g_x est une fonction croissante.

D'autre part on peut calculer la dérivée par rapport à cette nouvelle variable :

$$\frac{\partial \tilde{x}}{\partial x} = 1 - j\frac{g'_x(x)}{k_{ix}} \Rightarrow \frac{\partial}{\partial \tilde{x}} = \frac{1}{1 - j\frac{g'_x(x)}{k_{ix}}} \frac{\partial}{\partial x} = \frac{1}{S_x} \frac{\partial}{\partial x}.$$

On étend alors ce résultat aux trois directions de l'espace et on réécrit les équations de Maxwell Harmonique en introduisant l'opérateur rotationel modifié :

$$\tilde{\nabla}\times = \begin{pmatrix} 0 & \frac{-1}{S_z}\partial z & \frac{1}{S_y}\partial y \\ \frac{1}{S_z}\partial z & 0 & \frac{-1}{S_x}\partial x \\ \frac{-1}{S_y}\partial y & \frac{1}{S_x}\partial x & 0 \end{pmatrix} \Rightarrow \tilde{\nabla}\times = N((\nabla\times)L)$$

Avec

$$L = \begin{pmatrix} S_x & 0 & 0 \\ 0 & S_y & 0 \\ 0 & 0 & S_z \end{pmatrix} \ , \ N = \frac{1}{S_x S_y S_z} L.$$

Pour plus de clarté, on rappelle les équations de Maxwell :

$$\begin{cases} \text{Maxwell Faraday} \ \ \boldsymbol{\nabla} \times \mathbf{E} - j\omega\mu\mathbf{H} = \mathbf{0}, \\ \text{Maxwell Ampère} \ \ \boldsymbol{\nabla} \times \mathbf{H} = -j\omega\epsilon\mathbf{E}. \end{cases}$$

En posant :

$$\underline{\mathbf{E}} = L\mathbf{E}, \ \underline{\mathbf{H}} = L\mathbf{H}, \ N^{-1}L^{-1} = T = \begin{pmatrix} \frac{S_y S_z}{S_x} & 0 & 0 \\ 0 & \frac{S_x S_z}{S_y} & 0 \\ 0 & 0 & \frac{S_x S_y}{S_z} \end{pmatrix}$$

Les équations de Maxwell deviennent :

$$\boldsymbol{\nabla} \times \underline{\mathbf{E}} = j\omega[T\mu]\underline{\mathbf{H}} \ , \ \boldsymbol{\nabla} \times \underline{\mathbf{H}} = -j\omega[T\epsilon]\underline{\mathbf{E}}$$

$$\Rightarrow [T\mu]^{-1}\boldsymbol{\nabla} \times \underline{\mathbf{E}} = j\omega\underline{\mathbf{H}} \Rightarrow \boldsymbol{\nabla} \times ([T\mu]^{-1}\boldsymbol{\nabla} \times \underline{\mathbf{E}}) = \omega^2[T\epsilon]\underline{\mathbf{E}}.$$

On voit dans cette écriture, que pour implémenter ces PML, il faut multiplier la permittivité et la perméabilité par la matrice T dont les coefficients varies en fonction de la position spatiale. Cette formulation permet de construire une zone parfaitement adaptée et progressivement absorbante contrairement aux PML de la section III.3.3 qui sont constantes dans tous le milieu absorbant. Cependant une zone absorbante constante permet de définir plus facilement une fréquence minimum d'absorption dans le cas d'une résolution "mode propre".

III.3.5 Conditions aux limites *Maître-esclave*

III.3.5.1 Introduction

Les conditions aux limites de Floquet pour la méthode des éléments finis[35] sont aussi connues sous le nom de conditions *Maître-esclave*. Le domaine d'étude est réduit à la cellule de base Ω qui est contenue dans un parallélèpipède rectangle représenté Fig 3.III.

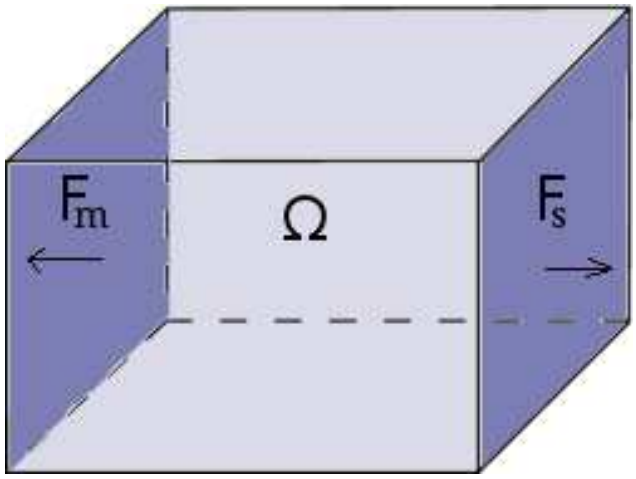


Fig 3.III Représentation du domaine Ω contenant le motif de base de la structure périodique.

Les conditions limites "*Maître-esclave*" se traduisent par un déphasage entre les éléments des interfaces en regard F_m et F_s de Ω dans les directions de périodicité.

III.3.5.2 Élimination des inconnues

On rappelle comme introduit en (8.III) que :

$$\mathbf{E} \simeq \tilde{\mathbf{E}} = \sum_{p=1}^{DL} e_p \mathbf{W}_p, \; e_p = \int_{\mathbf{Ar}_p} \tilde{\mathbf{E}} \tag{20.III}$$

Les conditions *Maître-esclave* appliquées dans la cellule de base traduisent une périodicité infinie selon la direction de $\mathbf{T}$ Fig 4.III. Le déphasage entre chaque éléments des faces en regard est égal à $\phi = \mathbf{k}_i.\mathbf{T}$. Où $\mathbf{k}_i$ est le vecteur d'onde de l'onde incidente sur la structure périodique. Soit alors e_m la circulation du champ $\tilde{\mathbf{E}}$ sur une arête $\mathbf{Ar}_m$. Sur l'arête $\mathbf{Ar}_s$ en vis à vis sur la Fig 4.III, la circulation du champ $\tilde{\mathbf{E}}$ a pour valeur $e_s = e^{-j(\mathbf{k}_i.\mathbf{T})} e_m$:

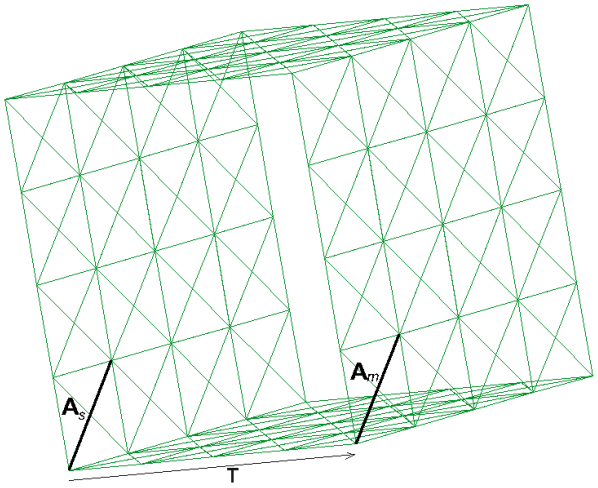


Fig 4.III Déphasage entre les éléments Maître-esclave.

Autrement dit, on supprime l'inconnue associée à l'arête esclave $\mathbf{Ar}_s$. Pour pouvoir appliquer ces relations de déphasage, le maillage surfacique de Ω_h doit avoir des éléments identiques à une translation près $\mathbf{T}$ sur les faces "Maître-esclave" en regard comme sur la Fig 4.III.

III.3.5.3 Déphasages multiples

Soit $\mathbf{T}_1$ le vecteur de translation de la face esclave $1(\mathrm{F_{1s}}$ sur le dessin) à la face maître $1(\mathrm{F_{1m}})$ et $\mathbf{T}_2$ le vecteur de translation de la face esclave $2(\mathrm{F_{2s}})$ à la face maître $2(\mathrm{F_{2m}})$.

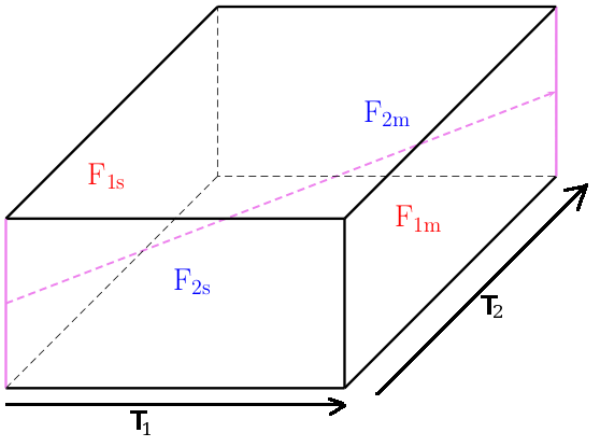


Fig 5.III Conditions aux limites "*Maître-esclave*"
sur les faces du parallélépipède contenant la cellule de base.

On considère une structure bipériodique contenue dans un parallélépipède représenté Fig 5.III. Les inconnues e_{1s}(respectivement e_{2s}) sur les arêtes de la face $\mathrm{F_{1s}}$ (respectivement $\mathrm{F_{2S}}$) sont égales à $e^{-j\phi_1} * e_{1m}$ (respectivement $e^{-j\phi_2} * e_{2m}$), où e_{1m} (respectivement e_{2m}) sont les inconnues du champ sur la face $\mathrm{F_{1m}}$ (respectivement $\mathrm{F_{2m}}$). Les déphasages ϕ_1 et ϕ_2 sont respectivement égaux à $\mathbf{k}_i.\mathbf{T}_1$ et $\mathbf{k}_i.\mathbf{T}_2$. Les inconnues de $\mathrm{F_{1s}} \cap \mathrm{F_{2s}}$ sont liées aux inconnues de $\mathrm{F_{1m}} \cap \mathrm{F_{2m}}$ par un déphasage égal à $-(\phi_1 + \phi_2)$ [42]. Il y a plusieurs possibilités pour définir les relations entre les inconnues lorsqu'une arête se trouve sur $\mathrm{F_{1s}} \cap \mathrm{F_{2s}}$ (zone violette sur

la figure Fig 5.III), mais la façon que nous présentons est celle qui permet déliminer le plus d'inconnues et de ne pas confondre d'éléments.

Remarque III.8. Si une arête est à l'intersection d'une face 'maître' et d'une face 'esclave' alors par convention c'est une arête 'Slave' puisque l'on veut éliminer le maximum d'inconnues. Donc, par convention, les arêtes à l'intersection des faces maître et esclave $\mathrm{F}_{1\mathrm{s}} \cap \mathrm{F}_{2\mathrm{m}}$ (respectivement $\mathrm{F}_{2\mathrm{s}} \cap \mathrm{F}_{1\mathrm{m}}$) sont des arêtes esclaves déphasées de $-\phi_1$ (respectivement $-\phi_2$) avec les arêtes $\mathrm{F}_{1\mathrm{m}} \cap \mathrm{F}_{2\mathrm{m}}$ (respectivement $\mathrm{F}_{1\mathrm{m}} \cap \mathrm{F}_{2\mathrm{m}}$).

Comme nous l'avons vu dans le paragraphe précédent III.3.5.2 le maillage surfacique doit avoir des éléments identiques qui se correspondent d'après les translations $\mathbf{T}_1$ et $\mathbf{T}_2$:

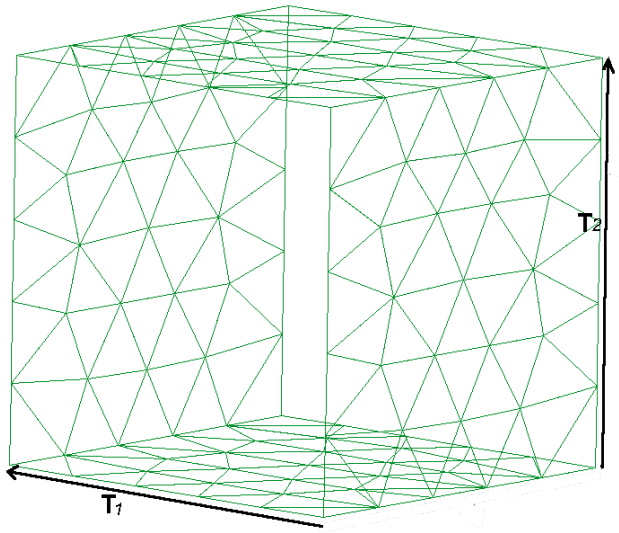


Fig 6.III Maillage contraint non régulier sur les faces ou l'on veut appliquer les conditions "*Maître-esclave*".

Sur la figure Fig 6.III les faces du maillage sont structurées de façon à ce que chaque élément en regard soit image l'un de l'autre par la translation des vecteurs $\mathbf{T}_1$ et $\mathbf{T}_2$.

Remarque III.9. On aurait aussi pu introduire un déphasage dans une troisième direction mais l'écriture du problème est plus lourde et le principe reste le même.

Soit S_k les indices des arêtes sur la face esclaves et M_k les indices des arêtes sur les faces maîtres. Soit N_M le nombre d'arêtes maîtres et N_S le nombre d'arêtes esclaves. Soit I_i les indices des arêtes internes et N_I leur nombre. Sur les interfaces ou ne sont pas appliquées les déphasages (faces du dessus et du dessous) on met une condition métallique PEC. Donc toutes les inconnues associées au PEC sont nulles (cf remarque III.7). Comme $\tilde{\mathbf{E}}$ s'écrit comme la combinaison linéaire des fonctions de base associées aux arêtes, d'après l'équation (20.III) on a alors :

$$\tilde{\mathbf{E}} = \sum_{i=1}^{N_I} e_{I_i} \mathbf{W}_{I_i} + \sum_{k=1}^{N_S} e_{S_k} \mathbf{W}_{S_k} + \sum_{k=1}^{N_M} e_{M_k} \mathbf{W}_{M_k}$$

Et donc en appliquant les relations de déphasages on a :

$$\tilde{\mathbf{E}} = \sum_{i=1}^{N_I} e_{I_i}\mathbf{W}_{I_i} + \sum_{k=1}^{N_M} e_{M_k}(\mathbf{W}_{M_k} + e^{-j\phi}\mathbf{W}_{S_k})$$

Ce qui revient à changer l'espace d'approximation [77]. En effet les fonctions de base $\mathbf{W}_{I_i}$ correspondantes aux arêtes internes ne changeront pas tandis que les fonctions de base correspondantes aux arêtes maîtres $\mathbf{W}_{M_k}$ s'écriront :

$$\tilde{\mathbf{W}}_{M_k} = \mathbf{W}_{M_k} + e^{-j\phi}\mathbf{W}_{S_k}$$

Si on pose :

$$\begin{gathered}\tilde{\mathbf{W}}_{I_i} = \mathbf{W}_{I_i}, \quad \tilde{\mathbf{W}}_{M_k} = \mathbf{W}_{M_k} + e^{-j\phi}\mathbf{W}_{S_k} \\ i \in [1 \dots N_I], \quad k \in [1 \dots N_M]\end{gathered} \tag{21.III}$$

Le problème variationnel défini équation (9.III) s'écrit alors :

$$\begin{gathered}\sum_{i=1}^{N_I} e_{I_i} * a(\mathbf{W}_{I_i}, \tilde{\mathbf{W}}_q) + \sum_{k=1}^{N_M} e_{M_k} * \big(a(\mathbf{W}_{M_k}, \tilde{\mathbf{W}}_q) + e^{-j\phi} * a(\mathbf{W}_{S_k}, \tilde{\mathbf{W}}_q)\big) = \\ k_0^2\big[\sum_{i=1}^{N_i} e_{I_i} * m(\mathbf{W}_{I_i}, \tilde{\mathbf{W}}_q) + \sum_{k=1}^{N_M} e_{M_k} * \big(m(\mathbf{W}_{M_k}, \tilde{\mathbf{W}}_q) + e^{-j\phi} * m(\mathbf{W}_{S_k}, \tilde{\mathbf{W}}_q)\big)\big]\end{gathered}$$

$$(q = I_i,\ i \in [1 \dots N_I]) \cup (q = M_k, k\ \in [1 \dots N_M]) \tag{22.III}$$

$$\text{avec } a(\mathbf{E}, \boldsymbol{\varphi}) = \int_\Omega \boldsymbol{\nabla} \times \boldsymbol{\varphi}^*.\frac{\boldsymbol{\nabla} \times \mathbf{E}}{\mu_r} + \int_{\partial\Omega} \big(\boldsymbol{\nu} \times (\frac{\boldsymbol{\nabla} \times \mathbf{E}}{\mu_r})\big).\boldsymbol{\varphi}^*,$$

$$\text{et } m(\mathbf{E}, \boldsymbol{\varphi}) = \int_\Omega \epsilon_r \mathbf{E}.\boldsymbol{\varphi}^*$$

Remarque III.10. L'espace d'approximation est muni d'une structure hermitienne (cf remarque III.1) donc :

$$a(\mathbf{W}_p, \tilde{\mathbf{W}}_{M_k}) = a(\mathbf{W}_p, \mathbf{W}_{M_k}) + e^{j\phi}a(\mathbf{W}_p, \mathbf{W}_{S_k})$$

Remarque III.11. Soit alors T_m un triangle sur la face maître et T_s le triangle correspondant sur la face esclave ($T_m = T_s + \mathbf{T}$) alors on a :

$$\int_{T_m} \big(\boldsymbol{\nu}_m \times (\frac{\boldsymbol{\nabla} \times \mathbf{W}_{qm}}{\mu_r})\big).\mathbf{W}_{pm} = -\int_{T_s} \big(\boldsymbol{\nu}_s \times (\frac{\boldsymbol{\nabla} \times \mathbf{W}_{qs}}{\mu_r})\big).\mathbf{W}_{ps}$$

Puisqu'il y a continuité des composantes tangentielles des fonctions de Nedelec. **Par conséquent les intégrales surfaciques des faces maîtres et esclaves vont s'annuler deux à deux, lors de l'assemblage des matrices.** Plus généralement, pour chaque terme

$a(\mathbf{W}_p, \tilde{\mathbf{W}}_{M_k})$ contenant une arête maître $\mathbf{Ar}_{M_k}$, on trouve toujours une ou plusieurs contributions de $a(\mathbf{W}_p, \tilde{\mathbf{W}}_{S_k})$ qui annulent l'intégrale surfacique de $a(\mathbf{W}_p, \tilde{\mathbf{W}}_{M_k})$.

Les inconnues seront alors toutes les arêtes moins toutes les arêtes sur les faces PEC moins toutes les arêtes qui sont sur les faces esclaves.

III.3.5.4 Conséquences sur la constructions des matrices

Regardons maintenant les conséquences sur les matrices A er M construites équation (10.III).

Pour clarifier les relations entre les éléments, on associe un déphasage ϕ à chaque arête des faces esclaves des faces $\mathrm{F_{1S}}$ et $\mathrm{F_{2s}}$. On note par F_{1s} la face Ouest, F_{2s} la face Sud, F_{1m} la face Est et F_{2m} la face Nord. Si l'arête se trouve sur la face Sud ($\mathrm{F_{2s}}$) le déphasage est $\phi = -\phi_1$, Ouest($\mathrm{F_{1s}}$) $\phi = -\phi_2$, Sud-Ouest($\mathrm{F_{1s}} \cap \mathrm{F_{2s}}$) $\phi = -(\phi_1 + \phi_2)$.

On note par I l'indice des arêtes internes, par E les arêtes Est(sur $\mathrm{F_{1m}}$), par O(sur $\mathrm{F_{1S}}$) les arêtes ouest, par N(sur $\mathrm{F_{2m}}$) les arêtes nord et par S(sur $\mathrm{F_{2s}}$) les arêtes sud. NE l'intersection des arêtes des faces Nord-Est, SO l'intersection des arêtes des faces Sud-Ouest, SE l'intersection des faces Sud-Est et NO l'intersection des faces Nord-Est. Comme nous l'avons déjà précisé (cf remarque III.8), si une arête est à l'intersection d'une face maître et d'une face esclave alors c'est une arête esclave. Donc, par convention, les arêtes de la face NO (respectivement de la face SE) sont des arêtes de la face O (respectivement de la face S). On note par NS le nombre d'inconnues sur la face Nord moins le nombre d'inconnues sur le segment NE. Autrement dit $NS = N - NE$ est égal au nombre d'inconnues sur $S - SO$. On désigne ensuite par EO le nombre d'inconnues sur $E - NE$ égale au nombre d'inconnues sur $O - SO$, par INS le nombre d'inconnues sur les segments Sud-Ouest et Nord-Est et par Ni le nombre d'inconnues correspondant aux arêtes internes. En regroupant les composantes, on pose :

$$\mathbf{e}_I = [e_{I_1}, \ldots, e_{I_{Ni}}], \quad \mathbf{e}_S = [e_{S_1}, \ldots, e_{S_{NS}}], \quad \mathbf{e}_E = [e_{E_1}, \ldots, e_{E_{EO}}], \quad \mathbf{e}_O = [e_{O_1}, \ldots, e_{O_{EO}}],$$

$$\mathbf{e}_{SO} = [e_{SO_1}, \ldots, e_{SO_{INS}}], \ \mathbf{e}_N = [e_{N_1}, \ldots, e_{N_{NS}}], \quad \mathbf{e}_{NE} = [e_{NE_1}, \ldots, e_{NE_{INS}}]$$

$$\mathbf{X} = (\mathbf{e}_I, \mathbf{e}_S, \mathbf{e}_{SO}, \mathbf{e}_O, \mathbf{e}_N, \mathbf{e}_{NE}, \mathbf{e}_E).$$

Ensuite on considère les différents blocs :

$$(A_{I,I})_{I_p I_q} = a(\mathbf{W}_{I_p}, \mathbf{W}_{I_q})_{(p,q)\in[1\ldots NI]\times[1\ldots NI]}, \ (A_{I,S})_{I_q,S_p} = a(\mathbf{W}_{I_p}, \mathbf{W}_{S_q})_{(p,q)\in[1\ldots NI]\times[1\ldots NS]} \ \ldots$$

Le produit $A\mathbf{X}$ s'écrit initialement :

$$\begin{bmatrix} A_{I,I} & A_{I,S} & A_{I,SO} & A_{I,O} & A_{I,N} & A_{I,NE} & A_{I,E} \\ A_{S,I} & A_{S,S} & A_{S,SO} & A_{S,O} & A_{S,N} & A_{S,NE} & A_{S,E} \\ A_{SO,I} & A_{SO,S} & A_{SO,SO} & A_{SO,O} & A_{SO,N} & A_{SO,NE} & A_{SO,E} \\ A_{O,I} & A_{O,S} & A_{O,SO} & A_{O,O} & A_{O,N} & A_{O,NE} & A_{O,E} \\ A_{N,I} & A_{N,S} & A_{N,SO} & A_{N,O} & A_{N,N} & A_{N,NE} & A_{N,E} \\ A_{NE,I} & A_{NE,S} & A_{NE,SO} & A_{NE,O} & A_{NE,N} & A_{NE,NE} & A_{NE,E} \\ A_{E,I} & A_{E,S} & A_{E,SO} & A_{E,O} & A_{E,N} & A_{E,NE} & A_{E,E} \end{bmatrix} \begin{bmatrix} \mathbf{e}_I \\ \mathbf{e}_S \\ \mathbf{e}_{SO} \\ \mathbf{e}_O \\ \mathbf{e}_N \\ \mathbf{e}_{NE} \\ \mathbf{e}_E \end{bmatrix}$$

Dans ce cas les relations de déphasage s'écrivent :

$$\mathbf{e}_S = e^{-j\phi_1}\mathbf{e}_N, \quad \mathbf{e}_O = e^{-j\phi_2}\mathbf{e}_E, \quad \mathbf{e}_{SO} = e^{-j(\phi_1+\phi_2)}\mathbf{e}_{NE}, \tag{23.III}$$

Et donc le vecteur $\mathbf{X}$ s'écrit :

$$\mathbf{X} = \begin{bmatrix} \mathbf{e}_I, & e^{-j\phi_1}\mathbf{e}_N, & e^{-j(\phi_1+\phi_2)}\mathbf{e}_{NE}, & e^{-j\phi_2}\mathbf{e}_E, & \mathbf{e}_N, & \mathbf{e}_{NE}, & \mathbf{e}_E \end{bmatrix}$$

On définit alors des nouveaux blocs $\tilde{A}_{P,Q}$ à partir des blocs $A_{P,Q}$ et des relations de déphasages entre les inconnues de l'équation (23.III) :

$$\tilde{A}_{P,N} = A_{P,N} + e^{-j\phi_1} * A_{P,S}, \ \tilde{A}_{E,P} = A_{P,E} + e^{-j\phi_2} * A_{P,O},$$

$$\tilde{A}_{P,NE} = A_{P,NE} + e^{-j(\phi_1+\phi_2)} * A_{P,SO} + e^{-j\phi_2} * A_{P,NO} + e^{-j\phi_1} * A_{P,SE} \tag{24.III}$$

$$P \in [N, S, E, O, SO, NE]$$

Ainsi le produit $A\mathbf{X}$ devient :

$$A\mathbf{X} = \begin{bmatrix} A_{I,I} & \tilde{A}_{I,N} & \tilde{A}_{I,NE} & \tilde{A}_{I,E} \\ A_{S,I} & \tilde{A}_{S,N} & \tilde{A}_{S,NE} & \tilde{A}_{S,E} \\ A_{SO,I} & \tilde{A}_{SO,N} & \tilde{A}_{SO,NE} & \tilde{A}_{SO,E} \\ A_{O,I} & \tilde{A}_{O,N} & \tilde{A}_{O,NE} & \tilde{A}_{O,E} \\ A_{NO,I} & \tilde{A}_{NO,N} & \tilde{A}_{NO,NE} & \tilde{A}_{NO,E} \\ A_{SE,I} & \tilde{A}_{SE,N} & \tilde{A}_{SE,NE} & \tilde{A}_{SE,E} \\ A_{N,I} & \tilde{A}_{N,N} & \tilde{A}_{N,NE} & \tilde{A}_{N,E} \\ A_{NE,I} & \tilde{A}_{NE,N} & \tilde{A}_{NE,NE} & \tilde{A}_{NE,E} \\ A_{E,I} & \tilde{A}_{E,N} & \tilde{A}_{E,NE} & \tilde{A}_{E,E} \end{bmatrix} \begin{bmatrix} \mathbf{e}_I \\ \mathbf{e}_N \\ \mathbf{e}_{NE} \\ \mathbf{e}_E \end{bmatrix}$$

D'après [77] l'espace d'approximation a changé selon les transformations décrites équation (21.III) ce qui revient à éliminer toutes les lignes correspondantes aux arêtes esclaves en effectuant les transformations par blocs suivantes :

$$\overleftrightarrow{A}_{N,I} = A_{N,I} + e^{j\phi_1} * A_{S,I}, \quad \overleftrightarrow{A}_{E,I} = A_{E,I} + e^{j\phi_2} * A_{O,I},$$

$$\overleftrightarrow{A}_{NE,I} = A_{NE,I} + e^{j(\phi_1+\phi_2)} * A_{SO,I} + e^{j\phi_2} * A_{NO,I} + e^{-j\phi_1} * A_{SE,I}$$

$$\overleftrightarrow{\overleftrightarrow{A}}_{N,Q} = \tilde{A}_{N,Q} + e^{j\phi_1} * \tilde{A}_{S,Q}, \quad \overleftrightarrow{\overleftrightarrow{A}}_{Q,E} = \tilde{A}_{E,Q} + e^{j\phi_2} * \tilde{A}_{O,Q}, \tag{25.III}$$

$$\overleftrightarrow{\overleftrightarrow{A}}_{NE,Q} = \tilde{A}_{NE,Q} + e^{j(\phi_1+\phi_2)} * \tilde{A}_{NE,Q} + e^{j\phi_2} * \tilde{A}_{NO,Q} + e^{j\phi_1} * \tilde{A}_{SE,Q}$$

$$Q \in [N, NE, E]$$

Ainsi en posant $\tilde{\mathbf{X}} = (\mathbf{e}_I, \mathbf{e}_N, \mathbf{e}_{NE}, \mathbf{e}_E)$ et $\overleftrightarrow{A}$ la matrice A avec les nouveaux blocs construit

par les relations (24.III) et (25.III) le produit $A\mathbf{X}$ devient :

$$A\mathbf{X} = \overleftrightarrow{A}\tilde{\mathbf{X}} = \begin{bmatrix} A_{I,I} & \tilde{A}_{I,N} & \tilde{A}_{I,NE} & \tilde{A}_{I,E} \\ \overleftrightarrow{A}_{N,I} & \overleftrightarrow{A}_{N,N} & \overleftrightarrow{A}_{N,NE} & \overleftrightarrow{A}_{N,E} \\ \overleftrightarrow{A}_{NE,I} & \overleftrightarrow{A}_{NE,N} & \overleftrightarrow{A}_{NE,NE} & \overleftrightarrow{A}_{NE,E} \\ \overleftrightarrow{A}_{E,I} & \overleftrightarrow{A}_{E,N} & \overleftrightarrow{A}_{E,NE} & \overleftrightarrow{A}_{E,E} \end{bmatrix} \begin{bmatrix} \mathbf{e}_I \\ \mathbf{e}_N \\ \mathbf{e}_{NE} \\ \mathbf{e}_E \end{bmatrix}$$

Remarque III.12. Si les matrices initiales A et M sont symétriques semi-définies positives (ce qui est en général le cas). Alors les matrices transformées $\overleftrightarrow{A}$ et $\overleftrightarrow{M}$ par les relations (24.III) et (25.III) sont hermitiennes. Autrement dit il existe une matrice de transformation R tel que[78] :

$$\overleftrightarrow{A} = RAR^*, \quad \overleftrightarrow{M} = RMR^* \tag{26.III}$$

En appelant INT l'ensemble des inconnues correspondants aux arêtes qui ne sont pas sur les faces maîtres, esclaves ou métalliques on peut alors expliciter R en gardant les conventions établies dans cette section :

$$R = \begin{array}{c} \\ INT \\ \mathrm{F}_{1\mathrm{m}} \\ \mathrm{F}_{2\mathrm{m}} \\ \mathrm{F}_{1\mathrm{m}} \cap \mathrm{F}_{2\mathrm{m}} \end{array} \begin{array}{c} \begin{array}{ccccccc} INT & \mathrm{F}_{1\mathrm{s}} & \mathrm{F}_{1\mathrm{s}} \cap \mathrm{F}_{2\mathrm{s}} & \mathrm{F}_{2\mathrm{s}} & \mathrm{F}_{1\mathrm{m}} & \mathrm{F}_{2\mathrm{m}} & \mathrm{F}_{1\mathrm{m}} \cap \mathrm{F}_{2\mathrm{m}} \end{array} \\ \begin{bmatrix} I_d & 0 & \dots & \dots & & 0 & 0 \\ 0 & e^{j\phi_1} I_d & 0 & 0 & I_d & 0 & 0 \\ 0 & 0 & e^{j(\phi_1+\phi_2)} I_d & 0 & 0 & I_d & 0 \\ 0 & \dots & 0 & e^{j\phi_2} I_d & \dots & 0 & I_d \end{bmatrix} \end{array}$$

Où I_d désigne l'application identité.

Remarque III.13. On a choisit ici d'inclure la structure dans un parallélépipède rectangle mais les conséquences des déphasages entre les faces maîtres et esclaves sur les matrices éléments finis s'écrivent de la même façon si la structure est contenue dans un un parallélépipède non rectangle :

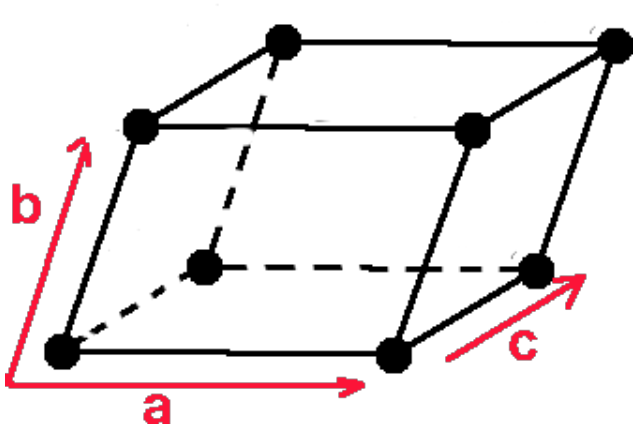


Fig 7.III Maille de base non rectangulaire. **a**, **b**, **c** sont les vecteurs de translations des faces esclaves vers les faces maîtres.

Remarque III.14. Lors de l'assemblage des matrices il faut bien vérifier que chacune des arêtes maîtres est orientée dans le même sens que son arête esclave correspondante.

III.4 Intérêts de la méthode des éléments finis

- Les domaines peuvent avoir des tailles très contrastées et des formes quelconques.
- Les matériaux étudiés peuvent être de natures très différentes : anisotropes, métaux parfaitement conducteurs, hétérogènes etc.
- Les matériaux peuvent être dispersifs : leurs paramètres peuvent dépendre de la fréquence de calcul (cf section VI.2 et VI.2.2 remarque VI.7).
- L'espace mémoire et le temps de calcul peuvent être optimisés, car les matrices stockées sont creuses (cf section IV.6 et section IV.6.5).

III.5 Limitations

- Contraintes géométriques sur le maillage dues aux conditions *Maître-esclave* (cf section III.3.5).
- Nécessité de mailler très finement dans les zones étroites (cf section IV.4.6).
- Quand les matériaux sont dispersifs : la taille du problème de calcul de modes propres est multipliée par un entier qui dépend de l'expression de la permittivité (cf section VI.2).

III.6 Conclusion

Cette section a traité de la modélisation du problème de modes propres à partir des équations de Maxwell fréquentielles. La formulation variationnelle des équations de Maxwell est abordée section III.2. On énonce dans les perspectives section VI.2 la possibilité de modifier cette formulation pour modéliser des structures dispersives. Les conditions limites particulières abordées dans la section III.3 permettent de traiter des structures périodiques infinies qui peuvent être composées de matériaux parfaitement conducteurs ou absorbants sur certaines faces de la frontière du domaine de résolution. La description détaillée des conditions limites *Maître-esclave* nous a permis de clarifier l'implémentation des problèmes éléments finis périodiques. Nous avons aussi vu que, quelles que soient les conditions aux limites imposées aux bords du domaine, les problèmes de modes propres aboutissent à la construction d'un système matriciel 10.III que nous allons chercher à résoudre de façon optimale dans la section suivante.

. . . . . .

IV Résolution du problème aux valeurs propres

IV.1 Définition du problème

Dans la section précédente nous avons posé les bases théoriques nécessaires pour modéliser les problèmes périodiques avec la méthode des éléments finis. Dans cette section nous allons décrire les méthodes de résolution du problème qui s'écrit :

$$A\mathbf{X} = k_0^2 M\mathbf{X} \tag{1.IV}$$

Où A et M sont les matrices éléments finis précédemment définies en fonction des conditions limites imposées au bord du domaine d'étude et $\mathbf{X}$ représente la circulation du champ électrique $\tilde{\mathbf{E}}$ sur chaque arête (cf équation (20.III)). Pour le calcul de diagrammes de bandes (cf section II.1.8) les matrices A et M dépendent des valeurs des vecteurs d'ondes incidents $\mathbf{k}_i$ qui permettent d'appliquer les conditions de Floquet au bord du domaine (cf section III.3.5).

Remarque IV.1. Comme nous l'avons déjà évoqué en remarque III.6 le noyau de A est non nul et il il existe un vecteur $\mathbf{U} \in (H^1(\Omega))^3$ dérivant d'un potentiel vecteur $\mathbf{V} \in H(curl, \Omega)$ et un potentiel scalaire $\phi \in H^1(\Omega)$ tel que :

$$\tilde{\mathbf{E}} = \mathbf{U} + \boldsymbol{\nabla}\phi, \text{ tel que } \boldsymbol{\nabla}.\mathbf{U} = 0 \text{ et } \mathbf{U} = \boldsymbol{\nabla} \times \mathbf{V}.$$

Le noyau de A contient alors les champs $\tilde{\mathbf{E}} = \mathbf{X}^T\underline{\mathbf{W}}$ qui ont une composante rotationnelle nulle (i.e $\mathbf{U} \equiv \mathbf{0}$). Il existe donc des solutions $(k_0^2, \tilde{\mathbf{E}}) = (0, \mathbf{X}^T\underline{\mathbf{W}})$ avec $\mathbf{X}$ non nul et solution de (1.IV).

Les modes propres recherchés sont les couples $(k_0^2, \mathbf{X})$. La recherche de modes propres correspond à une recherche de valeurs propres et vecteurs propres. En effet, supposons que la matrice A soit inversible, le système (1.IV) peut alors s'écrire :

$$MA^{-1}\mathbf{X} = \frac{1}{k_0^2}\mathbf{X} \tag{2.IV}$$

Bien que la matrice A ne soit pas inversible d'après la remarque (IV.1), nous avons volontairement choisi d'écrire l'équation (2.IV) de cette façon puisque nous devons calculer les valeurs k_0^2 dans l'ordre croissant et les algorithmes classiques de recherche de valeurs propres (cf section IV.2 et annexe A.6), nous permettent de calculer les solutions $\frac{1}{k_0^2}$ dans l'ordre décroissant. De cette façon, les premières valeurs propres calculées seront les plus petites. Or toujours d'après la remarque IV.1, les premières valeurs propres solutions de (1.IV) sont nulles. Pour ces raisons, le système (2.IV) devra subir la transformation *Shift and Invert* abordé section IV.3, pour calculer les premières valeurs propres non nulles et pour éviter d'inverser la matrice A. Dans la section IV.4, on donne une description originale des différents paramètres nécessaires à la résolution du problème (1.IV).

Dans la section IV.5.1, on classe les valeurs propres nulles en deux groupes :

- Les valeurs propres nulles k_0^2 solutions de (1.IV) dont le champ solution correspondant $\tilde{\mathbf{E}} = \mathbf{X}^T\underline{\mathbf{W}}$ est à divergence nulle ($\boldsymbol{\nabla}.\tilde{\mathbf{E}} = 0$).
- Les valeurs propres nulles k_0^2 solutions de (1.IV) dont le champ solution correspondant $\tilde{\mathbf{E}} = \mathbf{X}^T\underline{\mathbf{W}}$ est à divergence non nulle ($\boldsymbol{\nabla}.\tilde{\mathbf{E}} \neq 0$).

Dans la section IV.5.2, nous décrivons l'algorithme d'élimination des valeurs propres nulles dont le champ correspondant est à divergence non nulle, que nous avons développé pour les problèmes de cavités métalliques. Les algorithmes de recherches de valeurs propres utilisés dans nos travaux intègrent les solveurs issus des librairies ARPACK[33] et PARDISO[37]. Nous allons expliquer dans cette section comment les librairies sont combinées en détaillant chaque algorithme utilisé. Dans la section IV.6 on explique comment nous avons manipulé la structure des matrices A et M construites afin de minimiser les temps de calculs et l'espace mémoire utilisé.

IV.2 L' algorithme de calcul des valeurs propres

IV.2.1 Introduction

La méthode que nous allons présenter dans cette section nous permet de calculer les plus grandes valeurs propres d'une matrice de grande taille. On introduit alors les définitions nécessaires à la bonne compréhension des algorithmes que nous allons décrire.

Définition. Soit $\mathcal{O}$ une matrice carrée à coefficients complexes de dimension N et $\mathbf{v}_0$ un vecteur quelconque de $\mathbb{C}^N$. L'espace de Krilov d'ordre p noté K_p est l'espace vectoriel engendré par les vecteurs $\mathbf{v}_0, \mathcal{O}\mathbf{v}_0, \mathcal{O}^2\mathbf{v}_0, \ldots, \mathcal{O}^{p-1}\mathbf{v}_0$:

$$K_p = \text{span}\,\{\mathbf{v}_0, \mathcal{O}\mathbf{v}_0, \mathcal{O}^2\mathbf{v}_0, \ldots, \mathcal{O}^{p-1}\mathbf{v}_0\}.$$

On va noter p_{max} la dimension maximale de K_p, pour un $\mathbf{v}_0$ donné. D'après le théorème de Cayley Hamilton[79] $p_{max} \leq n$.

Proprieté IV.1. Si $\mathcal{O}^p\mathbf{v}_0 \in K_p$ alors $\mathcal{O}^{p+q}\mathbf{v}_0 \in K_p$ pour tout $q > 0$ [80].

Démonstration. On démontre ce résultat par récurrence. Si pour un $q \geq 0$, $\mathcal{O}^{p+q}\mathbf{v}_0 \in K_p$, alors :

$$\mathcal{O}^{p+q}\mathbf{v}_0 = \sum_{k=0}^{p-1} \alpha_k \mathcal{O}^k \mathbf{v}_0$$

$$\Rightarrow \mathcal{O}^{p+q+1}\mathbf{v}_0 = \sum_{k=0}^{p-2} \alpha_k \mathcal{O}^{k+1}\mathbf{v}_0 + \alpha_{p-1}\mathcal{O}^p\mathbf{v}_0$$

$$= \sum_{k=0}^{p-2} \alpha_k \mathcal{O}^{k+1}\mathbf{v}_0 + \alpha_{p-1}\sum_{k=0}^{p-1} \beta_k \mathcal{O}^k \mathbf{v}_0 = \sum_{k=0}^{p-1} \gamma_k \mathcal{O}^k \mathbf{v}_0$$

■

IV.2.1.1 Conséquence :

Si p est le plus petit entier pour lequel $\mathcal{O}^p\mathbf{v}_0$ est dépendant des vecteurs précédents, alors, les vecteurs

$$(\mathbf{v}_0, \mathcal{O}\mathbf{v}_0, \mathcal{O}^2\mathbf{v}_0, \ldots, \mathcal{O}^{p-1}\mathbf{v}_0)$$

sont linéairement indépendants et donc K_q est de dimension q, pour tout $q \leq p$. En particulier K_p est de dimension p [80].

IV.2.2 Construction de la base d'Arnoldi

Nous allons aborder dans cette section l'algorithme de construction de la base d'Arnoldi permettant de calculer les plus grandes valeurs propres de la matrice $\mathcal{O}$. Soit $\mathbf{v}_1$ un vecteur quelconque normé de $\mathbb{C}^n$. Voici l'algorithme de construction du vecteur $\mathbf{v}_{j+1}$ à partir des j

premiers vecteurs de la base d'Arnoldi :

$$\begin{aligned}
&\mathbf{w} = \mathcal{O}\mathbf{v}_j \\
&\text{for } i = 1 \text{ to } j \text{ do} \\
&h_{ij} =< \mathbf{w}, \mathbf{v}_i >= \mathbf{w}.\mathbf{v}_i^* \\
&\mathbf{w} = \mathbf{w} - h_{ij}\mathbf{v}_i \\
&\text{end do} \\
&\mathbf{v}_{j+1} = \frac{\mathbf{w}}{||\mathbf{w}||} \\
&h_{j+1j} = ||\mathbf{w}||
\end{aligned} \tag{3.IV}$$

Les vecteurs d'Arnoldi $\mathbf{v}_j$ sont construits par orthogonalisation des vecteurs construits par l'algorithme de puissance itérée[81](cf annexe A.6.2). En effet, les vecteurs calculés par l'algorithme au pas $j \geq 2$ sont orthogonalisés par rapport à $[\mathbf{v}_1, ..., \mathbf{v}_{j-1}]$. Si on désire calculer les $\bar{k}$ premières valeurs propres (dans l'ordre décroissant), on construit d'abord $\bar{p}$ vecteurs d'Arnoldi, avec $\bar{p}$ suffisamment grand. Ensuite les vecteurs $[\mathbf{v}_{\bar{p}}, \ldots, \mathbf{v}_{\bar{p}-\bar{k}+1}]$ constituent une approximation des $\bar{k}$ premiers vecteurs propres de $\mathcal{O}$ car chaque vecteur est orthogonal aux autres et donc deux vecteurs propres différents sont associés à deux valeurs propres différentes. Équation (3.IV) h_{ij} est le coefficient d'orthogonalisation de $\mathcal{O}\mathbf{v}_j$ par rapport à $\mathbf{v}_i$ et h_{j+1j} est la norme du vecteur $\mathbf{w}$ obtenue par orthogonalisation du vecteur $\mathcal{O}\mathbf{v}_j$ par rapport aux vecteurs $\mathbf{v}_i,\ i \in [1, \ldots j+1]$, on en déduit [80] :

$$\mathcal{O}\mathbf{v}_j = \sum_{i=1}^{j+1} h_{ij}\mathbf{v}_i \tag{4.IV}$$

On définit ensuite la matrice

$$V_{\bar{p}} = (\mathbf{v}_1, \ldots, \mathbf{v}_{\bar{p}}). \tag{5.IV}$$

$V_{\bar{p}}$ est une base orthonormale formée des $\bar{p}$ premiers vecteurs d'Arnoldi et donc on a [80] :

$$V_{\bar{p}}^* V_{\bar{p}} = I_{\bar{p}} \tag{6.IV}$$

Dans l'algorithme ci-dessus on remarque que la base $V_{\bar{p}}$ est construite par orthonormalisation de l'espace $[\mathbf{v}_1, \mathcal{O}\mathbf{v}_1, \mathcal{O}^2\mathbf{v}_1, \ldots, \mathcal{O}^{\bar{p}}\mathbf{v}_1]$ obtenue avec le procédé de Gram-Schmidt. Aussi, l'équation (4.IV), qui définie les relations entre les vecteurs d'Arnoldi, peut s'écrire [80] :

$$\mathcal{O}V_{\bar{p}} = V_{\bar{p}+1}H_{\bar{p}+1\bar{p}} \tag{7.IV}$$

où $H_{\bar{p}+1,\bar{p}}$ représente la matrice à $\bar{p}$ colonnes et $\bar{p}+1$ lignes et h_{ij} sont les coefficients d'or-

thonormalisation de l'équation (4.IV), pour $i \leq j+1$, les autres coefficients étant nuls [80] :

$$H_{\bar{p}+1,\bar{p}} = \begin{pmatrix} h_{1,1} & h_{1,2} & \vdots & \vdots & h_{1j} & \vdots & h_{1,\bar{p}} \\ h_{2,1} & h_{2,2} & \vdots & \vdots & h_{2,j} & \vdots & h_{2,\bar{p}} \\ 0 & h_{3,2} & \vdots & \vdots & h_{3,j} & \vdots & h_{3,\bar{p}} \\ \vdots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots \\ \vdots & & \ddots & \ddots & \vdots & \vdots & \vdots \\ \vdots & & & \ddots & h_{j+1,j} & \vdots & h_{j+1,p} \\ \vdots & & & & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & \dots & \dots & 0 & h_{\bar{p}+1,\bar{p}} \end{pmatrix}$$

La matrice $H_{\bar{p}}$ qui est la matrice $H_{\bar{p}+1,\bar{p}}$ sans la dernière ligne vérifie d'après (6.IV) et (7.IV) :

$$V_{\bar{p}}^* \mathcal{O} V_{\bar{p}} = H_{\bar{p}} = \begin{pmatrix} h_{1,1} & h_{1,2} & \vdots & \vdots & h_{1,\bar{p}-1} & h_{1,\bar{p}} \\ h_{2,1} & h_{2,2} & \vdots & \vdots & h_{2,\bar{p}-1} & h_{2,\bar{p}} \\ 0 & h_{3,2} & \ddots & \vdots & h_{3,\bar{p}-1} & h_{3,\bar{p}} \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & 0 & h_{\bar{p},\bar{p}-1} & h_{\bar{p},\bar{p}} \end{pmatrix} \tag{8.IV}$$

À l'exception des termes situés sur la première sous-diagonale, la partie triangulaire inférieure $H_{\bar{p}}$ est nulle. Autrement dit $H_{\bar{p}}$ est une matrice de Hessenberg supérieure [80]. Soit $\mathbf{e}_{\bar{p}}$ le vecteur de $\mathbb{C}^{\bar{p}}$ tel que $\mathbf{e}_{\bar{p}} = (0, 0, \dots, 0, 1)$. En utilisant le produit tensoriel, l'équation (7.IV) peut aussi s'écrire :

$$\mathcal{O} V_{\bar{p}} = V_{\bar{p}} H_{\bar{p}} + h_{\bar{p}+1,\bar{p}} \mathbf{v}_{\bar{p}+1} \otimes \mathbf{e}_{\bar{p}} \tag{9.IV}$$

et sachant que :

$$[\mathbf{v}_{\bar{p}+1} \otimes \mathbf{v}_{\bar{p}}^*] \mathbf{v}_j = \mathbf{v}_{\bar{p}+1} < \mathbf{v}_j, \mathbf{v}_{\bar{p}} >,$$

l'identité (6.IV) nous permet alors de réécrire l'identité (9.IV) de la façon suivante :

$$(\mathcal{O} - h_{\bar{p}+1,\bar{p}} \mathbf{v}_{\bar{p}+1} \otimes \mathbf{v}_{\bar{p}}^*) V_{\bar{p}} = V_{\bar{p}} H_{\bar{p}} \tag{10.IV}$$

La matrice $V_{\bar{p}}$ est alors une base d'un sous-espace invariant associé à $\mathcal{O} - h_{\bar{p}+1,\bar{p}} \mathbf{v}_{\bar{p}+1} \otimes \mathbf{v}_{\bar{p}}^*$.

Remarque IV.2. Nous avons vu (propriété IV.1) qu'il existe un rang p à partir duquel la famille des vecteurs de Krilov $[\mathbf{v}_1, \mathcal{O}\mathbf{v}_1, \mathcal{O}^2\mathbf{v}_1, \dots, \mathcal{O}^{\bar{p}}\mathbf{v}_1]$ est liée, ce qui est équivalent à dire que le facteur $h_{\bar{p}+1,\bar{p}}$ est nul. En effet si $h_{\bar{p}+1,\bar{p}} = 0$ on aura :

$$\mathcal{O} V_{\bar{p}} = V_{\bar{p}} H_{\bar{p}} \tag{11.IV}$$

L'égalité (11.IV) signifie que l'espace engendré par les vecteurs de $V_{\bar{p}}$ est stable par l'ap-

plication linéaire associée à la matrice $\mathcal{O}$. Soit alors $\mathbf{Z}$ un vecteur propre associé à une valeur propre λ de $H_{\bar{p}}$, $V_{\bar{p}}\mathbf{Z}$ est alors une approximation du vecteur propre de $\mathcal{O}$ associé à la valeur propre λ. En effet l'identité (9.IV) nous permet d'écrire :

$$\|\mathcal{O}V_{\bar{p}}\mathbf{Z} - \lambda V_{\bar{p}}\mathbf{Z}\| = \|(\mathcal{O}V_{\bar{p}} - V_{\bar{p}}H_{\bar{p}})\mathbf{Z}\| \leq |h_{\bar{p}+1,\bar{p}}|\|\mathbf{v}_{\bar{p}+1}\|_\infty|z_{\bar{p}}| \tag{12.IV}$$

Donc la convergence de chaque valeur propre est contrôlée par le réel $|h_{\bar{p}+1,\bar{p}}|\|\mathbf{v}_{\bar{p}+1}\|_\infty|z_{\bar{p}}|$, où $z_{\bar{p}}$ est la dernière composante du vecteur propre $\mathbf{Z}$. Soit $\bar{tol}$ une grandeur positive réelle fixée. On considère qu'une valeur propre est acceptable si

$$|h_{\bar{p}+1,\bar{p}}|\|\mathbf{v}_{\bar{p}+1}\|_\infty|z_{\bar{p}}| \leq \bar{tol}. \tag{13.IV}$$

Remarque IV.3. Les valeurs propres de $H_{\bar{p}}$ peuvent être calculées avec l'algorithme QR[82] (cf annexe A.6.4) puisque $H_{\bar{p}}$ n'est pas une matrice de grande taille. La décomposition QR de la matrice $H_{\bar{p}}$ peut se faire facilement avec les rotations de Givens[83] (cf propriété IV.4).

IV.2.3 Cas où la matrice $\mathcal{O}$ est symétrique réelle, ou hermitienne : l'algorithme de Lanczos.

Dans le cas particulier où la matrice $\mathcal{O}$ est hermitienne, la matrice $H_{\bar{p}}$ l'est aussi, car :

$$(V_{\bar{p}}^*\mathcal{O}V_{\bar{p}})^* = (H_{\bar{p}})^* = V_{\bar{p}}^*\mathcal{O}^*V_{\bar{p}} = H_{\bar{p}}.$$

Comme $H_{\bar{p}}$ est aussi une matrice de Hessenberg supérieure, on en déduit alors qu'elle est tri-diagonale hermitienne. Par conséquent les coefficients h_{ij} sont nuls, si $i < j-1$ et le coefficient $h_{j+1,j}$ est égal à $h_{j,j+1}^*$, précédemment calculé. L'algorithme de construction d'une base orthonormée des espaces de Krylov (cf définition IV.2.1) s'appelle dans ce cas l'algorithme de Lanczos [84]. Soit $\mathbf{v}_1$ un vecteur quelconque normé $\mathbb{C}^n$ l'algorithme d'Arnoldi (3.IV) devient :

$$\begin{aligned}
&\mathbf{w} = \mathcal{O}\mathbf{v}_j \\
&h_{j-1,j}^* = h_{j,j-1} = h_{j,j-1}^* \\
&\mathbf{w} = \mathbf{w} - h_{j-1,j}\mathbf{v}_{j-1} \\
&h_{jj} = < \mathbf{w}, \mathbf{v}_j > \\
&\mathbf{w} = \mathbf{w} - h_{jj}\mathbf{v}_j \\
&h_{j+1,j} = ||\mathbf{w}|| \\
&v_{j+1} = \frac{\mathbf{w}}{h_{j+1,j}}
\end{aligned} \tag{14.IV}$$

Cet algorithme permet la construction de la base $V_{\bar{p}}$ définie équation (5.IV), en calculant $\mathbf{v}_{j+1}$ à partir des seuls vecteurs $\mathbf{v}_{j-1}$ et $\mathbf{v}_j$. Cette méthode de Lanczos associée aux algorithmes QR et puissance itérée inverse (cf annexes A.6.3 et A.6.4) est utilisée par l'ONERA pour le calcul de modes guidés TE et TM dans des sections de guides de grande taille [85, 86].

Remarque IV.4. En se référent à la construction utilisant le produit tensoriel établie dans la

remarque (III.6), on peut facilement voir que lorsque ϵ et μ sont des matrices hermitiennes complexes A et M sont des matrices hermitiennes et donc $A-\sigma M$ est aussi une matrice hermitienne. Dans ce cas, l'opérateur $\mathcal{O}$ de l'équation (14.IV) est hermitien et tous les coefficients de la matrice $H_{\bar{p}+1,\bar{p}}$ sont réels par construction. Malheureusement l'algorithme de Lanczos est proposé par la bibliothèque ARPACK que dans le cas où A et M sont symétriques réels ce qui arrive quand ϵ et μ sont aussi symétriques et réels. C'est dans ce cas précis que nous avons utilisé l'algorithme de Lanczos.

Nous avons vu remarques IV.2 et IV.3 que le calcul des valeurs propres de $\mathcal{O}$ nécessite d'abord le calcul des valeurs propres de la matrice de Hessenberg supérieur $H_{\bar{p}}$ d éfinie équation (8.IV). Pour calculer les valeurs propres de ce type de matrices on utilise l'algorithme QR translaté que nous allons définir dans la section suivante.

IV.2.4 Algorithme QR translaté (QR shifted)

L' algorithme QR translaté [87] est une version optimisée de l'algorithme QR[82] classique(cf annexe A.6.4). Il est utilisé par ARPACK [33] lors du procédé de construction du calcul des valeurs propres de la matrice de Hessenberg $H_{\bar{p}}$ (8.IV). Soit une matrice H à valeur dans $\mathbb{C}^n$. On construit une suite de matrices $(H_k)_{k\geq 0}$ de la façon suivante :

$$\begin{cases} H-\alpha_0 I = Q_0 R_0 \\ H_1 = R_0 Q_0 + \alpha_0 Id \\ H_1 - \alpha_1 Id = Q_1 R_1 \\ \vdots \\ H_k - \alpha_k Id = Q_k R_k \\ H_{k+1} = R_k Q_k + \alpha_k Id = Q_{k+1} R_{k+1} \end{cases} \tag{15.IV}$$

Où α_k est un réel à choisir à chaque pas. Un choix classique de α_k est le terme matriciel $(H_k)_{nn}$.

Proprieté IV.2. Si H est trigonalisable (toujours dans $\mathbb{C}$), alors quand $k\to\infty$, les matrices H_k construites par cet algorithme tendent vers une matrice triangulaire supérieure qui a ses éléments diagonaux égaux aux valeurs propres de H. Cet algorithme à un intérêt pour sa rapidité de convergence mais les valeurs propres de la matrice triangulaire ne sont en général pas dans l'ordre décroissant[87].

Proprieté IV.3. Les matrices ainsi construites sont toutes orthogonalement semblables entre elles.

Démonstration.

$$H_{k+1} = R_k Q_k + \alpha_k Id = Q_k^* Q_k R_k Q_k + \alpha_k Id = Q_k^*[H_k - \alpha_k Id]Q_k + \alpha_k Id = Q_k^* H_k Q_k.$$

■

Proprieté IV.4. Si H est une matrice de Hessenberg complexe alors toutes les matrices H_k construites par l'algorithme QR translaté sont de Hessenberg.

Démonstration. Étant donné une matrice de Hessenberg H complexe disposant d'une factorisation QR, $H = QR$, montrons que $\tilde{H} = RQ$ est encore une matrice de Hessenberg. On utilise les matrices de rotations de Givens complexes[88, 83, 89] définies par :

$$G(i,j) = \begin{pmatrix} 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \dots & c & \dots & s & \dots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \dots & -s^* & \dots & c & \dots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix} \begin{matrix} \\ \\ \leftarrow i \\ \\ \leftarrow j \\ \\ \\ \end{matrix}$$
$$\begin{matrix} & & \uparrow & & \uparrow & & \\ & & i & & j & & \end{matrix}$$

Comme $G(i,j)$ est une matrice unitaire on a $|c|^2 + |s|^2 = 1$ et il existe θ_1, θ_2 tel que $s = \sin(\theta_1)e^{j\theta_2}$, $c = \cos(\theta_1)$. On veut pouvoir éliminer les coefficients nuls de la matrice de Hessenberg. Or, si on prend un vecteur

$$\mathbf{X} = (x_1 e^{j\phi_1}, x_2 e^{j\phi_2}, \dots, x_n e^{j\phi_n}) \in \mathbb{C}^n$$

on a :

$$G(i,j)\mathbf{X} = \mathbf{Y} = \begin{Bmatrix} cx_i e^{j\phi_i} + sx_j e^{j\phi_j}, & k = i \\ -s^* x_i e^{j\phi_i} + cx_j e^{j\phi_j}, & k = j \\ x_k e^{j\phi_k} & k \neq i, j \end{Bmatrix}$$

Et on peut faire en sorte que y_j soit nul en choisissant :

$$\theta_1 = \tan^{-1}(\frac{x_j}{x_i}),\ \theta_2 = \phi_i - \phi_j \Rightarrow c = \frac{x_i}{\sqrt{x_i^2 + x_j^2}},\ s = \frac{x_j}{\sqrt{x_i^2 + x_j^2}} e^{j(\phi_i - \phi_j)}.$$

Appliquons alors une à une les rotations de Givens à une matrice 4×4 de Hessenberg pour montrer que l'on peut annuler les coefficients sous diagonaux.

$$H = \begin{pmatrix} X & X & X & X \\ X & X & X & X \\ 0 & X & X & X \\ 0 & 0 & X & X \end{pmatrix} \quad \begin{matrix} G(1,2) \\ \longrightarrow \end{matrix} \quad = \begin{pmatrix} X & X & X & X \\ 0 & X & X & X \\ 0 & X & X & X \\ 0 & 0 & X & X \end{pmatrix}$$

$$\underset{\longrightarrow}{G(2,3)} = \begin{pmatrix} X & X & X & X \\ 0 & X & X & X \\ 0 & 0 & X & X \\ 0 & 0 & X & X \end{pmatrix} \quad \underset{\longrightarrow}{G(3,4)} = \begin{pmatrix} X & X & X & X \\ 0 & X & X & X \\ 0 & 0 & X & X \\ 0 & 0 & 0 & X \end{pmatrix}$$

Ainsi on peut écrire :

$$G(3,4)G(2,3)G(1,2)H = R \Longleftrightarrow H = QR \ , \text{ avec } Q = G(1,2)G(2,3)G(3,4)$$

Et donc

$$\tilde{H} = RQ = RG(1,2)G(2,3)G(3,4)$$

et $\tilde{H}$ est une matrice de Hessenberg par construction.

■

Proprieté IV.5. Soit H une matrice de Hessenberg irréductible de dimension N

$$\text{i.e } h_{i+1,i} \neq 0 \ \forall \ i = 1, \dots n-1.$$

Alors pour tous les éléments diagonaux r_{kk} de la matrice R de la décomposition QR, on a :

$$r_{kk} > 0 \quad \forall k < n.$$

Donc si H est singulière alors $r_{nn} = 0$ [89].

Démonstration. La multiplication du terme diagonal et du terme sous diagonal par la matrice de rotation de Givens d'ordre k $G(k, k+1)$ s'écrit :

$$\begin{pmatrix} \cos(\theta_1) & \sin(\theta_1)e^{j\theta_2} \\ -\sin(\theta_1)e^{-j\theta_2} & \cos(\theta_1) \end{pmatrix} \begin{pmatrix} h_{kk} \\ h_{k+1,k} \end{pmatrix} = \begin{pmatrix} r_{kk} \\ 0 \end{pmatrix}$$

Comme les rotations de Givens conservent la norme, on a :

$$|r_{kk}|^2 = |h_{kk}|^2 + |h_{k+1,k}|^2 \geq |h_{k+1,k}|^2 > 0$$

■

Remarque IV.5. En conséquence de la propriété IV.5, si on connaît une valeur propre λ de H, alors

$$H - \lambda Id = QR \text{ avec } r_{nn} = 0.$$

Ainsi la matrice construite à l'étape suivante $\tilde{H} = RQ + \lambda Id$ est une matrice de Hessenberg (cf remarque IV.4) et on a :

$$\tilde{H} = RQ + \lambda Id \text{ avec } r_{nn} = 0 \Rightarrow \tilde{h}_{n-1,n} = 0 \Rightarrow \tilde{H} = \begin{pmatrix} \tilde{H}_1 & \clubsuit \\ 0 & \lambda \end{pmatrix} \tag{16.IV}$$

Où ♣ représente un élément quelconque de $\mathbb{C}^{n-1}$. Si on applique l'algorithme QR translaté [82] en choisissant un coefficient α_k égal à une valeur propre de H, on peut directement continuer l'algorithme sur une sous matrice de Hessenberg $\tilde{H}_1$ représentée équation (16.IV).

IV.2.5 Itérations successives et redémarrage

L'algorithme d'Arnoldi avec redémarrage, appelé IRAM (Implicitly Restarted Arnoldi Method [90]) permet de maîtriser la convergence de l'algorithme d'Arnoldi en optimisant l'espace mémoire utilisé. Cette algorithme utilise la décomposition QR translaté que nous avons vu dans la section précédente. Soit $\bar{tol}$ le réel fixé qui contrôle la convergence de chaque valeur propre définie par le critère (13.IV). Soit $\bar{p}$ le nombre de vecteurs d'Arnoldi, $\bar{k}$ le nombre de valeur propres voulues et $h_{\bar{p}+1,\bar{p}}\mathbf{v}_{\bar{p}+1} = \mathbf{f}_{\bar{p}}$.

Étape initiale. Construire une base d'Arnoldi de dimension $\bar{p}$ à partir d'un vecteur $\mathbf{v}_1 \in \mathbb{C}^n$ (cf section IV.2.2) . On obtient alors la factorisation d'Arnoldi :

$$\mathcal{O}V_{\bar{p}} = V_{\bar{p}}H_{\bar{p}} + \mathbf{f}_{\bar{p}} \otimes \mathbf{e}_{\bar{p}}$$

1. Calculer les valeurs propres $\{\lambda_j : j = 1, 2, \ldots, \bar{p}\}$ et les vecteurs propres associés $\{\mathbf{Z}_j : j = 1, 2, \ldots, \bar{p}\}$ de la matrice $H_{\bar{p}}$ en utilisant un algorithme de QR classique (cf annexe A.6.4), qui utilise les rotations de Givens[83].
2. Si au moins $\bar{k}$ valeurs propres vérifient le critère (13.IV) arrêter l'algorithme, sinon faire $\bar{p} - \bar{k} = m$ itérations de l'algorithme QR translaté [87] sur la matrice $H_{\bar{p}}$ en utilisant les valeurs propres non désirées $\{\lambda_j : j = \bar{k}+1, \bar{k}+2, \ldots, \bar{p}\}$ (cf équation (15.IV) et remarque IV.5 de la section IV.2.4) dans l'ordre croissant de leur module[90] pour obtenir
$$H_{\bar{p}}Q_{\bar{p}} = Q_{\bar{p}}H^+_{\bar{p}} \Rightarrow \mathcal{O}V_{\bar{p}}Q_{\bar{p}} = V_{\bar{p}}Q_{\bar{p}}H^+_{\bar{p}} + \mathbf{f}^+_{\bar{p}} \otimes \mathbf{e}_{\bar{p}}Q_{\bar{p}}$$
3. Redémarrage : Garder les $\bar{k}$ premières colonnes $Q_{\bar{k}}$ de la matrice $Q_{\bar{p}}$. Calculer les $\bar{k}$ nouveaux vecteurs d'Arnoldi $V_{\bar{k}}$ en posant : $V_{\bar{k}} \leftarrow V_{\bar{p}}Q_{\bar{k}}$.
4. Étendre la factorisation de longueur $\bar{k}$ à une factorisation de longueur $\bar{p}$. Autrement dit calculer les $\bar{p} - \bar{k}$ vecteurs restants de $V_{\bar{p}}$ à partir des vecteurs $V_{\bar{k}}$. Retourner en 1.

Remarque IV.6. À chaque redémarrage, les multiplications successives par $\mathcal{O}$ forcent les vecteurs à prendre la direction des vecteurs propres (cf remarque A.2) et donc l'algorithme IRAM à plus de chance de converger. Pour la même raison, il apparaît aussi que, quand le nombre de vecteurs d'Arnoldi $\bar{p}$ est grand plus le nombre de redémarrage nécessaire pour faire converger l'algorithme IRAM est petit.

IV.2.6 Conclusion

Nous avons détaillé dans cette section les différentes étapes et paramètres de l'algorithme de calcul des valeurs propres d'une matrice $\mathcal{O}$ de grande taille. La base de cet algorithme

développé dans la librairie ARPACK[33] est l'algorithme d'Arnoldi[91] qui va subir les transformations nécessaires pour résoudre nos problèmes. En effet comme nous l'avons évoqué dans la section IV.1, si la matrice A est inversible, le problème (2.IV) suggère que $\mathcal{O} = MA^{-1}$. Ainsi, à chaque itération de l'algorithme d'Arnoldi(3.IV) une inversion de la matrice A est nécessaire et la construction de la base d'Arnoldi évoquée en (3.IV) devient :

$$\begin{aligned} & M\mathbf{w} = A\mathbf{v}_j \\ & \text{for } i = 1 \text{ to } j \text{ do} \\ & h_{ij} =< \mathbf{w}, \mathbf{v}_i >= \mathbf{w}.\mathbf{v}_i^* \\ & \mathbf{w} = \mathbf{w} - h_{ij}\mathbf{v}_i \\ & \text{end do} \\ & \mathbf{v}_{j+1} = \frac{\mathbf{w}}{||\mathbf{w}||} \\ & h_{j+1j} = ||\mathbf{w}|| \end{aligned} \tag{17.IV}$$

Équation (17.IV), par inversion de la matrice A on entend résolution du système linéaire $M\mathbf{w} = A\mathbf{v}_j$. D'après la remarque IV.1, nous savons que la matrice A n'est pas inversible et nous avons évoqué section IV.1 qu'il existe des valeurs propres nulles de (1.IV). Par conséquent il est nécessaire de transformer l'équation (2.IV) et de changer la valeur de l'opérateur $\mathcal{O}$ comme cela est décrit dans la section IV.3.

IV.3 Transformation "*Shift and Invert*"

IV.3.1 Introduction

Dans la définition du problème section IV.1, nous avons remarqué que la matrice A a un noyau non nul, ce qui explique la présence de valeurs propres nulles. Ainsi, le problème de valeurs propres (2.IV) ne peut pas être directement résolu. Nous allons voir dans cette section la transformation nécessaire pour capter des solutions proches d'une quantité réelle appelée "*shift*" et notée σ. Cette transformation nous permet de calculer les premières solutions non nulles de (1.IV) en choisissant la valeur de σ proche d'une quantité que nous allons détailler dans la section IV.4 et de construire un opérateur qui ne nécessite pas l'inversion de la matrice A.

IV.3.2 Transformation

On suppose que la matrice $A - \sigma M$ est inversible ce qui est vrai dans tous les cas que nous étudions (cf remarque III.6). On peut alors écrire :

$$(A-\sigma M)^{-1}M = \frac{1}{\sigma}(A-\sigma M)^{-1}(\sigma M - A + A) = \frac{-Id}{\sigma} + \frac{1}{\sigma}(A-\sigma M)^{-1}A$$

Comme le problème initial s'écrit $A\mathbf{X} = k_0^2 M\mathbf{X}$, on en déduit :

$$(A-\sigma M)^{-1}M\mathbf{X} = \frac{-\mathbf{X}}{\sigma} + \frac{k_0^2}{\sigma}(A-\sigma M)^{-1}M\mathbf{X} \tag{18.IV}$$

Et donc le problème de modes propres devient :

$$(A-\sigma M)^{-1}M\mathbf{X} = \mu\mathbf{X} \quad \text{avec} \quad \mu = \frac{1}{k_0^2 - \sigma} \tag{19.IV}$$

Le problème (19.IV) est décalé et nécessite une inversion part rapport au problème initial, d'où l'appellation "*Shift and Invert*". On obtient ainsi un nouveau problème aux valeurs propres que l'on peut résoudre par la méthode d'Arnoldi (cf section IV.2). On calcule alors les valeurs propres k_0^2 à partir des valeurs propres μ du problème (19.IV) grâce à la relation :

$$k_0^2 = \sigma + \frac{1}{\mu} \tag{20.IV}$$

L'algorithme présenté section IV.2 permet de calculer la plus grande valeur propre μ qui dans ce cas correspond à la valeur propre k_0^2 tel que $|\sigma - k_0^2|$ soit minimum. Ce qui nous conduit à la propriété importante suivante :

Proprieté IV.6. *La* transformation "*Shift and Invert*" permet de calculer des valeurs propres proches du *shift* σ.

IV.3.3 Conclusion

Nous venons de décrire la transformation nécessaire pour capter les solutions qui correspondent à des solutions proches d'une valeur σ non nulle pour éviter d'inverser la matrice A et de capter des solutions nulles. Nous allons maintenant aborder l'algorithme de calcul de valeurs propres afin d'introduire les paramètres de résolution du problème transformé (19.IV). Le choix du *shift* σ et l'ensemble des paramètres de résolutions seront repris et explicités dans la section suivante IV.4.

IV.4 Paramétrisations et précision numérique

IV.4.1 Rappel

L'algorithme global doit permettre de résoudre le problème (1.IV) en évitant de calculer trop de valeurs propres nulles. Pour ce faire on utilise la transformation "*Shift and Invert*"(cf section IV.3) et on applique l'algorithme IRAM (cf section IV.2.5) à l'opérateur $\mathcal{O} = (A - \sigma M)^{-1}M$. À chaque itération de la construction de la base d'Arnoldi (cf sectionIV.2.2), il est nécessaire d'inverser la matrice creuse $(A - \sigma M)$. En effet L'algorithme de construction du vecteur numéro $j + 1$ de la base d'Arnoldi s'écrit :

$$\begin{array}{l} (A - \sigma M)\mathbf{w} = M\mathbf{v}_j \\ \text{for } i = 1 \text{ to } j \text{ do} \\ h_{ij} =< \mathbf{w}, \mathbf{v}_i > \\ \mathbf{w} = \mathbf{w} - h_{ij}\mathbf{v}_i \\ \text{end do} \\ \mathbf{v}_{j+1} = \frac{\mathbf{w}}{||\mathbf{w}||} \\ h_{j+1j} = ||\mathbf{w}|| \end{array} \tag{21.IV}$$

À chaque itération, le système $(A-\sigma M)\mathbf{w} = M\mathbf{v}_j$ peut être résolu par plusieurs méthodes que nous avons développées et décrites dans la section (IV.6). La méthode que nous avons retenue pour nos calculs est la méthode de résolution directe utilisée par le solveur PARDISO[37, 92].

Les valeurs propres obtenues par l'algorithme IRAM[93] seront les valeurs les plus proches du *shift* σ(cf section IV.3). On rappelle dans le tableau suivant les paramètres de résolution du problème de modes propres (1.IV) :

2 matrices éléments finis creuses de dimension $n = DL$	A and M
Déphasage entre les faces maîtres et esclaves	$\phi_{ms} = \mathbf{k}_i.\mathbf{T}$
Nombre désiré de valeurs propres	$\bar{k}$
Nombre de vecteurs d'Arnoldi	$\bar{p}$
Critère de tolérance	$\bar{tol}$
Shift proche des valeurs cherchées	σ

Fig 1.IV Tableau récapitulatif des paramètres de résolution principaux.

IV.4.2 Initialisation du shift

Nous allons voir que le *shift* σ peut être changé au cours de la résolution, mais on peut deviner avec plus ou moins d'exactitude quelles seront les premières valeurs propres en fonctions des dimensions géométriques de la cellule de base. Dans tous les cas σ doit être suffisamment éloigné de ces premières solutions pour ne pas capter trop de valeurs propres nulles. De plus le rayon de convergence dépend du nombre de valeurs propres désirées $\bar{k}$. Autrement dit, plus $\bar{k}$ est grand plus on aura la possibilité de capter des solutions éloignées de σ. On peut approcher les premières solutions par les valeurs analytiques d'une cavité métallique de mêmes dimensions que le parallélépipède qui contient la cellule de base (cf section III.3.5.1). Dans un parallélépipède rectangle les valeurs propres solutions sont donnés par la formule [94] :

$$(k_0^2)_{mkp} = \frac{\pi^2}{\epsilon_r \mu_r}\left(\frac{p^2}{L^2} + \frac{m^2}{l^2} + \frac{k^2}{h^2}\right) \; (m,k,p) \in \mathbb{N}^3$$

Où L, l et h représentent respectivement la longueur, la largeur et la hauteur du parallélépipède rectangle.

Remarque IV.7. Comme la cellule de base est contenue dans un parallélépipède rectangle on

peut choisir le *shift* initial de la façon suivante :

$$\sigma = 3\pi^2 \min(\frac{1}{L^2}, \frac{1}{l^2}, \frac{1}{h^2}) \sum_{V_i \in V_{tot}} \frac{1}{\epsilon_{r_i}\mu_{r_i}} \frac{V_i}{V_{tot}} \tag{22.IV}$$

$\epsilon_{r_i}, \mu_{r_i}$ sont les permittivités et les perméabilités des différents sous domaines et V_i leurs volumes respectifs.

Le parallélépipède qui contient la cellule de base peut contenir plusieurs motifs de tailles différentes (cf section V.4.2) qui peuvent résonner à des fréquences plus hautes que les premières fréquences de la cavité. La longueur d'onde des premières valeurs propres est donc de l'ordre de la taille des éléments résonnants. On peut aussi définir le *shift* σ à partir de la taille caractéristique notée d_c des éléments résonnants de la cellule de base. d_c est la distance entre des motifs résonnants périodisés. Ainsi défini le *shift* σ s'écrit en fonction de cette distance :

$$\sigma = (\frac{2\pi}{d_c})^2 \tag{23.IV}$$

Dans le cas d'une structure périodique dans une direction, d_c correspond simplement au pas du réseau. Pour une structure périodique dans 2 ou 3 directions on choisit d_c égale à la diagonale de la cellule de base pour tenir compte dans cette distance caractéristique des éventuelles dissymétries selon les différentes directions :

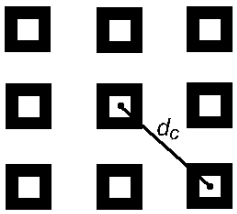


Fig 2.IV d_c distance caractéristique d'un réseau en 2 dimensions.

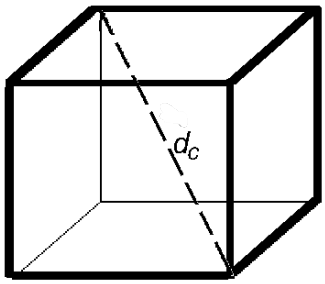


Fig 3.IV d_c : distance caractéristique d'un réseau en 3 dimensions.

La distance caractéristique peut être utilisée dans le cas où la cellule de base comporte plusieurs éléments de tailles différentes ce qui est le cas des super-cellules abordées dans la section V.4.2.

Remarque IV.8. La valeur initiale donnée à σ n'est qu'une estimation est sert de point de départ pour le solveur. C'est pourquoi nous avons laissé à l'utilisateur la possibilité de modifier la valeur du *shift* initial défini équation (22.IV) via le facteur multiplicatif K_σ (cf section V.2.2) ou en donnant une valeur de d_c exprimée en mètre. D'autre part sa valeur varie pendant le calcul en fonction de certains critères que nous allons énoncer dans la section suivante IV.4.5.

IV.4.3 Choix du nombre de vecteur d'Arnoldi $\bar{p}$

Les paramètres sont liés entre eux n'ont pas besoin de tous être renseignés. La précision numérique va dépendre

- du critère de convergence de chaque valeur propre (IV.2) contrôlé par le réel $\bar{tol}$,
- du nombre de vecteur d'Arnoldi $\bar{p}$
- et du nombre désiré de valeurs propres $\bar{k}$.

En raison de la structure de l'algorithme d'Arnoldi on peut facilement comprendre que $\bar{p}$ doit être supérieur à $\bar{k}$ et la relation (12.IV) nous permet de voir que $\bar{p}$ doit être suffisamment grand pour assurer que le critère de convergence (13.IV) soit satisfait (cf remarque IV.6). C'est la raison pour laquelle nous avons décidé de lier les paramètres $\bar{k}$ et $\bar{p}$ de la façon suivante :

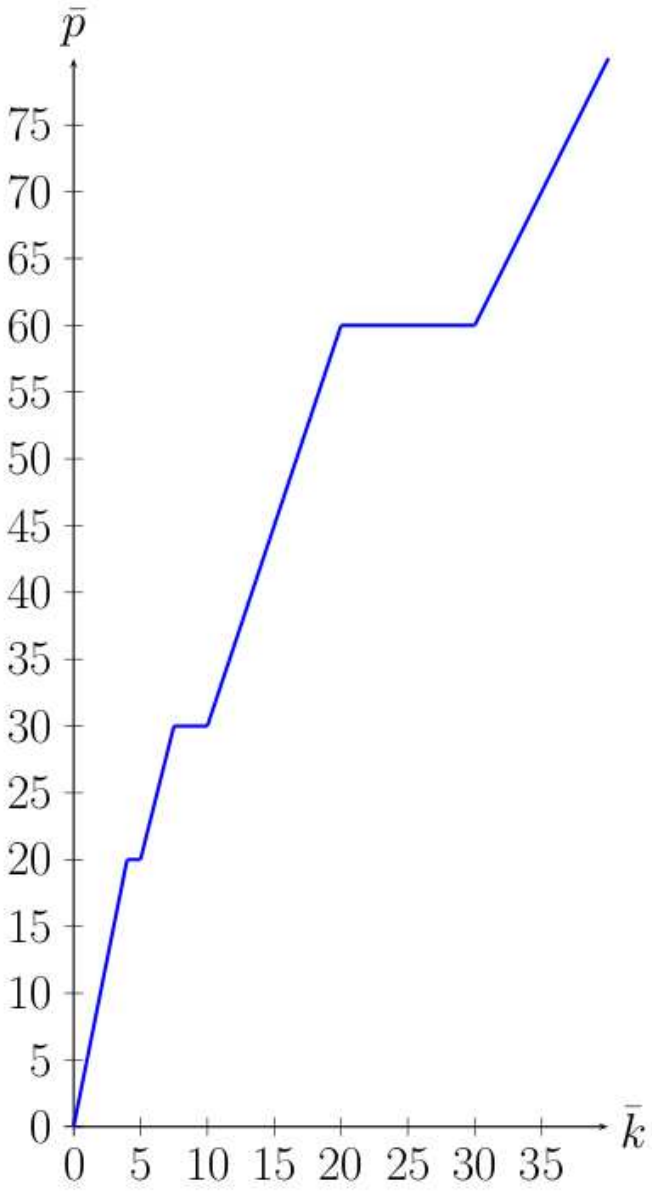


Fig 4.IV Relations entre $\bar{k}$ et $\bar{p}$.

Fig 4.IV $\bar{p}$ est choisi plus grand que $\bar{k}$, et $\bar{p}$ diminue quand $\bar{k}$ est grand pour optimiser l'espace mémoire utilisé. ARPACK laisse la possibilité à l'utilisateur de modifier $\bar{k}$ et $\bar{p}$ mais $\bar{k}$ doit toujours être strictement inférieur à $\bar{p}$. Le nombre de vecteur d'Arnoldi $\bar{p}$ doit être minimal pour optimiser l'espace mémoire utilisé, mais si $\bar{p}$ est choisi proche de $\bar{k}$, des redémarrages supplémentaires de l'algorithme IV.2.5 seront nécessaires pour qu'un nombre suffisant de valeurs propres vérifient le critère (13.IV). Le choix le plus optimal est donc celui qui va minimiser à la fois le nombre de redémarrage et le nombre de vecteurs d'Arnoldi $\bar{p}$ en vérifiant le critère (13.IV) pour les $\bar{k}$ premières valeurs propres. Dans les problèmes que nous adressons, nous avons du prendre une décision pragmatique pour gagner du temps et pour les résoudre de façon automatique.

IV.4.4 Critère d'élimination des valeurs propres nulles

Pour chaque mode propre solution $(\lambda, \mathbf{X})$, on introduit le résidu ϵ qui permettra d'éliminer les valeurs propres nulles selon le critère :

$$\text{si } \epsilon = \frac{\|A\mathbf{X} - \lambda M\mathbf{X}\|}{|\lambda|} \leq \bar{tol} \text{ alors } (\lambda, \mathbf{X}) \text{ est un mode acceptable} \tag{24.IV}$$

Ce critère vérifie à la fois que $(\lambda, \mathbf{X})$ est bien une solution de (1.IV) et que λ est non nulle. Le réel positif $\bar{tol}$ est le même que celui choisi pour sélectionner les valeurs propres convergentes dans équation (13.IV). Les équations (24.IV) et (13.IV) n'ont pas de lien direct mais dans les deux cas le réel positif $\bar{tol}$ doit être choisi inférieur à 10^{-3}, pour assurer d'une part la convergence des valeurs propres, et d'autre part l'élimination des valeurs propres nulles. Encore une fois il est sans doute possible d'optimiser le choix de $\bar{tol}$ mais nous avons du faire un choix pragmatique qui fonctionne dans tous les cas que nous résolvons.

IV.4.5 Réglage du *shift* pendant le calcul

Au cours du calcul, le *shift* σ doit être ajusté pour éviter de capter trop de valeurs propres nuls. On propose ici un algorithme de tri automatique qui permet de calculer tous les points d'un diagramme de bandes sans faire interagir un utilisateur extérieur. Cette méthode empirique fonctionne bien pour tous les cas que nous avons étudié et permet de minimiser le nombre de redémarrage pour $\bar{k}$ de l'ordre de quelques dizaines. Voici *l'algorithme de tri* que nous avons mis en œuvre :

- Pour chaque direction incidente $\mathbf{k}_i$, on calcule $\bar{k}$ valeurs propres de l'opérateur $\mathcal{O} = (A - \sigma(\bar{k})M)^{-1}M$.
- Parmi les solutions, n_s vont remplir le critère d'acceptation (24.IV). Ces n_s solutions sont stockées dans le tableau qu'on appelle Val.
- Si $n_s < 2 * \bar{k}/3$, σ est augmenté, le solveur est redémarré et les valeurs obtenues ne

sont pas conservées. On change l'ancien $\sigma = \sigma_{old}$ en nouveau $\sigma = \sigma_{new}$:

$$\sigma_{new} = \frac{1}{4}\max(Val) + \frac{3}{4}\sigma_{old}$$

– Si $n_s = \bar{k}$, σ est diminué, le solveur est redémarré et les valeurs obtenues ne sont pas conservées. On change l'ancien $\sigma = \sigma_{old}$ en nouveau $\sigma = \sigma_{new}$:

$$\sigma_{new} = \frac{1}{4}\min(Val) + \frac{3}{4}\sigma_{old}$$

– Sinon les valeurs obtenues sont conservées et affichées et on change l'ancien $\sigma = \sigma_{old}$ en nouveau $\sigma = \sigma_{new}$:$\sigma_{new} = \frac{n_s - 1}{n_s}\sigma_{old} + \frac{\bar{Val}}{n_s}$ ($\bar{Val}$ désigne la valeur moyenne de Val).

Pour chaque incidence $\mathbf{k}_i$, les valeurs conservées sont stockées dans un fichier de données dans l'ordre croissant (cf section V.2). Si le *shift* σ est choisi trop petit au départ il y aura un nombre de redémarrages nécessaires pour qu'il y ait suffisamment de solutions qui vérifient (24.IV). Si σ est choisi trop grand toutes les solutions vérifieront (24.IV), autrement dit n_s est égal à $\bar{k}$ et l'algorithme est redémarré jusqu'à ce que $n_s < \bar{k}$. Ainsi le choix de la valeur initiale du *shift* σ va jouer un rôle déterminant dans le temps de calcul.

Remarque IV.9. Cet algorithme a été mis en place pour automatiser le calcul de diagrammes de bandes. Lorsque $\bar{k}$ est de l'ordre de quelques dizaines cet algorithme ne génère pratiquement pas de redémarrage quand le shift initial est bien choisi comme on peut l'observer dans les résultats des sections V.3 et V.4. Cependant cette modification du *shift* pendant le calcul n'est pas la seule possible et nous avons fait un choix pragmatique basé sur l'expérience.

Remarque IV.10. Nous constatons que l'ajout de conditions PML conduit à réduire le nombre de valeurs propres nulles. Comme nous avons vu section III.3.3 l'ajout de conditions PML revient à multiplier la permittivité et la perméabilité par des coefficients complexes et à ajouter un volume d'absorption supplémentaire. Dans ce cas la structure des matrices influence la convergence des résultats.

IV.4.6 Considérations sur le maillage

La taille des tétraèdres joue un rôle essentiel dans l'approximation du résultat. Des résultats classiques[95] peuvent nous donner l'erreur commise localement dans le tétraèdre entre la solution véritable et la solution cherchée. Par exemple pour une triangulation $\mathcal{T}_h$ de l'espace Ω et une fonction $\boldsymbol{\phi} \in H(rot, \Omega) \cap H^s(\Omega)^3, 0 \leq s \leq 1$, on a pour un tétraèdre $K \in \mathcal{T}_h$:

$$\|\boldsymbol{\phi} - \boldsymbol{\phi}_h\| \leq Ch^s.$$

Avec $\boldsymbol{\phi}_h$ la solution approchée de la solution réelle $\boldsymbol{\phi}$. Ce type de résultat peut nous donner une idée de la taille maximale de maille à choisir mais ce n'est pas suffisant. En effet cette approximation ne permet pas de donner d'informations sur la taille des mailles à choisir dans les zones importantes de fort gradient de champs par exemple à proximité d'objets

métalliques pointus. Le raffinement dans les zones étroites est indispensable pour que la réalité physique du problème soit correctement modélisée. En effet un maillage grossier dans les zones étroites ou discontinues ne prendrait pas en compte les variations géométriques brutales et ainsi remettrait en cause la validité des résultats obtenus. Afin de savoir exactement quelle partie nous devons mailler plus finement, nous pouvons itérer sur le maillage comme le logiciel HFSS[96] et comparer après chaque calcul les variations locales des solutions. Ce processus, nous permet d'observer que les mailles plus fines doivent être situées aux endroits où la géométrie est discontinue ou varie brusquement. L'inconvénient de ce type de méthode est l'ajout d'itérations supplémentaires qui, dans certains cas, peuvent faire exploser le temps de calcul. C'est pourquoi nous avons décidé que tous nos calculs devaient se faire directement sur un maillage suffisamment fin pour garantir une précision physique du résultat. Plus précisément, quand on étudie la structure autour d'une longueur d'onde λ_0 donnée, les zones du maillage qui ont des tailles d'arêtes inférieures à $\lambda_0/10$ sont les zones qui comportent des matériaux fortement diffractants, ou les zones "étroites" où la géométrie varie fortement. Cette grandeur $\lambda_0/10$ n'est qu'une estimation globale insuffisante dans certains cas notamment quand la taille d'un obstacle est très petite devant λ_0. Prenons l'exemple de la surface haute impédance[20] :

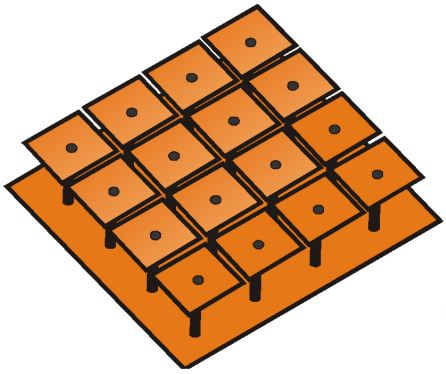


Fig 5.IV Surfaces Haute Impédance[22].

Le maillage est plus fin entre les champignons et sur les tiges qui sont des zones fortement diffractantes.

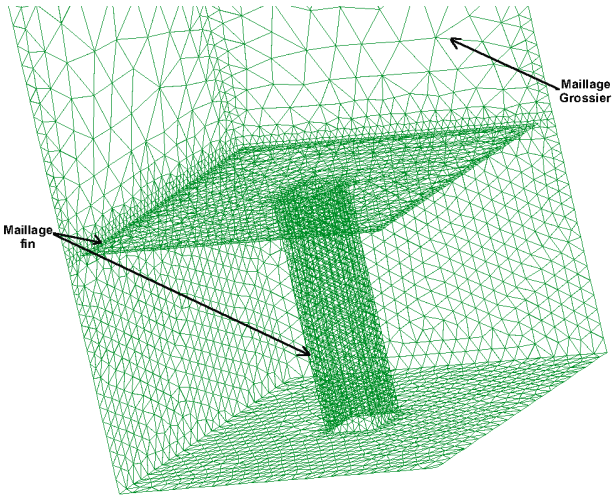


Fig 6.IV Maillage de surfaces de la cellule de base du réseau de champignons métalliques.

Remarque IV.11. Dans le cas d'une résolution en modes propres, on obtient un spectre de solutions comprises entre k_{min} et k_{max}. Le maillage doit être défini par rapport à la longueur d'onde minimale λ_{min} correspondant à la plus grande solution k_{max}. La précision du résultat est alors meilleure pour les premières valeurs proches de k_{min} correspondantes aux plus grandes longueurs d'ondes λ_{max}.

IV.5 Élimination automatique des solutions parasites

IV.5.1 Introduction

Comme nous l'avons évoqué dans la section IV.1, le problème (1.IV) admet des valeurs propres nuls que l'on a éliminées avec la transformation *Shift and Invert* abordée dans la section IV.3 et le critère de convergence (24.IV). Nous allons, dans cette section, décrire et classer ces valeurs propres nulles selon les critères de divergences énoncés en (26.IV). Ensuite dans la section (IV.5.2) nous décrivons l'algorithme permettant d'éliminer les valeurs propres nulles à divergences non nulles appelées modes propres parasites. Pour une meilleure compréhension du phénomène nous étudions le cas des cavités métalliques. Nous allons voir que dans certains cas, ces résultats sont transposables aux cas des structures périodiques. On se place dans une cavité métallique regroupant un domaine intérieur hétérogène non dispersif Ω, de frontière parfaitement conductrice $\partial\Omega$ et de normale extérieure $\boldsymbol{\nu}$. On suppose que la permittivité et la perméabilité relative sont des matrices 3×3 hermitiennes notées $[\epsilon_r]$ et $[\mu_r]$. L'espace de recherche des solutions $\mathcal{H}$ est noté $H_0(\text{curl}, \Omega)$ et l'espace d'approximation de $H_0(\text{curl}, \Omega)$ est noté $\mathcal{H}_h$. On rappelle que $H_0^1(\Omega)$ désigne le sous-espace vectoriel de $H^1(\Omega)$ obtenu par la complétion des fonctions $C^\infty(\Omega)$ à support compact dans Ω[69]. Soit alors la formulation dérivant des équations de Maxwell Harmonique dans Ω [34] :

$$\boldsymbol{\nabla} \times ([\mu]_r^{-1} \boldsymbol{\nabla} \times \mathbf{E}) - k_0^2 \epsilon_r \mathbf{E} = \mathbf{0}, \text{ dans } \Omega \tag{25.IV}$$

On cherche à calculer les solutions $(k_0^2, \mathbf{E})$. Si on suppose que le champ $\mathbf{E}$ dérive d'un potentiel (i.e $\mathbf{E} = -\boldsymbol{\nabla}\phi,\ \phi \in H_0^1(\Omega)$), alors la solution k_0^2 est nulle et en prenant la divergence de l'équation (25.IV) on a :

$$\boldsymbol{\nabla}.[\epsilon_r]\mathbf{E} = 0 \text{ dans } \Omega$$

Comme cela est indiqué dans [97], on peut classer les modes propres solutions en trois groupes :

$$\begin{cases} \text{Groupe.1 } (k_0^2, \mathbf{E}) \text{ solutions de (25.IV) tel que } k_0^2 \neq 0 \text{ et } \boldsymbol{\nabla}.[\epsilon_r]\mathbf{E} = 0 \\ \text{Groupe.2 } (k_0^2, \mathbf{E}) \text{ solutions de (25.IV) tel que } k_0^2 = 0 \text{ et } \boldsymbol{\nabla}.[\epsilon_r]\mathbf{E} = 0 \\ \text{Groupe.3 } (k_0^2, \mathbf{E}) \text{ solutions de (25.IV) tel que } k_0^2 = 0 \text{ et } \boldsymbol{\nabla}.[\epsilon_r]\mathbf{E} \neq 0 \end{cases} \tag{26.IV}$$

Les modes propres parasites sont les solutions du Groupe.3. La transformation que nous avons vu dans la section IV.3 permet, grâce au critère (24.IV) déliminer les solutions du Groupe.3 et du Groupe.2. Dans le cas des cavités métalliques il est possible, grâce à la construction d'une contrainte discrète abordée section IV.5.2 d'éliminer seulement les solutions du Groupe.3.

IV.5.2 Élimination des modes propres parasites dans les cavités métalliques

Les modes propres parasites sont les modes solution de l'équation (26.IV) qui appartiennent au Groupe.3 [34]. Pour éliminer ces solutions parasites, on transforme le problème

initial en ajoutant une contrainte de divergence à l'équation (25.IV) :

$$\begin{cases} \boldsymbol{\nabla} \times [\mu_r]^{-1} \boldsymbol{\nabla} \times \mathbf{E} - k_0^2 [\epsilon_r] \mathbf{E} = 0, \quad \text{dans } \Omega \\ \mathbf{E}_T = 0, \quad \text{dans } \partial\Omega \\ \boldsymbol{\nabla}.[\epsilon_r]\mathbf{E} = 0, \quad \text{dans } \Omega \end{cases} \tag{27.IV}$$

La formulation faible de l'équation (27.IV) avec contrainte (cf section III.2) s'écrit alors :

$$\begin{cases} a(\mathbf{E}, \boldsymbol{\varphi}) = k_0^2 m(\mathbf{E}, \boldsymbol{\varphi}) \ \ \forall\ \boldsymbol{\varphi} \in H_0(\text{curl}, \Omega) \\ \\ \displaystyle\int_\Omega \boldsymbol{\nabla} U.[\epsilon_r]\mathbf{E} = (\boldsymbol{\nabla} U.[\epsilon_r]\mathbf{E}) = 0 \ \ \forall U \in H_0^1(\Omega), \\ \\ \text{avec } a(\mathbf{E}, \boldsymbol{\varphi}) = \displaystyle\int_\Omega \boldsymbol{\nabla} \times \boldsymbol{\varphi}.\frac{\boldsymbol{\nabla} \times \mathbf{E}}{\mu_r}, \\ \\ \text{et } m(\mathbf{E}, \boldsymbol{\varphi}) = \displaystyle\int_\Omega \epsilon_r \mathbf{E}.\boldsymbol{\varphi} \end{cases} \tag{28.IV}$$

D'après [74] l'espace d'approximation $\mathcal{H}_h$ peut s'écrire sous la forme d'une somme directe :

$$\mathcal{H}_h = \boldsymbol{\nabla}\Pi \oplus Q \text{ avec } \boldsymbol{\nabla}\Pi = \{\boldsymbol{\nabla}\phi | \phi \in \mathcal{H}_{N_n}^0\} \text{ et } Q = (\boldsymbol{\nabla}\Pi)^\perp$$

Où $\mathcal{H}_{N_n}^0$ est l'espace de dimension fini qui est dense dans $H_0^1(\Omega)$. Sa dimension N_n est égale au nombre total de nœuds moins les nœuds de $\partial\Omega$. Dans $\mathcal{H}_{N_n}^0$, on approche la solution en chaque nœuds par un polynôme de degré 1. Autrement dit on choisit les éléments finis P_1 [69]. La dimension de Π est égale à N_n, qui est égale au nombre de modes propres parasites et les éléments de Q sont à divergence nulle [34]. Nous allons maintenant voir comment l'on peut construire une méthode numérique efficace qui traduit la contrainte de divergence nulle (28.IV) en s'appuyant sur le gradient discret[34] que nous allons décrire dans la section suivante.

IV.5.2.1 Le gradient discret

On désigne par DL et N_n le nombre d'arêtes et de nœuds du domaine qui ne sont pas sur $\partial\Omega$. Soit $\mathbf{u}_{DL} \in \boldsymbol{\nabla}\Pi \cap \mathcal{H}_h$ alors on peut écrire [34] :

$$\sum_{i=1}^{DL} u_i \mathbf{W}_i = \sum_{l=1}^{N_n} \xi_l \boldsymbol{\nabla}\lambda_l \tag{29.IV}$$

Où $\mathbf{W}_p$ est la fonction de Nedelec associée à l'arête $\mathbf{Ar}_p$, et λ_l la coordonnée barycentrique associée au nœud $\mathbf{N}_l$. Soit l'identité :

$$\int_{edge\ i_0=\{j,k\}} \sum_{i=1}^{DL} u_i \mathbf{W}_i = \int_{edge\ i_0=\{j,k\}} \sum_{l=1}^{N_n} \xi_l \boldsymbol{\nabla}\lambda_l \Rightarrow u_{i_0} = \xi_k - \xi_j. \tag{30.IV}$$

Si on note par $\mathbf{U}_{DL} = (u_1, \ldots, u_{DL})$, $\boldsymbol{\Xi} = (\xi_1, \ldots, \xi_{N_n})$, alors l'équation (30.IV) peut s'écrire [34] :

$$\mathbf{U}_{DL} = G\boldsymbol{\Xi} \tag{31.IV}$$

G est l'opérateur gradient discret. L'opérateur G est une matrice de taille $DL \times N_n$ qui ne comporte que des 1,-1 et des 0.

Considérons alors la matrice $A = (A)_{mn} = a(\mathbf{W}_n, \mathbf{W}_m)$,

la matrice $M = (M)_{mn} = m(\mathbf{W}_n, \mathbf{W}_m)$ et $\underline{\mathbf{W}} = (\mathbf{W}_1, \ldots, \mathbf{W}_{DL})$ on a [34] :

$$\begin{aligned} &\mathbf{u}_{DL} = \mathbf{U}_{DL}^T \underline{\mathbf{W}} \in \boldsymbol{\nabla}\Pi \cap \mathcal{H}_h \text{ donc } \boldsymbol{\nabla} \times \mathbf{u}_{DL} = \mathbf{0} \Rightarrow a(\mathbf{u}_{Dl}, \boldsymbol{\phi}) = 0, \ \forall \boldsymbol{\phi} \in \mathcal{H}_h \\ &\Rightarrow AG\boldsymbol{\Xi} = 0, \forall \boldsymbol{\Xi} \in \mathbb{R}^{N_n} \Rightarrow AG = 0. \\ &\text{et} \\ &\mathrm{Im}(G)^{\perp} = \ker(G^T) \Rightarrow G^T A = 0 \end{aligned} \tag{32.IV}$$

De plus, si on pose $\mathbf{u}_s = \mathbf{U}_s^T \underline{\mathbf{W}} \in \mathcal{H}_h$ un vecteur solution de (28.IV) alors on a [34] :

$$(\mathbf{s}, [\epsilon_r]\mathbf{u}_s) = \mathbf{S}^T M \mathbf{U}_s = 0, \quad \forall \mathbf{s} = \mathbf{S}^T \underline{\mathbf{W}} \in \boldsymbol{\nabla}\Pi \tag{33.IV}$$

Par définition de l'opérateur G il existe $\boldsymbol{\phi} \in \mathbb{R}^{N_n}$ tel que $\mathbf{S} = G\boldsymbol{\phi}$ donc la contrainte faible de divergence nulle implique [34] :

$$\boldsymbol{\phi}^T G^T M \mathbf{U}_s = 0 \ \forall \boldsymbol{\phi} \in \mathbb{R}^{N_n} \Rightarrow G^T M \mathbf{U}_s = 0 \tag{34.IV}$$

On en déduit d'après (34.IV) et (32.IV) qu'un vecteur solution de (28.IV) $\mathbf{u}_s = \mathbf{U}_s^T \underline{\mathbf{W}}$ vérifie [34] :

$$G^T(\alpha A + \beta M)\mathbf{U}_s = 0, \quad \alpha, \beta \in \mathbb{R}, \beta \neq 0 \tag{35.IV}$$

IV.5.2.2 Contrainte de divergence dans l'algorithme de Lanczos

Soit $(\mathbf{p}, \mathbf{q}) \in (\mathbb{C}^{DL})^2$. Chaque itération de l'algorithme d'Arnoldi abordé dans la section IV.2, combiné avec la transformation *Shift and Invert* (cf section IV.3), nécessite de résoudre le système (cf section IV.4 équation (21.IV)) :

$$(A - \sigma M)\mathbf{p} = \mathbf{q} \tag{36.IV}$$

Remarque IV.12. Comme la contrainte de divergence nulle va éliminer toutes les solutions du Groupe.3 (26.IV), alors le *shift* σ peut être négatif pour capter les premières solutions du Groupe.2 et du Groupe.1. Comme la matrice A est semi définie positive et M définie positive (cf remarque III.6), $(A - \sigma M)$ est définie positive et les algorithmes de résolution de système linéaire peuvent alors être optimisés. Par exemple pour résoudre l'équation (36.IV) on peut utiliser le gradient conjugué[98] qui n'a besoin que de multiplications matricielles, ou bien l'algorithme MDOA([99] et section IV.6.5.2). Pour équilibrer les coefficients des matrices on

choisit un *shift* [34] :

$$\sigma = -0.1k_{min}^2 \tag{37.IV}$$

Ou k_{min}^2 représente une estimation de la première solution non nulle. Une bonne approximation de cette solution est donnée par la relation (22.IV).

D'après l'équation (35.IV) la contrainte de divergence peut alors s'écrire [34] :

$$G^T(A - \sigma M)\mathbf{p} = G^T\mathbf{q} = 0. \tag{38.IV}$$

La matrice G n'est pas carrée on la sépare alors en deux parties :

$$G = \begin{bmatrix} G_{N_n} \\ G_{N_c} \end{bmatrix}$$

La matrice G_{N_n} est de dimension $N_n \times N_n$. Elle se construit en choisissant les N_n premières arêtes de l'identité (30.IV) de façon à ce qu'elle soit inversible. Cette construction est détaillée dans la section suivante IV.5.2.3. Ainsi la contrainte de divergence s'écrit [34] :

$$\begin{bmatrix} G_{N_n}^T & G_{N_c}^T \end{bmatrix} \begin{bmatrix} \mathbf{q}_{N_n} \\ \mathbf{q}_{N_c} \end{bmatrix} = 0. \tag{39.IV}$$

Et à chaque itération d'Arnoldi on remplace le vecteur $\mathbf{q}$ par le vecteur :

$$\mathbf{q}' = \begin{bmatrix} -(G_{N_n}^T)^{-1} G_{N_c}^T \mathbf{q}_{N_c} \\ \mathbf{q}_{N_c} \end{bmatrix} \tag{40.IV}$$

IV.5.2.3 Quelles arêtes choisir pour construire le gradient discret ?

On rappelle que le gradient discret G a pour dimension $DL \times N_n$ ou N_n est le nombre de nœuds et DL le nombre d'arêtes de Ω_h qui ne sont pas sur $\partial\Omega$. Pour construire la matrice G_{N_n} ci-dessus il faut sélectionner autant d'arêtes qu'il y a de nœuds à l'intérieur de Ω en choisissant le premier nœud sur une face PEC :

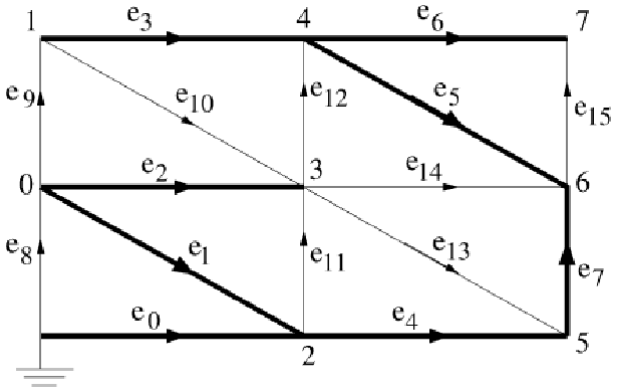


Fig 7.IV Choix des arêtes pour construire G [34].

Sur la Fig 7. IV les arêtes en noires sont correspondantes aux arêtes choisis pour construire

les N_n premières lignes de G. Les algorithmes utilisés pour le choix des arêtes en noires sont détaillés en annexe A.7. Un fois les arêtes choisies, il existe une façon simple de construire la matrice $G_{N_n}^{-1}$ qui, comme la matrice G_{N_n}, n'est constituée que de 0, de 1 et de -1 (41.IV). En effet la ligne i de $G_{N_n}^{-1}$ équation (40.IV) correspond au chemin à parcourir pour atteindre le i-ème nœud depuis le nœud de référence, en tenant compte de l'orientation des arêtes. Par exemple Fig 7.IV, pour atteindre le nœud 0 nous devons passer par l'arête e_0 et l'arête e_1. Ainsi la première ligne de la matrice $G_{N_n}^{-1}$ aura le coefficient 1 dans la première colonne et -1 dans la deuxième colonne :

$$G_{N_n}^{-1} = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 & 1 & -1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & -1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & -1 & 1 & 1 \end{bmatrix} \tag{41.IV}$$

Remarque IV.13. Malheureusement la contrainte de divergence nulle dans l'algorithme de Lanczos ne peut pas être appliquée pour les structures périodiques lorsque l'on utilise les fonctions d'arêtes de Nedelec d'ordre 0. En effet, les contraintes dues aux relations entre les arêtes maîtres-esclaves ajoutées aux contraintes nécessaires sur le choix des nøeuds pour construire G_{N_n} sont trop fortes et ne permettent pas de sélectionner les arêtes nécessaires. Mais il existe d'autres techniques qui utilisent notamment la montée en ordre pour palier à ce type de problème [36].

IV.6 Stockage des matrices

Par construction, les matrices éléments finis A et M définies avec la méthode de Galerkin (cf section III.2.2) sont creuses. En effet, pour un indice i, les indices j pour lesquels les coefficients $(A)_{ij}$ et $(M)_{ij}$ sont non nuls correspondent à des arêtes appartenant à un même tétraèdre K contenant l'arête $\mathbf{Ar}_i$. De plus nous avons vu section IV.4 qu'une combinaison linéaire des deux matrices doit être inversée. À partir de ces constats, nous devons alors trouver un moyen efficace de stocker ces matrices afin d'optimiser l'espace mémoire lors des calculs élémentaires notamment utilisés pour résoudre les systèmes linéaires de la construction de la base d'Arnoldi équation (21.IV). Pour ce faire, on doit d'abord construire le graphe associé aux matrices A et M. Par définition, le graphe de la matrice relie les indices des éléments non nuls des matrices A et M :

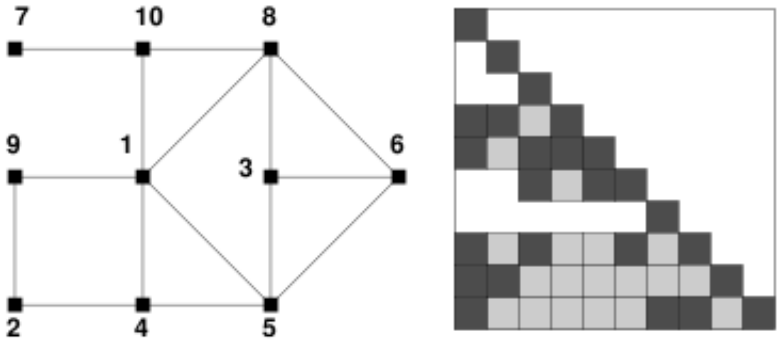


Fig 8.IV Graphe associé à la matrice [100].

En connaissant la structure du maillage on peut construire le graphe de la matrice. Nous allons maintenant présenter les deux méthodes que nous avons utilisées pour stocker ces matrices.

IV.6.1 Matrices de bandes

Les matrices creuses peuvent être stockées sous forme de matrices de bandes. C'est-à-dire que par une re-numérotation des inconnues on peut faire en sorte que tous les coefficients non nuls soient autour de la diagonale. Mais avant d'expliciter l'algorithme qui permet de faire cette transformation nous allons donner une définition plus précise d'une matrice de bandes.

Définition. Matrice de bandes
Une matrice de bandes est une matrice dont les coefficients non nuls sont stockés autour de la diagonale. Formellement : une matrice carrée $A = (a_{ij})$ d'ordre N est appelée matrice de bandes si tous les coefficients a_{ij} non nuls sont situés dans une bande délimitée par des parallèles à la diagonale principale. Une telle bande est déterminée par deux entiers k_1 et k_2

positifs ou nuls, tels que :

$$a_{ij} = 0 \quad \text{si} \quad j < i - k_1 \quad \text{ou} \quad j > i + k_2$$

Pour réduire l'espace mémoire, on stocke les éléments au-dessus et en-dessous de la diagonale sous la forme de tableaux rectangulaires.

Par exemple la matrice :

$$\begin{bmatrix} 4 & 0 & 0 & 0 & 0 \\ 4 & 9 & 0 & 0 & 0 \\ 0 & 8 & 9 & 0 & 0 \\ 0 & 0 & 2 & 13 & 0 \\ 0 & 0 & 0 & 5 & 16 \end{bmatrix} \text{ est stockée sous la forme : } \begin{bmatrix} 4 & 0 \\ 4 & 9 \\ 8 & 9 \\ 2 & 13 \\ 5 & 16 \end{bmatrix}$$

Les matrices creuses peuvent être transformées en matrices de bandes grâce à l'algorithme de Cuthill-Mc Kee inverse qui permet de trouver la largeur de bande la plus petite possible et ainsi stocker les éléments non nuls autour de la diagonale. Nous allons maintenant décrire cet algorithme.

IV.6.2 Algorithme de Cuthill et Mc Kee

L'algorithme de Cuthill et Mc Kee [101] donne une numérotation des éléments du graphe de la matrice permettant de pouvoir stocker les coefficients non nuls autour de sa diagonale. En voici une première version simplifiée [100] :

- On commence par choisir un sommet.
- On numérote ses voisins en choisissant en premier ceux qui ont le moins de voisins non encore numérotés.
- On numérote les voisins de ses voisins qui ne sont pas encore numérotés, ...

Plus précisément cet algorithme correspond à un parcours en largeur du graphe. Chaque sommet du graphe n'est pas visité plus d'une fois :

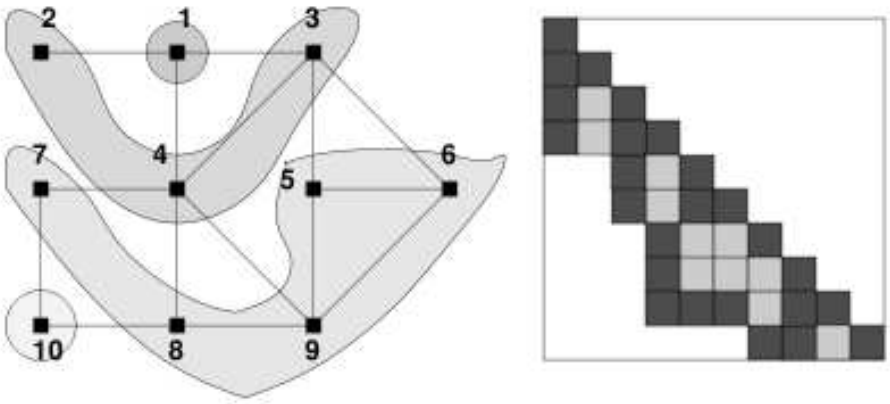


Fig 9.IV Matrice et graphe de la Fig 8.IV après application de Cuthill-Mc Kee direct[100].

La transformation de Cuthill et Mc Kee revient à minimiser les distances entre les indices des sommets voisins. D'autre part, la distance entre les indices des sommets peut encore être minimisée si l'on "renverse" leur numérotation [100] :

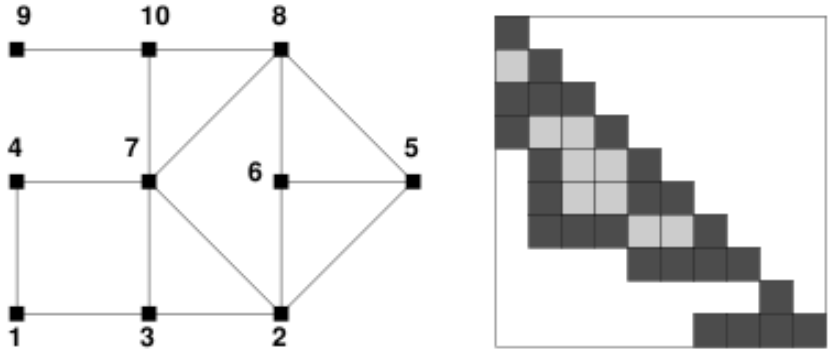


Fig 10.IV Matrice et graphe Fig IV.9 après inversion des sommets [100].

Une renumérotation dans l'ordre inverse revient à faire une symétrie par rapport à l'autre diagonale. Cette transformation ne crée aucun zéro supplémentaire.

IV.6.2.1 Choix d'un premier sommet

Soit $d(i,j) = |i-j|$ la distance du nœud i au nœud j, on définit l'excentricité d'un nœud i du graphe G par :

$$exc(i) = \max_{j \in G} d(i,j)$$

L'algorithme est d'autant plus efficace que l'excentricité du premier sommet est maximale. On construit alors un algorithme de recherche du sommet d'excentricité maximale :

1. Soit r le sommet construit à l'étape n.
2. On construit les successeurs (ou voisins) de r. Soit $D(r)$ l'ensemble des successeurs.

3. On choisit alors à l'étape $n+1$ le sommet $i \in D(r)$ non encore choisi, tel que $exc(i) > exc(j), \forall j \in D(r)$.

En appliquant l'Algorithme de Cuthill-Mc Kee au graphe des arêtes ou nœuds du domaine, les matrices élémentaires deviennent des matrices de bandes que l'on peut directement stocker sous forme de tableaux.

Remarque IV.14. Lorsque que l'on définit des conditions PEC sur l'un des bords du domaine (cf section III.3.2), on peut directement appliquer Cuthill-Mc Kee sur le graphe du maillage qui ne contient pas les arêtes PEC.

Le stockage de bandes a pour avantage de pouvoir inverser facilement les matrices en codant les algorithmes directs classiques du type "descente remontée LU"[102]. De plus le gain de mémoire est considérable lors d'un produit matrice vecteur. Bien que ce stockage soit efficace et facile à implémenter, le stockage YSM abordé dans la section suivante IV.6.3, qui consiste à ne ranger que les coefficients non nuls sous forme de tableaux est plus efficace autant dans le temps de calcul d'opérations élémentaires que dans l'espace mémoire utilisé. Cependant, les différents algorithmes LU appliqués aux matrices YSM sont plus complexes à coder (cf section IV.6.5). Durant le cours de la thèse, l'algorithme de Cuthill et Mc Kee a été implémenté dans un premier temps pour stocker et assembler les matrices creuses sous la forme de matrices de bandes et ainsi permettre une implémentation classique de l'algorithme de factorisation LU. Dans un second temps nous avons utilisé le solveur PARDISO[37] qui nous a permis une factorisation LU d'une matrice creuse stockée sous le format Yale Sparse Matrix que nous allons décrire dans la section suivante.

IV.6.3 Stockage au format *Yale Sparse Matrix*

Grâce au graphe de la matrice (cf Fig IV.8), on peut construire les tableaux IA et JA du format *Yale Sparse Matrix* [103] et ensuite assembler les coefficients non nuls stockés dans un tableau $\overleftrightarrow{A}$. La dimension de IA est égale à $N+1$ ou N est le nombre total d'inconnues tandis que les dimensions de IA et de JA sont égales aux nombres de coefficients non nuls. JA contient l'index de la colonne de chaque élément non nul tandis que $IA(i)$ contient la position dans $\overleftrightarrow{A}$ du premier coefficient non nul de la ligne i de A. Regardons cette construction sur un exemple :

$$A = \begin{bmatrix} 1 & 12 & 0 & 0 \\ 0 & 3 & 97 & 0 \\ 0 & 1 & 4 & 0 \\ 0 & 0 & 11 & 2 \end{bmatrix}$$

Est une matrice 4 par 4 avec 8 éléments non nuls ainsi :

$$\begin{aligned} \overleftrightarrow{A} &= [\ 1\ 12\ 3\ 97\ 1\ 4\ 11\ 2] \\ IA &= [\ 1\ 3\ 5\ 7\ 9] \\ JA &= [\ 1\ 2\ 2\ 3\ 2\ 3\ 3\ 4]. \end{aligned}$$

Le produit matrice de la matrice creuse A avec le vecteur Y peut alors s'écrire :

$$\begin{array}{l} Y = 0 \\ \text{for } i = 1 \text{ to } n \text{ do} \\ \text{for } j = IA(i) \text{ to } IA(i+1) - 1 \text{ do} \\ Y(i) = Y(i) + \overleftrightarrow{A}(j) * \mathbf{X}(JA(j)) \\ END\ DO \\ END\ DO \end{array} \tag{42.IV}$$

Par rapport au stockage de bandes on ne stocke ici que les coefficients non nuls des matrices creuses. Nous avons utilisé le format YSM dans notre code ce qui a permis un gain d'espace mémoire pendant l'assemblage des matrices. Les algorithmes directs classiques du type "descente remonté LU"[102] ont été optimisés grâce à l'usage du solveur PARDISO[37] qui fait partie de la bibliothèque MKL qui n'est pas open source. Pour la résolution de systèmes linéaires creux, il est aussi possible d'utiliser des méthodes itératives comme la méthode GMRES[104] si la matrice est quelconque où la méthode du gradient conjugué[102] si la matrice est hermitienne définie positive. Les méthodes itératives sont faciles à mettre en œuvre puisqu'elles utilisent essentiellement des opérations simples comme des produits matrices-vecteurs. Nous avons proposé une version qui utilise les méthodes itératives, mais les temps de calcul deviennent trop grands pour une précision similaire aux résultats obtenus avec les méthodes directes utilisées par PARDISO[37].

IV.6.4 En résumé

Nous avons utilisé deux types de stockage :

1. Le format matrices de bandes avec l'algorithme de Cuthill et Mc Kee.
2. Le format Yale Sparse Matrix.

Nous avons utilisé trois différents types algorithmes de résolution de système linéaire :

1. L'algorithme de "descente remonté LU"[102] pour les matrices de bandes.
2. L'algorithme de "descente remonté LU"[102] pour les matrices creuses (cf section IV.6.5).
3. Les algorithmes itératifs GMRES[104] et le gradient conjugué [98] pour les matrices creuses.

Notre choix définitif pour traiter nos problèmes s'est porté sur l'algorithme de "descente remonté LU"[102] pour les matrices creuses développé par le solveur PARDISO[37]. Bien que PARDISO ne soit pas libre et open source il permet des temps de résolution bien inférieurs aux deux autres algorithmes énoncés. Une description plus précise de cet algorithme est donnée dans la section suivante IV.6.5.

IV.6.5 Résolution des système linéaires pour les matrices creuses

IV.6.5.1 Introduction

Nous avons vu section IV.6 que les matrices construites sont creuses. D'autre part l'algorithme de calcul des modes propres (21.IV) nécessite de résoudre le système linéaire $(A - \sigma M)\mathbf{w} = M\mathbf{v}_j$ à chaque itération. La méthode GMRES (Generalized Minimal Residual Method[104]) est une méthode itérative qui permet de résoudre un tel système en utilisant la construction des vecteurs de base d'Arnoldi (cf section IV.2.2). Cette méthode est facile à coder parce qu'elle est construite à partir d'opérations simples tel que des produits matrice-vecteur. Mais nous avons choisi, dans un souci de recherche de performance, d'utiliser le solveur direct PARDISO[37, 92] issu de la librairie MKL (Math Kernel Library). Les algorithmes de factorisation LU utilisés par PARDISO sont des versions optimisées de l'algorithme général que nous allons présenter dans cette section.

Remarque IV.15. Habituellement, pour des matrices pleines le nombre d'opérations nécessaires à la diagonalisation de la matrice $(A - \sigma M)$ est proportionnel au nombre d'opérations nécessaires pour la factorisation LU de la matrice $(A - \sigma M)$ qui est proportionnel à n^3. Les opérations simples de calcul sur les lignes et les colonnes d'une matrice creuse requièrent une exigence de stockage informatique proportionnel au nombre de coefficients non nuls $nnz(A - \sigma M)$[105]. En se référant à l'algorithme classique "Minimum Degree Ordering Algorithm(MDOA)"[99] (cf section suivante IV.6.5.2), on peut déduire que le nombre d'opérations nécessaires à la factorisation des matrices creuses est proportionnel à $n \times nnz(L+U)$ où L et U désignent les matrices issues de la factorisation LU de la matrice $(A - \sigma M)$.

IV.6.5.2 Algorithme de factorisation LU pour les matrices creuses

L'algorithme direct de factorisation LU suivi d'une descente remontée [102], pour les matrices creuses, est l'algorithme utilisé lors de la résolution de grands systèmes linéaires creux. Nous traitons des matrices sous formats YSM (cf section IV.6.3). Lors de la résolution de grands systèmes, il est commun de faire précéder la factorisation numérique par une réorganisation des inconnues comme nous l'avons vu section IV.6.2. Les systèmes linéaires que nous résolvons se composent de matrices réelles symétriques non définies positives ou bien de matrices complexes non symétriques (cf remarqueIII.6). Le solveur PARDISO[37, 92] utilise des algorithmes optimisés pour ce type de matrice. La méthode du pivot de Gauss nécessaire pour la factorisation LU[102] génère des éléments non nuls dans les matrices construites à chaque étape. Prenons l'exemple suivant :

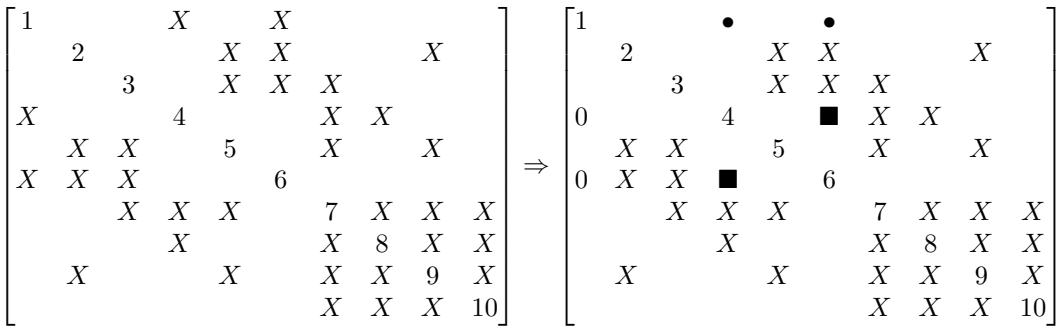


Fig 11.IV Matrice creuse après première élimination de Gauss.

Fig 11.IV X sont les éléments initiaux non nuls, $\bullet$ sont les facteurs qui resteront non nuls et $\blacksquare$ sont les éléments non nuls créés après avoir éliminé les coefficients de la colonne 1.

Les algorithmes utilisés par PARDISO proposent de re-numéroter les inconnues afin de créer le moins possible de 0 lors de la factorisation LU. Nous allons maintenant donner un exemple de renumérotation des inconnues à travers un algorithme classique de réduction d'éléments nuls appelé MDOA (Minimum Degree Ordering Algorithm[99]).

IV.6.5.3 Minimum Degree Ordering Algorithm(MDOA) [99]

Soit $\mathcal{A}$ une matrice carrée inversible de dimension N. On appelle degrés de l'inconnue i le cardinal de l'ensemble $\{j \in [1, N] \ / \ (\mathcal{A})_{ij} \neq 0, i \neq j\}$. À chaque étape K de l'algorithme MDOA, on permute l'inconnue de la ligne K avec l'inconnue de degrés minimum des lignes $[K \dots N]$. Nous allons illustrer les différentes étapes en prenant un exemple avec $N = 8$ et la matrice $\mathcal{A}$ définie par :

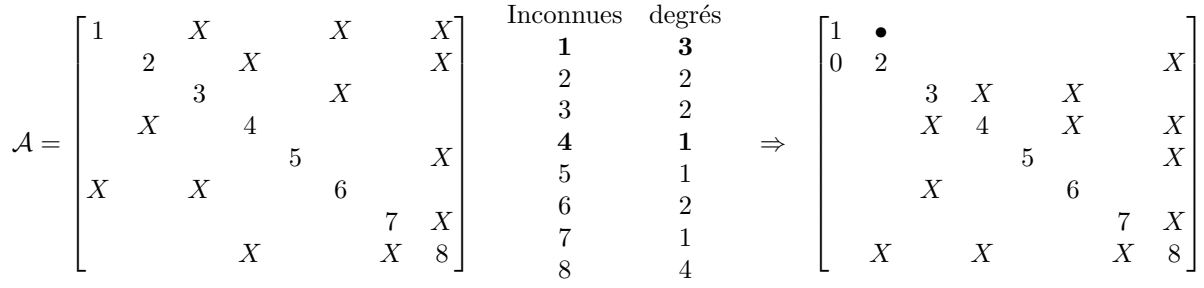


Fig 12.IV Étape 1 permutation de la ligne 1 avec la ligne 4 et élimination des coefficients de la colonne 1. On garde les conventions adoptées Fig 11.IV.

Fig 12.IV, on échange la ligne 1 avec la ligne 4 et l'itération de GAUSS ne génère pas d'élément non nul.

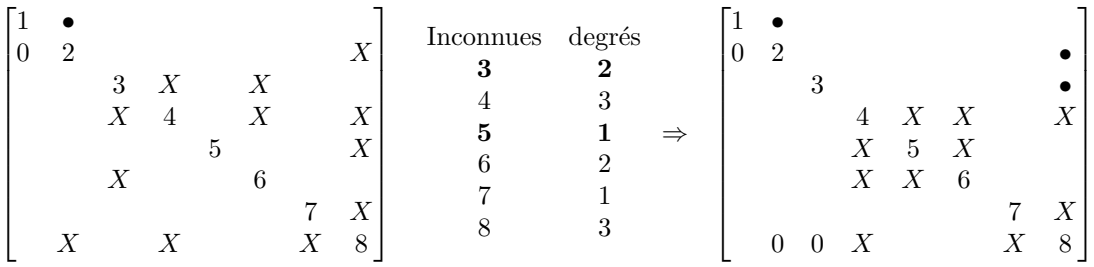


Fig 13.IV Étape 2 permutation de la ligne 5 avec la ligne 3 et élimination des coefficients des colonnes 2 et 3.

Fig 13.IV on échange la pivot 5 avec la ligne 3. Les itérations de GAUSS ne génèrent pas d'élément non nul.

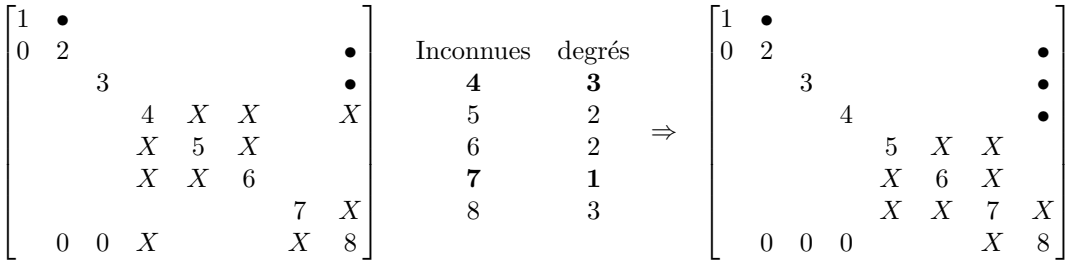


Fig 14.IV Étape 3 permutation de la ligne 4 avec la ligne 7 et élimination des coefficients de la colonne 4.

Fig 14.IV on échange la pivot 4 avec la ligne 7. L'itération de GAUSS ne génère pas d'élément non nul.

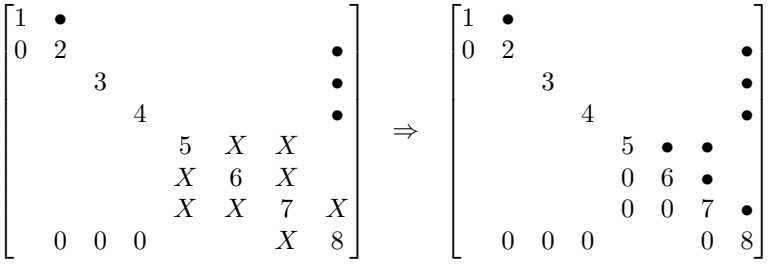


Fig 15.IV Étape 5 élimination des coefficients de la colonne 5,6,7 et 8.

Fig 15.IV Toutes les itérations de GAUSS restantes ne génèrent pas d'élément non nul, pas besoin de permuter d'inconnue.

L'algorithme MDOA n'est pas toujours celui qui crée le moins de zéro. En effet, considè-

rons la matrice $\mathcal{B}$:

$$\mathcal{B} = \begin{bmatrix} 1 & X & X & X & & & & & \\ X & 2 & X & X & & & & & \\ X & X & 3 & X & & & & & \\ X & X & X & 4 & X & & & & \\ & & & X & 5 & X & & & \\ & & & & X & 6 & X & X & X \\ & & & & & X & 7 & X & X \\ & & & & & X & X & 8 & X \\ & & & & & X & X & X & 9 \end{bmatrix} \underset{\text{MDOA}}{\longrightarrow} \tilde{\mathcal{B}} = \begin{bmatrix} 1 & & & \bullet & & \bullet & & & \\ & 2 & X & X & X & & & & \\ & X & 3 & X & X & & & & \\ 0 & X & X & 4 & X & \blacksquare & & & \\ & X & X & X & 5 & & & & \\ 0 & & & \blacksquare & & 6 & X & X & X \\ & & & & & X & 7 & X & X \\ & & & & & X & X & 8 & X \\ & & & & & X & X & X & 9 \end{bmatrix}.$$

La factorisation LU de $\mathcal{B}$ n'engendre pas de nouvel élément non nul alors que la matrice $\tilde{\mathcal{B}}$ construite après les permutations de l'algorithme MDOA si.

Remarque IV.16. Comme l'algorithme de Cuthill et McKee évoqué section IV.6.2, il existe plusieurs algorithmes de renumérotation à utiliser et à combiner suivant l'organisation des inconnues conditionnée par la structure du maillage. C'est le rôle du solveur PARDISO [92] qui peut traiter toutes les matrices complexes creuses non symétriques, mais on observe quand même des temps de calculs différents en fonction de la structure du maillage.

IV.7 Conclusion

Dans cette section nous avons présenté tous les détails des algorithmes que nous avons mis en œuvre par l'intermédiaire des librairies ARPACK et INTEL PARDISO pour résoudre le problème de modes propres (1.IV). Nous avons dans un premier temps, section IV.1, défini le problème numérique initial des modes propres. Ensuite dans la section section IV.5, nous avons expliqué qu'il existait des valeurs propres nulles que nous devons éliminer en transformant le problème initial. Une fois le problème initial transformé (cf section IV.3), on obtient un problème de valeurs propres que nous résolvons avec le solveur ARPACK [33] qui utilise des algorithmes détaillés dans la section IV.2. Les paramètres de résolution sont repris et détaillés dans la section IV.4 qui résume et analyse la méthode numérique dans son ensemble et présente l'influence du maillage sur la validité des résultats (sous-section IV.4.6). Dans cette section, nous insistons notamment sur le choix pragmatique des paramètres qui permet d'obtenir des résultats valables comme nous allons le voir dans les sections V.3 et V.4. Le problème de valeurs propres est composé de matrices creuses que l'on stocke sous format bande section IV.6.1 ou format YSM, IV.6.3, afin d'optimiser l'espace mémoire et le temps de calcul. Nous observons dans la section IV.6.5 que le profil des matrices creuses a une influence le temps de calcul qui est optimisé par l'utilisation du solveur PARDISO[37].

Dans le chapitre V suivant, on s'intéresse à l'utilisation de la méthode que nous avons développée pour des cas de validation et pour la caractérisation des modes de surface.

V Validations et extensions des outils numériques utilisés pour le calcul de diagrammes de bandes.

V.1 Introduction

Dans cette partie nous allons simuler plusieurs cas tests qui illustrent les possibilités de l'outil que nous avons développé. Dans un premier temps nous décrivons la structure des différents fichiers de données qui interviennent lors de l'exécution d'un programme section V.2, dans un second temps, dans la section validation V.3, nous comparons les résultats obtenus avec des cas de référence issus de la littérature ou d'autres méthodes. Ensuite nous abordons les calculs permettant d'identifier et de représenter les modes de surface section V.4, grâce au calcul du diagramme de bandes.

V.2 Structure des données

V.2.1 Introduction

Avant d'énumérer les différents cas tests, nous allons montrer comment sont construits les fichiers de données afin d'avoir une représentation générale des paramètres manipulés et de faciliter la compréhension d'un utilisateur du code. Par convention nous nommons les fichiers exécutables avec l'extension *.x*, les fichiers d'entrée avec l'extension *.don*, les fichiers de sortie avec l'extension *.dat* et la cartographie des champs avec une extension *.plt*. Les exécutables se compilent avec les librairies PARDISO[37] et ARPACK[33] une explication de la compilation est donnée en annexe A.8. Les fichiers de sorties sont des fichiers *.dat* qui contiennent les caractéristiques des modes propres en fréquence ou en rad/m dans l'ordre croissant pour chaque valeur de la phase (correspondant à la norme et à la direction du vecteur d'onde incident). Ces fichiers se lisent avec le logiciel libre Gnuplot.

V.2.2 Les fichiers d'entrée

Les fichiers d'entrée regroupent toutes les données du solveur modes propres vues section IV.4, les données sur les conditions limites (cf section III.3) imposées sur les différentes faces du maillage, les propriétés volumiques des matériaux (permittivité perméabilité), les noms des fichiers de maillage et des options relatives au diagramme de bandes. Les fichiers de maillage devront aussi être présents dans le dossier où est lancée l'exécution. Les fichiers de maillage ont l'extension *.bma* qui est reconnue par notre solveur. Cette extension est obtenue à partir du post-traitement du maillage sous format *.unv* avec le logiciel maison PPTM[1]. En connaissant la structure d'un fichier *.bma* il est cependant possible de générer des fichiers de maillage lisible par notre solveur sans utiliser PPTM. Lors du maillage de la structure on devra numéroter les zones volumiques et surfaciques afin de les identifier pour faire varier les propriétés du milieu ou de pouvoir appliquer des conditions limites sur les interfaces. Nous

donnons ici une explication des principaux paramètres de résolution à renseigner dans les fichiers *.don* :

Paramètres	**Signification**
DV	Nombre de directions du vecteur incident.
$\bar{k}$	Nombre de valeurs propres.
TB	Type du diagramme de bandes (directions d'incidence).
ND	Nombre de point par direction d'incidence (point tous les π/ND).
$TPML$	Type de PML : 1 PMLZ, 2 PMLX, 3 PMLY
K_σ	Facteur multiplicatif du *shift* σ (22.IV).
F_{im}	Fréquence minimum d'absorption de la PML en Gigahertz.
$\bar{tol}$	Critère de convergence $\|\|AX - MX\|\|/\|\lambda\| \leq \bar{tol}$
d_c	Distance caractéristique d'un élément de la structure.

Fig 1.V Paramètres principaux du solveur.

- TB permet de choisir parmi plusieurs chemins du vecteur d'onde incident. Soit les points du repère cartésien $\mathbf{G}(0,0,0)$; $\mathbf{X}(1,0,0)$;$\mathbf{M}(1,1,0)$; $\mathbf{P}(1,1,1)$; $\mathbf{Y}(0,1,0)$. Si $TB = 1$ le chemin est **GXMP**, 2→ **GXMG**. Des exemples seront détaillés dans les sections V.3.3 et V.4.2. Si TB=3 alors le vecteur incident varie dans la direction **GY** pour plusieurs valeurs de phases dans la direction **GX**. Autrement dit, c'est le diagramme de bandes projeté (cf section V.4.2) dans la direction **GY**.
- ND représente le pas du changement de phase $\phi = \mathbf{k}_i.\mathbf{T}$ où $\mathbf{k}_i$ est le vecteur d'onde incident et $\mathbf{T}$ le vecteur de translation entre les différentes faces *Master-Slave*. La variation de la phase ϕ correspond au "découpage" des abscisses du diagramme de bandes. Par exemple si $ND = 10$, on calcule les solutions tous les $\pi/10$ jusqu'à π ensuite on change la direction du vecteur incident DV fois selon les directions données par le paramètre TB.
- $TPML$ permet d'indiquer dans quelle direction l'on souhaite que la PML soit absor-

bante. Si il n'y a pas de PML alors la valeur renseignée n'est pas utilisée. Le fichier *.don* contient les codes des zones qui correspondent aux conditions limites que l'on souhaite imposer.

- K_σ le facteur multiplicatif du *shift* σ défini par la relation (22.IV).
- F_{im} est la fréquence minimum d'absorption définie à la section III.3.3 et par la relation (17.III).
- $\bar{tol}$ le facteur permettant à la fois de contrôler la convergence des modes propres (13.IV) et d'éliminer les solutions parasites (24.IV).
- d_c est la distance caractéristique d'un élément de la structure permettant de calculer le *shift* σ défini par la relation (23.IV). Elle est exprimée en mètre.

Remarque V.1. Soit λ_{im} la longueur d'onde correspondant à la fréquence F_{im}. Pour une absorption suffisante, la hauteur de la PML doit avoir une longueur minimum égale à $\lambda_{im}/4$ (cf relation (18.III)). La direction d'absorption réglée par le paramètre *TPML*, doit être confondue avec l'axe de la hauteur de la PML, telle que définie en (18.III).

Remarque V.2. Pour configurer le *shift* σ l'utilisateur a le choix entre les deux paramètres d_c et K_σ. Après multiplication par le facteur K_σ le nouveau *shift* σ_{new} sera lié avec d_c par la relation (23.IV) :

$$\sigma_{new} = K_\sigma \sigma = (\frac{2\pi}{d_c})^2$$

Si on choisit d'utiliser le paramètre K_σ, le *shift* est déjà calculé et il suffit de laisser $K_\sigma = 1$ dans la plupart des cas simples qui contiennent un seul élément à périodiser comme dans la section V.3.3. Si on choisit d_c il faut connaître au préalable les distances caractéristiques des éléments périodisés. Dans les deux cas il arrive que l'on fasse plusieurs réajustement avant de trouver une valeur qui limite le nombre d'itérations dans l'algorithme de tri (cf section IV.4.5).

V.2.3 Calcul et représentation des champs

Dans chaque tétraèdre le champ s'écrit comme combinaison linéaire des fonctions de base de Nedelec d'ordre 0[70]. Ces fonctions de base sont définies sur chacune des 6 arêtes du tétraèdre. Autrement dit, si on se place dans le tétraèdre T_K le champs local solution

$$\mathbf{E}_{T_k}(\mathbf{r}) = \sum_{p=1}^{6} e_p \mathbf{W}_p(\mathbf{r}).$$

Une fois le problème résolu, les valeurs de e_p sont connues en chaque point du tétraèdre et aussi en chaque point de l'espace. Pour représenter le champ sur un maillage d'un plan qui coupe le domaine périodique infini, il suffit de reporter la valeur du champ en chaque point du maillage surfacique en reproduisant par périodicité le champ de la cellule de base grâce aux relations de déphasage de Floquet (cf section II.1). Pour pouvoir facilement appliquer les relations de déphasage il est préférable que le maillage surfacique soit régulier.

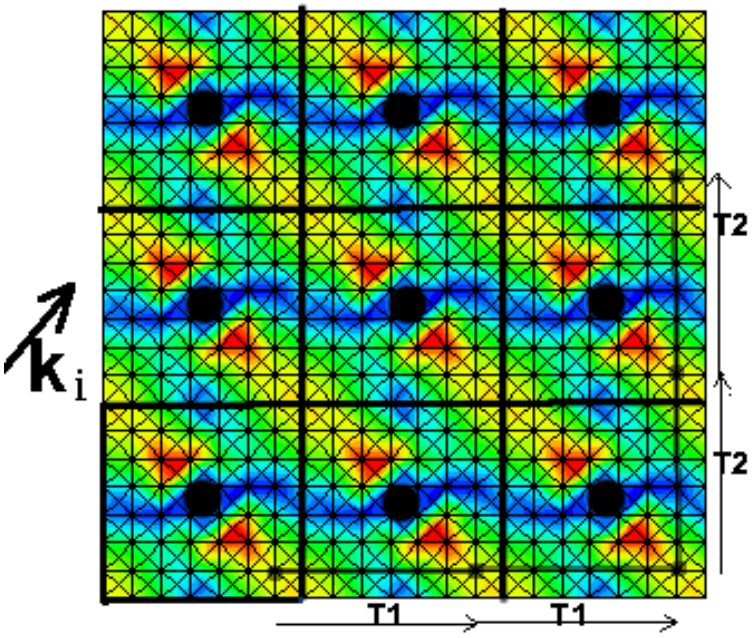


Fig 2.V Exemple du module du champ électrique d'un mode propre pour un réseau de tiges diélectriques (une coupe transverse).

Sur le dessin Fig 2. V, la cellule de base est la cellule encadrée en bas à gauche. Chaque point du maillage correspond à un point de la cellule de base après translations successives. Comme dans la relation (7.II) les champs des points images par translation sont déphasés d'un facteur $e^{-j\left(m\times\mathbf{T}_1+n\times\mathbf{T}_2\right).\mathbf{k}_i}$ où $\mathbf{T}_1$ et $\mathbf{T}_2$ représentent les vecteurs translation selon 2 directions de périodicité et $\mathbf{k}_i$ le vecteur d'onde incident. Sur les points choisis de la figure de droite $n = m = 2$. Les fichiers champs ont l'extension *.plt* et sont lisibles par le logiciel commercial de visualisation graphique Tecplot.

V.2.3.1 Conclusion

Dans cette partie nous avons présenté comment sont organisés les fichiers de données ainsi que les paramètres nécessaires pour la résolution de nos problèmes des modes propres. Dans la sections V.3 et V.4 les paramètres définis section V.2.2 sont utilisés.

V.3 Validation

V.3.1 Introduction

Dans cette section V.3, nous commençons par traiter des problèmes de cavités métalliques V.3.2, pour ensuite aborder des réseaux périodiques de cubes métalliques et de tiges diélectriques V.3.3. Cet ordre correspond à l'ordre dans lequel nous avons effectué les travaux de recherche. En effet avant de nous intéresser aux structures périodiques nous avons cherché à développer et à optimiser un solveur modes propres pour les cavités métalliques.

V.3.2 Cavités métalliques et solutions stationnaires

Dans les sections V.3.2.1 et V.3.2.2 les structures que nous étudions sont toutes fermées par des conditions métalliques PEC (cf section III.3.2). La résolution utilise l'élimination automatique des modes propres parasites (cf section IV.5). Pour des raisons de gains de mémoire et de temps de calcul, le *shift* σ est négatif (cf remarque IV.12 et l'équation (37.IV)). Le paramètre de convergence est $\bar{tol} = 10^{-3}$ et comme il n'y a pas de solution parasite, le nombre de valeurs propres désirées $\bar{k}$ est égal au nombre de valeurs propres convergentes. Autrement dit, on ne trie plus les solutions selon le critère d'élimination (24.IV). Les calculs sont réalisés sur notre machine (PC processeur Intel(64) Xeon avec une horloge de 2.80GHz).

V.3.2.1 Cavité inhomogène

Ce test est issu de l'article [32]. Pour valider nos résultats, nous allons comparer nos résultats les valeurs de k_0 obtenues dans cette publication.

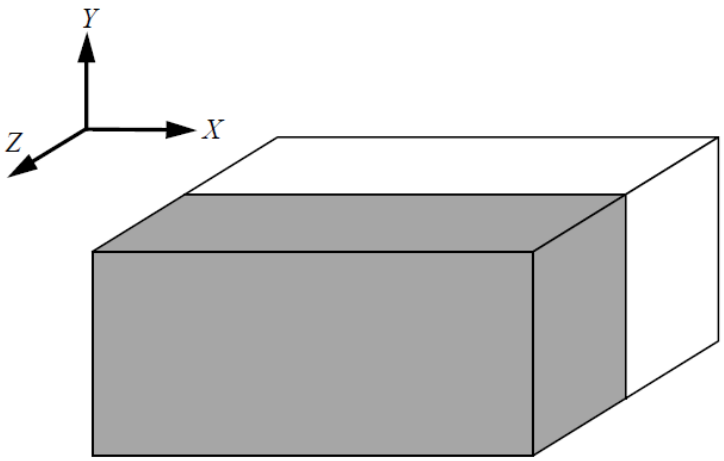


Fig 3.V Structure in-homogène.

Fig 3.V la structure étudiée est une cavité métallique qui est remplie de matériaux $\epsilon_r = 1$ pour $z \in [0, 0.005]$ et $\epsilon_r = 2$ pour $z \in [0.005, 0.01]$. La taille totale de la Cavité :10x1x10 milimètres, le nombre total d'arêtes est 1035, le nombre arêtes de bord PEC sont au nombre de 546. Le *shift* (37.IV) σ est égal à $-0.47 rad.cm^{-1}$ ce qui correspond à la valeur du paramètre $K_\sigma = -0.1$. Le temps de calcul de 10 modes ($\bar{k} = 10$) est de moins de 0.1 seconde et la

mémoire maximum utilisée est de moins de 5Mb. Les solutions sont exprimées en $rad.\text{cm}^{-1}$:

Analytical[106]	Article [32] 615 tétraèdres	Résulats obtenus 636 tétraèdres	Volakis[106] $\simeq$ 620 tétraèdres()*
$(k_0)_{101}$= 3.538	3.524	3.536	3.534
$(k_0)_{201}$=5.445	5.401	5.437	5.440
$(k_0)_{102}$=5.935	5.931	5.962	5.916
$(k_0)_{301}$= 7.503	7.382	7.4199	7.501
$(k_0)_{202}$ = 7.633	7.562	7.565	7.560
$(k_0)_{103}$ =8.096	8.003	8.0959	8.056

Fig 4.V Tableau comparatif de valeurs obtenues après un calcul de modes propres dans une cavité inhomogène.

Dans la Fig 4.V le symbole ()* signifie que le nombre de tétraèdre n'est pas précisé dans l'article [106], alors que le nombre d'inconnue est mentionné seulement dans la partie du volume de permittivité $\epsilon_r = 1$ parce que le calcul de modes propres est effectué en supprimant le milieu $\epsilon_r = 2$ et en introduisant un terme de bord à l'interface. En respectant les proportions de notre mailleur on arrive aux alentours de 620 tétraèdres. L'anisotropie de la structure et le maillage grossier sont responsables des erreurs qui deviennent plus grandes au fur et à mesure que l'on s'éloigne des premières solutions. On a aussi remarqué que la structure du maillage avait un rôle déterminant dans la précision des résultats obtenus. En effet, pour un même nombre de tétraèdre les résultats obtenus seront plus proches des résultats théoriques si le maillage est raffiné à l'interface entre les deux milieux $\epsilon_r = 1$ et $\epsilon_r = 2$. Dans le maillage que nous avons utilisé pour établir les résultats Fig 4.V les arêtes de l'interface entre les deux milieux sont deux fois plus petites que les autres arêtes. L'information du changement de milieu se trouve essentiellement à l'interface c'est pourquoi il est nécessaire de mailler plus finement à cet endroit. Observons maintenant l'allure des champs des deux premiers modes sur la Fig 5.V et la Fig 6.V. Elle est conforme à ce que l'on physiquement attendre pour ce type de structure.

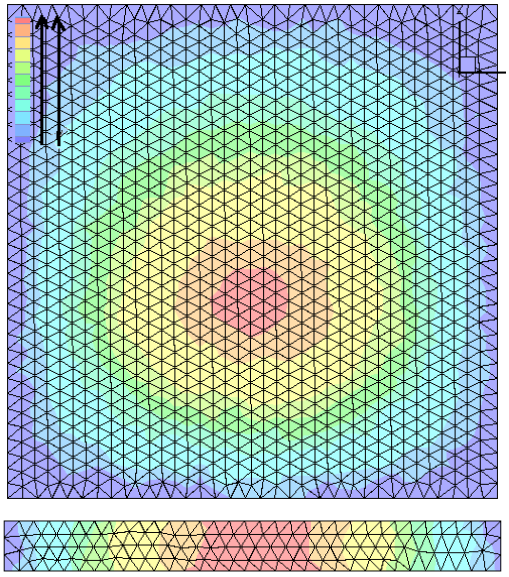


Fig 5.V Premier mode dans les plans Y=0.5mm et Z=5mm.

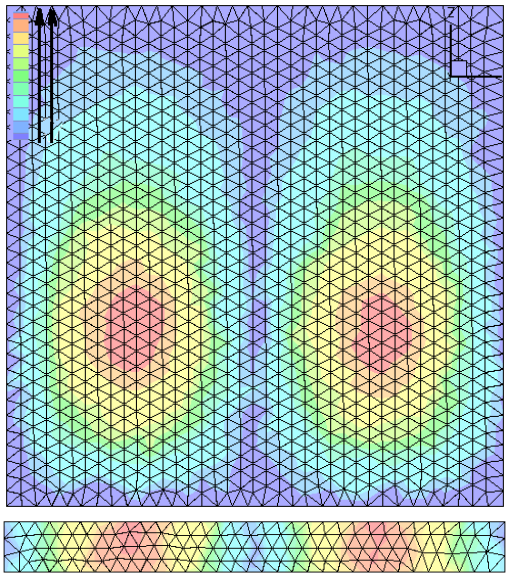


Fig 6.V Deuxième mode dans les plans Y=0.5mm et Z=5mm.

Grâce à la géométrie simple de la cavité étudiée, ce test a permis d'affirmer que notre outil fonctionne en comparant les résultats obtenus avec des résultats analytiques. Notre outil permet aussi de traiter des problèmes de cavités composées de géométries plus complexes. La possibilité de traiter des géométries complexes dépend uniquement des capacités du mailleur.

V.3.2.2 Cavité incluse dans une autre

Cet exemple est intéressant puisqu'il permet de visualiser un mode nul à divergence nulle autrement dit un mode du Groupe.2 (26.IV). Le nombre de modes du Groupe.2 est égal au nombre total de cavités moins un [34]. En fait, ces modes sont des champs statiques qui naissent de la différence de potentiel entre les parties métalliques indépendantes de la structure. Voici l'exemple d'une cavité de 0.5 m de coté incluse dans une cavité de 1 m de coté. Entre les deux cavités il y a un trou d'air :

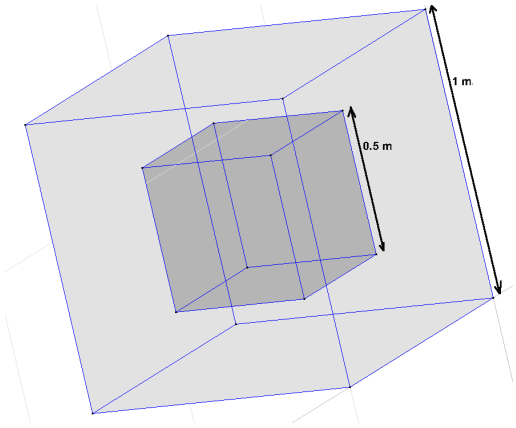


Fig 7.V Dimension de la structure métallique. nombre d'arêtes totales= 11024, arêtes interface PEC= 2316.

.

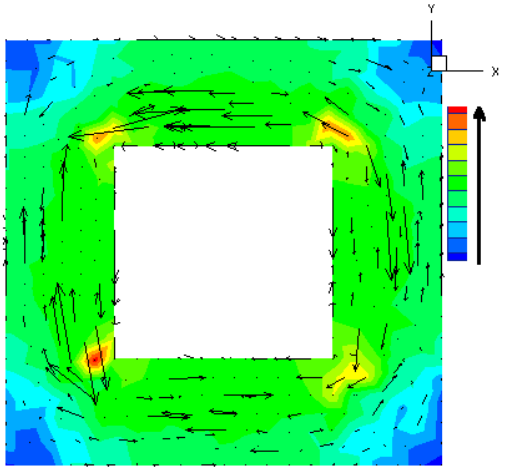


Fig 8.V Module du champ électrique $\tilde{\mathbf{E}}$ du mode nul du Groupe.2 (26.IV) et vecteur de Poynting dans le plan Z=0.5m. L'échelle des valeurs du champ n'est pas précisée, nous nous intéressons aux variations locales.

Dans la Fig 8.V on peut observer un mode statique avec un champ plus fort sur les coins du cube. Le temps de calcul constaté pour $\bar{k} = 15$ modes est de 3 secondes et la mémoire utilisée est 73 Mb. Le *shift* (37.IV) σ est égal à $-1.6 rad/m$ ce qui correspond aux valeurs $K_\sigma = -0.1$ et $d_c = 3.92$ m.

V.3.3 Réseaux périodiques

Dans les sections V.3.3.1 et V.3.3.2 les calculs sont effectués sur notre machine (PC processeur intel avec une horloge de 2.80GHz). Le résidu admissible $\bar{tol}$ pour chaque valeur propre équation (24.IV) est de 10^{-7}. À chaque itération $\bar{k} = 17$ modes propres sont calculés. La valeur du coefficient multiplicatif du *shift* K_σ est égale à 1.

V.3.3.1 Cubes Métalliques 3 D

On considère des cubes parfaitement conducteurs en espace libre répartis de façon périodique dans les trois directions. La cellule de dimension initiale est le cube de hauteur Dx=Dy=Dz=1m, la longueur du côté du cube est w= 0.5m. Les caractéristiques ϵ, μ sont celles de l'air soit $\epsilon_1 = \mu_1 = 1$, ce cas est donc la périodisation selon les trois directions propres de la cavité précédente :

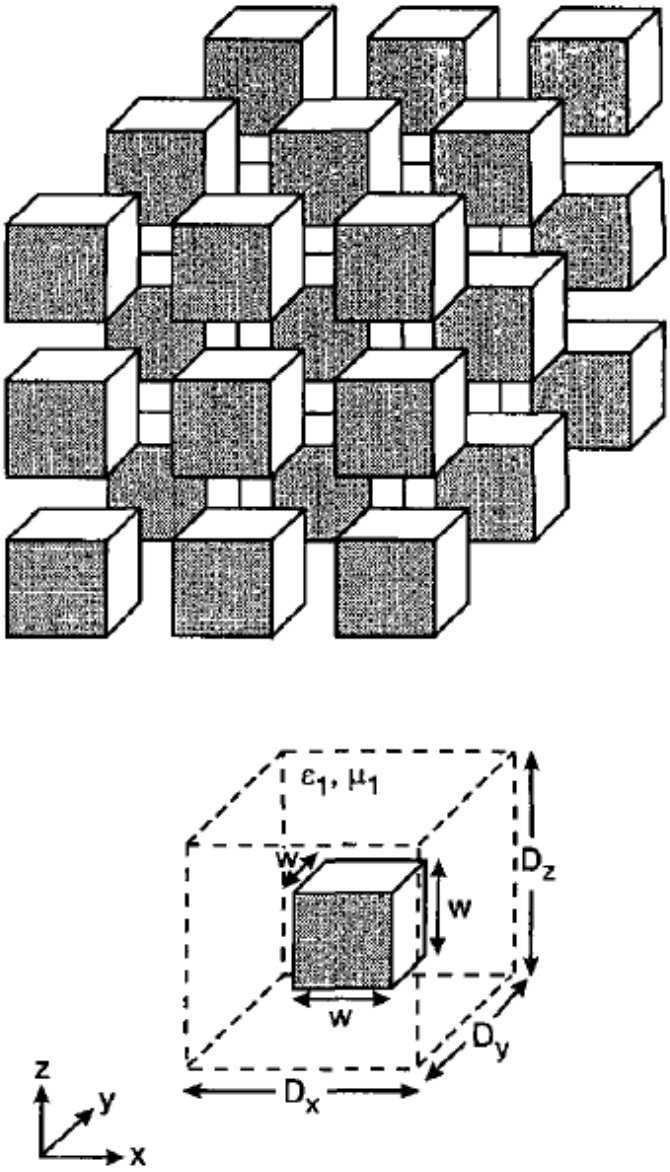


Fig 9.V Cube métalliques dans l'air [35].

Pour nos simulations on applique des conditions *Maître-esclave* (cf section III.3.5) dans les 3 directions de l'espace. D'abord, on fait varier le déphasage dans la direction Ox de 0 à π et les déphasages dans les autres directions sont nuls. Ensuite, on fixe le déphasage dans la direction Ox à π, dans la direction Oz à 0 et on fait varier le déphasage de la direction Oy de 0 à π (intervalle $[\pi, 2\pi]$ sur le schéma). Ensuite les déphasages Ox et Oy sont fixés à π et on fait varier le déphasage Oz de 0 à π(intervalle $[2\pi, 3\pi]$ sur le schéma). Autrement dit $DV = 3$ et $TB = 1$. Les résultats sont exprimés en radian par mètre et le déphasage est en radian :

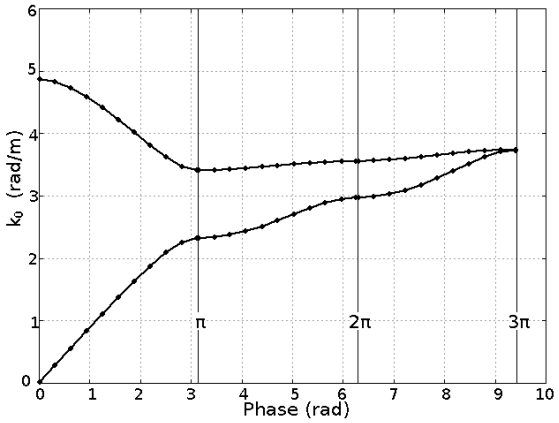


Fig 10.V Résultats issus de notre résolution. Nombre d'inconnues 9599. Le temps de calcul des 33 points($ND = 11,\ DV = 3$ et $TB = 1$) est de 4 minutes et 33 secondes. Mémoire maximum utilisée 152.7 Mb. Le *shift* σ (22.IV) est égal à 5 rad/m .

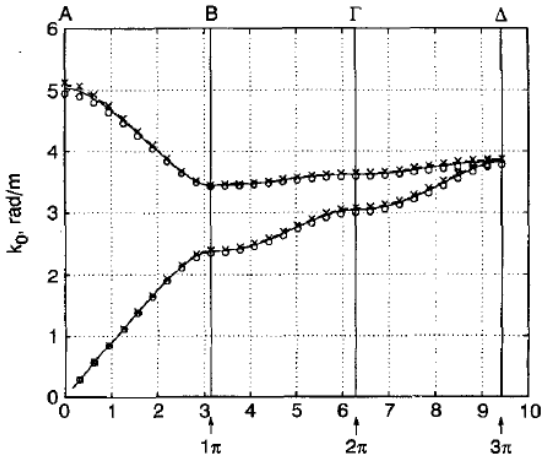


Fig 11.V Résultats tirés de [35]. Nombre d'inconnues (arêtes totales-arêtes PEC)= 10344. Le temps de calcul n'est pas précisé.

Les résultats obtenus par notre solveur Fig 10.V sont concordants avec ceux présentés dans [35]. On pourrait très bien remplacer les cubes métalliques par des sphères, des pyramides ou n'importe quelle structure 3 D que nous pouvons mailler.

V.3.3.2 Réseau de tiges diélectriques

On considère un réseau carré de tiges diélectriques infinies réparties[26] de façon périodique dans l'air selon l'axe Oz. L'écart a entre deux tiges est de $7mm$ et le diamètre des tiges est de $4mm$. Les caractéristiques ϵ_r, μ_r du diélectrique sont $\epsilon_r = 9.4,\ \mu_r = 1$:

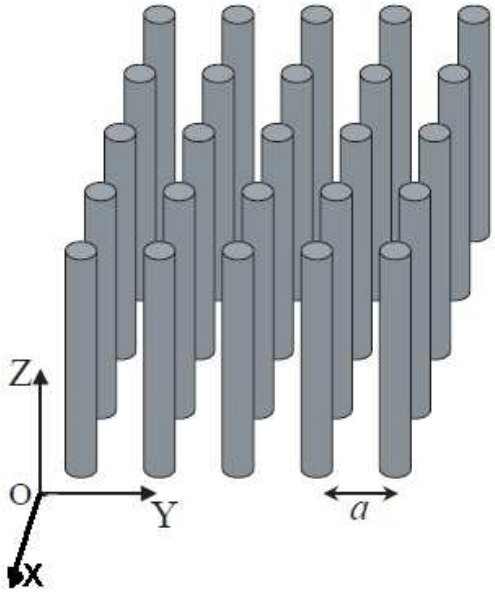


Fig 12.V Réseau de tiges diélectriques, $a = 7mm$.

Dans nos simulations nous appliquons des conditions "*Maître-esclave*" dans toutes les directions de l'espace avec un déphasage variable selon les 2 axes Ox et Oy. Le déphasage dans la direction Oz est fixé à 0.

Remarque V.3. Comme cela est illustré Fig 13.V, l'épaisseur de la cellule de base dans la direction OZ peut être fine par rapport aux autres dimensions de la cellule puisque le déphasage dans cette direction est fixé à 0. En effet la géométrie du motif ne varie pas dans cette direction. On pourrait aussi directement faire ce calcul avec des éléments finis en 2 dimensions [35].

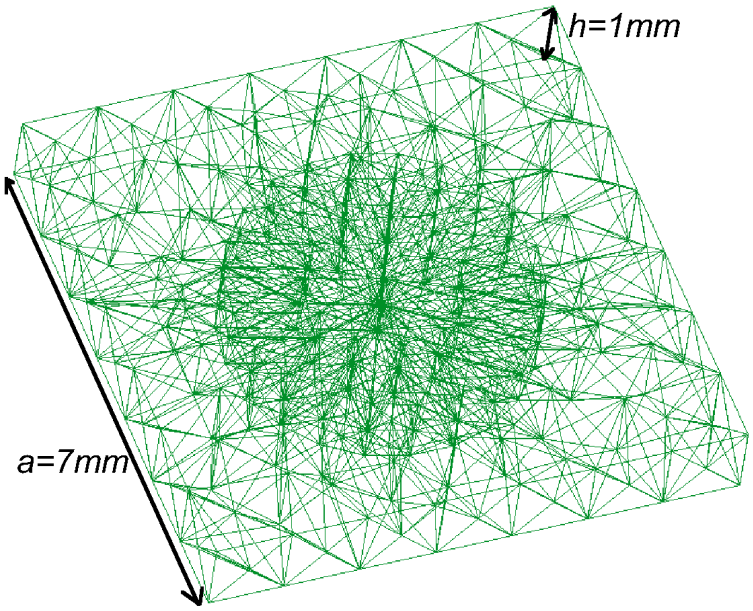


Fig 13.V Exemple de maillage de la cellule de base du réseau de tiges.
Le longueur de la cellule de base $a = 7mm$ et la hauteur $h = 1mm$.

Fig 14.V et Fig 15.V on fait varier le déphasage dans la direction Ox de 0 à π et les déphasages dans les autres directions sont nuls. Ensuite, on fixe le déphasage dans la direction Ox à π, dans la direction Oz à 0 et on fait varier le déphasage de la direction Oy de 0 à π (intervalle $[X, M]$ sur le schéma). Ensuite on fait varier les déphasages dans les directions Ox et Oy de π à 0 et on laisse le déphasage dans la direction Oz à 0 (intervalle $[M, \Gamma]$ sur le schéma). Autrement dit $DV = 3$ et $TB = 2$. Les résultats sont exprimés en unité normalisée a/λ :

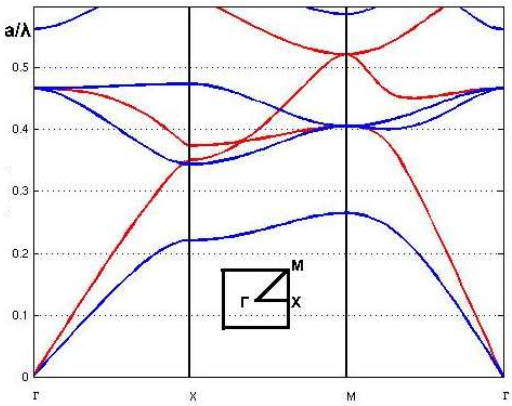


Fig 14.V Résultats issus de la méthode des ondes planes[49]. Les modes TE sont en rouge et les modes TM en bleu.

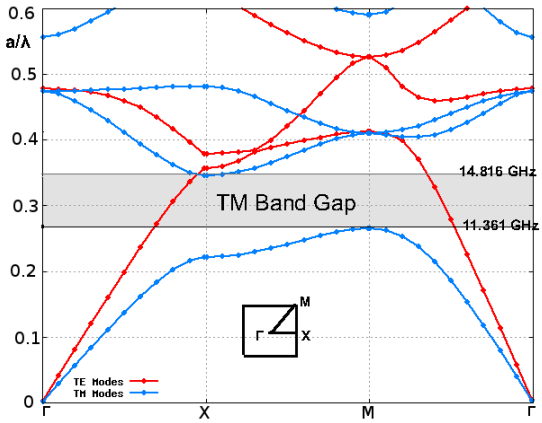


Fig 15.V Résultats issus de notre résolution.

Nombre d'inconnues= 7035, temps de calcul des 33 points ($ND = 11$, $DV = 3$ et $TB = 2$) est de 46 secondes. Mémoire maximum utilisée 49 Mb. Le *shift* σ (22.IV) est égal à 20.89 GHz=0.488 a/λ.

Fig 15.V et Fig 14.V les résultats sont identiques. Fig 15.V nous avons ajouté des informations sur la bande interdite des modes TM qui existe pour des fréquences variants entre 11.361 GHz et 14.816 GHz . Ces résultats ont été validés expérimentalement dans notre laboratoire[107] et le diagramme de transmission Fig 16.V confirme bien l'existence d'une bande interdite autour de 13 GHz. Grâce au dispositif Fig 17.V, deux mesures ont été réalisées à l'aide de plusieurs cornets adaptés : la première en l'absence de tiges (le support étant déjà en place), et la seconde en présence du réseau. Le résultat, présenté sous forme de différence, permet de minimiser les erreurs liées aux instruments de mesure et à la présence du support.

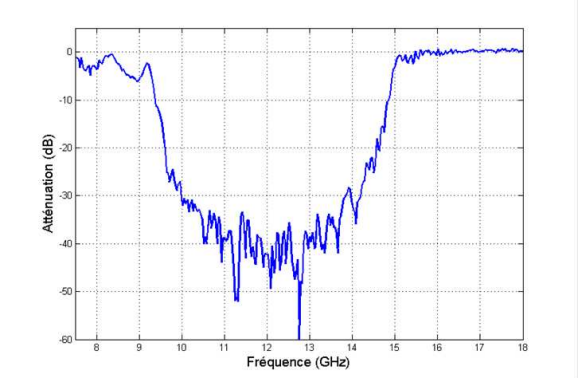


Fig 16.V Diagramme de transmission du réseau de tiges représenté Fig 17.V [107].

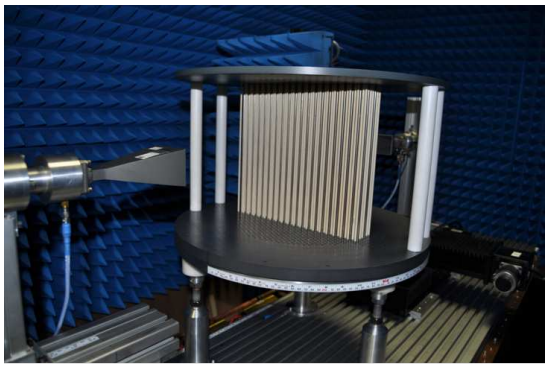


Fig 17.V Antenne cornet et et réseau fini de tiges sur le banc de mesure dans la chambre anéchoïque de l'ONERA [107].

Sur la Fig 17.V les tiges ont les mêmes caractéristiques (diamètre, permittivité, pas de réseau) que celles du réseau infini étudié dans cette section.

V.3.4 Conclusion

Dans cette partie nous montrons que les résultats obtenus avec notre outil sont fiables par comparaisons successives avec d'autres méthodes. Nous reprenons aussi certains paramètres de résolution, expliqués dans la section V.2, indispensables à la compréhension du déroulement d'un calcul. Dans la section V.3.3.2 une structure périodique dans 2 directions est abordée avec la méthode des ondes planes qui est très efficace dans ce cas puisque la géométrie est canonique et les modes se calculent sur une surface orthogonale aux tiges. Dans ce cas, notre résolution fonctionne mais nécessite plus de ressources que la méthode des ondes planes. Dans la section V.3.3.1 une géométrie 3 D canonique comportant des matériaux PEC est abordée. On montre que nos résultats sont concordants avec une autre mise en œuvre de la méthode des éléments finis et qu'il est possible de traiter des géométries non canoniques. Nous allons voir maintenant comment l'outil permet de résoudre des problèmes liés aux modes de surface.

V.4 Modes de surface et couplage inter-éléments dans un réseau

V.4.1 Introduction

Nous allons, dans la première partie V.4.2, décrire et tracer le diagramme de bandes de structures périodiques appelées super cellules qui sont construites à partir de plusieurs motifs inclus dans la cellule de base. La notion de super cellule a été introduite dans [4] et traitée avec la méthode des ondes planes dans [26]. Les super cellules sont utilisées pour observer des phénomènes de guidage d'énergie où de modes de surface. Dans la section V.4.2 nous proposons d'aborder les problématiques des super cellules avec la méthode des éléments finis en recalculant et en visualisant des modes particuliers déjà observés avec la méthode des ondes planes pour le réseau de tiges diélectriques. Ensuite certains résultats d'analyse des tiges diélectriques sont extrapolés aux surfaces hautes impédances section V.4.4 et à un réseau d'antennes patchs insérées dans des cavités métalliques (section V.4.5).

V.4.2 Super-cellules et diagrammes de bandes projetés

On appelle super cellule une cellule qui contient plusieurs motifs de base. Dans le cas classique [4, 26] les motifs de base sont les tiges infinies :

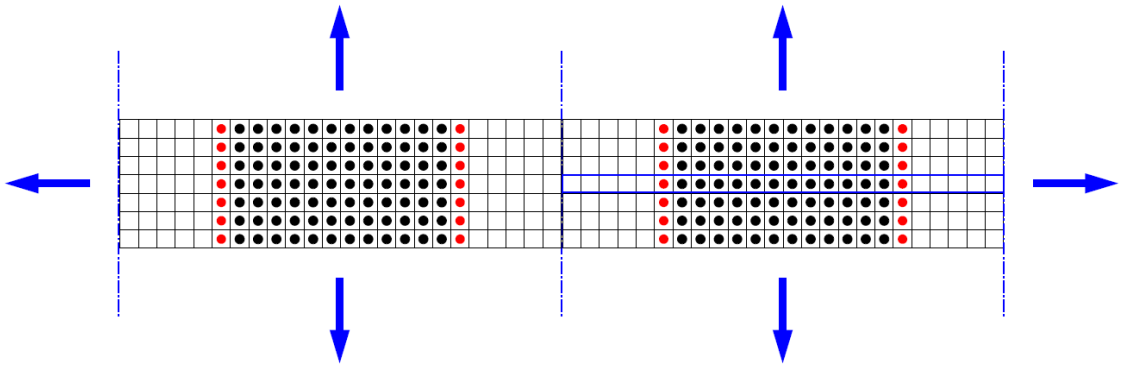


Fig 18.V Coupe transversale d'une superstructure obtenue par répétition de la super-cellule ([26] page 202), la super-cellule est encadrée en bleu.

Par exemple on peut construire une super-cellule avec un défaut linéaire qui est un trou et qui agit comme un guide à certaines fréquences [26] :

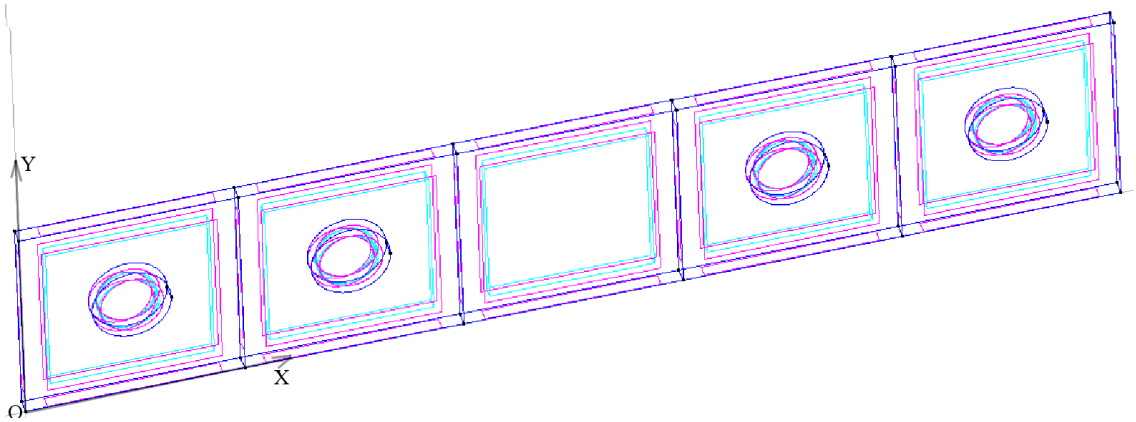


Fig 19.V Super cellule avec un défaut linéaire. La distance entre chaque élément est $a = 14mm$ et la longueur du trou d'air au milieu est égale à a. Les rayons des tiges sont tous égaux à $0.018a = 0.252mm$. La permittivité relative des tige est $\epsilon_r = 11.56$.

On peut aussi utiliser un défaut de surface pour fixer des ondes de surface [26] :

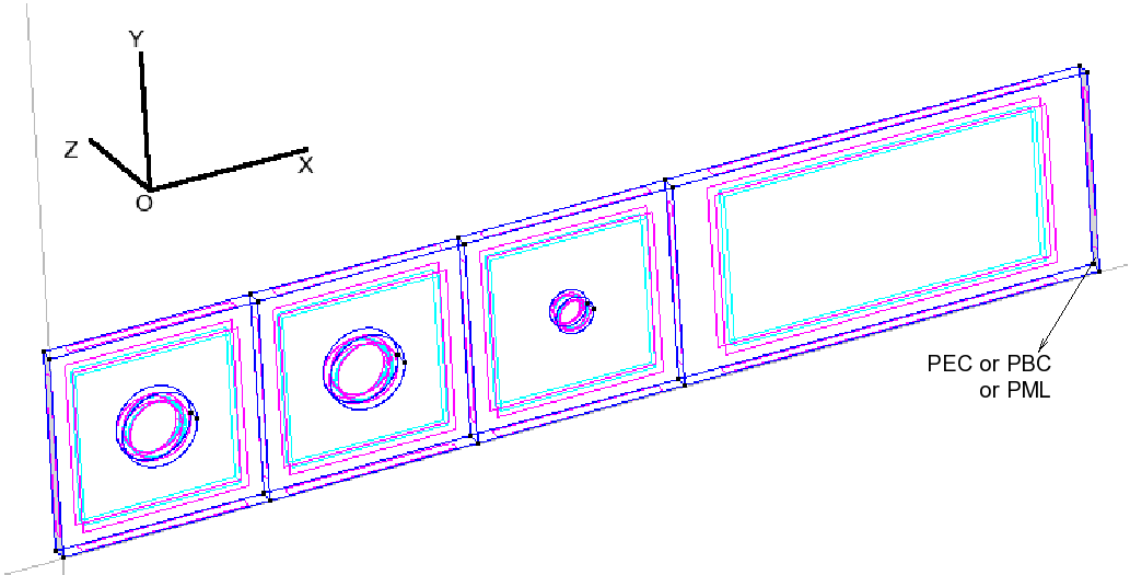


Fig 20.V Super-cellule avec un défaut de surface. La distance entre chaque élément est $a = 14mm$ et le trou d'air est égal à $2 * a$. Les rayons des 2 premières tiges sont $0.018a$ et le rayon de la tige de surface est 0.009a. La permittivité relative de la tige est $\epsilon_r = 11.56$.

Dans l'exemple Fig 20.V, on considère la structure périodique du réseau de tiges avec un défaut de surface. La cellule de base est constituée de 2 tiges diélectriques de tailles différentes de la tige de surface. En effet comme constaté en [26],[4] et [25], la modification de la tige de surface permet de faire apparaître des modes de surface (cf section I.2.4).

On considère une excitation parallèle à l'interface $k_{||}$=Oy . Si on appelle C la vitesse de la lumière dans le vide et ω la pulsation à laquelle on étudie du milieu, la région du diagramme de bande qui vérifie $\omega > Ck_{||}$ est appelée "cône de lumière". Toutes les solutions qui se situent en dessous du cône de lumière ($\omega \leq Ck_{||}$) sont soit des modes de surface soit des solutions qui se propagent dans le milieu périodique [4]. On peut rajouter que les modes qui se propagent dans l'air, et pas dans la structure, vérifient $\omega > Ck_{||}$ et peuvent être des modes guidés dans un défaut introduit dans la structure. Suffisamment loin de la structure

périodique, toutes les solutions peuvent être considérées comme des ondes planes. La valeur de la dernière composante du vecteur d'onde incident k_{iz} est fixée à 0. La structure est infinie dans la direction Oz et donc les super-cellules Fig 19.V et Fig 20.V peuvent être très fines dans cette direction (cf remarque V.3). Par définition le diagramme de bandes projeté dans la direction Oy est la superposition de plusieurs courbes $k_{iy} \in [0, \pi/dy]$ pour toutes les valeurs de $k_{ix} \in [0, \pi/dx]$ où dy et dx désignent respectivement la largeur et la longueur de la super-cellule étudiée. Les diagrammes de bandes projetés s'obtiennent avec le paramètre $TB = 3$. Pour localiser les modes de surfaces on doit d'abord calculer le diagramme de bandes projeté dans la direction $k_{||}$ =Oy du réseau infini de tiges Fig 21.V et ensuite comparer les résultats obtenus avec les diagrammes de bandes projeté des super-cellules Fig 22.V et Fig 25.V. Dans un premier temps, on identifie un certain nombre de zones au-dessus et en dessous de la ligne de lumière $\omega = Ck_{||}$ en rouge sur le diagramme Fig 21.V :

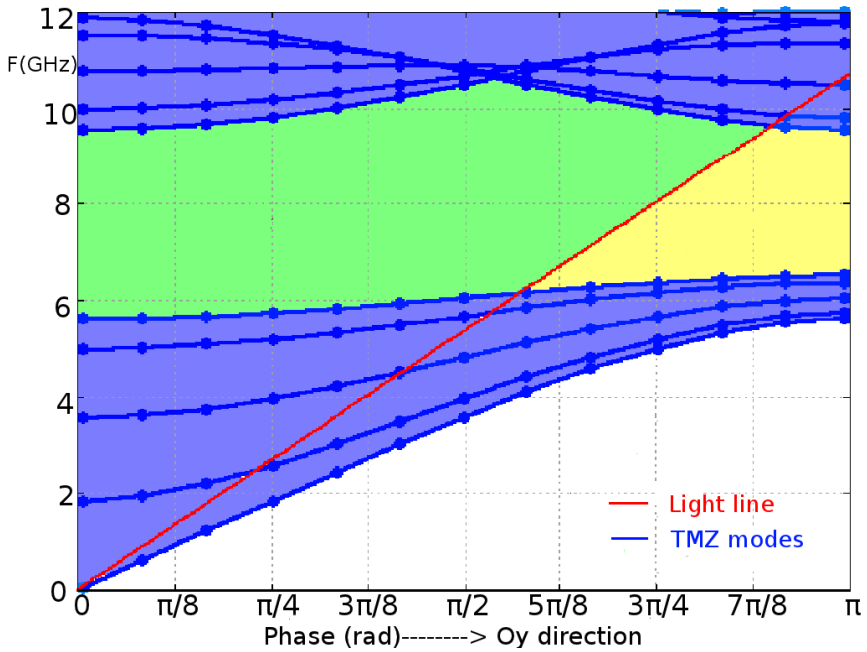


Fig 21.V Diagramme de bandes projeté du réseau de tiges infinies (Modes TM).
$(\phi_x, \phi_y) = (k_{ix}dx, k_{iy}dy) \in [0, \pi, \pi/2, 3 * \pi/4, \pi/4] \times [0, \pi/12, \ldots, \pi]$. $ND = 12$, $DV = 5$ et $TB = 3$.

Fig 21.V la région verte et la région jaune assemblées correspondent à la bande interdite du B.I.E. La zone bleue représente les solutions propagatives dans le B.I.E. Les modes guidés dans un défaut de la structure se propagent dans l'air mais sont évanescents dans le B.I.E. Ils sont donc situés dans la région verte, tandis que les modes de surface sont évanescents dans l'air et dans le B.I.E, ils sont donc localisés dans la région jaune. La cellule de base ne comporte qu'une seule tige (cf Fig 21. V). Dans toute cette partie Les calculs sont effectués sur notre machine (PC processeur intel avec une horloge de 2.80GHz). Le résidu admissible pour chaque valeur propre(24.IV) est de 10^{-7}. A chaque itération 17 modes propres sont calculés. Observons maintenant les diagrammes de bandes projetés des super cellules dans la direction Oy. Commençons par la super-cellule définie Fig 20.V :

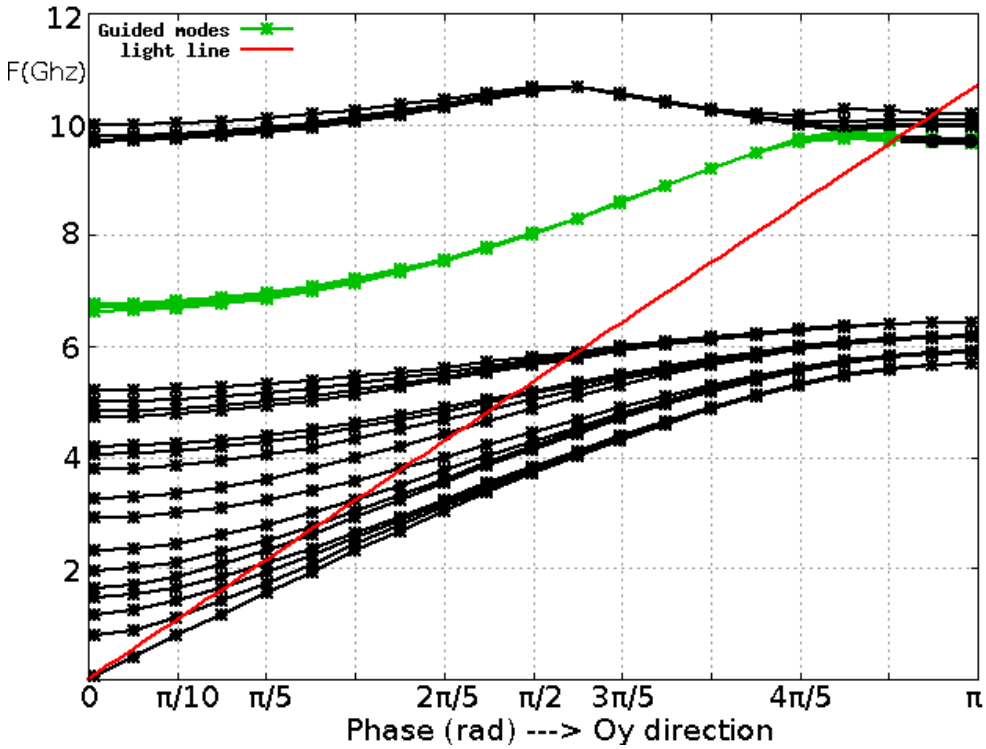


Fig 22.V Diagramme de bandes projeté de la super cellule Fig 19.V avec $\phi_x = k_{ix}dx = \in [0, \pi, \pi/2, 3 * \pi/4]$. $ND = 20,\ DV = 4$ et $TB = 3$. Le *shift* initial σ est égale à 8.61 Ghz ce qui correspond à $K_\sigma = 10$ et $d_c \simeq 35mm$.

Sur la Fig 22.V le diagramme de bandes projeté dans la direction Oy de la cellule avec le défaut linéaire représenté Fig 19.V avec $k_x \in [0, \pi, \pi/2, 3 * \pi/4]$. Le nombre d'arêtes totale est 12683. Le nombre d'arêtes esclave est de 3844. Le temps de calcul total pour 4*20 points du diagramme de Bandes est de 2 minutes et 48 secondes. La mémoire maximum utilisée est de 62 Mb. Les modes guidés apparaissent en vert dans la bande interdite.

Pour illustrer les résultats, on peut observer les champs résultants $Re(\tilde{E}_z)$ et $|\tilde{E}_z|$ d'un mode guidé dans le plan z=0.0005. La fréquence du mode observé est 8.58 Giga Hertz :

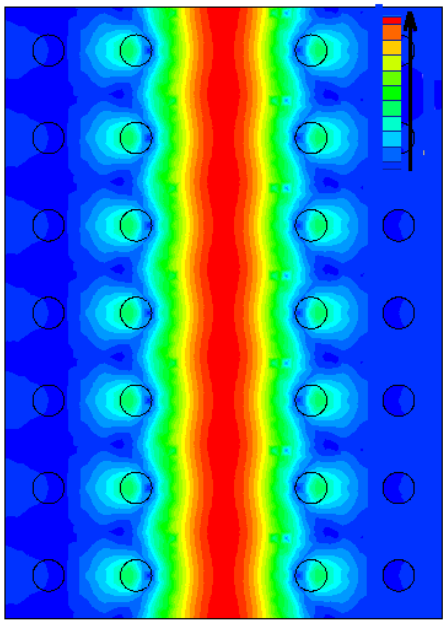


Fig 23.V
$|\tilde{E}_z|$ pour $k_{ix} = \pi/dx$, $k_{iy} = 0.6\pi/dy$.

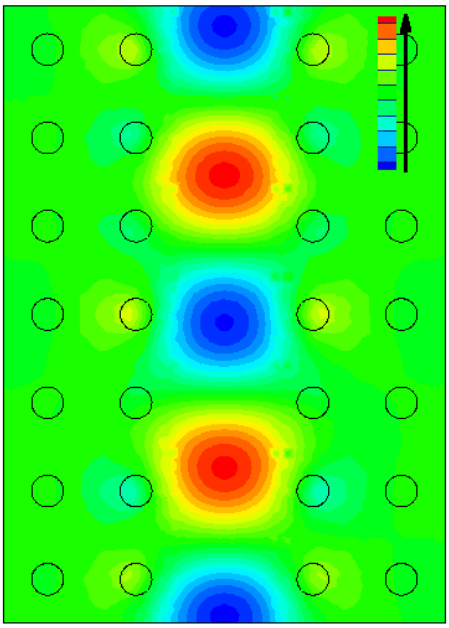


Fig 24.V
$Re(\tilde{E}_z)$ pour $k_{ix} = \pi/dx$, $k_{iy} = 0.6\pi/dy$.

Sur les Fig 23.V et Fig 24.V les variations des champs sont indiquées sans donner leurs valeurs. Sur la Fig 23.V le champ est localisé dans l'espace laissé par les tiges "manquantes". Sur la Fig 24.V la propagation du mode guidé se traduit par un déroulement de la phase selon la direction Oy, avec une évanescence de part et d'autre du trou selon Ox. Sur la Fig 25.V le diagramme de bandes de la cellule dans la direction Oy avec le défaut de surface Fig 20.V avec $\phi_x = k_{ix}dx \in [0, \pi, \pi/2, 3 * \pi/4]$. Le nombre total d'arêtes est 10471. Le nombre d'arêtes esclaves est de 4042. Le temps de calcul total 4*16 points du diagramme de bandes est de 1 minute 38 secondes. La mémoire maximum utilisée est de 60 Mb. Les modes de surface sont en jaune.

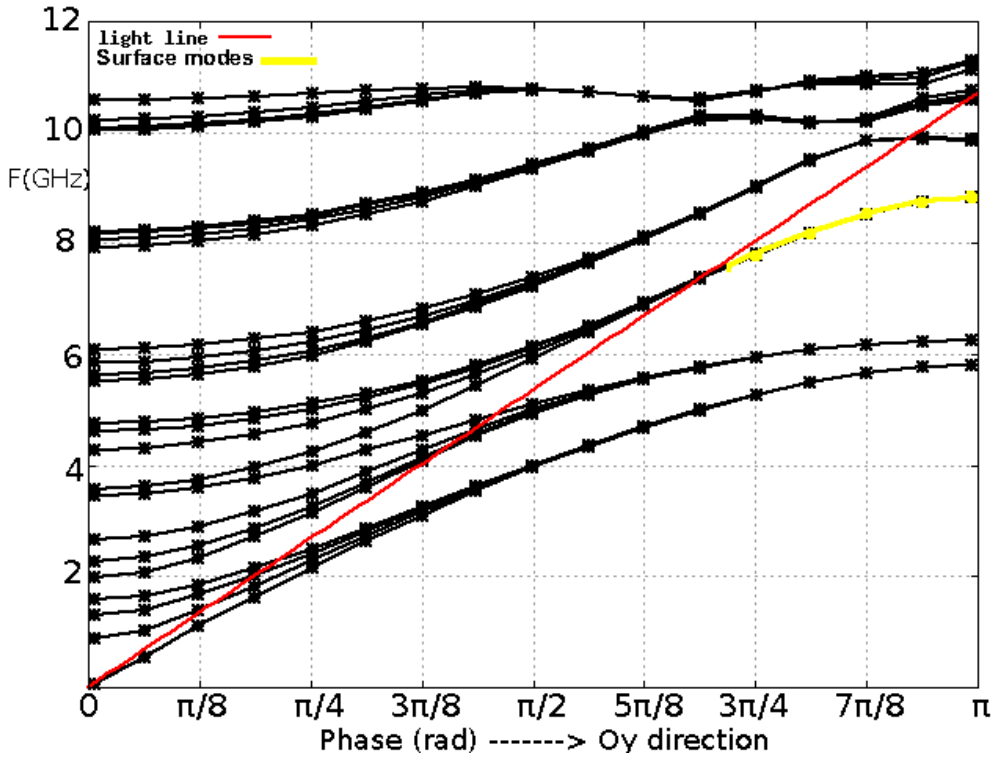


Fig 25.V Diagramme de bandes projeté de la super cellule Fig 20.V avec $\phi_x = k_{ix}dx \in [0, \pi, \pi/2, 3 * \pi/4]$. $ND = 16$, $DV = 4$ et $TB = 3$. Le *shift* initial σ est égal à 8.1 Ghz ce qui correspond à $K_\sigma = 7$ et $d_c \simeq 37mm$.

Pour confirmer les résultats, on peut observer les champs résultants $Re(\tilde{E}_z)$ and $|\tilde{E}_z|$ d'un mode de surface dans le plan z=0.0005. La fréquence du mode observé est 8.51 Giga Hertz :

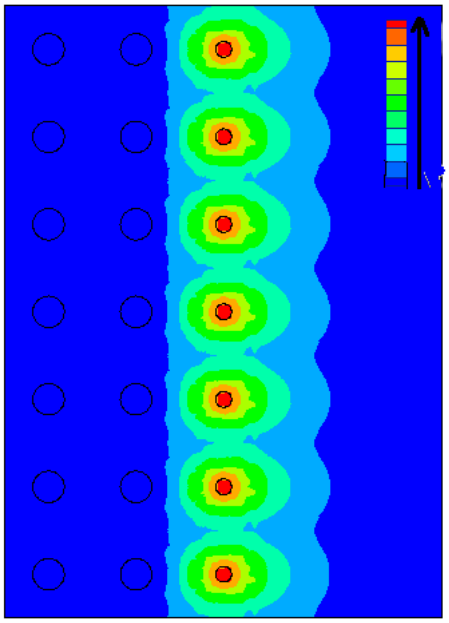


Fig 26.V $|\tilde{E}_z|$ pour $k_{ix} = 0$, $k_{iy} = 7\pi/8dy$.

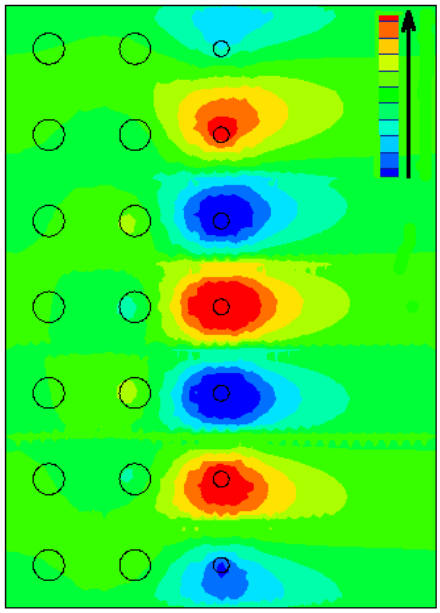


Fig 27.V $Re(\tilde{E}_z)$ pour $k_{ix} = 0$, $k_{iy} = 7\pi/8dy$.

Sur les Fig 26.V et Fig 27.V, d'une part le mode se propage le long de la surface en "s'accrochant" aux défauts et d'autre part le champ est évanescent de par et d'autre de la surface.

V.4.3 Outil complémentaire de caractérisation des ondes de surface

Comme nous venons de le voir, les ondes de surface sont des solutions se propageant à l'interface air-BIE. Ces solutions se situent toujours en dessous du cône de lumière dans le diagramme de bande. Mais nous avons vu Fig.21.V que toutes les solutions qui se situent en dessous du cône de lumière ne sont pas nécessairement des ondes de surface. Pour discriminer les cas où il y a une propagation significative de l'énergie électromagnétique le long des directions de périodisation des cas où l'énergie stagne, on propose d'évaluer les flux des vecteurs de Poynting à travers les surfaces où sont appliquées les conditions aux limites périodiques. Afin de rendre compte de cette propagation d'énergie, on doit exploiter ces flux en tenant compte de différentes considérations :

- Par construction le flux sortant sur une face maître est égal au flux entrant sur la face esclave correspondante. Il est donc inutile de définir une énergie à la fois sur les faces maîtres et les faces esclaves.
- Les différents flux des différentes faces maîtres ne doivent pas pouvoir se compenser. En effet comme le sens de propagation de l'énergie est pris en compte, sa valeur peut être négative sur une face maître et positive sur une autre. Pour calculer l'énergie totale, on effectue donc la somme de la valeur absolue des énergies de chaque face maître.
- On doit pouvoir comparer l'énergie propagée par rapport à une valeur d'énergie stockée contenue dans la cellule de base.

En prenant en compte tous ces critères on définit alors la quantité :

$$E\% = \frac{\sum_{k=1}^{N_M} \left| \int_{M_k} \Re(\mathbf{E} \times \mathbf{H}^*).\boldsymbol{\nu}_k \right|}{100 \int_V \mu |\mathbf{H}|^2 + \epsilon |\mathbf{E}|^2}. \tag{1.V}$$

N_M et $\boldsymbol{\nu}_k$ désignent respectivement le nombre de face maître (≤ 3) et la k-ième normale sortante de la face maître M_k. Cette quantité notée comme un pourcentage n'est pas nécessairement inférieure à 100. Elle traduit la quantité d'énergie transmise d'une cellule à l'autre par rapport à l'énergie stockée dans le volume représentée par $\int_V \mu |\mathbf{H}|^2 + \epsilon |\mathbf{E}|^2$.

Plus précisément sur chaque face maître $E\%$ représente le rapport entre la somme des flux des vecteurs de Poynting et l'énergie stockée dans le volume.

Remarque V.4. $E\%$ est non nul si $\Re(\mathbf{E} \times \mathbf{H}^*).\boldsymbol{\nu}$ a un signe dominant sur une des faces maîtres M_k. Autrement dit, les phénomènes de compensation ne sont pas suffisants pour annuler $E\%$ (cf Fig 28.V), ce qui permet à $|\int_{M_k} \Re(\mathbf{E} \times \mathbf{H}^*).\boldsymbol{\nu}|$ d'être non nulle. Par définition, les modes de surfaces ont une propagation d'énergie non nulle entre chaque élément. Ainsi, on fixe un seuil de $E\% = 1\%$ convenable pour discriminer les "modes de surfaces" des autres modes. En effet si $E\% > 1\%$, alors sur au moins une des faces maîtres M_k, il existe des vecteurs de Poynting non nuls, qui ont des directions non orthogonales à la normale $\boldsymbol{\nu}_k$ (cf Fig 29.V).

V.4.3.1 Cas où $E\%$ est nul

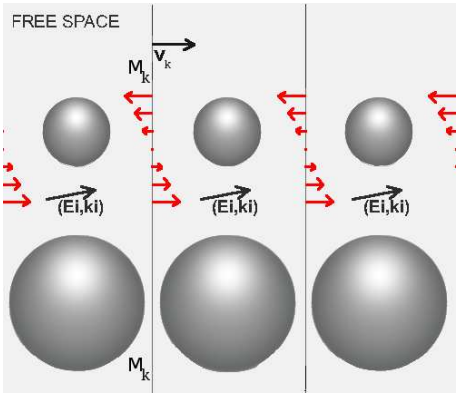


Fig 28.V Les vecteurs de Poynting représentés en rouge se compensent deux à deux.

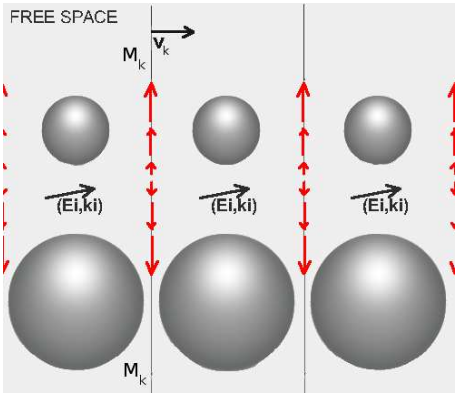


Fig 29.V Les vecteurs de Poynting représentés en rouge sont orthogonaux à la normale.

Les Fig 28.V et Fig 29.V représentent certaines directions possibles des vecteurs de Poynting de solutions se propageant à l'interface entre l'air et un réseau de billes métalliques plus petites en surface. Ces directions annulent la quantité $E\%$ définie en (1.V). Sur la Fig 28.V les vecteurs de Poynting se compensent. Sur la Fig 29.V, les vecteurs de Poynting sont orthogonaux à la normale $\boldsymbol{\nu}_k$. Pour annuler la quantité $E\%$, il y a aussi le cas évident des vecteurs de Poynting nuls sur M_k qui n'est pas représenté.

V.4.4 Surface Haute Impédance

Dans cette partie on s'intéresse au calcul et à la caractérisation des modes d'un réseau de champignons métalliques posés sur un plan de masse. Ce dispositif connu pour ses propriétés de bandes interdites pour des ondes de surface a fait l'objet d'études notamment dans [20] et [5].

Dessin de la cellule de base

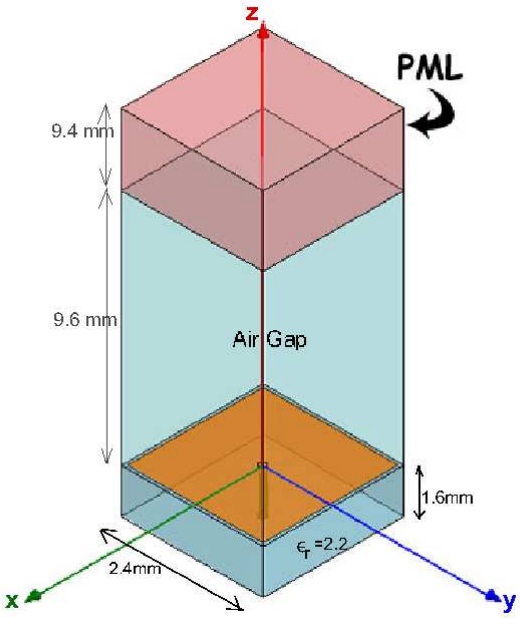


Fig 30.V Cellule de base du réseau de champignons métalliques.

Fig 30.V La pas du réseau est de 2.4 mm, l'espacement entre les patchs est de 0.15 mm et la largeur des vias métalliques est de 0,36 mm. Le substrat diélectrique en dessous des plaques carrées est du téflon ($\epsilon_r = 2,2$) et l'épaisseur totale de ces plaques est de 1,6 mm [20]. Le domaine est fermé par un matériau PEC ou PML[75] avec un trou d'air supérieur à 6 fois la hauteur des champignons pour simuler l'espace libre. Sa hauteur est égale 6*1.6 mm=9.6 mm. à La fréquence minimale absorbée F_{im} est égale à 8 GHz et la hauteur du volume PML est approximativement $\lambda_{im}/4 \simeq 9.4mm$ (cf section III.3.3). La PML est dans la direction OZ ($TPML$=1). Le calcul du diagramme de bandes comporte certaines difficultés. En effet premièrement l'espace entre les champignons est très petit comparé aux dimensions de la structure et donc cette zone devra être maillée très finement pour garantir la justesse des résultats comme nous l'avons vu dans la section IV.4.6. D'autre part, nous allons voir que les solutions varient très fortement au cours du calcul du diagramme de bande et la présence d'une PML peut modifier le nombre de solutions parasites par rapport aux cas sans PML(cf remarque IV.10). Ainsi le nombre de redémarrage de l'algorithme de tri vu section IV.4.5 peut devenir supérieur à 1 sur certains points du diagramme de bande et, par voie de conséquence, le temps de calcul peut lui aussi augmenter fortement. Le diagramme de bande obtenu avec une cellule de base fermée avec une condition PEC ou une condition PML est globalement identique à celui obtenu par Sievenpiper[20]. Nous retrouvons bien une bande interdite entre 12 et 15 GHz :

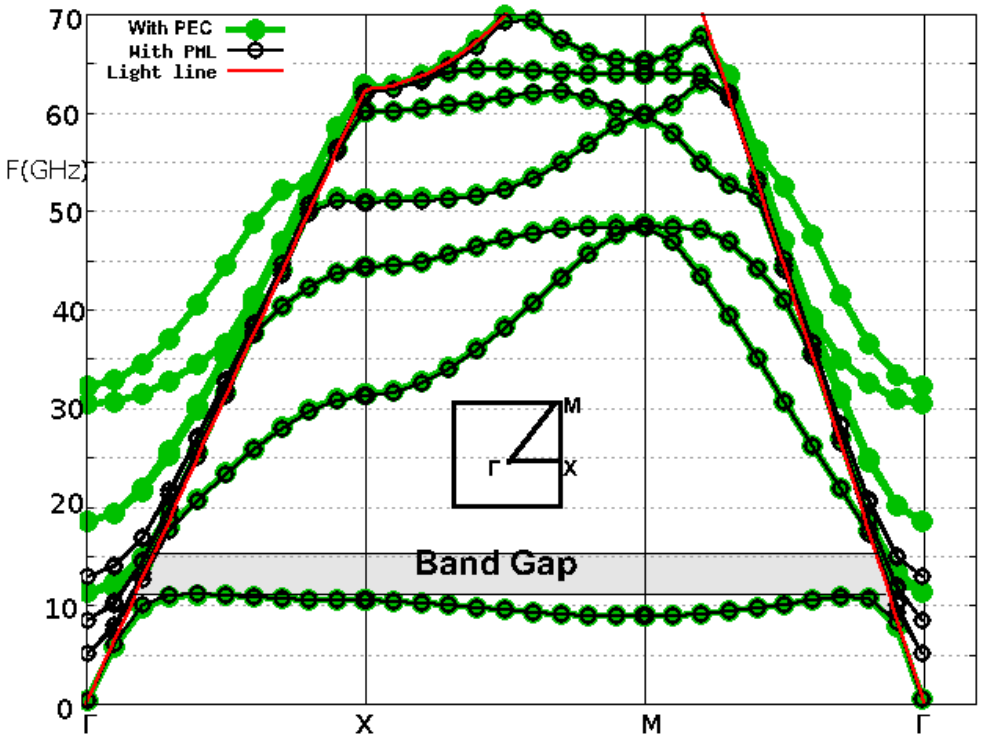


Fig 31.V Diagramme de bandes de la H.I.S

Sur la Fig 31.V nous pouvons observer que si l'on applique une condition PML , tous les modes radiatifs dans l'air ont tendance se rapprocher de la ligne de lumière, alors que tous les autres modes ne changent pas. Ce résultat est en bon accord avec la physique puisque

la PML absorbe les modes radiatifs dans l'air. Les valeurs sous le cône lumière sont en bon accord avec ceux présentés dans [20] .

Pour une étude plus complète observons maintenant l'allure des valeurs spatiales de $|\tilde{\mathbf{E}}|$ dans le plan de symétrie vertical ainsi que les vecteurs de Poynting associés quand le domaine est fermé par une condition PEC. On trace les quatre premiers modes du point M du diagramme de bandes Fig 31.V. Le terme $E\%$ (1.V) est aussi précisé :

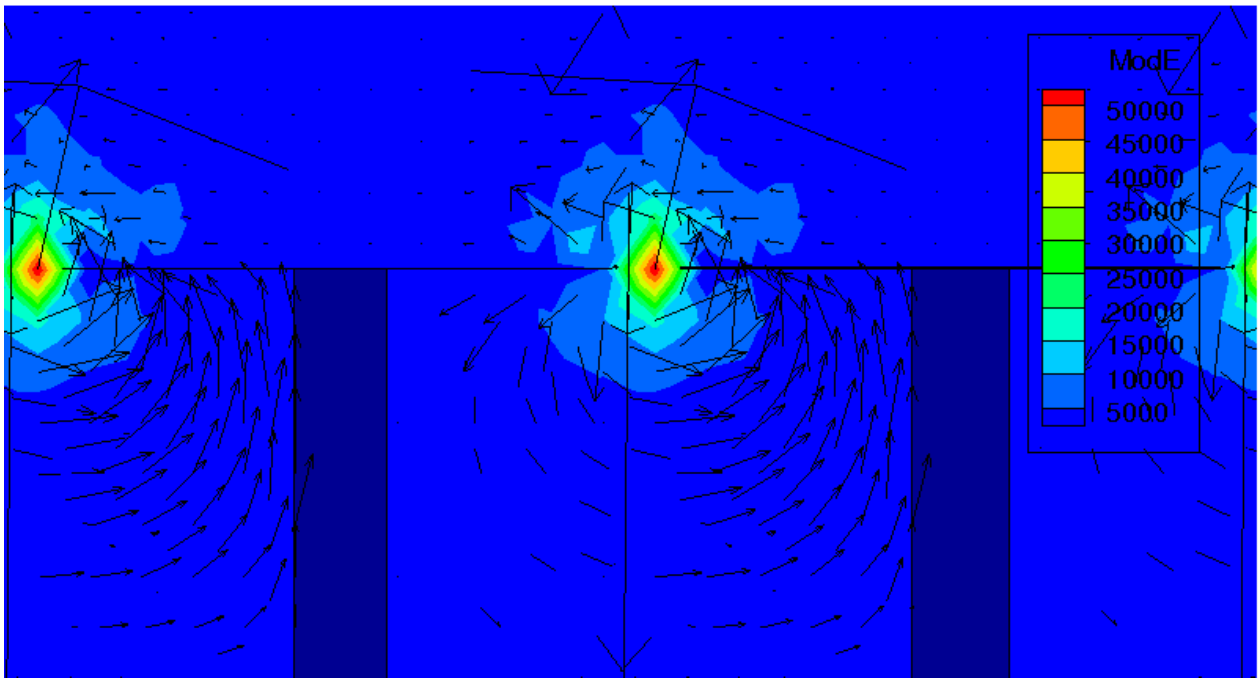


Fig 32.V La fréquence du mode est 8.9 GHz et $E\% = 1.4\%$.

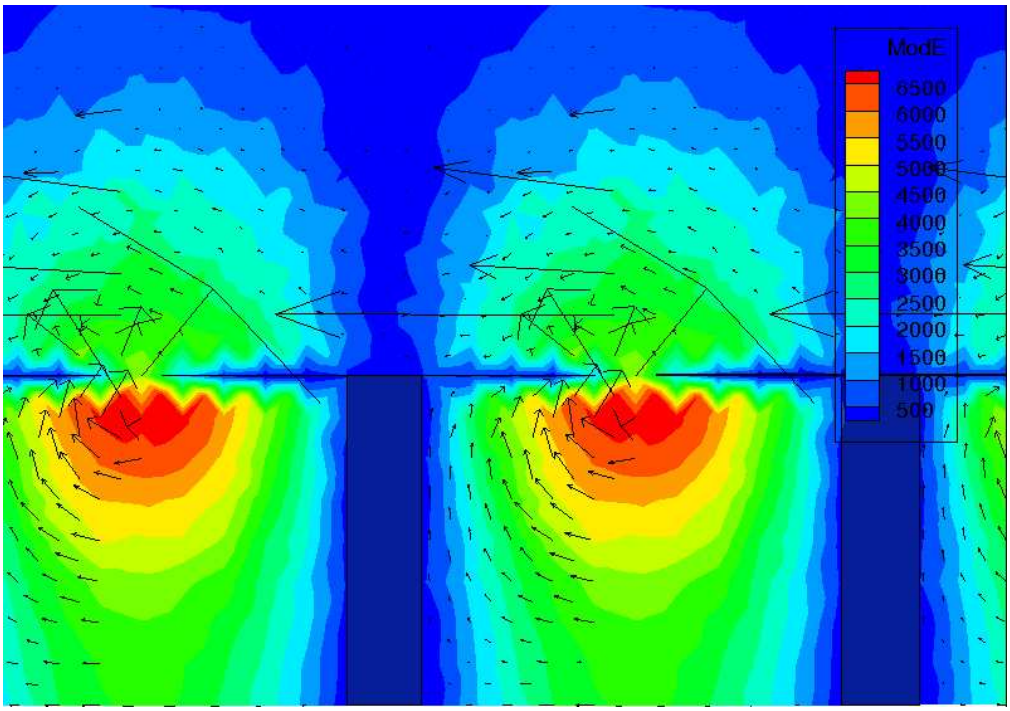


Fig 33.V La fréquence du mode est 48.63 GHz, et $E\% = 19.45\%$.

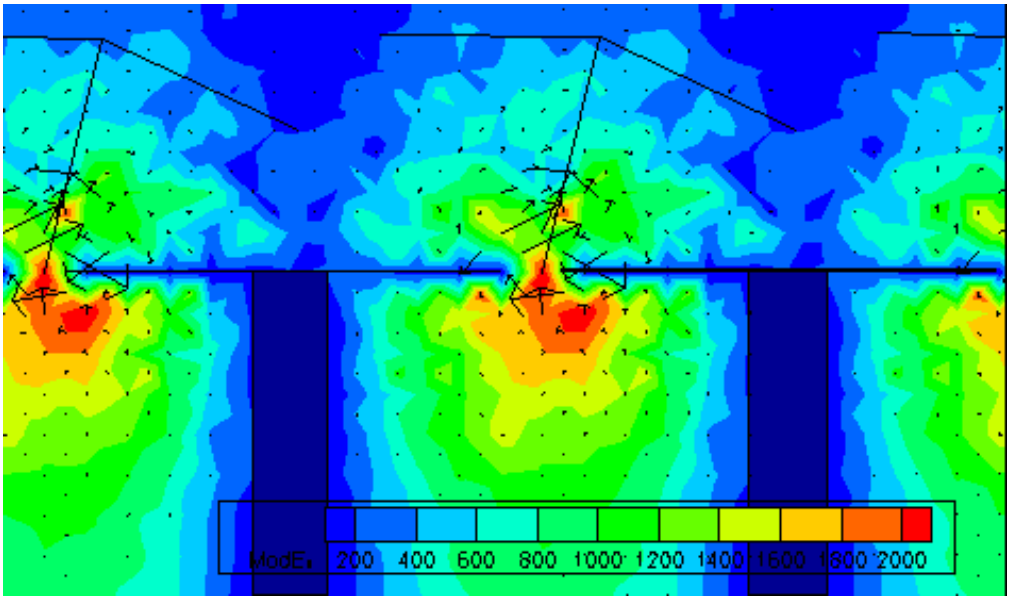


Fig 34.V La fréquence du mode est 48.52 GHz, et $E\% = 26\%$.

Sur les Fig 32.V l'énergie est localisée dans l'espace entre les surfaces métalliques. Sur les champs Fig 33.V et Fig 34.V l'énergie est plus étalée mais encore principalement localisée entre les surfaces métalliques. Les solutions observées sont situées en dessous du cône de lumière et donc elles ne sont pas propagatives dans l'air. Dans tous les cas la valeur de $E\% > 1$ indique un passage d'énergie entre les différents éléments du réseau comme nous l'avons vu remarque V.4. La faible valeur de $E\%$ sur la Fig 32.V s'explique par la compensation des différents vecteurs de Poynting présents sur chaque face maître.

V.4.5 Ondes de surfaces excités dans une antenne réseau constituée de patchs dans des cavités

On a pu observer que les ondes de surface conduisent à des phénomènes de couplage entre les éléments rayonnants d'un réseau, et sont la cause de directions aveugles dans les diagrammes de rayonnement. Dans cette partie on s'intéresse à une antenne réseau constituée de 2 patchs empilés dans une cavité parfaitement conductrice :

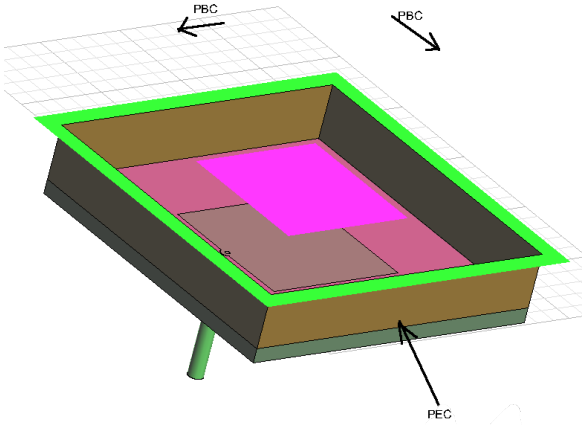


Fig 35.V Vue 3d de l'antenne.

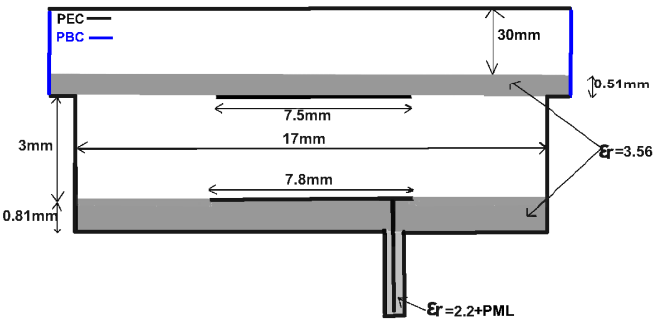


Fig 36.V Dimensions de l'antenne.

Comme dans le cas des surfaces HIS on pourrait fermer le domaine par une condition PML mais cela n'a pas d'intérêt dans notre cas puisque les solutions cherchées sont celles qui se propagent le long de l'interface air-BIE (cf Fig 32.V). Cependant, comme précédemment l'espace supérieur doit être supérieur à 6 fois la hauteur de l'antenne pour pouvoir appliquer les raisonnements établis section V.4.2. L'ajout d'une PML à la fin du câble permet d'empêcher l'apparition de solutions dues aux aller-retour de l'énergie dans le coax. Ici encore, on règle la longueur de la zone PML($\simeq 12cm = \lambda_{im}/4$) en fonction de la fréquence minimale d'absorption 2.5 GHz (cf section III.3.3). La PML est dans la direction Oz ($TPML$=1). Étant donné que les antennes sont piégées dans des cavités, toutes les solutions en dessous du cône de lumière ne peuvent pas se propager dans la structure périodique, sauf le long de la surface qui contient le patch supérieur. L'existence de modes de surface traduit clairement l'existence possible de couplages inter-éléments.

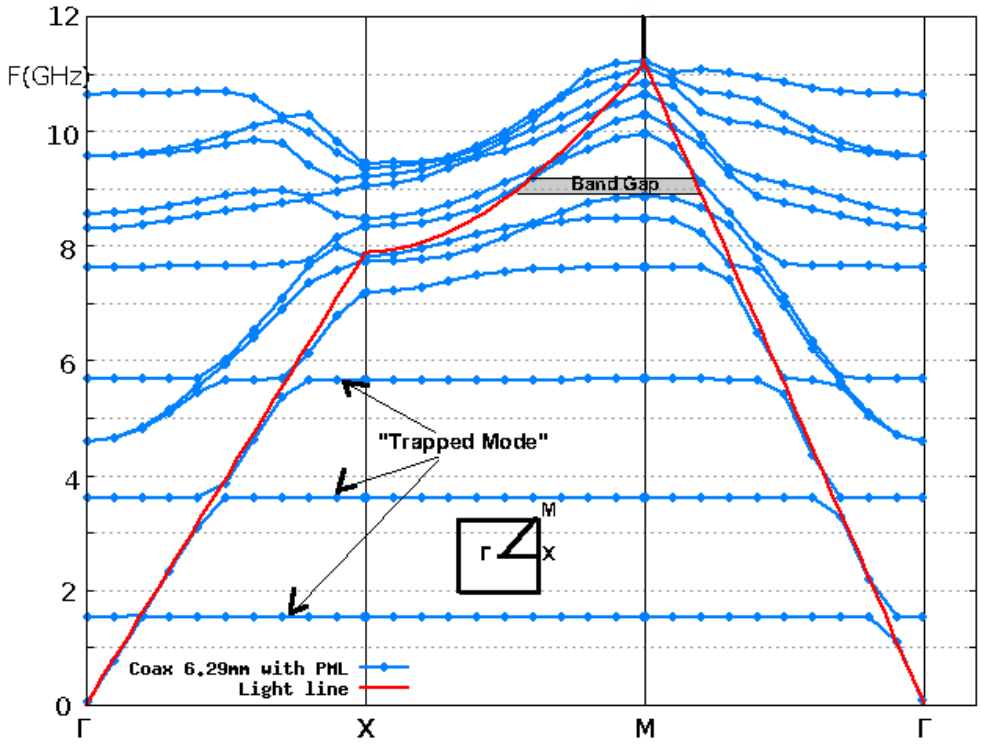


Fig 37.V Diagramme de bandes du réseau d'antennes patch.

Sur la Fig 37.V nous constatons qu'en dessous du cône de lumière il existe des modes de surface et des modes qui sont piégés dans la cavité qui ne voient pas le déphasage du aux conditions périodiques *Maître-esclave*(cf section III.3.5). Ils correspondent à des lignes droites horizontales dans le diagramme de bande Fig 37.V et leur valeur $E\%$ (1.V) est inférieure à 1. Sous le cône de lumière, nous pouvons observer une bande interdite entre 8,9 GHz et 9,2 GHz ce qui signifie théoriquement que, d'après la section V.4.2, il ne peut pas exister de modes de surface dans cette zone. Une étude expérimentale supplémentaire peut être nécessaire pour déterminer l'existence ou pas de modes de surface dans ce domaine. Maintenant nous pouvons observer la norme du champ électrique $|\tilde{\mathbf{E}}|$ de certains modes ainsi que le vecteur de Poynting dans le plan y=3,775mm pour le point M du diagramme de bandes Fig 37.V.

Le terme $E\%$ (1.V) est aussi précisé :

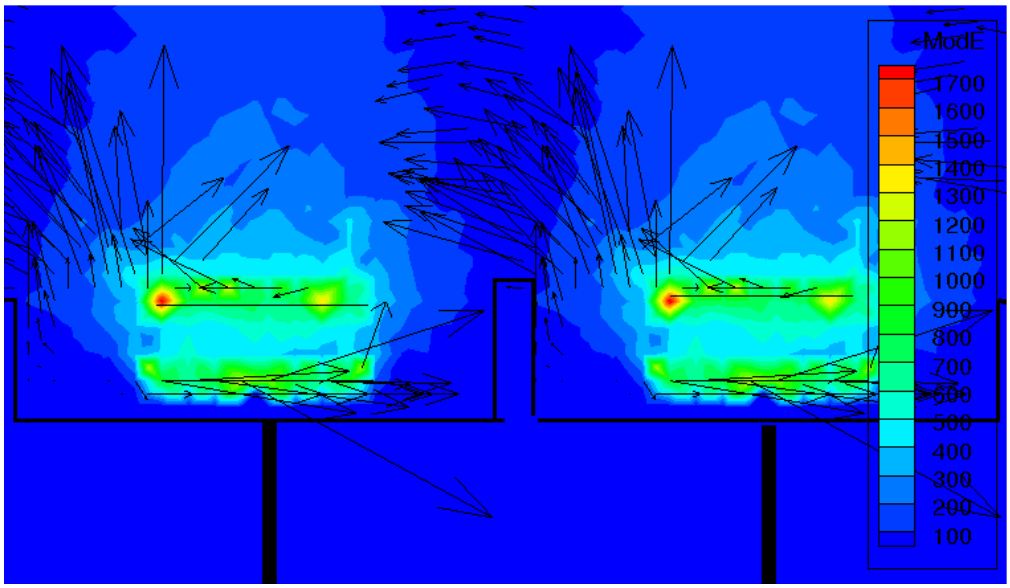


Fig 38.V 'Mode de surface'.
La fréquence du mode est 9.94 GHz et $E\% = 12.7\%$.

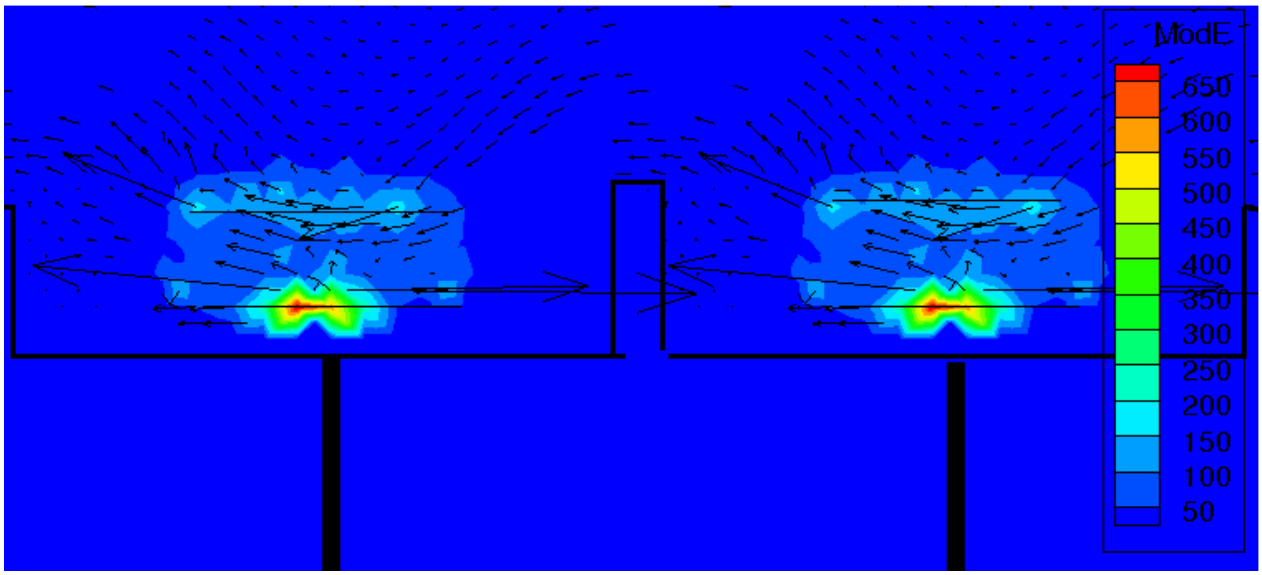


Fig 39.V '*Mode piégé*'.
La fréquence du mode est 5.67 GHz et $E\% \simeq 0\%$.

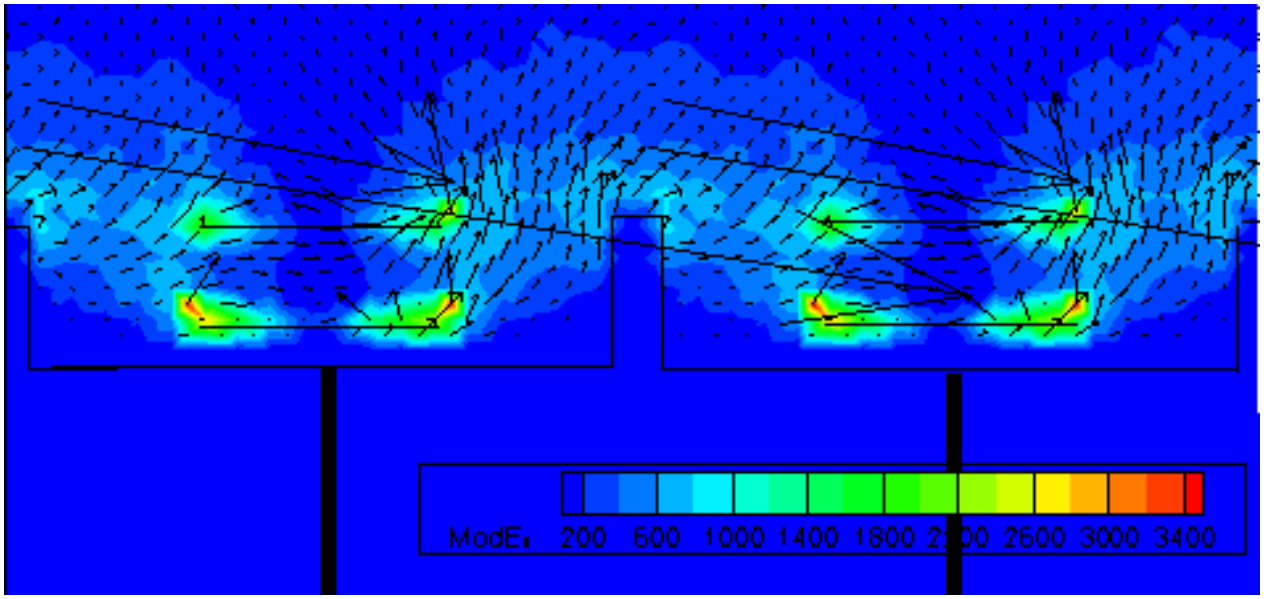


Fig 40.V 'Mode de surface'.
La fréquence du mode est 8.48 GHz et $E\% = 25.3\%$.

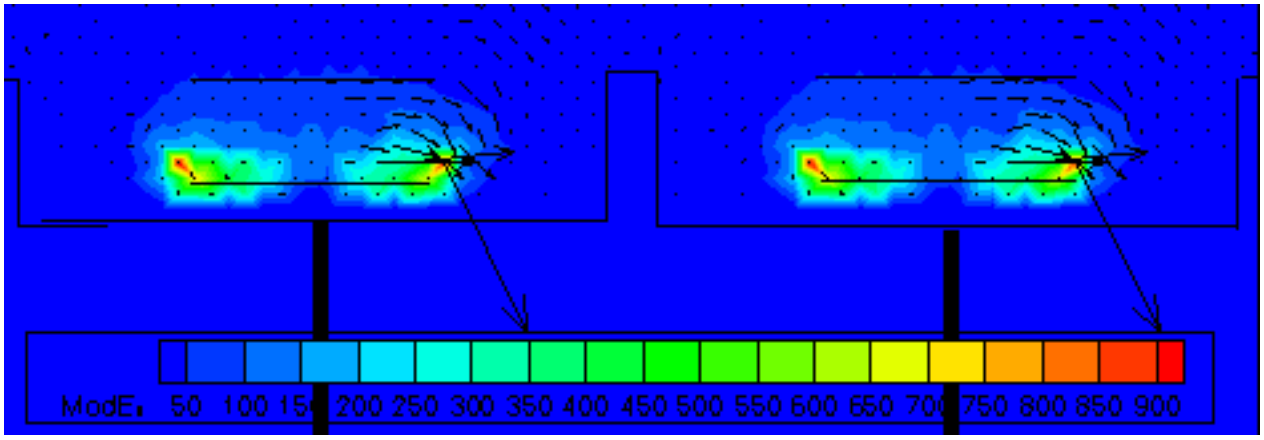


Fig 41.V '*Mode piégé*'.
La fréquence du mode est 3.6 GHz et $E\% \simeq 0\%$.

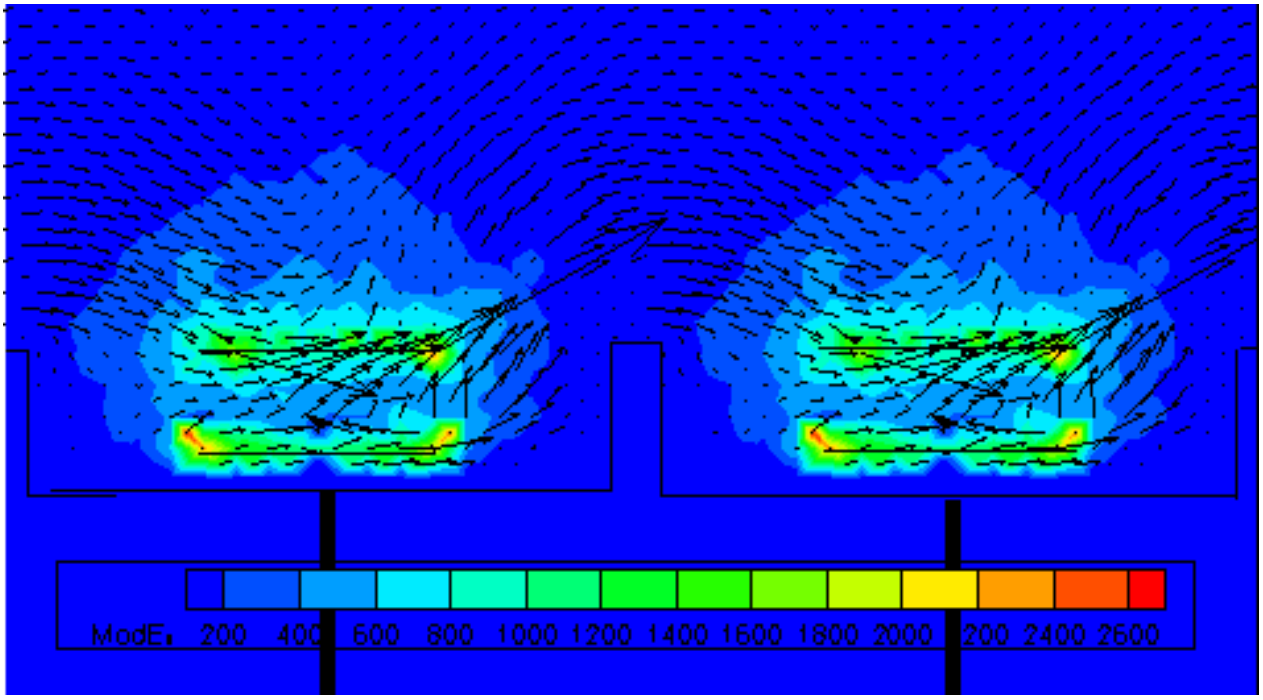


Fig 42.V 'Mode de surface'.
La fréquence du mode est 8.87 GHz et $E\% = 5.5\%$.

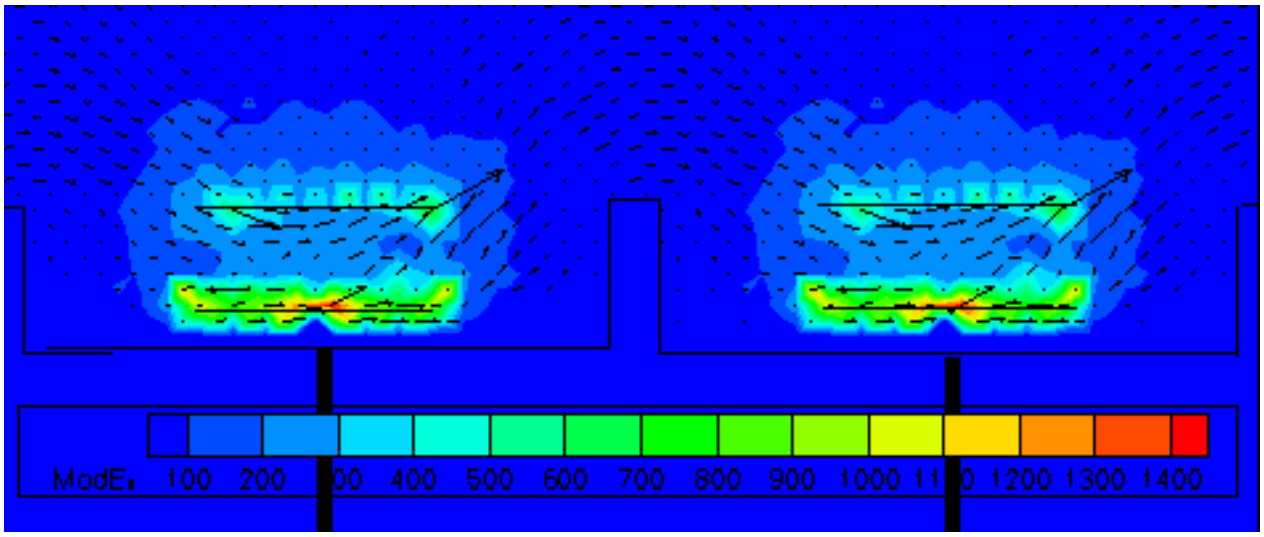


Fig 43.V '*Mode piégé*'.
La fréquence du mode est 7.63 GHz et $E\% \simeq 0.15\%$.

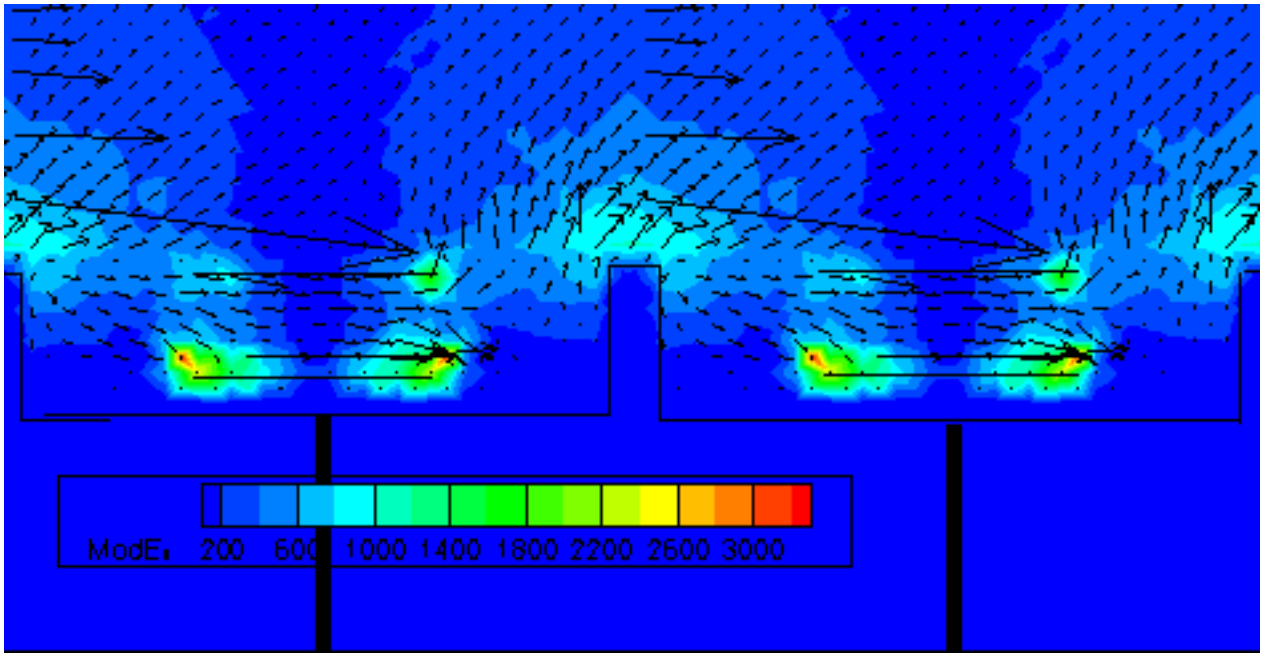


Fig 44.V 'Mode de surface'.
La fréquence du mode est 10.27 GHz et $E\% = 353\%$.

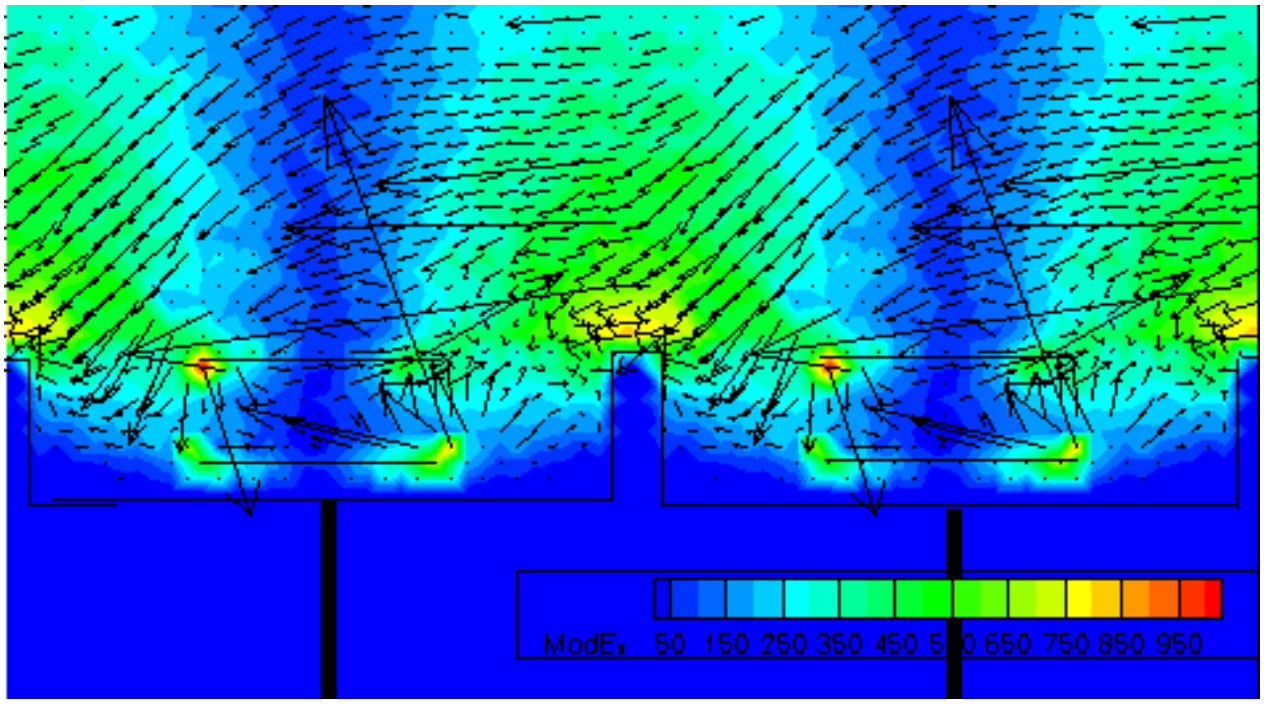


Fig 45.V 'Mode de surface'.
La fréquence du mode est 10.64 GHz et $E\% \simeq 465\%$.

Sur les champs observés Fig 38.V, Fig 40.V, Fig 42.V, Fig 44.V, Fig 45.V les modes de surface possèdent une valeur de $E\%$ non nulle et font apparaître un transfert d'énergie entre chaque élément contrairement aux '*Modes piégés*' Fig 39.V, Fig 41.V, qui ont tous une valeur $E\% \leq 0.1\%$. Sur la Fig 43.V on observe le premier mode qui n'est pas un '*Mode piégé*' puisqu'il n'est pas situé sur une ligne droite du diagramme de bande Fig 37.V, mais la quantité $E\% = 0.15\%$ est trop faible pour le considérer comme un mode de surface. Sur

les Fig 44.V et Fig 45.V les grandes valeurs de $E\%$ traduisent un fort couplage entre les éléments.

V.4.6 Résumé des performances de calcul obtenues pour la HIS et l'antenne réseau

Nous présentons les performances de calcul constatées pour les simulations HIS et antenne réseau. Nous avons utilisé un processeur Intel (R) Xeon 64 bits 2,80 GHz. Nous spécifions le temps CPU, le nombre total d'inconnues, le maximum de RAM utilisée et le nombre d'itérations nécessaires pour calculer 33 points du diagramme de bande. Les paramètres de résolution sont $\bar{tol} = 10^{-3}$et $\bar{k} = 15$ (cf section IV.4). Pour assurer l'exactitude, le nombre d'inconnues est grand et le maillage est fin dans les zones étroites du domaine géométrique. Pour la HIS la maille la plus fine mesure $10^{-1}mm$ (correspondant à 299 GHz) tandis que, pour l'antenne, la maille la plus fine est $5,10^{-2}mm$ (correspondant à 599 GHz). Le nombre d'inconnues est égal au nombre d'arêtes total moins le nombre d'arêtes PEC moins le nombre d'arêtes esclaves. Pour l'antenne et la HIS le nombre de points du diagramme de bande est 33 et les directions du vecteur incident sont les mêmes. Autrement dit $ND = 11$, $DV = 3$ et $TB = 2$ (cf section V.2). Le *shift* défini par (22.IV) multiplié par le K_σ est précisé en dernière colonne. Le choix des paramètres qui minimisent les temps de calcul n'a parfois pas été direct. C'est après plusieurs essais sur les premières valeurs obtenues que nous avons réussi à limiter le nombre d'itérations essentiellement conditionné par le choix du *shift* initial en Giga Hertz via la variable K_σ (cf remarque IV.10).

	Inconnues	RAM	Temps CPU	Itérations	$K_\sigma * \sigma$ (GHz)	d_c (en mm)
Antenne	119213	5Gb	1h29m55s	33	6.8 GHz ($K_\sigma = 8$)	44
HIS+PML	134486	6.8 Gb	7h34m28s	40	29.25 GHz ($K_\sigma = 6$)	10.2
HIS+PEC	113818	5.734 Gb	5h40m40s	33	45 Ghz ($K_\sigma = 4.5$)	6.66

Bien que le nombre d'inconnues du cas antenne soit supérieur à celui du cas HIS avec fermeture PEC, le temps de calcul est plus petit car la géométrie du maillage influe sur l'efficacité du solveur PARDISO[37](cf section IV.6.5.2). L'ajout d'une zone PML augmente le nombre d'inconnues du domaine et diminue le nombre de solutions parasites (cf remarque IV.10). Ceci peut créer un certain nombre d'itérations supplémentaires par rapport au nombre de points (33) nécessaire pour tracer le diagramme de bandes et conduit à une augmentation du temps de calcul (cf remarque IV.10).

V.5 Conclusion

Dans cette partie nous avons tout d'abord détaillé la structure des fichiers de données (section V.2) afin de clarifier l'utilisation des différents paramètres de résolution. Certains de ces paramètres ont été ensuite repris dans les deux sections suivantes V.3 et V.4. Section V.3 nous avons validé la fiabilité de notre outil par comparaison avec d'autres résultats classiques issus de méthodes alternatives. Enfin dans la section V.4 nous avons montré que

notre outil permet de fournir des informations supplémentaires sur les modes de surface. Notamment concernant le couplage mutuel entre plusieurs antennes patchs insérés dans une cavité. L'intérêt du paramètre $E\%$ permettant d'identifier les *Modes piégés* est également exposé. L'exploitation de ce paramètre peut être faite, en complément avec notre solveur dans le cadre d'une étude de conception d'antennes réseaux. De plus, les temps de calcul de dispositifs complexes restent abordables et on a pu constater que la forme géométrique, l'in-homogénéité des structures étudiées et la présence de métal ne sont pas des facteurs limitants.

. . . . . .

VI Conclusion et perspectives

VI.1 Conclusion générale

Reprenons les différents points qui ont été abordés au cours de la thèse.

Le contexte

Au début section I.1 nous introduisons les problèmes d'électromagnétisme liés aux structures périodiques.
Nous présentons section I.2 des applications possibles de l'utilisation des propriétés de bandes interdites des structures périodiques. Nous présentons comment ces propriétés peuvent être exploitées notamment pour la conception d'antennes et de radômes. Nous présentons notamment des dispositifs capables de générer des champs appelés ondes de surface, pouvant se propager à l'interface entre l'air et la structure périodique.

Objectifs (rappels)

L'objectif de la thèse a été de développer un outil de résolution par éléments finis en maîtrisant tous les paramètres d'entrée et de sortie afin de caractériser la propagation modale dans une structure périodique infinie avec le moins de facteurs limitants possibles. On doit pouvoir ainsi traiter des structures périodiques infinies hétérogènes, anisotropes, sans contraintes géométriques. Notre outil doit aussi permettre dans un second temps de caractériser les solutions pouvant se propager à l'interface air-structure qu'on appelle les "modes de surface".

Bilan

Après avoir établi le plan de travail, nous présentons la théorie de base de Floquet section II.1, permettant de comprendre le fonctionnement de notre outil et d'introduire le diagramme de bande, qui est une formalisation particulière des solutions. Il est utilisé notamment pour afficher les propriétés de bandes interdites.
Ensuite, section II.3,II.4 et II.5, nous abordons les différentes méthodes existantes avec leurs atouts et défauts respectifs.
Section III nous détaillons la modélisation de la méthode des éléments finis pour l'outil que nous avons développé. Nous indiquons les différentes conditions aux limites que nous pouvons combiner pour traiter le plus de cas possibles et nous énonçons les qualités et les défauts de la méthode. Les qualités principales énoncées sont les suivantes :

- Les domaines peuvent avoir des tailles très contrastées et des formes quelconques et les matériaux utilisés peuvent être de natures très différentes : anisotropes, métaux parfaitement conducteurs, etc.
- L'espace mémoire et le temps de calcul peuvent être optimisés, car les matrices stockées sont creuses.

– Il est envisageable de traiter les problèmes dispersifs (cf section VI.2).

Les défaults principaux sont :

– Les contraintes géométriques se reportent sur le maillage.
– Quand les matériaux sont dispersifs : la taille du problème de calcul de modes propres est multipliée par un entier qui dépend de l'expression de la permittivité (cf section VI.2).

Après, section IV, nous abordons les méthodes numériques nécessaires pour résoudre le problème éléments finis modélisé dans la section III. La méthode numérique est présentée ainsi que tous les paramètres de résolution nécessaires non seulement pour comprendre et optimiser les algorithmes utilisés, mais aussi pour pouvoir donner une interprétation pertinente des résultats obtenus. Nous constatons qu'il existe des valeurs propres nulles dans la section IV.1. Parmi ces valeurs propres nulles certaines sont identifiées comme solutions parasites. Nous indiquons d'abord comment éliminer toutes les valeurs propres nulles dans la section IV.4 et ensuite, dans la section IV.5, nous identifions les solutions parasites afin de les écarter du résultat final. Un algorithme d'élimination automatique de ces solutions est présenté pour les problèmes de cavités métalliques avec les fonctions de base que nous utilisons. Cet algorithme ne fonctionne pas pour les problèmes périodiques à moins de monter en ordre comme c'est évoqué dans l'article [36].

Dans la dernière section V nous nous sommes d'abord intéressés à la structure des fichiers de données afin de clarifier comment sont utilisés tous les paramètres de résolution. Ensuite nous démontrons que notre outil est fiable en comparant des résultats obtenus par diverses méthodes et en indiquant les performances de calculs qui permettent une utilisation sur un PC classique. Plus précisément l'outil de caractérisation des structures périodiques a été validé, optimisé et il est utilisable avec d'autres outils de l'ONERA (PPTM GID). Dans la section V.4, nous montrons comment nos outils avec l'aide d'une nouvelle quantité (1.V), peuvent être utilisés pour caractériser les modes de surface. À travers tous les exemples traités on a pu constater que la forme géométrique, l'in-homogénéité des structures étudiées et la présence de métal ne sont pas des facteurs limitants.

Ce travail a conduit à la production d'un code utilisable avec un ensemble d'outils de l'ONERA, et à plusieurs publications :

<u>En conférences</u>

– JNM 2011 Caractérisation de surfaces EBG par la méthode des éléments finis (Poster).
– EUCAP2011 Finite element method used to characterize surface modes in electromagnetic band gap structures (Présentation orale, article disponible dans la base IEEE).
– EUROEM 2012 Finite element method used to characterize surface modes in electromagnetic band gap structures (Présentation orale, article disponible dans la base IEEE).

<u>Dans une revue</u>

– Computer Physics : Efficient periodic band diagram computation using a finite element

method, Arnoldi eigensolver and sparse linear system solver.

Transition

Maintenant que nous venons de donner une conclusion générale sur l'ensemble du travail que nous avons effectué, nous allons aborder une perspective qui est la modélisation des problèmes dispersifs. Comme nous l'avons évoqué dans la section III.1, les propriétés de bande interdite des structures périodiques peuvent être obtenues de plusieurs façons avec la méthode des éléments finis.

1. La première méthode consiste à résoudre le problème périodique en calculant le champ solution à chaque fréquence donnée.
2. La deuxième méthode calcule les modes propres solutions dans la cellule de base du réseau périodisé à l'aide des conditions de Floquet.

Ainsi les milieux dispersifs peuvent être traités avec des résolutions de modes propres dans la section VI.2.1 et avec les problèmes bipériodiques qui comportent une source explicite section VI.2.2. Nous n'avons pas eu le temps de coder ces deux méthodes, mais nous allons voir que leur modélisation utilise les transformations décrites section III.3.5. Finalement dans la section, VI.3, nous présentons quelques perspectives qui pourraient s'inscrire dans la suite de cette thèse.

VI.2 Milieux dispersifs

VI.2.1 Modélisation du problème de modes propres pour les milieux dispersifs

VI.2.1.1 Introduction

Dans les problèmes de modes propres, la pulsation du mode qui se propage dans le milieu est une inconnue du problème. On comprend alors que si la permittivité dépend de cette pulsation, la formulation variationnelle (7.III) établie section III.2 est modifiée et l'équation (10.III) est plus compliquée à résoudre. Une transformation a déjà été proposée avec la méthode des ondes planes [29, 30, 26]. Nous proposons dans cette section d'adapter cette approche à la méthode des éléments finis. Une étude supplémentaire, que ne nous n'avons pas menée, serait nécessaire pour affirmer que les solutions approchées tendent vers des solutions physiques existantes.

VI.2.1.2 Milieux à pertes par conduction

On rappelle l'équation de Maxwell-Ampère[38] en régime harmonique (cf remarque II.1) :

$$\boldsymbol{\nabla}\,\mathbf{H} \;=\; \mathbf{J} \;-j\omega\mathbf{D}$$

Avec $\mathbf{D}$ l'induction électrique. Pour un milieu homogène, il existe un tenseur réel ε tel que :

$$\mathbf{D} \;=\; \varepsilon\mathbf{E} \tag{1.VI}$$

Remarque VI.1. On a choisi volontairement de prendre ε constant pour ne pas rendre le modèle encore plus compliqué, mais ε pourrait dépendre de la fréquence. Dans ce cas modèle plus général est traité dans la section VI.2.1.4.

D'autre part la loi d'Ohm(cf Annexe A.5) permet d'affirmer que :

$$\mathbf{J} = \sigma_e \mathbf{E} \tag{2.VI}$$

Où σ_e est la conductivité électrique du matériau à pertes. Donc l'équation de Maxwell Ampère devient :

$$\boldsymbol{\nabla} \times \mathbf{H} = j\omega\epsilon\mathbf{E} \tag{3.VI}$$

Avec

$$\epsilon = \varepsilon - j\frac{\sigma_e}{\omega}\delta_{cdc}, \tag{4.VI}$$

le terme ε représente l'effet diélectrique et $\frac{\sigma_e}{\omega}$, les pertes par conduction, *cdc* est la zone où le matériau comporte des pertes par conduction (δ est la fonction de Dirac). Étant donné que $k_0 = \frac{\omega}{C} = \omega\sqrt{\varepsilon_0\mu_0}$ et que $Z_0 = \sqrt{\frac{\mu_0}{\epsilon_0}}$ on a :

$$\epsilon_r = \frac{\epsilon}{\varepsilon_0} = \frac{\varepsilon}{\varepsilon_0} - j\frac{\sigma_e}{\omega\varepsilon_0}\delta_{cdc} = \varepsilon_r - j\frac{\sigma_e Z_0}{k_0}\delta_{cdc}$$

Par analogie avec la formulation établie (4.III), on peut écrire pour toute fonction test $\boldsymbol{\varphi} \in \mathcal{H}$:

$$\int_\Omega \boldsymbol{\nabla} \times \boldsymbol{\varphi}^* . \frac{\boldsymbol{\nabla} \times \mathbf{E}}{\mu_r} = k_0^2 \int_\Omega \epsilon_r \mathbf{E}.\boldsymbol{\varphi}^*$$

$$\Rightarrow \int_\Omega \boldsymbol{\nabla} \times \boldsymbol{\varphi}^* . \frac{\boldsymbol{\nabla} \times \mathbf{E}}{\mu_r} = k_0^2\varepsilon_r \int_\Omega \mathbf{E}.\boldsymbol{\varphi}^* - j\frac{\sigma_e k_0^2}{\omega\epsilon_0}\int_\Omega \mathbf{E}.\boldsymbol{\varphi}^*\delta_{cdc} \tag{5.VI}$$

La formulation variationnelle ne contient pas de terme de bord. Ils sont nuls dans les problèmes de modes propres (cf remarque III.3).

Remarque VI.2. Des termes de bords sont présents dans le cas des problèmes avec une source explicite tel que nous les abordons section VI.2.2. Dans ce cas, les pertes par conduction équation (4.VI) peuvent directement être intégrées dans l'expression de la permittivité qui est alors une constante complexe. En effet lorsqu'il y a une source explicite, nous calculons des solutions pour une pulsation ω donnée, alors que, dans le cas des problèmes de modes propres, la pulsation fait partie des inconnues.

L'équation (5.VI) peut aussi s'écrire :

$$\int_\Omega \boldsymbol{\nabla} \times \boldsymbol{\varphi}^* . \frac{\boldsymbol{\nabla} \times \mathbf{E}}{\mu_r} = k_0^2\varepsilon_r \int_\Omega \mathbf{E}.\boldsymbol{\varphi}^* - j\sigma_e Z_0 k_0 \int_\Omega \mathbf{E}.\boldsymbol{\varphi}^*\delta_{cdc}$$

Que l'on peut encore écrire :

$$k_0^2 U(\mathbf{E}, \boldsymbol{\varphi}) - k_0 V(\mathbf{E}, \boldsymbol{\varphi}) - S(\mathbf{E}, \boldsymbol{\varphi}) = 0$$

avec

$$U(\mathbf{E},\boldsymbol{\varphi})=\int_\Omega \varepsilon_r \mathbf{E}.\boldsymbol{\varphi}^*,\ V(\mathbf{E},\boldsymbol{\varphi})=j\sigma_e Z_0\int_\Omega \mathbf{E}.\boldsymbol{\varphi}^*\delta_{cdc},\ S(\mathbf{E},\boldsymbol{\varphi})=\int_\Omega \boldsymbol{\nabla}\times\boldsymbol{\varphi}^*.\frac{\boldsymbol{\nabla}\times\mathbf{E}}{\mu_r}$$

Après avoir appliqué la méthode de Galerkin on obtient alors le système matriciel suivant :

$$\begin{gathered}T(k_0)\mathbf{x}_1=(k_0^2\mathcal{A}-k_0\mathcal{B}-\mathcal{C})\mathbf{x}_1=0 \text{ avec } (\mathcal{A})_{qp}=U(\mathbf{W}_p,\mathbf{W}_q),\ (\mathcal{B})_{qp}=V(\mathbf{W}_p,\mathbf{W}_q),\\ (\mathcal{C})_{qp}=S(\mathbf{W}_p,\mathbf{W}_q),\quad \text{et } \tilde{\mathbf{E}}=\textstyle\sum_{p=1}^{DL} e_p\mathbf{W}_p,\ \mathbf{x}_1=[e_1,e_2,\ldots,e_{DL}].\end{gathered} \tag{6.VI}$$

Remarque VI.3. Le système (6.VI) est un problème de valeurs propres non linéaire. Il existe plusieurs méthodes pour le résoudre :

- Les méthodes itératives type Newton utilisées pour trouver les zéros de la fonctionnelle $F(\mathbf{x}_1,k_0)$ définie par :

$$F(\mathbf{x}_1,k_0)=\begin{bmatrix}T(k_0)\mathbf{x}_1\\ \mathbf{U}^*\mathbf{x}_1-1\end{bmatrix},\mathbf{U}\in\mathbb{C}^{DL}$$

  Le vecteur $\mathbf{U}$ est d'abord choisi comme un vecteur de la base canonique de $\mathbb{C}^{DL}$ et ensuite, par le procédé de Gram-Schmidt, on construit $\mathbf{U}$ orthogonal aux vecteurs vecteurs propres déjà claculés afin d'éviter de recalculer 2 fois le même[108].
- La méthode linéarisation du problème (6.VI) en un problème de valeur propre linéaire de dimension supérieure[108, 109] comme nous allons l'expliquer dans la suite du document.

En effet le système (6.VI) peut se mettre sous la forme d'un problème linéaire aux valeurs propres de dimension $2*DL$:

$$O\mathbf{X}=k_0P\mathbf{X},\ \mathbf{X}=\begin{bmatrix}\mathbf{x}_1\\ \mathbf{x}_2\end{bmatrix},\ O=\begin{bmatrix}0 & I_{DL}\\ \mathcal{C} & \mathcal{B}\end{bmatrix},\ P=\begin{bmatrix}I_{DL} & 0\\ 0 & \mathcal{A}\end{bmatrix} \tag{7.VI}$$

On peut conclure que pour les milieux définis par $(\varepsilon,\sigma_e,\mu)$, avec un modèle dispersif lié aux pertes conductives, on peut proposer une résolution qui double le nombre d'inconnues.

VI.2.1.3 Modèle de Drude

Dans certains cas, on peut expliciter la conductivité à l'aide du modèle de Drude [56] ce qui modifie la forme de la dispersion. Dans le modèle de Drude, on suppose qu'un électron du matériau étudié est soumis à une force électrique $e\mathbf{E}$. Ensuite, les collisions potentielles avec les atomes du réseau sont modélisées par la force de freinage $-\frac{1}{\tau}\mathbf{v}$ pour des électrons de masse m_e se propageant à la vitesse $\mathbf{v}$. L'équation du mouvement des électrons prend alors la forme suivante :

$$\frac{\partial\mathbf{v}}{\partial t}+\frac{\mathbf{v}}{\tau}=-\frac{e}{m_e}\mathbf{E} \tag{8.VI}$$

En prenant les solutions homogènes (sans second membre) de l'équation (8.VI), la vitesse des électrons en régime transitoire s'écrit :

$$\mathbf{v}(t) = \mathbf{v}_0 e^{\frac{-t}{\tau}}$$

La vitesse initiale des électrons notées $\mathbf{v}_0$ s'atténue en fonction du temps de manière exponentielle avec un facteur de décroissance $F_\nu = \frac{1}{\tau}$. τ est le temps de relaxation exprimé en secondes et donc F_ν est homogène à une fréquence que l'on l'appelle la fréquence de collision. En régime harmonique l'équation (8.VI) permet d'obtenir l'expression de la vitesse en fonction du champ électrique appliqué :

$$\mathbf{v} = \frac{-\frac{e}{m_e}}{j\omega + F_\nu}\mathbf{E} \tag{9.VI}$$

La densité de courant électrique $\mathbf{J} = -Ne\mathbf{v}$ se déduit alors de (9.VI) :

$$\mathbf{J} = \frac{\frac{Ne^2}{m_e}}{j\omega + F_\nu}\mathbf{E}$$

On en déduit l'expression de la conductivité :

$$\sigma_e(\omega) = \frac{1}{j\omega}\frac{\omega_p^2\epsilon_0}{1 - \frac{jF_\nu}{\omega}}, \text{ avec } \omega_p^2 = \frac{Ne^2}{m_e\epsilon_0} \tag{10.VI}$$

En injectant la conductivité définie (10.VI) dans la transformation initiale des équations de Maxwell on obtient :

$$\boldsymbol{\nabla} \times [\mu_r]^{-1}\boldsymbol{\nabla} \times \mathbf{E} - [k_0^2\varepsilon_r - \frac{\omega_p^2\delta_{cdc}}{C^2(1 - j\frac{F_\nu\delta_{cdc}}{\omega})}]\mathbf{E} = \mathbf{0}, \quad \text{dans } \Omega \tag{11.VI}$$

En multipliant chaque membre de léquation (11.VI) par le terme $(k_0 - j\frac{F_\nu}{C}\delta_{cdc})$ on obtient :

$$(k_0 - j\frac{F_\nu}{C}\delta_{cdc})\boldsymbol{\nabla} \times [\mu_r]^{-1}\boldsymbol{\nabla} \times \mathbf{E} - [k_0^2\varepsilon_r(k_0 - j\frac{F_\nu}{C}\delta_{cdc}) - k_0\frac{\omega_p^2\delta_{cdc}}{C^2}]\mathbf{E} = \mathbf{0}, \quad \text{dans } \Omega \tag{12.VI}$$

Remarque VI.4. *cdc* est la zone comportant des pertes par conduction. Quand δ_{cdc} est partout nulle, on retrouve bien la formulation classique non dispersive .

On multiplie ensuite par une fonction test $\boldsymbol{\varphi} \in \mathcal{H}$ et on intègre en utilisant les transformations indiquée dans la section III.2.1 sans tenir compte des termes de bords sur $\partial\Omega$ qui sont nuls dans les problèmes que nous traitons et on obtient :

$$k_0\int_\Omega \boldsymbol{\nabla} \times \boldsymbol{\varphi}^*.\frac{\boldsymbol{\nabla} \times \mathbf{E}}{\mu_r} - j\frac{F_\nu}{C}\int_\Omega \boldsymbol{\nabla} \times \boldsymbol{\varphi}^*.\frac{\boldsymbol{\nabla} \times \mathbf{E}}{\mu_r}\delta_{cdc} = \int_\Omega \left[k_0^3\varepsilon_r - j\frac{F_\nu k_0^2\varepsilon_r}{C}\delta_{cdc} - \frac{k_0\omega_p^2}{C^2}\delta_{cdc}\right]\mathbf{E}.\boldsymbol{\varphi}^* \tag{13.VI}$$

Que l'on peut aussi écrire :

$$k_0^3 U(\mathbf{E},\boldsymbol{\varphi}) - k_0^2 V(\mathbf{E},\boldsymbol{\varphi}) - k_0 W(\mathbf{E},\boldsymbol{\varphi}) - S(\mathbf{E},\boldsymbol{\varphi}) = 0 \tag{14.VI}$$

avec

$$U(\mathbf{E},\boldsymbol{\varphi}) = \int_\Omega \varepsilon_r \mathbf{E}.\boldsymbol{\varphi}^*,\ V(\mathbf{E},\boldsymbol{\varphi}) = j\frac{F_\nu}{C}\int_\Omega \varepsilon_r \mathbf{E}.\boldsymbol{\varphi}^* \delta_{cdc},$$

$$W(\mathbf{E},\boldsymbol{\varphi}) = \frac{\omega_p^2}{C^2}\int_\Omega \mathbf{E}.\boldsymbol{\varphi}^* \delta_{cdc} + \int_\Omega \boldsymbol{\nabla}\times\boldsymbol{\varphi}^*.\frac{\boldsymbol{\nabla}\times\mathbf{E}}{\mu_r},\ S(\mathbf{E},\boldsymbol{\varphi}) = -\frac{jF_\nu}{C}\int_\Omega \boldsymbol{\nabla}\times\boldsymbol{\varphi}^*.\frac{\boldsymbol{\nabla}\times\mathbf{E}}{\mu_r}\delta_{cdc}.$$

On peut alors calquer sur la méthode des éléments finis, les travaux effectués pour la méthode des ondes planes [29, 30].

Après avoir appliqué la méthode de Galerkin sur des fonctions de base $\mathbf{W}_p,\ p \in [1,\dots DL]$ on obtient le système matriciel suivant :

$$(k_0^3\mathcal{A} - k_0^2\mathcal{B} - k_0\mathcal{C} - \mathcal{D})\mathbf{x}_1 = 0 \text{ avec } (\mathcal{A})_{qp} = U(\mathbf{W}_p,\mathbf{W}_q),\ (\mathcal{B})_{qp} = V(\mathbf{W}_p,\mathbf{W}_q),$$

$$(\mathcal{C})_{qp} = W(\mathbf{W}_p,\mathbf{W}_q),\ (\mathcal{D})_{qp} = S(\mathbf{W}_p,\mathbf{W}_q) \text{ et } \tilde{\mathbf{E}} = \sum_{p=1}^{DL} e_p\mathbf{W}_p,\ \mathbf{x}_1 = [e_1,e_2,\dots,e_{DL}] \tag{15.VI}$$

Ce système peut se mettre sous la forme d'un problème aux valeurs propres de dimension $3 * DL$:

$$O\mathbf{X} = k_0 P\mathbf{X},\ \mathbf{X} = \begin{bmatrix}\mathbf{x}_1\\ \mathbf{x}_2\\ \mathbf{x}_3\end{bmatrix},\ O = \begin{bmatrix}0 & I_{DL} & 0\\ 0 & 0 & I_{DL}\\ \mathcal{D} & \mathcal{C} & \mathcal{B}\end{bmatrix},\ P = \begin{bmatrix}I_{DL} & 0 & 0\\ 0 & I_{DL} & 0\\ 0 & 0 & \mathcal{A}\end{bmatrix}$$

Remarque VI.5. Pour ajouter des conditions limites périodiques il suffit d'appliquer les transformations de Floquets définies par la relation (26.III) aux matrices $\mathcal{A}$,$\mathcal{B}$,$\mathcal{C}$,$\mathcal{D}$ définies (15.VI).

Remarque VI.6. Encore une fois la solution du problème (15.VI) peut aussi se calculer avec une méthode itérative brièvement présentée dans la remarque VI.3.

On peut conclure que pour les milieux définis par $(\varepsilon,\omega_p,F_\nu,\mu)$, on peut encore proposer une résolution qui triple le nombre d'inconnues.

VI.2.1.4 Cas général

Les cas énoncés précédemment peuvent se généraliser lorsque que ε défini en (1.VI) et σ_e défini en (2.VI) sont des fractions rationnelles en ω. À partir de la formulation variationnelle (5.VI) on obtient un système multilinéaire polynomial en k_0 de degré d. Par exemple dans l'équation (14.VI) $d = 3$. Le degré du système dépend de l'expression de ε et de σ_e et le nombre d'inconnues du système matriciel obtenu après avoir appliqué la méthode de Galerkin

est exactement égal à $d*DL$. Dans ce cas, il est aussi toujours possible d'utiliser les méthodes itératives présentées remarque VI.3. Ce type de résolution est donc possible si la permittivité peut être exprimée à l'aide d'une fraction rationnelle en ω. Dans le cas où il n'existe pas d'expression formelle de la permittivité, on peut par exemple envisager de faire correspondre un ensemble de mesures expérimentales à une fraction rationnelle.

VI.2.2 Introduction d'une source explicite dans une structure bi-périodique

VI.2.2.1 Introduction

On cherche à modéliser un problème bi-périodique avec une source explicite. Cette méthode n'a pas été développée pendant la thèse, mais sa construction est très similaire à celle que nous avons décrite dans la section III. Un avantage significatif de cette méthode est la possibilité de traiter les cas dispersifs (cf remarque VI.7) sans changement d'écriture du système linéaire.

VI.2.2.2 Problème bi-périodique

Pour la description, nous avons choisi de nous placer dans un milieu bi-périodique dont la cellule de base est illustrée Fig 1.VI.

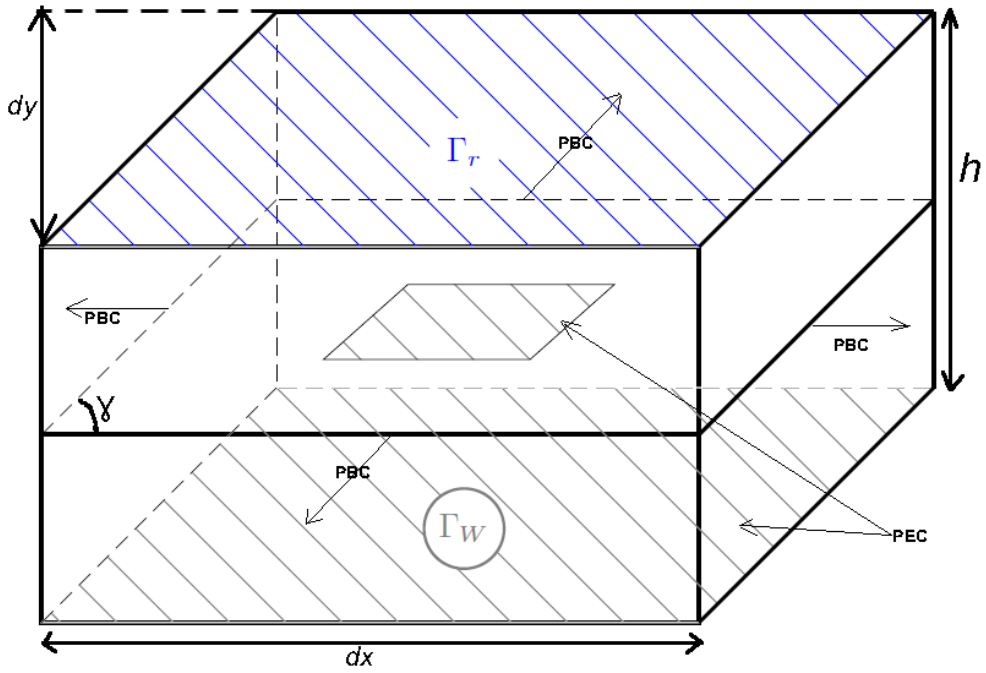


Fig 1.VI Cellule de base d'un réseau périodique de patchs métalliques excités par un mode guidé. La cellule de base est contenue dans le domaine Ω de hauteur h.

Sur la Fig 1.VI, on impose sur la face supérieure Γ_r une condition de radiation et la solution y est développée en série de Floquet pseudo périodique. La source vient d'un mode guidé introduit sur Γ_W, le reste de la face inférieure est métallisé. Rappel : PEC= Perfect Electric Conductor et PBC= Periodic Boundary Conditions. Les conditions de Floquet annulent les termes surfaciques sur les interfaces PBC Fig 1.VI où sont imposées les conditions limites *Maîtres-esclaves* (cf remarque III.11). Pour toute fonction test $\boldsymbol{\varphi}$ définie dans un espace approprié noté $\mathcal{H}$, on peut écrire la formulation variationnelle des équations de Maxwell dans la cellule de base Ω de la façon suivante (cf section III.2 équation (6.III)) :

$$\int_\Omega \boldsymbol{\nabla}\times\boldsymbol{\varphi}^*.\frac{\boldsymbol{\nabla}\times\mathbf{E}}{\mu_r}-k_0^2\int_\Omega \epsilon_r\mathbf{E}.\boldsymbol{\varphi}^*+jk_0Z_0\int_{\Gamma_r}(\nu\times\mathbf{H}).\boldsymbol{\varphi}^*+jk_0Z_0\int_{\Gamma_W}(\boldsymbol{\nu}\times\mathbf{H}).\boldsymbol{\varphi}^*=0. \quad (16.\text{VI})$$

Remarque VI.7. Les cas dispersifs peuvent se traiter facilement puisque pour chaque pulsation ω de l'onde incidente, on peut faire dépendre la permittivité et la perméabilité relative $\epsilon_r,\ \mu_r$ de ω dans l'équation (16.VI). On peut par exemple intégrer des termes de pertes par conductions dans l'expression de la permittivité comme nous l'avons vu dans le section VI.2 équations (4.VI) et (10.VI).

On rappelle, comme introduit en (8.III), que :

$$\mathbf{E} \simeq \tilde{\mathbf{E}} = \sum_{p=1}^{DL} e_p \mathbf{W}_p,\ e_p = \int_{\mathbf{Ar}_p} \mathbf{E} \tag{17.VI}$$

On note $\mathbf{X} \in \mathbb{C}^{DL} = (e_1, \ldots, e_{DL})$. Les fonctions tests étant choisies dans l'espace d'approximation $\mathcal{H}_h$, le problème variationnel (16.VI), peut s'écrire comme un système linéaire en appliquant la méthode Galerkin, comme dans l'équation (22.III) :

$$R(S^i + S^r + S^W)R^*\mathbf{X} = \mathbf{E}^{inc}$$

Où R est la matrice de transformation explicitée dans l'équation (26.III), elle est nécessaire pour appliquer les conditions de Floquet. $\mathbf{E}^{inc} = (E_1^{inc}, E_k^{inc}, \ldots),\ k \in [1, \ldots, DL]$ est le champ incident sur Γ_W représenté par un vecteur de dimension DL que nous explicitons plus tard équation (22.VI). S^i désigne les termes matriciels volumiques, S^r les termes matriciels surfaciques sur Γ_r et S^W les termes matriciels surfaciques sur Γ_W. Pour deux arêtes $(\mathbf{Ar}_k, \mathbf{Ar}_l)$, on peut déjà facilement expliciter les termes matriciels $(S^i)_{kl}$:

$$(S^i)_{kl} = \int_\Omega \boldsymbol{\nabla} \times \mathbf{W}_k . \frac{\boldsymbol{\nabla} \times \mathbf{W}_l}{\mu_r} - k_0^2 \int_\Omega \epsilon_r \mathbf{W}_k . \mathbf{W}_l$$

VI.2.2.3 Détail des termes sources sur Γ_W

Il est nécessaire d'avoir une représentation précise de la source sur Γ_W. C'est pourquoi nous choisissons d'introduire une excitation via un guide d'onde rectangulaire ou cylindrique. Dans un tel guide les solutions modales se calculent facilement de façon analytique[94]. Et donc, sur Γ_W, le champ électrique transversal $\mathbf{E}^t$ s'écrit comme combinaison linéaire de ces solutions modales[110] :

$$\mathbf{E}^t = e^{-\gamma_0 z}\mathbf{g}_0 + \sum_{i=0}^{+\infty} C_i e^{\gamma_i z}\mathbf{g}_i \tag{18.VI}$$

Où l'exposant t dénote le champ tangent à Γ_W et les $\mathbf{g}_i$ sont les vecteurs de la base modale orthonormale avec leurs constantes de propagation respectives γ_i. Par exemple, dans le cas d'un guide d'onde rectangulaire on a [94] :

$$\gamma_{mn} = \sqrt{(\frac{m\pi}{l})^2 + (\frac{n\pi}{L})^2 - k_0^2}$$

Où l et L désignent la largeur et la longueur du rectangle de la coupe transversale du guide. Le terme $e^{-\gamma_0 z}\mathbf{g}_0$ est le mode dominant qui représente le champ incident et C_0 le coefficient

de réflection de ce champs. On remarque alors que

$$C_i = \int_{\Gamma_W} \mathbf{E}^t.\mathbf{g}_i ds - \delta_{0i}.$$

Où δ_{0i} est le symbôle de Kronecker. En injectant le champ transversal $\mathbf{E}^t$ dans les équations de Maxwell et en notant Y_i les admittances modales, le champ $\mathbf{H}^t$ peut s'écrire :

$$\mathbf{H}^t = Y_0 e^{-\gamma_0 z}\boldsymbol{\nu} \times \mathbf{g}_0 - \sum_{i=0}^{+\infty} C_i Y_i e^{\gamma_i z}\boldsymbol{\nu} \times \mathbf{g}_i \tag{19.VI}$$

En évaluant $\mathbf{H}_t$ sur Γ_W $(z = 0)$ on en déduit que :

$$\mathbf{H}^t \times \boldsymbol{\nu}|_{\Gamma_W} = 2Y_0\mathbf{g}_0 - \sum_{i=0}^{+\infty} Y_i\mathbf{g}_i \int_{\Gamma_W} \mathbf{E}^t.\mathbf{g}_i \tag{20.VI}$$

Or $\mathbf{E}_t$ s'écrit aussi comme combinaison linéaire de fonction d'arêtes de Nedelec (cf équation (17.VI)), on en déduit :

$$C_i = \sum_{p=1}^{DL} \int_{\Gamma_W} \mathbf{W}_p.\mathbf{g}_i ds - \delta_{0i}$$

On peut alors expliciter les termes matriciels $(S^W)_{kl}$, où (k, l) correspondent aux indices des arêtes $(\mathbf{Ar}_k, \mathbf{Ar}_l)$:

$$(S^W)_{kl} = jk_0Z_0 \sum_{i=0}^{+\infty} Y_i \int_{\Gamma_W} \mathbf{W}_k.\mathbf{g}_i \int_{\Gamma_W} \mathbf{W}_l.\mathbf{g}_i,\ (\mathbf{Ar}_k, \mathbf{Ar}_l) \in \Gamma_r^2 \tag{21.VI}$$

Et le terme source $\mathbf{E}^{inc}$ s'écrit :

$$E_k^{inc} = 2jk_0Y_0 \int_{\Gamma_W} \mathbf{W}_k.\mathbf{g}_0,\ \mathbf{Ar}_k \in \Gamma_W \tag{22.VI}$$

VI.2.2.4 Conditions de radiation

Chaque patch du réseau infini est supposé être excité par un champ incident d'amplitude unitaire dans le guide d'ondes à partir de $z < 0$. Les déphasages entre les sources issues des guides d'ondes produisent un faisceau dirigé selon les angles θ_0 et ϕ_0 en coordonnées sphériques. Autrement dit le vecteur d'onde incident s'écrit :

$$\mathbf{k}_i = (k_{ix}, k_{iy}, k_{iz}) = (k_0 \sin\theta_0 \cos\phi_0, k_0 \sin\theta_0 \cos\phi_0, k_0 \cos\theta_0)$$

Sur Γ_r situé à la hauteur h le champ $\tilde{\mathbf{E}} \in \mathcal{H}_h$ est pseudo périodique et s'écrit comme somme infinie de modes de Floquet [42] dans les deux directions de périodicité. On peut alors comme dans la section II.1.5.1, exprimer le champ sur Γ_r lorsque les directions de périodicités ne sont pas confondues avec les directions des axes du repère cartésien comme

sur la figure Fig 1.VI :

$$\tilde{\mathbf{E}}(x,y,h) = \sum_{n=-\infty}^{+\infty}\sum_{m=-\infty}^{+\infty} \hat{\mathbf{E}}_{nm}(h) e^{j\left((G_{x_{nm}}-k_{ix})x+(G_{y_{nm}}-k_{iy})y)\right)} \tag{23.VI}$$

Avec :

$$\begin{cases} \hat{\mathbf{E}}_{nm}(h) = \sum_{p=1}^{DL} e_p \dfrac{1}{dxdy}\displaystyle\int_{\Gamma_r} \mathbf{W}_p(x,y,h) e^{j\left((-G_{x_{nm}}+k_{ix})x+(-G_{y_{nm}}+k_{iy})y)\right)} dxdy \\ \qquad\qquad = \sum_{p=1}^{DL} e_p \hat{\mathbf{E}}^p_{nm}(h) \\ \text{avec } \hat{\mathbf{E}}^p_{nm}(h) = \dfrac{1}{dxdy}\displaystyle\int_{\Gamma_r} \mathbf{W}_p(x,y,h) e^{j\left((-G_{x_{nm}}+k_{ix})x+(-G_{y_{nm}}+k_{iy})y)\right)} dxdy \end{cases} \tag{24.VI}$$

Avec les conventions définies sur la figure Fig 1.VI et comme dans le cas traité section II.1.5.1 $\mathbf{G}_{nm} = (\frac{2\pi.m}{dx}, \frac{2\pi.n}{dy} - \frac{2\pi.m\cot(\gamma)}{dx}, 0)$.

Par analogie avec l'argument développé pour la méthode RCWA (cf section II.3.3 remarque II.6), comme Γ_r est contenu dans un domaine situé dans l'air ($\epsilon_r = \epsilon_0$ et $\mu_r = \mu_0$), $\tilde{\mathbf{E}}|_{\Gamma_r}$ vérifie l'équation de Helmoltz :

$$\begin{cases} \triangle\Big(\displaystyle\sum_{n=-\infty}^{+\infty}\sum_{m=-\infty}^{+\infty} \hat{\mathbf{E}}_{nm}(z) e^{j\left((G_{x_{nm}}-k_{ix})x+(G_{y_{nm}}-k_{iy})y)\right)}\Big) + \\ k_0^2\Big(\displaystyle\sum_{n=-\infty}^{+\infty}\sum_{m=-\infty}^{+\infty} \hat{\mathbf{E}}_{nm}(z) e^{j\left((G_{x_{nm}}-k_{ix})x+(G_{y_{nm}}-k_{iy})y)\right)}\Big) = \mathbf{0} \end{cases} \tag{25.VI}$$

Où $k_0^2 = \omega^2\epsilon_0\mu_0$. Comme l'opérateur $\triangle$ est linéaire, on en déduit que chaque coefficient $\hat{\mathbf{E}}_{nm}(z)$ de la série de (25.VI) vérifie une équation de Helmoltz dont les solutions s'écrivent soit comme un terme incident ou "entrant" noté $\hat{\mathbf{E}}_{i_{nm}}(z)$, soit comme un terme réfléchi ou "sortant" noté $\hat{\mathbf{E}}_{r_{nm}}(z)$:

$$\begin{cases} \hat{\mathbf{E}}_{i_{nm}}(z) = \hat{\mathbf{E}}_{i_{nm}} e^{-j\gamma_{nm}z},\ \hat{\mathbf{E}}_{r_{nm}}(z) = \hat{\mathbf{E}}_{r_{nm}} e^{+j\gamma_{nm}z} \\ \text{avec } \gamma_{nm} = \sqrt{k_0^2 - (G_{x_{nm}} - k_{ix})^2 - (G_{y_{nm}} - k_{iy})^2}. \end{cases} \tag{26.VI}$$

Or on ne considère que les ondes qui arrivent sur la paroi Γ_r sont des ondes "sortantes" se propageant selon l'axe Oz car chaque patch du réseau infini est supposé être excité par un champ unitaire incident dans un guide d'ondes situé sous Γ_W (z<0). Le cas des ondes entrantes se traite de façon similaire et le champ $\tilde{\mathbf{E}}$ s'écrit comme une combinaison linéaire du champ entrant et sortant si la source est située au dessus de Γ_r. On en déduit que :

$$\hat{\mathbf{E}}_{nm}(z) = \hat{\mathbf{E}}_{r_{nm}} e^{+j\gamma_{nm}z} \tag{27.VI}$$

Le champ $\tilde{\mathbf{E}}$ peut alors s'écrire d'une autre façon encore sur Γ_r :

$$\tilde{\mathbf{E}}(x,y,z)|_{\Gamma_r} \simeq \sum_{n=-\infty}^{+\infty}\sum_{m=-\infty}^{+\infty} \hat{\mathbf{E}}_{r_{nm}} e^{j\left((G_{x_{nm}}-k_{ix})x+(G_{y_{nm}}-k_{iy})y+\gamma_{nm}z\right)} \tag{28.VI}$$

En remarquant que $\hat{\mathbf{E}}_{nm}(h)$ se décompose selon les vecteurs $\hat{\mathbf{E}}^p_{nm}(h),\ p \in [1,\dots DL]$ dans l'équation (24.VI) et en utilisant la relation (27.VI), on peut écrire :

$$\hat{\mathbf{E}}_{r_{nm}} = \hat{\mathbf{E}}_{nm}(h)e^{-j\gamma_{nm}h} \Rightarrow\ \hat{\mathbf{E}}_{r_{nm}} = \sum_{p=1}^{DL} e_p \hat{\mathbf{E}}^p_{r_{nm}} \text{ avec } \hat{\mathbf{E}}^p_{r_{nm}} = \hat{\mathbf{E}}^p_{nm}(h)e^{-j\gamma_{nm}h} \tag{29.VI}$$

En posant $\mathbf{Q}_{mn} = (G_{x_{nm}} - k_{ix}, G_{y_{nm}} - k_{iy}, \gamma_{nm})$ et en utilisant les équations de Maxwell reliant les champs $\mathbf{E}$ et $\mathbf{H}$ (1.III), on peut, à partir de l'expression de l'expression de $\tilde{\mathbf{E}}$ sur Γ_r (28.VI), exprimer le champ approché $\boldsymbol{\nu} \times \tilde{\mathbf{H}}$ sur Γ_r qui est contenu dans l'air :

$$\begin{cases} \boldsymbol{\nu} \times \tilde{\mathbf{H}}(x,y,z)|_{\Gamma_r} = \dfrac{1}{\mu_0\omega}\displaystyle\sum_{n=-\infty}^{+\infty}\sum_{m=-\infty}^{+\infty} \boldsymbol{\nu} \times (\mathbf{Q}_{mn} \times \hat{\mathbf{E}}_{r_{nm}}) e^{j\mathbf{Q}_{mn}.\mathbf{r}} = \\ \displaystyle\sum_{n=-\infty}^{+\infty}\sum_{m=-\infty}^{+\infty} -\gamma_{nm}\hat{\mathbf{E}}_{r_{nm}} e^{j\mathbf{Q}_{mn}.\mathbf{r}} \end{cases} \tag{30.VI}$$

Remarque VI.8. Les coordonnées selon z des fonctions de base $\mathbf{W}_p(h)$ peuvent être choisies nulles parce que les inconnues sont définies par la circulation du champ le long des arêtes de Γ_r qui est perpendiculaire à l'axe Oz (cf remarque III.7). On en déduit que $(\hat{\mathbf{E}}_{nm}(h))_z = 0$ et $\boldsymbol{\nu}.\hat{\mathbf{E}}_{r_{nm}} = 0$.

On peut alors expliciter les termes matriciels $(S^r)_{kl}$ en utilisant les équations (16.VI), (29.VI), (30.VI) et en choisissant (k,l) correspondant aux indices des arêtes $(\mathbf{Ar}_k, \mathbf{Ar}_l)$:

$$\begin{cases} (S^r)_{kl} = -j\displaystyle\sum_{n=-\infty}^{+\infty}\sum_{m=-\infty}^{+\infty} \gamma_{nm}\hat{\mathbf{E}}^k_{r_{nm}} \int_{\Gamma_r} \hat{\mathbf{E}}^l_{r_{nm}} e^{j\mathbf{Q}_{mn}.\mathbf{r}} = \\ -j\displaystyle\sum_{n=-\infty}^{+\infty}\sum_{m=-\infty}^{+\infty} \gamma_{nm}\hat{\mathbf{E}}^k_{nm}(h)\hat{\mathbf{E}}^l_{nm}(h)^* e^{j\left((G_{x_{nm}}-k_{ix})x+(G_{y_{nm}}-k_{iy})y\right)} \end{cases} \tag{31.VI}$$

Où $\hat{\mathbf{E}}^l_{nm}(h)^*$ est le conjugué de $\hat{\mathbf{E}}^l_{nm}(h)$ défini équation (24.VI).

L'article [42] présente une méthode de calcul similaire des termes sur Γ_r et Γ_W. L'existence et l'unicité des solutions pour cette catégorie de problème sont expliquées dans la thèse de Pedro Ferreira[77], pour des cas plus simples en dimension 2.

VI.2.3 Conclusion

Nous venons de décrire la possibilité d'introduire des modèles dispersifs dans la méthode des éléments finis, via une augmentation du nombre d'inconnues pour la résolution modes

propres dans la section VI.2 et l'introduction d'une source explicite dans un milieu ouvert bi-périodique section VI.2.2. L'introduction d'une source explicite permet de traiter les milieux dispersifs sans modification de la formulation variationnelle ce qui la rend plus efficace que la résolution de modes propres qui nécessite l'augmentation du nombre d'inconnues. Cependant, contrairement à cette résolution, l'introduction d'une source explicite ne permet pas de traiter les milieux périodiques dans trois directions non colinéaires de l'espace pour des raisons évidentes d'accessibilité des sources. Nous allons maintenant énoncer quelques perspectives.

VI.3 Perspectives

- Comme nous venons de l'énoncer, il est envisageable de traiter des structures dispersives en modifiant le problème variationnel (cf section VI.2) où en introduisant une source explicite (section VI.2.2).
- On peut optimiser le code d'une part par la montée en ordre (cf remarque III.4) et d'autre part par un choix des paramètres permettant de minimiser l'espace mémoire utilisé par l'algorithme IRAM (section IV.2.5).
- On peut envisager de coder l'élimination automatique des modes propres parasites pour les structures périodiques (cf remarque IV.13) une fois la montée en ordre effectuée.
- On peut envisager d'essayer d'adapter les outils développés aux métamatériaux phononiques et à certains problèmes de sismologie [9] dont la modélisation peut conduire à des problèmes de modes propres.
- Les algorithmes de factorisation des matrices creuses peuvent être optimisés afin de gagner en temps de calcul (cf remarque IV.16).
- Nous avons montré comment l'outil diagramme de bandes à l'aide d'une nouvelle quantité(1.V), peut être utilisé pour caractériser les modes de surface. Une perspective est maintenant considérée pour valider notre approche par une étude expérimentale.

Pour aider à clarifier certains passages nous avons regroupé des résultats classiques et une explication détaillée de certains calculs en annexe.

ANNEXE

A Compléments et outils : sommaire

A.1 Conditions d'interface entre deux milieux

Soit $\overline{\mathbb{R}^3} = \overline{\Omega_1} \cup \overline{\Omega_2}$ et $\Sigma = \partial\Omega_1 = \partial\Omega_2$

$$\epsilon(x) = \begin{cases} \epsilon_1(x) & \text{dans } \Omega_1 \\ \epsilon_2(x) & \text{dans } \Omega_2 \end{cases} \qquad \mu(x) = \begin{cases} \mu_1(x) & \text{dans } \Omega_1 \\ \mu_2(x) & \text{dans } \Omega_2 \end{cases}$$

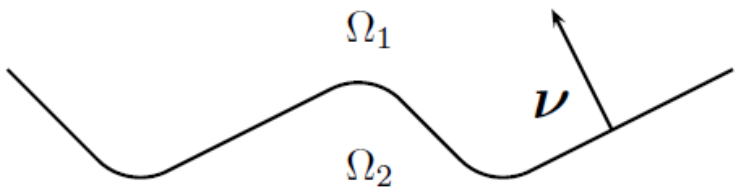


Fig 1.A $\boldsymbol{\nu}$ normale sortante (respectivement entrante) à $\partial\Omega_2$ (respectivement $\partial\Omega_1$) .

Soit $(\mathbf{E}, \mathbf{B})$ solutions des équations de Maxwell dans $\mathbb{R}_t \times \mathbb{R}^3$. On suppose que les champs solutions dans chaque région $(\mathbf{E}_p, \mathbf{H}_p) \in \mathcal{C}^1(\mathbb{R}_t^+ \times \Omega_j)^2 \quad p = 1, 2$.

Soit l'équation de Maxwell Faraday :

$$\boldsymbol{\nabla} \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = \mathbf{0}, \quad \text{dans } \mathbb{R}_t \times \mathbb{R}^3 \tag{1.A}$$

Soit $\boldsymbol{\varphi}$ une fonction à support compact dans $\mathbb{R}_t \times \mathbb{R}^3$ et à valeur dans $\mathbb{R}^3$. Après multiplication scalaire par $\boldsymbol{\varphi}$ et intégration de l'équation (1.A), on obtient :

$$0 = \int_{\mathbb{R}_t} \int_{\mathbb{R}^3} \frac{\partial \mathbf{B}}{\partial t}.\boldsymbol{\varphi} + \boldsymbol{\nabla} \times \mathbf{E}.\boldsymbol{\varphi}$$

D'une part $\boldsymbol{\nabla}.(\mathbf{E} \times \boldsymbol{\varphi}) = (\boldsymbol{\nabla} \times \mathbf{E}).\boldsymbol{\varphi} - (\boldsymbol{\nabla} \times \boldsymbol{\varphi}).\mathbf{E}$. D'autre part d'après la formule de Green-Ostrogradski, on a :

$$\int_{\Omega_1} \boldsymbol{\nabla}.\mathbf{E}_1 = -\int_{\partial\Omega_1} \mathbf{E}_1.\boldsymbol{\nu} \text{ et } \int_{\Omega_2} \boldsymbol{\nabla}.\mathbf{E}_2 = \int_{\partial\Omega_2} \mathbf{E}_2.\boldsymbol{\nu}$$

On obtient donc en séparant les domaines et en intégrant par partie :

$$\mathbf{0} = \int_{\mathcal{R}_t} \int_{\Omega_1} (\frac{\partial \mathbf{B}_1}{\partial t} + \boldsymbol{\nabla} \times \mathbf{E}_1).\boldsymbol{\varphi} + \int_{\mathcal{R}_t} \int_{\Omega_2} (\frac{\partial \mathbf{B}_2}{\partial t} + \boldsymbol{\nabla} \times \mathbf{E}_2).\boldsymbol{\varphi} + \int_{\mathcal{R}_t} \int_{\Sigma} \boldsymbol{\nu}.\mathbf{E}_1 \times \boldsymbol{\varphi} - \boldsymbol{\nu}.\mathbf{E}_2 \times \boldsymbol{\varphi}$$

Comme les équations de Maxwell sont vérifiées respectivement dans les milieux Ω_1 et Ω_2 par les champs $(\mathbf{E}_1, \mathbf{B}_1)$ et $(\mathbf{E}_2, \mathbf{B}_2)$ on en déduit que les deux premiers termes sont nuls et que

$$\mathbf{E}_1 \times \boldsymbol{\nu} = \mathbf{E}_2 \times \boldsymbol{\nu}. \tag{2.A}$$

Soit maintenant le théorème de Gauss et l'équation de conservation du champ magnétique :

$$\nabla.(\epsilon\mathbf{E}) = \rho, \qquad \nabla.(\mu\mathbf{H}) = 0$$

Soit φ une fonction test définie sur $\mathbb{R}_t \times \mathbb{R}^3$ à valeur dans $\mathbb{R}$. Un raisonnement analogue à ce qui précede nous conduit à :

$$0 = \int_{\mathbb{R}_t} \int_{\mathbb{R}^3} \varphi\rho - \boldsymbol{\nabla}.(\epsilon\mathbf{E})\varphi$$

Ceci donne en intégrant par parties (φ à support compacte dans Ω) :

$$0 = \int_{\mathbb{R}_t} \int_{\mathbb{R}^3} \varphi\rho + \epsilon\mathbf{E}.\nabla\varphi$$

Puis en séparant les domaines on à :

$$-\int_{\mathbb{R}_t} \int_{\Omega_1 \cup \Omega_2} \varphi\rho = \int_{\mathbb{R}_t} \int_{\Omega_1} \epsilon_1\mathbf{E}_1.\nabla\varphi + \int_{\mathcal{R}_t} \int_{\Omega_2} \epsilon_2\mathbf{E}_2.\nabla\varphi$$

En appliquant la formule de Green-Ostrogradski et le théorème de Gauss dans chaque domaine respectif, on obtient alors :

$$\int_{\mathbb{R}_t} \int_{\Sigma} \varphi(\epsilon_1\mathbf{E}_1.\boldsymbol{\nu} - \epsilon_2\mathbf{E}_2.\boldsymbol{\nu}) = 0$$

Soit

$$\epsilon_1\mathbf{E}_1.\boldsymbol{\nu} = \epsilon_2\mathbf{E}_2.\boldsymbol{\nu}$$

De même la conservation du flux magnétique nous donne :

$$\mu_1\mathbf{H}_1.\boldsymbol{\nu} = \mu_2\mathbf{H}_2.\boldsymbol{\nu}$$

Récapitulatif des relations établies[38] :

$$\begin{aligned} \epsilon_1\mathbf{E}_1.\boldsymbol{\nu} &= \epsilon_2\mathbf{E}_2.\boldsymbol{\nu} \\ \mu_1\mathbf{H_1}.\boldsymbol{\nu} &= \mu_2\mathbf{H_2}.\boldsymbol{\nu} \end{aligned} \tag{3.A}$$

A.2 Comportement face à une excitation électrique ou magnétique

A.2.1 Permittivité et polarisation d'un diélectrique

Un matériau est qualifié de diélectrique s'il ne contient pas de charges électriques pouvant se déplacer de façon macroscopique. Autrement dit, c'est un milieu dans lequel ne se propage pas de courant électrique. Parmi ces milieux on compte : le verre, le vide et de nombreux plastiques. Bien qu'ils soient non conducteurs, les diélectriques peuvent interagir avec un champ électrique. En effet on peut considérer que les atomes du matériau présente des dipôles électrostatiques susceptibles de s'aligner selon les lignes de champs électriques. La densité de présence de ces dipôles s'appelle la polarisation. Les causes de la présence d'effets de polarisation peuvent être diverses :

* Le déplacement et la déformation des nuages électroniques des atomes.
* Les déplacements d'atomes ou bien d'ions au sein du matériau.

La grandeur la plus utilisée pour caractériser un diélectrique est la permittivité relative notée ϵ_r. On peut la définir comme étant le rapport entre la capacité C_x de deux électrodes plongées dans le diélectrique et la capacité C_v des mêmes électrodes dans le vide :

$$\epsilon_r = \frac{C_x}{C_v} \tag{4.A}$$

ϵ_r est donc un nombre réel sans dimension contrairement à la permittivité absolue ϵ qui est le produit de la constante diélectrique du vide ϵ_0 par la permittivité relative :

$$\epsilon = \epsilon_0 \epsilon_r$$

avec :

$$\epsilon_0 = 8,854187.10^{-12} \quad \text{F.m}^{-1}$$

On a mesuré que la permittivité relative des gaz est très légèrement supérieure 1 (1,00053 pour l'air) tandis que certains diélectriques pour condensateurs peuvent avoir une permittivité relative supérieure à 1000. En général les diélectriques ont une permittivité relative n'excédant pas la dizaine.

A.2.2 Causalité de la polarisation

Dans une modélisation réaliste, le temps intervient dans les mécanismes de polarisation. En effet un matériau ne peut pas réagir instantanément sous l'effet d'un champ électrique. Si on note $\mathbf{P}(t)$ la polarisation d'un matériau à l'instant t on a :

$$\mathbf{P}(t) = \epsilon_0 \int_{-\infty}^{t} \chi(t - t')\mathbf{E}(t')\, dt'. \tag{5.A}$$

Autrement dit, la polarisation à l'instant t est le produit de convolution entre le champ électrique aux instants passés et la susceptibilité électrique dépendante du temps $\chi(\Delta t)$. Le principe de causalité nous dit que la polarisation ne peut que dépendre du champ électrique passé. Une conséquence du principe de causalité est que la limite supérieure de cette intégrale peut être étendue à l'infini à moins que $\chi(\Delta t) = 0$ si $\Delta t < 0$. Pour une réponse instantanée on écrit la susceptibilité électrique avec la fonction de Dirac :

$$\chi(\Delta t) = \chi * \delta(\Delta t)$$

D'autre part la transformée de Fourier d'un produit de convolution est le produit des transformées de Fourier des fonctions respectives. On peut donc exprimer la polarisation en fonction de la pulsation :

$$\mathbf{P}(\omega) = \epsilon_0 \chi(\omega) \mathbf{E}(\omega). \tag{6.A}$$

La courbe de la susceptibilité en fonction de la pulsation caractérise les propriétés de dispersion du matériau. De plus, le fait que la polarisation ne puisse que dépendre du champ électrique à moment antérieur (i.e $\chi(\Delta t) = 0$ si $\Delta t < 0$), impose les contraintes de Kramers-Kronig sur la susceptibilité $\chi(0)$ [111].

A.2.3 Permittivité électrique et champ d'induction

Le champ d'induction électrique noté $\mathbf{D}$ représente l'influence du champ électrique $\mathbf{E}$ sur l'organisation des charges électriques d'un matériau. Soit $\mathbf{P}$ la polarisation et χ la susceptibilité électrique. Le champ d'induction électrique est par définition :

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P} = \epsilon_0 \mathbf{E} + \chi \epsilon_0 \mathbf{E} = \epsilon \mathbf{E} \tag{7.A}$$

L'effet d'induction vient donc de la polarisation du vide et du matériau étudié. Ainsi, pour une pulsation donnée, d'après l'équation (6.A), la susceptibilité électrique est reliée à la permittivité électrique d'un matériaux par la relation :

$$\epsilon(\omega) = \big(1 + \chi(\omega)\big)\epsilon_0 = \epsilon_r(\omega)\epsilon_0$$

* Si le matériau est anisotrope, la permittivité est une matrice (ϵ). Dans ce cas, le champ de vecteur $\mathbf{D}$ n'est pas colinéaire $\mathbf{E}$.
* Si le matériau est in-homogène, la valeur de (ϵ) n'est pas constante dans le matériau.
* Si le matériau est non linéaire, la relation $\mathbf{D} = \epsilon.\mathbf{E}$ n'est plus valable. Cependant, d'après l'équation (5.A) on peut écrire la relation (7.A) sous une forme plus générale :

$$\mathbf{D}(\mathbf{r}, t) = \epsilon_0 \big[\mathbf{E}(\mathbf{r}, t) + \int_{-\infty}^{t} \chi(t - t') \mathbf{E}(\mathbf{r}, t')\, dt'\big].$$

A.2.4 Permittivité complexe

Dans une modélisation réaliste en régime harmonique, la permittivité est complexe. En effet, la partie imaginaire permet de modéliser un phénomène d'absorption ou d'émission du champ électromagnétique. Par opposition à la réponse du vide, la réponse d'un matériau normal à un champ électrique extérieur dépend généralement de la fréquence de ce champ. Cette dépendance en fréquence reflète le fait que la polarisation d'un matériau ne répond pas de manière instantanée à un champ appliqué. La réponse peut arriver après que le champ a été appliqué ce qui peut être représenté par un déphasage. Pour cette raison la permittivité contient une partie imaginaire non nulle.

La contribution de la partie réelle de la permittivité est liée à l'énergie stockée dans le matériau.

$$\varepsilon(\omega) = \varepsilon'(\omega) - i\varepsilon''(\omega)$$

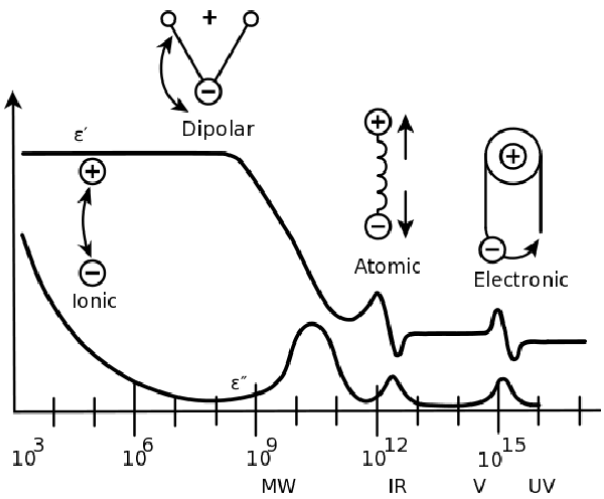


Fig 2.A Réponse du diélectrique en fonction de la fréquence $\omega/2\pi$[112].

Fig 2.A $\varepsilon^{'}$ et $\varepsilon^{''}$ représentent respectivement la partie réelle et la partie imaginaire de la permittivité, en fonction de la fréquence. Plusieurs phénomènes sont représentés Fig 2.A :

Réponse ionique et dipolaire pour les fréquences basses.

Réponse atomique et électronique pour les fréquences plus fortes.

ε'' représente ce qu'on appelle les pertes diélectriques. Ces pertes sont souvent très faibles. La partie imaginaire est donc très petite devant la partie réelle. On parle alors parfois d'angle de perte, exprimé en pourcentage et défini par :

$$\delta_e \approx \tan\delta_e = \frac{\varepsilon''}{\varepsilon'}$$

Cette appellation s'explique par le fait que cet angle δ_e est l'angle formé par les vecteurs champ électrique et déplacement électrique dans le plan complexe. Les parties réelles et imaginaires de la permittivité ne sont pas complètement indépendantes. Elles sont reliées par les relations de Kramers-Kronig [111].

A.2.5 Relaxation de Debye

La relaxation de Debye [57] est une réponse de la relaxation d'un ensemble de dipôles idéaux qui n'interagissent pas entre eux, à un champ électrique alternatif extérieur. Elle s'exprime comme une fonction de la fréquence ω du champ extérieur à travers la permittivité complexe $\hat{\varepsilon}(\omega)$ d'un matériau :

$$\hat{\varepsilon}(\omega) = \varepsilon_\infty + \frac{\Delta\varepsilon}{1 + i\omega\tau},$$

Où ε_∞ est la permittivité limite à quand la fréquence tend vers l'infinie, $\Delta\varepsilon = \varepsilon_s - \varepsilon_\infty$ où ε_s est la permittivité statique à basse fréquence et τ est la constante de temps caractéristique de relaxation du diélectrique (délai entre le champ extérieur et l'établissement d'un état d'équilibre dans le milieu).

A.3 Éléments finis d'arêtes de Nedelec

A.3.1 Définition locale et propriétés

Nous allons présenter les fonctions de base qui vont approcher le champ solution dans un tétraèdre :

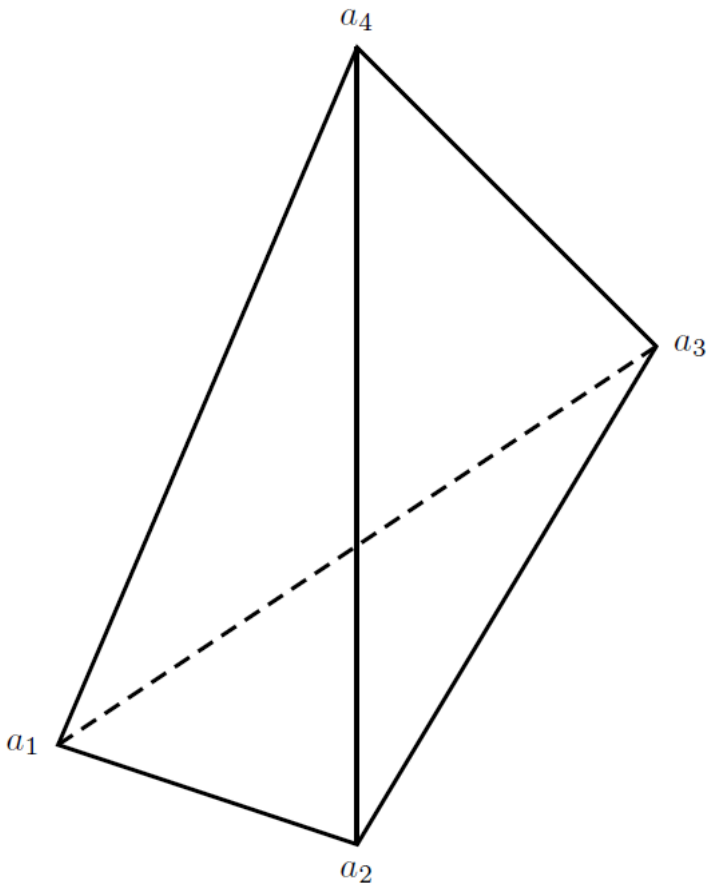


Fig 3.A Tétraèdre.

Le tétraèdre illustré Fig 3.A est constitué des quatre sommets $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4$. On définit alors les fonctions de base $\mathbf{W}_{ij}\ i, j \in [1 \dots 4],\ i \neq j,\ i < j$ de Nedelec d'ordre 0 par[70] :

$$\mathbf{W}_{ij} = (\lambda_i \nabla \lambda_j - \lambda_j \nabla \lambda_i). \tag{8.A}$$

Où λ_i représente la coordonnée barycentrique du point $\mathbf{a}_i$ par rapport aux points $\mathbf{a}_j = (a_{1j}, a_{2j}, a_{3j})\ j \in [1 \dots 4],\ i \neq j$ [69] (page 82). On peut exprimer les coordonnées barycentriques à l'aide du déterminant par exemple :

$$\lambda_4(x, y, z) = \frac{1}{6vol(K)} \begin{vmatrix} a_{11} & a_{12} & a_{13} & x \\ a_{21} & a_{22} & a_{23} & y \\ a_{31} & a_{32} & a_{33} & z \\ 1 & 1 & 1 & 1 \end{vmatrix}$$

Plus précisément on a :

$$\lambda_4(\mathbf{X}) = A_4x + B_4y + C_4z + D_4 = \nabla\lambda_4.\mathbf{X} + \lambda_4(0)\ (\tilde{P})$$

Avec :

$$A_4 = \frac{1}{6vol(K)}\begin{vmatrix} a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ 1 & 1 & 1 \end{vmatrix}, \quad B_4 = \frac{1}{6vol(K)}\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \\ 1 & 1 & 1 \end{vmatrix}$$

$$C_4 = \frac{1}{6vol(K)}\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 1 & 1 & 1 \end{vmatrix}, \quad D_4 = \lambda_4(0) = \frac{1}{6vol(K)}\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

Remarque A.1. Soit l_{ij} la longueur du segment $[\mathbf{a}_i, \mathbf{a}_j]$ $i, j \in [1 \dots 4],\ i \neq j,\ i < j$. On peut normaliser les fonctions de base en multipliant $\mathbf{W}_{ij}$ par l_{ij}. Cette transformation est utilisée pour changer l'unité du champ solution $\tilde{\mathbf{E}}$ qui sera alors homogène à des volts au lieu d'être homogène à des volts par mètres.

En posant $\mathbf{X} = (x, y, z)$ et :

$$\lambda_i(\mathbf{X}) = A_ix + B_iy + C_iz + D_i = \nabla\lambda_i.\mathbf{X} + \lambda_i(0),$$

on a :

$$\mathbf{W}_{ij}(\mathbf{X}) = \begin{pmatrix} (A_jB_i - A_iB_j)y + (A_jC_i - A_iC_j)z + (A_jD_i - A_iD_j) \\ (B_jA_i - B_iA_j)x + (B_jC_i - B_iC_j)z + (B_jD_i - B_iD_j) \\ (C_jA_i - C_iA_j)x + (C_jB_i - C_iB_j)y + (C_jD_i - C_iD_j). \end{pmatrix}$$

On en déduit que $\mathbf{W}_{ij}$ peut s'écrire sous la forme : $\mathbf{N}_{ij} \times \mathbf{X} + \mathbf{B}$. Où :

$$\mathbf{N}_{ij} = \begin{pmatrix} (B_iC_j - B_jC_i) \\ (A_jC_i - A_iC_j) \\ (B_jA_i - B_iA_j) \end{pmatrix} = \boldsymbol{\nabla}\lambda_i \times \boldsymbol{\nabla}\lambda_j, \text{ Et } \mathbf{B} = \begin{pmatrix} (A_jD_i - A_iD_j) \\ (B_jD_i - B_iD_j) \\ (C_jD_i - C_iD_j) \end{pmatrix}.$$

En appliquant les formules :

$$\boldsymbol{\nabla}.(\mathbf{u} \times \mathbf{v}) = -\mathbf{u}.(\boldsymbol{\nabla} \times \mathbf{v}) + \mathbf{u}.(\boldsymbol{\nabla} \times \mathbf{v})$$

$$\boldsymbol{\nabla} \times (\mathbf{u} \times \mathbf{v}) = \mathbf{u}.(\boldsymbol{\nabla}.\mathbf{v}) - (\mathbf{u} \cdot \boldsymbol{\nabla})\mathbf{v} - \mathbf{v}.(\boldsymbol{\nabla}.\mathbf{u}) + (\mathbf{v}.\boldsymbol{\nabla})\mathbf{u},$$

on obtient les propriétés suivantes :

$$\nabla \times \mathbf{W}_{ij} = 2\mathbf{N}_{ij} \quad \text{et} \quad \nabla.\mathbf{W}_{ij} = 0. \tag{9.A}$$

D'après [113] l'espace $\mathcal{H}_h$ engendré par les $\mathbf{W}_p,\ p \in [1 \dots DL]$ est dense dans $H(curl, \Omega)$

quand la longueur des arêtes tend vers 0. Rappel :

$$H(curl, \Omega) = \{\mathbf{u} \in L^2(\Omega)^3,\ \boldsymbol{\nabla} \times \mathbf{u}\ \in L^2(\Omega)^3\}.$$

A.4 Calcul des matrices élémentaires

Nous allons maintenant calculer de façon explicite les matrices élémentaires construites avec les fonctions de bases de Nedelec et les coordonnées barycentriques. Les fonctions de bases sont calculées localement dans le tétraèdre. Ces résultats utilisent les formules classiques d'intégration de Gauss [114].

A.4.1 Intégration dans le tétraèdre

On se place dans un tétraèdre K. Dans un premier temps, on cherche les valeurs $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ tel que la formule :

$$\int_K P(x,y,z)dxdydz = \alpha_1 P(\mathbf{a}_1) + \alpha_2 P(\mathbf{a}_2) + \alpha_3 P(\mathbf{a}_3) + \alpha_4 P(\mathbf{a}_4)$$

soit exacte pour les fonctions $P_1 = 1,\ P_2 = x,\ P_3 = y,\ P_4 = z$.

En remplaçant P par les $P_i,\ i \in [1,2,3,4]$ on obtient un système linéaire 4×4 qui a pour solutions :

$$\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = \frac{vol(K)}{4}.$$

On en déduit alors les termes volumiques de la matrice A ($\mu_r = 1$) définie équation (7.III) :

$$\int_K (\boldsymbol{\nabla} \times \mathbf{W}_{ij}).(\boldsymbol{\nabla} \times \mathbf{W}_{kl}) = 4vol(K)\mathbf{N}_{ij}.\mathbf{N}_{kl}$$

Ensuite, pour calculer le terme volumique de la matrice M ($\epsilon_r = 1$) définie équation (7.III) :

$$\int_K \mathbf{W}_{ij}.\mathbf{W}_{kl} =$$

$$\nabla\lambda_j.\nabla\lambda_l \int_K \lambda_i\lambda_k - \nabla\lambda_j.\nabla\lambda_k \int_K \lambda_i\lambda_l - \nabla\lambda_i.\nabla\lambda_l \int_K \lambda_j\lambda_k + \nabla\lambda_i.\nabla\lambda_k \int_K \lambda_j\lambda_l$$

Il faut intégrer des polynômes de degré 2 en utilisant la formule :

$$\int_K \lambda_1^{n_1}\lambda_2^{n_2}\lambda_3^{n_3}\lambda_4^{n_4} = 6vol(K)\frac{n_1!n_2!n_3!n_4!}{(n_1+n_2+n_3+n_4+3)!}.$$

En particulier on a :

$$\int_K \lambda_i\lambda_k = \frac{vol(K)}{20}\ (i \neq k) \text{ et } \int_K \lambda_i^2 = \frac{vol(K)}{10}.$$

Idée de la preuve

Il faut ajouter les points milieux des arêtes pour que la formule d'intégration de Gauss soit exacte pour les polinômes de degrés 2 de trois variables. Autrement dit on doit vérifier :

$$\int_K P(x,y,z)dxdydz = \alpha_1 P(\mathbf{a}_1) + \alpha_2 P(\bar{\mathbf{a}}_1) + \alpha_3 P(\mathbf{a}_2) + \alpha_4 P(\bar{\mathbf{a}}_2) + \alpha_5 P(\mathbf{a}_3) + \alpha_6 P(\bar{\mathbf{a}}_3)$$

$$+\alpha_7 P(\mathbf{a}_4) + \alpha_8 P(\bar{\mathbf{a}}_4) + \alpha_9 P(\bar{a}_5) + \alpha_{10} P(\bar{a}_6)$$

Où $\bar{\mathbf{a}}_1, \bar{\mathbf{a}}_2, \bar{\mathbf{a}}_3, \bar{\mathbf{a}}_4, \bar{a}_5, \bar{a}_6$ sont les milieux respectifs des segments :

$$[\mathbf{a}_1, \mathbf{a}_2], [\mathbf{a}_2, \mathbf{a}_3], [\mathbf{a}_3, \mathbf{a}_1], [\mathbf{a}_1, \mathbf{a}_4], [\mathbf{a}_2, \mathbf{a}_4], [\mathbf{a}_3, \mathbf{a}_4]$$

On calcul alors les valeurs $\alpha_i,\ i \in [1 \dots 10]$ en remplaçant P par :

$$X, Y, Z, X^2, Y^2, Z^2, XY, YZ, XZ$$

On obtient alors un systèmpe 10×10 dont les solutions sont :

$$\alpha_1 = \alpha_3 = \alpha_5 = \alpha_7 = \frac{-Vol(K)}{20},\ \alpha_2 = \alpha_4 = \alpha_6 = \alpha_8 = \alpha_9 = \alpha_{10} = \frac{Vol(K)}{5}$$

Testons par exemple ce résultat avec la fonction λ_1^2. Comme les coordonnées barycentriques sont des applications affines on obtient :

$$\lambda_1^2(\bar{\mathbf{a}}_1) = \lambda_1^2(\bar{\mathbf{a}}_3) = \lambda_1^2(\bar{\mathbf{a}}_4) = \frac{1}{4},\ \lambda_1^2(\bar{\mathbf{a}}_2) = \lambda_1^2(\bar{a}_5) = \lambda_1^2(\bar{a}_6) = 0,\ \lambda_1^2(a_i) = \delta_{1i}$$

Donc :

$$\int_K \lambda_1^2 = \frac{Vol(K)}{10}.$$

Donc finalement on obtient si $i \neq k$ et $j \neq l$:

$$\int_K \mathbf{W}_{ij}.\mathbf{W}_{kl} = \frac{vol(K)}{20}\big(\nabla\lambda_j.\nabla\lambda_l - \nabla\lambda_j.\nabla\lambda_k - \nabla\lambda_i.\nabla\lambda_l + \nabla\lambda_i.\nabla\lambda_k\big).$$

Si $i = k$ et $j = l$:

$$\int_K \mathbf{W}_{ij}.\mathbf{W}_{kl} = \int_K \|\mathbf{W}_{ij}\|^2 = \frac{vol(K)}{10}(\|\nabla\lambda_j\|^2 + \|\nabla\lambda_i\|^2) - 2\nabla\lambda_j.\nabla\lambda_i\frac{vol(K)}{20}.$$

Si $i = k$ et $j \neq l$:

$$\int_K \mathbf{W}_{ij}.\mathbf{W}_{kl} = \frac{vol(K)}{20}\big(-\nabla\lambda_j.\nabla\lambda_i - \nabla\lambda_i.\nabla\lambda_l + \|\nabla\lambda_i\|^2\big) + \frac{vol(K)}{10}\nabla\lambda_j.\nabla\lambda_l.$$

Si $i \neq k$ et $j = l$:

$$\int_K \mathbf{W}_{ij}.\mathbf{W}_{kl} = \frac{vol(K)}{20}\big(-\nabla\lambda_j.\nabla\lambda_k - \nabla\lambda_i.\nabla\lambda_j + \|\nabla\lambda_j\|^2\big) + \frac{vol(K)}{10}\nabla\lambda_i.\nabla\lambda_k.$$

A.5 Énoncé locale de la loi d'Ohm

Rappelons d'abord ce qu'est la mobilité des porteurs de charge. Soit une charge q soumise à un champ électrique $\mathbf{E}$. La relation fondamentale de la dynamique nous permet d'affirmer que :

$$m\frac{d\mathbf{v}}{dt} = q\mathbf{E}$$

d'où

$$\mathbf{v} = \frac{q\mathbf{E}}{m}t + \mathbf{v}_0$$

Par conséquent la vitesse diverge et donc le vecteur densité de courant aussi car $\mathbf{j} = qn < \mathbf{v} >$ et donc l'intensité diverge, ce qui est impossible. De plus, les interactions avec le milieu contraignent la norme de la vitesse à tendre vers une valeur limite finie. Aussi, $\mathbf{v}$ est colinéaire à $\mathbf{E}$ (en l'absence d'induction magnétique $\mathbf{B}$), et donc on a :

$$\mathbf{v} = \gamma\mathbf{E}$$

γ est est appelée la mobilité de la charge q, elle dépend du matériau, du porteur de charge, de la température, de $||\mathbf{E}||$... En posant $\sigma_e = qn\gamma$ on en déduit que :

$$\mathbf{J} = qn\mathbf{v} = qn\gamma\mathbf{E} \Rightarrow \mathbf{J} = \sigma_e\mathbf{E}$$

A.6 Algorithmes de calcul des valeurs propres

A.6.1 Introduction

Nous énonçons dans cette annexe les algorithmes classiques de calcul des valeurs propres pour permettre une meilleure compréhension de l'algorithme d'Arnoldi décrit dans la section IV.2.

A.6.2 Méthode de la puissance itérée

La méthode de la puissance itérée [81] est une méthode itérative de calcul de la valeur propre de module maximal d'une matrice et du vecteur propre associé. Soit A une matrice complexe (n, n). On suppose d'abord que A est diagonalisable, on note $\lambda_i,\ (i = 1, \ldots, n)$ ses valeurs propres et $\mathbf{u}_i, (i = 1, \ldots, n)$ ses vecteurs propres associés. Les valeurs propres seront supposées ordonnées :

$$|\lambda_1| > |\lambda_2| > \ldots |\lambda_n|$$

Soit $\mathbf{v}_0$ un vecteur quelconque de $\mathbb{C}^n$. $\mathbf{v}_0$ se décompose dans la base des vecteurs propres sous la forme :

$$\mathbf{v}_0 = \sum_{i=1}^{n} \alpha_i \mathbf{u}_i, \quad \alpha_i \in \mathbb{R}$$

Soit

$$\mathbf{w}_k = A^k \mathbf{v}_0 = \sum_{i=1}^{n} \alpha_i \lambda_i^k \mathbf{u}_i$$

Ce qui peut s'écrire :

$$\mathbf{w}_k = \lambda_1^k (\alpha_1 \mathbf{u}_1 + \sum_{i=2}^{n} \alpha_i (\frac{\lambda_i}{\lambda_1})^k \mathbf{u}_i) \tag{10.A}$$

Comme $|\frac{\lambda_i}{\lambda_1}| < 1$ par hypothèse, le vecteur $\mathbf{w}_k$ converge, lorsque $k \to \infty$, vers un vecteur de même direction que $\mathbf{u}_1$, donc vers un vecteur propre associé à la valeur propre de module maximal. À chaque étape du calcul, on normalise les vecteurs w_k par leur norme infinie. On a alors le processus itératif de la méthode de la puissance itérée :

$$\begin{cases} \mathbf{v}_0 \in \ \mathbb{C}^n \text{ donné} \\ \mathbf{w}_{k+1} = A\mathbf{v}_k, \\ \mathbf{v}_{k+1} = \frac{\mathbf{w}_{k+1}}{||\mathbf{w}_{k+1}||_\infty} \end{cases} \tag{11.A}$$

On obtient à la limite en utilisant (10.A) :

$$\begin{cases} ||\mathbf{w}_k||_\infty \to \ |\lambda_1| \text{ quand } \ k \to \infty \\ \mathbf{v}_k \to \ \mathrm{u}_1 \text{ quand } \ k \to \infty \end{cases}$$

On peut le voir en remarquant que pour k suffisamment grand :

$$||\mathbf{w}_{k+1}||_\infty = \frac{||A\mathbf{w}_k||_\infty}{||\mathbf{w}_k||_\infty} \simeq \frac{\lambda_1^{k+1}\alpha_1||\mathbf{u}_1||_\infty}{\lambda_1^k\alpha_1||\mathbf{u}_1||_\infty}$$

Remarque A.2. Le cas général quand A est trigonalisable se démontre en écrivant $\mathbf{v}_0$ dans la base formée des vecteurs des sous espaces caractéristiques de A associés aux valeurs propres respectives λ_i avec un ordre de multiplicité égale à μ_i :

$$\mathbf{v}_0 = \sum_{i=1}^{n}\sum_{k=1}^{\mu_i}\alpha_{ik}\mathbf{u}_{ik}, \quad \alpha_{ik} \in \mathbb{C}$$

Soit λ_1 la plus grande valeur propre de A. Comme les sous espaces caractéristiques sont stables par A et en utilisant la réduction de Jordan de A on peut voir que, en factorisant par λ_1^n, $A^n\mathbf{v}_0$ prendra la direction d'un vecteur $\mathbf{U}$ qui est une combinaison linéaire des vecteurs du sous espace caractéristique associé à λ_1. On aura pour k suffisamment grand, en appliquant l'algorithme ci-dessus :

$$||\mathbf{w}_{k+1}||_\infty = \frac{||A\mathbf{w}_k||_\infty}{||\mathbf{w}_k||_\infty} \simeq \frac{\lambda_1^{k+1}||\mathbf{U}||_\infty}{\lambda_1^k||\mathbf{U}||_\infty} \simeq \lambda_1$$

Et comme $A\mathbf{U}$ à même direction que $\mathbf{U}$ (puisque $\mathbf{U}$ est une direction limite), on en déduit que $\frac{\mathbf{U}}{\|\mathbf{U}\|}$ est un vecteur propre de A.

Remarque A.3. Si A est hermitienne, on peut continuer à chercher la deuxième plus grande valeur propre en calculant la plus grande valeur propre de la matrice A_1 définie par :

$$A_1 = A - \sigma\mathbf{u}_1 \otimes \mathbf{v}^*$$

Avec $\mathbf{v}$ choisit tel que $\mathbf{v}\mathbf{u}_1^* = 1$. En effet, on vérifie que les valeurs propres de A_1 sont :

$$(\lambda_1 - \sigma, \lambda_2, \ldots, \lambda_n)$$

Où les $\lambda_i \;\; i \in [1 \ldots n]$ sont les valeurs propres de A. Une méthode classique consiste à choisir $\sigma = \lambda_1$ et $\mathbf{v} = \mathbf{u}_1$ pour calculer la deuxième plus grande valeur propre λ_2 et son vecteur propre $\mathbf{u}_2$ et ainsi de suite pour les vecteurs suivants ce qui permet de calculer les couples (valeurs-propres, vecteurs propres) dans l'ordre décroissant. Plus précisément on calcul successivement les matrices de déflation :

$$A_{j+1} = A_j - \sigma_i\mathbf{u}_i \otimes \mathbf{u}_i$$

Ce procéder s'appelle la déflation de Wiedlandt [115].

A.6.3 Puissance itérée inverse

Si on utilise ce même algorithme pour chercher les valeurs propres de la matrice $(A-\sigma I)^{-1}$ au lieu de la matrice A ou σ est appelé *shift*. Alors, l'algorithme de la puissance itéré va

converger vers valeur propre la plus proche de σ au lieu de converger vers la valeur propre de module maximale.
Tout d'abord remarquons que si A est diagonalisable alors A et $(A-\sigma I)^{-1}$ ont les mêmes vecteurs propres. En effet si A est diagonalisable et P est la matrice formée des vecteurs propres de A alors :

$$A = PDP^{-1} \Rightarrow A - \sigma I = P(D-\sigma I)P^{-1} \Rightarrow (A-\sigma I)^{-1} = P(D-\sigma I)^{-1}P^{-1},$$

et les valeurs propres de $(A-\sigma I)^{-1}$ sont $(\lambda_j - \sigma)^{-1}$ où $(\lambda_j)_{j\in[1\dots n]}$ sont les valeurs propres de A. Ainsi l'algorithme de la puissance itéré appliqué à $(A-\sigma I)^{-1}$ converge vers la plus grande valeur propre de $(A-\sigma I)^{-1}$ donc le nombre σ tel que $|\lambda_j - \sigma|$ soit le plus petit possible. L'algorithme de la puissance itérée inverse s'écrit :

$$\begin{cases} \mathbf{v}_0 \in \mathbb{R} \text{ donné} \\ \text{À chaque itération on cherche } \mathbf{w}_{k+1} \text{ tel que } (A-\sigma I)\mathbf{w}_{k+1} = \mathbf{v}_k, \\ \mathbf{v}_{k+1} = \frac{\mathbf{w}_{k+1}}{||\mathbf{w}_{k+1}||_\infty} \end{cases}$$

A.6.4 Calcul des valeurs propres par l'algorithme QR

L'algorithme QR de recherche de valeurs propres se construit à partir de la décomposition QR[72] qui utilise le procédé de Gramm-Smidt. Soit A une matrice carrée de dimension n à coefficients dans $\mathbb{C}$. Soit la décomposition QR de $A = A_0 = Q_0R_0$. On construit une suite de matrice $(A_k)_{k\geq 0}$ de la façon suivante[82] :

$$\begin{cases} A_0 = Q_0R_0 \\ A_1 = R_0Q_0 = Q_1R_1 \\ \vdots \\ A_k = Q_kR_k \\ A_{k+1} = R_kQ_k = Q_{k+1}R_{k+1} \end{cases} \tag{12.A}$$

Proprieté A.1. Toutes les matrices A_k définies équation (12.A) sont orthogonalement semblables à A_0. De plus, on a $A_{k+1} = (Q_k^* \dots Q_0^*)A(Q_0 \dots Q_k)$ et $A^k = (Q_1 \dots Q_k)(R_k \dots R_1)$. Par conséquent les matrices A_k construites par l'algorithme (12.A) ont toutes les mêmes valeurs propres.

Démonstration. A_0 est semblable à elle même et :

$$A_1 = R_0Q_0 = Q_0^*Q_0R_0Q_0 = Q_0^*A_0Q_0$$

On suppose la propriété vrai au rang k. D'après (12.A) on a :

$$A_{k+1} = R_kQ_k = Q_k^*Q_kR_kQ_k = Q_k^*A_kQ_k.$$

Et on a démontré les deux premières assertions.
On suppose ensuite que $A^k = (Q_0 \dots Q_k)(R_k \dots R_0)$. On a la propriété suivante :

$$Q_k A_{k+1} = Q_k R_k Q_k = A_k Q_k.$$

On en déduit :

$$(Q_0 \dots Q_{k+1})(R_{k+1} \dots R_0) = (Q_0 \dots Q_k) A_{k+1} (R_k \dots R_0) = A(Q_0 \dots Q_k)(R_k \dots R_0) = A^{k+1}$$

Comme

$$det(Q_k^* A_k Q_k - \lambda I) = det(Q_k^*(A_k - \lambda I)Q_k) = det(A_k - \lambda I),$$

les matrices A_k on toutes les mêmes valeurs propres que A ; donc si A_k converge vers une matrice diagonale, les éléments diagonaux sont les valeurs propres de A. Soit $\mathbf{X} \in \mathbb{C}^n$ un vecteur propre de A_k associé à la valeur propre $\lambda \in \mathbb{C}$. On a les relations suivantes :

$$A_k \mathbf{X} = \lambda \mathbf{X} \Rightarrow A_{k+1} Q_k^* \mathbf{X} = Q_k^* A_k Q_k Q_k^* \mathbf{X} = \lambda Q_k^* \mathbf{X}$$

Donc si $\mathbf{X}$ est un vecteur propre de A_k alors $Q_k^* \mathbf{X}$ est un vecteur propre de A_{k+1}.

■

Proprieté A.2. Les matrices de A_k définies équation (12.A) s'écrivent $A_k = \tilde{Q}_k^* A \tilde{Q}_k$, où les matrices $\tilde{Q}_k$ sont construites par l'algorithme :

$$\begin{cases} \tilde{Q}_0 = I \\ Y_{k+1} = A\tilde{Q}_k \\ Y_{k+1} = \tilde{Q}_{k+1} R_{k+1}. \end{cases} \tag{13.A}$$

Autrement dit $\tilde{Q}_{k+1} R_{k+1}$ est la décomposition QR de la matrice $A\tilde{Q}_k$

Démonstration. Par récurrence.

$$A = \tilde{Q}_0^* A \tilde{Q}_0 = A = A_0.$$

On suppose que $A_k = \tilde{Q}_k^* A \tilde{Q}_k$. D'après (13.A) $A\tilde{Q}_k = \tilde{Q}_{k+1} R_{k+1}$, donc $\tilde{Q}_k^* A \tilde{Q}_k = \tilde{Q}_k^* \tilde{Q}_{k+1} R_{k+1}$ et le produit d'une matrice orthogonale $Q = \tilde{Q}_k^* \tilde{Q}_{k+1}$ et d'une matrice triangulaire supérieur $R = R_{k+1} = \tilde{Q}_{k+1}^* A \tilde{Q}_k$. Comme la décomposition QR est unique (Hormis le fait que l'on peut multiplier chaque colonne de Q et chaque ligne de R par -1) alors on a d'après (12.A) :

$$A_{k+1} = R_{k+1} \tilde{Q}_k^* \tilde{Q}_{k+1} = \tilde{Q}_{k+1}^* A \tilde{Q}_k \tilde{Q}_k^* \tilde{Q}_{k+1} = \tilde{Q}_{k+1}^* A \tilde{Q}_{k+1}$$

■

Proprieté A.3. Si $A = A_0$ est symétrique alors la matrice A_k définie par l'algorithme (12.A) est symétrique $\forall k$.

Démonstration. D'après la proposition A.2 il existe $\tilde{Q}_k$ orthogonale tel que $A_k = \tilde{Q}_k^* A \tilde{Q}_k$ donc

$$\forall (\mathbf{X}, \mathbf{Y}) \in (\mathbb{C}^n \times \mathbb{C}^n),\ \mathbf{X}^* A_k \mathbf{Y} = (\tilde{Q}_k \mathbf{X})^* A (\tilde{Q}_k \mathbf{Y}) = (\tilde{Q}_k \mathbf{Y})^* A (\tilde{Q}_k \mathbf{X}) = \mathbf{Y}^* A_k \mathbf{X}$$

■

Proprieté A.4. Si A est trigonalisable (toujours dans $\mathbb{C}$), quand $k \to \infty$ alors la suite A_k construite par l'algorithme 12.A tend vers une matrice triangulaire supérieur qui a ses éléments diagonaux égaux aux valeurs propres de A. De plus, les valeurs propres de la matrice triangulaire supérieur sont dans l'ordre décroissant le long de la diagonale. Si A est réelle et non trigonalisable, alors A_k tends vers une matrice de même forme que la matrice de la décomposition de Schur[72] dans $\mathbb{R}$.

Démonstration. (euristique) Soit $\tilde{Q}_k$ la matrice construite par l'algorithme (13.A) alors $A_k = \tilde{Q}_k^* A \tilde{Q}_k$.
Si on écrit $\tilde{Q}_k = [\tilde{Q}_{1k}, \tilde{Q}_{2k}]$ avec $\tilde{Q}_{1k}$ ayant p colonnes ($p \leq n$) alors on a :

$$A_k = \tilde{Q}_k^* A \tilde{Q}_k = \begin{bmatrix} \tilde{Q}_{1k}^* A \tilde{Q}_{1k} & \tilde{Q}_{1k}^* A \tilde{Q}_{2k} \\ \tilde{Q}_{2k}^* A \tilde{Q}_{1k} & \tilde{Q}_{2k}^* A \tilde{Q}_{2k} \end{bmatrix}$$

Pour plus de visibilité on rappelle l'algorithme (13.A) :

$$\begin{cases} \tilde{Q}_0 = I \\ Y_{k+1} = A \tilde{Q}_k \\ Y_{k+1} = \tilde{Q}_{k+1} R_{k+1} \end{cases}$$

On voit alors que les espaces engendrés par les vecteurs des matrices $\tilde{Q}_k$ vérifient :

$$vect(\tilde{Q}_{k+1}) = vect(Y_{k+1}) = vect(A \tilde{Q}_k) = vect(A^k \tilde{Q}_0)$$

Et donc si $\tilde{Q}_{1k}$ converge vers un sous-espace de $\mathbb{C}^n$ alors $A\tilde{Q}_{1k}$ converge vers le même sous espace et $\tilde{Q}_{2k}^* A \tilde{Q}_{1k}$ converge vers $\tilde{Q}_{2k}^* \tilde{Q}_{1k} = 0$.
Si A est trigonalisable on montre que ceci est vrai pour tout $p < n$ et donc tous les éléments sous-diagonaux tendent vers 0. Et comme on a vu que toutes les matrices A_k ont les mêmes valeurs propres (proposition A.1) alors on en déduit que la limite de A_k quand $k \to \infty$ à les valeurs propres de A sur sa diagonale. L'ordre décroissant s'explique par la méthode la puissance itérée (cf annexe A.6.2) qui indique que les premiers vecteurs de Q^{k+1} sont orthogonaux entre eux et prennent la direction des vecteurs de $A^k \tilde{Q}_0$ qui sont les premiers vecteurs propres de A (associé aux plus grandes valeurs propres). En particulier $\lim_{k\to\infty} \tilde{Q}_{1k}$

contient un vecteur propre de A. Dans le cas où la matrice est symétrique ou hermitienne, une transformation orthogonale conserve la symétrie dans $\mathbb{R}$ et l'hermiticité dans $\mathbb{C}$ (Car $(QAQ^*)^* = QAQ^*$) donc si les coefficients sous diagonaux tendent vers 0 alors les coefficients sur diagonaux aussi. ■

A.7 Arbre couvrant de poids minimal (Minimum spanning tree)

A.7.1 Introduction

Nous allons présenter des exemples d'algorithme permettant de choisir les arêtes nécessaires pour la construction du gradient discret G défini précédemment (cf section IV.5.2.1).

A.7.2 Définitions

Soit Gr un graphe connexe et non orienté, un arbre couvrant de Gr est un sous-ensemble d'arêtes constituant un arbre qui connecte tous les sommets de Gr ensemble.

Chaque arête est ensuite pondérée d'un poids qui est une grandeur réelle positive. On peut par exemple, identifié le poids d'une arête à sa longueur. Ensuite, la somme des poids des arêtes d'un arbre couvrant représente le poids de cet arbre. Un arbre couvrant de poids minimal est un arbre couvrant possédant un poids inférieur ou égal à celui de tous les autres arbres couvrants du graphe.

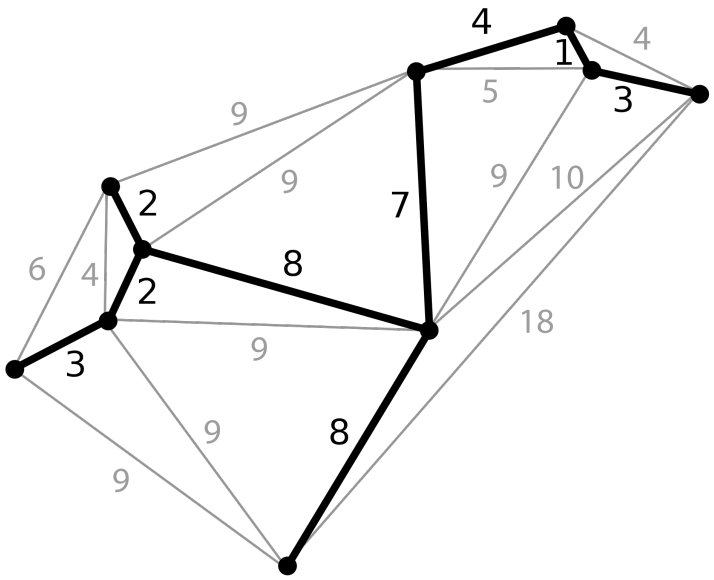


Fig 4.A L'arbre couvrant de poids minimal d'un graphe planaire [116].

Plus précisément si on appelle S l'ensemble des sommets d'un graphe Gr, A l'ensemble des arêtes et w l'ensemble des poids des arêtes et si $Gr = (S, A, w)$ est connexe alors un arbre couvrant pour Gr est un sous-ensemble T de A tel que :

- (S, T) est un graphe connexe et acyclique,
- tout sommet $u \in S$ apparaît dans une arête de T ;

Et puisque T est connexe acyclique

- $|T| = |S| - 1$
- l'ajout d'une nouvelle arête dans T crée un cycle,
- le retrait d'une arête de T le rend non connexe,
- c'est un arbre et donc entre tout couple de sommets il existe un chemin unique.

Un Arbre Couvrant de Poids Minimal (ACPM) de Gr est un arbre couvrant de poids $w_s = \sum_{(u,v)\in T} w(u,v)$ minimal. Pour construire le gradient discret G (cf section IV.5.2.1) on doit trouver un l'arbre couvrant de poids minimal au graphe du maillage en éléminant les éléments des faces PEC, PMC et esclaves. Les poids des arêtes peuvent être égaux à la différence des indices des 2 nœuds qui la constituent mais ce n'est pas la seul possibilité. Pour que les dimensions correspondent, on doit choisir le premier nœud de l'arbre dans les éléments des faces PEC, PMC ou esclaves. Ce premier nœuds correspond au potentiel nul : c'est le nœud de référence. Voici les algorithmes classiques permettant de construire un ACPM.

A.7.3 Algorithme générique

Soit $Gr = (S, A, w)$ un graphe pondéré. On veut construire un ACPM T.

1. $T = \emptyset$,
2. tant que T nest pas un arbre couvrant pour Gr faire
3. Choisir une arête (u, v) telle qu'il existe un ACPM qui inclus $T \cup \{(u, v)\}$
4. $T = T \cup \{(u, v)\}$
5. retourner en 2.

A.7.4 Algorithme de Prim

On désigne par A_i les arêtes du graphe Gr.

1. Choisir une arête A_1 de poids minimal et ajouter l'arête à l'arbre T.
2. Étape n :tant que $|T| \neq |S| - 1$,
   Choisir un sommet u_n qui ne crée pas de cycle (i.e $u_n \notin T$)
   tel qu'il soit un voisin de tous les sommets déjà ajoutés à l'arbre,
   et tel qu'il crée une arête de poids minimale.
3. Ajouter le sommet u_n et l'arête A_n correspondante à l'arbre T et retourner en 2.

A.7.5 Algorithme de Kruskal

On désigne par $\mathcal{V}(u_n)$ tous les voisins de u_n qui sont dans l'arbre T, autrement dit ce sont tous les points des sous-arbres qui contiennent u_n.

1. Choisir une arête A_1 de poids minimal puis l'ajouter à l'arbre T.
2. Étape n :tant que $|T| \neq |S| - 1$,
   Choisir une arête $A_n = \{(u_n, v_n)\}$ de poids minimal qui n'est pas dans l'arbre et qui ne créer pas de cycle i.e :
   $$\mathcal{V}(u_n) \cap \mathcal{V}(v_n) = \emptyset$$

3. Ajouter les sommets $\{(u_n, v_n)\}$ et l'arête A_n à l'arbre et retourner en 2.

A.8 Compilation des exécutables

On donne dans cette annexe quelques indications pour compiler un exécutable permettant d'effectuer les calculs que nous avons vus section V.3 et V.4. Avant de compiler les différents exécutables il faut s'assurer de posséder les librairies nécessaires :

- La bibliothèque MKL qui contient le solveur PARDISO[37] et les routines associées.
- La bibliothèque ARPACK[33] qui contient les routines de calcul de valeurs propres par l'algorithme d'Arnoldi.
- La bibliothèque Import_bma [1] qui permet de lire les fichiers de maillage sous format *.bma* .

Remarque A.4. Les bibliothèque s ARPACK et MKL se trouvent sur internet alors que Import_bma est une bibliothèque développée par l'ONERA . Ainsi un utilisateur hors ONERA devra développer lui-même des routines de post-traitement et de lecture du maillage.

Remarque A.5. Les bibliothèques ARPACK et Import_bma sont des librairies statiques. Autrement dit ces librairies se lieront à l'exécutable lors de la compilation. La bibliothèque MKL contient des librairies dynamiques qui viendront se greffer à l'exécutable au moment de l'exécution du programme. Et donc le chemin du répertoire où se trouve la bibliothèque MKL devra être indiqué à la variable LD_LIBRARY_PATH où via la commande module.

On rappelle que l'exécutable à une extension *.x* par convention. Les sources de l'exécutable sont l'ensemble des fichiers nécessaires pour compiler l'exécutable et les règles de compilation les indications nécessaires à la compilation de l'exécutable. Les noms des fichiers sources se trouvent dans les règles de compilation. Le nom des sources et les règles de compilation s'éditent en général dans un fichier *Makefile*. Le chemin d'accès des librairies statiques doit être indiqué dans les règles de compilation via l'option du compilateur -L. Aussi les librairies statiques qui ont l'extension l'extension *.a*, devront être associées à l'exécutable sans l'extension *.a* en ajoutant -l et sans faire apparaître le préfixe *lib*, sinon tout le chemin du fichier *.a* doit être indiqué. Par exemple si la librairie s'appelle libmkl_solver.a , on écrira -lmkl_solver dans les règles de compilation ou bien on écrira ...chemin_de_libmkl_solver.a/libmkl_solver.a.

Voici quelques lignes de code d'un fichier *Makefile* utilisé pour compiler l'exécutable ./Pard_FEM.x. (Les lignes commençants par le symbole # sont des lignes de commentaires) :

```
# Déclaration de variables et attribution d'un nom de répertoire .
LIBDIR = ../lib
DIRMKL=/usr/local/intel/cmkl/11.1.064/mkl/lib/em64t
# Déclaration de variables et attribution d'un nom de librairie .
ARPACKLIB = ../lib/libarpack_SUN4.a
```

LIBMKL = -lmkl_blas95_lp64 -lmkl_lapack95_lp64 -Wl,–start-group -lmkl_intel_lp64 -lmkl_sequential -lmkl_core -Wl,–end-group -lpthread

Liste des fichiers sources dont le l'exécutable dépend.

./Pard_FEM.x : Basics.o Solver_Pard.o DonneesPhysiques.o jeu_de_parametres.o mots_cles_ELSEM3D.o groupe_direct.o KeyWords_tools.o nedelec.o geometrie_tetra_6ddl.o Maillage.o Allocation.o Graphe_RG.o Assemblage.o Pard_FEM.o

Règles de compilation

ifort -L $LIBDIR -L $DIRMKL -o ./Pard_FEM.x Pard_FEM.o Basics.o Solver_Pard.o DonneesPhysiques.o jeu_de_parametres.o mots_cles_ELSEM3D.o groupe_direct.o KeyWords_tools.o nedelec.o geometrie_tetra_6ddl.o Maillage.o Allocation.o Graphe_RG.o Assemblage.o -lImport_bma $(ARPACKLIB) $(LIBMKL)

Références

[1] Thibault Volpert. Environnement pame outils de post traitement et de lecture du maillage. *ONERA*, 1999.

[2] Juslan Lo. *ÉTUDE DE LA RECONFIGURABILITÉ D'UNE STRUCTURE À BANDE INTERDITE ÉLECTROMAGNÉTIQUE (BIE) MÉTALLIQUE PAR PLASMAS DE DÉCHARGE*. PhD thesis, l'Université Toulouse III - Paul Sabatier Discipline ou spécialité : Micro-ondes, Électromagnétisme, et Optoélectronique, 14 mai 2012.

[3] L. Rayleigh. On the maintenance of vibrations by forces of double frequency, and on the propagation of waves through a medium endowed with a periodic structure. *Philosophical Magazine*, 24 :145–159, 1887.

[4] John D. Joannopoulos, Steven G. Johnson, Joshua N. Winn, and Robert D. Meade. Photonic cristal book. molding the flow of light. *PRINCETON UNIVERSITY PRESS*, November 2007.

[5] Yahya Rahmat-Samii. Ebg structures for low profile antenna designs : What have we learned ? *Proceedings of the second European Conference on Antennas and Propagation (EUCAP)*, 2007.

[6] V. G. Veselago and E. E. Narimanov. The left hand of brightness : past, present and future of negative index materials. *Nature Materials*, 5 :759–762, 2006.

[7] Viktor G Veselago. The electrodynamics of substances with simultaneously negative values of ϵ and μ. *Soviet Physics Uspekhi*, 10(4) :509, 1968.

[8] Alain Priou. *Matériaux composites en électromagnétisme*. Techniques de l'Ingénieur, (E1165), 1968.

[9] Mohamed Farhat, Sebastien Guenneau, and Stefan Enoch. Ultrabroadband elastic cloaking in thin plates. *Phys. Rev. Lett.*, 103(2) :024301, July 2009.

[10] P. de Maagt, R. Gonzalo, Y.C. Vardaxoglou, and J.-M. Baracco. Electromagnetic bandgap antennas and components for microwave and (sub)millimeter wave applications. antennas and propagation. *IEEE Transactions on*, 51(10) :2667–2677, October 2003.

[11] B.A. Munk. Frequency-selective surfaces : Theory and design. *Wiley, New York*, 2000.

[12] J. C. Vardaxoglou. Frequency-selective surfaces : Analysis and design. *Research Studies Press, Ltd., Taunton, UK*, 1997.

[13] Farhad Bayatpur. *Metamaterial-Inspired Frequency-Selective Surfaces*. PhD thesis, The University of Michigan, 2009.

[14] T. K. Wu. Frequency-selective surface and grid array. *Wiley, New York*, 1995.

[15] R. Ulrich. Far-infrared properties of metallic mesh and its complementary structure. *Infrared Phys*, 7 :37–55, 1967.

[16] G. Marconi and C.S. Franklin. Reflector for use in wireless telegraphy and telephony,. *U.S. Patent*, pages 1,301,473, 1914.

[17] Stanford University California. Railroad track for rotating the antenna, 2012. http://www.reeve.com/RadioScience/RadioAstronomyFacilities/Stanford-CA/Stanford-CA.htm.

[18] Mats Halldin. Radome (wikipedia), 2005. http://fr.wikipedia.org/wiki/Fichier:Navy-Radome.jpg.

[19] Grimlock. F-117 furtif (wikipedia), 2004. http://fr.wikipedia.org/wiki/Fichier:US_Air_Force_F-117_Nighthawk.jpg.

[20] Dan Sievenpiper, Lijun Zhang, Romulo F. Jimenez Broas, Nicholas G. Alex opolous, and Eli Yablonovitch. High-impedance electromagnetic surfaces with a forbidden frequency band. *IEEE Transactions on Microwave Theory and Techniques*, VOL. 47, NUMBER 11, NOVEMBER 1999.

[21] Nicolas Capet. *Amélioration du découpage inter éléments par Surface Haute Impédance pour des antennes réseaux GNSS compactes*. PhD thesis, Université Paul Sabatier III, ONERA, Toulouse, Décembre 2010.

[22] Roy Sambles, Alastair Hibbins, and Matthew Lockyear. Manipulating microwaves with spoof surface plasmons, 2009. http://spie.org/x33849.xml.

[23] T-H. Vu, A-C. Tarot, S. Collardey, and K. Mahdjoubi. Bandwidth enlargement of planar ebg antennas. *Loughborough Antennas and Propagation Conference*, April 2007.

[24] Thai-Hung VU, Anne-Claude TAROT Sylvain COLLARDEY, and Kourosh MAHDJOUBI. Bandwidth enlargement of ebg antennas by a combined prs. *IETR Laboratories UMR 6164 University of Rennes 1, France*, Mai 2008.

[25] Esteban Moreno, F. J. Garcia-Vidal, and L. Martin Moreno. Enhanced transmission and beaming of light via photonic crystal surface modes. *PHYSICAL REVIEW B 69*, 2004.

[26] Stéphane Varault. *Modélisation et études expérimentales de structures à bande interdite électromagnétique reconfigurables intégrant des capillaires plasmas pour applications micro-onde*. PhD thesis, Université Paul Sabatier III, ONERA, Toulouse, Février 2011.

[27] Guy EYMIN PETOT TOURTOLLET, Tan-Phu Vuong (IMEP-LAHC), Pierre Lemaitre-Auger (LCIS), and AHLSTROM. Dossier de presse metapapier. *Centre terchnique du papier Domaine Universitaire BP251 38044 GRENOBLE CEDEX 09 FRANCE www.webCTP.com*, Mars 2012.

[28] Marianne Imperor-Clerc. PhD thesis, Laboratoire de Physique des Solides d'Orsay(LPS), 2000.

[29] V. Kuzmiak, A. A. Maradudin, and F. Pincemin. Photonic band structures of twodimensional systems containing metallic components. *Phys. Rev. B*, 50(23) :1683516844, December 1994.

[30] V. Kuzmiak and A. A. Maradudin. Photonic band structures of one- and two-dimensional periodic systems with metallic components in the presence of dissipation. *Phys. Rev. B*, 50(12) :74277444, Mars 1997.

[31] André Barka. Logiciel onera factopo, 2001. Dépot à l'agence de protection des programmes (APP) IDDN.FR.001.400003.00.S.P.2001.000.31235.

[32] C. J. Reddy, Manohar D. Deshpande, C. R. Cockrell, and Fred B. Beck. Finite element method for eigenvalue problems in electromagnetics. *NASA Technical Paper 3485*, december 1994.

[33] Lehoucq, D.C Sorensen, and C.Yang. *ARPACK USER'S GUIDE : Solution of large scale Eigenvalues problems with implicitly restarted Arnoldi Methods*. SIAM, 1998.

[34] Neelakantam Venkatarayalu and J. F. Lee. Removal of spurious dc modes in edge element solutions for modeling three-dimensional resonators. *IEEE Transactions on Microwave Theory and Techniques*, VOL. 54, NO. 7, July 2006.

[35] C. Mias, J. P. Webb, and R. L. Ferrari. Finite element modelling of electromagnetic waves in doubly and triply periodic structures. *IEE*, march 1999.

[36] Li Xu, Zhen Ye, Jian-Qing Li, and Bin Li. A novel approach of removing spurious dc modes in finite-element solution for modeling microwave tubes. *IEEE Transactions on Electron Devices*, Vol. 57, no. 11, November 2010.

[37] O. Schenk and K. Gärtner. On fast factorization pivoting methods for symmetric indefinite systems. *JElec. Trans. Numer. Anal.*, 23 :158–179, 2006.

[38] Roger f. harrington. *Time-harmonic electromagnetic fields*. McGraw-Hill, 1961.

[39] Swartz L. *Méthodes mathématiques pour les sciences physiques*. Herman, Paris, 1961.

[40] Jean-Marie BRÉBEC. *H Prépa Électromagnetisme* 2^{nde} *année* $PC-PC^*$,$PSI-PSI^*$. HACHETTE supérieur, 1998.

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

[42] Daniel T. McGrath and Vittal P. Pyati. Phased array antenna analysis with the hvbrid finite element method. *IEEE Transactions on Antennas and Propagation*, VOL. 42, NO. 12, december 1994.

[43] A. Bravais. Mémoire sur les systèmes formés par les points distribués régulièrement sur un plan ou dans l'espace. *Journal de l'Ecole Polytechnique : 1-128*, 19 :1–128, 1850.

[44] Dr. A. RASKIN. Exemple de mailles élémentaires, 2010. http://umvf.univ-nantes.fr/odontologie/enseignement/chap1/site/html/3_31_311_1.html.

[45] Christian Nolleke. Zur konstruktion einer wigner-seitz-zelle (wikipedia), 2007. http://fr.wikipedia.org/wiki/Fichier:Wigner-Seitz-Zelle.png.

[46] Grimlock. Zone de brillouin (wikipedia), 2007. http://fr.wikipedia.org/wiki/Zone_de_Brillouin.

[47] R. CHiniard. *contribution à la modélisation de la surface équivalente radar de grandes antennes réseaux par une approche multidomaine / Floquet*. PhD thesis, Université Toulouse Paul Sabatier, 2007.

[48] R. CHiniard, A. Barka, and O. Pascal. Hybride fem / floque modes / po technique for the rcs prediction of array antennas. *IEEE Transactions on Antennas and Propagation*, 56 no.6, June 2008.

[49] Jean-Michel lourtioz, Henri Benisty, and Vincent Berger. Les cristaux photoniques ou la lumière en cage. *Edition Hermes Lavoisier*, pages 26–40, 2003.

[50] M. G. Moharam and T. K. Gaylord. "rigorous coupled-wave analysis of planargrating diffraction". *J. Opt. Soc. Am. A*, 71(7) :811–818, 1981.

[51] Samuel Nosal. *Modélisation électromagnétique de structures périodiques et matériaux artificiels. Application à la conception d'un radôme passe bande*. PhD thesis, Laboratoire dEnergétique Moléculaire et Macroscopique, Combustion (E.M2.C.) UPR 288, CNRS et École Centrale Paris, 3 septembre 2009.

[52] Nicolas Chateau and Jean-Paul Hugonin. Algorithm for the rigorous coupled-wave analysis of grating diffraction. *J. Opt. Soc. Am. A*, 11(4) :13211331, Avril 1994.

[53] Philippe Lalanne and G. Michael Morris. Highly improved convergence of the coupled-wave method for tm polarization. *J. Opt. Soc. Am. A*, 13(4) :779784, 1996.

[54] Lifeng Li. New formulation of the fourier modal method for crossed surface-relief gratings. *J. Opt. Soc. Am. A*, 14(10) :27582767, October 1997.

[55] Kazuaki Sakoda, Noriko Kawai, Takunori Ito, Alongkarn Chutinan, Susumu Noda, Tsuneo Mitsuyu, and Kazuyuki Hirao. Photonic bands of metallic systems. i. principle of calculation and accuracy. *Phys. Rev. B*, 64(4) :045116, July 2001.

[56] N. W. Ashcroft and N. D. Mermin. *Solid State Physics*. Saunders College Publishing, Fort Worth, 1976.

[57] P. Debye. reprinted 1954 in collected papers of peter j.w. debye interscience, new york. *Ox Bow Press, 1954*, Ver. Deut. Phys. Gesell. 15, 777, 1913.

[58] Allen Taflove and Morris E.Brodwin. Numerical solution of steady state electromagnetic scattering problems using the time-dependent maxwell's equation. *IEEE transactions on microwave theory and techniques*, Vol 23 no.8, August 1975.

[59] Kane Yee. Numerical solution of initial boundary value problems involving maxwell's equations in isotropic media. *IEEE Transactions on Antennas and Propagation*, 14, 1966.

[60] Richard M. and R.F Warming. An implicit finite-difference algorithm for hyperbolic systems in conservation-law form. *Computational Fluid Dynamics Branch Ames Research Center NASA Mojfett Field California 94035 USA*, 22 :87–110, September 1976.

[61] Belkhir Abderrahmane. *Extension de la modélisation par FDTD en nano-optique*. PhD thesis, Université de Franche-Comté, École doctorale SPIM, November 2008.

[62] Roger F. Harrington and Robert E. Krieger. *Field Computation by Moment Methods*. Publishing Co., Malabar, Florida, 1968.

[63] A. J. Poggio and E. K. Miller. Integral equation solutions of three-dimensional scattering problems. in raj mittra. *Pergamon Press, New York*, 1973.

[64] George Green. An essay on the application of mathematical analysis to the theory of electricity and magnetism, journal. *für die reine und angewandte Mathematik (connu aussi sous le nom de "Journal de Crelle ")*, 39 :73–79, 1850.

[65] SADASIVA M. RAO, DONALD R. WILTON, and ALLEN W. GLISSON. Electromagnetic scattering by surfaces of arbitrary shape. *IEEE TRANSACTIONS ON ANTENNAS AND PROPAGATION*, VOL. AP-30, NO. 3, May 1982.

[66] El Hadji KONÉ. *Equations intégrales volumiques pour la diffraction d'ondes électromagnétiques par un corps diélectrique*. PhD thesis, lUMR 6625 CNRS - IRMAR Institut de Recherche Mathématique de Rennes U.F.R. de Mathématique, 23 Juin 2010.

[67] A. Barka. *Habilitation à diriger des recherches (HDR)*. PhD thesis, Institut National Polytechnique de Toulouse, 2008.

[68] Jianming Jin. *The Finite Element Method in Electromagnetics*. New York Wiley, 1993.

[69] Pierre Arnaud Raviart and Jean-Marie Thomas. *Introduction à l'analyse numérique des équations aux dérivées partielles*. Dunod, july 2004.

[70] J.C. Nedelec. Mixed finite elements in $\mathbb{R}^3$, numer. math. 35. *Numerische Mathematik Springer-Verlag*, pages 315–341, 1980.

[71] J. Scott Savage and Andrew F. Peterson. Higher-order vector finite elements for tetrahedral cells. *IEEE TRANSACTIONS ON MICROWAVE THEORY AND TECHNIQUES*, 44 ,NO.6, June 1996.

[72] Horn, Roger A., Johnson, and Charles R. *Matrix Analysis*. Cambridge University Press, 1985.

[73] J.F. Lee and D. K. Sun. pmus (p-type multiplicative schwarz) method with vector finite elements for modeling three-dimensional waveguide discontinuities. *IEEE Trans. Microw. Theory Tech*, 52, no. 3 :864–870, March 2004.

[74] P. Monk. A simple proof of convergence for an edge element discretization of maxwell's equations. *Comput. Electromagn., Lecture Notes Comput. Sci. Eng.*, vol. 28 :127–142, 2003.

[75] Zachary S. Sacks, David M. Kingsland, Robert Lee, and Jin-Fa Lee. A perfectly matched anisotropic absorber for use as an absorbing boundary condition. *IEEE TRANSACTIONS ON ANTENNAS AND PROPAGATION*, Vol 43, no 12, December 1995.

[76] Weng Cho Chew and William H. Weedon. A 3-d perfectly matched medium from modified maxwell's equations with streched coordinates. weng cho chew and william h. weedon. *Micro. Opt. Lett.*, vol. 7, no. 13 :599–604, September 1994.

[77] Pedro Ferreira. *Étude numérique de quelques problèmes de diffraction d'ondes par des réseaux périodiques en dimension 2*. PhD thesis, Ecole polytechnique Palaiseau, 1998.

[78] Amir Ali Tavallaee and Jon P. Webb. Finite-element modeling of evanescent modes in the stopband of periodic structures. *IEEE Transactions on Magnetic*, VOL. 44, NO. 6, JUNE 2008.

[79] Warwick and Kevin. "using the cayley hamilton theorem with n partitioned matrices". *IEEE Transactions on Automatic Control*, AC 28 No.12, :1127–1128, December 1983.

[80] François Xavier Roux. ONERA. Analyse numérique matricielle avancée et calcul parallèle : chapitre 6, 2012. http://www.ljll.math.upmc.fr/MathModel/polycopies/fxroux6.pdf.

[81] B. N. Parlett. Presentation geometrique des methodes de calcul des valeurs propres. *Numer. Math.*, 21 :223–233, 1973.

[82] Vera N. Kublanovskaya. On some algorithms for the solution of the complete eigenvalue problem. *USSR Computational Mathematics and Mathematical Physics*, 1, no.3 :637–657, 1963.

[83] Alexander Maltsev, Vladimir Pestretsov, Roman Maslennikov, and Alexey Khoryaev. Triangular systolic array with reduced latency for qr-decomposition of complex matrices. *IEEE ISCAS*, pages 385–388, 2006.

[84] Cullum and Willoughby. *Lanczos Algorithms for Large Symmetric Eigenvalue Computations, Vol.1*. SIAM, 2002.

[85] A. Barka, A. Cosnuau, and FX. Roux. Parallel organization of air intake electromagnetic mode computation on a distributed memory machine. *La recherche Aérospatiale*, no.6 :389–404, 1995.

[86] André Barka. Logiciel onera guide, 2001. Dépot à l'agence de protection des programmes (APP) IDDN.FR.001.400004.00.S.P.2001.000.31235.

[87] T. J. Dekker and J. F. Traub. The shifted qr algorithm for hermitian matrices. *University of Washington, Seattle Technical Report No. 70-11-07, Computer Science Group*, November 1970.

[88] Anderson Edward. Discontinuous plane rotations and the symmetric eigenvalue problem. *LAPACK Working Note 150, University of Tennessee, UT-CS-00-454*, December 2000.

[89] Peter Arbenz. ETH. Numerical methods for solving large scale eigenvalue problems : Chapter 3, 2012. http://people.inf.ethz.ch/arbenz/ewp/Lnotes/chapter3.pdf.

[90] D.C Sorensen. Implicit application of polynomial filters in a k-step arnoldi method. *Siam journal on matrix analysis and applications*, 13 :357–385, 1992.

[91] W. E. Arnoldi. "the principle of minimized iterations in the solution of the matrix eigenvalue problem". *Quarterly of Applied Mathematics*, 9 :17–29, 1951.

[92] O. Schenk and K. Gärtner. Solving unsymmetric sparse systems of linear equations with pardiso. *Journal of Future Generation Computer Systems*, 20(3) :475–487, 2004.

[93] Lehoucq, D.C Sorensen, and C.Yang. Implicitly restarted arnoldi method, 1998. http://www.caam.rice.edu/software/ARPACK/UG/node56.html#23chap5.

[94] Paul F.Combes. *Micro-ondes 2. Circuits passifs, propagation, antennes*. Dunod Université, 1999.

[95] Christine Bernardi and Frédéric Hecht. Quelques propriétés d'approximation des éléments finis de nédélec, application à l'analyse à posteriori. *hal-00139231v version 1-30*, Mars 2007.

[96] HFFS and version 13(ANSYS). Setting the maximum delta frequency per pass. october 2010.

[97] S. Perepelitsa, R. Dyczij-Edlinger, and J. F. Lee. Finite-element analysis of arbitrarily shaped cavity resonator using h^1(curl) elements. *IEEE Trans. Magn.*, vol. 3, no. 3 :pp. 17761779, March 2004.

[98] Hestenes Magnus R. and Stiefel Eduard. "methods of conjugate gradients for solving linear systems". *Journal of Research of the National Bureau of Standards*, 49(6), December 1952.

[99] Patrick R. Amestoy, Timothy A., Davisy Iain, and S. Duff. An approximate minimum degree ordering algorithm. *SIAM J. Matrix Analysis and Applic*, Vol 17, no 4 :886–905, December 1996.

[100] Pham Quang Cuong. Résolution de grands système linéaires, 2003. http://www.normalesup.org/~pham/docs/tipe.pdf.

[101] Cuthill and J. McKee. Reducing the bandwidth of sparse symmetric matrices in proc. 24th nat. conf. acm. *IEEE Transactions on Antennas and Propagation*, pages 157–172, 1969.

[102] Philippe Ciarlet. *Introduction à l'analyse numérique matricielle et à l'optimisation*. Masson, coll. Math. Appl. pour la Maîtrise, 2001.

[103] S. C. Eisenstat, M. C. Gursky, M. H. Schultz, and A. H. Sherman. Yale sparse matrix package. *the symmetric codes, Int. J. Numer. Methods in Engin*, pages 1145–1151, April 1982.

[104] Y. Saad. Iterative methods for sparse linear systems. *Society for Industrial and Applied Mathematics*, 2nd edition, 2003.

[105] John R. Gilbert, Cleve Moler, and Robert Schreiber. Sparse matrices in matlab :design and implementation. *Xerox Corporation, Research Institute for Advanced Computer Science, and The MathWorks Incorporated. All rights reserved*, 1991.

[106] Chatterjee A., Jin J. M., and Volakis J. L. Computation of cavity resonances using edge-based finite elements. *IEEE Trans. Microw. Theory and Techni.*, vol. 40, no. 11 :2106–2108, Nov 1992.

[107] Benjamin Gabard. Spectre de transmission d'un réseau carré orienté selon γx, maille de 7 mm, 2012. http://www.onera.fr/toulouse/.

[108] AXEL RUHEF. Algorithms for the nonlinear eigenvalue problem. *SIAM J. NUMER. ANAL.*, 10, No.4 :674–700, September 1973.

[109] Philippe Guillaume. Nonlinear eigenproblems. *SIAM J. NUMER. ANAL.APPL*, 20, No.3 :575–595, 1999.

[110] N. Marcuvitz. *Waveguide Handbook*. New York : McGraw-Hill, 1951.

[111] XuH.A. Kramers. La diffusion de la lumiere par les atomes. *Atti Cong. Intern. Fisica, (Transactions of Volta Centenary Congress)*, 2 :547–557, 1927.

[112] Webster's Online Dictionary. Extended definition : permittivity, 2012. http://www.websters-online-dictionary.org/definitions/permittivity?cx=partner-pub-0939450753529744%3Av0qd01-tdlq&cof=FORID%3A9&ie=UTF-8&q=permittivity&sa=Search#906.

[113] J.Dodzuk. Finite difference approach to the hodge theory of harmonic form. *American Journal of Mathematics*, 98 :79–104, 1976.

[114] Carl Friedrich Gauss. *Methodus nova integralium valores per approximationem inveniendi*. Gottingae, 1815.

[115] E. Pereira and J.Vitòria. Deflation for block eigenvalues of block partitioned matrices with an application to matrix polynomials of commuting matrices. *ELSEVIER*, 42 :1177–1188, 1973.

[116] Dcoetzee. The minimum spanning tree of a planar graph, 2005. http://en.wikipedia.org/wiki/File:Minimum_spanning_tree.svg.