

Méthodes de type Galerkin discontinu pour les équations de Maxwell en régime harmonique: flux numériques et algorithmes multigrille

Victorita Dolean, Stephane Lanteri, Ronan Perrussel

▶ **To cite this version:**

Victorita Dolean, Stephane Lanteri, Ronan Perrussel. Méthodes de type Galerkin discontinu pour les équations de Maxwell en régime harmonique: flux numériques et algorithmes multigrille. [Rapport de recherche] RR-6805, INRIA. 2009, pp.45. inria-00354510

HAL Id: inria-00354510
https://inria.hal.science/inria-00354510

Submitted on 20 Jan 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Méthodes de type Galerkin discontinu pour les équations de Maxwell en régime harmonique: flux numériques et algorithmes multigrille

Victorita Dolean — Stéphane Lanteri — Ronan Perrussel

N° 6805

20 janvier 2009

Thème NUM



ISSN 0249-6399 ISRN INRIA/RR--6805--FR+ENG

INRIA
SOPHIA ANTIPOLIS

Méthodes de type Galerkin discontinu pour les équations de Maxwell en régime harmonique: flux numériques et algorithmes multigrille

Victorita Dolean [*] , Stéphane Lanteri [†] , Ronan Perrussel [‡]

Thème NUM — Systèmes numériques
Projets Nachos

Rapport de recherche n° 6805 — 20 janvier 2009 — 44 pages

Résumé : Ce rapport traite de la résolution des équations de Maxwell en régime harmonique par une méthode de type Galerkin discontinu. Il rappelle tout d'abord la formulation du problème et propose quelques éléments de discussion pour le choix du flux numérique utilisé dans la méthode de discrétisation. Il présente ensuite quelques pistes pour la résolution des systèmes linéaires obtenus. En particulier, on étudie la possibilié de résoudre ces systèmes linéaires par des algorithmes multigrille algébriques et des résultats préliminaires sont obtenus dans le cas des équations de Maxwell en deux dimensions d'espace.

Mots-clés : électromagnétisme numérique, équations de Maxwell en régime harmonique, méthode de type Galerkin discontinu, méthode multigrille.

[*] Laboratoire J.A. Dieudonné, Université de Nice-Sophia Antipolis, UMR 6621
[†] INRIA Sophia Antipolis - Méditerranée, équipe-projet Nachos
[‡] Laboratoire Ampère, Ecole Centrale de Lyon, UMR 5005

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles, BP 93, 06902 Sophia Antipolis Cedex (France)
Téléphone : +33 4 92 38 77 77 — Télécopie : +33 4 92 38 77 65

Discontinuous Galerkin methods for the time-harmonic Maxwell equations: numerical fluxes and multigrid algorithms

Abstract: This report deals with the solution of time-harmonic Maxwell's equations by a discontinuous Galerkin method. First, the formulation of the problem is introduced and a few elements are discussed for the choice of the numerical flux which is used in the discretisation. Then, some ideas for solving the resulting linear systems are presented. In particular, a preliminary study is conducted for solving these linear systems by algebraic multilevel algorithms and preliminary results are obtained in the case of the two-dimensional Maxwell's equations.

Key-words: computational electromagnetism, time-harmonic Maxwell's equations, Discontinuous Galerkin method, multilevel method.

Sommaire

1 Introduction **4**

2 Formulation et discrétisation **4**
2.1 Problème considéré 4
2.1.1 Équations de Maxwell et conditions aux limites 4
2.1.2 Notations pour la présentation du problème 5
2.1.3 Cas bidimensionnel utilisé pour les exemples numériques 6
2.2 Discrétisation 7
2.2.1 Problème local 7
2.2.2 Quelques flux numériques possibles 9
2.2.3 Problème global 11
2.2.4 Formulation mixte du problème 12

3 Éléments de discussion pour le choix du flux numérique **14**
3.1 Résultats concernant l'ordre de convergence 14
3.1.1 Résultats existants pour la convergence 14
3.1.2 Estimation de l'ordre de convergence asymptotique sur un cas simple . 15
3.1.3 Convergence avec les maillages obtenus par raffinement uniforme . . . 18
3.1.4 Convergence avec la série de maillages distincts 20
3.2 Résultats concernant le problème spectral 21
3.2.1 Passage en revue de l'existant 21
3.2.2 Illustration numérique 23

4 Résolution et préconditionnement du système linéaire **26**
4.1 Résultats génériques pour les systèmes de type point-selle 26
4.1.1 Approche découplée 27
4.1.2 Approche couplée 28
4.2 Mise en oeuvre 30
4.2.1 L'approche et le préconditionnement découplé 30
4.2.2 Méthodes multigrille avec le lisseur couplé 33

5 Bilan **40**

1 Introduction

Ce rapport traite de la résolution numérique des équations de Maxwell en régime harmonique par une méthode de type Galerkin discontinu. Il vient compléter le rapport [DFLP06] notamment en présentant différents choix de flux numérique et en proposant des approches pour une résolution efficace des systèmes linéaires obtenus. Ce document ne contient pas de démonstration théorique mais discute principalement de résultats présents dans la littérature et d'exemples numériques à caractère illustratif en deux dimension d'espace. Les quelques idées qui sont développées s'appuient pour l'instant uniquement sur une démarche heuristique.

Les points suivants sont abordés :

- présentation de la **discrétisation** des équations de Maxwell en régime harmonique par des méthodes de type Galerkin discontinu. Certaines écritures du problème sont redondantes mais elles facilitent la comparaison avec d'autres travaux où la description du problème peut différer.
- Discussion autour de plusieurs **choix possibles pour les flux numériques** et de quelques propriétés pour ces différents choix.
- Présentation de la forme générale des systèmes algébrique résultant de la discrétisation des équations de Maxwell en régime harmonique par des méthodes de type Galerkin discontinu — nous verrons que l'on obtient des **systèmes de type point-selle** — et de quelques règles générales pour la résolution de cette catégorie de système. Plusieurs pistes sont alors envisagées pour mettre en œuvre les énoncés généraux et on s'intéresse plus particulièrement à la possibilité de construire des algorithmes de résolution multigrille algébriques pour ces systèmes.

2 Formulation et discrétisation

2.1 Problème considéré

2.1.1 Équations de Maxwell et conditions aux limites

Le système des équations de Maxwell adimensionnées en régime harmonique peut s'écrire sous la forme suivante :

$$\begin{cases} \mathrm{i}\omega\varepsilon\mathbf{E} - \mathrm{rot}\,\mathbf{H} = -\mathbf{J}, \\ \mathrm{i}\omega\mu\mathbf{H} + \mathrm{rot}\,\mathbf{E} = 0. \end{cases} \tag{1}$$

Les champs électrique et magnétique $\mathbf{E}$ et $\mathbf{H}$ sont les inconnues à déterminer et le vecteur $\mathbf{J}$ désigne une source de courant imposée. Les paramètres ε et μ sont respectivement la permittivité diélectrique relative (incluant les données sur la conductivité électrique) et la perméabilité magnétique relative ; on considère ici des matériaux linéaires et isotropes. La fréquence angulaire du problème est donnée par le paramètre ω.

Considéré sur un domaine Ω de frontière $\partial\Omega = \Gamma_a \cup \Gamma_m$, on adjoint au système (1) les conditions aux limites :

$$\begin{cases} \mathbf{n} \times \mathbf{E} = 0 \text{ sur } \Gamma_m, \\ \mathbf{n} \times \mathbf{E} - \mathbf{n} \times (\mathbf{H} \times \mathbf{n}) = \mathbf{n} \times \mathbf{E}^{\text{inc}} - \mathbf{n} \times (\mathbf{H}^{\text{inc}} \times \mathbf{n}) \text{ sur } \Gamma_a. \end{cases} \tag{2}$$

La première est rattachée à la notion de conducteur parfait (on parle aussi souvent de condition métallique). La seconde est une condition absorbante approchée où les champs $\mathbf{E}^{\text{inc}}$ et $\mathbf{H}^{\text{inc}}$ représentent les composantes d'une possible onde incidente.

2.1.2 Notations pour la présentation du problème

Présenter sous une autre forme, on cherche à résoudre le problème de formulation forte suivante (pour simplifier, on considère un problème de diffraction où le terme $\mathbf{J}$ est nul) :

$$\begin{cases} \text{Trouver le champ de vecteurs } \mathbf{W} \text{ tel que :} \\ \mathrm{i}\omega G_0 \mathbf{W} + G_x \partial_x \mathbf{W} + G_y \partial_y \mathbf{W} + G_z \partial_z \mathbf{W} = 0 \text{ dans } \Omega, \\ (M_{\Gamma_m} - G_{\mathbf{n}})\mathbf{W} = 0 \text{ sur } \Gamma_m, \\ (M_{\Gamma_a} - G_{\mathbf{n}})(\mathbf{W} - \mathbf{W}_{\text{inc}}) = 0 \text{ sur } \Gamma_a. \end{cases} \tag{3}$$

Les notations utilisées ont alors la signification suivante :

- le champ de vecteurs inconnu $\mathbf{W}$ représente le champ électromagnétique, soit :

$$\mathbf{W} = \begin{pmatrix} \mathbf{E} \\ \mathbf{H} \end{pmatrix}.$$

- la matrice G_0 rassemble les propriétés des milieux et s'écrit :

$$G_0 = \begin{pmatrix} \varepsilon \,\mathrm{Id}_3 & 0_{3\times 3} \\ 0_{3\times 3} & \mu \,\mathrm{Id}_3 \end{pmatrix}.$$

  où Id_k désigne une matrice identité à k lignes et k colonnes et $0_{l\times m}$ une matrice de zéros à l lignes et m colonnes.

- la matrice $G_{\mathbf{n}}$ représente matriciellement le produit vectoriel par le vecteur $\mathbf{n}$ (qui désigne ici la normale unitaire sortante du domaine) pour les deux composantes du champ électromagnétique :

$$G_{\mathbf{n}} = \begin{pmatrix} 0_{3\times 3} & N_{\mathbf{n}} \\ N_{\mathbf{n}}^t & 0_{3\times 3} \end{pmatrix} \text{ avec } N_{\mathbf{n}} = \begin{pmatrix} 0 & \mathbf{n}_z & -\mathbf{n}_y \\ -\mathbf{n}_z & 0 & \mathbf{n}_x \\ \mathbf{n}_y & -\mathbf{n}_x & 0 \end{pmatrix}.$$

La notation N^t désigne la transposée de N. Les matrices $G_{\mathbf{n}}^+$ et $G_{\mathbf{n}}^-$ représentent respectivement les parties positive et négative de la matrice[1]. On définit aussi $|G_{\mathbf{n}}|= G_{\mathbf{n}}^+ - G_{\mathbf{n}}^-$. On peut ainsi déduire ces différentes matrices de l'expression de $G_{\mathbf{n}}$ et de celle de $|G_{\mathbf{n}}|$ donnée par :

$$|G_{\mathbf{n}}|= \begin{pmatrix} N_{\mathbf{n}} N_{\mathbf{n}}^t & 0_{3\times 3} \\ 0_{3\times 3} & N_{\mathbf{n}}^t N_{\mathbf{n}} \end{pmatrix}.$$

- les matrices G_l avec l dans $\{x, y, z\}$ représentent les produits vectoriels avec les vecteurs de base $\mathbf{e}_l$ et s'écrivent :

$$G_l = \begin{pmatrix} 0_{3\times 3} & N_{\mathbf{e}_l} \\ N_{\mathbf{e}_l}^t & 0_{3\times 3} \end{pmatrix}.$$

- les matrices M_{Γ_m} et M_{Γ_a} sont utilisées pour la prise en compte des conditions aux limites, respectivement métallique imposée sur Γ_m et absorbante imposée sur Γ_a :

$$M_{\Gamma_m} = \begin{pmatrix} 0_{3\times 3} & N_{\mathbf{n}} \\ -N_{\mathbf{n}}^t & 0_{3\times 3} \end{pmatrix} \text{ et } M_{\Gamma_a} = |G_{\mathbf{n}}|.$$

Les détails pour relier l'écriture de (3) à la formulation (1) + (2) sont donnés dans [DFLP06].

2.1.3 Cas bidimensionnel utilisé pour les exemples numériques

Pour le cas bidimensionnel, on considère le problème transverse électrique dans le plan (O, x, y). Il n'y a donc plus de dépendance suivant z et les composantes $\mathbf{E}_z$, $\mathbf{H}_x$ et $\mathbf{H}_y$ sont nulles. Le problème obtenu est formellement identique au cas tridimensionnel si l'on admet les abus de notation suivants :

- le vecteur $\mathbf{W}$ est désormais égal à $\begin{pmatrix} \mathbf{E}_x & \mathbf{E}_y & \mathbf{H}_z \end{pmatrix}^t$.
- la matrice G_0 prend la forme réduite :

$$G_0 = \begin{pmatrix} \varepsilon \,\mathrm{Id}_2 & 0_{2\times 1} \\ 0_{1\times 2} & \mu \end{pmatrix}.$$

- la matrice $N_{\mathbf{n}}$ prend aussi une forme réduite :

$$N_{\mathbf{n}} = \begin{pmatrix} -\mathbf{n}_y \\ \mathbf{n}_x \end{pmatrix}.$$

[1] Si $G_{\mathbf{n}} = T\Lambda T^{-1}$ avec Λ une matrice diagonale contenant les valeurs propres de $G_{\mathbf{n}}$ alors $G_{\mathbf{n}}^{\pm} = T\Lambda^{\pm}T^{-1}$ où Λ^+ (respectivement Λ^-) regroupe les valeurs propres positives (respectivement négatives) de la matrice $G_{\mathbf{n}}$.

– les matrices G_x et G_y sont égales à :

$$G_x = \begin{pmatrix} 0_{2\times 2} & N_{\mathbf{e}_x} \\ N_{\mathbf{e}_x}^t & 0 \end{pmatrix} \text{ et } G_y = \begin{pmatrix} 0_{2\times 2} & N_{\mathbf{e}_y} \\ N_{\mathbf{e}_y}^t & 0 \end{pmatrix}.$$

– les matrices $G_{\mathbf{n}}$ et $G_{\mathbf{n}}^+$ sont égales pour un vecteur $\mathbf{n}$ donné à :

$$G_{\mathbf{n}} = \begin{pmatrix} 0_{2\times 2} & N_{\mathbf{n}} \\ N_{\mathbf{n}}^t & 0 \end{pmatrix} \text{ et } G_{\mathbf{n}^+} = \frac{1}{2}\begin{pmatrix} N_{\mathbf{n}} N_{\mathbf{n}}^t & N_{\mathbf{n}} \\ N_{\mathbf{n}}^t & 1 \end{pmatrix}.$$

– les matrices M_{Γ_m} et M_{Γ_a} s'écrivent toujours :

$$M_{\Gamma_m} = \begin{pmatrix} 0_{2\times 2} & N_{\mathbf{n}} \\ -N_{\mathbf{n}}^t & 0 \end{pmatrix} \text{ et } M_{\Gamma_a} = |G_{\mathbf{n}}| \text{ soit } M_{\Gamma_a} = \begin{pmatrix} N_{\mathbf{n}_F} N_{\mathbf{n}_F}^t & 0_{2\times 1} \\ 0_{1\times 2} & 1 \end{pmatrix}.$$

2.2 Discrétisation

2.2.1 Problème local

Si l'on effectue le produit scalaire de l'équation aux dérivées partielles de (3) par un champ de vecteurs régulier $\mathbf{V}$ et on intègre formellement sur un sous-ensemble K du domaine Ω, on obtient :

$$\int_K (\mathrm{i}\omega G_0 \mathbf{W})^t \, \overline{\mathbf{V}} dv + \int_K \left(\sum_{l\in\{x,y,z\}} G_l \partial_l(\mathbf{W}) \right)^t \overline{\mathbf{V}} dv = 0.$$

En passant à l'opérateur adjoint, un terme de bord apparaît et l'équation devient :

$$\int_K (\mathrm{i}\omega G_0 \mathbf{W})^t \, \overline{\mathbf{V}} dv - \int_K \mathbf{W}^t \left(\sum_{l\in\{x,y,z\}} G_l \partial_l(\overline{\mathbf{V}}) \right) dv + \int_{\partial K} (G_{\mathbf{n}} \mathbf{W})^t \, \overline{\mathbf{V}} ds = 0. \qquad (4)$$

Cette formulation faible de l'équation aux dérivées partielles sur le sous-ensemble K est le point de départ de la méthode de type Galerkin discontinu. Supposons maintenant que le domaine de calcul est découpé en un ensemble d'éléments (tétraèdres ou hexaèdres) ; on considère ici uniquement des maillages conformes. Ce maillage du domaine de calcul Ω est noté $\mathcal{T}_h$ et on a :

$$\bigcup_{K\in\mathcal{T}_h} \overline{K} = \overline{\Omega_h}.$$

L'approximation numérique $\mathbf{W}_h = \begin{pmatrix} \mathbf{E}_h \\ \mathbf{H}_h \end{pmatrix}$ de la solution du problème (3) est recherchée dans l'espace $V_h \times V_h$ où V_h est défini par :

$$V_h = \left\{ \mathbf{V} \in [L^2(\Omega)]^3 \mid \forall K \in \mathcal{T}_h,\ Ve_{|K} \in \mathcal{P}(K) \right\}. \tag{5}$$

Le terme $\mathcal{P}(K)$ désigne un espace de champs à composantes polynomiales sur l'élément K. Les champs de vecteurs tests appartiennent aussi à $V_h \times V_h$.

Flux principal. On part de l'équation (4) considérée sur un élément K de $\mathcal{T}_h$ mais désormais l'approximation numérique $\mathbf{W}_h$ remplace $\mathbf{W}$ et le terme de bord est une application à définir $\Phi_{\partial K}$, appelée flux principal pour reprendre les termes employés par Ern et Guermond dans [EG06a, EG06b]. On souhaite alors vérifier :

$$\int_K (\mathrm{i}\omega G_0 \mathbf{W}_h)^t \overline{\mathbf{V}} dv - \int_K \mathbf{W}_h^t \left(\sum_{l \in \{x,y,z\}} G_l \partial_l(\overline{\mathbf{V}}) \right) dv + \int_{\partial K} (\Phi_{\partial K}(\mathbf{W}_h))^t \overline{\mathbf{V}} ds = 0, \quad \forall \mathbf{V} \in V_h \times V_h. \tag{6a}$$

La définition de $\Phi_{\partial K}$ doit permettre d'assurer la consistence asymptotique et la stabilité de la méthode d'approximation utilisée. Pour respecter ces contraintes, le flux peut être alors défini sur une face F de ∂K par :

$$\Phi_{\partial K}(\mathbf{W}_h) = \begin{cases} I_{FK} S_F [\![\mathbf{W}_h]\!]_F + I_{FK} G_{\mathbf{n}_F} \{\mathbf{W}_h\} \text{ si } F \in \Gamma^0, \\ \dfrac{1}{2}(M_{F,K} + I_{FK} G_{\mathbf{n}_F}) \mathbf{W}_h \text{ si } F \in \Gamma^m, \\ \dfrac{1}{2}(M_{F,K} + I_{FK} G_{\mathbf{n}_F}) \mathbf{W}_h - \dfrac{1}{2}(M_{F,K} - I_{FK} G_{\mathbf{n}_F}) \mathbf{W}^{\mathrm{inc}} \text{ si } F \in \Gamma^a. \end{cases} \tag{6b}$$

Pour comprendre l'expression de ce flux principal, voici les éléments manquants :

- on introduit une matrice I pour prendre en compte les conventions d'orientation entre les faces et les éléments ; on l'appelle généralement matrice d'incidence face-élément. Le nombre de lignes de I est égal au nombre de faces du maillage et son nombre de colonnes au nombre d'éléments. La normale $\mathbf{n}_K$ sortante de l'élément K induit une orientation pour la frontière de cet élément et chaque face F a une orientation propre qui détermine la direction de sa normale $\mathbf{n}_F$. Les entrées de I sont alors définies la manière suivante :

$$I_{FK} = \begin{cases} 0 \text{ si la face } F \text{ n'appartient pas à l'élément } K, \\ 1 \text{ si } F \in K \text{ et les orientations coïncident } (\mathrm{signe}(\mathbf{n}_F^t \mathbf{n}_K) = 1), \\ -1 \text{ si } F \in K \text{ et les orientations ne coïncident pas } (\mathrm{signe}(\mathbf{n}_F^t \mathbf{n}_K) = -1). \end{cases}$$

– pour la face F intersection de K et $\tilde{K}$, on définit respectivement le saut du champ de vecteurs $\mathbf{V}$ et sa moyenne sur la face F :

$$[\![\mathbf{V}]\!]_F = I_{FK}\mathbf{V}_{|K} + I_{F\tilde{K}}\mathbf{V}_{|\tilde{K}} \text{ et } \{\mathbf{V}\} = \frac{1}{2}(\mathbf{V}_{|K} + \mathbf{V}_{|\tilde{K}}).$$

La notation $\mathbf{V}_{|K}$ désigne la restriction du champ de vecteurs $\mathbf{V}$ à l'élément K.

– la matrice S_F est une matrice permettant de pénaliser le saut du champ ou de certaines de ses composantes sur la face F séparant deux éléments. La matrice $M_{F,K}$ est une matrice qui assure la consistence asymptotique avec les conditions aux limites du problème continu. Des définitions possibles de ces matrices sont donnés en sous-section 2.2.2.

– les ensembles Γ^0, Γ^m et Γ^a désignent respectivement l'ensemble des faces internes, des faces appartenant à Γ_a et des faces appartenant à Γ_m.

Flux adjoint. Si l'on revient à l'opérateur initial, on peut aussi introduire la notion de flux adjoint $\tilde{\Phi}_{\partial K}$ (voir [EG06a, EG06b]) et dans ce cas on souhaite vérifier :

$$\int_K (\mathrm{i}\omega G_0 \mathbf{W}_h)^t \overline{\mathbf{V}} dv + \int_K \left(\sum_{l\in\{x,y,z\}} G_l \partial_l(\mathbf{W}_h) \right)^t \overline{\mathbf{V}} dv + \int_{\partial K} \left(\tilde{\Phi}_{\partial K}(\mathbf{W}_h) \right)^t \overline{\mathbf{V}} = 0, \quad \forall \mathbf{V} \in V_h \times V_h. \tag{7a}$$

Le flux adjoint est défini sur une face F de ∂K par :

$$\tilde{\Phi}_{\partial K}(\mathbf{W}_h) = \begin{cases} I_{FK} S_F [\![\mathbf{W}_h]\!]_F - \frac{1}{2} G_{\mathbf{n}_F} [\![\mathbf{W}_h]\!]_F \text{ si } F \in \Gamma^0 \\ \frac{1}{2}(M_{F,K} - I_{FK} G_{\mathbf{n}_F})\mathbf{W}_h \text{ si } F \in \Gamma^m, \\ \frac{1}{2}(M_{F,K} - I_{FK} G_{\mathbf{n}_F})(\mathbf{W}_h - \mathbf{W}^{\mathrm{inc}}) \text{ si } F \in \Gamma^a. \end{cases} \tag{7b}$$

2.2.2 Quelques flux numériques possibles

On rappelle ici quelques flux numériques simples dont l'utilisation est comparée dans la suite pour approcher la solution du problème (3) :

– le **flux centré** [FLLP05] correspond au choix pour toutes faces F :

$$S_F = 0, \tag{8a}$$

et au choix pour les faces F appartenant à Γ^m ou à Γ^a :

$$M_{F,K} = \begin{cases} I_{FK} \begin{pmatrix} 0_{3\times 3} & N_{\mathbf{n}_F} \\ -N^t_{\mathbf{n}_F} & 0_{3\times 3} \end{pmatrix} \text{ si } F \text{ appartient à } \Gamma^m, \\ |G_{\mathbf{n}_F}| \text{ si } F \text{ appartient à } \Gamma^a. \end{cases} \tag{8b}$$

– le **flux décentré du premier ordre** [EG06a, Pip00] correspond au choix :

$$S_F = \begin{pmatrix} \alpha^E_F N_{\mathbf{n}_F} N^t_{\mathbf{n}_F} & 0_{3\times 3} \\ 0_{3\times 3} & \alpha^H_F N^t_{\mathbf{n}_F} N_{\mathbf{n}_F} \end{pmatrix}, \tag{9a}$$

avec $\alpha^E_F = 1/2$ et $\alpha^H_F = 1/2$ si le matériau est homogène (sinon la forme donnée ici n'est qu'une variante simplifiée du flux décentré du premier ordre), et au changement de $M_{F,K}$ sur Γ_m pour prendre la forme suivante :

$$M_{F,K} = \begin{pmatrix} \eta_F N_{\mathbf{n}_F} N^t_{\mathbf{n}_F} & I_{FK} N_{\mathbf{n}_F} \\ -I_{FK} N^t_{\mathbf{n}_F} & 0_{3\times 3} \end{pmatrix}. \tag{9b}$$

– on peut choisir de décentrer uniquement suivant l'une des variables (cas particulier de la méthode Galerkin discontinue locale [CS98]). L'amplitude du coefficient de pénalisation utilisé est alors beaucoup plus importante, généralement proportionnelle à h_F^{-1} où h_F désigne la mesure de la face F. On a ainsi :

$$S_F = \tau_F h_F^{-1} \begin{pmatrix} N_{\mathbf{n}_F} N^t_{\mathbf{n}_F} & 0_{3\times 3} \\ 0_{3\times 3} & 0_{3\times 3} \end{pmatrix}, \tag{10a}$$

et pour F sur Γ^m on effectue aussi le changement :

$$M_{F,K} = \begin{pmatrix} \tau_F h_F^{-1} N_{\mathbf{n}_F} N^t_{\mathbf{n}_F} & I_{FK} N_{\mathbf{n}_F} \\ -I_{FK} N^t_{\mathbf{n}_F} & 0_{3\times 3} \end{pmatrix}. \tag{10b}$$

2.2.3 Problème global

Formulation avec l'opérateur principal. À partir de la formulation du flux adjoint (7), on effectue la somme sur tous les éléments K de $\mathcal{T}^h$ et la formulation faible du problème s'écrit alors :

$$\left\{\begin{array}{l} \text{Trouver } \mathbf{W}_h \text{ dans } V_h \times V_h \text{ tel que :} \\ \displaystyle\int_{\Omega_h} (\mathrm{i}\omega G_0 \mathbf{W}_h)^t \overline{\mathbf{V}} dv + \sum_{K\in\mathcal{T}_h} \int_K \left(\sum_{l\in\{x,y,z\}} G_l \partial_l(\mathbf{W}_h) \right)^t \overline{\mathbf{V}} dv \\ \displaystyle + \sum_{F\in\Gamma^m\cup\Gamma^a} \int_F \left(\frac{1}{2}(M_{F,K} - I_{FK} G_{\mathbf{n}_F}) \mathbf{W}_h \right)^t \overline{\mathbf{V}} ds \\ \displaystyle - \sum_{F\in\Gamma^0} \int_F (G_{\mathbf{n}_F} [\![\mathbf{W}_h]\!]_F)^t \{\overline{\mathbf{V}}\} ds + \sum_{F\in\Gamma^0} \int_F (S_F [\![\mathbf{W}_h]\!]_F)^t [\![\overline{\mathbf{V}}]\!]_F ds \\ \displaystyle = \sum_{F\in\Gamma^a} \int_F \left(\frac{1}{2}(M_{F,K} - I_{FK} G_{\mathbf{n}_F}) \mathbf{W}^{\mathrm{inc}} \right)^t \overline{\mathbf{V}} ds, \ \forall \mathbf{V} \in V_h \times V_h. \end{array}\right. \tag{11}$$

Cette formulation est la reformulation de [EG06a, Équation (4.12)] pour le problème (3).

Formulation avec l'opérateur adjoint. À partir de la formulation du flux principal (6), on effectue la somme sur tous les éléments K du $\mathcal{T}_h$ et la formulation faible du problème s'écrit alors aussi :

$$\left\{\begin{array}{l} \text{Trouver } \mathbf{W}_h \text{ dans } V_h \times V_h \text{ tel que :} \\ \displaystyle\int_{\Omega_h} (\mathrm{i}\omega G_0 \mathbf{W}_h)^t \overline{\mathbf{V}} dv - \sum_{K\in\mathcal{T}_h} \int_K \mathbf{W}_h^t \left(\sum_{l\in\{x,y,z\}} G_l \partial_l(\overline{\mathbf{V}}) \right) dv \\ \displaystyle + \sum_{F\in\Gamma^m\cup\Gamma^a} \int_F \left(\frac{1}{2}(M_{F,K} + I_{FK} G_{\mathbf{n}_F}) \mathbf{W}_h \right)^t \overline{\mathbf{V}} ds \\ \displaystyle + \sum_{F\in\Gamma^0} \int_F (G_{\mathbf{n}_F} \{\mathbf{W}_h\})^t [\![\overline{\mathbf{V}}]\!]_F ds + \sum_{F\in\Gamma^0} \int_F (S_F [\![\mathbf{W}_h]\!]_F)^t [\![\overline{\mathbf{V}}]\!]_F ds \\ \displaystyle = \sum_{F\in\Gamma^a} \int_F \left(\frac{1}{2}(M_{F,K} - I_{FK} G_{\mathbf{n}_F}) \mathbf{W}^{\mathrm{inc}} \right)^t \overline{\mathbf{V}} ds, \ \forall V \in V_h \times V_h. \end{array}\right. \tag{12}$$

Cette formulation est la reformulation de [EG06a, Équation (4.26)] pour le problème (3).

2.2.4 Formulation mixte du problème

Écriture directe depuis (11) et (12). Afin d'énoncer les techniques envisagées pour la résolution du problème et s'appuyant sur les résultats existants pour des systèmes de type point-selle, on présente la formulation mixte du problème à résoudre. Pour écrire celle-ci, on peut utiliser des équations de la formulation avec l'opérateur principal et avec l'opérateur adjoint. Si l'on considère tout d'abord tous les champs tests de la forme $\begin{pmatrix} 0_{3\times 1} \\ \mathbf{G} \end{pmatrix}$ dans la formulation avec l'opérateur principal (11), les champs solutions $(\mathbf{E}_h, \mathbf{H}_h)$ doivent vérifier :

$$a(\mathbf{H}_h, \mathbf{G}) + \overline{b(\mathbf{G}, \mathbf{E}_h)} = \sum_{F\in\Gamma^a} \int_F \frac{1}{2} \left(N^t_{\mathbf{n}_F} N_{\mathbf{n}_F} \mathbf{H}^{\text{inc}} - N^t_{\mathbf{n}_F} \mathbf{E}^{\text{inc}} \right)^t \overline{\mathbf{G}} ds, \ \forall \mathbf{G} \in V_h, \tag{13}$$

avec les formes sesquilinéaires a et b définies pour $(\mathbf{F}, \mathbf{G}) \in V_h \times V_h$ par :

$$\begin{aligned} a(\mathbf{F}, \mathbf{G}) = & \int_{\Omega_h} \mathrm{i}\omega\mu \mathbf{F}^t \overline{\mathbf{G}} dv + \sum_{F\in\Gamma^0} \int_F \left(\alpha^H_F N^t_{\mathbf{n}_F} N_{\mathbf{n}_F} [\![\mathbf{F}]\!]_F \right)^t [\![\overline{\mathbf{G}}]\!]_F ds \\ & + \sum_{F\in\Gamma^a} \int_F \left(\frac{1}{2} N^t_{\mathbf{n}_F} N_{\mathbf{n}_F} \mathbf{F} \right)^t \overline{\mathbf{G}} ds, \end{aligned} \tag{14}$$

et aussi :

$$\begin{aligned} b(\mathbf{F}, \mathbf{G}) = & \sum_{K\in\mathcal{T}_h} \int_K \mathbf{F}^t \left(\sum_{l\in\{x,y,z\}} N^t_{\mathbf{e}_l} \partial_l(\overline{\mathbf{G}}) \right) dv - \sum_{F\in\Gamma^0} \int_F \left(N_{\mathbf{n}_F} \{\mathbf{F}\} \right)^t [\![\overline{\mathbf{G}}]\!]_F ds \\ & - \sum_{F\in\Gamma^m} \int_F (I_{KF} N_{\mathbf{n}_F} \mathbf{F})^t \overline{\mathbf{G}} ds - \sum_{F\in\Gamma^a} \int_F \left(\frac{1}{2} I_{KF} N_{\mathbf{n}_F} \mathbf{F} \right)^t \overline{\mathbf{G}} ds. \end{aligned} \tag{15}$$

Ensuite, si l'on considère toutes les fonctions tests de la forme $\begin{pmatrix} \mathbf{F} \\ 0_{3\times 1} \end{pmatrix}$ dans la formulation avec l'opérateur adjoint (12), les solutions $(\mathbf{E}_h, \mathbf{H}_h)$ doivent vérifier :

$$c(\mathbf{E}_h, \mathbf{F}) - b(\mathbf{H}_h, \mathbf{F}) = \sum_{F\in\Gamma^a} \int_F \frac{1}{2} \left(N_{\mathbf{n}_F} N^t_{\mathbf{n}_F} \mathbf{E}^{\text{inc}} - N_{\mathbf{n}_F} \mathbf{H}^{\text{inc}} \right)^t \overline{\mathbf{F}} ds, \ \forall \mathbf{F} \in V_h, \tag{16}$$

avec la forme sesquilinéaire c définie pour $(\mathbf{F}, \mathbf{G}) \in V_h \times V_h$ par :

$$\begin{aligned} c(\mathbf{F}, \mathbf{G}) = & \int_{\Omega_h} \mathrm{i}\omega\varepsilon \mathbf{F}^t \overline{\mathbf{G}} dv + \sum_{F\in\Gamma^0} \int_F \left(\alpha^E_F N_{\mathbf{n}_F} N^t_{\mathbf{n}_F} [\![\mathbf{F}]\!]_F \right)^t [\![\overline{\mathbf{G}}]\!]_F ds \\ & + \sum_{F\in\Gamma^m} \int_F \left(\frac{1}{2} \eta_F N_{\mathbf{n}_F} N^t_{\mathbf{n}_F} \mathbf{F} \right)^t \overline{\mathbf{G}} ds + \sum_{F\in\Gamma^a} \int_F \left(\frac{1}{2} N_{\mathbf{n}_F} N^t_{\mathbf{n}_F} \mathbf{F} \right)^t \overline{\mathbf{G}} ds. \end{aligned} \tag{17}$$

Le problème peut donc s'écrire sous la forme du problème mixte suivant :

Trouver $(\mathbf{E}_h, \mathbf{H}_h) \in V_h \times V_h$ tel que :

$$\begin{cases} a(\mathbf{H}_h, \mathbf{G}) + \overline{b(\mathbf{G}, \mathbf{E}_h)} &= \displaystyle\sum_{F\in\Gamma^a} \int_F \frac{1}{2} \left(N_{\mathbf{n}_F}^t N_{\mathbf{n}_F} \mathbf{H}^{\text{inc}} - N_{\mathbf{n}_F}^t \mathbf{E}^{\text{inc}}\right)^t \overline{\mathbf{G}} ds, \ \forall \mathbf{G} \in V_h, \\ -b(\mathbf{H}_h, \mathbf{F}) + c(\mathbf{E}_h, \mathbf{F}) &= \displaystyle\sum_{F\in\Gamma^a} \int_F \frac{1}{2} \left(N_{\mathbf{n}_F} N_{\mathbf{n}_F}^t \mathbf{E}^{\text{inc}} - N_{\mathbf{n}_F} \mathbf{H}^{\text{inc}}\right)^t \overline{\mathbf{F}} ds, \ \forall \mathbf{E} \in V_h. \end{cases}$$

C'est autour de cette écriture du problème que nous allons travailler dans la section 4.

Notations classiques. Pour faciliter l'analogie avec les résultats des articles tels que [BP06, BHI07], on écrit aussi le problème avec des notations vectorielles plus usuelles. Soit un champ $\mathbf{G}$ définit sur le domaine Ω. Pour la face interne F séparant deux éléments K et $\tilde{K}$ de $\mathcal{T}_h$, on définit le saut de la composante tangentielle de $\mathbf{G}$ sur la face F :

$$[\![\mathbf{G}]\!]_T = \mathbf{G}_{|K} \times \mathbf{n}_K + \mathbf{G}_{|\tilde{K}} \times \mathbf{n}_{\tilde{K}} = N_{\mathbf{n}_F} [\![\mathbf{G}]\!]_F = -N_{\mathbf{n}_F}^t [\![\mathbf{G}]\!]_F.$$

Pour les faces situées sur la frontière du domaine, cette grandeur est aussi définie et devient : $[\![\mathbf{G}]\!]_T = \mathbf{G} \times \mathbf{n}$. De même, on prolonge la définition de la valeur moyenne du champ pour les faces F de la frontière : $\{\mathbf{G}\} = \mathbf{G}$. Les différentes formes bilinéaires définies précédemment admettent donc aussi l'écriture suivante :

$$\begin{aligned} a(\mathbf{F}, \mathbf{G}) = &\int_{\Omega_h} \mathrm{i}\omega\mu \mathbf{F}^t \overline{\mathbf{G}} dv + \sum_{F\in\Gamma^0} \int_F \alpha_F^H [\![\mathbf{F}]\!]_T^t [\![\overline{\mathbf{G}}]\!]_T ds \\ &+ \sum_{F\in\Gamma^a} \int_F \frac{1}{2} [\![\mathbf{F}]\!]_T^t [\![\overline{\mathbf{G}}]\!]_T ds, \end{aligned} \tag{18}$$

$$\begin{aligned} b(\mathbf{F}, \mathbf{G}) = &\sum_{K\in\mathcal{T}_h} \int_K \mathbf{F}^t \operatorname{rot}(\overline{\mathbf{G}}) dv + \sum_{F\in\Gamma^0} \int_F \{\mathbf{F}\}^t [\![\overline{\mathbf{G}}]\!]_T ds \\ &+ \sum_{F\in\Gamma^m} \int_F \{\mathbf{F}\}^t [\![\overline{\mathbf{G}}]\!]_T ds + \sum_{F\in\Gamma^a} \int_F \frac{1}{2} \{\mathbf{F}\}^t [\![\overline{\mathbf{G}}]\!]_T ds, \end{aligned} \tag{19}$$

et enfin :

$$\begin{aligned} c(\mathbf{F}, \mathbf{G}) = &\int_{\Omega_h} \mathrm{i}\omega\varepsilon \mathbf{F}^t \overline{\mathbf{G}} dv + \sum_{F\in\Gamma^0} \int_F \alpha_F^E [\![\mathbf{F}]\!]_T^t [\![\overline{\mathbf{G}}]\!]_T ds \\ &+ \sum_{F\in\Gamma^m} \frac{1}{2} \int_F \eta_F [\![\mathbf{F}]\!]_T^t [\![\overline{\mathbf{G}}]\!]_T ds + \sum_{F\in\Gamma^a} \int_F \frac{1}{2} [\![\mathbf{F}]\!]_T^t [\![\overline{\mathbf{G}}]\!]_T ds. \end{aligned} \tag{20}$$

3 Éléments de discussion pour le choix du flux numérique

Pour le choix des flux numériques lors de la discrétisation du problème (3), on va privilégier les méthodes qui garantissent sur un maillage **non structuré** :

- le meilleur ordre de convergence asymptotique de l'erreur en norme L^2 pour le champ électrique et/ou le champ magnétique (avec un ordre d'approximation polynomiale fixe),
- une dispersion numérique minimale.

L'ordre de convergence semble être le critère prépondérant, l'étude de la dispersion numérique venant dans un second temps pour optimiser l'approximation. Concernant la dispersion numérique pour les méthodes de type Galerkin discontinu, on pourra consulter [Ain04]. L'ordre de convergence asymptotique de l'erreur en norme L^2 entre la solution exacte $(\mathbf{E}, \mathbf{H})$ et son approximation numérique $(\mathbf{E}_h, \mathbf{H}_h)$ correspond à connaître les coefficients réels δ et γ les plus grands possibles tels que :

$$\exists C_1, C_2 > 0,\ \exists h_0 > 0,\ \forall h > h_0,\ \|\mathbf{E} - \mathbf{E}_h\|_{L^2(\Omega)} \leq C_1 h^{\delta} \text{ et } \|\mathbf{H} - \mathbf{H}_h\|_{L^2(\Omega)} \leq C_2 h^{\gamma}. \quad (21)$$

Le paramètre h désigne le *pas* du maillage $\mathcal{T}^h$ utilisé pour le calcul. Les coefficients δ et γ décrivent donc à quelle vitesse on converge vers la solution exacte.

3.1 Résultats concernant l'ordre de convergence

3.1.1 Résultats existants pour la convergence

On choisit $\mathcal{P}(K) = [P_k(K)]^3$ où l'entier k correspond à l'ordre de l'espace polynomial utilisé. Cet ordre k est constant sur l'ensemble des éléments du maillage. On donne ici quelques résultats liés à l'ordre de convergence et aux équations de Maxwell (mais pas nécéssairement en régime harmonique). On rappelle en particulier dans le tableau 1 les ordres de convergence asymptotiques théoriques pour le système de Maxwell elliptique [EG06a, EG06b] quand la solution recherchée est suffisament régulière.

flux	centré (8)	decentré (9)	pénalisation de $\mathbf{E}$ (10)
champ $\mathbf{E}$	k	$k+1/2$	$k+1$
champ $\mathbf{H}$	k	$k+1/2$	k

TAB. 1 – Ordre de convergence théorique en norme L^2 pour chaque champ avec la discrétisation du système de Maxwell elliptique par une méthode de type Galerkin discontinu et les flux numériques de la sous-section 2.2.2.

Les résultats actuels pour le problème harmonique se limitent à des problèmes de cavité (une conditions métallique est appliquée à toute la frontière du domaine de calcul). Dans ce cas, les ordres de convergence sont identiques au cas elliptique pour des méthodes de

pénalisation intérieure et de pénalisation du champ $\mathbf{E}$ sur la formulation mixte (flux (10)) ; voir [HPSS05, BP06]. Les résultats donnés dans le tableau 1 sont valables en deux et trois dimensions.

3.1.2 Estimation de l'ordre de convergence asymptotique sur un cas simple

On souhaite valider numériquement les résultats de convergence théoriques ou, si ces résultats n'existent pas, estimer les ordres de convergence que l'on peut s'attendre à trouver.

Le problème. On effectue un calcul pour simuler la propagation d'une onde plane dans le vide. L'onde plane incidente retenue est de la forme :

$$\begin{pmatrix} \mathbf{E}_x^{\text{inc}} \\ \mathbf{E}_y^{\text{inc}} \\ \mathbf{H}_z^{\text{inc}} \end{pmatrix} = \exp(-\mathrm{i}\omega x) \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}. \tag{22}$$

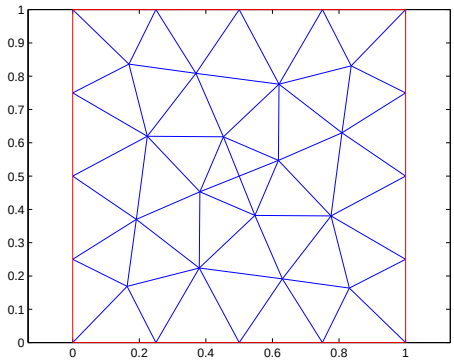
Le domaine est le carré unité, c'est-à-dire $\Omega =]0;1[^2$ et des conditions de Silver-Müller sont imposées sur l'ensemble de la frontière, c'est-à-dire $\Gamma_a = \partial\Omega$ et $\Gamma_m = \emptyset$. Comme on s'est placé dans le vide, les paramètres ε et μ valent 1 sur l'ensemble du domaine et la conductivité est nulle. Enfin, pour les essais effectués, on choisit $\omega = 2\pi$ et la longueur d'onde est donc égale au côté du domaine.

La méthode. On estime numériquement les taux de convergence asymptotique dans l'exemple évoqué ci-dessus. On suppose pour l'évolution de l'erreur en norme L^2 un comportement asymptotique en $C_{\mathbf{E}} h^{\delta}$ pour le champ $\mathbf{E}$ et en $C_{\mathbf{H}} h^{\gamma}$ pour le champ $\mathbf{H}$. Les coefficients $C_{\mathbf{E}}$ et $C_{\mathbf{H}}$ sont des constantes et δ et γ sont les ordres de convergence asymptotiques recherchés. Numériquement il s'agit donc de faire varier le paramètre h et d'estimer ces coefficients δ et γ à partir de l'évolution de l'erreur en norme L^2.

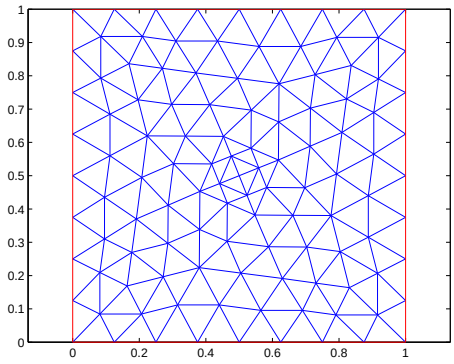
Remarque : au lieu d'observer l'évolution suivant le paramètre h, on peut, pour des maillages quasi-uniformes, regarder de manière équivalente l'évolution suivant la racine carré du nombre de degrés de liberté en deux dimensions. Ce choix a été fait pour les résultats numériques dans la suite.

On va réaliser cette démarche pour deux séries de maillages :

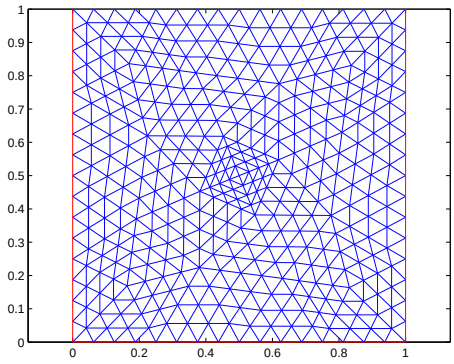
Cas 1. On utilise pour le premier calcul le maillage initial non structuré présenté à la figure 1(a). Ce maillage va être ensuite raffiné de manière uniforme ; les premiers raffinements sont donnés aux figures 1(b) et 1(c). Notons déjà qu'avec cette taille de maille initial, le problème sera sous-résolu pour les ordres polynomiaux les plus faibles (particulièrement pour le cas P_0 en fait). Néanmoins, ceci est sans conséquence puisque c'est le comportement asymptotique qui nous intéresse.



(a) Maillage initial.



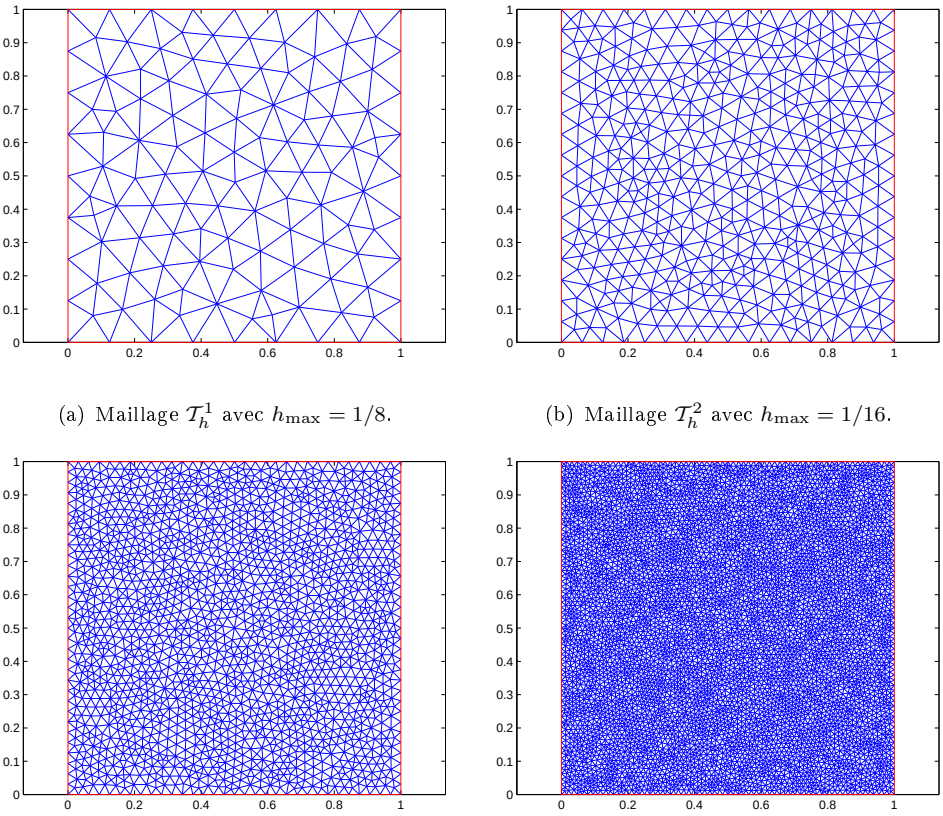
(b) Premier raffinement.



(c) Second raffinement.

FIG. 1 – Maillage initial du carré unité et les deux premiers raffinements uniformes.

Cas 2. On utilise quatre maillages non structurés quasi-uniformes dont on a imposé le pas maximal $h_{\max}$; ces quatre maillages sont présentés à la figure 2. Ce paramètre $h_{\max}$ est divisé par deux pour passer d'un maillage à l'autre. Ces maillages sont notés $\mathcal{T}_h^i$ pour i allant de 1 à 4 avec $h_{\max}$ décroissant. Ainsi aucun maillage $\mathcal{T}_h^{i+1}$ ne correspond au raffinement du maillage $\mathcal{T}_h^i$.



(a) Maillage $\mathcal{T}_h^1$ avec $h_{\max} = 1/8$.

(b) Maillage $\mathcal{T}_h^2$ avec $h_{\max} = 1/16$.

(c) Maillage $\mathcal{T}_h^3$ avec $h_{\max} = 1/32$.

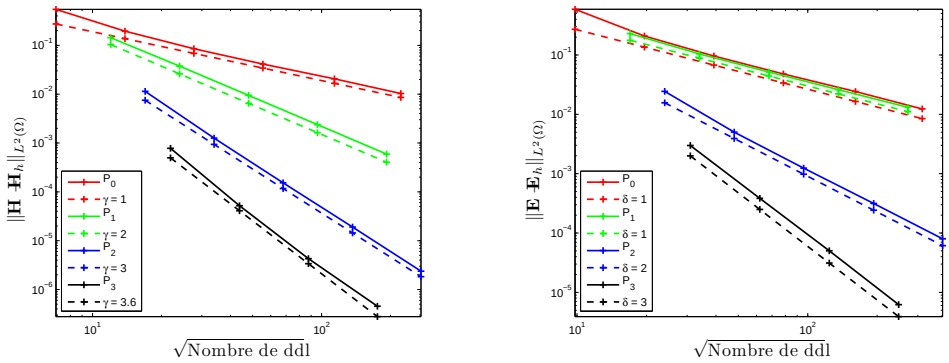
(d) Maillage $\mathcal{T}_h^4$ avec $h_{\max} = 1/64$.

FIG. 2 – Quatre maillages non structurés utilisés pour le calcul.

Remarque : pour des questions d'intégration numérique concernant la prise en compte de l'onde incidente, les résultats obtenus pour le cas P_3 sont sujets à caution ; il sont néanmoins cohérents avec le reste des résultats.

3.1.3 Convergence avec les maillages obtenus par raffinement uniforme

Résultats avec l'utilisation de flux centrés. Les courbes des résultats obtenus avec l'utilisation des flux centrés sont données à la figure 3. Elles représentent l'évolution de l'erreur en norme L^2 pour les calculs des champs $\mathbf{E}$ et $\mathbf{H}$ en fonction de la racine carrée du nombre de degrés de liberté (ddl). À partir de la figure 3, on perçoit l'intérêt de l'ordre élevé : on constate ainsi qu'une augmentation de l'ordre permet une réduction drastique du nombre de degrés de liberté pour atteindre une même précision. Partant ainsi d'un même maillage, la convergence vers la solution est beaucoup plus rapide (convergence exponentielle) en augmentant l'ordre qu'en raffinant le maillage (convergence polynomiale à degré fixe). Bien évidemment, cela n'est vrai que si la solution recherchée présente une régularité suffisante. En outre, on ne débat pas non plus ici de l'impact de l'ordre d'approximation sur la matrice du problème (matrice moins creuse, de conditionnement différent mais de bien plus petite dimension aussi). Une régression linéaire nous permet d'estimer numériquement l'ordre de convergence asymptotique pour les différents ordres polynomiaux ; les résultats sont donnés dans le tableau 2.



(a) Évolution de $\|\mathbf{H}-\mathbf{H}_h\|_{L^2}$ en fonction de la racine carré du nombre de degrés de liberté. Échelle logarithmique.

(b) Évolution de $\|\mathbf{E}-\mathbf{E}_h\|_{L^2}$ en fonction de la racine carré du nombre de degrés de liberté. Échelle logarithmique.

FIG. 3 – Résultats de convergence avec l'utilisation de flux centrés. Les courbes en ligne continu correspondent aux calculs pour les différents ordres polynomiaux. Les courbes en pointillés reprennent l'ordre de convergence calculé par une régression linéaire (voir le tableau 2).

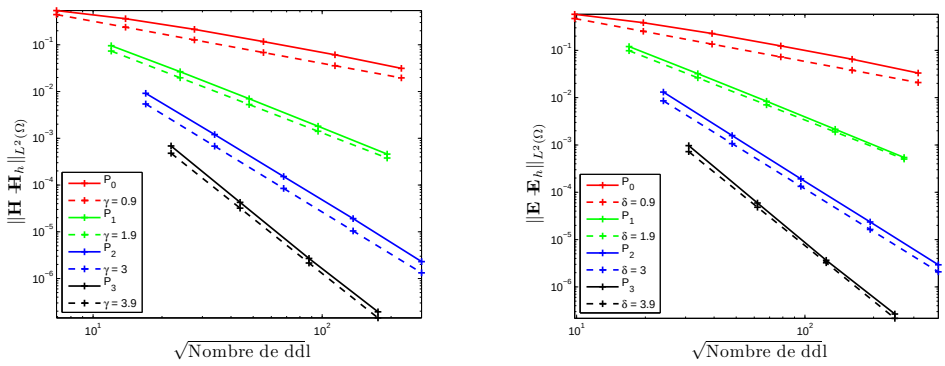
	P_0	P_1	P_2	P_3
$\mathbf{E}$	1	1	2	3
$\mathbf{H}$	1	2	3	3.6

TAB. 2 – Estimation numérique de l'ordre de convergence asymptotique. Cas flux centré.

Le cas P_0 est un peu particulier : l'ordre de convergence est optimal pour les champs $\mathbf{E}$ et $\mathbf{H}$, c'est-à-dire égal à $k+1$. Cela est sans doute dû à l'utilisation d'une série de maillages particulière obtenue par raffinement uniforme (cette affirmation est corroborée par les résultats obtenus dans la sous-section 3.1.4).

Pour les ordres polynomiaux d'approximation strictement positifs, l'ordre de convergence pour la norme L^2 de l'erreur sur $\mathbf{E}$ coïncide avec l'ordre de convergence annoncé pour le système de Maxwell elliptique, c'est-à-dire k. L'ordre de convergence pour $\mathbf{H}$ est par contre optimal. Pour cet exemple particulier et avec le choix des flux centrés, le champ magnétique est donc mieux approché numériquement que le champ électrique.

Résultats avec l'utilisation de flux décentrés. On a utilisé dans (9) les paramètres $\alpha_F^H = \alpha_F^E = \eta_F = 1$ pour toutes les faces F. Les courbes des résultats obtenus avec l'utilisation des flux décentrés sont donnés à la figure 4. On peut faire des remarques similaires au cas des flux centrés. On note en outre que les propriétés de convergence pour les interpolations P_0 et P_1 sont cette fois clairement distinctes contrairement au cas des flux centrés. Une régression linéaire permet d'estimer numériquement l'ordre de convergence asymptotique pour les différents ordres polynomiaux ; les résultats sont donnés dans le tableau 3. L'ordre de convergence est cette fois similaire pour les deux champs ; cela coïncide avec les résultats théoriques pour le système de Maxwell elliptique et correspond à la symétrie de la démarche. Les ordres de convergence sont optimaux pour les deux variables (dans ce cas particulier) sauf pour le cas P_0 où l'on est néanmoins au-dessus des estimations théoriques.



(a) Évolution de $\|\mathbf{H}-\mathbf{H}_h\|_{L^2}$ en fonction de la racine carré du nombre de degrés de liberté. Échelle logarithmique.

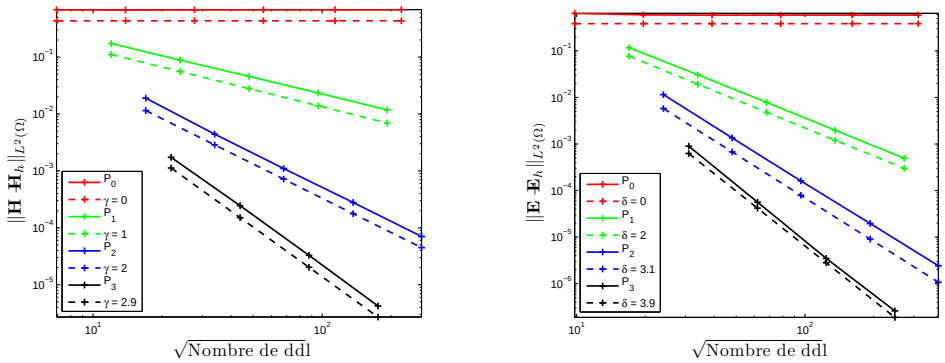
(b) Évolution de $\|\mathbf{E}-\mathbf{E}_h\|_{L^2}$ en fonction de la racine carré du nombre de degrés de liberté. Échelle logarithmique.

FIG. 4 – Résultats de convergence avec l'utilisation de flux décentrés. Les courbes en ligne continu correspondent aux calculs pour les différents ordres polynomiaux. Les courbes en pointillés reprennent l'ordre de convergence calculé par une régression linéaire (voir le tableau 3).

	P_0	P_1	P_2	P_3
E	0.9	1.9	3	3.9
H	0.9	1.9	3	3.9

TAB. 3 – Estimation numérique de l'ordre de convergence asymptotique. Cas flux décentré.

Résultats avec l'utilisation de flux avec pénalisation sur E. On a utilisé dans (10) les paramètres $\tau_F = \eta_F = 1$ pour toutes les faces F. Les courbes des résultats obtenus avec l'utilisation des flux avec pénalisation sur **E** sont donnés à la figure 5. On note en particulier l'absence de convergence pour le cas P_0. Une régression linéaire permet d'estimer numériquement l'ordre de convergence asymptotique pour les différents ordres polynomiaux ; les résultats sont donnés dans le tableau 4. Outre la non-convergence (attendue) pour le cas P_0, on note pour les $(P_k)_{k>0}$, un comportement complémentaire du cas avec flux centrés puisque cette fois l'ordre de convergence est optimal pour le champ **E** mais plus pour le champ **H**.



(a) Évolution de $\|\mathbf{H}-\mathbf{H}_h|_{L^2}$ en fonction de la racine carré du nombre de degrés de liberté. Échelle logarithmique.

(b) Évolution de $\|\mathbf{E}-\mathbf{E}_h\|_{L^2}$ en fonction de la racine carré du nombre de degrés de liberté. Échelle logarithmique.

FIG. 5 – Résultats de convergence avec l'utilisation de flux pénalisés sur **E**. Les courbes en ligne continu correspondent aux calculs pour les différents ordres polynomiaux. Les courbes en pointillés reprennent l'ordre de convergence calculé par une régression linéaire (voir le tableau 4).

3.1.4 Convergence avec la série de maillages distincts

On condense l'étude de l'ordre de convergence pour cette série de maillages. Soit err_i^E l'erreur en norme L^2 pour le calcul du champ **E** sur la maillage $\mathcal{T}_h^i$. De même, err_i^H désigne le même type d'erreur pour le champ **H**. On propose deux tableaux donnant une estimation

	P_0	P_1	P_2	P_3
E	X	2	3.1	3.9
H	X	1	2.02	2.9

TAB. 4 – Estimation numérique de l'ordre de convergence asymptotique. Cas flux pénalisé sur **E**. X signifie que la méthode ne converge pas.

de l'ordre de convergence à partir des calculs sur deux maillages consécutifs, c'est-à-dire la quantité :

$$\frac{\log\left(\frac{\mathrm{err}_{i+1}^E}{\mathrm{err}_i^E}\right)}{\log(0.5)} \text{ pour le champ } \mathbf{E} \text{ et } \frac{\log\left(\frac{\mathrm{err}_{i+1}^H}{\mathrm{err}_i^H}\right)}{\log(0.5)} \text{ pour le champ } \mathbf{H}.$$

Les résultats dans le cas de l'utilisation de flux décentrés sont donnés dans les tableaux 5. Ils coïncident avec ceux obtenus avec les maillages raffinés uniformément. Le comportement semble donc non lié au raffinement uniforme utilisé précédemment. Les résultats pour les flux centrés sont donnés dans les tableaux 6. On note l'absence de convergence pour le cas P_0. Pour les autres cas, cela correspond à ce que l'on avait obtenu avec le raffinement uniforme.

(a) Champ **E**

	$\mathcal{T}_h^1 \to \mathcal{T}_h^2$	$\mathcal{T}_h^2 \to \mathcal{T}_h^3$	$\mathcal{T}_h^3 \to \mathcal{T}_h^4$
P_0	0.78	0.87	0.94
P_1	1.92	1.97	1.99
P_2	3.00	3.0	3.04

(b) Champ **H**

	$\mathcal{T}_h^1 \to \mathcal{T}_h^2$	$\mathcal{T}_h^2 \to \mathcal{T}_h^3$	$\mathcal{T}_h^3 \to \mathcal{T}_h^4$
P_0	0.78	0.86	0.94
P_1	1.90	1.95	1.98
P_2	2.92	3.00	3.02

TAB. 5 – Ordre de convergence de $\mathcal{T}_h^i$ à $\mathcal{T}_h^{i+1}$. Flux décentré.

(a) Champ **E**

	$\mathcal{T}_h^1 \to \mathcal{T}_h^2$	$\mathcal{T}_h^2 \to \mathcal{T}_h^3$	$\mathcal{T}_h^3 \to \mathcal{T}_h^4$
P_0	0.37	0.08	0
P_1	1.02	0.98	0.99
P_2	2.03	2.02	2.03

(b) Champ **H**

	$\mathcal{T}_h^1 \to \mathcal{T}_h^2$	$\mathcal{T}_h^2 \to \mathcal{T}_h^3$	$\mathcal{T}_h^3 \to \mathcal{T}_h^4$
P_0	1.27	0.37	-0.06
P_1	2.00	1.96	2.00
P_2	2.97	3.02	3.02

TAB. 6 – Ordre de convergence de $\mathcal{T}_h^i$ à $\mathcal{T}_h^{i+1}$. Flux centré.

3.2 Résultats concernant le problème spectral

3.2.1 Passage en revue de l'existant

Plusieurs articles traitent de l'utilisation des méthodes de type Galerkin discontinu pour la résolution numérique de problèmes de modes propres, notamment dans le cas des équations

de Maxwell. Dans le cas de ces équations, le problème spectral continu peut s'écrire sous la forme suivante :

$$\begin{cases} \text{Trouver les triplets } (\mathbf{E}, \mathbf{H}, \omega) \text{ dans } \mathbf{H}(\mathrm{rot}, \Omega) \times \mathbf{H}(\mathrm{rot}, \Omega) \times \mathbb{R}^* \text{ tels que :} \\ \mathrm{i}\omega\varepsilon\mathbf{E} - \mathrm{rot}\,\mathbf{H} = 0, \\ \mathrm{i}\omega\mu\mathbf{H} + \mathrm{rot}\,\mathbf{E} = 0, \\ \mathbf{E} \times \mathbf{n} = 0 \text{ sur } \partial\Omega. \end{cases} \tag{23}$$

Le premier des articles concernant la résolution du problème (23) par une méthode de type Galerkin discontinu est dû à Hesthaven et Warburton [HW04]. Leur étude est basée essentiellement sur une série de tests numériques et montre les difficultés pour résoudre (23) lorsque l'on utilise des flux centrés. Pour comprendre ces difficultés, nous donnons brièvement quelques compléments sur l'approximation numérique du problème (23). Une méthode de discrétisation *correcte* de ce problème doit, selon [BP06, Introduction], vérifier les propriétés suivantes :

1. **Isolement du noyau discret**, c'est-à-dire que les valeurs propres discrètes approchant le spectre essentiel $\sigma^{\mathrm{ess}} = \{0\}$ sont séparées des autres valeurs propres.
2. **Non-pollution du spectre**, c'est-à-dire qu'il n'y a pas de valeurs propres discrètes non-physiques.
3. **Complétude du spectre**, c'est-à-dire que toutes les valeurs propres physiques plus petites qu'un nombre arbitraire fixé sont approchées, pour un maillage suffisament fin.
4. **Non-pollution et complétude des sous-espaces propres**, c'est-à-dire qu'il n'y a pas fonctions propres non-physiques et que l'approximation des sous-espaces propres dont les valeurs propres n'approchent pas le spectre essentiel ont la bonne dimension.

L'utilisation des flux centrés pour approcher (23) conduit à certains résultats dans [HW04] qui ne vérifient pas les conditions 2 ou 4. Les auteurs de [HW04] proposent comme alternative de pénaliser plus fortement le saut des composantes tangentielles des champs sur les interfaces entre éléments voisins et dans ce cas les résultats montrent que les modes propres discrets non-physiques disparaissent. Une étude plus étendue sur l'utilisation de cette pénalisation est présentée dans [WE06]. En particulier, les auteurs de [WE06] montrent les analogies entre l'approximation du spectre par une méthode de Galerkin discontinu avec la pénalisation et l'approximation de ce même spectre par une famille d'éléments finis conformes dans $\mathbf{H}(\mathrm{rot}, \Omega)$, la deuxième famille de Nédélec [N86].

Des résultats théoriques sont présentés dans [BP06, CN06] et expliquent pourquoi certaines méthodes de de type Galerkin discontinu fonctionnent correctement pour approcher numériquement le problème (23). Les résultats numériques illustrant ces démonstrations théoriques sont montrés dans [BHI07] où l'on considère aussi des exemples avec des maillages non conformes, qui ne sont pas pris en compte dans la théorie développée en [BP06]. Notons au passage que l'article des mêmes auteurs correspondant au calcul du spectre du Laplacien à l'aide de méthodes de type Galerkin discontinu [ABP06] traite indistinctement des cas de

maillages conformes et non-conformes ; il y a donc une difficulté supplémentaire à traiter le cas des équations de Maxwell.

3.2.2 Illustration numérique

Nous avons repris certains de ces calculs car il semble que ces résultats aient un lien étroit avec le choix d'une bonne méthode pour discrétiser les équations de Maxwell en régime harmonique (voir [Buf05] pour des résultats théoriques et [Dur06] pour des résultats numériques avec plusieurs méthodes de discrétisation conformes ou de type Galerkin discontinu sur des maillages quadrangulaires et hexaédriques). Dans l'exemple numérique considéré ici, le domaine Ω correspond au carré unité et on choisit $\mu = \varepsilon = 1$. On peut alors déterminer les valeurs propres et vecteurs propres qui sont de la forme :

$$\omega_{m,n}^{\pm} = \pm\sqrt{m^2+n^2}\pi \text{ et } \begin{pmatrix} (\mathbf{E}_{m,n}^{\pm})_x \\ (\mathbf{E}_{m,n}^{\pm})_y \\ (\mathbf{H}_{m,n}^{\pm})_z \end{pmatrix} = \begin{pmatrix} -\dfrac{m}{\sqrt{m^2+n^2}}\cos(n\pi x)\sin(m\pi y) \\ [0.35cm]\dfrac{n}{\sqrt{m^2+n^2}}\sin(n\pi x)\cos(m\pi y) \\ [0.35cm] \pm \mathrm{i}\cos(n\pi x)\cos(m\pi y) \end{pmatrix}, \quad (24)$$

$$\forall m,n \in \mathbb{N} \text{ avec } (m \neq 0 \text{ ou } n \neq 0).$$

On va utiliser une suite de maillage en drapeau anglais (voir la figure 6) et discrétisé le problème en utilisant une approximation polynomiale P_1 et des flux centrés. Nous utilisons ces maillages en drapeau anglais car ils ont été étudiés à plusieurs reprises pour mettre en évidence la présence de modes propres discrets non-physiques lorsque la méthode de discrétisation n'était pas adaptée à l'approximation du problème spectral [BDG99, BFGP99, BBG00].

Les résultats obtenus pour l'approximation numérique des valeurs propres sont rassemblés dans le tableau 7. On constate l'existence de deux modes propres non-physiques qui sont mis en valeur dans le tableau (ligne en gras). On s'attend à ce que ces modes disparaissent avec le raffinement du maillage mais ils demeurent présents même lorsque l'approximation des modes propres physiques de fréquence angulaire plus élevée est déjà de bonne précision comme cela est le cas ici pour les maillages les plus fins. Des représentations graphiques du champ électrique $\mathbf{E}$ du mode propre correspondant à $\omega_{m,n}^2 = 8\pi^2$ et d'un des modes propres correspondant à $\omega_{m,n}^2 = 9\pi^2$ sont montrées aux figures 7(a) et 7(b). On peut les comparer aux composantes $\mathbf{E}_x$ et $\mathbf{E}_y$ correspondant au premier mode propre non-physique et présentées à la figure 7. C'est un mode propre spatialement très oscillant contrairement aux modes physiques.

Dans [HW04], la résolution du même problème bidimensionnel est réalisée avec des approximations polynomiales à partir de l'ordre 2 et les auteurs ne mettent pas en évidence des modes propres discrets non-physiques. Ils en concluent que l'utilisation de flux centrés convient sans doute pour des problèmes bidimensionnels ce que semblent démentir les calculs

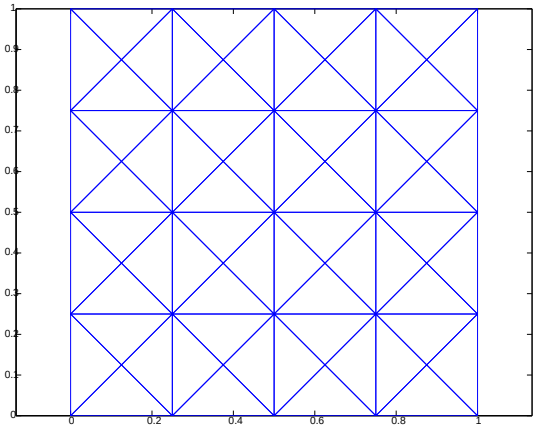
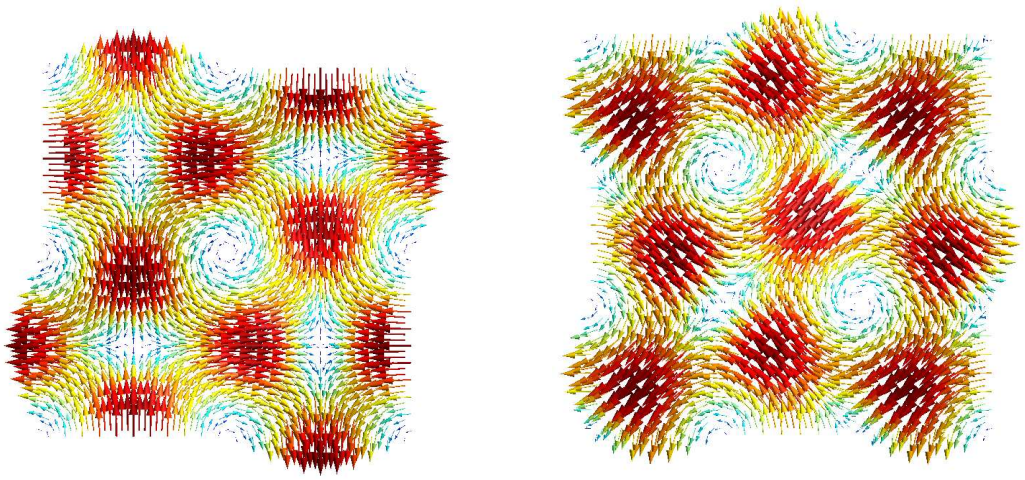


FIG. 6 – Maillage en drapeau anglais le plus grossier.

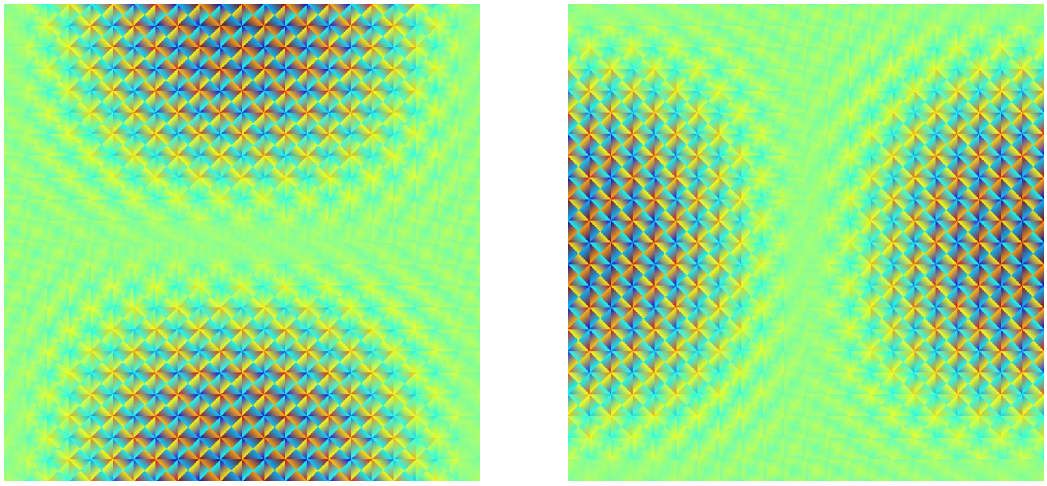
	64	256	1024	1936
1	1.0085	1.0021	1.0005	1.0003
1	1.0085	1.0021	1.0005	1.0003
2	2.0429	2.0109	2.0027	2.0014
4	4.1295	4.0338	4.0085	4.0045
4	4.1295	4.0338	4.0085	4.0045
5	5.2364	5.0619	5.0157	5.0083
5	5.2364	5.0619	5.0157	5.0083
8	7.9372	8.1715	8.0434	8.0230
	8.5857	**8.3390**	**8.4378**	**8.4532**
9	9.5992	9.1682	9.0430	9.0228
9	9.5992	9.1682	9.0430	9.0228
	8.6581	**9.0057**	**9.1087**	**9.1248**
10	10.8060	10.2270	10.0583	10.0310
10	10.8060	10.2270	10.0583	10.0310

TAB. 7 – Approximation numérique des plus petites valeurs propres non nulles sur le carré unité. Utilisation d'une méthode de Galerkin discontinu P_1 avec flux centrés. Le tableau contient les valeurs de $\omega_{m,n}^2/\pi^2$ pour différents maillages et de petites valeurs de m et n. La première colonne reprend les valeurs analytiques et les suivantes les valeurs calculées. La première ligne reprend le nombre d'éléments sur chaque maillage.



(a) $\omega_{m,n}^2 = 8\pi^2$ (b) $\omega_{m,n}^2 = 9\pi^2$

FIG. 7 – Représentation de deux modes propres physiques.



(a) $\mathbf{E}_x$ (b) $\mathbf{E}_y$

FIG. 8 – Représentation des composantes du mode non-physique correspondant à $\omega_{m,n}^2 \approx 8.5\pi^2$.

réalisés avec une approximation polynomiale P_1. Les calculs réalisés ici corroborent néanmoins la conclusion des résultats tridimensionnels fournis dans [HW04, WE06] et indique que si on considère un problème de cavité en utilisant une discrétisation de type Galerkin discontinu et des flux centrés, il est a priori impossible de montrer que le problème est bien posé. Cela ne présage par contre en rien de ce qui peut se passer avec l'utilisation de conditions aux limites absorbantes mais cela soulève des interrogations sur l'approximation numérique des équations de Maxwell en régime harmonique avec l'utilisation de flux centrés.

4 Résolution et préconditionnement du système linéaire

On présente dans cette section des techniques existantes pour la résolution des systèmes linéaires de type point-selle et certaines spécificités dont il faut tenir compte pour les équations de Maxwell. Les stratégies proposées peuvent être considérées comme des points de départ pour mettre en place des méthodes de résolution plus efficaces dans une perspective de calcul séquentiel ou à titre de solveur local dans une méthode de décomposition de domaine.

4.1 Résultats génériques pour les systèmes de type point-selle

Pour obtenir une écriture matricielle, prenons $(\mathbf{V}_i)_{1\leq i\leq N_h}$ une base de V_h. Les vecteurs $\underline{\mathbf{E}}$ et $\underline{\mathbf{H}}$, contenant les coefficients de $\mathbf{E}_h$ et de $\mathbf{H}_h$ dans cette base, sont alors solution du système :

$$\begin{pmatrix} A & \overline{B}^t \\ -B & C \end{pmatrix} \begin{pmatrix} \underline{\mathbf{H}} \\ \underline{\mathbf{E}} \end{pmatrix} = \begin{pmatrix} \underline{\mathbf{Q}}^H \\ \underline{\mathbf{Q}}^E \end{pmatrix}, \tag{25}$$

où les matrices et vecteurs sont définis de la manière suivante :

$$\forall i,j \in [\![1;N_h]\!],\ (A)_{ij} = a(\mathbf{V}_j,\mathbf{V}_i),\ (B)_{ij} = b(\mathbf{V}_j,\mathbf{V}_i),\ (C)_{ij} = c(\mathbf{V}_j,\mathbf{V}_i),$$
$$\forall i \in [\![1;N_h]\!],\ (\underline{\mathbf{Q}}^H)_i = \sum_{F\in\Gamma^a} \int_F \frac{1}{2} \left(N^t_{\mathbf{n}_F} N_{\mathbf{n}_F} \mathbf{H}^{\text{inc}} - N^t_{\mathbf{n}_F} \mathbf{E}^{\text{inc}}\right)^t \overline{\mathbf{V}}_i ds.$$
$$(\underline{\mathbf{Q}}^E)_i = \sum_{F\in\Gamma^a} \int_F \frac{1}{2} \left(N_{\mathbf{n}_F} N^t_{\mathbf{n}_F} \mathbf{E}^{\text{inc}} - N_{\mathbf{n}_F} \mathbf{H}^{\text{inc}}\right)^t \overline{\mathbf{V}}_i ds.$$

Dans le cas qui nous intéresse, on utilise des vecteurs de base à valeurs réelles et aucun coefficient complexe n'apparaît dans la définition de b en (15), on a donc $\overline{B} = B$. Cette structure est bien entendu valable que le problème soit bi- ou tridimensionnel. On se trouve ainsi face à un système de type point-selle, c'est-à-dire de la forme suivante :

$$M \begin{pmatrix} \underline{\mathbf{H}} \\ \underline{\mathbf{E}} \end{pmatrix} = \begin{pmatrix} A & B^t \\ -B & C \end{pmatrix} \begin{pmatrix} \underline{\mathbf{H}} \\ \underline{\mathbf{E}} \end{pmatrix} = \begin{pmatrix} \underline{\mathbf{Q}}^H \\ \underline{\mathbf{Q}}^E \end{pmatrix}. \tag{26}$$

qui peut aussi être considéré sous une forme symétrique :

$$M_s \begin{pmatrix} \underline{\mathbf{H}} \\ \underline{\mathbf{E}} \end{pmatrix} = \begin{pmatrix} A & B^t \\ B & -C \end{pmatrix} \begin{pmatrix} \underline{\mathbf{H}} \\ \underline{\mathbf{E}} \end{pmatrix} = \begin{pmatrix} \underline{\mathbf{Q}}^H \\ -\underline{\mathbf{Q}}^E \end{pmatrix}. \tag{27}$$

Pour résoudre ce type de systèmes, des approches couplée ou découplée peuvent être envisagées ; l'article [BGL05] présente un inventaire des méthodes qui ont été utilisées pour résoudre ce type de système. On en rappelle quelques éléments généraux qui pourraient être utilisés pour notre cas.

4.1.1 Approche découplée

Dans le cas d'une approche découplée, on réalise une élimination de Gauss par blocs pour résoudre d'abord un système d'inconnue $\underline{\mathbf{E}}$ et ensuite un système d'inconnue $\underline{\mathbf{H}}$ (ou inversement). L'élimination par blocs conduit ainsi à un système de la forme :

$$\begin{pmatrix} A & B^t \\ 0 & S \end{pmatrix} \begin{pmatrix} \underline{\mathbf{H}} \\ \underline{\mathbf{E}} \end{pmatrix} = \begin{pmatrix} \underline{\mathbf{Q}}^H \\ \underline{\mathbf{Q}}^E + \overline{B}A^{-1}\underline{\mathbf{Q}}^H \end{pmatrix} \tag{28}$$

où S désigne le complément de Schur et est égal à $C + BA^{-1}B^t$. On résout ainsi par une méthode itérative le système :

$$\mathrm{i}\omega S\underline{\mathbf{E}} = \mathrm{i}\omega(\underline{\mathbf{Q}}^E + BA^{-1}\underline{\mathbf{Q}}^H). \tag{29}$$

La multiplication par le scalaire $\mathrm{i}\omega$ est réalisée ici pour que le système puisse être vue comme une discrétisation *particulière* du problème du second ordre :

$$\mathrm{rot}\,\frac{1}{\mu}\,\mathrm{rot}\,\mathbf{E} - \omega^2\varepsilon\mathbf{E} = < \text{source} >, \tag{30}$$

avec les conditions aux limites adéquates. Une fois l'inconnue $\underline{\mathbf{E}}$ calculée, l'inconnue $\underline{\mathbf{H}}$ est la solution du système :

$$A\underline{\mathbf{H}} = \underline{\mathbf{Q}}^H - B^t\underline{\mathbf{E}}. \tag{31}$$

Un avantage dans le cas des méthodes de type Galerkin discontinu est que la matrice A^{-1} est creuse, diagonale par blocs et peut se calculer de manière explicite lors d'une phase de pré-traitement (ce n'est pas le cas néanmoins pour les flux décentrés présentés en (9)) ; la matrice du complément de Schur peut donc être complètement assemblée et la résolution d'un système de matrice A correspond juste à un produit matrice-vecteur.

On peut aussi envisager de ne pas assembler la matrice du complément de Schur, l'essentiel étant de savoir tout de même effectuer efficacement un produit matrice-vecteur avec B et B^t et des résolutions de systèmes de matrice A. L'inconvénient majeur de ne pas disposer explicitement de la matrice est la difficulté ensuite pour la construction d'un préconditionneur.

Le choix de la méthode par sous-espaces de Krylov peut s'orienter soient vers des méthodes génériques telles GMRES [SS86] ou vers des méthodes prenant en compte le caractère symétrique de la matrice telles COCG [VdVM90] ou QMR symétrique [FN94] (cela impose aussi d'utiliser un préconditionnement symétrique). Un préconditionnement efficace demeure néanmoins nécéssaire pour résoudre le problème restreint au complément de Schur.

4.1.2 Approche couplée

Préconditionnement découplé. La plupart des préconditionnements pour des systèmes de type point-selle s'appuie sur une vision par blocs et amène aussi à une résolution découplée mais uniquement dans l'étape de préconditionnement. Des résultats de Murphy et collab. [MGW00], généralisés par Ipsen [Ips01] dans le cas où $C \neq 0$, ont permis de définir des stratégies pour de bons préconditionnements par blocs. Ces auteurs ont en particulier montrer les deux résultats suivants :

- si P_d est la matrice diagonale par blocs définie par :

$$P_d = \begin{pmatrix} A & 0 \\ 0 & -S \end{pmatrix}, \tag{32}$$

  alors le polynôme $(\lambda - 1)(\lambda^2 - \lambda - 1)$ est un polynôme annulateur pour les matrices $P_d^{-1}M$ et MP_d^{-1}, M étant le matrice définie en (26).
- si P_t est la matrice triangulaire inférieure par blocs définie par :

$$P_t = \begin{pmatrix} A & 0 \\ -B & S \end{pmatrix}. \tag{33}$$

  alors le polynôme $(\lambda - 1)^2$ est un polynôme annulateur pour les matrices $P_t^{-1}M$ et MP_t^{-1}.

Ces résultats impliquent qu'une méthode telle que GMRES convergera en au plus 3 itérations dans le cas d'un préconditionnement à droite ou à gauche avec P_d et en 2 itérations avec P_t. Cependant, comme il est souligné dans [MGW00], on peut constater que si un tel préconditionneur fonctionne c'est parce que les matrices préconditionnées ont un nombre limité de valeurs propres, 3 en utilisant (32) et 1 en utilisant (33), et que ces valeurs propres sont clairement séparées de l'origine.

Pour mettre en oeuvre ces résultats, on remplace donc généralement A par une approximation $\hat{A}$ et S par une approximation $\hat{S}$ de façon à ce qu'une itération de préconditionnement ait un coût numérique raisonnable[2] tout en s'assurant que les valeurs propres de $\tilde{P}_d^{-1}M$ ou $\tilde{P}_t^{-1}M$ restent regroupées et à distance suffisante de l'origine (par exemple avec $\tilde{P}_t^{-1}M$, on s'attend à ce que les valeurs propres soient toutes distinctes mais restent très proches de 1). Cela ne prouve pas nécéssairement que la méthode itérative va converger rapidement [BGL05, Chapitre 9] mais cela est souvent suffisant en pratique.

[2] Le terme raisonnable signifie ici que le coût de calcul reste proportionnel aux nombre de degrés de liberté même si l'on raffine la discrétisation du problème.

Dans le cas de matrices issues de méthodes de type Galerkin discontinu (excepté le cas des flux décentrés (9)), on peut raisonnablement prendre $\hat{A} = A$ et dans ce cas la difficulté est de déterminer une matrice de préconditionnement pour S. Par ailleurs, dans le cas de l'utilisation d'un préconditionnement triangulaire de la forme :

$$\hat{P}_t = \begin{pmatrix} A & 0 \\ -B & \hat{S} \end{pmatrix}, \tag{34}$$

il est montré par Kanschat dans [Kan03, Lemme 5.1] que le spectre de $M\hat{P}_t^{-1}$ est égal à la réunion de 1 et du spectre de $\hat{S}^{-1}S$. Connaître une bonne approximation $\hat{S}$ de S permet donc d'envisager une convergence rapide. Des résultats numériques sont ainsi donnés dans [Kan03] dans le cas du Laplacien et du problème de Stokes discrétisés par des méthodes Galerkin discontinu localisé [CS98].

Multigrille avec lisseur couplé. Des méthodes multigrille ont aussi été proposées, conservant la structure particulière de type point-selle du niveau le plus fin au niveau le plus grossier. Les lisseurs sont alors quelque peu différents des lisseurs usuels pour les méthodes multigrille. On propose ici de présenter le lisseur de type Vanka [Van86] qui est parmi les plus utilisés pour les algorithmes multigrille appliqués au système de type point-selle.

On considère N_s ensembles formant une partition des indices de $\underline{\mathbf{E}}$: $(E_i)_{i=1,\dots,N_s}$. Partant de cette partition, on construit des ensembles $(H_i)_{i=1,\dots,N_s}$ des indices de $\underline{\mathbf{H}}$ de la manière suivante :

$$\forall i \in \{1,\dots,N_s\},\ \forall j \in \{1,\dots N_h\},\ (j \in H_i) \Leftrightarrow (\exists k \in E_i,\ B_{jk} \neq 0).$$

L'ensemble H_i contient ainsi tous les indices de $\underline{\mathbf{H}}$ qui interviennent dans une équation concernant l'un des indices de $\underline{\mathbf{E}}$ dans E_i. La notation $|E|$ désigne le cardinal de E. Soit i dans $\{1,\dots,N_s\}$ avec $E_i = \{i_1,\dots i_{E_i}\}$, on introduit l'opérateur de projection qui conserve uniquement les composantes indexées par E_i :

$$\begin{aligned} &P_i : \mathbb{R}^{N_h} \to \mathbb{R}^{|E_i|}, \\ &(P_i x)_k = x_{i_k},\ \forall x \in \mathbb{R}^{N_h},\ \forall k \in \{1,\dots,|E_i|\}. \end{aligned}$$

La notation x_{i_k} désigne la i_k-ième composante de x. Il est évident alors que l'opération transposée P_i^t est définie par :

$$\begin{aligned} &P_i^t : \mathbb{R}^{|E_i|} \to \mathbb{R}^{N_h}, \\ &(P_i^t y)_i = \begin{cases} 0 \text{ if } i \notin E_i \\ y_k \text{ if } i = i_k \in E_i \end{cases},\ \forall y \in \mathbb{R}^{|E_i|}. \end{aligned}$$

De la même manière, on peut définir l'opérateur de projection $Q_i : \mathbb{R}^{N_h} \to \mathbb{R}^{|H_i|}$ et sa transposée.

Les matrices des problèmes locaux pour i dans $\{1, \ldots, N_s\}$ ont une forme identique au problème initial (27), c'est-à-dire de type point-selle, mais sont de plus petite taille. Ces matrices prennent ainsi la forme suivante :

$$M_i = \begin{pmatrix} A_i & B_i^t \\ B_i & -C_i \end{pmatrix}, \tag{35}$$

où $A_i = P_i^t A P_i$, $B_i = Q_i^t B P_i$ et $C_i = Q_i^t C Q_i$. Dans des versions plus avancées, les matrices locales sont généralement des versions approchées de celle décrite en (35) [SZ03].

Une itération du lisseur de de type Vanka consiste à effectuer une itération d'un algorithme de Schwarz multiplicatif où les matrices des problèmes locaux sont les $(M_i)_{i=1,\ldots,N_s}$ et les opérateurs de restriction sont de la forme :

$$R_i = \begin{pmatrix} P_i & Q_i \end{pmatrix}.$$

L'algorithme 1 correspond à une itération de cette méthode.

Data : r, $(M_i)_{i=1,\ldots,N_s}$, $(R_i)_{i=1,\ldots,N_s}$
Result : g
begin
  for $i = 1 \ldots N_s$ **do**
    $r_i = R_i(r - M_s g)$;
    Résoudre $M_i g_i = r_i$;
    $g \leftarrow g + R_i^t g_i$;
  end
end

Algorithme 1 : Algorithme pour le lisseur de type Vanka.

Dans une version additive issue de ces algorithmes, on peut effectuer une analyse de la méthode en tant que solveur seul et en tant que lisseur pour la méthode multigrille. Aucune analyse par contre n'existe pour le cas multiplicatif qui est néanmoins le cas le plus couramment rencontré.

Remarque : des méthodes multigrille algébriques ont été étudiées dans [Wab06] avec des problèmes de type point-selle et des lisseurs de type Vanka. Le système d'équations considérés était Oseen et les espaces d'éléments finis étaient conformes mais cela peut être un point de départ intéressant pour nos applications.

4.2 Mise en oeuvre

4.2.1 L'approche et le préconditionnement découplé

Approximation du complément de Schur. Des résultats sont proposés uniquement pour la formulation utilisant des flux numériques pénalisés sur **E**. L'idée provient des similitudes démontrées dans [WE06] entre l'approximation du spectre par une méthode d'éléments

finis conformes, utilisant la 2ème famille de Nédélec [N86] et une méthode de type Galerkin discontinu utilisant un terme de pénalisation pour la composante tangentielle du champ **E**.

Pour approcher le complément de schur S, on a utilisé un algorithme à deux niveaux avec :

- une itération de Jacobi par blocs sur le système global. Chaque bloc regroupe toutes les inconnues pour $\underline{\mathbf{E}}$ associées à un même élément,
- une projection sur l'espace d'éléments finis conformes inclus dans l'espace d'approximation V_h, c'est à dire l'espace d'éléments finis de la 2ème famille de Nédélec. La résolution est exacte sur ce sous-espace.

Le problème traité est posé sur le carré unité ; c'est un problème de cavité, c'est-à-dire $\Gamma^m = \partial\Omega$. On choisit $\omega = 1.1\pi$ pour ne pas tomber sur une fréquence de résonance. Par ailleurs, pour la méthode de type Galerkin discontinu, on choisit une approximation polynomiale P_1.

On utilise pour la résolution une méthode GMRES(40) préconditionnée par l'algorithme à deux niveaux. La notation *(40)* signifie que la construction de l'espace de Krylov est redémarrée toutes les 40 itérations de la méthode GMRES. La méthode itérative s'arrête lorsque la norme 2 du résidu a été divisée par 10^6.

On donne les résultats obtenus avec deux valeurs du coefficient de pénalisation τ_h de (10) dans les tableaux 7(a) et 7(b). La première ligne indique le nombre d'éléments du maillage ; dans chaque élément, six degrés de liberté sont associés au champ **E** et trois degrés de liberté au champ **H**. L'approche *Schur* correspond à la résolution du système où la matrice est le complément de Schur (approche totalement découplée). L'approche *Complet* correspond au préconditionnement par une matrice triangulaire inférieure par blocs de la forme (34) qui utilise la même approximation du complément de Schur. Entre parenthèses est donnée d'abord l'itération externe en cours pour la méthode GMRES puis le nombre d'itérations internes au cours de cette itération externe.

(a) $\tau_h = 1/\pi$

	64	256	1024
Schur	(3, 35)	(4, 35)	(6, 15)
Complet	(4, 5)	(5, 5)	(6, 25)

(b) $\tau_h = 10/\pi$

	64	256	1024
Schur	(1, 36)	(2, 7)	(2, 24)
Complet	(1, 39)	(2, 13)	(2, 31)

TAB. 8 – Nombre d'itérations pour la méthode GMRES(40) préconditionnée par un algorithme à deux niveaux pour réduire la norme 2 du résidu par 10^6. Méthode GD-P1, flux décentré.

Commentaires. Le comportement des méthodes itératives est similaire entre l'approche totalement découplée (Schur) et l'approche où l'on considère un préconditionnement triangulaire par blocs pour le système global (Complet). L'évolution du nombre d'itérations se rapproche d'un comportement optimal lorsque l'on accentue le coefficient de pénalisation. Quelques justications sont possibles :

- avec la pénalisation sur le champ $\mathbf{E}$, la convergence spectrale est quasi-semblable pour les valeurs propres d'amplitude peu élevée avec la méthode d'éléments finis conformes ;
- en particulier, on retrouve un nombre identique de modes propres discrets pour la valeur propre 0 et le sous-espace propre correspondant peut être décrit de manière identique.

***Noyau* discret de l'opérateur rotationnel.** Pour expliquer la dernière affirmation, on identifie les modes propres correspondant à la valeur propre $\omega = 0$. On se place pour cela dans la cas où $\Gamma_a = \emptyset$. On considère le problème mixte pour lequel $\omega = 0$ avec les formes bilinéaires a, b et c données en (18)–(20). Un mode propre discret de valeur propre nulle vérifie alors les équations suivantes :

$$\begin{cases} \displaystyle\sum_{F\in\Gamma^0}\int_F \alpha_F^H [\![\mathbf{H}_h]\!]_T^t [\![\overline{\mathbf{G}}]\!]_T ds + \sum_{K\in\mathcal{T}_h}\int_K \overline{\mathbf{G}}^t \operatorname{rot}\mathbf{E}_h dv + \\ \displaystyle\sum_{F\in\Gamma^0\cup\Gamma^m}\int_F \{\overline{\mathbf{G}}\}^t [\![\mathbf{E}_h]\!]_T ds = 0,\ \forall \mathbf{G}\in V_h, \\ \displaystyle\sum_{F\in\Gamma^0}\int_F \alpha_F^E [\![\mathbf{E}_h]\!]_T^t [\![\overline{\mathbf{F}}]\!]_T ds + \sum_{F\in\Gamma^m}\int_F \eta_F [\![\mathbf{E}_h]\!]_T^t [\![\overline{\mathbf{F}}]\!]_T ds - \sum_{K\in\mathcal{T}_h}\int_K \mathbf{H}_h^t \operatorname{rot}\overline{\mathbf{F}} dv - \\ \displaystyle\sum_{F\in\Gamma^0\cup\Gamma^m}\int_F \{\mathbf{H}_h\}^t [\![\overline{\mathbf{F}}]\!]_T ds = 0,\ \forall \mathbf{F}\in V_h. \end{cases}$$

En utilisant pour vecteur test $\begin{pmatrix}\mathbf{H}_h\\ \mathbf{E}_h\end{pmatrix}$, on obtient :

$$\begin{cases} \displaystyle\sum_{F\in\Gamma^0}\int_F \alpha_F^H [\![\mathbf{H}_h]\!]_T^t [\![\overline{\mathbf{H}}_h]\!]_T ds + \sum_{K\in\mathcal{T}_h}\int_K \overline{\mathbf{H}}_h^t \operatorname{rot}\mathbf{E}_h dv + \\ \displaystyle\sum_{F\in\Gamma^0\cup\Gamma^m}\int_F \{\overline{\mathbf{H}}_h\}^t [\![\mathbf{E}_h]\!]_T ds = 0, \\ \displaystyle\sum_{F\in\Gamma^0}\int_F \alpha_F^E [\![\mathbf{E}_h]\!]_T^t [\![\overline{\mathbf{E}}_h]\!]_T ds + \sum_{F\in\Gamma^m}\int_F \eta_F [\![\mathbf{E}_h]\!]_T^t [\![\overline{\mathbf{E}}_h]\!]_T ds - \sum_{K\in\mathcal{T}_h}\int_K \mathbf{H}_h^t \operatorname{rot}\overline{\mathbf{E}}_h dv - \\ \displaystyle\sum_{F\in\Gamma^0\cup\Gamma^m}\int_F \{\mathbf{H}_h\}^t [\![\overline{\mathbf{E}}_h]\!]_T ds = 0. \end{cases}$$

En sommant ces deux équations et en prenant la partie réelle, on a :

$$\sum_{F\in\Gamma^0}\int_F \alpha_F^H [\![\mathbf{H}_h]\!]_T^t [\![\overline{\mathbf{H}}_h]\!]_T ds + \sum_{F\in\Gamma^0}\int_F \alpha_F^E [\![\mathbf{E}_h]\!]_T^t [\![\overline{\mathbf{E}}_h]\!]_T ds + \sum_{F\in\Gamma^m}\int_F \eta_F [\![\mathbf{E}_h]\!]_T^t [\![\overline{\mathbf{E}}_h]\!]_T ds = 0.$$

En conséquence :
- si $\alpha_F^H > 0$ alors $[\![\mathbf{H}_h]\!]_T = 0$, $\forall F \in \Gamma^0$,
- si $\alpha_F^E > 0$ alors $[\![\mathbf{E}_h]\!]_T = 0$, $\forall F \in \Gamma^0$,
- si $\eta_F > 0$ alors $[\![\mathbf{E}_h]\!]_T = 0$, $\forall F \in \Gamma^m$.

Ainsi, si η_F et α_F^E sont strictement positifs, il découle de la première équation que :

$$\int_K \overline{\mathbf{G}}^t \operatorname{rot} \mathbf{E}_h = 0,\ \forall K \in \mathcal{T}^h,\ \forall \mathbf{G} \in \mathcal{P}(K) \implies \operatorname{rot} \mathbf{E}_h = 0 \text{ sur } \mathcal{P}(K).$$

En outre, si α_F^H est strictement positif, on obtient alors à partir de la seconde équation :

$$\int_K \overline{\mathbf{F}}^t \operatorname{rot} \mathbf{H}_h = 0,\ \forall K \in \mathcal{T}^h,\ \forall \mathbf{F} \in \mathcal{P}(K) \implies \operatorname{rot} \mathbf{H}_h = 0 \text{ sur } \mathcal{P}(K).$$

En conséquence, on a une caractérisation des modes propres correspondant à $\omega = 0$ quels que soient le maillage et l'espace d'éléments finis utilisés avec le choix de flux décentrés (9). En particulier, dans le cas d'un maillage conforme, de l'utilisation pour l'espace local de $\mathcal{P}(K) = [P_k(K)]^3$ et du choix de flux décentrés, cela coïncide avec l'espace $V_{h,0}^{\mathbf{E}} \times V_{h,0}^{\mathbf{H}}$ où :

$$V_{h,0}^{\mathbf{E}} = \left\{ \operatorname{grad}(\phi) \mid \phi \in H_0^1(\Omega) \text{ et } \forall K \in \mathcal{T}_h,\ \phi_{|K} \in P_{k+1}(K) \right\}. \tag{36a}$$

et :

$$V_{h,0}^{\mathbf{H}} = \left\{ \operatorname{grad}(\phi) \mid \phi \in H^1(\Omega) \text{ et } \forall K \in \mathcal{T}_h,\ \phi_{|K} \in P_{k+1}(K) \right\}. \tag{36b}$$

4.2.2 Méthodes multigrille avec le lisseur couplé

Comparaison numérique de différentes approches sur un problème elliptique. Pour comprendre les propriétés de la méthode multigrille, on se ramène dans le cadre le plus simple possible :

- le domaine est le carré unité, le milieu est homogène et on utilise uniquement des conditions aux limites métalliques.
- on considère le système de Maxwell elliptique le plus simple, c'est-à-dire que l'on prend $\omega = -\mathrm{i}$ dans le modèle harmonique.
- on considère le cas d'un maillage uniforme. On part d'une grille avec peu d'éléments , 8 dans le cas présent (voir figure 9).
- on utilise l'inclusion des espaces d'éléments finis pour définir les opérateurs de prolongement. En effet, les problèmes sont définis sur une suite de maillages emboîtés.
- on commence par l'étude d'une méthode deux grilles avant de passer à l'algorithme multigrille.

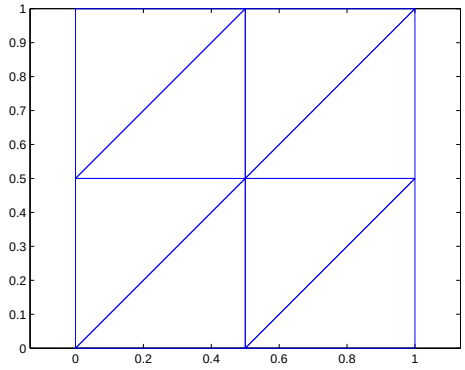


FIG. 9 – Maillage initial pour tester les méthodes multigrille.

Algorithme de lissage. Inspiré des résultats connus pour les méthodes multigrille appliquées aux éléments finis conformes [Hip99, AFW00], on se propose d'utiliser dans l'algorithme de lissage un traitement spécifique pour les composantes dans le *noyau* discret de l'opérateur rotationnel.

Ainsi, si M_s est la matrice par blocs définie en (27) et P_ϕ la matrice des coefficients des fonctions de $V_{h,0}^{\mathbf{E}} \times V_{h,0}^{\mathbf{H}}$ dans la base d'éléments finis discontinue, on note M_ϕ la matrice $P_\phi^t M_s P_\phi$ qui correspond à la forme bilinéaire du problème évaluée uniquement sur l'espace des fonctions mises en exergue plus haut en (36a) et (36b). Pour être plus précis, sur le sous-espace $V_{h,0}^{\mathbf{E}} \times V_{h,0}^{\mathbf{H}}$, on obtient deux problèmes découplés pour les champs $\mathbf{H}$ et $\mathbf{E}$ de forme bilinéaire $\tilde{a}$ sur la partie concernant $\mathbf{H}$:

$$\begin{aligned} &\tilde{a}(\phi_h, \psi_h) = \int_{\Omega_h} \mathrm{i}\omega\mu \,\mathrm{grad}(\phi_h)^t \,\mathrm{grad}(\psi_h) dv, \\ &\forall \phi_h, \psi_h \in \left\{ \phi \in H^1(\Omega) \mid \forall K \in \mathcal{T}_h,\ \phi_{|K} \in P_{k+1}(K) \right\}. \end{aligned} \tag{37a}$$

et $\tilde{c}$ pour la partie concernant $\mathbf{E}$:

$$\begin{aligned} &\tilde{c}(\phi_h, \psi_h) = \int_{\Omega_h} \mathrm{i}\omega\varepsilon \,\mathrm{grad}(\phi_h)^t \,\mathrm{grad}(\psi_h) dv, \\ &\forall \phi_h, \psi_h \in \left\{ \phi \in H_0^1(\Omega) \mid \forall K \in \mathcal{T}_h,\ \phi_{|K} \in P_{k+1}(K) \right\}. \end{aligned} \tag{37b}$$

Précisons que dans le cas 2D tranverse électrique, la chose est plus simple car $V_{h,0}^{\mathbf{H}}$ est l'espace des constantes ; on oublie donc de faire la projection sur ce sous-espace et on se concentre sur $V_{h,0}^{\mathbf{E}}$. Comme $\tilde{a}(.,.)$ et $\tilde{c}(.,.)$ sont des formes bilinéaires d'opérateur de type

$-\operatorname{div} k \operatorname{grad}$, on utilise une ou plusieurs itérations de Gauss-Seidel symétrique pour lisser l'erreur dans ces sous-espaces. Pour traiter les composantes de l'erreur qui ne sont pas dans le sous-espace $V_{h,0}^{\mathbf{E}} \times V_{h,0}^{\mathbf{H}}$, on va utiliser un lissage de type Vanka sur le système global. L'algorithme 2 correspond au lisseur utilisé pour les essais numériques.

Data : r, M_s, M_ϕ, P_ϕ, $(M_i)_{i=1,\dots,N_s}$, $(R_i)_{i=1,\dots,N_s}$
Result : g
begin
  $r_\phi = P_\phi^t r$, $g_\phi = 0$;
  ν itérations de Gauss-Seidel symétrique sur $M_\phi g_\phi = r_\phi$;
  $g \leftarrow r + P_\phi r_\phi$;
  for $j = 1 \dots \nu$ **do**
    for $i = 1 \dots N_s$ **do**
      $r_i = R_i(r - M_s g)$;
      Résoudre $M_i g_i = r_i$;
      $g \leftarrow g + R_i^t g_i$;
    end
  end
  $r_\phi = P_\phi^t g$, $g_\phi = 0$;
  ν itérations de Gauss-Seidel symétrique sur $M_\phi g_\phi = r_\phi$;
  $g \leftarrow g + P_\phi r_\phi$;
end

Algorithme 2 : Algorithme utilisé pour le lisseur.

Construction des sous-espaces pour le lisseur de type Vanka. Comme indiqué sous-section 4.1 dans le sous-paragraphe concernant les lisseurs de type Vanka, il est nécéssaire de définir des sous-espaces où vont être effectuées les corrections. Nous avons pour cela utilisé divers groupements :

Cas 1 Les composantes du champ **E** associées à chaque élément **fin** du maillage forment les ensembles de la partition des indices. Nous associons ensuite les indices de **H** qui sont en relation avec chacun de ces ensembles.

Cas 2 Les composantes du champ **E** associées à chaque élément **grossier** du maillage forment les ensembles de la partition des indices.

Cas 3 les composantes du champ **E** associées aux deux éléments partageant une face sont regroupées (ce n'est plus alors une partition ; un recouvrement existe déjà au niveau des composantes du champ **E**).

Méthode pour évaluer le comportement des algorithmes. Pour évaluer les propriétés des algorithmes, on suit la démarche suivante :

- l'algorithme est démarré avec un second membre nul et le vecteur initial est aléatoire,
- la tolérance est fixée à 10^{-20} et le nombre maximal d'itérations est fixé à 30. L'estimation du taux de convergence numérique correspond au rapport de la norme 2 de l'erreur à la dernière itération sur la norme 2 de l'erreur à l'avant-dernière itération.
- devant la faible (mais néanmoins non nulle) variabilité obtenue dans l'estimation du taux de convergence, on répète plusieurs fois l'opération et on prend la moyenne des résultats obtenus.

Ces opérations sont effectuées sur une série de maillages : le maillage initial est structuré et les suivants sont obtenus par raffinement uniforme. On note τ_h^0 le maillage initial et τ_h^i le maillage correspondant à i raffinements.

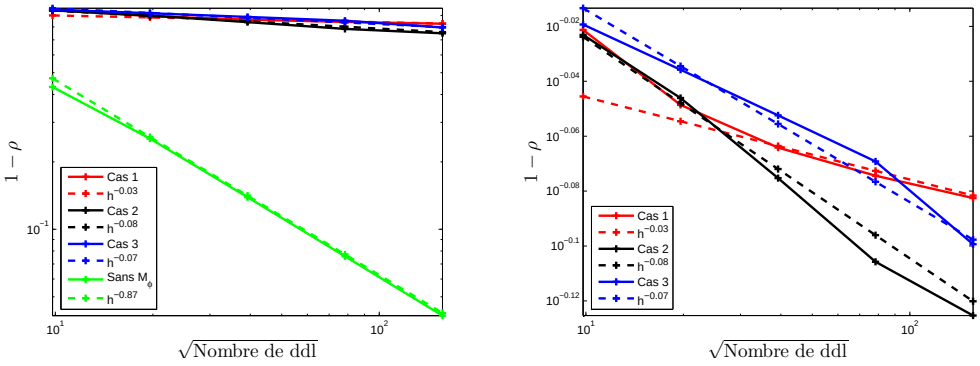
Résultats. Pour le cas de la méthode deux grilles et d'une approximation P_0, l'évolution du taux de convergence avec les différents maillages est donnée dans le tableau 9. On a utilisé $\nu = 1$ dans l'algorithme 2, les Cas 1, 2 et 3 correspondent aux différentes manières de construire les sous-espaces et la dénomination sans M_ϕ désigne le cas où l'on utilise uniquement une itération de l'algorithme 1 pour le lissage. Le comportement obtenu est proche de l'optimal, à savoir que le taux de convergence évalué numériquement est quasiment indépendant du nombre d'inconnues comme on peut le constater sur la figure 10, en particulier dans la légende la figure 10(b). On retiendra particulièrement les Cas 1 et 3. On note l'importance de la correction sur le noyau discret puisque la comparaison est très nettement favorable aux méthodes utilisant cette correction supplémentaire.

	τ_h^1 (96)	τ_h^2 (384)	τ_h^3 (1 536)	τ_h^4 (6 144)	τ_h^5 (24 576)
Cas 1	0.0478	0.106	0.137	0.157	0.173
Cas 2	0.0518	0.101	0.159	0.216	0.251
Cas 3	0.0435	0.0791	0.114	0.147	0.205
Sans M_ϕ	0.568	0.746	0.861	0.924	0.959

TAB. 9 – Estimation de l'évolution du taux de convergence en fonction du maillage pour diverses méthodes deux grilles. Entre parenthèses le nombre de degrés de liberté pour chaque problème. Méthode GD-P0, flux décentré.

L'utilisation de l'algorithme deux grilles est d'un intérêt limité puisque le coût numérique et le stockage mémoire pour la grille grossière deviennent rapidement prohibitif. Le tableau 10 montre une comparaison de l'algorithme deux grilles avec deux versions multigrille :

- un algorithme V-cycle(1, 1) avec un pas de pré- et un pas de post-lissage ; ses performances se dégradent par rapport à la méthode deux grilles,
- un V-cycle variable, c'est-à-dire utilisant 2^{j-1} pré- et post-lissages au niveau j ; son coût est comparable à celui d'un W-cycle. On retrouve avec cette méthode le même comportement que l'algorithme deux grilles.



(a) Ensemble complet des courbes. (b) Détail pour les Cas 1 à 3.

FIG. 10 – Évolution du taux de convergence en fonction de la racine carrée du nombre de degrés de liberté. Les algorithmes sont ceux annoncés dans le tableau 9. Méthode GD-P0, flux décentré.

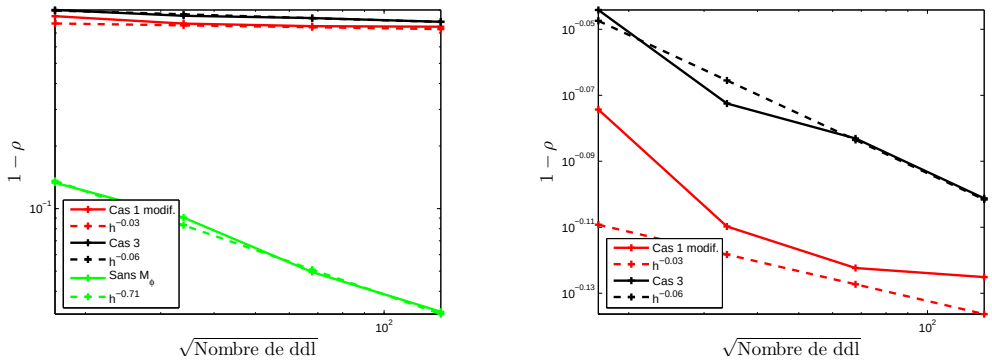
	τ_h^1	τ_h^2	τ_h^3	τ_h^4	τ_h^5
Deux grilles	0.0478	0.1059	0.1373	0.1574	0.1731
Multigrille V-cycle(1, 1)	0.0478	0.0920	0.1534	0.2094	0.2537
Multigrille V-cycle variable	0.0478	0.1045	0.1383	0.1569	0.1724

TAB. 10 – Estimation de l'évolution du taux de convergence en fonction du maillage. Comparaison des méthodes deux grilles et multigrille pour le Cas 1. Méthode GD-P0, flux décentré.

Pour le cas de la méthode deux grilles et de l'utilisation d'une méthode d'approximation P_1, on observe une dégradation globale des taux de convergence. Le Cas 3 conserve néanmoins un comportement asymptotique sensiblement équivalent au cas P_0 ; on peut le constater en comparant les comportements asymptotiques donnés sur les figures 10 et 11. Le comportement pour le Cas 1 n'est plus équivalent à son comportement pour une approximation P_0. Cependant si l'on accentue la correction au niveau du noyau discret ($\nu+1$ itérations de Gauss-Seidel symétrique sur $M_\phi g_\phi = r_\phi$ au lieu de ν), on retrouve un comportement proche de celui observé pour l'approximation P_0. Le taux de convergence de la méthode sans M_ϕ est très proche de 1 dès le premier maillage.

	τ_h^1 (288)	τ_h^2 (1152)	τ_h^3 (4 608)	τ_h^4 (18 432)
Cas 1	0.3189	0.4182	0.5060	0.7209
Cas 1 modifié	0.1573	0.2233	0.2455	0.2503
Cas 3	0.0966	0.1537	0.1741	0.2079
Sans M_ϕ	0.8667	0.9095	0.9505	0.9681

TAB. 11 – Estimation de l'évolution du taux de convergence en fonction du maillage pour diverses méthodes deux grilles. Entre parenthèses le nombre de degrés de liberté sur chaque grille. Méthode GD-P1, flux décentré.



(a) Ensemble complet des courbes. (b) Détail pour les Cas 1 à 3.

FIG. 11 – Évolution du taux de convergence en fonction de la racine carrée du nombre de degrés de liberté. Les algorithmes sont ceux annoncés dans le tableau 11. Méthode GD-P1, flux décentré.

Pour conclure sur ces problèmes modèles, on discute brièvement de l'occupation mémoire des algorithmes proposés. Il s'agit de vérifier que l'espace mémoire occupé par la structure de données contenant les informations multiniveau évolue à peu près linéairement avec le nombre d'inconnues. Les résultats sont contenus dans le tableau 12. Dans la première colonne :

- le terme *multiniveau* désigne la structure de données qui contient toute l'information multiniveau. Elle n'est pas réellement optimisée et sans doute un peu redondante ; elle contient en particulier M, M_ϕ, P_ϕ et les factorisations des matrices $(M_i)_{i=1,\dots,N_{ss}}$ à tous les niveaux ainsi que les opérateurs de transfert.
- le terme *infossd* désigne la structure de données qui contient les factorisations des $(M_i)_{i=1,\dots,N_{ss}}$ et d'autres informations liées à celles-ci. On constate que c'est la structure de données prépondérante pour la consommation mémoire de multiniveau.
- L, U, P, Q désigne les structures de données contenant la factorisation LU pour matrices creuses délivrée par UMFPACK version 4.3 [Dav04] sous Matlab. C'est la meilleure méthode que l'on puisse utiliser pour factoriser des matrices creuses sans propriétés particulières sous Matlab 7.1.
- A désigne l'espace occupé par la matrice seul.

Le comportement est bien quasi-linéaire pour la structure de données multiniveau. Même peu optimisé, le comportement asymptotique de la consommation mémoire est bien supérieur à un solveur direct pour matrice creuse très optimisé comme UMFPACK.

	τ_h^1	τ_h^2	τ_h^3	τ_h^4
multiniveau	323 430	1 543 270	6 576 446	27 025 278
infossd	165 280	705 696	2 913 952	11 840 160
L, U, P, Q	111 284	638 128	3 500 800	19 080 088
M	47 236	194 308	787 972	3 173 380

TAB. 12 – Évolution de l'espace mémoire (en octets) occupé par différentes structures de données pour la résolution du problème modèle en fonction des différents maillages. Méthode GD-P1, flux décentré.

Comportement numérique sur le problème harmonique. On résout de nouveau notre problème modèle où le domaine est le carré unité avec une onde plane incidente de fréquence angulaire $\omega = 2\pi$ et des conditions de Silver-Müller sur l'ensemble de la frontière.

On observe le comportement de la méthode GMRES(10) préconditionnée par une itération d'un V-cycle variable appuyé sur la décomposition en sous-espace du Cas 1 pour la résolution du problème. En effet, le problème étant en régime harmonique, l'algorithme multigrille seul peut ne pas converger ou difficilement et l'utilisation d'une méthode de Krylov permet d'assurer cette convergence tout en maintenant un taux de convergence intéressant.

Le tableau 13 donne pour les différents maillage et un ordre d'approximation P_0 le nombre d'itérations nécéssaires à la méthode itérative pour réduire par 10^{10} la norme 2 du résidu. Le maillage grossier utilisé correspond à τ_h^0 même si le problème harmonique est sans doute très mal approché sur ce maillage ; la méthode demeure robuste même en utilisant ce maillage grossier. Le nombre d'itérations pour atteindre le critère de convergence varie peu en particulier à partir du maillage τ_h^3. À titre indicatif, nous avons repris l'étude du stockage

mémoire avec les structures du tableau 12 mais appliquée à notre problème harmonique avec une approximation P_0 dans le tableau 14.

Maillage	τ_h^1	τ_h^2	τ_h^3	τ_h^4	τ_h^5
Nombre d'itérations	(1, 8)	(2, 1)	(2, 3)	(2, 4)	(2, 4)

TAB. 13 – Nombre d'itérations de la méthode GMRES(10) préconditionnée par un V-cycle variable pour réduire par 10^{10} la norme 2 du résidu pour les différents maillages. Le premier chiffre entre parenthèses est l'itération externe, le second l'itération interne. Méthode GD-P0, flux décentré.

	τ_h^1	τ_h^2	τ_h^3	τ_h^4	τ_h^5
multiniveau	85 774	404 806	1 722 342	7 079 070	28 682 558
infossd	42 960	184 144	761 424	3 095 632	12 482 640
L, U, P, Q	30 416	180 816	1 021 016	5 484 936	29 397 336
M	13 188	55 300	226 308	915 460	3 682 308

TAB. 14 – Évolution de l'espace mémoire (en octets) occupé par différentes structures de données pour la résolution du problème modèle en fonction des différents maillages. À relier au tableau 13. Méthode GD-P0, flux décentré.

Le tableau 15 donne pour les différents maillage et un ordre d'approximation P_1 le nombre d'itérations nécéssaires à la méthode itérative pour réduire par 10^{10} la norme 2 du résidu. Notons ici que le V-cycle est la version modifiée du Cas 1 référencé dans le tableau 11. Le nombre d'itérations pour atteindre le critère de convergence varie peu en particulier à partir du maillage τ_h^2. À titre indicatif, nous avons repris l'étude du stockage mémoire avec les structures du tableau 12 mais appliqué à notre problème harmonique avec une approximation P_0 dans le tableau 16.

Maillage	τ_h^1	τ_h^2	τ_h^3	τ_h^4
Nombre d'itérations	(3, 4)	(2, 6)	(2, 7)	(3, 4)

TAB. 15 – Nombre d'itérations de la méthode GMRES(10) préconditionnée par un V-cycle variable pour réduire par 10^{10} la norme 2 du résidu pour les différents maillages. Le premier chiffre entre parenthèses est l'itération externe, le second l'itération interne. Méthode GD-P1, flux décentré.

5 Bilan

Dans ce rapport, nous nous sommes intéressés à la résolution numérique des équations de Maxwell en régime harmonique par une méthode de type Galerkin discontinu. L'introduction

	τ_h^1	τ_h^2	τ_h^3	τ_h^4
multiniveau	519 366	2 455 750	10 393 310	42 533 598
infossd	256 768	1 102 080	4 562 176	18 560 256
L, U, P, Q	177 616	1 032 816	5 711 776	31 308 616
M	77 316	319 492	1 298 436	5 234 692

TAB. 16 – Évolution de l'espace mémoire (en octets) occupé par différentes structures de données pour la résolution du problème modèle en fonction des différents maillages. À relier au tableau 15. Méthode GD-P1, flux décentré.

de quelques paramètres permet d'obtenir une formulation couvrant différents choix de flux numérique. Pour évaluer ces différentes options, une comparaison des ordres de convergence a été menée sur un problème simple en dimension deux. Ces résultats indiquent que la méthode utilisant des flux décentrés permet de garantir une convergence des deux champs comparable et proche d'un taux optimal. Les deux autres approches envisagées, avec flux centrés et flux avec pénalisation sur le champ **E**, ne permettent pas d'obtenir ces perfomances.

D'autre part, nous rappelons que la discrétisation d'un problème de modes propres ne peut être réalisées de manière fiable avec une méthode Galerkin discontinue utilisant des flux centrés. Cela conduit à penser que l'on ne peut pas faire confiance à cette technique pour la résolution des équations de Maxwell en régime harmonique. On privilégie donc des approches avec flux centrés ou flux avec pénalisation sur le champ **E**.

Pour la résolution des systèmes linéaires obtenus, on propose ici de s'appuyer sur les techniques existantes pour le traitement des systèmes de type point-selle. Néanmoins, des spécificités propres aux équations de Maxwell, notamment le noyau discret de l'opérateur rot, doivent être soigneusement prises en compte. Partant de ces considérations, une méthode multigrille est proposée qui fournit des résultats quasi-optimaux pour la résolution des systèmes linéaires en question.

Références

[ABP06] P.F. Antonietti, A. Buffa, and I. Perugia. Discontinuous Galerkin approximation of the Laplace eigenproblem. *Comput. Methods Appl. Mech. Engrg.*, 195(25-28) :3483–3503, 2006.

[AFW00] D.N. Arnold, R.S. Falk, and R. Winther. Multigrid in $H(\mathrm{div})$ and $H(\mathrm{curl})$. *Numer. Math.*, 85(2) :197–217, 2000.

[Ain04] M. Ainsworth. Dispersive and dissipative behaviour of high order discontinuous Galerkin finite element methods. *J. Comput. Phys.*, 198(1) :106–130, 2004.

[BBG00] D. Boffi, F. Brezzi, and L. Gastaldi. On the problem of spurious eigenvalues in the approximation of linear elliptic problems in mixed form. *Math. Comp.*, 69(229) :121–140, 2000.

[BDG99] D. Boffi, R. G. Duran, and L. Gastaldi. A remark on spurious eigenvalues in a square. *Appl. Math. Lett.*, 12(3) :107–114, 1999.

[BFGP99] D. Boffi, P. Fernandes, L. Gastaldi, and I. Perugia. Computational models of electromagnetic resonators : analysis of edge element approximation. *SIAM J. Numer. Anal.*, 36(4) :1264–1290, 1999.

[BGL05] M. Benzi, G.H. Golub, and J. Liesen. Numerical solution of saddle point problems. *Acta Numer.*, 14 :1–137, 2005.

[BHI07] A. Buffa, P. Houston, and Perugia I. Discontinuous Galerkin computation of the Maxwell eigenvalues on simplicial meshes. *J. Comput. Appl. Math.*, 204(2) :317–333, 2007.

[BP06] A. Buffa and I. Perugia. Discontinuous Galerkin approximation of the Maxwell eigenproblem. *SIAM J. Numer. Anal.*, 44(5) :2198–2226, 2006.

[Buf05] A. Buffa. Remarks on the discretization of some noncoercive operator with applications to heterogeneous Maxwell equations. *SIAM J. Numer. Anal.*, 43(1) :1–18, 2005.

[CN06] E. Creuse and S. Nicaise. Discrete compactness for a discontinuous Galerkin approximation of Maxwell's system. *M2AN : Math. Model. Numer. Anal.*, 40(2) :413–430, 2006.

[CS98] B. Cockburn and C.-W. Shu. The local discontinuous Galerkin method for time-dependent convection-diffusion systems. *SIAM J. Numer. Anal.*, 35(6) :2440–2463, 1998.

[Dav04] T.A. Davis. Algorithm 832 : UMFPACK V4.3—an unsymmetric-pattern multifrontal method. *ACM Trans. Math. Software*, 30(2) :196–199, 2004.

[DFLP06] V. Dolean, H. Fol, S. Lanteri, and S. Piperno. Méthodes de type Galerkin discontinu pour la résolution numérique des équations de Maxwell en régime fréquentiel. Technical Report 5904, INRIA, 2006.

[Dur06] M. Durufle. *Intégration numérique et éléments finis d'ordre élevé appliqués aux équations de Maxwell en régime harmonique*. PhD thesis, Université Paris Dauphine - Paris IX, 2006.

[EG06a] A. Ern and J.-L. Guermond. Discontinuous Galerkin methods for Friedrichs' systems. I. General theory. *SIAM J. Numer. Anal.*, 44(2) :753–778, 2006.

[EG06b] A. Ern and J.-L. Guermond. Discontinuous Galerkin methods for Friedrichs' systems. II. Second-order elliptic pde's. *SIAM J. Numer. Anal.*, 44(6) :2363–2388, 2006.

[FLLP05] L. Fezoui, S. Lanteri, S. Lohrengel, and S. Piperno. Convergence and stability of a discontinuous Galerkin time-domain method for the 3D heterogeneous Maxwell equations on unstructured meshes. *M2AN Math. Model. Numer. Anal.*, 39(6) :1149–1176, 2005.

[FN94] R.W. Freund and N.M. Nachtigal. An implementation of the QMR method based on coupled two-term recurrences. *SIAM J. Sci. Comput.*, 15(2) :313–337, 1994.

[Hip99] R. Hiptmair. Multigrid method for Maxwell's equations. *SIAM J. Numer. Anal.*, 36(1) :204–225, 1999.

[HPSS05] P. Houston, I. Perugia, A. Schneebeli, and D. Schötzau. Interior penalty method for the indefinite time-harmonic Maxwell equations. *Numer. Math.*, 100(3) :485–518, 2005.

[HW04] J.S. Hesthaven and T. Warburton. High-order nodal discontinuous Galerkin methods for the Maxwell eigenvalue problem. *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 362(1816) :493–524, 2004.

[Ips01] I.C.F. Ipsen. A note on preconditioning nonsymmetric matrices. *SIAM J. Sci. Comput.*, 23(3) :1050–1051, 2001.

[Kan03] G. Kanschat. Preconditioning methods for local discontinuous Galerkin discretizations. *SIAM J. Sci. Comput.*, 25(3) :815–831, 2003.

[MGW00] M.F. Murphy, G.H. Golub, and A.J. Wathen. A note on preconditioning for indefinite linear systems. *SIAM J. Sci. Comput.*, 21(6) :1969–1972, 2000.

[Néd86] J. C. Nédélec. A new family of mixed finite elements in $\mathbb{R}^3$. *Num. Math.*, 50 :57–81, 1986.

[Pip00] S. Piperno. L^2-stability of the upwind first order finite volume scheme for the Maxwell equations in two and three dimensions on arbitrary unstructured meshes. *M2AN : Math. Model. Numer. Anal.*, 34(1) :139–158, 2000.

[SS86] Y. Saad and M.H. Schultz. GMRES : a generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM J. Sci. Statist. Comput.*, 7(3) :856–869, 1986.

[SZ03] J. Schöberl and W. Zulehner. On Schwarz-type smoothers for saddle point problems. *Numer. Math.*, 95(2) :377–399, 2003.

[Van86] S. P. Vanka. Block-implicit multigrid solution of Navier-Stokes equations in primitive variables. *J. Comput. Phys.*, 65(1) :138–158, 1986.

[VdVM90] H.A. Van der Vorst and J.B.M. Melissen. A Petrov-Galerkin type method for solving $Ax = b$, where A is symmetric complex. *IEEE Trans. Mag.*, 26(2) :706–708, 1990.

[Wab06] M. Wabro. AMGe—coarsening strategies and application to the Oseen equations. *SIAM J. Sci. Comput.*, 27(6) :2077–2097, 2006.

[WE06] T. Warburton and M. Embree. The role of the penalty in the local discontinuous Galerkin method for Maxwell's eigenvalue problem. *Comput. Methods Appl. Mech. Engrg.*, 195(25-28) :3205–3223, 2006.



Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Unité de recherche INRIA Futurs : Parc Club Orsay Université - ZAC des Vignes
4, rue Jacques Monod - 91893 ORSAY Cedex (France)
Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399