


Année : 2010

N° attribué par la bibliothèque

THÈSE

présentée à

UNIVERSITÉ PARIS DAUPHINE

pour obtenir le titre de

DOCTEUR EN SCIENCES

Spécialité

Mathématiques appliquées

soutenue par

Morgane BERGOT

le 22 novembre 2010

Titre

Éléments finis d'ordre élevé pour maillages hybrides

Application à la résolution de systèmes hyperboliques linéaires en régimes harmonique et temporel

Directeur de thèse : Gary COHEN

Jury

Rapporteurs :	Mme	**Christine**	**BERNARDI**
	Mme	**Nilima**	**NIGAM**
Suffragants :	M.	**Patrick**	**CIARLET**
	M.	**Gary**	**COHEN**
	M.	**Marc**	**DURUFLÉ**
	M.	**Xavier**	**FERRIÈRES**
	M.	**Gabriel**	**TURINICI**

« L'Université n'entend donner aucune approbation, ni improbation aux opinions émises dans les thèses : ces opinions doivent être considérées comme propres à leurs auteurs. »

À ma mère.

Remerciements

Je tiens tout d'abord à remercier mon directeur de thèse, Gary Cohen, qui est à l'origine de ce travail. Je reste admirative devant ses talents de persuasion qui m'ont permis d'obtenir un cofinancement entre l'INRIA et le CEA-Gramat. Nos rapports ayant souvent été plus que tendus, je tiens néanmoins à le remercier pour sa patience et sa bienveillance.

Marc Duruflé a dirigé l'essentiel de ma thèse, notamment la partie la plus numérique. Son aide et son soutien ont été primordiaux, et ce travail n'aurait pas pu aboutir sans ses interventions. J'ai travaillé sur son redoutable code, Montjoie, et je ne le remercierai jamais assez de m'avoir rendue possible son utilisation en effectuant sa maintenance et son évolution jour après jour, chaque fois que j'en ai exprimé le besoin. Je le remercie enfin pour ses remarques lors de la rédaction, ainsi que pour la relecture patiente et précise qu'il a faite de ce manuscrit.

Je tiens également à remercier Nilima Nigam d'avoir accepté d'être rapporteur de cette thèse en français. Ses remarques concernant la partie la plus théorique m'ont grandement aidée pour adopter une plus grande rigueur mathématique dans les problèmes d'estimation d'erreur.

Je remercie chaleureusement Xavier Ferrières pour sa bonne humeur et sa diplomatie.

Je sais gré à Christine Bernardi d'avoir bien voulu rapporter ma thèse, et à Gabriel Turinici d'avoir accepté de faire partie de mon jury.

Merci également à Patrick Ciarlet d'avoir accepté d'être président de mon jury, merci aussi pour ses encouragements et ses conseils, ainsi que pour sa relecture de mon premier article.

Mes remerciements vont aussi à tous les membres actuels et anciens du projet POems : la bonne ambiance qu'ils développent au sein du bâtiment 13 et, plus généralement, pour la bonne ambiance de travail qui règne à l'INRIA. Je remercie plus particulièrement Patrick Joly pour ses conseils avisés en diplomatie et ses remarques lors de la préparation de ma soutenance. Merci bien sûr à tous les doctorants de POems : Adrien pour son incroyable gestion de l'administration informatique du projet, Bérangère pour ses discussions philosophiques et les sorties culturelles dans la capitale, Juliette pour m'avoir initiée au chant, Alexandre pour m'avoir supportée comme co-bureau pendant 3ans, Sébastien pour son enthousiasme et sa bonne humeur, Julien pour sa rigueur et sa précision mathématique, Sonia pour son soutien et ses conseils, Aliénor pour nos rigolades dans la navette, et bien sûr tous les stagiaires et post-doc qui sont passés dans le projet.

Je n'oublie pas notre assistance, Nathalie. Merci pour sa disponibilité et son efficacité, sa bonne humeur, et les discussions que nous avons pu avoir.

Je finis cette page par un grand merci à ma famille, mon père pour m'avoir montré un jour la poésie des mathématiques, ma mère pour m'avoir encouragée dans cette voie qu'elle n'a jamais pu comprendre, et à mes amis pour toutes les joies qu'il m'a été donné de vivre pendant ces trois ans.

Table des matières

Introduction

Ce mémoire a pour objet la construction d'élément finis d'ordre élevé, en particulier d'éléments pyramidaux, adaptés à des maillages hybrides pour la résolution de systèmes hyperboliques linéaires en régimes harmonique et temporel. Les travaux ont été menés au sein du projet POems (Propagation des Ondes : Étude Mathématique et Simulation) de l'INRIA (Institut National de Recherche en Informatique et Automatique), laboratoire spécialisé dans l'étude mathématique des problèmes de propagation d'onde. Les études ont été menées pour le CEA-Gramat (Commissariat à l'Énergie Atomique de Gramat), en partenariat avec l'ONERA Toulouse.

Bien que l'objectif du présent travail soit de développer une méthode efficace pour la résolution des équations de Maxwell en régime temporel, nous étudierons l'équation des ondes et l'équation de Helmholtz, ainsi que les équations de Maxwell en régime harmonique. Nous rappelons ici la problématique initiale.

Problématique

La complexité géométrique des dispositifs étudiés en électromagnétisme tend naturellement à rechercher des méthodes de résolution sur des maillages non-structurés. Ceux-ci permettent en effet de décrire de manière très précise les détails des objets, et d'éviter le phénomène de diffraction par des marches d'escalier observé lorsqu'on utilise des maillages structurés pour approcher des géométries complexes. De nombreuses méthodes numériques ont été étudiées pour la résolution des équations de Maxwell en maillage non-structuré.

La modélisation des phénomènes électromagnétiques et ondulatoires est également un sujet ardu en raison des longueurs caractéristiques pouvant être petites devant la taille du domaine. Pour assurer une bonne précision, on doit donc utiliser des méthodes numériques sophistiquées, comme les éléments finis d'ordre élevé. En effet, les éléments finis permettent d'approcher correctement les géométries complexes, et l'ordre élevé permet d'assurer une grande précision pour un coût raisonnable.

Des méthodes d'éléments finis d'ordre élevé utilisant des maillages hexaédriques ont été mises au point par Cohen [17] et ses proches collaborateurs (Fauqueux [19], Pernet et Ferrières [20], [62], Duruflé [28]), méthodes qui ont montré toute leur efficacité par rapport à des méthodes de différences finies ou des éléments finis tétraédriques. Malheureusement, à ce jour, il s'avère difficile de générer de manière automatique un maillage non-structuré purement hexaédrique d'une géométrie complexe. Une technique existante est de mailler la géométrie avec des tétraèdres, et en découpant chaque tétraèdre en quatre hexaèdres, mais cette technique s'avère très pénalisante. En effet, on multiplie alors inutilement le nombre de degrés de libertés, et les éléments obtenus sont déformés, ce qui détériore leurs propriétés numériques (Cohen, Duruflé et Grob [29]).

Une approche pragmatique est d'« utiliser des hexaèdres quand on le peut, et des tétraèdres quand on le doit » (voir Lee, Wong et Lie [50]). Dans cette optique, il existe des outils qui permettent de mailler des géométries quelconques avec un maximum d'hexaèdres, et des tétraèdres en faibles proportions. Afin de faire la transition entre les faces quadrangulaires des hexaèdres et les faces triangulaires des tétraèdres, et ainsi conserver un **maillage conforme**, il faut ajouter des éléments comportant les deux types de faces : des prismes, mais également des pyramides qui sont obligatoires pour générer un maillage hybride conforme (voir par exemple Owen et Saigal [60], Tapp [70], en allemand). Notre but est donc d'étudier des méthodes d'éléments finis sur ces maillages hybrides afin de préserver les excellentes performances des méthodes développées par Cohen *et al.*

Nous rappelons maintenant les méthodes pratiquées sur les tétraèdres, prismes et pyramides.

État de l'art

Approximation continue

Pour l'espace fonctionnel H^1

Pour les équations telles que l'équation des ondes et l'équation de Helmholtz, on peut utiliser une formulation H^1 qui nécessite alors des éléments finis continus adaptés.

Une approche performante pour les problèmes instationnaires serait de construire des **éléments avec condensation de masse** sur tous les types d'éléments. Or si les éléments finis avec condensation de masse sont bien connus pour les hexaèdres (voir Cohen [17]), ils sont moins aboutis pour les autres types d'éléments (voir Mulder et al. [54] pour les tétraèdres). Les éléments proposés requièrent en effet un nombre élevé de degrés de liberté supplémentaires et conduisent à une condition de stabilité plus restrictive. De plus, ces éléments comportent des degrés de liberté supplémentaires sur les faces : dans l'optique de construire des pyramides et des prismes condensés, il faudrait par exemple imposer les points des tétraèdres sur les faces triangulaire et les points de Gauss sur les faces quadrangulaires. L'ensemble de ces contraintes nous semble ainsi compromettre la précision de la formule de quadrature 3D résultante. Pour finir, sur l'équation des ondes, des expériences numériques montrent en outre que ces tétraèdres condensés sont en pratique moins efficaces qu'un schéma implicite (par exemple en utilisant un θ-schéma) utilisant des tétraèdres. Cette approche n'a ainsi pas été explorée

Les **éléments finis nodaux** classiques pour l'approximation continue sont détaillés dans Hesthaven et Teng [44] pour les tétraèdres. En ce qui concerne les prismes, ils sont obtenus de manière classique en formant le produit tensoriel entre une arête avec points de Legendre-Gauss-Lobatto (LGL) sur le segment $[0,1]$, et un triangle avec points « électrostatiques » de Hesthaven [44] comportant des points de LGL sur les arêtes. La construction de fonctions de base pyramidales préservant la conformité avec les autres types d'éléments est plus délicate. Deux approches ont été envisagées.

La première approche pour construire des éléments pyramidaux consiste à utiliser des fonctions de base contenant des fractions rationnelles.

- Bedrosian [5] propose des fonctions de base rationnelles pour des approximations du premier et du second ordre. Zgainski *et al.* [76] conduisent des expériences numériques avec les fonctions de Bedrosian et proposent une famille modifiée de fonctions de bases du second ordre. La même idée est reprise par Graglia *et al.* [38] qui propose une autre famille de fonctions de base du second ordre.
- Chatzi et Preparata [13] introduisent une généralisation des fonctions de base de Bedrosian à un ordre quelconque pour des degrés de liberté régulièrement distribués sur la pyramide

La seconde approche consiste à découper la pyramide en tétraèdres afin d'éviter d'utiliser des fractions rationnelles qui ont la réputation (discutable) d'être difficils à utiliser.

- Au premier ordre, Wieners [74], Knabner et Summ [49], ainsi que Bluck et Walker [7] donnent une famille consistante de fonctions de base qui permet d'assurer la conformité avec les hexaèdres et les tétraèdres, en découpant une pyramide en deux tétraèdres.
- Liu *et al.* [51] proposent une version symétrisée des fonctions de base de Wieners en découpant la pyramide en quatre tétraèdres.

Une alternative aux éléments finis nodaux est l'utilisation d'**éléments finis hiérarchiques**. L'approche *hp* est détaillée par Szabó et Babuška [68], avec par exemple Šolín *et al.* [71] pour les hexaèdres, les tétraèdres et les prismes. Plusieurs articles étendent le concept d'élément fini *hp* aux éléments pyramidaux.

- Warburton [72], Sherwin [65], Sherwin *et al.* [66], ainsi que Karniadakis et Sherwin [47] donnent une famille de fonctions de bases tensorielles pour tous les types d'éléments à partir de la dégénérescence d'un cube.
- Nigam et Phillips [58] proposent une autre famille de fonctions de base en utilisant une pyramide infinie comme élément de référence, et un second espace de dimension inférieure dans un article ultérieur [59].
- Demkowicz *et al.* [23] et Zaglmayr [75] proposent la construction de fonctions de base partiellement orthogonales pour la matrice de rigidité avec des tétraèdres, des hexaèdres et des prismes affines afin d'améliorer le conditionnement de cette matrice. Pour construire l'espace d'approximation pyramidal, les auteurs utilisent la dégénérescence d'un cube.

Pour l'espace fonctionnel $H(rot)$

Les éléments finis pour la formulation $H(rot)$ dédiée à la résolution des équations de Maxwell ont fait l'objet de nombreux travaux. Les **éléments finis d'arête** sont introduits par Nédélec qui propose une première famille dans [56] pour les tétraèdres et les hexaèdres, puis une seconde famille dans [57] pour les tétraèdres, les hexaèdres et les prismes. Monk [55] construit quant à lui des éléments prismatiques de la première famille.

En ce qui concerne les pyramides, l'approche la plus générale consiste à tenter de construire les éléments de la première famille de Nédélec, plus rarement la seconde famille. Beaucoup d'auteurs utilisent pour cela les formes de Whitney.

- Coulomb, Zgainski et Maréchal [22] construisent une première famille pour les ordres 1 et 2, et une seconde famille à l'ordre 1.
- En utilisant les formes de Whitney, Gradinaru et Hiptmair [36] construisent des fonctions de base sur les arêtes pour les éléments d'ordre 1, et Doucet *et al.* [25] proposent un espace d'ordre 1.

- Graglia et Gheorma [38] poursuivent l'étude de Graglia *et al.* [37] sur les tétraèdres et les hexaèdres en construisant des fonctions de base nodale d'ordre quelconque sur les pyramides à partir de fonctions d'arête d'ordre 1 et de points d'interpolation régulièrement répartis.
- Des travaux très théoriques ont été réalisés par Bossavit [10] qui construit des éléments finis sur tous les types d'éléments en utilisant deux opérations simples sur les formes de Whitney [73].
- Zaglmayr citée par Demkowicz [23] propose un espace d'ordre quelconque sur les pyramides en considérant un cube dégénéré. Nigam et Phillips partent d'une pyramide infinie pour construire des fonctions de base $H(rot)$-conformes pour un ordre quelconque dans [58], et un second espace de dimension plus petite que le premier dans [59].

Comme en H^1, certains auteurs préfèrent découper la pyramide en tétraèdres afin d'éviter d'utiliser des fractions rationnelles. Pour les éléments finis d'arête, une technique est proposée par Marais et Davidson [53].

Sur des maillages hexaédriques constitués d'éléments non affines, on observe une perte de l'ordre de convergence pour les éléments de la première comme pour ceux la deuxième famille (voir Duruflé [28] pour les illustrations numériques). Ainsi, Arnold *et al.* [3] pour les quadrangles, et Falk *et al.* [32] pour les hexaèdres d'ordre 1, tentent d'y remédier en construisant des éléments d'arête permettant d'éviter cette perte d'ordre.

Approximation discontinue

Les méthodes de Galerkin discontinues permettent une très grande flexibilité dans le choix des éléments : on peut ainsi aisément mixer les éléments au sein d'un même maillage, traiter des maillages non-conformes, ou encore utiliser un ordre d'approximation différent par élément. Elles induisent de plus naturellement une matrice de masse diagonale par blocs, où un bloc correspond à un élément, ce qui permet d'éluder toute problématique de condensation de masse pour des éléments tétraédriques, prismatiques et pyramidaux.

Les méthodes de Galerkin discontinues ont été largement développées pour des maillages tétraédriques, par exemple par Hesthaven et Warburton [46] pour les équations de Maxwell. Les travaux de Cohen et de ses collaborateurs (Fauqueux [19], Pernet et Ferrières [20], [62], Duruflé [28], [29]) ont quant à eux mis en avant l'efficacité obtenue sur des maillages hexaédriques en exploitant la **tensorisation des fonctions de base**, permettant d'obtenir un gain important par rapport aux tétraèdres sur des ordres élevés.

Concernant les éléments prismatiques et pyramidaux, il est possible d'utiliser les éléments finis développés pour la formulation continue avec l'espace H^1 pour les méthodes de Galerkin discontinues. Cependant, puisque la continuité n'est pas requise, on peut considérer d'autres familles de fonctions de base qui peuvent avoir de meilleures propriétés.

- Un choix attractif de fonctions de base pour tous les types d'éléments est celui proposé par Kirby *et al.* [48] et Warburton [72] qui considèrent différentes fonctions de base orthogonales. L'utilisation de fonctions orthogonales et tensorisées permet en effet de creuser la matrice de masse élémentaire, ce qui induit un gain de temps de calcul important lorsque l'on considère un ordre d'approximation suffisamment élevé.
- Dans des travaux récents, Gassner *et al.* [33] proposent une approche originale de construction d'éléments finis nodaux sur maillages hybrides pour les méthodes de Galerkin discontinues ne nécessitant pas le passage par un élément de référence. Les fonctions de base sont construites directement sur l'élément du maillage.

Réalisation

Pour construire des éléments finis adaptés à chaque formulation, on part de l'idée de Arnold, Boffi et Falk [3] sur les quadrangles pour $H(div)$: pour tout type d'élément K d'un maillage d'arête de longueur moyenne h, on choisit comme espace d'approximation sur K l'espace d'approximation de dimension minimale permettant d'obtenir une erreur d'interpolation en $O(h^r)$ pour la norme de l'espace d'approximation considéré à la fois sur l'élément et sur tout le maillage. Pour trouver cet espace optimal, pour la formulation H^1, on part de la constatation que la démonstration des estimations d'erreur utilise le fait que l'espace $\mathbb{P}_r$ est inclus dans l'espace d'approximation sur l'élément K. Comme tous les calculs sont faits sur l'élément de référence $\hat{K}$ via une transformation F, on va chercher l'espace d'approximation minimal sur $\hat{K}$ tel que cet espace contient $\mathbb{P}_r$ pour tous les types d'éléments. Pour la formulation $H(rot)$, il s'agit d'inclure l'espace $\mathcal{R}_r$, l'espace de la première famille de Nédélec sur les tétraèdres.

Pour la résolution des équations modélisant la propagation d'ondes, on peut considérer deux types d'approximations de Galerkin : une approximation continue (H^1 ou $H(rot)$) et une approximation discontinue. La formulation continue a entre autre l'avantage de nécessiter un nombre moins important de degrés de liberté, mais la matrice de masse, bien que creuse, est coûteuse à inverser ou factoriser. La formulation discontinue, quant à elle, permet d'obtenir une matrice de masse diagonale par blocs, et la matrice de rigidité fait intervenir des termes de flux supplémentaires. Elle nécessite en général des termes de stabilisation (ou pénalisation) pour obtenir des solutions

non parasitées qui font intervenir un paramètre à choisir (voir Hesthaven et Warburton [46] pour les équations de Maxwell).

Pour chaque équation considérée, on peut choisir une formulation en temps d'ordre 1 ou d'ordre 2. Dans le cas d'une formulation d'ordre 1, les PML (Perfect Matching Layers) sont plus faciles à mettre en oeuvre, les schémas temporels sont plus répandus et la formulation discontinue, est moins lourde à implémenter. La formulation d'ordre 2 fait quant à elle intervenir moins de degrés de liberté, et la matrice de rigidité est classique à préconditionner.

Sans prétendre qu'il s'agit du meilleur, notre choix sera donc le suivant :

- schéma d'ordre 1 et formulation discontinue en régime temporel
- schéma d'ordre 2 et formulation continue en régime harmonique

En ce qui concerne la formulation de Galerkin discontinu d'ordre 1, nous utiliserons le schéma Local Discontinuous Galerkin introduit par Cockburn et Shu [16], et développé par Hesthaven et Warburton [45] pour les équations de Maxwell.

Plan

Cette thèse est divisée en cinq parties et quinze chapitres.

Dans la partie I, qui contient le chapitre 1, on donne les notations et rappels théoriques qui serviront dans tout le reste du présent travail. Y seront données en particulier les formulations variationnelles utilisées pour tous les types de systèmes hyperboliques linéaires en régimes harmonique et temporel. L'application du schéma général à l'équation des ondes et aux équations de Maxwell est précisée dans le cas des deux régimes.

La partie II traite de la formulation continue H^1. Dans le chapitre 2, on présente des éléments finis d'ordre r quelconque et « optimaux » au sens de la convergence en norme H^1. Pour les tétraèdres, les prismes et les hexaèdres, on retrouve les espaces polynomiaux classiques utilisés dans la littérature. Pour les éléments pyramidaux, l'espace trouvé est le même que celui de Zaglmayr et Demkowicz [23] et celui de Nigam et Phillips [59]. La preuve d'optimalité est détaillée dans le cas pyramidal. Le chapitre 3 traite des formules de quadrature à utiliser pour calculer les intégrales intervenant dans la construction des matrices du problème et présente les estimations d'erreur pour tous les types d'éléments, en particulier pour les éléments pyramidaux. Dans le chapitre 4, on dégage les propriétés numériques des éléments construits précédemment que l'on compare finalement dans le chapitre 5 à ceux trouvés dans la littérature.

Des éléments finis hybrides pour une formulation discontinue sont construits dans la partie III en partant des éléments présentés dans la partie II. La définition des éléments et un algorithme de construction de la matrice de masse rapide sont donnés dans le chapitre 6, et un produit matrice-vecteur rapide utilisant la structure des éléments est décrit dans le chapitre 7. Une étude numérique des éléments construits fait l'objet du chapitre 8, tandis que la comparaison de différents types d'éléments adaptés à la structure discontinue est faite dans le chapitre 9.

La partie IV traite de la formulation continue $H(rot)$. Les éléments finis « optimaux » en terme de convergence en norme $H(rot)$ sont présentés dans le chapitre 10. L'espace trouvé pour les tétraèdres est l'espace classique des éléments finis d'arête de la première famille, tandis que, pour les prismes et hexaèdres, les espaces trouvés sont nouveaux. Concernant les pyramides, l'espace proposé est également nouveau, et en particulier différent de ceux proposés par Nigam et Phillips [58] [59]. L'optimalité de l'espace est prouvée dans le cas des pyramides. Le chapitre 11 traite des formules de quadrature et des estimations d'erreur pour tous les types d'éléments. On dégage dans le chapitre 12 les propriétés numériques de nos éléments que l'on compare finalement dans le chapitre 13 aux éléments trouvés dans la littérature.

La partie V est dédiée aux résultats numériques. Pour le régime harmonique dans le chapitre 14, et pour le régime temporel dans le chapitre 15, on effectue des expériences numériques dans des cas réels afin de constater l'efficacité des éléments décrits dans les trois parties précédentes, et mettre en valeur les atouts de l'hybride. Dans le cas temporel, deux stratégies combinées permettent d'augmenter la CFL, et par conséquent d'utiliser des pas de temps plus élevés pour accélérer les calculs.

Première partie

Rappels théoriques

Chapitre 1

Équations et formulations variationnelles

On donne ici le cadre général des équations modélisant la propagation d'ondes de différentes natures, en régime harmonique et temporel. On présente également les formulations variationnelles associées pour chacun des régimes. L'application à l'équation des ondes et aux équations de Maxwell, dont il sera plus particulièrement question dans la suite, est détaillée dans chaque cas.

Sommaire

1.1 Espaces fonctionnels et notations

On désigne par Ω, ouvert lipschitzien de $\mathbb{R}^3$, le milieu de propagation, par (x,y,z) le point courant de Ω, et t désigne la variable temporelle. Lorsque $\Omega \neq \mathbb{R}^3$, on désignera par Γ la frontière de Ω, et n la normale sortante de Γ.

On définit les opérateurs suivants

Définition 1.1.1

$$grad\, u = \nabla u = \begin{bmatrix} \dfrac{\partial u}{\partial x} \\ \dfrac{\partial u}{\partial y} \\ \dfrac{\partial u}{\partial z} \end{bmatrix}, \quad u(x,y,z) \in \mathbb{R}$$

$$div\, u = \frac{\partial u_1}{\partial x} + \frac{\partial u_2}{\partial y} + \frac{\partial u_3}{\partial z}, \quad u(x,y,z) \in \mathbb{R}^3$$

et

$$rot\, u = \begin{bmatrix} \dfrac{\partial u_3}{\partial y} - \dfrac{\partial u_2}{\partial z} \\ \dfrac{\partial u_1}{\partial z} - \dfrac{\partial u_3}{\partial x} \\ \dfrac{\partial u_2}{\partial x} - \dfrac{\partial u_1}{\partial y} \end{bmatrix}, \quad u(x,y,z) \in \mathbb{R}^3$$

On rappelle la définition d'espaces de Sobolev qui seront utilisés par la suite

Définition 1.1.2 *On définit*

$$L^2(\Omega) = \{u \in L^2_{loc} \mid \int_\Omega |u|^2 < +\infty\}$$

$$H^m(\Omega) = \{u \in L^2(\Omega) \mid \frac{\partial^\alpha u}{\partial x^\alpha} \in L^2(\Omega),\ |\alpha| \leqslant m\}$$

$$H(rot,\Omega) = \{u \in \left(L^2(\Omega)\right)^3 \mid rot\, u \in \left(L^2(\Omega)\right)^3\}$$

$$H(m,rot,\Omega) = \{u \in (H^m(\Omega))^3 \mid rot\, u \in (H^m(\Omega))^3\}$$

Ces espaces sont équipés des normes usuelles

$$\|u\|^2_{m,\Omega} = \sum_{|\alpha|\leqslant m} \int_\Omega \left|\frac{\partial^\alpha u}{\partial x^\alpha}\right|^2$$

$$\|u\|^2_{rot,\Omega} = \int_\Omega u^2 + \int_\Omega (rot\, u)^2$$

$$\|u\|^2_{m,rot,\Omega} = \sum_{|\alpha|\leqslant m} \int_\Omega \left|\frac{\partial^\alpha u}{\partial x^\alpha}\right|^2 + \sum_{|\alpha|\leqslant m} \int_\Omega \left|\frac{\partial^\alpha rot\, u}{\partial x^\alpha}\right|^2$$

et la semi-norme usuelle de $H^m(\Omega)$ *est*

$$|u|^2_{m,\Omega} = \sum_{|\alpha|=m} \int_\Omega \left|\frac{\partial^\alpha u}{\partial x^\alpha}\right|^2.$$

On rappelle que, de manière évidente, on a

$$\forall u \in H^m(\Omega),\ |u|^2_{m,\Omega} \leqslant \|u\|^2_{m,\Omega}. \tag{1.1.1}$$

Définition 1.1.3 *On définit également les espaces suivants*

$$H^m_0(\Omega) = \textit{fermeture de } \mathcal{C}^\infty_0(\Omega) \textit{ dans } H^m(\Omega)$$
$$H_0(rot,\Omega) = \textit{fermeture de } \mathcal{C}^\infty_0(\Omega) \textit{ dans } H(rot,\Omega)$$

1.2 Systèmes hyperboliques linéaires en régime temporel

1.2.1 Définition du problème

On considère le problème hyperbolique linéaire représentatif suivant (voir par exemple Godlewski et Raviart [30]) avec notre choix de formulation

$$\begin{cases} M\dfrac{\partial u}{\partial t} + \displaystyle\sum_{1\leqslant i\leqslant d} A_i \frac{\partial u}{\partial x_i} + \sum_{1\leqslant i\leqslant d} B_i \frac{\partial u}{\partial x_i} = 0 \qquad (M, A_i, B_i) \in (\mathcal{M}_{n_s}(\mathbb{R}))^3,\ u \in \mathbb{R}^{n_s} \quad \text{dans } \Omega \\ \sum (A_i + B_i)\, n_i\, u = Nu \qquad N \in \mathcal{M}_{n_s}(\mathbb{R}) \quad \text{sur } \Gamma \\ u(x, y, z, t = 0) = u_0(x, y, z) \end{cases} \tag{1.2.1}$$

où n_s est le nombre d'inconnues scalaires de l'équation, et d est la dimension. La matrice N est une matrice qui décrit la condition aux limites, et u_0 est la donnée initiale.

Lorsque le système est symétrique, on a

$$B_i = A_i^*.$$

Remarque 1.2.1 *Cette formulation est très utile pour exhiber l'antisymétrie de la matrice de rigidité lorsque l'on utilise des flux centrés pour une méthode de Galerkin discontinue et sans conditions absorbantes, ce qui est essentiel pour avoir la stabilité.*

1.2.2 Approximation spatiale

1.2.2.1 Formulation variationnelle

Soit Ω un ouvert de $\mathbb{R}^3$, composé de n_e éléments K_i

$$\Omega = \bigcup_{1\leqslant i\leqslant n_e} K_i.$$

Pour tout élément K, on note ∂K la frontière de K, de normale sortante n. La formulation variationnelle s'écrit

$$\begin{cases} \text{Trouver } u \in V \text{ tel que} \\ \forall v \in V, \quad \dfrac{\mathrm{d}}{\mathrm{d}t}\displaystyle\int_K M\, u \cdot v\, dx - \int_K \sum_{1\leqslant i\leqslant d} \left(A_i\, u \cdot \frac{\partial v}{\partial x_i} - B_i \frac{\partial u}{\partial x_i} \cdot v \right) dx \\ \qquad\qquad + \displaystyle\int_{\partial K} (N_1\, \{u\} + N_2\, [u]) \cdot v\, ds = 0 \end{cases} \tag{1.2.2}$$

avec $V = \left(L^2(\Omega)\right)^{n_s}$ et

$$N_1 = \sum_{1\leqslant i\leqslant d} A_i\, n_i \qquad N_2 = \sum_{1\leqslant i\leqslant d} B_i\, n_i$$

La moyenne $\{u\}$ est définie par

$$\{u\} = \frac{1}{2}(u_1 + u_2) \tag{1.2.3}$$

et $[u]$ est défini par

$$[u] = \frac{1}{2}(u_2 - u_1) + \frac{1}{2}\,\alpha\, C(u_2 - u_1), \tag{1.2.4}$$

où C est une matrice symétrique positive, u_1 la valeur de u sur l'élément K et u_2 la valeur de u sur un élément voisin de K, et $\alpha \leqslant 0$ est un facteur de pénalisation. Lorsque $\partial K \in \Gamma$, u_2 n'étant pas défini, on remplace $(N_1 + N_2)u_2$ par Nu_1.

Remarque 1.2.2 *Concernant la pénalisation, le lecteur pourra se référer à Pernet [61], le résultat essentiel étant que l'on a une dispersion en $O(h^{2r})$ sans pénalisation ($\alpha = 0$), et une dispersion en $O(h^{2r+1})$ avec pénalisation ($\alpha < 0$). Pour le choix de α, on se réfèrera à Castel, Cohen et Duruflé [12], l'idée générale étant que le schéma est d'autant plus précis que $|\alpha|$ est grand, mais que la condition de stabilité (CFL) est d'autant plus restrictive lorsque l'on utilise un schéma explicite sur les termes de pénalisation.*

En pratique, on prend $\alpha = -1$ qui correspond au flux de Lax-Friedrichs $C = |\displaystyle\sum_{1\leqslant i\leqslant d} (A_i + B_i)\, n_i|$ lorsque l'on utilise un schéma de Runge-Kutta, et $\alpha = -0.1$ dans le cas d'un schéma de saute-mouton, à cause du problème sur la condition de stabilité soulevé précédemment.

1.2.2.2 Discrétisation

Pour un sous espace d'approximation V_h de dimension $n = n_s\, n_r$ de l'espace V, le problème discret s'écrit

$$\left\{\begin{array}{l} \text{Trouver } u_h \in V_h \text{ tel que} \\ \forall v_h \in V_h, \quad \dfrac{\mathrm{d}}{\mathrm{d}t}\displaystyle\int_K M\, u_h \cdot v_h\, dx - \int_K \sum_{1\leqslant i\leqslant d} \left(A_i\, u_h \cdot \frac{\partial v_h}{\partial x_i} - B_i\, \frac{\partial u_h}{\partial x_i} \cdot v_h \right) dx \\ \qquad\qquad + \displaystyle\int_{\partial K} \left(N_1\, \{u_h\} + N_2\, [u_h]\right) \cdot v_h\, ds = 0. \end{array}\right. \tag{1.2.5}$$

De manière classique, on écrit les intégrales sur un élément de référence noté $\hat{K}$ en utilisant une transformation F (voir par exemple Ciarlet [14]). L'espace d'approximation V_h sur Ω est alors

$$V_h = \left\{ u \in V(\Omega) \mid u_{|K} \in \left(P_r^F(K)\right)^{n_s} \right\}, \tag{1.2.6}$$

où P_r^F est l'espace d'approximation réel d'ordre r sur un élément K du maillage, défini par

$$P_r^F(K) = \left\{ u \mid u \circ F \in \left(\hat{P}_r(\hat{K})\right)^{n_s} \right\}.$$

L'espace d'élément fini $\hat{P}_r$ d'ordre r sur $\hat{K}$ sera défini sur chaque type d'élément dans le chapitre 6. En attendant, on définit les matrices du problème.

Définition 1.2.3 *Soit $(\varphi_i)_{\leqslant i\leqslant n_r}$ une base de V_h, on définit la matrice de masse M_h par*

$$(M_h)_{i,j} = \int_K M\, \varphi_i \cdot \varphi_j\, dx. \tag{1.2.7}$$

La matrice de masse R_h est telle que

$$(R_h)_{i,j} = \int_K \sum_{1\leqslant k\leqslant d} \left(A_k\, \frac{\partial \varphi_j}{\partial x_k} \cdot \varphi_i - B_k\, \varphi_j \cdot \frac{\partial \varphi_i}{\partial x_k} \right) dx \tag{1.2.8}$$

et la matrice de flux S_h est définie par

$$(S_h)_{i,j} = \int_{\partial K} \left(N_1\, \{\varphi_j\} + N_2\, [\varphi_j]\right) \cdot \varphi_i\, ds. \tag{1.2.9}$$

On définit également la matrice K_h

$$K_h = R_h - S_h \tag{1.2.10}$$

Avec la définition 1.2.3, la discrétisation spatiale s'écrit finalement

$$\frac{\mathrm{d}}{\mathrm{d}t} M_h U - K_h U = 0. \tag{1.2.11}$$

1.2.3 Discrétisation temporelle

Pour la discrétisation en temps, on utilise un schéma explicite. Parmi l'ensemble des schémas explicites, on utilisera en particulier le schéma saute-mouton classique, rapide, et le schéma de Runge-Kutta d'ordre 4, plus précis.

Soit Δt le pas de discrétisation en temps, les deux schémas considérés sont

Le schéma saute-mouton classique :

- Si K_h est antisymétrique ($\alpha = 0$, pas de condition absorbante)

$$U^{n+1} = U^{n-1} - 2\,\Delta t\, M_h^{-1} K_h U^n \tag{1.2.12}$$

- Sinon

$$U^{n+1} = U^{n-1} - 2\,\Delta t\, M_h^{-1} (K_h U^n + L_h U^{n-1}) \tag{1.2.13}$$

où L_h est une matrice positive contenant la partie du flux associée à α et aux conditions absorbantes.

Le schéma de Runge-Kutta scheme d'ordre 4 : voir Carpenter et Kennedy [11]

$$\begin{array}{l} U^{n+1} = U^n \\ \rho = U^n \\ \text{for } i = 1 \text{ to } 5 \\ \qquad \rho = \alpha_i \rho + \Delta t (M_h)^{-1} (R_h - S_h)(U^{n+1}) \\ \qquad U^{n+1} = U^{n+1} + \beta_i \rho \\ \text{end for} \end{array} \tag{1.2.14}$$

Dans les deux cas, à chaque pas de temps, on doit
- effectuer les produits matrice-vecteur $R_h U^n$ et $S_h U^n$;
- résoudre le système linéaire $M_h X = Y$.

La matrice de masse étant diagonale par blocs avec des blocs par élément relativement petits, la résolution du système linéaire se fait en utilisant une décomposition de Cholesky LL^* comme il en sera question dans le chapitre 6. Comme on le verra dans le chapitre 7, le produit matrice-vecteur peut se faire de manière rapide avec un choix judicieux de base d'éléments finis.

1.2.4 Applications aux équations

1.2.4.1 Équation des ondes acoustiques

On note $p(x,y,z,t) \in \mathbb{R}$ la pression du fluide, et $v(x,y,z,t) \in \mathbb{R}^3$ la vitesse moyenne des particules dans un volume élémentaire centré autour du point (x,y,z) à l'instant t. La formulation dite « mixte » de l'équation des ondes acoustiques est

$$\left\{ \begin{array}{l} \chi \dfrac{\partial p}{\partial t} + div\, v = f \\ \rho_0 \dfrac{\partial v}{\partial t} + grad\, p = 0 \end{array} \right. \tag{1.2.15}$$

où ρ_0 est la masse volumique du fluide considéré, et $\chi = \dfrac{1}{\rho_0 c^2}$ le coefficient de compressibilité adiabatique du fluide, avec c la célérité des ondes acoustiques.

La formulation variationnelle LDG correspondante est la suivante

$$\left\{ \begin{array}{l} \text{Trouver } (p_h, v_h) \in V_h \text{ tel que} \\ \\ \forall (\varphi, \psi) \in V_h, \quad \left\{ \begin{array}{l} \dfrac{\mathrm{d}}{\mathrm{d}t} \displaystyle\int_K \chi\, p_h\, \varphi\, dx - \int_K v_h \cdot \nabla \varphi\, dx + \int_{\partial K} \{v_h\} \cdot n\, \varphi\, ds = \int f \varphi \\ \dfrac{\mathrm{d}}{\mathrm{d}t} \displaystyle\int_K \rho_0\, v_h \cdot \psi\, dx + \int_K \nabla p_h \cdot \psi\, dx + \int_{\partial K} \psi \cdot n\, [p_h]\, ds = 0 \end{array} \right. \end{array} \right. \tag{1.2.16}$$

où V_h est défini par l'équation 1.2.6, et $\{\varphi\} = \dfrac{\varphi_1 + \varphi_2}{2}$ et $[\varphi] = \dfrac{\varphi_2 - \varphi_1}{2}$.

Avec le formalisme introduit par l'équation 1.2.1, on a donc
- $d = 3$
- $n_s = 4$
- Le vecteur des inconnues u est

$$u = \begin{bmatrix} p \\ v_x \\ v_y \\ v_z \end{bmatrix}$$

- Les matrices M et A_i sont

$$M = \begin{bmatrix} \chi & 0 & 0 & 0 \\ 0 & \rho_0 & 0 & 0 \\ 0 & 0 & \rho_0 & 0 \\ 0 & 0 & 0 & \rho_0 \end{bmatrix}$$

et

$$A_1 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad A_2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad A_3 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Le système étant symétrique, les matrices B_i sont obtenues en prenant la transposée des A_i.

– Concernant les conditions au bord, on a

Condition de Dirichlet : condition du type $u = 0$

$$N = \begin{bmatrix} 0 & n^* \\ -n & 0 \end{bmatrix}$$

Condition de Neumann : condition du type $\dfrac{\partial u}{\partial n} = 0$

$$N = \begin{bmatrix} 0 & -n^* \\ n & 0 \end{bmatrix}$$

Condition absorbante : condition du type $p - Zv \cdot n = 0$, où $Z = \sqrt{\dfrac{\rho_0}{\chi}}$ est l'impédance

$$N = \begin{bmatrix} -Z^{-1} & 0 \\ 0 & -Z\, nn^* \end{bmatrix}$$

1.2.4.2 Équations de Maxwell

Soit $H \in \mathbb{R}^3$ le champ magnétique et $E \in \mathbb{R}^3$ le champ électrique. En l'absence de charges et de courants, les équations de Maxwell sont

$$\begin{cases} \mu \dfrac{\partial H}{\partial t} + rot\, E = 0 \\ \varepsilon \dfrac{\partial E}{\partial t} - rot\, H = f \end{cases} \tag{1.2.17}$$

où μ est la perméabilité magnétique du milieu, et ε est la permittivité diélectrique du milieu.

On utilise la formulation LDG suivante

$$\begin{cases} \text{Trouver } (E, H) \in V_h \text{ tels que} \\ \forall (\varphi, \psi) \in V_h, \begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t} \displaystyle\int_K \varepsilon\, E \cdot \varphi\, dx - \int_K H \cdot rot\, \varphi\, dx - \int_\Gamma n \wedge \{H\} \cdot \varphi\, ds + \alpha \int_\Gamma n \wedge [E] \cdot (n \wedge \varphi) = \int_K f \cdot \varphi\, dx \\ \dfrac{\mathrm{d}}{\mathrm{d}t} \displaystyle\int_K \mu\, H \cdot \psi\, dx + \int_K rot\, E \cdot \psi\, dx + \int_\Gamma n \wedge [E] \cdot \psi\, ds + \alpha \int_\Gamma n \wedge [H] \cdot (n \wedge \psi) = 0 \end{cases} \end{cases} \tag{1.2.18}$$

avec V_h défini par l'équation 1.2.6, $\{\varphi\} = \dfrac{\varphi_1 + \varphi_2}{2}$ et $[\varphi] = \dfrac{\varphi_2 - \varphi_1}{2}$.

Avec le formalisme introduit par l'équation 1.2.1, on a donc

– $d = 3$

– $n_s = 6$

– Le vecteur des inconnues u est

$$u = \begin{bmatrix} E_x \\ E_y \\ E_z \\ H_x \\ H_y \\ H_z \end{bmatrix}$$

– Les matrices M et A_i sont

$$M = \begin{bmatrix} \varepsilon & 0 & 0 & 0 & 0 & 0 \\ 0 & \varepsilon & 0 & 0 & 0 & 0 \\ 0 & 0 & \varepsilon & 0 & 0 & 0 \\ 0 & 0 & 0 & \mu & 0 & 0 \\ 0 & 0 & 0 & 0 & \mu & 0 \\ 0 & 0 & 0 & 0 & 0 & \mu \end{bmatrix}$$

et

$$A_1 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \quad A_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad A_3 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Le système étant symétrique, les matrices B_i sont obtenues en prenant la transposée des A_i.

– Concernant les conditions au bord, en posant

$$S = \begin{bmatrix} 0 & n_3 & -n_2 \\ -n_3 & 0 & n_1 \\ n_2 & -n_1 & 0 \end{bmatrix}$$

et I la matrice identité de $\mathcal{M}_3(\mathbb{R})$, on a

Condition de Dirichlet : condition de conducteur parfait du type $E \wedge n = 0$

$$N = \begin{bmatrix} 0 & S \\ S & 0 \end{bmatrix}$$

Condition de Neumann : condition du type $H \wedge n = 0$

$$N = \begin{bmatrix} 0 & -S \\ -S & 0 \end{bmatrix}$$

Condition absorbante : condition de Silver-Müller du type $H \wedge n - Z^{-1}(n \wedge E) \wedge n = 0$ où $Z = \sqrt{\frac{\mu}{\varepsilon}}$ est l'impédance

$$N = \begin{bmatrix} -Z^{-1}(I - nn^*) & 0 \\ 0 & -Z(I - nn^*) \end{bmatrix}$$

1.3 Systèmes hyperboliques linéaires en régime harmonique

1.3.1 Définition du problème

Le cas du régime harmonique se met moins facilement sous une forme générale permettant d'englober le cas H^1 et le cas $H(rot)$. On considère le problème représentatif sous la forme suivante, avec notre choix de formulation

$$\begin{cases} -\omega^2 M u + C^* A C u = f \qquad (M, A) \in (\mathcal{M}_{n_s}(\mathbb{R}))^2, \, u \in \mathbb{R}^{n_s} \\ \\ + \text{ Conditions aux bords} \end{cases} \tag{1.3.1}$$

où C est un opérateur du premier ordre, n_s est le nombre d'inconnues scalaires de l'équation, et d est la dimension.

1.3.2 Approximation spatiale

1.3.2.1 Formulation variationnelle

Soit Ω un ouvert de $\mathbb{R}^3$, composé de n_e éléments K_i

$$\Omega = \bigcup_{1 \leqslant i \leqslant n_e} K_i.$$

Pour tout élément K, on note ∂K la frontière de K, de normale sortante n. La formulation variationnelle s'écrit

$$\begin{cases} \text{Trouver } u \in V \text{ tel que} \\ \\ \forall v \in V, \quad -\omega^2 \displaystyle\int_K M u \cdot v \, dx + \int_K A C u \cdot C v \, dx = 0 \end{cases} \tag{1.3.2}$$

avec $V = \left(H^1(\Omega)\right)^{n_s}$ ou $V = (H(rot, \Omega))$ suivant l'équation considérée.

Remarque 1.3.1 *Dans le cas d'une condition de Dirichlet homogène sur* Γ*, on prend* $V = \left(H_0^1(\Omega)\right)^{n_s}$ *ou* $V = (H_0(rot, \Omega))$ *suivant l'équation considérée.*

1.3.2.2 Discrétisation

Pour un sous espace d'approximation V_h de dimension $n = n_s\, n_r$ de l'espace V, le problème discret s'écrit, pour une formulation continue

$$\begin{cases} \text{Trouver } u_h \in V_h \text{ tel que} \\ \forall v_h \in V_h, \quad -\omega^2 \int_K M\, u_h \cdot v_h\, dx + \int_K A\, C\, u_h \cdot C\, v_h\, dx = 0. \end{cases} \tag{1.3.3}$$

Comme dans le cas temporel, on écrit alors les intégrales sur un élément de référence noté $\hat{K}$ en utilisant une transformation F. L'espace d'approximation V_h sur Ω est alors

$$V_h = \left\{ u \in V(\Omega) \mid u_{|K} \in \left(P_r^F(K)\right)^{n_s} \right\},$$

où P_r^F est l'espace d'approximation réel d'ordre r sur un élément K du maillage. On définit P_r^F à l'aide d'une transformation conforme pour chaque type d'espace (voir Monk [55]) par

- Si $V = \left(H^1(\Omega)\right)^{n_s}$

$$P_r^F(K) = \left\{ u \mid u \circ F \in \left(\hat{P}_r(\hat{K})\right)^{n_s} \right\}$$

- Si $V = (H(rot, \Omega))^{n_s}$

$$P_r^F(K) = \left\{ u \mid DF^* u \circ F \in \left(\hat{P}_r(\hat{K})\right)^{n_s} \right\}$$

où DF est le jacobien de la transformation F.

Là encore, l'espace d'élément fini $\hat{P}_r$ d'ordre r sur $\hat{K}$ sera détaillé sur chaque type d'élément dans le chapitre 2 pour H^1, et 10 pour $H(rot)$. En attendant, on définit les matrices du problème.

Définition 1.3.2 *Soit $(\varphi_i)_{\leqslant i \leqslant n_r}$ une base de V_h, on définit la matrice de masse M_h par*

$$(M_h)_{i,j} = \int_K M\, \varphi_i \cdot \varphi_j\, dx. \tag{1.3.4}$$

La matrice de masse R_h est telle que

$$(R_h)_{i,j} = \int_K A\, C\, \varphi_i \cdot C\, \varphi_j\, dx \tag{1.3.5}$$

Avec la définition 1.3.2 des matrices, on obtient le système discret suivant

$$-\omega^2 M_h U + R_h U = 0, \tag{1.3.6}$$

1.3.3 Résolution du système linéaire

Pour l'équation de Helmholtz, le système à résoudre étant de grande taille, on utilise une méthode de résolution itérative. Concernant la théorie générale de la résolution des grands systèmes linéaires creux, on pourra se reporter aux travaux de Hackbusch [39], [40] et Saad [64]. Pour résoudre le système linéaire, on utilise les solveurs COCG ou BICGCR de Clemens-Weiland [15]) auxquels on peut adjoindre ou non une étape de préconditionnement.

L'étape de préconditionnement est faite par une itération p-multigrille, en utilisant l'équation de Helmholtz avec terme d'amortissement (voir Erlangga [31] pour les différences finies, et Duruflé [28] pour les éléments finis).

Concernant les équations de Maxwell, on utilise exclusivement un solveur direct (Pastix) pour factoriser la matrice éléments finis.

Différents types de préconditionneurs peuvent être utilisés sur les équations de Maxwell avec terme d'amortissement pour obtenir des algorithmes stables. Le lecteur pourra consulter le chapitre 5 de la thèse de Duruflé [28] dans lequel sont étudiés différents types de préconditionneurs spécifiquement adaptés aux équations de Maxwell.

1.3.4 Applications aux équations

1.3.4.1 Équation de Helmholtz

$$-\omega^2 \chi p - div \frac{1}{\rho_0} grad\, p = f \tag{1.3.7}$$

où ρ_0 est la masse volumique du fluide considéré, et $\chi = \dfrac{1}{\rho_0 c^2}$ le coefficient de compressibilité adiabatique du fluide, avec c la célérité des ondes acoustiques.

La formulation variationnelle est donc la suivante

$$\left\{ \begin{array}{l} \text{Trouver } p \in H^1(\Omega) \text{ tel que} \\ \forall \varphi \in H^1(\Omega),\ -\omega^2 \displaystyle\int_K \chi p\, \varphi\, dx + \int_K \frac{1}{\rho_0} \nabla p \nabla \varphi\, dx = \int_K f\, \varphi\, dx \end{array} \right. \tag{1.3.8}$$

Avec le formalisme utilisé,

$$C = grad$$
$$M = \chi, \qquad A = \frac{1}{\rho_0}$$

On considèrera les conditions au bord suivantes

Condition de Dirichlet : du type $u = 0$ (condition essentielle, qui s'intègre dans l'espace d'approximation)

Condition de Neumann : du type $\dfrac{\partial u}{\partial n} = 0$ (condition naturelle)

Condition absorbante : du type $\dfrac{\partial p}{\partial n} - \dfrac{i\omega}{c} p = 0$ (condition faisant apparaitre le terme $-i\omega \displaystyle\int_\Gamma \frac{1}{c\rho_0} p\varphi\, dx$)

1.3.4.2 Équations de Maxwell

$$-\omega^2 \varepsilon E + rot \frac{1}{\mu} rot\, E = f \tag{1.3.9}$$

La formulation variationnelle est donc la suivante

$$\left\{ \begin{array}{l} \text{Trouver } E \in H(rot, \Omega) \text{ tel que} \\ \forall \varphi \in H(rot, \Omega),\ -\omega^2 \displaystyle\int_K \varepsilon\, E\, \varphi\, dx + \int_K \frac{1}{\mu}\, rot\, E\, rot\, \varphi\, dx = \int_K f\, \varphi\, dx \end{array} \right. \tag{1.3.10}$$

Avec le formalisme utilisé,

$$C = rot$$
$$M = \begin{bmatrix} \varepsilon & 0 & 0 \\ 0 & \varepsilon & 0 \\ 0 & 0 & \varepsilon \end{bmatrix}, \qquad A = \begin{bmatrix} \mu^{-1} & 0 & 0 \\ 0 & \mu^{-1} & 0 \\ 0 & 0 & \mu^{-1} \end{bmatrix}$$

On considèrera les conditions au bord suivantes

Condition de Dirichlet : du type $E \wedge n = 0$ (condition essentielle, qui s'intègre dans l'espace d'approximation)

Condition de Neumann : du type $rot\, E \wedge n = 0$ (condition naturelle)

Condition absorbante (Silver-Müller) : du type $rot\, E \wedge n - \dfrac{i\omega}{c}(n \wedge E) \wedge n = 0$ (condition faisant apparaitre le terme $-i\omega \displaystyle\int_\Gamma \frac{1}{c\mu}(E \wedge n) \cdot (\varphi \wedge n)\, dx$)

Deuxième partie

Éléments finis pour une formulation continue

Chapitre 2

Éléments finis d'ordre arbitrairement élevé

Le but est ici de construire des éléments (K, P_r^F, Σ_r) sur chaque type d'élément (hexaèdre, prisme, pyramide et tétraèdre), tels que l'on ait continuité aux interfaces des éléments du maillage. Nous nous proposons plus particulièrement de donner l'espace d'approximation « optimal » au sens de la convergence, et des degrés de liberté permettant de préserver la conformité entre les différents types d'élément pour les approches nodale et hp, et ce pour chaque type d'élément. L'accent sera porté sur les pyramides qui sont des éléments relativement nouveaux.

Sommaire

2.1 Définition des éléments

2.1.1 Élément droit

La première question que l'on se pose en étudiant un élément géométrique non régulier est de savoir comment on peut le définir. Or, si la réponse est simple pour le cas du tétraèdre, à savoir que quatre points non coplanaires dans l'espace forment un tétraèdre, elle est moins évidente pour un hexaèdre, un prisme ou une pyramide. En effet, comme l'illustre la figure 2.1, il existe plus d'une façon de relier 8 points dans l'espace pour former ce que l'on s'imagine être un hexaèdre : intuitivement, l'exemple 1 est un « meilleur hexaèdre » que celui de l'exemple 2, mais quantifier ce « meilleur » n'est pas évident.

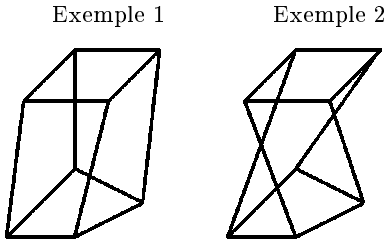

FIG. 2.1 – Exemples illustrant le problème lié à la définition de l'élément hexaédrique

Une réponse à été donnée pour les hexaèdres par Duruflé *et al.* [29], et c'est l'idée de cette définition que nous garderons pour construire nos éléments.

Définition 2.1.1 *On définit un élément $K(x,y,z)$, tétraèdre, pyramide, prisme ou hexaèdre, comme l'image de l'élément référence $\hat{K}(\hat{x},\hat{y},\hat{z})$ par la transformation F définie par :*

$$F = \sum_{1 \leqslant i \leqslant n} S_i \, \hat{\varphi}_i^1, \tag{2.1.1}$$

où les $S_i = (x_i, y_i, z_i)$ désignent les sommets de l'élément K, n est le nombre de sommets de l'élément, et les $\hat{\varphi}_i^1$ sont les fonctions de base correspondant aux fonctions de base de l'espace d'approximation d'ordre 1, lorsque F est inversible.

Lorsque F est inversible, on dit que la transformation est admissible.

- ***Tétraèdre :*** *(cf figure 2.2)*
  - *$n = 4$;*
  - *$\hat{K}$ est le tétraèdre droit unitaire ;*
  - *Les fonctions de base d'ordre 1 sont :*

$$\begin{cases} \hat{\varphi}_1^1 = (1 - \hat{x} - \hat{y} - \hat{z}) \\ \hat{\varphi}_2^1 = \hat{x} \\ \hat{\varphi}_3^1 = \hat{y} \\ \hat{\varphi}_4^1 = \hat{z} \end{cases}$$

- ***Pyramide :*** *(cf figure 2.3)*
  - *$n = 5$;*
  - *$\hat{K}$ est la pyramide unité symétrique centrée à l'origine ;*
  - *Les fonctions de base d'ordre 1 sont :*

$$\begin{cases} \hat{\varphi}_1^1 = \dfrac{1}{4}\left(1 - \hat{x} - \hat{y} - \hat{z} + \dfrac{\hat{x}\hat{y}}{1-\hat{z}}\right) \\ \hat{\varphi}_2^1 = \dfrac{1}{4}\left(1 + \hat{x} - \hat{y} - \hat{z} - \dfrac{\hat{x}\hat{y}}{1-\hat{z}}\right) \\ \hat{\varphi}_3^1 = \dfrac{1}{4}\left(1 + \hat{x} + \hat{y} - \hat{z} + \dfrac{\hat{x}\hat{y}}{1-\hat{z}}\right) \\ \hat{\varphi}_4^1 = \dfrac{1}{4}\left(1 - \hat{x} + \hat{y} - \hat{z} - \dfrac{\hat{x}\hat{y}}{1-\hat{z}}\right) \\ \hat{\varphi}_5^1 = \hat{z} \end{cases}$$

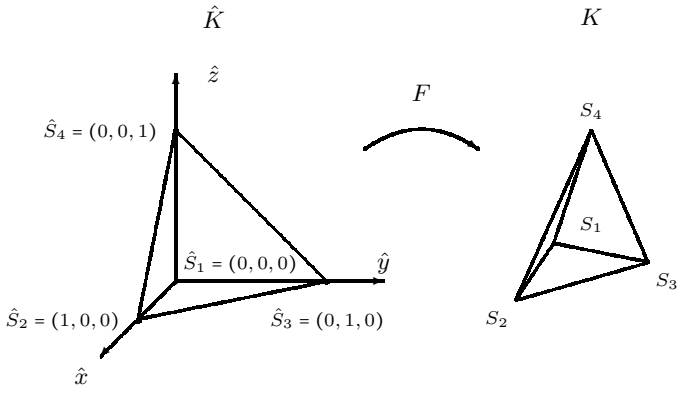

FIG. 2.2 – Passage du tétraèdre de référence $\hat{K}$ vers le tétraèdre K via la transformation F

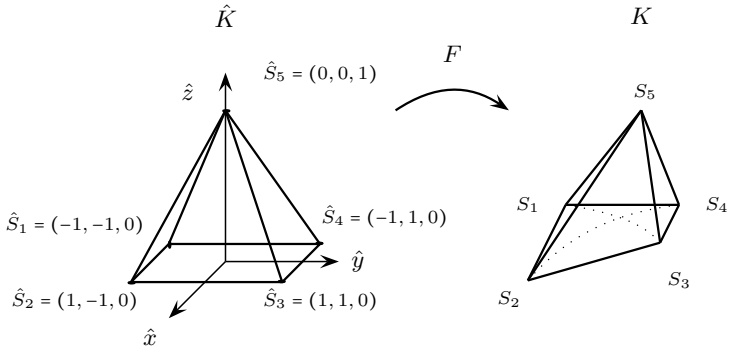

FIG. 2.3 – Passage de la pyramide de référence $\hat{K}$ à la pyramide K via la transformation F

- ***Prisme :*** *(cf figure 2.4)*
  - $n = 6$;
  - $\hat{K}$ *est le prisme droit à base triangulaire ;*
  - *Les fonctions de base d'ordre 1 sont :*

$$\begin{cases} \hat{\varphi}_1^1 = (1-\hat{x}-\hat{y})(1-\hat{z}) \\ \hat{\varphi}_2^1 = \hat{x}(1-\hat{z}) \\ \hat{\varphi}_3^1 = \hat{y}(1-\hat{z}) \\ \hat{\varphi}_4^1 = (1-\hat{x}-\hat{y})\hat{z} \\ \hat{\varphi}_5^1 = \hat{x}\hat{z} \\ \hat{\varphi}_6^1 = \hat{y}\hat{z} \end{cases}$$

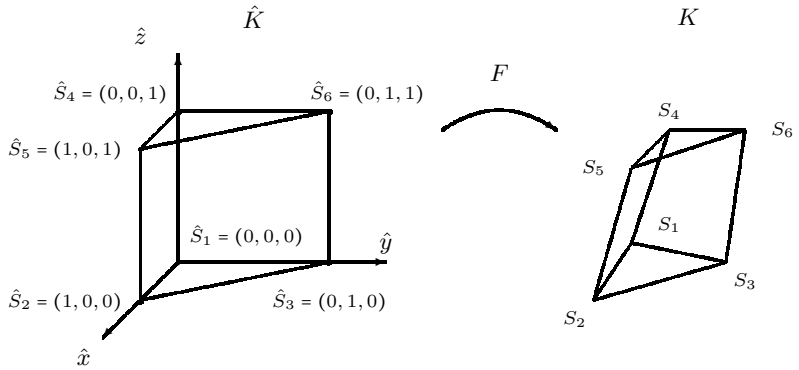

FIG. 2.4 – Passage du prisme de référence $\hat{K}$ vers le prisme K via la transformation F

- ***Hexaèdre :*** *(voir figure 2.5)*
  - $n = 8$;
  - $\hat{K}$ *est le cube unité ;*
  - *Les fonctions de base d'ordre 1 sont :*

$$\begin{cases} \hat{\varphi}_1^1 = (1-\hat{x})(1-\hat{y})(1-\hat{z}) \\ \hat{\varphi}_2^1 = \hat{x}(1-\hat{y})(1-\hat{z}) \\ \hat{\varphi}_3^1 = \hat{x}\hat{y}(1-\hat{z}) \\ \hat{\varphi}_4^1 = (1-\hat{x})\hat{y}(1-\hat{z}) \\ \hat{\varphi}_5^1 = (1-\hat{x})(1-\hat{y})\hat{z} \\ \hat{\varphi}_6^1 = \hat{x}(1-\hat{y})\hat{z} \\ \hat{\varphi}_7^1 = \hat{x}\hat{y}\hat{z} \\ \hat{\varphi}_8^1 = (1-\hat{x})\hat{y}\hat{z} \end{cases}$$

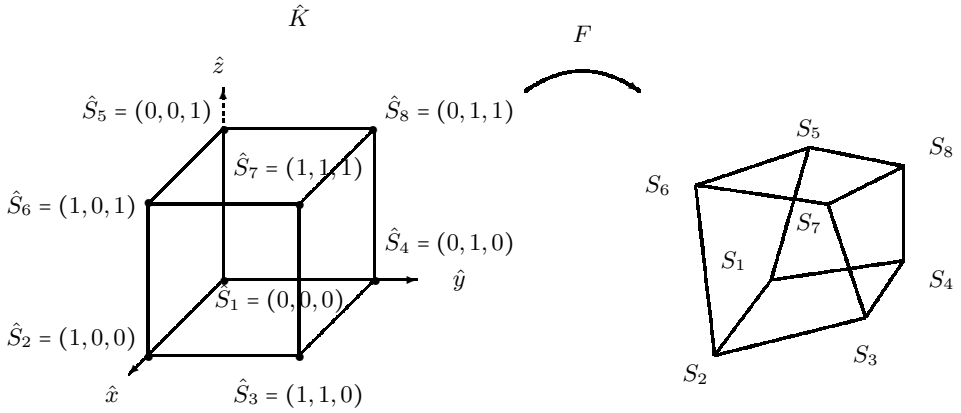

FIG. 2.5 – Passage de l'hexaèdre de référence $\hat{K}$ vers l'hexaèdre K via la transformation F

Pour en revenir à l'exemple de la figure 2.1, le cas de l'exemple 2 correspondrait à une transformation non inversible, ce qui vient confirmer l'intuition selon laquelle plus un élément est « droit », « meilleur » il est. Le cas d'une transformation non inversible peut par ailleurs arriver dans le cas d'un élément dégénéré, comme par exemple lorsque plus de 4 sommets sont coplanaires. Une étude de l'inversibilité du jacobien a été faite en trois parties par Zhang [77], [78], [79], mais la caractérisation des éléments pour lesquelles F est inversible est une question ouverte pour les prismes et les pyramides. Dans la suite, nous supposerons toujours que F est inversible.

Remarque 2.1.2 *Pour les pyramides, la transformation F utilisant les fractions rationnelles est donnée par Bedrosian [5]. On peut retrouver ces fonctions de base en définissant une transformation T permettant de passer du cube unité $\widetilde{Q}$ à la pyramide de référence $\hat{K}$*

$$T : \begin{cases} \hat{x} = (1-\widetilde{z})(2\widetilde{x}-1) \\ \hat{y} = (1-\widetilde{z})(2\widetilde{y}-1) \\ \hat{z} = \widetilde{z}. \end{cases} \tag{2.1.2}$$

Pour les fonctions de base de l'hexaèdres $\varphi(\widetilde{x},\widetilde{y},\widetilde{z}) = (1-\widetilde{x})(1-\widetilde{y})(1-\widetilde{z})$, la transformation T donne en effet

$$\varphi \circ T^{-1}(\hat{x},\hat{y},\hat{z}) = \frac{1}{4}\frac{(1-\hat{x}-\hat{z})(1-\hat{y}-\hat{z})}{1-\hat{z}} = \hat{\varphi}_1^1(\hat{x},\hat{y},\hat{z}).$$

De la même manière, on peut trouver les autres fonctions de Bedrosian.

Nous allons maintenant faire quelques remarques sur la transformation des différents éléments en utilisant la définition 2.1.1.

Définition 2.1.3 *On définit les espaces polynomiaux suivants*

$$\begin{aligned}
\mathbb{P}_r(x,y,z) &= \left\{x^i y^j z^k, 0 \leqslant i,j,k \leqslant r, i+j+k \leqslant r\right\} \\
\mathbb{Q}_r(x,y,z) &= \left\{x^i y^j z^k, 0 \leqslant i,j,k \leqslant r\right\} \\
\mathbb{Q}_{m,n,p}(x,y,z) &= \left\{x^i y^j z^k, 0 \leqslant i \leqslant m, 0 \leqslant j \leqslant n, 0 \leqslant k \leqslant p\right\}
\end{aligned}$$

Proposition 2.1.4 *La transformation F pour chaque type d'élément est dans les espaces suivants :*

- ***Hexaèdres :***

$$F \in (\mathbb{Q}_1(x,y,z))^3$$

F est affine lorsque K est un parallélépipède;

- ***Prismes :***

$$F \in (\mathbb{P}_1(x,y) \otimes \mathbb{P}_1(z))^3$$

F est affine lorsque les bases triangulaires de K sont translatées l'une par rapport à l'autre;

- ***Pyramides :***

$$F \in \left(\mathbb{P}_1(x,y,z) + \left\{\frac{xy}{1-z}\right\}\right)^3$$

F est affine lorsque la base de K est un parallélogramme;

- ***Tétraèdres :***

$$F \in (\mathbb{P}_1(x,y,z))^3$$

F est affine dans tous les cas.

Preuve. Avec la définition 2.1.1, les transformations F peuvent être écrites explicitement comme suit :

- **Hexaèdres :**

$$\begin{aligned}
F = \; & S_1 + (-S_1+S_2)\,\hat{x} + (-S_1+S_4)\,\hat{y} + (-S_1+S_5)\,\hat{z} \\
& + (S_1-S_4-S_5+S_8)\,\hat{y}\hat{z} + (S_1-S_2+S_3-S_4)\,\hat{x}\hat{y} + (S_1-S_2-S_5+S_6)\,\hat{x}\hat{z} \\
& + (-S_1+S_2-S_3+S_4+S_5-S_6+S_7-S_8)\,\hat{x}\hat{y}\hat{z}.
\end{aligned}$$

Lorsque

$$S_1 - S_2 + S_3 - S_4 = S_1 - S_4 - S_5 + S_8 = S_1 - S_2 - S_5 + S_6 = 0,$$

on a

$$-S_1 + S_2 - S_3 + S_4 + S_5 - S_6 + S_7 - S_8 = 0,$$

i.e. la transformation est affine lorsque toutes les faces de K sont des parallélogrammes.

- **Prismes :**

$$\begin{aligned}
F = \; & S_1 + (-S_1+S_2)\,\hat{x} + (-S_1+S_3)\,\hat{y} + (-S_1+S_4)\,\hat{z} \\
& + (S_1-S_2-S_4+S_5)\,\hat{x}\hat{z} + (S_1-S_3-S_4+S_6)\,\hat{y}\hat{z}.
\end{aligned}$$

La transformation est affine lorsque

$$S_1 + S_5 = S_2 + S_4 \text{ et } S_1 + S_6 = S_3 + S_4$$

i.e. lorsque les trois faces quadrangulaires du prisme sont des parallélogrammes.

- **Pyramides :**

$$\begin{aligned}
4F = \; & (S_1+S_2+S_3+S_4) + \hat{x}\,(-S_1+S_2+S_3-S_4) + \hat{y}\,(-S_1-S_2+S_3+S_4) \\
& + \hat{z}\,(4S_5-S_1-S_2-S_3-S_4) + \frac{\hat{x}\hat{y}}{1-\hat{z}}(S_1+S_3-S_2-S_4).
\end{aligned}$$

La transformation est affine lorsque

$$S_1 + S_3 = S_2 + S_4,$$

i.e. lorsque la base de la pyramide est un parallélogramme.

- **Tétraèdres :**

$$F = S_1 + \hat{x}\,(-S_1+S_2) + \hat{y}\,(-S_1+S_3) + \hat{z}\,(-S_1+S_4)$$

qui est affine, ce qui achève la démonstration. □

2.1.2 Élément courbe isoparamétrique

La construction d'éléments courbes est bien connue pour les hexaèdres, les prismes et les tétraèdres (voir Šolín *et al.* [71]), et l'extension aux éléments pyramidaux est immédiate en suivant le même principe. On ne considère ici que le cas isoparamétrique, c'est à dire que, comme pour les éléments droits, F s'écrit sous la forme

$$F = \sum_{1 \leqslant i \leqslant n_s} a_i \, \hat{\varphi}_i^r$$

où les $\hat{\varphi}_i^r$ sont les fonctions de base de l'espace d'approximation d'ordre r (voir section 2.3).

2.2 Espace d'approximation d'ordre r

2.2.1 Espace d'approximation optimal sur l'élément de référence

L'espace d'approximation V_h sur un ouvert Ω de $\mathbb{R}^3$ est donné par

$$V_h = \{u \in H^1(\Omega) \mid u_{|K} \in P_r^F(K)\},$$

où P_r^F est l'espace d'approximation d'ordre r de l'espace réel restreint à un élément K quelconque du maillage. En utilisant la transformation H^1-conforme donnée par Monk [55],

$$\hat{\varphi}_i = \varphi_i \circ F^{-1}, \tag{2.2.1}$$

cet espace est défini, par

$$P_r^F(K) = \{u \mid u \circ F \in \hat{P}_r(\hat{K})\}.$$

L'objectif est de construire un espace d'approximation $\hat{P}_r$ permettant d'avoir une convergence optimale.

Définition 2.2.1 *Pour un élément K d'un maillage d'arête moyenne de longueur h, l'espace d'approximation P_r^F* ***optimal*** *est l'espace d'approximation de dimension minimale tel que $\mathbb{P}_r \subset P_r^F$.*

Or on a le théorème suivant

Theorème 2.2.2 *L'espace d'approximation P_r^F optimal pour un élément K du maillage permet, pour une solution suffisamment régulière, d'obtenir une erreur d'interpolation sur l'élément en $O(h^r)$ pour la norme H^1.*

Preuve. Voir Chapitre 3 sur les estimations d'erreur.

On cherche donc, pour chaque élément, l'espace optimal $\hat{P}_r$ sur l'élément de référence tel que l'on ait $\mathbb{P}_r \subset P_r^F$ sur l'élément du maillage, via la transformation F.

Theorème 2.2.3 *L'espace d'approximation optimal $\hat{P}_r$ d'ordre r tel que l'on a $\mathbb{P}_r \subset P_r^F$ est*

- ***Tétraèdre et transformation F affine :***

$$\boxed{\hat{P}_r = \mathbb{P}_r(\hat{x}, \hat{y}, \hat{z})} \tag{2.2.2}$$

dont la dimension est

$$dim\ \mathbb{P}_r(x, y, z) = \frac{(r+1)(r+2)(r+3)}{6}$$

- ***Hexaèdres :***

$$\boxed{\hat{P}_r = \mathbb{Q}_r(\hat{x}, \hat{y}, \hat{z})} \tag{2.2.3}$$

dont la dimension est

$$dim\ \mathbb{Q}_r(x, y, z) = (r+1)^3$$

- ***Prismes :***

$$\boxed{\hat{P}_r = \mathbb{P}_r(\hat{x}, \hat{y}) \otimes \mathbb{P}_r(\hat{z})} \tag{2.2.4}$$

dont la dimension est

$$dim\ P_r(x, y, z) = \frac{(r+1)^2(r+2)}{2}$$

– ***Pyramides :***

$$\boxed{\hat{P}_r = \mathbb{P}_r(\hat{x},\hat{y},\hat{z}) \oplus \sum_{0\leqslant k\leqslant r-1} \left(\frac{\hat{x}\hat{y}}{1-\hat{z}}\right)^{r-k} \mathbb{P}_k(\hat{x},\hat{y})} \tag{2.2.5}$$

dont la dimension est

$$dim\ \hat{P}_r = \frac{(r+1)(r+2)(2r+3)}{6}.$$

Définition 2.2.4 *On notera*

$$\begin{aligned}\mathbb{W}_r(x,y,z) &= \mathbb{P}_r(x,y)\otimes\mathbb{P}_r(z)\\ \mathbb{B}_r(x,y,z) &= \mathbb{P}_r(x,y,z) \oplus \sum_{0\leqslant k\leqslant r-1}\left(\frac{xy}{1-z}\right)^{r-k}\mathbb{P}_k(x,y)\end{aligned}$$

Preuve. Lorsque F est affine, il est immédiat que

$$\hat{P}_r(\hat{K}) = \mathbb{P}_r(\hat{K}) \Longleftrightarrow P_r^F(K) = \mathbb{P}_r(K).$$

Lorsque l'élément n'est pas affine, on considère un monôme de $\mathbb{P}_r(x,y,z)$: $p = x^i y^j z^k$, avec $0 \leqslant i+j+k \leqslant r$ et on applique la transformation F de chaque élément en utilisant la proposition 2.1.4. On détaille ici le cas de la pyramide, les autres espaces étant beaucoup plus simples à traiter.

On considère le cas $p = x^i$, avec $0 \leqslant i \leqslant r$, le cas général pouvant se déduire aisément à partir de ce cas simple. En utilisant 2.1.4 et en développant, on peut écrire

$$x^i = \left(a_1 + b_1\hat{x} + c_1\hat{y} + d_1\hat{z} + e_1\frac{\hat{x}\hat{y}}{1-\hat{z}}\right)^i = \sum_{0\leqslant m_i+n_i+p_i+q_i\leqslant i} C_i(a_1,b_1,c_1,d_1,e_1)\hat{x}^{m_i}\,\hat{y}^{n_i}\,\hat{z}^{p_i}\left(\frac{\hat{x}\hat{y}}{1-\hat{z}}\right)^{q_i}$$

En écrivant $\hat{z} = 1-(1-\hat{z})$ et en développant $\hat{z}^{p_i}$, on a

$$\hat{z}^{p_i} = \sum_{0\leqslant l_i\leqslant p_i}(-1)^{l_i}\,C_{p_i}^{l_i}\,(1-\hat{z})^{l_i}$$

où les $C_{p_i}^{l_i}$ sont les coefficients binomiaux. On a donc

$$x^i = \sum_{0\leqslant m_i+n_i+p_i+q_i\leqslant i}\ \sum_{0\leqslant l_i\leqslant p_i} C_i(a_1,b_1,c_1,d_1,e_1)\,(-1)^{l_i}\,C_{p_i}^{l_i}\,\hat{x}^{m_i+l_i}\,\hat{y}^{n_i+l_i}\left(\frac{\hat{x}\hat{y}}{1-\hat{z}}\right)^{q_i-l_i}$$

On sépare à présent la partie polynomiale et la partie rationnelle.

– Pour $q_i - l_i \leqslant 0$, on pose $i_1 = m_i + q_i$, $i_2 = n_i + q_i$ et $i_3 = l_i - q_i$ et on a

$$\hat{x}^{m_i+l_i}\,\hat{y}^{n_i+l_i}\left(\frac{\hat{x}\hat{y}}{1-\hat{z}}\right)^{q_i-l_i} = \hat{x}^{i_1}\,\hat{y}^{i_2}\,(1-\hat{z})^{i_3}$$

avec $0 \leqslant i_1 + i_2 + i_3 \leqslant i$.

– Si $q_i - l_i > 0$, on pose $i'_1 = m_i + l_i$, $i'_2 = n_i + l_i$ et $i'_3 = q_i - l_i$ et on a

$$\hat{x}^{m_i+l_i}\,\hat{y}^{n_i+l_i}\left(\frac{\hat{x}\hat{y}}{1-\hat{z}}\right)^{q_i-l_i} = \hat{x}^{i'_1}\,\hat{y}^{i'_2}\left(\frac{\hat{x}\hat{y}}{1-\hat{z}}\right)^{i'_3}$$

où $0 \leqslant i'_1 + i'_2 + i'_3 \leqslant i$ et $i'_3 > 0$.

Ainsi, en posant

$$C_{i_1,i_2,i_3} = \sum_{\substack{0\leqslant m_i+n_i+p_i+q_i\leqslant i\\ 0\leqslant l_i\leqslant p_i\\ i_1=m_i+q_i\\ i_2=n_i+q_i\\ i_3=l_i-q_i}} C_i(a_1,b_1,c_1,d_1,e_1)\,(-1)^{l_i}\,C_{p_i}^{l_i}$$

$$C'_{i'_1,i'_2,i'_3} = \sum_{\substack{0\leqslant m_i+n_i+p_i+q_i\leqslant i\\ 0\leqslant l_i\leqslant p_i\\ i'_1=m_i+l_i\\ i'_2=n_i+l_i\\ i'_3=q_i-l_i}} C_i(a_1,b_1,c_1,d_1,e_1)\,(-1)^{l_i}\,C_{p_i}^{l_i}$$

on obtient

$$\begin{aligned}x^i &= \sum_{0\leqslant i_1+i_2+i_3\leqslant i} C_{i_1,i_2,i_3}\,\hat{x}^{i_1}\,\hat{y}^{i_2}\,(1-\hat{z})^{i_3} + \sum_{1\leqslant i'_3\leqslant i}\ \sum_{0\leqslant i'_1+i'_2\leqslant i-i'_3} C'_{i'_1,i'_2,i'_3}\,\hat{x}^{i'_1}\,\hat{y}^{i'_2}\left(\frac{\hat{x}\hat{y}}{1-\hat{z}}\right)^{i'_3}\\ &= p_i(\hat{x},\hat{y},\hat{z}) + f_i(\hat{x},\hat{y},\hat{z})\end{aligned}$$

- Lorsque p décrit tout $\mathbb{P}_i(\hat{x},\hat{y},\hat{z})$, la partie polynomiale est telle que $p_i \in \mathbb{P}_i(\hat{x},\hat{y},\hat{z})$ et tous les monômes de $\mathbb{P}_i$ apparaissent. Pour $0 \leqslant i \leqslant r$, on obtient donc tous les monômes de $\mathbb{P}_r(\hat{x},\hat{y},\hat{z})$.
- Concernant la partie rationnelle, en posant $i"_3 = i - i'_3$, on a

$$f_i(\hat{x},\hat{y},\hat{z}) = \sum_{1\leqslant i'_3\leqslant i} \sum_{0\leqslant i'_1+i'_2\leqslant i-i'_3} C'_{i'_1,i'_2,i'_3}\,\hat{x}^{i'_1}\,\hat{y}^{i'_2}\left(\frac{\hat{x}\hat{y}}{1-\hat{z}}\right)^{i'_3} = \sum_{1\leqslant i"_3\leqslant i} \sum_{0\leqslant i'_1+i'_2\leqslant i"_3} C'_{i'_1,i'_2,i'_3}\,\hat{x}^{i'_1}\,\hat{y}^{i'_2}\left(\frac{\hat{x}\hat{y}}{1-\hat{z}}\right)^{i-i"_3}$$

c'est à dire $f_i \in \sum_{0\leqslant k\leqslant i-1} \mathbb{P}_k(\hat{x},\hat{y})\left(\frac{\hat{x}\hat{y}}{1-\hat{z}}\right)^{i-k}$. Lorsque p décrit tout $\mathbb{P}_i(\hat{x},\hat{y},\hat{z})$, $f_i \in \sum_{0\leqslant k\leqslant i-1} \mathbb{P}_k(\hat{x},\hat{y})\left(\frac{\hat{x}\hat{y}}{1-\hat{z}}\right)^{i-k}$ et tous les monômes de cet espace apparaissent. Pour $0 \leqslant i \leqslant r$, on obtient donc tous les monômes de $\sum_{0\leqslant k\leqslant r-1} \mathbb{P}_k(\hat{x},\hat{y})\left(\frac{\hat{x}\hat{y}}{1-\hat{z}}\right)^{r-k}$.

On a donc montré que la condition $\mathbb{B}_r \subset \hat{P}_r$ était suffisante pour obtenir l'inclusion $\mathbb{P}_r \subset P_r^F$. Montrons à présent que c'est une condition nécessaire.

On a nécessairement $\mathbb{P}_r \subset \hat{P}_r$ de manière immédiate grâce à la partie affine de F. Concernant la partie rationnelle, on peut procéder par récurrence comme le font Arnold *et al.* dans [2] pour les quadrangles en considérant la transformation suivante

$$F = A + B\hat{x} + C\hat{y} + D\frac{\hat{x}\hat{y}}{1-\hat{z}}$$

On note

$$\mathbb{F}_n(\hat{x},\hat{y},\hat{z}) = \sum_{0\leqslant k\leqslant n-1} \mathbb{P}_k(\hat{x},\hat{y})\left(\frac{\hat{x}\hat{y}}{1-\hat{z}}\right)^{n-k}$$

On procède par récurrence sur l'ordre

- $r = 1$: de manière triviale, $\mathbb{F}_1 = \dfrac{\hat{x}\hat{y}}{1-\hat{z}}$ est nécessaire
- Supposons que $\mathbb{F}_{n-1}$ est nécessaire. Dans ce cas, on a

$$\begin{aligned} x^n &= \left(A_1 + B_1\hat{x} + C_1\hat{y} + D_1\frac{\hat{x}\hat{y}}{1-\hat{z}}\right)^n \\ &= \sum_{0\leqslant k\leqslant n-1} C_n^k\,(A_1 + B_1\hat{x} + C_1\hat{y})^k\,D_1^{n-k}\left(\frac{\hat{x}\hat{y}}{1-\hat{z}}\right)^{n-k} \end{aligned}$$

La partie de plus haut degré qui n'est pas dans $\mathbb{F}_{n-1}$ est alors

$$\sum_{0\leqslant k\leqslant n-1} C_n^k\,(B_1\hat{x} + C_1\hat{y})^k\,D_1^{n-k}\left(\frac{\hat{x}\hat{y}}{1-\hat{z}}\right)^{n-k}$$

En développant, on obtient

$$\sum_{0\leqslant k\leqslant n-1} \sum_{0\leqslant p\leqslant k} C_n^k\,C_k^p\,B_1^p\,C_1^{k-p}\,D_1^{n-k}\,\hat{x}^p\hat{y}^{k-p}\left(\frac{\hat{x}\hat{y}}{1-\hat{z}}\right)^{n-k}$$

On obtient finalement une combinaison linéaire des termes de plus haut degré de $\mathbb{F}_n$ avec des coefficients $C_n^k\,C_k^p\,B_1^p\,C_1^{k-p}\,D_1^{n-k}$ où B_1, C_1 et D_1 sont des constantes pouvant être choisies arbitrairement en modifiant les sommets de la pyramide. Les coefficients sont polynomiaux en B_1, C_1 et D_1, et linéairement indépendants car ils forment une famille libre de $\mathbb{P}_n(B_1, C_1, D_1)$.

On a montré que la condition $\mathbb{B}_r \subset \hat{P}_r$ était nécessaire pour obtenir l'inclusion $\mathbb{P}_r \subset P_r^F$. L'espace de dimension minimale est donc $\hat{P}_r = \mathbb{B}_r$, ce qui achève la démonstration de l'optimalité de l'espace.

Concernant les dimensions, celles de $\mathbb{P}_r(x,y,z)$ et $\mathbb{Q}_r(x,y,z)$ sont des résultats classiques. Reste à traiter les deux autres espaces en utilisant la propriété des sommes directes.

- $\mathbb{W}_r(x,y,z)$:

$$\dim \mathbb{P}_r(\hat{z}) = r+1, \dim \mathbb{P}_r(\hat{x},\hat{y}) = \frac{(r+1)(r+2)}{2},$$

i.e.

$$\dim \mathbb{W}_r(\hat{x},\hat{y},\hat{z}) = \frac{(r+1)^2(r+2)}{2}$$

– $\mathbb{B}_r(x,y,z)$:

$$\dim \sum_{0\leqslant k\leqslant r-1} \left(\frac{\hat{x}\hat{y}}{1-\hat{z}}\right)^{r-k} \mathbb{P}_k(\hat{x},\hat{y}) = \sum_{0\leqslant k\leqslant r-1} \dim \mathbb{P}_k(\hat{x},\hat{y}) = \sum_{0\leqslant k\leqslant r-1} \frac{(k+1)(k+2)}{2} = \frac{r(r+1)(r+2)}{6},$$

i.e.

$$\dim \mathbb{B}_r(\hat{x},\hat{y},\hat{z}) = \dim \mathbb{P}_r(\hat{x},\hat{y},\hat{z}) + \frac{r(r+1)(r+2)}{6} = \frac{(r+1)(r+2)(2r+3)}{6}$$

ce qui achève la démonstration. □

Remarque 2.2.5 *Une preuve similaire a été proposée par Falk, Gatto et Monk dans [32] pour les hexaèdres $H(rot)$ et $H(div)$ d'ordre 1. Pour construire leur espace, les auteurs identifient un certain nombre de coefficients indépendants pour déterminer l'espace. Nous avons procédé de même dans un premier temps pour pouvoir établir une conjecture sur la forme de l'espace, la preuve n'ayant été établie que plus tard. À l'aide du logiciel Maple, il a été possible d'identifier les coefficients indépendants pour les trois premiers ordres et ainsi conjecturer la forme de l'espace pour tout ordre.*

On remarque que l'espace pyramidal n'est pas polynomial, ce qui fait toute la particularité des éléments pyramidaux. En effet, on a le théorème suivant

Theorème 2.2.6 *Il n'existe pas d'espace d'approximation polynomiale pour la formulation H^1 qui permette d'obtenir des restrictions sur les faces égales à celles du tétraèdres pour les faces triangulaires, égales à celles de l'hexaèdre pour la face quadrangulaire.*

Preuve. Voir par exemple Nigam et Phillips [58].

2.2.2 Espace d'approximation optimal sur le cube unité

Comme nous le verrons dans la section 2.3.1.2, il est souvent pratique de considérer chacun des éléments comme un cube dégénéré.

Définition 2.2.7 *Pour chacun des éléments, la transformation permettant de passer du cube unité $\widetilde{Q}$ à l'élément de référence $\hat{K}$ est*

– ***Hexaèdre :***

$$T = Id;$$

– ***Prisme :***

$$T : \begin{cases} \hat{x} = (1-\widetilde{y})\,\widetilde{x} \\ \hat{y} = \widetilde{y} \\ \hat{z} = \widetilde{z} \end{cases}$$

– ***Pyramide :***

$$T : \begin{cases} \hat{x} = (1-\widetilde{z})\,(2\widetilde{x}-1) \\ \hat{y} = (1-\widetilde{z})\,(2\widetilde{y}-1) \\ \hat{z} = \widetilde{z} \end{cases}$$

– ***Tétraèdre :***

$$T : \begin{cases} \hat{x} = (1-\widetilde{y})(1-\widetilde{z})\,\widetilde{x} \\ \hat{y} = (1-\widetilde{z})\,\widetilde{y} \\ \hat{z} = \widetilde{z} \end{cases}$$

Ce changement de variable définit un difféomorphisme de l'ouvert $\widetilde{Q}$ vers l'ouvert $\hat{K}$, et pour toute fonction f, on note

$$\widetilde{f}(\widetilde{x},\widetilde{y},\widetilde{z}) = \hat{f}(\hat{x},\hat{y},\hat{z}),$$

et le changement de variable donne

– Pour le tétraèdre et une transformation F affine :

$$\int_{\hat{K}} \hat{f}(\hat{x},\hat{y},\hat{z})\,d\hat{x}\,d\hat{y}\,d\hat{z} = \int_{\widetilde{Q}} \widetilde{f}(\widetilde{x},\widetilde{y},\widetilde{z})\,(1-\widetilde{y})\,(1-\widetilde{z})^2\,d\widetilde{x}\,d\widetilde{y}\,d\widetilde{z}. \tag{2.2.6}$$

- Pour l'hexaèdre :
$$\int_{\hat K} \hat f(\hat x,\hat y,\hat z)\, d\hat x\, d\hat y\, d\hat z = \int_{\widetilde Q} \widetilde f(\widetilde x,\widetilde y,\widetilde z)\, d\widetilde x\, d\widetilde y\, d\widetilde z. \tag{2.2.7}$$
- Pour le prisme :
$$\int_{\hat K} \hat f(\hat x,\hat y,\hat z)\, d\hat x\, d\hat y\, d\hat z = \int_{\widetilde Q} \widetilde f(\widetilde x,\widetilde y,\widetilde z)\,(1-\widetilde y)\, d\widetilde x\, d\widetilde y\, d\widetilde z. \tag{2.2.8}$$
- Pour la pyramide :
$$\int_{\hat K} \hat f(\hat x,\hat y,\hat z)\, d\hat x\, d\hat y\, d\hat z = \int_{\widetilde Q} 4\,\widetilde f(\widetilde x,\widetilde y,\widetilde z)\,(1-\widetilde z)^2\, d\widetilde x\, d\widetilde y\, d\widetilde z. \tag{2.2.9}$$

On peut ainsi écrire chaque espace d'approximation $\hat P_r$ sur le cube unité après transformation par T.

Proposition 2.2.8 *L'espace optimal d'approximation C_r d'ordre r sur le cube unité $\widetilde Q$ est*
- ***Tétraèdre et transformation F affine :***
$$C_r = \mathbb{P}_r(\hat x,\hat y,\hat z)\circ T = \sum_{0\leqslant k\leqslant r} \mathbb{P}_k(\widetilde x(1-\widetilde y),\widetilde y)\,(1-\widetilde z)^k$$
- ***Hexaèdre :***
$$C_r = \mathbb{Q}_r(\hat x,\hat y,\hat z)\circ T = \mathbb{Q}_r(\widetilde x,\widetilde y,\widetilde z)$$
- ***Prisme :***
$$C_r = \mathbb{W}_r(\hat x,\hat y,\hat z)\circ T = \mathbb{W}_r((1-\widetilde y)\widetilde x,\widetilde y,\widetilde z)$$
- ***Pyramide :***
$$C_r = \mathbb{B}_r(\hat x,\hat y,\hat z)\circ T = \sum_{0\leqslant k\leqslant r} \mathbb{Q}_k(\widetilde x,\widetilde y)(1-\widetilde z)^k$$

Preuve. Le résultat est immédiat par l'application de T pour l'hexaèdre et le prisme.
Vérifions tout d'abord ce qui concerne les pyramides. Avec la transformation T, la partie polynomiale $\mathbb{B}_r$ s'écrit

$$\begin{aligned}\{\hat x^m\hat y^n\hat z^p,\quad 0\leqslant m+n+p\leqslant r\}\circ T &= \left\{(2\widetilde x-1)^m(2\widetilde y-1)^n(1-\widetilde z)^{m+n}\widetilde z^p,\quad 0\leqslant m+n+p\leqslant r\right\}\\ &= \left\{\widetilde x^m\widetilde y^n\widetilde z^p(1-\widetilde z)^{m+n},\ 0\leqslant m+n+p\leqslant r\right\}\subset C_r\end{aligned}$$

La partie rationnelle devient quant à elle

$$\begin{aligned}\left\{\hat x^i\hat y^j\left(\frac{\hat x\hat y}{1-\hat z}\right)^{r-p},\quad 0\leqslant i+j\leqslant p\leqslant r\right\}\circ T &= \left\{(2\widetilde x-1)^{r-p+i}(2\widetilde y-1)^{r-p+j}(1-\widetilde z)^{r-p+i+j},\quad 0\leqslant i+j\leqslant p\leqslant r\right\}\\ &= \left\{\widetilde x^{r-p+i}\widetilde y^{r-p+j}(1-\widetilde z)^{r-p+i+j},\quad 0\leqslant i+j\leqslant p\leqslant r\right\}\subset C_r\end{aligned}$$

c'est à dire $\mathbb{B}_r\circ T\subset C_r$. Or

$$\dim C_r = \sum_{0\leqslant k\leqslant r}(k+1)^2 = \frac{1}{6}(r+1)(r+2)(2r+3) = \dim\mathbb{B}_r = \dim\mathbb{B}_r\circ T.$$

En ce qui concerne les tétraèdres, la transformation T donne

$$C_r = \{\hat x^m\hat y^n\hat z^p,\quad 0\leqslant m+n+p\leqslant r\}\circ T = \{x^m(1-\widetilde y)^m(1-\widetilde z)^m y^n(1-\widetilde z)^n\widetilde z^p,\quad 0\leqslant m+n+p\leqslant r\}$$

Or,

$$\sum_{0\leqslant k\leqslant r}\mathbb{P}_k(\widetilde x(1-\widetilde y),\widetilde y)\,(1-\widetilde z)^k = \left\{\widetilde x^i\,(1-\widetilde y)^i\widetilde y^j\,(1-\widetilde z)^k, 0\leqslant i+j\leqslant k\leqslant r\right\}$$

c'est à dire, en notant $k=i+j+l$, $0\leqslant l\leqslant r$,

$$\sum_{0\leqslant k\leqslant r}\mathbb{P}_k(\widetilde x(1-\widetilde y),\widetilde y)\,(1-\widetilde z)^k = \left\{\widetilde x^i\,(1-\widetilde y)^i(1-\widetilde z)^i\widetilde y^j(1-\widetilde z)^j\,(1-\widetilde z)^l, 0\leqslant i+j+l\leqslant r\right\}$$

qui est précisément C_r, ce qui achève la preuve de la proposition. □

Remarque 2.2.9 *Pour tous les types d'éléments, on a $C_r\subset\mathbb{Q}_r$.*

2.3 Degrés de liberté et fonctions de base

2.3.1 Éléments finis nodaux

2.3.1.1 Localisation des degrés de liberté

On souhaite lier continûment les éléments pyramidaux avec les autres éléments du maillage. Pour cela, on opte pour les éléments les plus performants existants pour les éléments de base (hexaèdres et tétraèdres), et on s'arrange pour que les autres éléments possèdent les mêmes degrés de liberté que les hexaèdres sur les faces quadrangulaire, et les mêmes degrés de liberté que les tétraèdres sur les faces triangulaires.

Les choix suivants ont été faits :
- des hexaèdres avec points de Gauss-Lobatto (GL) (Cohen [17]) ;
- des tétraèdres avec les points calculés partir d'un problème électrostatique par Hesthaven, avec des points de GL sur les arêtes (Hesthaven and Teng [44]) ;
- des prismes obtenus à partir du produit tensoriel entre une face de tétraèdre de Hesthaven (qui correspond à un triangle de Hesthaven [43]) et une arête avec des points de GL : on a ainsi les bonnes propriétés de placement sur les faces, et on bénéficie en outre de la tensorisation dans une direction.

Restent les degrés de liberté sur la pyramide, que l'on place sur les points de GL pour la base quadrangulaire de la pyramide, et sur les points de Hesthaven pour les faces triangulaires. Le nombre de degrés de liberté n_f sur les faces est alors

$$n_f = 3r^2 + 2.$$

On ajoute n_i degrés de liberté à l'intérieur de la pyramide

$$n_i = \frac{1}{6}(r-1)(r-2)(2r-3) = \sum_{1 \leqslant k \leqslant r-2} k^2.$$

On place ces degrés de liberté sur $(r-2)$ plans parallèle, chaque plan contenant k^2 degrés de liberté, comme montré sur la figure 2.6.

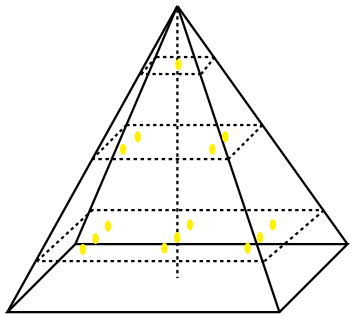

FIG. 2.6 – Localisation des degrés de liberté à l'intérieur de l'élément pyramidal d'ordre 5

Le nombre total de degrés de liberté est alors

$$n_r = n_i + n_f = \frac{1}{6}(r+1)(r+2)(2r+3).$$

qui est la dimension de $\mathbb{B}_r$.

Les degrés de liberté peuvent ainsi être placés de manière systématique sur la pyramide, à n'importe quel degré. Chaque catégorie de point est représentée par une couleur sur la figure 2.7 pour les éléments pyramidaux d'ordre deux à quatre.

2.3.1.2 Fonctions de base

On cherche à présent la base de Lagrange de chaque espace $\hat{P}_r$ sur les éléments de référence $\hat{K}$, les fonctions réelles sur K étant déduites de celles-ci à l'aide de la transformation H^1-conforme (voir Monk [55]), ou transformation de Piolat

$$\hat{p} = p \circ F. \tag{2.3.1}$$

Les fonctions de base $(\hat{\varphi}_i)$ sur l'élément de référence $\hat{K}$ sont obtenues comme suit.

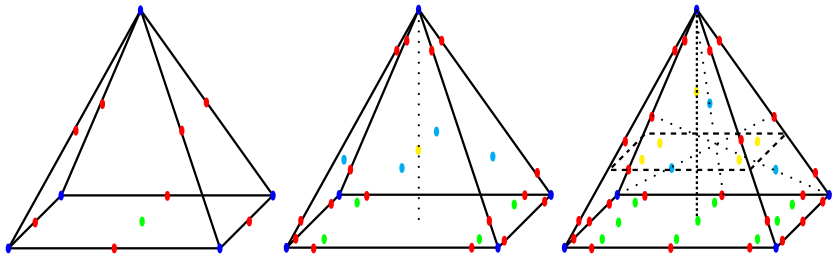

FIG. 2.7 – Localisation des degrés de liberté pour les éléments pyramidaux d'ordre 2, 3 et 4

Définition 2.3.1 *Soit $(\hat{M}_i)_{1\leqslant i\leqslant n_r}$ les coordonnées des points d'interpolation sur l'élément $\hat{K}$, et $(\hat{\psi}_i)_{1\leqslant i\leqslant n_r}$ une base de $\hat{P}_r$. La matrice de Vandermonde $VDM \in \mathcal{M}_{n_r}(\mathbb{R})$ est définie par*

$$VDM_{i,j} = \hat{\psi}_i(\hat{M}_j), \quad 1 \leqslant i,j \leqslant n_r,$$

et la fonction de base $\hat{\varphi}_i$ liée au point d'interpolation $\hat{M}_i$ est alors définie par

$$\hat{\varphi}_i = \sum_{1\leqslant j\leqslant n_r} (VDM^{-1})_{i,j}\, \hat{\psi}_j. \tag{2.3.2}$$

Remarque 2.3.2 *La caractérisation de l'inversibilité de la matrice de Vandermonde reste une question ouverte, mais on observe qu'avec notre choix pour le positionnement des degrés de liberté, la matrice de Vandermonde est inversible, c'est à dire que l'élément est unisolvant.*

On a le choix des $(\hat{\psi}_i)$: on peut prendre les monômes « classiques » de $\hat{P}_r$, mais pour avoir un meilleur conditionnement de la matrice de Vandermonde, on va plutôt chercher une base orthogonale de $\hat{P}_r$.

Proposition 2.3.3 *Les familles de fonctions de base suivantes sont une base « orthogonale » de l'espace $\hat{P}_r$ correspondant au type d'élément considéré*

- ***Hexaèdre :***

$$\hat{\varphi}^{GL}_{i_1}(\hat{x})\,\hat{\varphi}^{GL}_{i_2}(\hat{y})\,\hat{\varphi}^{GL}_{i_3}(\hat{z}), \quad 0 \leqslant i_1,i_2,i_3 \leqslant r,$$

  où

$$\hat{\varphi}^{GL}_i(\hat{x}) = \frac{\prod_{j\neq i} \hat{x} - \xi^{GL}_j}{\prod_{j\neq i} \xi^{GL}_i - \xi^{GL}_j}$$

- ***Prisme :***

$$P^{0,0}_{i_1}\left(\frac{2\hat{x}}{1-\hat{y}} - 1\right)(1-\hat{y})^{i_1} P^{2i_1+1,0}_{i_2}(2\hat{y}-1)\,\varphi^{GL}_{i_3}(\hat{z}), \quad 0 \leqslant i_1+i_2, i_3 \leqslant r,$$

- ***Pyramide :***

$$P^{0,0}_{i_1}\left(\frac{\hat{x}}{1-\hat{z}}\right) P^{0,0}_{i_2}\left(\frac{\hat{y}}{1-\hat{z}}\right)(1-\hat{z})^{max(i_1,i_2)} P^{2max(i_1,i_2)+2,0}_{i_3}(2\hat{z}-1), \quad 0 \leqslant i_1,i_2 \leqslant r,\ 0 \leqslant i_3 \leqslant r - max(i_1,i_2),$$

- ***Tétraèdre :***

$$P^{0,0}_{i_1}\left(\frac{2\hat{x}}{1-\hat{y}-\hat{z}} - 1\right) P^{2i_1+1,0}_{i_2}\left(\frac{2\hat{y}}{1-\hat{z}} - 1\right)(1-\hat{y}-\hat{z})^{i_1} P^{2(i_1+i_2)+2,0}_{i_3}(2\hat{z}-1)\,(1-\hat{z})^{i_2}, \quad 0 \leqslant i_1+i_2+i_3 \leqslant r.$$

où les ξ^{GL}_j désignent les points de Gauss-Lobatto (cf Stroud [67]), et $P^{i,j}_m(x)$ les polynômes de Jacobi ***orthonormalisés*** *d'ordre m, orthogonaux pour les poids $(1-x)^i(1+x)^j$*

Remarque 2.3.4 *En ce qui concerne les hexaèdres et les prismes, on appelle abusivement la base proposée « orthogonale » bien qu'elle ne soit que « pseudo-orthogonale » : la base est orthogonale pour le produit scalaire évalué avec une formule de quadrature avec points de Gauss-Lobatto, c'est à dire*

$$\oint^{GL} \varphi_i\varphi_j = \delta_{ij}$$

Preuve. Pour tous les éléments, on vérifie que les familles sont orthogonales en écrivant les intégrales sur le cube unité $\widetilde{Q}$ par les transformations T associées (cf. définition 2.2.7) .

- **Hexaèdres :** La famille est orthogonale (pseudo-orthogonale) par construction. On a l'inclusion $C_r = \mathbb{Q}_r$ de manière immédiate du fait que l'on a $\varphi_i^{GL}(\eta) \in \mathbb{Q}_r(\eta)$. L'égalité entre les deux espaces finalement s'obtient par un argument de dimension.
- **Prismes :** Pour $0 \leqslant i+j, k \leqslant r$, on note

$$\hat{\psi}_{i,j,k}(\hat{x},\hat{y},\hat{z}) = P_i^{0,0}\left(\frac{2\hat{x}}{1-\hat{y}} - 1\right)(1-\hat{y})^i P_j^{2i+1,0}(2\hat{y}-1)\,\varphi_k^{GL}(\hat{z})$$

En utilisant l'équation 2.2.8, avec une formule de quadrature de Gauss-Lobatto pour la partie en $\widetilde{z}$, on a

$$\begin{aligned}\int_{\hat{K}} \hat{\psi}_{i,j,k}(\hat{x},\hat{y},\hat{z}) \quad & \hat{\psi}_{i',j',k'}(\hat{x},\hat{y},\hat{z})\,d\hat{x}\,d\hat{y}\,d\hat{z} = \\ & \underbrace{\int_0^1 P_i^{0,0}(2\widetilde{x}-1)P_{i'}^{0,0}(2\widetilde{x}-1)\,d\widetilde{x}}_{=\delta_{i\,i'}} \underbrace{\int_0^1 \varphi_k^{GL}(\hat{z})\,\varphi_{k'}^{GL}(\hat{z})\,d\widetilde{z}}_{\approx \omega_k^{GL}\delta_{k\,k'}} \\ & \underbrace{\int_0^1 (1-\widetilde{y})^{i+i'+1} P_j^{2i+1,0}(2\widetilde{y}-1)P_{j'}^{2i'+1,0}(2\widetilde{y}-1)\,d\widetilde{y}}_{=\delta_{j\,j'}\text{ lorsque } i=i'}\end{aligned}$$

ce qui prouve que la famille est orthogonale.

Il est clair que, pour $0 \leqslant i+j, k \leqslant r$,

$$\hat{\psi}_{i,j,k}(\hat{x},\hat{y},\hat{z}) \circ T^{-1} = P_i^{0,0}(2\widetilde{x}-1)\,(1-\widetilde{y})^i P_j^{2i+1,0}(2\widetilde{y}-1)\,\varphi_k^{GL} \in C_r,$$

soit, en utilisant un argument de dimension,

$$Span\,\left\{\hat{\psi}_{i,j,k}(\hat{x},\hat{y},\hat{z}),\ 0 \leqslant i+j,k \leqslant r\right\} \circ T^{-1} = C_r,$$

- **Pyramides :** Pour $0 \leqslant i,j \leqslant r$, $0 \leqslant k \leqslant r - max(i,j)$, on note

$$\hat{\psi}_{i,j,k}(\hat{x},\hat{y},\hat{z}) = P_i^{0,0}\left(\frac{\hat{x}}{1-\hat{z}}\right) P_j^{0,0}\left(\frac{\hat{y}}{1-\hat{z}}\right)(1-\hat{z})^{max(i,j)} P_k^{2max(i,j)+2,0}(2\hat{z}-1).$$

En utilisant la transformation (2.1.2) sur $\tilde{Q}_r$, on a

$$\begin{aligned}\int_{\hat{K}} \hat{\psi}_{i,j,k}(\hat{x},\hat{y},\hat{z}) \quad & \hat{\psi}_{i',j',k'}(\hat{x},\hat{y},\hat{z})\,d\hat{x}\,d\hat{y}\,d\hat{z} = \\ & \underbrace{\int_0^1 P_i^{0,0}(2\widetilde{x}-1)P_{i'}^{0,0}(2\widetilde{x}-1)\,d\widetilde{x}}_{=\delta_{i\,i'}} \underbrace{\int_0^1 P_j^{0,0}(2\widetilde{y}-1)P_{j'}^{0,0}(2\widetilde{y}-1)\,d\widetilde{y}}_{=\delta_{j\,j'}} \\ & \underbrace{4\int_0^1 (1-\widetilde{z})^{max(i,j)+max(i',j')+2} P_k^{2max(i,j)+2,0}(2\widetilde{z}-1)P_{k'}^{2max(i',j')+2,0}(2\widetilde{z}-1)\,d\widetilde{z}}_{=\delta_{k\,k'}\text{ lorsque } i=i' \text{ et } j=j'},\end{aligned}$$

la famille est donc bien orthogonale.

D'après l'équation 2.2.9, on a

$$\hat{\psi}_{i,j,k}(\hat{x},\hat{y},\hat{z}) \circ T^{-1} = P_i^{0,0}(2\widetilde{x}-1)\,P_j^{0,0}(2\widetilde{y}-1)\,(1-\widetilde{z})^{max(i,j)} P_k^{2max(i,j)+2,0}(2\widetilde{z}-1) \in C_r,$$

pour $0 \leqslant i,j \leqslant r$, $k \leqslant r - max(i,j)$, c'est à dire, en utilisant un argument de dimension,

$$Span\,\left\{\hat{\psi}_{i,j,k}(\hat{x},\hat{y},\hat{z}),\ 0 \leqslant i,j \leqslant r,\ k \leqslant r - max(i,j)\right\} \circ T^{-1} = C_r,$$

- **Tétraèdres :** Pour $0 \leqslant i+j+k \leqslant r$, on note

$$\hat{\psi}_{i,j,k}(\hat{x},\hat{y},\hat{z}) = P_i^{0,0}\left(\frac{2\hat{x}}{1-\hat{y}-\hat{z}} - 1\right) P_j^{2i+1,0}\left(\frac{2\hat{y}}{1-\hat{z}} - 1\right)(1-\hat{y}-\hat{z})^i P_k^{2(i+j)+2,0}(2\hat{z}-1)\,(1-\hat{z})^j$$

D'après 2.2.6, on a

$$\int_{\hat{K}} \hat{\psi}_{i,j,k}(\hat{x},\hat{y},\hat{z})\quad \hat{\psi}_{i',j',k'}(\hat{x},\hat{y},\hat{z})\,d\hat{x}\,d\hat{y}\,d\hat{z} = \underbrace{\int_0^1 P_i^{0,0}(2\widetilde{x}-1)P_{i'}^{0,0}(2\widetilde{x}-1)\,d\widetilde{x}}_{=\delta_{i\,i'}} \underbrace{\int_0^1 (1-\widetilde{y})^{i+i'+1} P_j^{2i+1,0}(2\widetilde{y}-1)P_{j'}^{2i'+1,0}(2\widetilde{y}-1)\,d\widetilde{y}}_{=\delta_{j\,j'}}$$
$$\underbrace{\int_0^1 (1-\widetilde{z})^{i+i'+j+j'+2} P_j^{2(i+j)+2,0}(2\widetilde{z}-1)P_{k'}^{2(i'+j')+2,0}(2\widetilde{z}-1)\,d\widetilde{z}}_{=\delta_{k\,k'}\ \text{lorsque}\ i=i'\ \text{et}\ j=j'},$$

i.e. la famille est orthogonale.

Il est clair que, pour $0 \leqslant i+j+k \leqslant r$,

$$\hat{\psi}_{i,j,k}(\hat{x},\hat{y},\hat{z}) \circ T^{-1} = P_i^{0,0}(2\widetilde{x}-1)\,(1-\widetilde{y})^i P_j^{2i+1,0}(2\widetilde{y}-1)\,(1-\widetilde{z})^{i+j}\,P_k^{2(i+j)+2,0}(2\widetilde{z}-1) \in C_r,$$

soit, en utilisant un argument de dimension,

$$Span\ \left\{\hat{\psi}_{i,j,k}(\hat{x},\hat{y},\hat{z}),\ 0 \leqslant i+j+k \leqslant r\right\} \circ T^{-1} = C_r,$$

ce qui termine la preuve de la proposition. □

La figure 2.8 présente la comparaison entre le conditionnement de la matrice de Vandermonde pour la base monomiales classiques de $\hat{P}_r$, et celui pour la base orthogonale, dans le cas d'éléments tétraédriques, pyramidaux et prismatiques. On remarque que le conditionnement de la matrice de Vandermonde augmente plus vite pour les tétraèdres que pour les pyramides lorsque l'on utilise la base monomiale, tandis que l'inverse se produit pour les bases orthogonales. En outre, l'utilisation de bases orthogonales améliore de beaucoup le conditionnement de la matrice de Vandermonde.

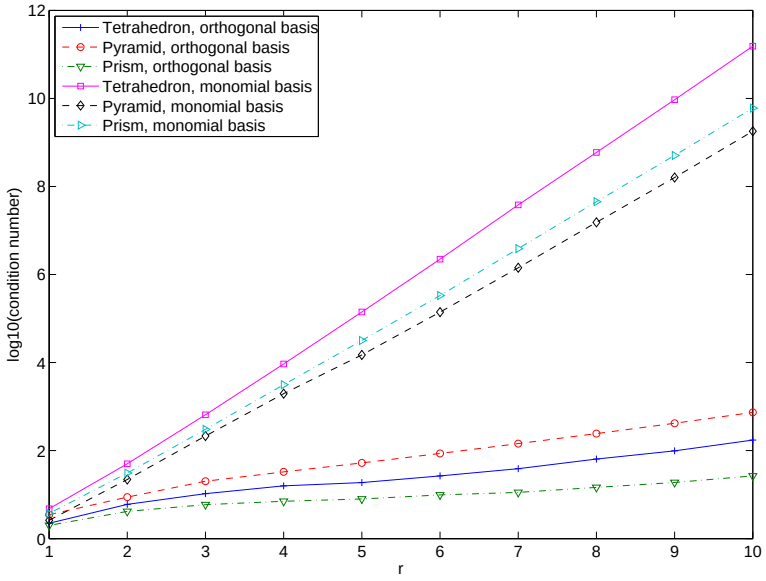

FIG. 2.8 – Conditionnement de la matrice de Vandermonde en fonction de l'ordre pour les éléments tétraédriques, pyramidaux et prismatiques, pour des bases monomiales et orthogonales

2.3.2 Éléments finis hiérarchiques

On donne ici des fonctions de base hiérarchiques de $\hat{P}_r$ conformes pour une formulation H^1, pour tous les types d'éléments. À partir des travaux de Šolín *et al.* [71]) et Warburton [72], on construit des fonctions de base
- invariantes par rotation pour les fonctions liées aux arêtes ;
- orthogonales pour les fonctions liées aux faces ;
- vérifiant une certaine orthogonalité pour les fonctions intérieures ;
- tensorisées dès que possible.

Le but en prenant des fonctions orthogonales et tensorisées est de creuser la matrice de masse, et potentiellement d'avoir un meilleur conditionnement.

Proposition 2.3.5 *Les fonctions suivantes forment une base hiérarchique de $\hat{P}_r$ préservant la continuité*
- ***Hexaèdre :*** *On considère les paramètres suivants*

$$\begin{cases} \lambda_1 = \hat{x} \\ \lambda_2 = \hat{y} \\ \lambda_3 = \hat{z} \\ \lambda_4 = 1 - \hat{x} \\ \lambda_5 = 1 - \hat{y} \\ \lambda_6 = 1 - \hat{z} \end{cases}$$

FONCTIONS H^1 HIÉRARCHIQUES POUR L'HEXAÈDRE

Pour un sommet s

$$\lambda_{s_1}\lambda_{s_2}\lambda_{s_3}, \quad 1 \leqslant s \leqslant 8$$

où s_1, s_2 et s_3 désignent les faces ne contenant pas s ($s_1 < s_2 < s_3$)

Pour une arête a

Si a est orientée selon e_x

$$\lambda_1\,\lambda_4\,P^{1,1}_{i-1}(\lambda_1 - \lambda_4)\,\lambda_{a_1}\,\lambda_{a_2}, \quad 1 \leqslant i \leqslant r-1, \quad 1 \leqslant a \leqslant 4$$

Si a est orientée selon e_y

$$\lambda_2\,\lambda_5\,P^{1,1}_{i-1}(\lambda_2 - \lambda_5)\,\lambda_{a_1}\,\lambda_{a_2}, \quad 1 \leqslant i \leqslant r-1, \quad 1 \leqslant a \leqslant 4$$

Si a est orientée selon e_z

$$\lambda_3\,\lambda_6\,P^{1,1}_{i-1}(\lambda_3 - \lambda_6)\,\lambda_{a_1}\,\lambda_{a_2}, \quad 1 \leqslant i \leqslant r-1, \quad 1 \leqslant a \leqslant 4$$

où a_1 et a_2 sont les faces ne contenant aucun sommet de a ($a_1 < a_2$)

Pour une face f

Si f est dans le plan (e_x, e_y)

$$\lambda_{f_1}\,\lambda_1\,\lambda_2\,\lambda_4\,\lambda_5\,P^{1,1}_{i-1}(\lambda_1 - \lambda_4)\,P^{1,1}_{j-1}(\lambda_2 - \lambda_5), \quad 1 \leqslant i,j \leqslant r-1, \quad 1 \leqslant f \leqslant 2$$

Si f est dans le plan (e_y, e_z)

$$\lambda_{f_1}\,\lambda_2\,\lambda_3\,\lambda_5\,\lambda_6\,P^{1,1}_{i-1}(\lambda_2 - \lambda_5)\,P^{1,1}_{j-1}(\lambda_3 - \lambda_6), \quad 1 \leqslant i,j \leqslant r-1, \quad 1 \leqslant f \leqslant 2$$

Si f est dans le plan (e_x, e_z)

$$\lambda_{f_1}\,\lambda_1\,\lambda_3\,\lambda_4\,\lambda_5\,P^{1,1}_{i-1}(\lambda_3 - \lambda_6)\,P^{1,1}_{j-1}(\lambda_1 - \lambda_4), \quad 1 \leqslant i,j \leqslant r-1, \quad 1 \leqslant f \leqslant 2$$

où f_1 désigne la face directement opposée à f.

Pour les fonctions intérieures

$$\lambda_1\,\lambda_2\,\lambda_3\,\lambda_4\,\lambda_5\,\lambda_6\,P^{1,1}_{i-1}(\lambda_1 - \lambda_4)\,P^{1,1}_{j-1}(\lambda_2 - \lambda_5)\,P^{1,1}_{k-1}(\lambda_3 - \lambda_6), \quad 1 \leqslant i,j,k \leqslant r-1$$

– ***Prisme :*** *On considère les paramètres suivants*

$$\begin{cases} \lambda_1 = \lambda_4 = 1 - \hat{x} - \hat{y} \\ \lambda_2 = \lambda_5 = \hat{x} \\ \lambda_3 = \lambda_6 = \hat{y} \end{cases} \qquad \begin{cases} \beta_1 = 1 - \hat{z} \\ \beta_2 = \hat{z} \end{cases} \qquad \textit{et} \qquad \begin{cases} \gamma_1 = \dfrac{\lambda_2 - \lambda_1}{\lambda_2 + \lambda_1} \\ \gamma_2 = \lambda_3 - \lambda_2 - \lambda_1 \end{cases}$$

FONCTIONS H^1 HIÉRARCHIQUES POUR LE PRISME

Pour un sommet s

$$\lambda_s \, \beta_{s'}, \quad 1 \leqslant s \leqslant 6$$

où s' *désigne la face triangulaire ne touchant pas* s*.*

Pour une arête horizontale a

$$\lambda_{a_1} \, \lambda_{a_2} \, \beta_{a'} \, P_{i-1}^{1,1}(\lambda_{a_2} - \lambda_{a_1}), \quad 1 \leqslant i \leqslant r-1, \quad 1 \leqslant a \leqslant 6$$

où a_1 *et* a_2 *désignent les sommets de* a *(*$a_1 < a_2$*), et* a' *désigne la face triangulaire ne touchant pas* a*.*

Pour une arête verticale a

$$\lambda_{a_1} \, \beta_1 \, \beta_2 \, P_{i-1}^{1,1}(\beta_2 - \beta_1), \quad 1 \leqslant i \leqslant r-1, \quad 1 \leqslant a \leqslant 3$$

où a_1 *désigne l'un des sommets de* a *(dans ce cas,* $\lambda_{a_1} = \lambda_{a_2}$*).*

Pour une face triangulaire f

$$\lambda_1 \, \lambda_2 \, \lambda_3 \, \beta_{f'} \, (\lambda_1 + \lambda_2)^{i-1} \, P_{i-1}^{1,1}(\gamma_1) \, P_{j-1}^{2i+1,1}(\gamma_2), \quad 1 \leqslant i + j \leqslant r-1, \quad 1 \leqslant f \leqslant 2$$

Pour une face quadrangulaire f

Si $\lambda_1 = 0$

$$\lambda_2 \, \lambda_3 \, \beta_1 \, \beta_2 \, P_{i-1}^{1,1}(\gamma_2) \, P_{j-1}^{1,1}(\beta_2 - \beta_1), \quad 1 \leqslant j, k \leqslant r-1$$

Si $\lambda_2 = 0$

$$\lambda_1 \, \lambda_3 \, \beta_1 \, \beta_2 \, P_{i-1}^{1,1}(\gamma_2) \, P_{j-1}^{1,1}(\beta_2 - \beta_1), \quad 1 \leqslant j, k \leqslant r-1$$

Si $\lambda_3 = 0$

$$\lambda_1 \, \lambda_2 \, \beta_1 \, \beta_2 \, (\lambda_1 + \lambda_2)^{i-1} \, P_{i-1}^{1,1}(\gamma_1) \, P_{j-1}^{1,1}(\beta_2 - \beta_1), \quad 1 \leqslant j, k \leqslant r-1$$

Pour les fonctions intérieures

$$\lambda_1 \, \lambda_2 \, \lambda_3 \, \beta_1 \, \beta_2 \, (\lambda_1 + \lambda_2)^{i-1} \, P_{i-1}^{1,1}(\gamma_1) P_{j-1}^{2i+1,1}(\gamma_2) P_{k-1}^{1,1}(\beta_2 - \beta_1), \quad 1 \leqslant i + j, k \leqslant r-1$$

– ***Pyramide :*** *On considère les paramètres suivants*

$$\begin{cases} \lambda_1 = \dfrac{1-\hat{x}-\hat{z}}{2} \\ \lambda_2 = \dfrac{1+\hat{y}-\hat{z}}{2} \\ \lambda_3 = \dfrac{1+\hat{x}-\hat{z}}{2} \\ \lambda_4 = \dfrac{1-\hat{y}-\hat{z}}{2} \\ \lambda_5 = \hat{z} \end{cases} \qquad \begin{cases} \gamma_1 = \gamma_3 = \dfrac{\hat{x}}{1-\hat{z}} \\ \gamma_2 = \gamma_4 = \dfrac{\hat{y}}{1-\hat{z}} \end{cases} \qquad et \qquad \begin{cases} \beta_1 = \dfrac{2\hat{z}+\hat{x}+\hat{y}}{2} \\ \beta_2 = \dfrac{2\hat{z}-\hat{x}+\hat{y}}{2} \\ \beta_3 = \dfrac{2\hat{z}-\hat{x}-\hat{y}}{2} \\ \beta_4 = \dfrac{2\hat{z}+\hat{x}-\hat{y}}{2} \end{cases}$$

FONCTIONS H^1 HIÉRARCHIQUES POUR LA PYRAMIDE

Pour un sommet s

$$\frac{\lambda_{s_1}\lambda_{s_2}}{1-\hat{z}}, \quad 1 \leqslant s \leqslant 4$$

où s_1 *et* s_2 *désignent les faces ne touchant pas* s $(s_1 < s_2)$

Pour l'apex

$$\lambda_5$$

Pour l'arête horizontale a

$$\frac{\lambda_{a_1}\lambda_{a_2}\lambda_{a_3}}{1-\hat{z}}(1-\hat{z})^{i-1} P_{i-1}^{1,1}(\gamma_a), \quad 1 \leqslant i \leqslant r-1, \quad 1 \leqslant a \leqslant 4$$

où a_1, a_2 *et* a_3 *sont les faces ne touchant pas* a $(a_1 < a_2 < a_3)$

Pour l'arête verticale a

$$\frac{\lambda_{a_1}\lambda_{a_2}\lambda_{a_3}}{1-\hat{z}} P_{i-1}^{1,1}(\beta_a), \quad 1 \leqslant i \leqslant r-1, \quad 1 \leqslant a \leqslant 4$$

où a_1, a_2 *et* a_3 *sont les faces ne touchant pas* a $(a_1 < a_2 < a_3)$

Pour une face triangulaire f

$$\frac{\lambda_{f_1}\lambda_{f_2}\lambda_{f_3}\lambda_{f_4}}{1-\hat{z}}(1-\hat{z})^{i-1} P_{i-1}^{1,1}(\gamma_f) P_{j-1}^{2i+1,1}(2\hat{z}-1), \quad 1 \leqslant i+j \leqslant r-1, \quad 1 \leqslant f \leqslant 4$$

où f_1, f_2, f_3 *et* f_4 *sont les autres faces* $(f_1 < f_2 < f_3 < f_4)$

Pour la face quadrangulaire

$$\frac{\lambda_1\lambda_2\lambda_3\lambda_4}{(1-\hat{z})^2}(1-\hat{z})^{\max(i,j)-1} P_{i-1}^{1,1}(\gamma_1) P_{j-1}^{1,1}(\gamma_2), \quad 1 \leqslant i,j \leqslant r-1$$

où i_1 *et* i_2 *désignent les faces quadrangulaires opposées*

Pour les fonctions intérieures

$$\frac{\lambda_1\lambda_2\lambda_3\lambda_4\lambda_5}{(1-\hat{z})^2}(1-\hat{z})^{max(i,j)-1} P_{i-1}^{1,1}(\gamma_1) P_{j-1}^{1,1}(\gamma_2) P_{k-1}^{2\max(i,j)+2,1}(2\hat{z}-1), \qquad \begin{matrix} 1 \leqslant i,j \leqslant r-1, \\ 1 \leqslant k \leqslant r-1-\max(i,j) \end{matrix}$$

- ***Tétraèdre :*** *On considère les paramètres suivants*

$$\begin{cases} \lambda_1 = 1 - \hat{x} - \hat{y} - \hat{z} \\ \lambda_2 = \hat{x} \\ \lambda_3 = \hat{y} \\ \lambda_4 = \hat{z} \end{cases} \quad et \quad \begin{cases} \gamma_1 = \dfrac{\lambda_2 - \lambda_1}{\lambda_2 + \lambda_1} \\ \gamma_2 = \dfrac{\lambda_3 - \lambda_2 - \lambda_1}{\lambda_3 + \lambda_2 + \lambda_1} \\ \gamma_3 = \lambda_4 - \lambda_3 - \lambda_2 - \lambda_1 \end{cases}$$

FONCTIONS H^1 HIÉRARCHIQUES POUR LE TÉTRAÈDRE

Pour un sommet s

$$\lambda_s, \quad 1 \leqslant s \leqslant 4$$

Pour une arête a

$$\lambda_{a_1}\,\lambda_{a_2}\,P_{i-1}^{1,1}(\lambda_{a_2} - \lambda_{a_1}), \quad 1 \leqslant i \leqslant r-1, \quad 1 \leqslant a \leqslant 6$$

où a_1 *et* a_2 *désignent les sommets de* a $(a_1 < a_2)$

Pour une face triangulaire f

Si $\lambda_1 = 0$

$$\lambda_2\,\lambda_3\,\lambda_4\,(\lambda_1+\lambda_2+\lambda_3)^{i-1}\,P_{i-1}^{1,1}(\gamma_2)\,P_{j-1}^{2i+1,1}(\gamma_3), \quad 1 \leqslant i+j \leqslant r-1$$

Si $\lambda_2 = 0$

$$\lambda_1\,\lambda_3\,\lambda_4\,(\lambda_1+\lambda_2+\lambda_3)^{i-1}\,P_{i-1}^{1,1}(\gamma_2)\,P_{j-1}^{2i+1,1}(\gamma_3), \quad 1 \leqslant i+j \leqslant r-1$$

Si $\lambda_3 = 0$

$$\lambda_1\,\lambda_2\,\lambda_4\,(\lambda_1+\lambda_2)^{i-1}\,P_{i-1}^{1,1}(\gamma_1)\,P_{j-1}^{2i+1,1}(\gamma_3), \quad 1 \leqslant i+j \leqslant r-1$$

Si $\lambda_4 = 0$

$$\lambda_1\,\lambda_2\,\lambda_3\,(\lambda_1+\lambda_2)^{i-1}\,P_{i-1}^{1,1}(\gamma_1)\,(\lambda_1+\lambda_2+\lambda_3)^{j-1}\,P_{j-1}^{2i+1,1}(\gamma_2), \quad 1 \leqslant i+j \leqslant r-1$$

Pour les fonctions intérieures

$$\lambda_1\,\lambda_2\,\lambda_3\,\lambda_4\,(\lambda_1+\lambda_2)^{i-1}\,P_{i-1}^{1,1}(\gamma_1)\,(\lambda_1+\lambda_2+\lambda_3)^{j-1}\,P_{j-1}^{2i+1,1}(\gamma_2)\,P_{k-1}^{2(i+j)+1,1}(\gamma_3), \quad 1 \leqslant i+j+k \leqslant r-1$$

Preuve. Concernant les hexaèdres, les prismes et les tétraèdres, voir Warburton [72]. Pour les pyramides, le principe est le même, aussi la démonstration ne sera-t-elle pas détaillée.

Par construction, l'ensemble des fonctions forme une base de $\hat{P}_r$ pour chaque élément et les restrictions aux arêtes, aux faces triangulaires et quadrangulaires sont égales (à une rotation près pour les faces triangulaires, à prendre en compte lors de la programmation) pour tous les éléments.

Remarque 2.3.6 *Soit* B *la fonction « bulle » qui annule les arêtes et les faces, on remarque que toute fonction intérieure s'écrit sous la forme*

$$\hat{\varphi}_{i,j,k}(\hat{x},\hat{y},\hat{z}) = B(\hat{x},\hat{y},\hat{z})\,P_i(\hat{x})\,P_j^i(\hat{y})\,P_k^{i,j}(\hat{z})$$

Or, grâce à la structure des P_i, P_j^j *et* $P_k^{i,j}$ *choisis, on vérifie en écrivant les intégrales sur* $\widetilde{Q}$ *que*

$$\forall \hat{\varphi}_{i,j,k}(\hat{x},\hat{y},\hat{z}),\ \forall \hat{p}(\hat{x},\hat{y},\hat{z}) \in \hat{P}_{ijk}(\hat{K}), \quad \int_{\hat{K}} \hat{\varphi}_{i,j,k}(\hat{x},\hat{y},\hat{z})\,\hat{p}(\hat{x},\hat{y},\hat{z})\,d\hat{x}\,d\hat{y}\,d\hat{z} = 0$$

avec

- ***Hexaèdres*** *:* $\hat{P}_{ijk}(\hat{K}) = P_{\max(i,j,k)-2}(\hat{K})$
- ***Prisme*** *:* $\hat{P}_{ijk}(\hat{K}) = P_{\max(i+j,k)-3}(\hat{K})$
- ***Pyramide*** *:* $\hat{P}_{ijk}(\hat{K}) = P_{max(i,j)+k-3}(\hat{K})$
- ***Tétraèdre*** *:* $\hat{P}_{ijk}(\hat{K}) = P_{i+j+k-4}(\hat{K})$

À partir de l'ordre 4, outre B *qui n'est orthogonale à aucune fonction, les fonctions intérieures sont donc orthogonales à toutes les fonctions de bas degré, ce qui permet de creuser la matrice de masse pour les ordres élevés.*

2.4 Conformité

On rappelle le théorème concernant les conditions de conformité H^1.

Theorème 2.4.1

$$\begin{cases} \forall K \in \Omega,\ P_r^F(K) \subset H^1(K) \\ V_h \subset \mathcal{C}^0(\bar{\Omega}) \end{cases} \Longrightarrow V_h \subset H^1(\Omega)$$

Preuve. Voir par exemple Monk [55].

Theorème 2.4.2 *Avec les espaces P_r^F construits dans la section 2.2 et le choix de degré de liberté de la section 2.3, on a*

$$V_h(\Omega) \subset H^1(\Omega).$$

Preuve.

– Vérifions tout d'abord le premier point du théorème 2.4.1. Soit $p \in P_r^F$. Lorsque p est polynomiale, ce qui est le cas pour les tétraèdres, les prismes et les hexaèdres, on a de manière évidente $p \in \mathcal{C}^0(K)$ et $\nabla p \in \mathcal{C}^0(K)^3$. Puisque K est borné, on a alors immédiatement $p \in L^2(K)$ et $\nabla p \in L^2(K)^3$.

Concernant les pyramides, $p \in C^\infty(\bar{K} \backslash S_5)$ comme toute fraction rationnelle dont le pôle n'appartient pas au domaine. On prouve la continuité en S_5 en considérant quatre pseudo-faces F_ε^i, $0 \leqslant i \leqslant 4$, $0 \leqslant \varepsilon \leqslant 1$ parcourant un quart, noté Q_i, de la pyramide.
On considère une seule face, les trois autres étant similaire par symétrie. Sur le quart Q_2 représenté en bleu sur la figure 2.9, la pseudo-face F_ε^2, représentée en rouge, est telle que

$$\begin{cases} \hat{x} = (1-z)(1-\varepsilon) \\ -(1-\hat{z})(1-\varepsilon) \leqslant y \leqslant (1-\hat{z})(1-\varepsilon) \\ 0 \leqslant \hat{z} \leqslant 1, \end{cases}$$

et on a

$$\forall M = (\hat{x}, \hat{y}, \hat{z}) \in Q_2,\ \exists \varepsilon \in [0,1],\ M \in F_\varepsilon^2.$$

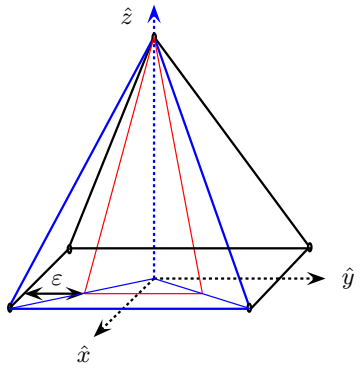

FIG. 2.9 – Pseudo face F_ε^2

On traite ici le cas le plus difficile $p = \dfrac{\hat{x}\hat{y}}{1-\hat{z}}$. Pour un point M de Q_2 attaché à une face F_ε^2, on a

$$p(M) = \frac{(1-\hat{z})(1-\varepsilon)\hat{y}}{1-\hat{z}} = (1-\varepsilon)\hat{y} \xrightarrow{z \to 1} 0$$

puisque lorsque $z \to 1$, on a $y \to 0$. Finalement, $p \in C^0(\bar{K})$, et, K étant borné, $p \in L^2(K)$.
On considère à présent ∇p dans le cas où $p = \dfrac{xy}{1-z}$, et un point M de Q_2 attaché à une face F_ε^2. Il vient

$$-(1-\varepsilon)^2 \leqslant \frac{\partial p}{\partial z}(M) = -\frac{(1-\hat{z})(1-\varepsilon)\hat{y}}{(1-\hat{z})^2} \leqslant (1-\varepsilon)^2.$$

c'est à dire que $\dfrac{\partial p}{\partial z}$ est bornée. La même méthode est appliquée pour $\dfrac{\partial p}{\partial x}$ et $\dfrac{\partial p}{\partial y}$, ce qui permet de conclure que ∇p est borné dans K. On a finalement $\nabla p \in L^2(K)$ puisque K est borné.

- Concernant le deuxième point du théorème 2.4.1, il s'agit de vérifier que les restrictions des espaces P_r^F sur chaque type d'élément sur chaque type de face sont identiques, ce qui revient à étudier les restrictions des espaces $\hat{P}_r$ et à vérifier que la transformation F assure la conformité.
  - Commençons par vérifier que les restrictions des espaces $\hat{P}_r$ à chaque type de face sont les mêmes pour tous les éléments. Pour un paramétrage (η,ξ) de la face considérée, il est immédiat que la restriction de $\mathbb{P}_r(\hat{x},\hat{y},\hat{z})$ à toute face triangulaire du tétraèdre est dans $\mathbb{P}_r(\eta,\xi)$, et que la restriction de $\mathbb{Q}_r(\hat{x},\hat{y},\hat{z})$ à toute face quadrangulaire de l'hexaèdre est dans $\mathbb{Q}_r(\eta,\xi)$. Or retrouve aisément ces résultats sur le prismes, mais nous allons le détailler pour la pyramide.
    Sur la pyramide $\hat{K}$, toute fonction $p \in \hat{P}_r$ peut s'écrire de la manière suivante

$$p(\hat{x},\hat{y},\hat{z}) = p_r(\hat{x},\hat{y},\hat{z}) + \sum_{0\leqslant k\leqslant r-1} p_k(\hat{x},\hat{y})\left(\frac{\hat{x}\hat{y}}{1-\hat{z}}\right)^{r-k},$$

    avec $p_r \in \mathbb{P}_r(\hat{x},\hat{y},\hat{z})$ and $p_k \in \mathbb{P}_k(\hat{x},\hat{y})$.
    Sur une face triangulaire, par exemple sur la face $\hat{x} = (1-\hat{z})$, on a $p_r(1-\hat{z},\hat{y},\hat{z})$ qui appartient de manière évidente à $\mathbb{P}_r(\hat{y},\hat{z})$, et la partie rationnelle devient

$$p_k(\hat{x},\hat{y})\left(\frac{\hat{x}\hat{y}}{1-\hat{z}}\right)^{r-k} = p_k(1-\hat{z},\hat{y})\, y^{r-k},\ \ 0 \leqslant k \leqslant r-1,$$

    soit $p_k(1-\hat{z},\hat{y})y^{r-k} \in \mathbb{P}_r(\hat{y},\hat{z})$. Finalement $p \in \mathbb{P}_r(\hat{y},\hat{z})$. La même simplification peut être effectuée sur les autres faces.
    Sur la base quadrangulaire, on a $p_r(\hat{x},\hat{y},0)$ qui appartient clairement à $\mathbb{Q}_r(\hat{x},\hat{y})$, et la partie rationnelle devient

$$p_k(\hat{x},\hat{y})\left(\frac{\hat{x}\hat{y}}{1-\hat{z}}\right)^{r-k} = p_k(\hat{x},\hat{y})\, x^{r-k}y^{r-k},\ 0 \leqslant k \leqslant r-1,$$

    soit $p_k(\hat{x},\hat{y})x^{r-k}y^{r-k} \in \mathbb{Q}_r(\hat{x},\hat{y})$. Finalement $p \in \mathbb{Q}_r(\hat{x},\hat{y})$.
    En utilisant un argument de dimension, on en déduit que

$$\begin{array}{rcl} \hat{P}_{r|\ \hat{x}=1-\hat{z} \text{ ou } \hat{x}=\hat{z}-1} & = & \mathbb{P}_r(\hat{y},\hat{z}) \\ \hat{P}_{r|\ \hat{y}=1-\hat{z} \text{ ou } \hat{y}=\hat{z}-1} & = & \mathbb{P}_r(\hat{x},\hat{z}) \\ \hat{P}_{r|\ \hat{z}=0} & = & \mathbb{Q}_r(\hat{x},\hat{y}), \end{array} \tag{2.4.1}$$

    ce qui correspond bien aux mêmes restrictions d'espaces sur les faces des autres éléments.

  - Ne considérant que le cas isoparamétrique, on a par construction $F \in \left(\hat{P}_r\right)^3$ (avec $r = 1$ dans le cas d'éléments droits). On applique le résultat précédent : les restrictions de F à chaque type de face sont dans les mêmes espaces pour les quatre type d'éléments.

Finalement, on a les mêmes espaces et les mêmes degrés de liberté sur les faces, localement on a donc unisolvance sur chaque face, ce qui signifie que l'on a égalité des fonctions de chaque élément sur cette face, ce qui achève la démonstration □

Remarque 2.4.3 *Dans le cas de la pyramide, c'est la fraction rationnelle de F qui permet d'obtenir une restriction conforme sur la face quadrangulaire sans perdre la conformité sur les faces triangulaire. Cela n'aurait pas été possible si F avait été polynomiale comme le montre Bedrosian [5].*

Une transformation polynomiale par morceaux assurant la conformité des pyramides avec les autres types d'éléments a été proposée par Knabner et Summ [49]. Pour cela, ils découpent la pyramide de référence en deux tétraèdres et considèrent une transformation pour chaque tétraèdre permettant d'assurer la conformité (voir chapitre 5).

Concernant les hexaèdres, les prismes et les tétraèdres, on a immédiatement plus précisément $P_r^F \subset H^m(K)$, pour tout $m \geqslant 0$. En revanche, à cause de la fraction rationnelle, il est difficile d'obtenir un résultat plus précis que $P_r^F \subset H^1(K)$ sur la pyramide. En effet, comme le remarquent Nigam et Phillips [59], on a le théorème suivant

Theorème 2.4.4 *Sur la pyramide de référence $\hat{K}$, on a*

$$\forall \varepsilon > 0,\ \left\{ \frac{\hat{x}^i\hat{y}^j}{(1-\hat{z})^{i+j-k}},\ 0 \leqslant i,j \leqslant k \right\} \subset H^{k+3/2-\varepsilon}(\hat{K}).$$

Preuve. Soit $p_k = \dfrac{\hat{x}^i \hat{y}^j}{(1-\hat{z})^{i+j-k}}$ avec $0 \leqslant i, j \leqslant k$

$$\begin{aligned}
\frac{\mathrm{d}^{m_1} p_k}{\mathrm{d}\hat{x}^{m_1}} &= C_1 \frac{\hat{x}^{i-m_1} \hat{y}^j}{(1-\hat{z})^{i+j-k}} \\
\frac{\mathrm{d}^{m_2} p_k}{\mathrm{d}\hat{y}^{m_2}} &= C_2 \frac{\hat{x}^i \hat{y}^{j-m_2}}{(1-\hat{z})^{i+j-k}} \\
\frac{\mathrm{d}^{m_3} p_k}{\mathrm{d}\hat{z}^{m_3}} &= C_3 \frac{\hat{x}^i \hat{y}^j}{(1-\hat{z})^{i+j+m_3-k}}
\end{aligned} \tag{2.4.2}$$

où C_1, C_2 et C_3 sont les constantes qui apparaissent lors de la dérivation, dépendant de m_1, m_2, m_3 et k.

En intégrant sur le cube unité $\widetilde{Q}$ via la transformation T, on a

$$\begin{aligned}
\int_{\hat{K}} \left(\frac{\mathrm{d}^{m_1} p_k}{\mathrm{d}\hat{x}^{m_1}} \right)^2 &= \int_0^1 \widetilde{x}^{2(i-m_1)}\, d\widetilde{x} \int_0^1 \widetilde{y}^{2j}\, d\widetilde{y} \int_0^1 (1-\widetilde{z})^{2(k-m_1+1)}\, d\widetilde{z} \\
\int_{\hat{K}} \left(\frac{\mathrm{d}^{m_2} p_k}{\mathrm{d}\hat{y}^{m_1}} \right)^2 &= \int_0^1 \widetilde{x}^{2i}\, d\widetilde{x} \int_0^1 \widetilde{y}^{2(j-m_2)}\, d\widetilde{y} \int_0^1 (1-\widetilde{z})^{2(k-m_2+1)}\, d\widetilde{z} \\
\int_{\hat{K}} \left(\frac{\mathrm{d}^{m_3} p_k}{\mathrm{d}\hat{z}^{m_3}} \right)^2 &= \int_0^1 \widetilde{x}^{2i}\, d\widetilde{x} \int_0^1 \widetilde{y}^{2j}\, d\widetilde{y} \int_0^1 (1-\widetilde{z})^{2(k-m_3+1)}\, d\widetilde{z}
\end{aligned}$$

Or pour $0 \leqslant i, j \leqslant k$, ces trois intégrales sont finies si et seulement si $2(k-m_i+1) > -1$, c'est à dire $m_i < k + \dfrac{3}{2}$, pour $i = 1, 2, 3$, d'où le résultat. □

Une conséquence directe de ce théorème est la suivante

Theorème 2.4.5

$$\forall \varepsilon > 0,\ \mathbb{B}_r \subset H^{5/2-\varepsilon}(\hat{K}).$$

Preuve. Comme pour tout $r \geqslant 1$, on a $\dfrac{\hat{x}\,\hat{y}}{1-\hat{z}} \in \mathbb{B}_r$, le résultat est immédiat en utilisant le théorème 2.4.4. □

Lorsque la pyramide K est affine, on a ainsi immédiatement $P_r^F \subset H^{5/2-\varepsilon}(K)$. Lorsque F n'est pas affine, le résultat est plus délicat à obtenir car nous n'avons que peu d'informations sur F^{-1}.

Chapitre 3

Formule de quadrature et estimations d'erreur

Nous cherchons à présent des formules de quadrature permettant l'évaluation exacte des intégrales intervenant dans le calcul des matrices du problème discret considéré, et ce pour chaque type d'élément. Nous donnons également la formule de quadrature minimale permettant d'obtenir une erreur globale d'ordre r. Nous effectuons un calcul d'estimation de l'erreur commise par rapport à la solution exacte avec les éléments définis précédemment afin de justifier notre choix d'espace.

Sommaire

3.1 Intégration par formule de quadrature

3.1.1 Introduction

Pour évaluer les intégrales intervenant dans la construction des matrices du problème, une manière simple de faire consiste à construire une formule de quadrature utilisant des points de type Gauss (-Legendre, -Jacobi ou -Lobatto) sur le cube unité $\widetilde{Q}$ de coordonnées $(\widetilde{x},\widetilde{y},\widetilde{z})$, et de considérer leur image sur l'élément de référence $\hat{K}$ de coordonnées $(\hat{x},\hat{y},\hat{z})$ en utilisant le changement de variables T donné par la définition (2.2.7).

On rappelle (cf Stroud [67]) que la méthode de quadrature de Gauss est une méthode de quadrature 1D exacte à n points pris sur l'intervalle (a,b), telle que

$$\int_a^b f(x)\varpi(x)\,dx \approx \sum_{1\leqslant i\leqslant n} w_i f(x_i), \tag{3.1.1}$$

où $\varpi(\cdot)$ est une fonction de pondération sur (a,b), les w_i sont les coefficients ou poids de quadrature et les x_i sont les points de quadrature, réels, distincts, uniques et sont les racines de polynômes orthogonaux, choisis conformément au domaine d'intégration et à la fonction de pondération. On rajoute une extrémité de l'intervalle d'intégration pour les formules de type Gauss-Radau, les deux extrémités de l'intervalle pour les formules de type Gauss-Lobatto. Les poids et les points de quadrature sont choisis de façon à obtenir des degrés d'exactitude les plus grands possibles. Les différentes méthodes de quadrature de Gauss à $n = r+1$ points sont

- **Formule de Gauss-Legendre** : L'intervalle considéré est $]-1,1[$ avec $\varpi(x)=1$. Les $r+1$ points sont dans $]-1,1[$ et la formule de quadrature obtenue est exacte pour les polynômes de $\mathbb{Q}_{2r+1}$
- **Formule de Gauss-Radau** : L'intervalle considéré est $]-1,1[$ avec $\varpi(x)=1$. On prend r points dans $]-1,1[$ et on ajoute un point à l'une des extrémités -1 ou 1. La formule de quadrature obtenue est exacte pour les polynômes de $\mathbb{Q}_{2r}$
- **Formule de Gauss-Lobatto** : L'intervalle considéré est $]-1,1[$ avec $\varpi(x)=1$. On prend $r-1$ points dans $]-1,1[$ et on ajoute les deux extrémités à l'ensemble des points de quadrature. La formule de quadrature obtenue est exacte pour les polynômes de $\mathbb{Q}_{2r-1}$
- **Formule de Gauss-Jacobi** : On considère $\varpi(x)=(1-x)^\alpha\,(1+x)^\beta$, α et β dans $\mathbb{Z}$. La formule de quadrature obtenue est exacte pour les polynômes de
  - $(1-x)^\alpha(1+x)^\beta\mathbb{Q}_{2r+1}$ pour une formule de Gauss-Legendre-Jacobi ;
  - $(1-x)^\alpha(1+x)^\beta\mathbb{Q}_{2r}$ pour une formule de Gauss-Radau-Jacobi ;
  - $(1-x)^\alpha(1+x)^\beta\mathbb{Q}_{2r-1}$ pour une formule de Gauss-Lobatto-Jacobi.

  (voir Gautschi [34] pour ces deux dernières formules)

Les formules de quadratures 2D et 3D sont alors obtenues par tensorisation de formules 1D.

On notera $\oint^G$ une intégrale approchée par une formule de quadrature de Gauss.

3.1.2 Intégration exacte

On rappelle l'expression des matrices de masse M_h et de rigidité R_h sur un élément K du maillage

$$\begin{aligned} M_{h\,i,j} &= \int_K \varphi_i\varphi_j\,dx\,dy\,dz = \int_{\hat{K}} |DF|\,\hat{\varphi}_i\hat{\varphi}_j\,d\hat{x}\,d\hat{y}\,d\hat{z} \\ R_{h\,i,j} &= \int_K \nabla\varphi_i\cdot\nabla\varphi_j\,dx\,dy\,dz = \int_{\hat{K}} |DF|DF^{-1}DF^{*-1}\hat{\nabla}\hat{\varphi}_i\cdot\hat{\nabla}\hat{\varphi}_j\,d\hat{x}\,d\hat{y}\,d\hat{z}, \end{aligned} \tag{3.1.2}$$

où DF et $|DF|$ désignent respectivement la jacobienne et le jacobien de F.

En utilisant la définition 2.2.7 de la transformation T permettant de passer de l'élément de référence $\hat{K}$ au cube unité $\widetilde{Q}$, on a

$$\begin{aligned} M_{h\,i,j} &= \int_{\widetilde{Q}} \widetilde{|DF|}\,\widetilde{|T|}\,\widetilde{\varphi}_i\widetilde{\varphi}_j\,d\widetilde{x}\,d\widetilde{y}\,d\widetilde{z} \\ R_{h\,i,j} &= \int_{\widetilde{Q}} \widetilde{|DF|}\,\widetilde{|T|}\,\widetilde{DF^{-1}}\,\widetilde{DF^{*-1}}\widetilde{\nabla}\widetilde{\varphi}_i\cdot\widetilde{\nabla}\widetilde{\varphi}_j\,d\widetilde{x}\,d\widetilde{y}\,d\widetilde{z}, \end{aligned} \tag{3.1.3}$$

où $|T|$ désigne le déterminant du changement de variables T.

On cherche à intégrer exactement les matrices dès que cela est possible. Pour cela, on cherche l'espace de polynômes auquel appartient le contenu des intégrales et pour lequel il faudra avoir une intégration exacte. Comme l'on va généralement utiliser des formules de quadratures issues de la tensorisation de formules 1D, nous cherchons les inclusions dans des espaces de type $\mathbb{Q}_{m,n,p}$.

Lemme 3.1.1 *La matrice jacobienne, le jacobien et la comatrice de la matrice jacobienne de la transformation F pour chaque type d'élément sont dans les espaces suivants :*

- ***Hexaèdre :***

$$\widetilde{DF} \in \left(\mathbb{Q}^3_{0,1,1} \times \mathbb{Q}^3_{1,0,1} \times \mathbb{Q}^3_{1,1,0}\right)(\widetilde{x},\widetilde{y},\widetilde{z}),$$
$$\widetilde{|DF|} \in \mathbb{Q}_{2,2,2}(\widetilde{x},\widetilde{y},\widetilde{z}),$$
$$\widetilde{|DF|}\widetilde{DF}^{*-1} \in \left(\mathbb{Q}^3_{2,1,1} \times \mathbb{Q}^3_{1,2,1} \times \mathbb{Q}^3_{1,1,2}\right)(\widetilde{x},\widetilde{y},\widetilde{z}).$$

- ***Prisme :***

$$\widetilde{DF} \in \left(\mathbb{Q}^3_{0,0,1} \times \mathbb{Q}^3_{0,0,1} \times \mathbb{Q}^3_{1,1,0}\right)(\widetilde{x},\widetilde{y},\widetilde{z})$$
$$\widetilde{|DF|} \in \mathbb{Q}_{1,1,2}(\widetilde{x},\widetilde{y},\widetilde{z}),$$
$$\widetilde{|DF|}\widetilde{DF}^{*-1} \in \left(\mathbb{Q}^3_{1,1,1} \times \mathbb{Q}^3_{1,1,1} \times \mathbb{Q}^3_{0,0,2}\right)(\widetilde{x},\widetilde{y},\widetilde{z}).$$

- ***Pyramide :***

$$\widetilde{DF} \in \left(\mathbb{Q}^3_{0,1,0} \times \mathbb{Q}^3_{1,0,0} \times \mathbb{Q}^3_{1,1,0}\right)(\widetilde{x},\widetilde{y},\widetilde{z}),$$
$$\widetilde{|DF|} \in \mathbb{Q}_{1,1,0}(\widetilde{x},\widetilde{y},\widetilde{z}),$$
$$\widetilde{|DF|}\widetilde{DF}^{*-1} \in \left(\mathbb{Q}^3_{1,1,0}\right)^3(\widetilde{x},\widetilde{y},\widetilde{z}).$$

- ***Tétraèdre :***

$$\widetilde{DF} \in \left(\mathbb{Q}^3_{0,0,0} \times \mathbb{Q}^3_{0,0,0} \times \mathbb{Q}^3_{0,0,0}\right)(\widetilde{x},\widetilde{y},\widetilde{z}),$$
$$\widetilde{|DF|} \in \mathbb{Q}_{0,0,0}(\widetilde{x},\widetilde{y},\widetilde{z}),$$
$$\widetilde{|DF|}\widetilde{DF^{*-1}} \in \left(\mathbb{Q}^3_{0,0,0}\right)^3(\widetilde{x},\widetilde{y},\widetilde{z})$$

Preuve. En utilisant la proposition 2.1.4 et la transformation T, on a :

- **Hexaèdres :**

$$\begin{aligned}
\frac{\partial F}{\partial \hat{x}} &= (-S_1+S_2)+(S_1-S_2-S_3+S_4)\hat{y}+(S_1-S_2-S_5+S_6)\hat{z}+(-S_1+S_2-S_3+S_4+S_5-S_6+S_7-S_8)\hat{y}\hat{z}\\
&= (-S_1+S_2)+(S_1-S_2-S_3+S_4)\widetilde{y}+(S_1-S_2-S_5+S_6)\widetilde{z}+(-S_1+S_2-S_3+S_4+S_5-S_6+S_7-S_8)\widetilde{y}\widetilde{z}\\
&= A_1+C_1\widetilde{y}+C_2\widetilde{z}+D\widetilde{y}\widetilde{z},\\
\frac{\partial F}{\partial \hat{y}} &= (-S_1+S_4)+(S_1-S_2-S_3+S_4)\hat{x}+(S_1-S_4-S_5+S_8)\hat{z}+(-S_1+S_2-S_3+S_4+S_5-S_6+S_7-S_8)\,\hat{x}\hat{z}\\
&= (-S_1+S_4)+(S_1-S_2-S_3+S_4)\widetilde{x}+(S_1-S_4-S_5+S_8)\widetilde{z}+(-S_1+S_2-S_3+S_4+S_5-S_6+S_7-S_8)\widetilde{x}\widetilde{z}\\
&= A_2+C_1\widetilde{x}+C_3\widetilde{z}+D\widetilde{x}\widetilde{z},\\
\frac{\partial F}{\partial \hat{z}} &= (-S_1+S_5)+(S_1-S_2-S_5+S_6)\hat{x}+(S_1-S_4-S_5+S_8)\hat{y}+(-S_1+S_2-S_3+S_4+S_5-S_6+S_7-S_8)\hat{x}\hat{y}\\
&= (-S_1+S_5)+(S_1-S_2-S_5+S_6)\widetilde{x}+(S_1-S_4-S_5+S_8)\widetilde{y}+(-S_1+S_2-S_3+S_4+S_5-S_6+S_7-S_8)\widetilde{x}\widetilde{y}\\
&= A_3+C_2\widetilde{x}+C_3\widetilde{y}+D\widetilde{x}\widetilde{y},
\end{aligned}$$

et

$$\widetilde{|DF|}(\widetilde{x},\widetilde{y},\widetilde{z}) = det(A_1+C_1\widetilde{y}+C_2\widetilde{z}+D\widetilde{y}\widetilde{z}, A_2+C_1\widetilde{x}+C_3\widetilde{z}+D\widetilde{x}\widetilde{z}, A_3+C_2\widetilde{x}+C_3\widetilde{y}+D\widetilde{x}\widetilde{y}),$$

soit

$$\begin{aligned}
\widetilde{|DF|}(\widetilde{x},\widetilde{y},\widetilde{z}) = \;& det(A_1,A_2,A_3)+[det(A_1,C_3,A_3)+det(C_2,A_2,A_3)]\,\widetilde{z}\\
&+[det(A_1,A_2,C_2)+det(A_1,C_1,A_3)]\,\widetilde{x}+[det(A_1,A_2,C_3)+det(C_1,A_2,A_3)]\,\widetilde{y}\\
&+[det(A_1,A_2,D)+det(A_1,C_1,C_3)+det(C_1,A_2,C_2)]\,\widetilde{x}\widetilde{y}+det(C_2,C_3,A_3)\,\widetilde{z}^2\\
&+[det(A_1,D,A_3)+det(C_2,C_1,A_3)+det(A_1,C_3,C_2)]\,\widetilde{x}\widetilde{z}+det(C_1,A_2,C_3)\,\widetilde{y}^2\\
&+[det(C_1,C_3,A_3)+det(C_2,A_2,C_3)+det(D,A_2,A_3)]\,\widetilde{y}\widetilde{z}+det(A_1,C_1,C_2)\,\widetilde{x}^2\\
&+2det(C_1,C_3,C_2)\,\widetilde{x}\widetilde{y}\widetilde{z}\\
&+det(A_1,C_1,D)\,\widetilde{x}^2\widetilde{y}+det(A_1,D,C_2)\,\widetilde{x}^2\widetilde{z}+det(C_2,C_1,D)\,\widetilde{x}^2\widetilde{y}\widetilde{z}\\
&+det(C_1,A_2,D)\,\widetilde{x}\widetilde{y}^2+det(D,A_2,C_3)\,\widetilde{y}^2\widetilde{z}+det(C_1,C_3,D)\,\widetilde{x}\widetilde{y}^2\widetilde{z}\\
&+det(C_2,D,A_3)\,\widetilde{x}\widetilde{z}^2+det(D,C_3,A_3)\,\widetilde{y}\widetilde{z}^2++det(C_3,C_2,D)\,\widetilde{x}\widetilde{y}\widetilde{z}^2
\end{aligned}$$

Pour $\widetilde{|DF|}\widetilde{DF}^{*-1}$, on a

$$\widetilde{|DF|}\widetilde{DF}^{*-1} = \left[\frac{\partial F}{\partial \hat{y}} \wedge \frac{\partial F}{\partial \hat{z}}, \frac{\partial F}{\partial \hat{z}} \wedge \frac{\partial F}{\partial \hat{x}}, \frac{\partial F}{\partial \hat{x}} \wedge \frac{\partial F}{\partial \hat{y}}\right]$$

et le résultat est obtenu en sommant les degrés de polynômes.

– **Prismes :**

$$\begin{aligned}
\frac{\partial F}{\partial \hat{x}} &= (-S_1+S_2)+(S_1-S_2-S_4+S_5)\hat{z}\\
&= (-S_1+S_2)+(S_1-S_2-S_4+S_5)\widetilde{z}\\
&= A_1+C_1\widetilde{z},\\
\frac{\partial F}{\partial \hat{y}} &= (-S_1+S_3)+(S_1-S_3-S_4+S_6)\hat{z}\\
&= (-S_1+S_3)+(S_1-S_3-S_4+S_6)\widetilde{z}\\
&= A_2+C_2\widetilde{z},\\
\frac{\partial F}{\partial \hat{z}} &= (-S_1+S_4)+(S_1-S_2-S_4+S_5)\hat{x}+(S_1-S_3-S_4+S_6)\hat{y}\\
&= (-S_1+S_4)+(S_1-S_2-S_4+S_5)\widetilde{x}(1-\widetilde{y})+(S_1-S_3-S_4+S_6)\widetilde{y}\\
&= A_3+C_1\widetilde{x}\,(1-\widetilde{y})+C_2\widetilde{y},
\end{aligned}$$

et

$$\widetilde{|DF|}(\widetilde{x},\widetilde{y},\widetilde{z}) = det(A_1+C_1\widetilde{z},A_2+C_2\widetilde{z},A_3+C_1\widetilde{x}(1-\widetilde{y})+C_2\widetilde{y}),$$

soit

$$\begin{aligned}
\widetilde{|DF|}(\widetilde{x},\widetilde{y},\widetilde{z}) = \;& det(A_1,A_2,A_3)+\widetilde{x}\,(1-\widetilde{y})\,det(A_1,A_2,C_1)+\widetilde{y}\,det(A_1,A_2,C_2)\\
&+\widetilde{z}\,[det(C_1,A_2,A_3)+det(A_1,C_2,A_3)]+\widetilde{x}\,(1-\widetilde{y})\,\widetilde{z}\,det(A_1,C_2,C_1)+\widetilde{y}\widetilde{z}\,det(C_1,A_2,C_2)\\
&+\widetilde{z}^2\,det(C_1,C_2,A_3).
\end{aligned}$$

Comme $\widetilde{|DF|}\widetilde{DF}^{*-1}$ est la comatrice de $\widetilde{DF}$, on a

$$\begin{aligned}
\widetilde{|DF|}\widetilde{DF}^{*-1}(\widetilde{x},\widetilde{y},\widetilde{z}) = \;& [A_2+C_2\widetilde{z})\wedge(A_3+C_1\widetilde{x}\,(1-\widetilde{y})+C_2\widetilde{y}),\\
&(A_3+C_1\widetilde{x}\,(1-\widetilde{y})+C_2\widetilde{y})\wedge(A_1+C_1\widetilde{z}),(A_1+C_1\widetilde{z})\wedge(A_2+C_2\widetilde{z})],
\end{aligned}$$

soit

$$\begin{aligned}
\widetilde{|DF|}\widetilde{DF}^{*-1} = \;& (A_2\wedge A_3, A_3\wedge A_1, A_1\wedge A_2)\\
&+(A_2\wedge C_1, C_1\wedge A_1, 0)\;\widetilde{x}(1-\widetilde{y})+(A_2\wedge C_2, C_2\wedge A_1, 0)\;\widetilde{y}\\
&+(C_2\wedge A_3, A_3\wedge C_1, (A_1\wedge C_2+C_1\wedge A_2))\;\widetilde{z}\\
&+(C_2\wedge C_1, 0, 0)\;\widetilde{x}(1-\widetilde{y})\widetilde{z}+(0, C_2\wedge C_1, 0)\;\widetilde{y}\widetilde{z}\\
&+(0,0,C_1\wedge C_2)\;\widetilde{z}^2.
\end{aligned}$$

– **Pyramide :**

$$\begin{aligned}
\frac{\partial F}{\partial \hat{x}} &= \frac{1}{4}(-S_1+S_2+S_3-S_4)+\frac{1}{4}(S_1-S_2+S_3-S_4)\frac{\hat{y}}{1-\hat{z}}\\
&= \frac{1}{4}(-S_1+S_2+S_3-S_4)+\frac{1}{4}(S_1-S_2+S_3-S_4)(2\widetilde{y}-1)\\
&= A_1+C\widetilde{y},\\
\frac{\partial F}{\partial \hat{y}} &= \frac{1}{4}(-S_1-S_2+S_3+S_4)+\frac{1}{4}(S_1-S_2+S_3-S_4)\frac{\hat{x}}{1-\hat{z}}\\
&= \frac{1}{4}(-S_1-S_2+S_3+S_4)+\frac{1}{4}(S_1-S_2+S_3-S_4)(2\widetilde{x}-1)\\
&= A_2+C\widetilde{x},\\
\frac{\partial F}{\partial \hat{z}} &= \frac{1}{4}(4S_5-S_1-S_2-S_3-S_4)+\frac{1}{4}(S_1-S_2+S_3-S_4)\frac{\hat{x}\hat{y}}{(1-\hat{z})^2}\\
&= \frac{1}{4}(4S_5-S_1-S_2-S_3-S_4)+\frac{1}{4}(S_1-S_2+S_3-S_4)(2\widetilde{x}-1)(2\widetilde{y}-1)\\
&= A_3+2C\,\widetilde{x}\,\widetilde{y},
\end{aligned}$$

(3.1.4)

et

$$\widetilde{|DF|}(\widetilde{x},\widetilde{y},\widetilde{z}) = det(A_1+C\widetilde{y},A_2+C\widetilde{x},A_3+2C\,\widetilde{x}\,\widetilde{y}),$$

soit

$$\widetilde{|DF|}(\widetilde{x},\widetilde{y},\widetilde{z}) = det(A_1,A_2,A_3)+\widetilde{x}\,det(A_1,C,A_3)+\widetilde{y}\,det(C,A_2,A_3)+2\,\widetilde{x}\,\widetilde{y}\,det(A_1,A_2,C),$$

Comme $\widetilde{|DF|}\widetilde{DF}^{*-1}$ est la comatrice de $\widetilde{DF}$, on a

$$\widetilde{|DF|}\widetilde{DF}^{*-1}(\widetilde{x},\widetilde{y},\widetilde{z}) = [(A_2+C\widetilde{x})\wedge(A_3+2C\widetilde{x}\,\widetilde{y}),(A_3+2C\widetilde{x}\,\widetilde{y})\wedge(A_1+C\widetilde{y}),(A_1+C\widetilde{y})\wedge(A_2+C\widetilde{x})]\,,$$

soit

$$\begin{aligned}\widetilde{|DF|}\widetilde{DF}^{*-1} &= (A_2\wedge A_3, -A_1\wedge A_3, A_1\wedge A_2)\\ &+(C\wedge A_3, 0, A_1\wedge C)\,\widetilde{x}+(0,-C\wedge A_3, C\wedge A_2)\,\widetilde{y}\\ &-2\,(-A_2\wedge C, A_1\wedge C, 0)\,\widetilde{x}\widetilde{y}\end{aligned}$$

– **Tétraèdre :** Les résultats sont immédiats d'après la définition de F, ce qui achève la preuve du lemme. □

Lemme 3.1.2 *Pour tous les types d'éléments,*

$$\forall i\in[\![1,n_r]\!],\ \ \widetilde{\varphi}_i\in\mathbb{Q}_r(\widetilde{x},\widetilde{y},\widetilde{z}).$$

Preuve. Conséquence de la remarque 2.2.9. □

Lemme 3.1.3 *Pour chaque type d'élément, on a :*

– ***Hexaèdre*** *:*

$$\forall i\in[\![1,n_r]\!],\ \ \hat{\nabla}\widetilde{\varphi}_i\in\mathbb{Q}_{r-1,r,r}\times\mathbb{Q}_{r,r-1,r}\times\mathbb{Q}_{r,r,r-1}\,(\widetilde{x},\widetilde{y},\widetilde{z})$$

– ***Prisme*** *:*

$$\forall i\in[\![1,n_r]\!],\ \ \hat{\nabla}\widetilde{\varphi}_i\in\mathbb{Q}_{r-1,r-1,r}\times\mathbb{Q}_{r,r-1,r}\times\mathbb{Q}_{r,r,r-1}\,(\widetilde{x},\widetilde{y},\widetilde{z})$$

– ***Pyramide*** *:*

$$\forall i\in[\![1,n_r]\!],\ \ \hat{\nabla}\widetilde{\varphi}_i\in\mathbb{Q}_{r-1,r,r-1}\times\mathbb{Q}_{r,r-1,r-1}\times\mathbb{Q}_{r,r,r-1}\,(\widetilde{x},\widetilde{y},\widetilde{z})$$

– ***Tétraèdre*** *:*

$$\forall i\in[\![1,n_r]\!],\ \ \hat{\nabla}\widetilde{\varphi}_i\in\mathbb{Q}_{r-1,r-1,r-1}\times\mathbb{Q}_{r,r-1,r-1}\times\mathbb{Q}_{r,r,r-1}\,(\widetilde{x},\widetilde{y},\widetilde{z})$$

Preuve. Le résultat est relativement aisé à démontrer pour l'hexaèdre, le prisme et le tétraèdre. Nous faisons ici la démonstration pour la pyramide qui est le cas le plus particulier à traiter. On décompose $\hat{\varphi}_i(\hat{x},\hat{y},\hat{z})$ dans la base des monômes $\hat{\psi}_j(\hat{x},\hat{y},\hat{z})$ de $\hat{P}_r$ et on traite successivement les différents cas.

On considère d'abord la dérivée en x, le cas de la dérivée en y étant traité de manière similaire par symétrie : si $\hat{\psi}_j(\hat{x},\hat{y},\hat{z})\in\mathbb{P}_r(\hat{x},\hat{y},\hat{z})$, on a

$$\frac{\partial\hat{\psi}_j}{\partial\hat{x}}(\hat{x},\hat{y},\hat{z})=\hat{x}^{m-1}\hat{y}^n\hat{z}^p=(2\widetilde{x}-1)^{m-1}(2\widetilde{y}-1)^n(1-\widetilde{z})^{m+n-1}\widetilde{z}^p,\ \ m+n+p\leqslant r.$$

Sinon,

$$\frac{\partial\hat{\psi}_j}{\partial\hat{x}}(\hat{x},\hat{y},\hat{z})=\frac{\hat{x}^{r-p+i-1}\hat{y}^{r-p+j}}{(1-\hat{z})^{r-p}}=(2\widetilde{x}-1)^{r-p+i-1}(2\widetilde{y}-1)^{r-p+j}(1-\widetilde{z})^{r-p+i+j-1},\ \ i+j\leqslant p\leqslant r-1,$$

c'est à dire $\dfrac{\partial\widetilde{\psi}_j}{\partial\hat{x}}(\widetilde{x},\widetilde{y},\widetilde{z})\in\mathbb{Q}_{r-1,r,r-1}(\widetilde{x},\widetilde{y},\widetilde{z})$ dans les deux cas.

De la même manière, pour la dérivée en z, soit

$$\frac{\partial\hat{\psi}_j}{\partial\hat{z}}(\hat{x},\hat{y},\hat{z})=\hat{x}^m\hat{y}^n\hat{z}^{p-1}=(2\widetilde{x}-1)^m(2\widetilde{y}-1)^n\widetilde{z}^{p-1}(1-\widetilde{z})^{m+n},\ \ m+n+p\leqslant r$$

soit

$$\frac{\partial\hat{\psi}_j}{\partial\hat{z}}(\hat{x},\hat{y},\hat{z})=\frac{\hat{x}^{r-p+i}\hat{y}^{r-p+j}}{(1-\hat{z})^{r-p+1}}=(2\widetilde{x}-1)^{r-p+i}(2\widetilde{y}-1)^{r-p+j}(1-\widetilde{z})^{r-p+i+j-1},\ \ i+j\leqslant p\leqslant r-1,$$

c'est à dire $\dfrac{\partial\widetilde{\psi}_j}{\partial\hat{z}}(\widetilde{x},\widetilde{y},\widetilde{z})\in\mathbb{Q}_{r,r,r-1}(\widetilde{x},\widetilde{y},\widetilde{z})$ dans les deux cas. □

Considérant le lemme 3.1.1, on remarque que $\widetilde{|DF|}$ est dans un espace de type $\mathbb{Q}_{s_1,s_1,s_2}$ pour tous les éléments, avec

F affine :	$s_1=0,\ s_2=0$
Hexaèdre :	$s_1=2,\ s_2=2$
Prisme :	$s_1=1,\ s_2=2$
Pyramide :	$s_1=1,\ s_2=0$

On rappelle de plus que, d'après la définition 2.2.7, on a

Hexaèdre :	$\widetilde{\lvert T\rvert} = 1$
Prisme :	$\widetilde{\lvert T\rvert} = (1-\widetilde{y})$
Pyramide :	$\widetilde{\lvert T\rvert} = 4(1-\widetilde{z})^2$
Tétraèdre :	$\widetilde{\lvert T\rvert} = (1-\widetilde{y})(1-\widetilde{z})^2$

On peut à présent écrire les résultats suivants sur la matrice de masse et la matrice de rigidité.

Proposition 3.1.4 *:*

- *Pour avoir l'intégration exacte de la matrice de masse, la formule de quadrature utilisée doit être exacte pour les polynômes de* $\widetilde{\lvert T\rvert}\,\mathbb{Q}_{2r+s_1,2r+s_1,2r+s_2}(\widetilde{x},\widetilde{y},\widetilde{z})$
- *Pour avoir l'intégration exacte de la matrice de rigidité,* ***lorsque F est affine****, la formule de quadrature utilisée doit être exacte pour les polynômes de*
  - $\widetilde{\lvert T\rvert}\,\mathbb{Q}_{2r}(\widetilde{x},\widetilde{y},\widetilde{z})$ *pour les hexaèdres et les prismes*
  - $\widetilde{\lvert T\rvert}\,\mathbb{Q}_{2r,2r,2r-2}(\widetilde{x},\widetilde{y},\widetilde{z})$ *pour les pyramides et les tétraèdres*
- *À cause de la fraction rationnelle qui apparait dans le terme* $\widetilde{DF}^{-1}$*, la matrice de rigidité ne peut être intégrée exactement dans le cas d'une transformation non affine*

Preuve. Pour la matrice de masse, le résultat est obtenu en utilisant les lemmes 3.1.1 et 3.1.2, via l'écriture de la matrice de masse sur le cube unité $\widetilde{Q}$ donnée par l'équation 3.1.3, et en additionnant les puissances.

Concernant la matrice de rigidité dans le cas affine, en utilisant les lemmes 3.1.1 et 3.1.3, via l'écriture de la matrice de rigidité sur le cube unité $\widetilde{Q}$ donnée par l'équation 3.1.3, le résultat est obtenu en additionnant les puissances. □

3.1.3 Formule de quadrature

Pour intégrer exactement la matrice de masse, on suit l'idée de Hammer, Marlowe et Stroud [41] pour les cônes en incluant $\widetilde{\lvert T\rvert}$ à la formule de quadrature. On prend ainsi les formules de quadrature suivantes pour chacun des types d'éléments :

- **Hexaèdre :**
$$(\xi_r^{GL},\xi_r^{GL},\xi_r^{GL}),(\omega_r^{GL}\omega_r^{GL}\omega_r^{GL}),$$
- **Prisme :**
$$(\xi_r^{G},\xi_r^{GJ1},\xi_r^{GL}),(\omega_r^{G}\omega_r^{GJ1}\omega_r^{GL}),$$
ou
$$(\xi_r^{tri},\xi_r^{GL}),(\omega_r^{tri},\omega_r^{GL})$$
- **Pyramide :**
$$(\xi_r^{G},\xi_r^{G},\xi_r^{GJ2}),(\omega_r^{G}\omega_r^{G}\omega_r^{GJ2}),$$
- **Tétraèdre :**
$$(\xi_r^{G},\xi_r^{GJ1},\xi_r^{GJ2}),(\omega_r^{G}\omega_r^{GJ1}\omega_r^{GJ2}),$$
ou
$$(\xi_r^{tetra}),(\omega_r^{tetra})$$

où

- (ξ_r^{G},ω_r^{G}) est la formule de quadrature de Gauss d'ordre r, exacte pour les polynômes de $\mathbb{Q}_{2r+1}$,
- $(\xi_r^{GJa},\omega_r^{GJa})$ est la formule de quadrature de Gauss-Jacobi d'ordre r, exacte pour les polynômes de $(1-x)^a\mathbb{Q}_{2r+1}$
- $(\xi_r^{tri},\omega_r^{tri})$ est une formule de quadrature d'ordre r pour le triangle, exacte pour les polynômes de $\mathbb{P}_{2r}$, comme par exemple celle décrite par Dunavant [27].
- $(\xi_r^{tetra},\omega_r^{tetra})$ est une formule de quadrature d'ordre r pour le tétraèdre d'ordre, exacte pour les polynômes de $\mathbb{P}_{2r}$, comme par exemple celle décrite par Šolín *et al.* [71].

Finalement, on utilise au maximum $(r+1)^3$ points d'intégration.

Remarque 3.1.5 *Les formules de quadrature adaptées au triangle et au tétraèdre respectent les symétries des éléments et nécessitent un nombre moins élevé de points de quadrature que des formules de quadrature de type Gauss tensorisées.*

3.2 Estimation d'erreur abstraite

3.2.1 Présentation du problème

On considère le problème variationnel standard suivant

$$\left\{ \begin{array}{l} \text{Trouver } u \in V \text{ tel que} \\ \forall v \in V,\ a(u,v) = f(v), \end{array} \right. \tag{3.2.1}$$

où $V = H^1(\Omega)$, et où $a(.,.)$ désigne une forme bilinéaire continue et coercive, et $f(.)$ un forme linéaire continue. Pour un sous-espace de dimension finie V_h de l'espace V, le problème discret s'écrit alors

$$\left\{ \begin{array}{l} \text{Trouver } u_h \in V_h \text{ tel que} \\ \forall v_h \in V_h,\ a_h(u_h,v_h) = f_h(v_h), \end{array} \right. \tag{3.2.2}$$

où $a_h(.,.)$ désigne une forme bilinéaire définie sur l'espace V_h, uniformément V_h-elliptique, et $f_h(.)$ une forme linéaire définie sur l'espace V_h.

On considèrera le cas simple suivant

$$a(u,v) = \int_\Omega uv + \int_\Omega \nabla u \cdot \nabla v.$$

3.2.2 Lemme de Strang

On considère la version suivante du lemme de Strang

Lemme 3.2.1 *(Lemme de Strang). Si u est solution de (3.2.1) et u_h est solution de (3.2.2), il existe une constante $C > 0$ ne dépendant pas du pas d'espace h telle que*

$$\|u - u_h\|_1 \leqslant C \quad \underbrace{\inf_{v_h \in V_h} \left\{ \|u - v_h\|_1 \right.}_{erreur\ d'interpolation} \quad + \underbrace{\sup_{w_h \in V_h} \frac{|a(v_h,w_h) - a_h(v_h,w_h)|}{\|w_h\|_1} \right\}}_{erreur\ d'intégration\ numérique}.$$

Preuve. En remarquant que $V_h \subset V$, la preuve de cette version du lemme de Strang est similaire à la preuve proposée par Ciarlet [14]. □

On étudie à présent séparément les deux termes du membre de droite de l'inégalité du lemme de Strang, c'est à dire l'erreur d'interpolation et l'erreur de quadrature.

On supposera que u est dans $H^{r+1}(\Omega) \cap \mathcal{C}^0(\Omega)$ pour un domaine Ω suffisamment régulier.

3.2.3 Erreur d'interpolation

Soit Ω un ouvert lipschitzien de $\mathbb{R}^3$ composé de n_e éléments K

$$\Omega = \bigcup_K K.$$

On note

$$h_K = \text{diam } (K) = \sup_{(x,y)\in K} |x - y|, \quad h = \max_K h_K$$
$$\rho_K = \sup_B \{\text{diam } (B),\ B \text{ boule incluse dans } K\}$$

et on suppose que le maillage est tel qu'il existe $\sigma > 0$ tel que

$$\frac{h_K}{\rho_K} \leqslant \sigma \tag{3.2.3}$$

Définition 3.2.2 *Pour E et F deux espaces vectoriels normés, on note $\mathcal{L}(E,F)$ l'ensemble des applications linéaires continues de E dans F.*

Pour estimer l'erreur d'interpolation, on utilise la version suivante du lemme de Bramble-Hilbert

Lemme 3.2.3 *(Lemme de Bramble-Hilbert). Soit $0 \leqslant m < s+1$. Pour $\check{K}$ lipschitzien, on note $\check{\Pi} \in \mathcal{L}(H^{s+1}(\check{K}), H^m(\check{K}))$ un projecteur vérifiant*

$$\forall \check{p} \in \mathbb{P}_s(\check{K}),\ \check{\Pi}\check{p} = \check{p}.$$

Pour tout ouvert K affinement équivalent à $\check{K}$ par une transformation F, on définit un projecteur Π tel que, pour toutes fonctions $\check{u} \in H^{s+1}(\check{K})$ et $u \in H^m(K)$ telles que $\check{u} = u \circ F$,

$$(\Pi \check{u}) = \check{\Pi}\check{u}$$

Il existe alors une constante $C(\check{\Pi}, \check{K}) > 0$ telle que

$$\forall u \in H^{s+1}(K),\ \|u - \Pi u\|_{m,K} \leqslant C(\check{\Pi}, \check{K})\, {h_K}^{s+1-m}\, |u|_{s+1,K}.$$

Preuve. On utilise le théorème 3.1.4 de Ciarlet [14] avec $p = q = 2$. Pour $0 \leqslant k \leqslant m$, on a

$$\forall u \in H^{s+1}(K),\ |u - \Pi u|_{k,K} \leqslant C_k(\check{\Pi}, \check{K})\, \frac{{h_K}^{s+1}}{{\rho_K}^k}\, |u|_{s+1,K}.$$

En utilisant la propriété du maillage 3.2.3, on obtient

$$\forall u \in H^{s+1}(K),\ |u - \Pi u|_{k,K} \leqslant \sigma\, C_k(\check{\Pi}, \check{K})\, {h_K}^{s+1-m}\, |u|_{s+1,K}.$$

Pour h_K suffisamment petit, on a

$$\sigma \sum_{0 \leqslant k \leqslant m} C_k(\check{\Pi}, \check{K})\, {h_K}^{s+1-m} \leqslant C(\check{\Pi}, \check{K})\, {h_K}^{s+1-m}$$

d'où le résultat avec la norme H^m. □

Définition 3.2.4 *Soit $I_h \in \mathcal{L}(H^{r+1}(\Omega), H^1(\Omega))$ un opérateur tel que*

$$\forall u_h \in V_h,\ I_h u_h = u_h$$

et

$$\begin{aligned} &\forall u \in L^2(\Omega), \quad \|I_h u\|_{0,\Omega} \leqslant C_{0,\Omega} \|u\|_{0,\Omega} \\ &\forall u \in H^1(\Omega), \quad \|I_h u\|_{1,\Omega} \leqslant C_{1,\Omega} \|u\|_{1,\Omega} \end{aligned}$$

On notera I_h^K la restriction de I_h à un élément K.

Remarque 3.2.5 *L'interpolant de Clément convient (voir par exemple Monk [55])*

Proposition 3.2.6 *Pour $u \in H^{r+1}(\Omega)$, il existe une constante $C_\Omega > 0$ ne dépendant que de r telle que*

$$\|u - I_h u\|_{1,\Omega} \leqslant C_\Omega\, h^r\, \|u\|_{r+1,\Omega}.$$

où h désigne la longueur caractéristique maximale sur tous les éléments K du maillage.

Preuve. On a

$$\|u - I_h u\|_{1,\Omega}^2 = \sum_K \|u - I_h^K u\|_{1,K}^2$$

Puisque $\mathbb{P}_r \subset P_r^F$, on a

$$\forall p \in \mathbb{P}_r(K),\ I_h^K p = p$$

c'est à dire que l'on peut appliquer le lemme de Bramble-Hilbert 3.2.3 avec $s = r$, $m = 1$ et en choisissant $\check{K} = \dfrac{K - S_1}{h_K}$ qui est bien affinement équivalent à K. On prend $\Pi = I_h^K$, $\check{\Pi}$ se déduisant de Π par la transformation affine. On obtient ainsi

$$\|u - I_h^K u\|_{1,K} \leqslant C(\check{I}_h, \check{K})\, {h_K}^r\, |u|_{r+1,K}.$$

En utilisant l'inégalité de Cauchy-Schwarz discrète et en notant $C_\Omega = \max_K C(\check{I}_h, \check{K})$ et $h = \max_K h_K$, on a donc

$$\|u - I_h u\|_{1,\Omega} \leqslant C_\Omega\, h^r\, |u|_{r+1,\Omega}.$$

Grâce à l'inégalité des normes 1.1.1, on obtient finalement le résultat annoncé. □

Remarque 3.2.7 *La constante $C(\check{I}_h, \check{K})$ ne dépend donc plus de h_K et ρ_K, mais dépend néanmoins de la forme de K, si bien que C_Ω dépend toujours de la géométrie du maillage.*

La condition $\max_K C_K$ borné est vérifiée dans le cas d'un maillage périodique dans les études numériques qui suivront puisque le nombre de formes de chaque type d'élément utilisé dans le maillage est fini. Dans un cas plus général, on conjecture que l'aspect borné de C_K est lié à l'existence d'une borne supérieure pour l'inverse de la matrice jacobienne DF, comme dans le cas des hexaèdres (Girault et Raviart [35]).

3.2.4 Erreur de quadrature

Puisque, d'après le lemme 3.2.6, même si les matrices de masse et de rigidité étaient intégrées exactement, l'erreur globale serait en $O(h^r)$, on cherche donc la formule de quadrature minimale qui nous permet d'avoir une erreur de quadrature en $O(h^r)$ également.

Définition 3.2.8 *Pour $(v_h, w_h) \in P_r^F$ et $K \in \Omega$, on note l'erreur d'intégration sur un élément K*

$$E_K(v_h, w_h) = \int_K v_h\, w_h\, dx\, dy\, dz - \oint_K v_h w_h\, dx\, dy\, dz,$$

où $\oint_K v_h w_h\, dx\, dy\, dz$ est l'intégrale approchée exacte pour les polynômes de $\widetilde{|T|}\,\mathbb{Q}_{m,m,n}$.

Pour $(v_h, w_h) \in V_h$, l'erreur d'intégration sur Ω est

$$E(v_h, w_h) = \sum_K E_K(v_h, w_h).$$

Pour $s \geqslant 1$, on définit également $\pi_s \in \mathcal{L}(H^{s+1}(K), H^1(K))$ le projecteur orthogonal sur $\mathbb{P}_s(K)$.

3.2.4.1 Matrice de masse

On cherche tout d'abord l'erreur de quadrature commise pour le calcul de la matrice de masse.

Lemme 3.2.9 *Pour une formule de quadrature exacte pour les polynômes de $\widetilde{|T|}\,\mathbb{Q}_{m,m,n}$, on a*

$$\forall (v_h, w_h) \in P_r^F,\ E_K(v_h, w_h) = E_K(v_h - \pi_p v_h, w_h - \pi_q w_h)$$

pour $m \geqslant r + \max(p,q) + s_1$ et $n \geqslant r + \max(p,q) + s_2$, avec $p + q \leqslant r$.

Preuve. On a

$$\forall (v_h, w_h) \in P_r^F,\ E_K(v_h - \pi_p v_h, w_h - \pi_q w_h) = E_K(v_h, w_h) - E_K(v_h, \pi_q w_h) - E_K(\pi_p v_h, w_h) + E_K(\pi_p v_h, \pi_q w_h)$$

On considère tout d'abord l'intégrale

$$\int_K \pi_p v_h\, w_h\, dx\, dy\, dz.$$

Après changement de variable, on obtient

$$\int_{\widetilde{Q}} \widetilde{|T|}\widetilde{|DF|}\,(\widetilde{\pi_p v_h})\ \widetilde{w}_h\, d\widetilde{x}\, d\widetilde{y}\, d\widetilde{z}.$$

D'après les lemmes 3.1.2 et 3.1.1, on a $\widetilde{|DF|}(\widetilde{\pi_p v_h})\ \widetilde{w}_h \in \mathbb{Q}_{p+r+s_1, p+r+s_1, p+r+s_2}(\widetilde{x}, \widetilde{y}, \widetilde{z})$, si bien que pour une formule de quadrature exacte pour les polynômes de $\widetilde{|T|}\,\mathbb{Q}_{m,m,n}$, on a donc

$$E_K(\pi_p v_h, w_h) = 0, \tag{3.2.4}$$

dès que $m \geqslant r + p + s_1$ et $n \geqslant r + p + s_2$.

De la même manière, pour $m \geqslant r + q + s_1$ et $n \geqslant r + q + s_2$, on a

$$E_K(v_h, \pi_q w_h) = 0$$

Lorsque $m \geqslant r + \max(p,q) + s_1$ et $n \geqslant r + \max(p,q) + s_2$, avec $p + q \leqslant r$, on a donc

$$E_K(\pi_p v_h, \pi_q w_h) = 0,$$

ce qui prouve le résultat avancé. □

Proposition 3.2.10 *Pour une formule de quadrature exacte pour les polynômes de $\widetilde{|T|}\,\mathbb{Q}_{m,m,n}$*

$$\forall w_h \in V_h,\ |E(I_h u, w_h)| \leqslant C'_\Omega\, h^r\, \|u\|_{r+1,\Omega} \|w_h\|_{1,\Omega}$$

avec $m \geqslant 2r - 2 + s_1$ et $n \geqslant 2r - 2 + s_2$.

Preuve. Soit $w_h \in V_h$ et $v_h = I_h u \in V_h$. On a

$$|E(v_h, w_h)| \leqslant \sum_K |E_K(v_h, w_h)|$$

D'après le lemme 3.2.9 avec $p = r-2$ et $q = 0$,

$$E_K(v_h, w_h) = E_K(v_h - \pi_{r-2} v_h, w_h - \pi_0 w_h).$$

pour une formule de quadrature exacte au moins pour des polynômes de $|\widetilde{T}|\mathbb{Q}_{2r-2+s_1,2r-2+s_1,2r-2+s_2}$.

La norme définie par l'intégrale approchée est équivalente à la norme H^1 avec une constante C_N lorsque les poids de quadrature sont positifs, ce qui est le cas pour les quadratures utilisées ici. On a donc

$$|E_K(u,u)| \leqslant (1 + C_N^2)\, ||u||_{1,K}^2$$

Soit $C_n = \sqrt{1 + C_N^2}$, en utilisant l'inégalité de Cauchy-Schwarz, on obtient ainsi

$$|E_K(v_h, w_h)| \leqslant C_n\, ||v_h - \pi_{r-2} v_h||_{0,K} ||w_h - \pi_0 w_h||_{0,K}.$$

Comme $w_h \in H^1(K)$, le lemme de Bramble-Hilbert pour $s = 0$, $m = 0$ et $\Pi = \pi_0$ donne

$$||w_h - \pi_0 w_h||_{0,K} \leqslant C_K^w\, h_K\, ||w_h||_{1,K}$$

En revanche, puisque $v_h = I_h^K u$ n'est a priori pas plus régulier que H^1, on ne peut appliquer le lemme de Bramble-Hilbert avec $s = r$ directement. On a cependant

$$||v_h - \pi_{r-2} v_h||_{0,K} = ||I_h^K u - \pi_{r-2} I_h^K u||_{0,K} \leqslant ||I_h^K u - I_h^K \pi_{r-2} u||_{0,K} + ||I_h^K \pi_{r-2} u - \pi_{r-2} I_h^K u||_{0,K}$$

Les interpolants I_h^K et π_{r-2} sont bornés dans $L^2(K)$, on note donc

$$\begin{aligned} ||I_h^K u||_{0,K} &\leqslant C_{0,h,K} ||u||_{0,K} \\ ||\pi_{r-2} u||_{0,K} &\leqslant C_{r-2,K} ||u||_{0,K} \end{aligned}$$

Puisque $\mathbb{P}_{r-2} \subset P_r^F$, on a

$$\forall u \in H^{r+1}(K),\ I_h^K \pi_{r-2} u = \pi_{r-2} u$$

d'où

$$||v_h - \pi_{r-2} v_h||_{0,K} \leqslant C_{0,h,K} ||u - \pi_{r-2} u||_{0,K} + C_{r-2,K} ||u - I_h^K u||_{0,K}$$

Pour $u \in H^{r+1}$, les deux normes sont traitées grâce au lemme de Bramble-Hilbert

- avec $s = r$, $m = 0$ et $\Pi = I_h^K$
$$||u - I_h^K u||_{0,K} \leqslant C_{0,K}\, {h_K}^{r+1}\, ||u||_{r+1,K}$$
- avec $s = r-2$ et $m = 0$ et $\Pi = \pi_{r-2}$,
$$||u - \pi_{r-2} u||_{0,K} \leqslant C_K^u h^{r-1} ||u||_{r-1,K} \leqslant C_K^u\, h_K^{r-1}\, ||u||_{r+1,K}$$

Pour h_K suffisamment petit, il existe $C_K > 0$ telle que

$$C_n C_K^w\, (C_{0,h,K} C_K^u\, {h_K}^r + C_{r-2,K} C_{0,K}\, {h_K}^{r+2}) \leqslant C_K {h_K}^r$$

On a donc

$$\forall w_h \in V_h,\ |E_K(v_h, w_h)| \leqslant C_K\, h_K^r\, ||u||_{r+1,K} ||w_h||_{1,K}$$

En sommant sur les éléments et en prenant $C'_\Omega = \max_{K \in \Omega} C_K$ et $h = \max_{K \in \Omega} h_K$, on obtient finalement

$$\forall w_h \in V_h,\ |E_K(v_h, w_h)| \leqslant C'_\Omega\, h^r\, ||u||_{r+1,\Omega} ||w_h||_{1,\Omega}$$

ce qui achève la démonstration. □

3.2.4.2 Matrice de rigidité

On cherche ensuite une estimation de l'erreur commise pour le terme de rigidité de a.

Proposition 3.2.11 *Pour une formule de quadrature exacte pour les polynômes de* $\widetilde{|T|}\,\mathbb{Q}_{m,m,n}$

$$\forall w_h \in V_h,\ |E(\nabla I_h u, \nabla w_h)| \leqslant C"_\Omega\, h^r\, ||u||_{r+1,\Omega} ||w_h||_{1,\Omega}$$

avec $m \geqslant 2r-1+t_1$ *et* $n \geqslant 2r-2+t_2$.

Preuve. Soit $w_h \in V_h$ et $v_h = I_h u \in V_h$, on a

$$|E(\nabla v_h, \nabla w_h)| \leqslant \sum_K |E_K(\nabla v_h, \nabla w_h)|$$

Or, pour tout $w_h \in V_h$, l'inégalité triangulaire donne

$$E_K(\nabla I_h u, \nabla w_h) = E_K(\nabla I_h u - \nabla \pi_r u, \nabla w_h) + E_K(\nabla \pi_r u, \nabla w_h)$$

soit

$$E_K(\nabla I_h u, \nabla w_h) = E_K(\nabla (I_h u - \pi_r u), \nabla w_h) + E_K(\nabla \pi_r u, \nabla w_h)$$

On a

$$|E_K(\nabla (I_h u - \pi_r u), \nabla w_h)| \leqslant C_n ||\nabla (I_h u - \pi_r u)||_{0,K} ||\nabla w_h||_{0,K}$$

c'est à dire, en utilisant l'inégalité sur les normes,

$$|E_K(\nabla (I_h u - \pi_r I_h u), \nabla w_h)| \leqslant C_n ||I_h u - \pi_r u||_{1,K} ||w_h||_{1,K}$$

Or

$$\pi_r u = I_h \pi_r u$$

et l'interpolant I_h est borné dans $H^1(K)$, c'est à dire qu'il existe $C_{1,K} > 0$ telle que

$$\forall u \in H^1(K),\ ||I_h u||_{1,K} \leqslant C_{1,K}\, ||u||_{1,K}.$$

On a ainsi

$$|E_K(\nabla (I_h u - \pi_r u), \nabla w_h)| \leqslant C_n\, C_{1,K}\, ||u - \pi_r u||_{1,K} ||w_h||_{1,K}$$

Pour $u \in H^{r+1}$, on peut utiliser le lemme de Bramble-Hilbert avec $s = r$, $m = 1$ et $\Pi = \pi_r$ et obtenir de la sorte

$$||u - \pi_r u||_{1,K} \leqslant C_K\, {h_K}^r\, ||u||_{r+1,K}$$

On a ainsi, en notant $\mathcal{C}_K = C_n C_{1,K} C_K$

$$|E_K(\nabla (I_h u - \pi_r u), \nabla w_h)| \leqslant \mathcal{C}_K {h_K}^r ||u||_{r+1,K} ||w_h||_{1,K}$$

On traite à présent la partie $E_K(\nabla \pi_r u, \nabla w_h)$. Soit une formule de quadrature exacte pour les polynômes de $\widetilde{|T|}\,\mathbb{Q}_{m,m,n}$. En passant que l'élément de référence, on a

$$E_K(\nabla \pi_r u, \nabla w_h) = E_{\hat{K}}(\widetilde{\nabla \pi_r u}, |DF| DF^{*-1} \widetilde{\nabla w_h})$$

Or

$$\nabla \pi_r u \in \mathbb{P}^3_{r-1}$$

donc

$$\widetilde{\nabla \pi_r u} \in \mathbb{P}^3_{r-1}$$

En passant sur le cube unité, d'après le lemme 3.1.2, on obtient alors

$$\widetilde{\widetilde{\nabla \pi_r u}} \in \mathbb{Q}^3_{r-1}$$

D'après les lemmes 3.1.1 et 3.1.3, en sommant les degrés, on a finalement

- **Hexaèdre et prisme** :

$$\widetilde{|DF|} \widetilde{DF^{*-1}} \hat{\nabla} \widetilde{w_h} \cdot \widetilde{\widetilde{\nabla \pi_r u}} \in \mathbb{Q}_{2r}$$

- **Pyramide** :

$$\widetilde{|DF|} \widetilde{DF^{*-1}} \hat{\nabla} \widetilde{w_h} \cdot \widetilde{\widetilde{\nabla \pi_r u}} \in \mathbb{Q}_{2r,2r,2r-2}$$

– **Tétraèdre** :

$$\widetilde{|DF|}\widetilde{DF^{*-1}}\hat{\nabla}\widetilde{w_h}\cdot\widetilde{\widetilde{\nabla\pi_r u}}\in\mathbb{Q}_{2r-1,2r-1,2r-2}$$

Finalement, en utilisant les notations suivantes

Hexaèdre :	$t_1=1,\ t_2=2$
Prisme :	$t_1=1,\ t_2=2$
Pyramide :	$t_1=1,\ t_2=0$
Tétraèdre :	$t_1=0,\ t_2=0$

on peut écrire

$$\widetilde{|DF|}\widetilde{DF^{*-1}}\hat{\nabla}\widetilde{w_h}\cdot(\widetilde{\widetilde{\nabla\pi_r u}})\in\mathbb{Q}_{2r-1+t_1,2r-1+t_1,2r-2+t_2}$$

c'est à dire que $E_K(\nabla\pi_r u,\nabla w_h)=0$ pour une formule de quadrature exacte pour les polynômes de $\mathbb{Q}_{2r-1+t_1,2r-1+t_1,2r-2+t_2}$.

En sommant sur les éléments et en prenant $C"_\Omega=\max_{K\in\Omega}\mathcal{C}_K$ et $h=\max_{K\in\Omega}h_K$, on obtient

$$\forall w_h\in V_h,\ |E_K(\nabla I_h u,\nabla w_h)|\leqslant C"_\Omega\, h^r\,||u||_{r+1,\Omega}||w_h||_{1,\Omega}$$

□

Remarque 3.2.12 *Comme dans le cas de l'erreur d'interpolation, on a utilisé le fait que, même si les constantes* $C(\check{\Pi},\check{K})$ *qui apparaissent dans le lemme de Bramble-Hilbert dépendent de la forme de* K*, on peut borner* $\max_K C(\check{\Pi},\check{K})$.

3.2.4.3 Estimation globale de l'erreur de quadrature

On peut à présent obtenir l'estimation d'erreur pour la quadrature.

Proposition 3.2.13 *Pour une formule de quadrature exacte pour les polynômes de* $\widetilde{|T|}\,\mathbb{Q}_{m,m,n}$*, on a*

$$\forall w_h\in V_h,\ \sup_{w_h\in V_h}\frac{|(a-a_h)(I_h u,w_h)|}{||w_h||_{1,\Omega}}\leqslant C\,h^r\,||u||_{r+1,\Omega}$$

avec $m\geqslant 2r-1+t_1,\ n\geqslant 2r-2+t_2$.

Preuve. On note $v_h=I_h u$. Pour une formule de quadrature exacte pour les polynômes de $\widetilde{|T|}\,\mathbb{Q}_{m,m,n}$, on utilise les propositions 3.2.10 et 3.2.11 en prenant le résultat le plus restrictif. En remarquant que, dans tous les cas, $2r-2+s_1\leqslant 2r-1+t_1$ et $2r-2+s_2\leqslant 2r-2+t_2$, on a alors

$$\forall w_h\in V_h,\ |E_K(v_h,w_h)|+|E_K(\nabla v_h,\nabla w_h)|\leqslant C'_\Omega\,h^r\,||u||_{r+1,\Omega}||w_h||_{1,\Omega}+C"_\Omega\,h^r\,||u||_{r+1,\Omega}||w_h||_{1,\Omega}$$

pour $m\geqslant 2r-1+t_1,\ n\geqslant 2r-2+t_2$.

Soit $C=\max(C'_\Omega,C"_\Omega)$, on obtient alors

$$\forall w_h\in V_h,\ |(a-a_h)(v_h,w_h)|\leqslant|E_K(v_h,w_h)|+|E_K(\nabla v_h,\nabla w_h)|\leqslant C\,h^r\,||u||_{r+1,\Omega}||w_h||_{1,\Omega}$$

Le résultat voulu est ainsi obtenu en divisant par $||w_h||_{1,\Omega}$ pour $w_h\neq 0$ et en prenant le supremum sur V_h. □

3.2.5 Estimation globale

On peut à présent écrire l'estimation globale.

Theorème 3.2.14 *Pour une formule de quadrature exacte pour les polynômes de* $\widetilde{|T|}\,\mathbb{Q}_{m,m,n}$*, avec* $m\geqslant 2r-1+t_1$ *et* $n\geqslant 2r-2+t_2$*, l'estimation finale est*

$$\inf_{v_h\in V_h}\left(||u-v_h||_{1,\Omega}+\sup_{w_h\in V_h}\frac{|(a-a_h)(v_h,w_h)|}{||w_h||_{1,\Omega}}\right)\leqslant\mathcal{C}\,h^r||u||_{r+1,\Omega}.$$

Preuve. En sommant les résultats respectifs des propositions 3.2.6 et 3.2.13, on obtient

$$\|u - I_h u\|_{1,\Omega} + \sup_{w_h \in V_h} \frac{|(a - a_h)(I_h u, w_h)|}{\|w_h\|_{1,\Omega}} \leqslant \mathcal{C}\, h^r \|u\|_{r+1,\Omega}$$

où $\mathcal{C} = \max(C_\Omega, C)$.

Or, puisque $I_h u \in V_h$, on a

$$\inf_{v_h \in V_h} \left(\|u - v_h\|_{1,\Omega} + \sup_{w_h \in V_h} \frac{|(a - a_h)(v_h, w_h)|}{\|w_h\|_{1,\Omega}} \right) \leqslant \|u - I_h u\|_{1,\Omega} + \sup_{w_h \in V_h} \frac{|(a - a_h)(I_h u, w_h)|}{\|w_h\|_{1,\Omega}}$$

d'où le résultat. □

Remarque 3.2.15 *Étant donné que l'inclusion $\mathbb{P}_r \subset P_r^F$ est nécessaire à l'application du lemme de Bramble-Hilbert, nous pensons que l'espace vérifiant cette inclusion est l'espace de dimension minimale permettant d'obtenir une erreur d'interpolation sur l'élément en $O(h^r)$ pour la norme H^1 pour une solution suffisamment régulière.*

Chapitre 4

Étude numérique des éléments continus

Afin de dégager des propriétés numériques des éléments construits dans le chapitre 2, on effectue une analyse de dispersion et une étude de stabilité des schémas obtenus à partir de ces éléments. On rappelle les principes de l'analyse de dispersion et l'expression de la condition de stabilité, le détail pouvant être trouvé dans Cohen [17], et on vérifie que l'on obtient les mêmes estimations d'erreurs que celles trouvées en théorie dans le chapitre 3. Un cas test numérique vient finalement confirmer le bon comportement des éléments au sein d'un maillage hybride.

Sommaire

4.1 Analyse de dispersion

4.1.1 Rappels théoriques

On considère l'équation de Helmholtz

$$-\omega^2 u - \Delta u = 0.$$

L'analyse de dispersion par une technique de type onde plane s'effectue dans un milieu infini homogène sur un maillage périodique, de préférence régulier. Elle consiste en l'étude de l'équation harmonique considérée par des ondes planes s'écrivant

$$u(x,t) = u_0(x)e^{i(kx-\omega t)}, \tag{4.1.1}$$

avec

$$k = (k_x, k_y, k_z) = (k\cos\theta\cos\phi, k\sin\theta\cos\phi, k\sin\phi),$$

où θ et ϕ désignent les angles d'incidence de l'onde plane. La relation de dispersion consiste à écrire la relation que doivent respecter la pulsation ω et le vecteur d'onde k pour que u soit solution de l'équation continue. Dans le cas de l'équation des ondes, la relation de dispersion est

$$\omega^2 = |k|^2,$$

$|k|$ désignant le nombre d'onde.

Une fois l'onde plane injectée dans le système « infini » (sur tout le maillage), on réduit le domaine de calcul à une seule cellule périodique de taille h sur laquelle on impose des conditions de quasi-périodicité

$$\begin{aligned} u(x+h,y,z) &= e^{ik_x h}u(x,y,z) \\ u(x,y+h,z) &= e^{ik_y h}u(x,y,z) \\ u(x,y,z+h) &= e^{ik_z h}u(x,y,z) \end{aligned}$$

On se ramène donc à un système discret du type

$$(-\omega^2 M_h + R_h)U = 0$$

où M_h et R_h sont calculées sur la cellule périodique.

Il s'agit donc de résoudre un problème aux valeurs propres pour la matrice $M_h^{-1}K_h$. La valeur propre numérique la plus proche de $|k|$ est noté ω_h qui est la pulsation approchée. L'analyse de la dispersion numérique du schéma consiste finalement en l'étude des variations de la vitesse de phase adimensionnelle q_h définie par

$$q_h = \frac{\omega_h}{\omega} = \frac{\omega_h}{|k|}.$$

Comme q_h doit être proche de 1, on peut écrire

$$q_h = 1 + C\,h^p + o(h^p),$$

où p désigne l'ordre de dispersion du schéma numérique.

On montre que cette quantité dépend de θ, ϕ et du paramètre adimensionnel K défini par

$$K = \frac{nkh}{2\pi r},$$

où h désigne le pas de maillage du motif périodique. L'ordre r donne ici une certaine relativité aux résultats obtenus par rapport à l'ordre, et le facteur n est le nombre de points par longueur d'onde que l'on souhaite avoir dans le maillage pour $K = 1$.

4.1.2 Résultats numériques

Pour cette étude, on prend les élément nodaux définis dans la section 2.3.1. Dans notre cas, le maillage périodique infini est obtenu à partir d'une cellule prise comme un cube découpé en un unique hexaèdre, deux prismes, deux pyramides et deux tétraèdres (hybride), six pyramides ou six tétraèdres comme le montre la figure 4.1.

L'analyse a également été faite sur des cellules périodiques faites à partir de cubes déformés afin de vérifier la consistance de nos méthodes lorsque les éléments ne sont pas affines : la figure 4.2 présente la cellule hybride déformée utilisée.

On obtient un ordre de dispersion de $2r$, aussi bien pour les maillages avec éléments affines que pour les maillages déformés (voir Babuska et Osborn [4] pour le facteur 2), ce qui coïncide avec les résultats d'estimation d'erreur

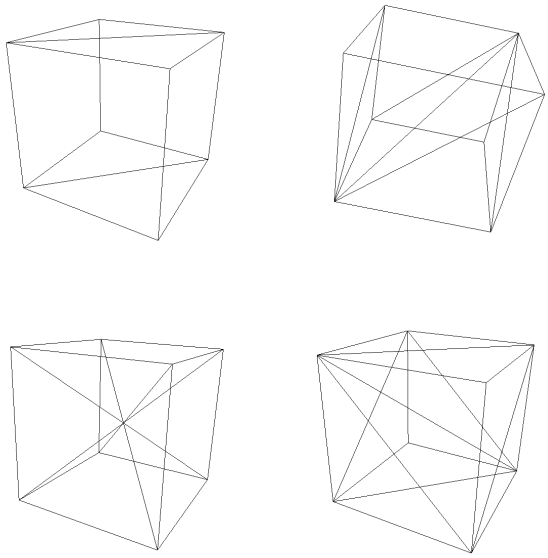

FIG. 4.1 – Cellules utilisées pour créer un maillage périodique infini : prismes (en haut à gauche), hybride (en haut à droite), pyramides (en bas à gauche), tétraèdres (en bas à droite)

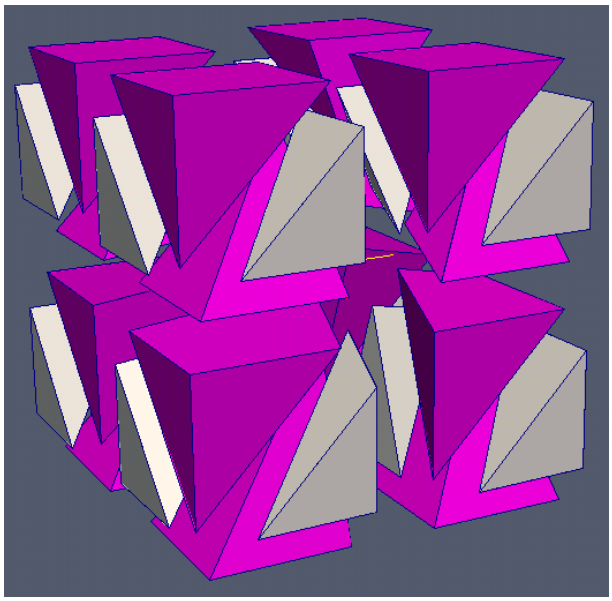

FIG. 4.2 – Motif périodique dans le cas hybrides, avec des pyramides non affines (violet) et des tétraèdres (gris)

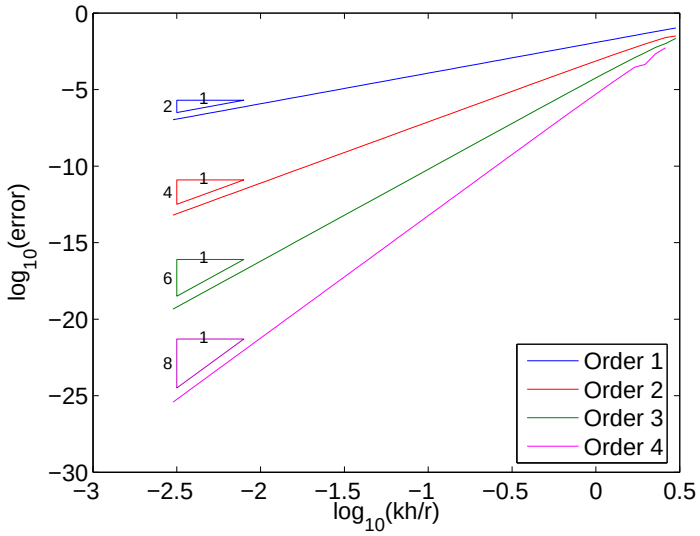

FIG. 4.3 – Erreur de dispersion en échelle logarithmique pour les éléments continus d'ordre 1 à 4 pour un maillage hybride déformé

théoriques obtenus précédemment. La courbe en échelle logarithmique de la figure. 4.3 montre ce résultat sur la cellule hybride déformée, qui est le cas le plus « difficile », pour les quatre premiers ordres.

Les courbes de dispersion pour les éléments réguliers d'ordre 1 à 3 sont présentées sur la figure 4.4. Le tétraèdre, le prisme et la cellule hybride donnent des dispersions très proches. Pour les éléments pyramidaux, l'approximation exacte et l'approximation minimale donnent des résultats très proches, l'élément le moins dispersif étant toujours l'élément pyramidal. La même étude a été menée sur les éléments déformés, comme le montre la figure 4.5, et mènent aux mêmes observations. Dans les deux cas, la dispersion pour tous les éléments diminue lorsque l'on monte en ordre.

4.2 Étude de stabilité

4.2.1 Condition CFL

Bien qu'utiliser les éléments finis nodaux continus en régime temporel ne soit pas la méthode la plus efficace, on souhaite étudier la condition de stabilité de ces éléments pour l'équation des ondes avec une discrétisation en temps par un schéma centré d'ordre deux. On considère là encore un maillage périodique infini.

Pour tout schéma temporel, la condition de Courant-Friedrichs-Lewy (CFL), pour laquelle on a la condition de stabilité $\Delta t \leqslant \mathrm{CFL}\ h$, est définie par

$$CFL = \frac{\alpha}{\sqrt{\max\limits_{|k|\leqslant\pi} \lambda(M_h^{-1}(k)\, R_h(k))}},$$

où α dépend du schéma de discrétisation en temps considéré. Pour un schéma centré d'ordre deux, $\alpha = 1$. Les matrices $M_h(k)$ et $R_h(k)$ sont les matrices de masse et de rigidité définies sur une cellule périodique, comme pour l'analyse de dispersion, et k le vecteur d'onde.

4.2.2 Résultats numériques

Pour chaque type d'élément, le tableau 4.1 donne la CFL obtenue jusqu'à l'ordre 4 sur des cellules régulières, et jusqu'à l'ordre 3 pour les cellules déformées. Pour les éléments pyramidaux, la CFL est recherchée avec la formule de quadrature $(\xi_k^{GJ2}, \omega_k^{GJ2})$ présentée dans la section 3.1.3 du chapitre 3 pour l'intégration exacte, et avec une formule de quadrature de Gauss $(\xi_k^{G}, \omega_k^{G})$ pour l'intégration approchée.

La condition CFL des éléments pyramidaux est clairement plus élevée lorsque l'on considère l'intégration exacte que lorsque l'on calcule les intégrales avec la formule de Gauss (60% plus élevée en moyenne). De manière générale,

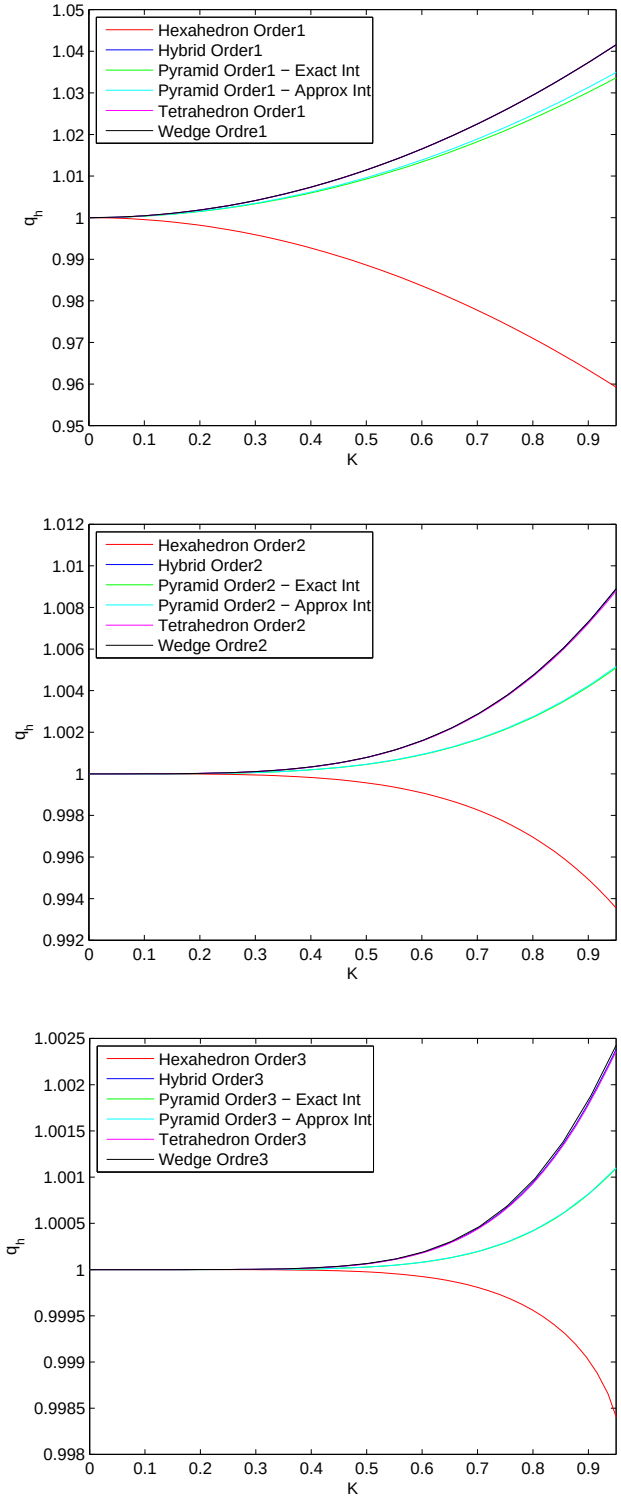

FIG. 4.4 – Courbes de dispersion pour les éléments finis nodaux d'ordre 1 à 3 sur un maillage régulier ($K = \frac{6kh}{2\pi r}$)

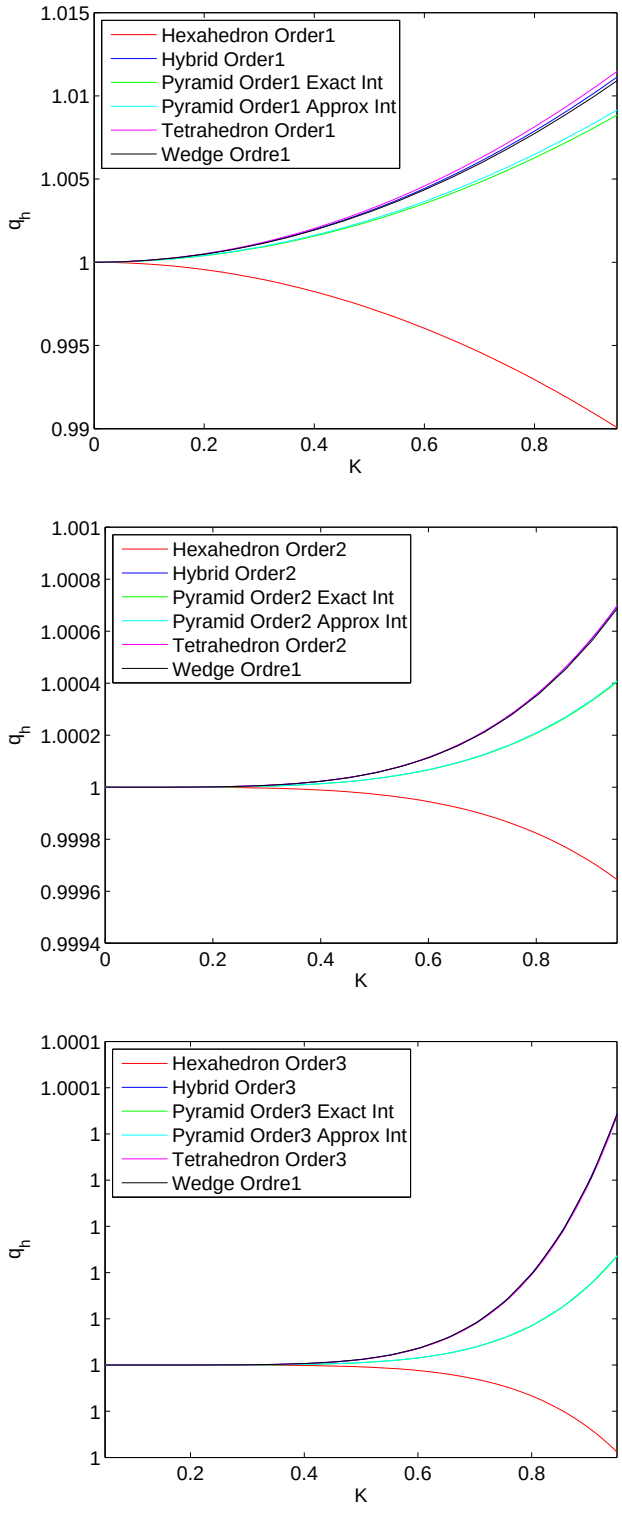

FIG. 4.5 – Courbes de dispersion pour les éléments finis nodaux d'ordre 1 à 3 sur un maillage déformé ($K = \frac{6kh}{2\pi r}$)

TAB. 4.1 – Stabilité des éléments continus pour un maillage régulier et un maillage déformé

Élément	Maillage régulier				Maillage déformé		
	Ordre 1	Ordre 2	Ordre 3	Ordre 4	Ordre 1	Ordre 2	Ordre 3
Hexaèdre	0.28868	0.11785	0.06697	0.04264	0.28306	0.11296	0.06192
Prisme	0.16666	0.07454	0.04426	0.02926	0.16390	0.07028	0.04120
Pyramide IntExacte	0.09682	0.04803	0.03083	0.02143	-	-	-
Pyramide IntApprox	0.07217	0.03335	0.01985	0.01316	0.07142	0.03296	0.01962
Hybride	0.14887	0.07251	0.04568	0.03191	0.14708	0.07056	0.04256
Tétraèdre	0.11180	0.05975	0.03815	0.02669	0.01372	0.04978	0.02640

pour tous les éléments, la CFL est plus basse dans le cas des maillages déformés, et pour les deux types de maillage, la condition CFL des différents types d'éléments se classe alors comme suit

$$CFL_{Hexa} > CFL_{Hybride} > CFL_{Prisme} > CFL_{Tetra} > CFL_{Pyr-IntExacte} > CFL_{Pyr-IntApprox}.$$

Le fait que la CFL sur les maillages hybrides soit meilleure que celle obtenue sur les tétraèdres et les pyramides constitue un résultat assez surprenant qui peut s'expliquer par la taille des éléments, plus gros dans le cas des maillages obtenus à partir d'une cellule hybride. De ce fait, il y a moins d'éléments dans la cellule hybride, donc moins de degrés de liberté.

Remarque 4.2.1 *Les CFL ont été vérifiées dans le cas instationnaire.*

4.3 Convergence

On souhaite vérifier l'ordre de convergence obtenu à l'aide de l'étude de dispersion. On considère l'équation de Helmholtz (voir équation 1.3.8) sur une cavité cubique $[-1,1]^3$ avec conditions de Dirichlet homogènes au bord. On prend $\omega = 1.92\,\pi$ et f est une source gaussienne centrée à l'origine.

On étudie la convergence sur un maillage hybride avec des motifs similiaires à ceux utilisés dans l'étude de dispersion et de stabilité.

On trace l'erreur obtenue en norme H^1 par rapport au pas du maillage h en échelle log-log sur la figure 4.6. On observe que l'erreur en norme H^1 est en $O(h^r)$ comme nous l'avons démontré lors des estimations d'erreur, et l'erreur en norme L^2 est en $O(h^{r+1})$. En effet, puisque le système hyperbolique est symétrique pour l'équation de Helmholtz, c'est à dire que le problème adjoint est également consistant, on a une convergence en $O(h^{r+1})$ pour la norme L^2 (conséquence du lemme de Aubin-Nitsche, voir Ciarlet [14]).

Pour ce cas test, l'intégration utilisée pour les pyramides est celle avec $r+1$ points de Gauss-Jacobi dans la direction $\tilde{z}$ et $r+1$ points de Gauss dans les directions $\tilde{x}$ et $\tilde{y}$.

4.4 Équation de Helmholtz sur un cone-sphère

Afin de vérifier le bon fonctionnement des éléments dans un maillage hybride, on présentera un cas-test standard sur l'équation de Helmholtz pour lequel les résultats sont bien connus.

On considère la diffraction par un cone-sphère de bord Γ placé dans une boite parallélépipédique Σ

$$\left\{ \begin{array}{rcl} -\omega^2\, u - \Delta u & = & 0 \quad \text{sur } \Omega \\ u & = & -u^{\text{inc}} \quad \text{sur } \Gamma \\ \dfrac{\partial u}{\partial n} & = & i\omega\, u \text{ sur } \Sigma \end{array} \right. \tag{4.4.1}$$

On considère le cas où le champ incident u^{inc} est une onde plane de type

$$u^{\text{inc}} = e^{i\,k\cdot x},$$

avec k le vecteur d'onde valant ici $(0,-\omega,0)$, c'est à dire que l'onde arrive sur le cone-sphère par la pointe.

La solution numérique obtenue pour un maillage hybride contenant plus d'un million de degrés de liberté et utilisant des éléments d'ordre 3 est donnée par la figure 4.7. Sur un maillage hybride plus grossier contenant 450 000 ddls, on observe une erreur de 2.6% par rapport à cette solution de référence.

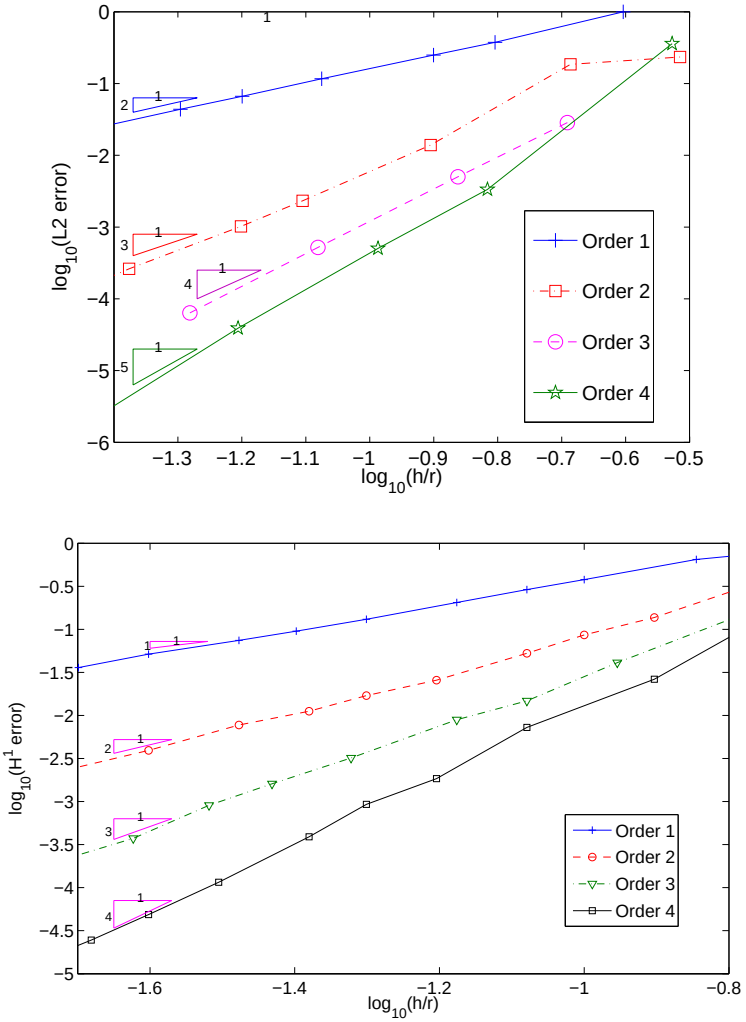

FIG. 4.6 – Erreur en norme L^2 (en haut) et en norme H^1 (en bas) par rapport au pas du maillage h pour une cavité cubique avec différents ordres d'approximation. Maillage hybride avec des pyramides non-affines

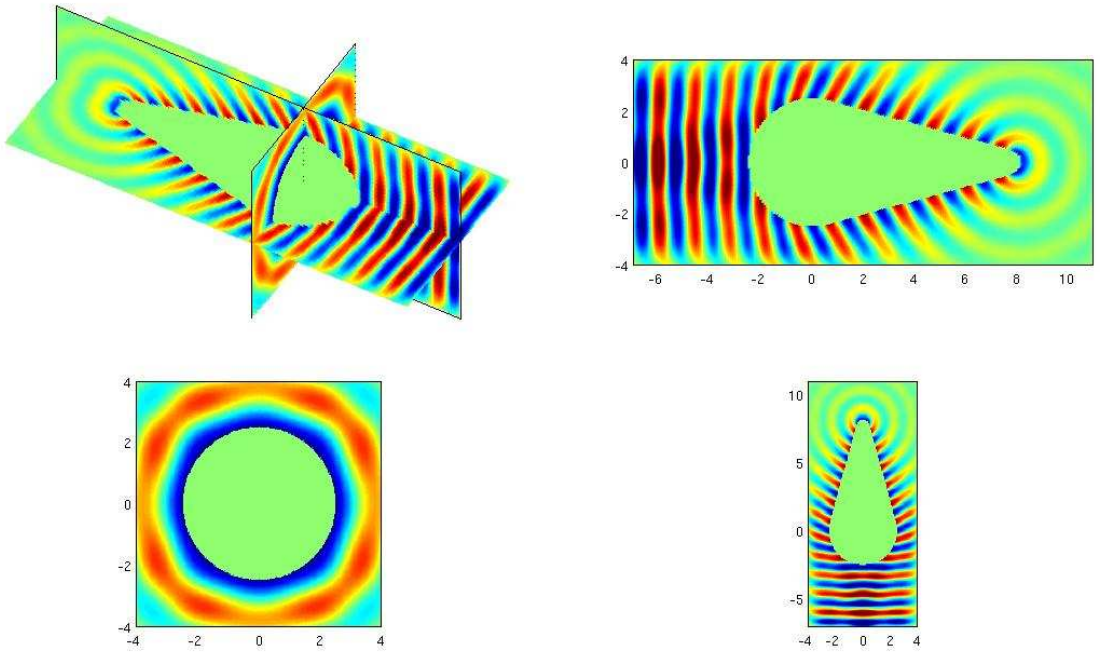

FIG. 4.7 – Partie réelle du champ diffracté par le cone-sphère sur un maillage hybride pour des éléments d'ordre 3

Remarque 4.4.1 *On voit nettement l'effet de la condition absorbante d'ordre 1 placée sur Σ : en plaçant des PML, une condition transparente ou une condition absorbante d'ordre plus élevé, on aurait obtenu une solution à symétrie de révolution.*

On lance la même simulation sur deux maillages différents (voir figure 4.8) avec des éléments d'ordre 3. On utilise volontairement des éléments droits et non des éléments courbes pour prendre en compte la géométrie. En effet, dans le cas de maillages tétraédriques, la projection du point milieu des arêtes sur la géométrie engendre parfois un élément dégénéré. Ce phénomène se retrouve donc sur les maillages composés de tétraèdres découpés, alors que le problème de dégénérescence se constate nettement moins pour les hexaèdres « normaux ».

Le nombre de degrés de liberté utilisés et l'erreur en norme H^1 par rapport à la solution de référence calculée sur un maillage hybride de 1 million de degrés de liberté sont présentés dans le tableau 4.2. On obtient donc une précision similaire avec 4 fois moins de degrés de liberté en utilisant un maillage hybride.

TAB. 4.2 – Nombre de degrés de liberté et erreur en norme H^1 par rapport à une solution de référence pour deux types de maillages

Type de maillage	Nombre de ddl	Erreur H^1
Tétras découpés	1 077 000 ddls	9.0%
Hybride	247 000 ddls	7.7%

4.5 Remarques générales

Les maillages hybrides, lorsqu'ils conservent la géométrie (ce qui n'est pas toujours le cas des mailleurs du commerce), permettent d'avoir un maillage très bien conditionné : les CFL sont plus élevées, ce qui permet d'utiliser des Δt plus élevés, et les éléments sont plus gros, on a donc moins de degrés de liberté. En outre, beaucoup d'éléments sont affines, ce qui permet, pour les hexaèdres, de ne stocker qu'un seul DF^{-1}. En exploitant le caractère constant du jacobien, on peut également accélérer les calculs sur les prismes et les pyramides.

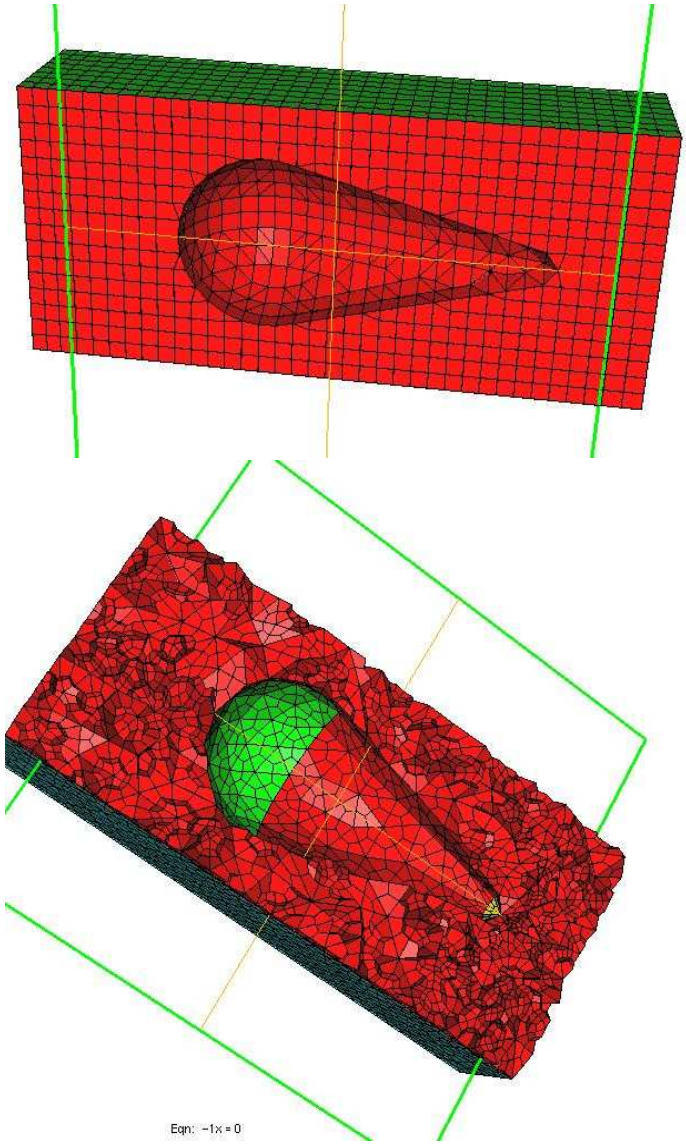

FIG. 4.8 – Maillages utilisés pour l'ordre 3 : maillage hybride (en haut) et maillage hexaédrique (en bas) obtenu avec des tétraèdres découpés

Les fonctions de base utilisant des fractions rationnelles s'avèrent particulièrement adaptées aux éléments finis pyramidaux, notamment en vertu de leur respect des conditions de restriction sur les faces de la pyramide, et donc de la continuité des fonctions de base aux interfaces avec les autres types d'éléments. Contrairement à ce qu'affirment Bluck et Walter [7], la présence d'un pôle au dénominateur ne pose pas de problème majeur dans leur manipulation numérique, et les matrices peuvent être intégrées de manière satisfaisante. Il est en outre possible de remédier à la non dérivabilité des fonctions de base à l'apex de la pyramide en imposant, comme le suggère Bedrosian [5], une valeur à la dérivée à l'apex de la pyramide (par exemple 0). On peut également prendre comme élément de référence cube unité, puisque sur le cube, DF est polynomial. En ne prenant aucun point de quadrature sur l'apex, le seul pb pouvant apparaitre est l'évaluation du gradient de la solution sur l'apex d'une pyramide du maillage. Dans ce cas, imposer 0 pour les dérivées des fonctions de base conduit à un jacobien non-inversible. Il est alors plus judicieux de prendre par exemple la limite quand z tend vers 1 en imposant $x = y = 0$, qui est l'option que nous avons retenue.

Chapitre 5

Comparaison entre différentes méthodes

Nous comparons ici les éléments optimaux obtenus dans le chapitre 2 avec ceux que l'on peut trouver dans la littérature. On s'intéresse en particulier aux éléments pyramidaux pour lesquels une comparaison numérique est également effectuée.

Sommaire

5.1 Introduction

Pour les hexaèdres, l'utilisation de fonctions nodales basées sur les points de Gauss-Lobatto permet de réduire de manière significative le temps de calcul et le stockage (Cohen et Fauqueux [19], Duruflé [28]).

Concernant les tétraèdres, on peut citer les travaux de Hesthaven et Teng [44] qui construisent des éléments tétraédriques en plaçant les degrés de liberté sur les « points électrostatiques » fournissant une bonne constante de Lebesgue. Les éléments tétraédriques que nous utilisons sont issus de ces travaux.

Les éléments finis construits sur des prismes (en anglais « triangular prism » ou « wedge ») sont obtenus de manière classique par la tensorisation d'un élément fini triangulaire par un élément 1D (Lunéville [9], Ciarlet [14] et Šolín [71]). L'intérêt de la tensorisation étant de diminuer de manière importante les calculs de quadrature effectués pour l'évaluation des intégrales, cette propriété est donc recherchée et exploitée dès que possible.

L'obtention d'une base appropriée pour les pyramides étant un point délicat lors de leur construction, plusieurs approches ont été considérées.

5.2 Éléments pyramidaux dans la littérature

5.2.1 Éléments nodaux à base rationnelle

La première approche pour construire des éléments pyramidaux consiste à utiliser des fonctions de base contenant des fractions rationnelles.

- Bedrosian dans [5] propose des fonctions de base rationnelles pour des approximations du premier et du second ordre. Cependant, au second ordre, Bedrosian n'ajoute pas de degré de liberté au centre de la face quadrangulaire de la pyramide, ce qui interdit toute conformité avec les éléments hexaédriques du second ordre.
- Doucet [24] retrouve les fonctions de base de Bedrosian d'ordre 1. Zgainski *et al.* [76] conduisent des expériences numériques avec les fonctions de Bedrosian et proposent une famille modifiée de fonctions de bases du second ordre, ajoutant cette fois un degré de liberté au centre de la base de la pyramide. Cependant, la fonction de base centrale ne satisfait pas la condition lagrangienne $\varphi_i(M_j) = \delta_{ij}$, et la modification n'améliore pas la précision puisque l'espace d'approximation généré par cette famille de fonctions de base ne contient pas l'espace $\mathbb{P}_2$. La même idée est reprise par Graglia *et al.* [38] qui parviennent à améliorer la précision avec leur propre fonction de base centrale du second ordre.
- Chatzi et Preparata [13] introduisent une généralisation des fonctions de base de Bedrosian à un ordre quelconque pour des degrés de liberté régulièrement distribués sur la pyramide, comme représenté sur la figure 5.1. Malheureusement, les fonctions de base proposées ne génèrent pas l'espace des polynômes dès l'ordre 3, et ne sont donc pas consistantes.

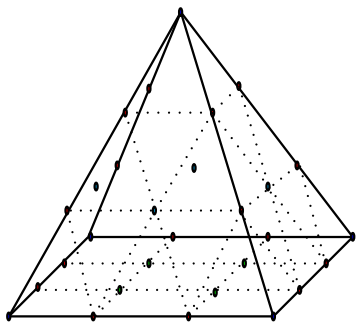

FIG. 5.1 – Pyramide régulière d'ordre 3

5.2.2 Pyramides découpées en tétraèdres

La seconde approche consiste à découper la pyramide en tétraèdres afin d'éviter d'utiliser des fractions rationnelles qui ont la réputation (discutable) d'être difficile à utiliser, et d'utiliser des fonctions de base polynomiales.

- Au premier ordre, Wieners [74], Knabner et Summ [49], ainsi que Bluck et Walker [7] donnent une famille consistante de fonctions de base qui permet d'assurer la conformité avec les hexaèdres et les tétraèdres, en découpant une pyramide en deux tétraèdres. Wieners propose une famille de fonctions de base du second ordre, et des fonctions de base d'ordre élevé sont proposées par Bluck et Walker. Cependant, les espaces

d'approximation générés par ces familles d'ordre élevé ne contiennent pas les espaces d'approximation d'ordre plus bas, ce qui conduit à des méthodes non consistantes pour des ordres élevés dans le cas de pyramides non affines. De plus, cette méthode nécessite des formules de quadrature coûteuses sur chaque tétraèdre.
- Liu *et al.* [51] proposent une version symétrisée des fonctions de base de Wieners, mais cette modification n'améliore qu'à peine la précision de la méthode originale.

5.2.3 Éléments *hp*

Une autre méthode populaire pour les éléments finis est l'approche *hp* (Szabó et Babuška [68]), avec par exemple Šolín *et al.* [71] pour les hexaèdres, les tétraèdres et les prismes. Plusieurs articles étendent le concept d'élément fini *hp* aux éléments pyramidaux.
- Warburton [72], Sherwin [65], Sherwin *et al.* [66], ainsi que Karniadakis et Sherwin [47] proposent une famille de fonctions de bases tensorielles pour tous les types d'éléments à partir de la dégénérescence d'un cube. Pour les tétraèdres, les hexaèdres et les prismes, les espaces d'approximation générés par ces familles sont les espaces classiques. Pour les pyramides, les espaces d'approximation proposés permettent d'obtenir une convergence optimale dans le cas de pyramides affines, mais concernant les pyramides déformées, ce n'est plus le cas au delà de l'ordre deux. De plus, la transition continue entre les pyramides et les autres types d'éléments n'est pas possible dans le cas général de maillages non structurés car les fonctions proposées ne sont pas invariantes par rotation.
- Nigam and Phillips [58] proposent un autre espace d'approximation en utilisant une pyramide infinie comme élément de référence. Avec l'espace d'approximation obtenu, la précision est conservée, mais la dimension de l'espace pourrait être réduite. Dans un article ultérieur [59], ils proposent une correction de leur espace initial pour obtenir une dimension optimale.
- Demkowicz *et al.* [23] et Zaglmayr [75] proposent la construction de fonctions de base partiellement orthogonales pour les tétraèdres, les hexaèdres et les prismes, et utilisent la dégénérescence du cube pour construire un espace d'approximation qui préserve la précision optimale, avec une dimension égale à celle de Nigam et Phillips [59].

5.3 Comparaison d'éléments pyramidaux

5.3.1 Comparaison théorique

- Les fonctions de base nodales présentées dans le chapitre 2 sont les mêmes que celles proposées par Bedrosian [5], Zgainski *et al.* [76], ainsi que Chatzi et Preparata [13] à l'ordre 1. À l'ordre 2, ce sont les mêmes que celles de Graglia *et al.* [38], et sont totalement nouvelles pour les ordres supérieurs.

- L'espace d'approximation C_r d'ordre r sur le cube unité $\widetilde{Q}$ défini dans la proposition 2.2.8 est le même que celui proposé par Zaglmayr, citée dans [23], et Nigam et Phillips dans leur second article [59].

Proposition 5.3.1 *Le sous espace C_r^0 de C_r dont la trace est nulle sur la frontière de $\widetilde{Q}$ est*

$$C_r^0(\widetilde{x},\widetilde{y},\widetilde{z}) = (1-\widetilde{z})^2\,\widetilde{x}(1-\widetilde{x})\,\widetilde{y}(1-\widetilde{y})\,\widetilde{z}\,\widetilde{C}_{r-3}.$$

Preuve. Les fonctions de base s'annulent de manière évidente sur la frontière de $\widetilde{Q}$, et appartiennent à C_r. La dimension de l'espace est dim $C_{r-3} = \dfrac{1}{6}(r-1)(r-2)(2r-3) = n_i$, ce qui achève la preuve de la proposition. □

- On définit la transformation $\bar{T}$ de la pyramide infinie $\bar{Q}$ vers le cube unité $\widetilde{Q}$.

$$\bar{T} : \begin{cases} \widetilde{x} = \bar{x} \\ \widetilde{y} = \bar{y} \\ \widetilde{z} = \dfrac{\bar{z}}{1+\bar{z}}. \end{cases} \tag{5.3.1}$$

Proposition 5.3.2 *L'espace d'approximation U_r proposé par Nigam et Phillips dans [58] sur la pyramide infinie $\bar{Q}$ vérifie*

$$U_r \supset C_r \circ \bar{T},$$

et contient plus de degrés de liberté que C_r puisque

$$dim\ U_r = 1 + 3k + k^3 > dim\ C_r.$$

Le sous espace U_r^0 de U_r dont la trace est nulle sur le bord de l'élément est

$$U_r^0(\bar{x},\bar{y},\bar{z}) = \left\{ \frac{\bar{x}(1-\bar{x})\bar{y}(1-\bar{y})\bar{z}}{(1+\bar{z})^r}\, u(\bar{x},\bar{y},\bar{z}),\ u \in \mathbb{Q}^{r-2}(\bar{x},\bar{y},\bar{z}) \right\},$$

et si l'on remplace U_0^r par $C_r^0 \circ \bar{T}$, on obtient l'espace optimal

$$\bar{U}_r = C_r \circ \bar{T}.$$

Preuve. On utilise la transformation (5.3.1) pour traiter les fonctions de base suivantes (les autres pouvant être obtenues de manière similaire par symétrie)

Pour les sommets : $\dfrac{(1-\bar{x})(1-\bar{y})}{(1+\bar{z})^r} \circ \bar{T}^{-1} = (1-\widetilde{x})(1-\widetilde{y})(1-\widetilde{z})^r \in C_r.$

Pour l'apex : $\dfrac{\bar{z}^r}{(1+\bar{z})^r} \circ \bar{T}^{-1} = \widetilde{z}^r \in C_r.$

Pour une arête verticale :
$$\left\{ \frac{(1-\bar{x})(1-\bar{y})\bar{z}^a}{(1+\bar{z})^r},\ 1 \leqslant a \leqslant r-1 \right\} \circ \bar{T}^{-1} = \left\{ (1-\widetilde{x})(1-\widetilde{y})(1-\widetilde{z})^{r-a}\widetilde{z}^a,\ 1 \leqslant a \leqslant r-1 \right\} \subset C_r.$$

Pour une arête de la base :
$$\left\{ \frac{(1-\bar{x})(1-\bar{y})\bar{x}^a}{(1+\bar{z})^r},\ 1 \leqslant a \leqslant r-1 \right\} \circ \bar{T}^{-1} = \{ (1-\widetilde{x})(1-\widetilde{y})\widetilde{x}^a(1-\widetilde{z})^r,\ 1 \leqslant a \leqslant r-1 \} \subset C_r.$$

Pour une face triangulaire : $\left\{ \dfrac{(1-\bar{x})(1-\bar{y})\bar{x}^a\bar{z}^b}{(1+\bar{z})^r},\ a,b \geqslant 1, a+b \leqslant r-1 \right\} \circ \bar{T}^{-1} =$
$$\left\{ (1-\widetilde{x})(1-\widetilde{y})\widetilde{x}^a(1-\widetilde{z})^{r-b}\widetilde{z}^b,\ 1 \leqslant a+b \leqslant r-1 \right\} \subset C_r.$$

Pour la base :
$$\left\{ \frac{(1-\bar{x})(1-\bar{y})\bar{x}^a\bar{y}^b}{(1+\bar{z})^r},\ 1 \leqslant a,b \leqslant r-1 \right\} \circ \bar{T}^{-1} = \left\{ (1-\widetilde{x})(1-\widetilde{y})\widetilde{x}^a\widetilde{y}^b(1-\widetilde{z})^r,\ 1 \leqslant a,b \leqslant r-1 \right\} \subset C_r.$$

Pour l'intérieur :
$$\left\{ \frac{\bar{x}(1-\bar{x})\bar{y}(1-\bar{y})\bar{z}}{(1+\bar{z})^r}\, u(\bar{x},\bar{y},\bar{z}),\ u \in \mathbb{Q}^{r-2}(\bar{x},\bar{y},\bar{z}) \right\} \circ \bar{T}^{-1} =$$
$$\left\{ \widetilde{x}^{i+1}(1-\widetilde{x})\widetilde{y}^{j+1}(1-\widetilde{y})\widetilde{z}^{k+1}(1-\widetilde{z})^{r-k-1},\ 0 \leqslant i,j,k \leqslant r-2 \right\} \supset C_r^0.$$

Le sous espace de U_r de trace nulle au bord de l'élément est

$$U_r^0 = \left\{ \frac{\bar{x}(1-\bar{x})\bar{y}(1-\bar{y})\bar{z}}{(1+\bar{z})^r}\, u(\bar{x},\bar{y},\bar{z}),\ u \in \mathbb{Q}^{r-2}(\bar{x},\bar{y},\bar{z}) \right\}$$

dont la dimension est

$$\dim U_r^0 = \dim \mathbb{Q}_{r-2} = (r-1)^3.$$

Puisqu'il y a $n_f = 3r^2 + 2$ fonctions de base associées à la frontière, on a

$$\dim U_r = 3r^2 + 2 + (r-1)^3 = 1 + 3r + r^3 > \dim C_r.$$

Si l'on remplace U_r^0 par $C_r^0 \circ \bar{T}$, le nouvel espace d'approximation $\bar{U}_r$ vérifie

$$\dim \bar{U}_r = \dim C_r,$$

et

$$\bar{U}_r \supset C_r \circ \bar{T},$$

i.e. il y a égalité entre les deux espaces. □

- On définit la transformation $\widehat{T}$ du cube $\widehat{Q}(a,b,c) = [-1,1]^3$ vers le cube unité $\widetilde{Q}$

$$\widehat{T} : \begin{cases} \widetilde{x} = \dfrac{1+a}{2} \\ \widetilde{y} = \dfrac{1+b}{2} \\ \widetilde{z} = \dfrac{1+c}{2}. \end{cases} \tag{5.3.2}$$

Proposition 5.3.3 *L'espace d'approximation W_r d'ordre r introduit par Warburton [72] sur le cube $[-1,1]^3$ n'est pas optimal en termes de dimension.*

Le sous espace de W_r dont la trace est nulle à la frontière de l'élément est

$$W_r^0 \circ \widehat{T}^{-1} = \left\{ \widetilde{x}(1-\widetilde{x})\widetilde{y}(1-\widetilde{y})\widetilde{z}(1-\widetilde{z})^2\, u(\widetilde{x},\widetilde{y},\widetilde{z}),\ u \in \mathbb{P}_{r-3}(\widetilde{x},\widetilde{y},\widetilde{z}) \right\}.$$

En remplaçant W_0^r par $C_r^0 \circ \widehat{T}$, et les fonctions de base liées à la face quadrangulaire par les fonctions suivantes

$$\left\{ \left(\frac{1-a}{2}\right)\left(\frac{1+a}{2}\right)\left(\frac{1-b}{2}\right)\left(\frac{1+b}{2}\right)\left(\frac{1-c}{2}\right)^{max(i,j)+1} P_{i-1}^{1,1}(a)P_{j-1}^{1,1}(b),\ 1 \leqslant i,j \leqslant r-1 \right\},$$

on obtient l'espace optimal

$$\widehat{W}_r = C_r \circ \widehat{T}.$$

Preuve. On utilise la transformation (5.3.2) pour traiter les fonctions de base suivantes (les autres pouvant être obtenues de manière similaire par symétrie)

Pour les sommets : $\left\{ \left(\frac{1-a}{2}\right)\left(\frac{1-b}{2}\right)\left(\frac{1-c}{2}\right) \right\} \circ \widehat{T}^{-1} = (1-\widetilde{x})(1-\widetilde{y})(1-\widetilde{z}) \in C_r.$

Pour l'apex : $\left\{ \frac{1+c}{2} \right\} \circ \widehat{T}^{-1} = \widetilde{z} \in C_r.$

Pour une arête verticale : $\left\{ \left(\frac{1-a}{2}\right)\left(\frac{1-b}{2}\right)\left(\frac{1-c}{2}\right)\left(\frac{1+c}{2}\right) P_{i-1}^{1,1}(c),\ 1 \leqslant i \leqslant r-1 \right\} \circ \widehat{T}^{-1} =$
$\left\{ (1-\widetilde{x})(1-\widetilde{y})(1-\widetilde{z})\widetilde{z}\, P_{i-1}^{1,1}(2\widetilde{z}-1),\ 1 \leqslant i \leqslant r-1 \right\} \subset C_r.$

Pour une arête de la base :
$\left\{ \left(\frac{1-a}{2}\right)\left(\frac{1+a}{2}\right)\left(\frac{1-b}{2}\right)\left(\frac{1-c}{2}\right)^{i+1} P_{i-1}^{1,1}(a),\ 1 \leqslant i \leqslant r-1 \right\} \circ \widehat{T}^{-1} =$
$\left\{ \widetilde{x}(1-\widetilde{x})(1-\widetilde{y})(1-\widetilde{z})^{i+1} P_{i-1}^{1,1}(2\widetilde{x}-1),\ 1 \leqslant i \leqslant r-1 \right\} \subset C_r.$

Pour une face triangulaire :
$\left\{ \left(\frac{1-a}{2}\right)\left(\frac{1+a}{2}\right)\left(\frac{1-b}{2}\right)\left(\frac{1-c}{2}\right)^{i+1}\left(\frac{1+c}{2}\right) P_{i-1}^{1,1}(a)P_{j-1}^{2i+1,1}(c),\ i+j \leqslant r-1, i,j \geqslant 1 \right\} \circ \widehat{T}^{-1} =$
$\left\{ (1-\widetilde{x})\widetilde{x}(1-\widetilde{y})(1-\widetilde{z})^{i+1}\widetilde{z}\, P_{i-1}^{1,1}(2\widetilde{x}-1)P_{j-1}^{2i+1,1}(2\widetilde{z}-1),\ i+j \leqslant r-1, i,j \geqslant 1 \right\} \subset C_r.$

Pour la base :
$\left\{ \left(\frac{1-a}{2}\right)\left(\frac{1+a}{2}\right)\left(\frac{1-b}{2}\right)\left(\frac{1+b}{2}\right)\left(\frac{1-c}{2}\right)^{i+j+1} P_{i-1}^{1,1}(a)P_{j-1}^{1,1}(b),\ 1 \leqslant i,j \leqslant r-1 \right\} \circ \widehat{T}^{-1} =$
$\left\{ \widetilde{x}(1-\widetilde{x})\widetilde{y}(1-\widetilde{y})(1-\widetilde{z})^{i+j+1} P_{i-1}^{1,1}(2\widetilde{x}-1)P_{j-1}^{1,1}(2\widetilde{y}-1),\ 1 \leqslant i,j \leqslant r-1 \right\} \not\subset C_r.$

Pour l'intérieur :
$\left\{ \left(\frac{1-a}{2}\right)\left(\frac{1+a}{2}\right)\left(\frac{1-b}{2}\right)\left(\frac{1+b}{2}\right)\left(\frac{1-c}{2}\right)^{i+j+1}\left(\frac{1+c}{2}\right) P_{i-1}^{1,1}(a)P_{j-1}^{1,1}(b)P_{k-1}^{2i+2j+1,1}(c),\ i+j+k \leqslant r-1,\ i,j,k \geqslant 1 \right\} \circ$
$\widehat{T}^{-1} = \left\{ \widetilde{x}(1-\widetilde{x})\widetilde{y}(1-\widetilde{y})\widetilde{z}(1-\widetilde{z})^{i+j+1} P_{i-1}^{1,1}(2\widetilde{x}-1)P_{j-1}^{1,1}(2\widetilde{y}-1)P_{k-1}^{2i+2j+1,1}(2\widetilde{z}-1),\ i+j+k \leqslant r-1,\ i,j,k \geqslant 1 \right\} \subset C_r.$

Le sous espace W_r^0 de W_r de trace nulle au bord de l'élément est

$$W_r^0 \circ \widehat{T}^{-1} = \left\{ \widetilde{x}(1-\widetilde{x})\widetilde{y}(1-\widetilde{y})\widetilde{z}(1-\widetilde{z})^2\, u(\widetilde{x},\widetilde{y},\widetilde{z}),\ u \in \mathbb{P}_{r-3}(\widetilde{x},\widetilde{y},\widetilde{z}) \right\},$$

dont la dimension est

$$\dim W_r^0 = \dim \mathbb{P}_{r-3} = \frac{(r-2)(r-1)r}{6}.$$

Comme il y a $3r^2+2$ fonctions de base associées à la frontière, on a

$$\dim W_r = \frac{(r-2)(r-1)r}{6} + 3r^2 + 2 = \frac{(r+1)(r+2)(r+3)}{6} + r^2 < \dim C_r.$$

En remplaçant les fonctions de base sur la face quadrangulaire par les fonctions suivantes

$$\left\{ \left(\frac{1-a}{2}\right)\left(\frac{1+a}{2}\right)\left(\frac{1-b}{2}\right)\left(\frac{1+b}{2}\right)\left(\frac{1-c}{2}\right)^{max(i,j)+1} P_{i-1}^{1,1}(a)P_{j-1}^{1,1}(b),\ 1 \leqslant i,j \leqslant r-1 \right\} \circ \widehat{T}^{-1} =$$
$$\left\{ \widetilde{x}(1-\widetilde{x})\widetilde{y}(1-\widetilde{y})(1-\widetilde{z})^{max(i,j)+1} P_{i-1}^{1,1}(2\widetilde{x}-1)P_{j-1}^{1,1}(2\widetilde{y}-1)\ 1 \leqslant i,j \leqslant r-1 \right\} \subset C_r,$$

et W_r^0 par $C_r^0 \circ \widehat{T}$, l'espace d'approximation $\widehat{W}_r$ obtenu vérifie

$$\widehat{W}_r \subset C_r \circ \widehat{T}$$

et

$$\dim \widehat{W}_r = \dim C_r,$$

i.e. il y a égalité entre les deux espaces. □

Remarque 5.3.4 *Comme $W_1 = \widehat{W}_1$ mais $W_r \not\supset \widehat{W}_2$, le fait d'utiliser W_r comme espace d'approximation pour les éléments pyramidaux nous assure de n'avoir qu'une convergence d'ordre 1 en norme H^1 pour les pyramides non-affines.*

5.3.2 Comparaison numérique

On trace les courbes de dispersion obtenues avec d'autres espaces d'éléments finis pyramidaux dont il vient d'être question, c'est à dire ceux de Sherwin *et al.* [66], Nigam et Phillips [58], Bluck et Walker [7]. Les résultats sont présentés sur la figure 5.2 pour les ordres 2 et 3. À l'ordre 1, toutes ces méthodes donnent le même ordre de convergence, c'est à dire en $O(h^2)$.

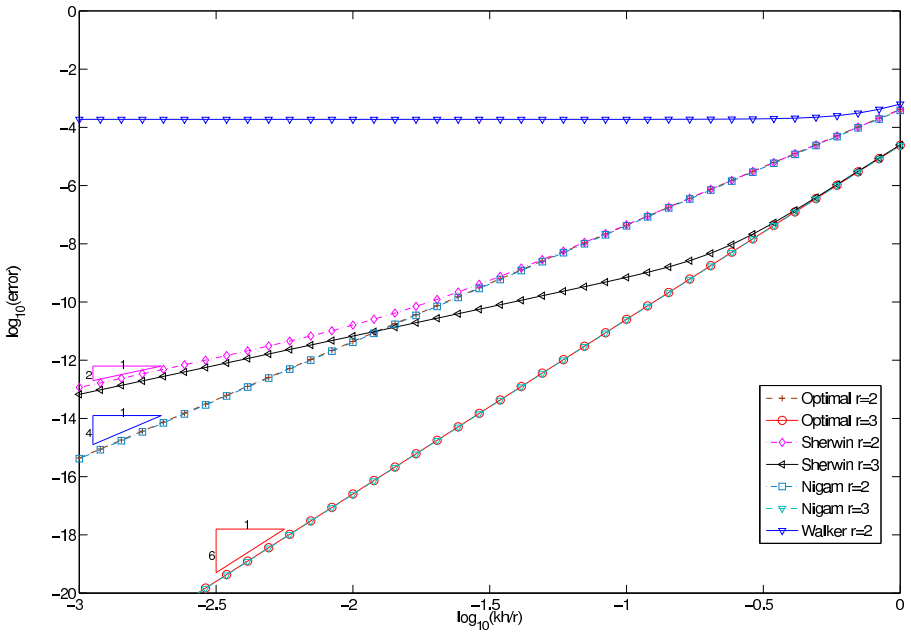

FIG. 5.2 – Erreur de dispersion pour différents types d'éléments aux ordres 2 et 3 sur un maillage hybride déformé

- La dispersion obtenue avec l'espace proposé par Sherwin *et al.* est d'ordre 2 quel que soit l'ordre d'approximation, puisque les fonctions de base sur la base et à l'intérieur de la pyramide ne sont pas suffisantes pour que l'espace d'approximation final contienne l'espace optimal. Dans le cas d'éléments affines, cependant, l'ordre de dispersion est bien $2r$ puisque l'espace d'approximation contient au moins les polynômes.
- La dispersion obtenue avec l'espace optimal est égale à celle obtenue pour le premier espace proposé par Nigam et Phillips, ce qui signifie que les degrés de liberté qu'ils ajoutent ne sont pas nécessaires puisqu'ils n'augmentent pas la précision.
- À l'ordre 2, il est clair que la méthode de Bluck et Walker n'est pas consistante (dispersion d'ordre 0) puisque l'espace qu'ils proposent ne contient pas leur espace d'ordre un. Cependant, la dispersion obtenue dans le cas de pyramides affines est bien en $O(h^4)$ à l'ordre 2.

Remarque 5.3.5 *La mise en oeuvre de la méthode de Hesthaven [43] pour trouver des points « électrostatiques » minimisant un potentiel électrostatique a été envisagée pour les points intérieurs de la pyramide. Cependant, notre objectif principal était d'obtenir des éléments avec une CFL la plus élevée possible : comme la CFL ne dépend que*

de l'espace d'approximation et de la formule de quadrature utilisée, l'étude d'une autre configuration des points intérieurs n'a pas été creusée.

5.4 Comparaison nodal/hiérarchique

5.4.1 Introduction

On considère l'équation de Helmholtz. Après discrétisation, on obtient une matrice de masse M_h et une matrice de rigidité R_h définies par

$$M_h = \int_\Omega \varphi_i \varphi_j \, dx$$

$$K_h = \int_\Omega \nabla\varphi_i \cdot \nabla\varphi_j \, dx$$

où les φ_i sont une base de l'espace d'approximation V_h et peuvent être soit
- les fonctions de base nodales
- les fonctions de base hiérarchiques

Il s'agit de construire la matrice $-\omega^2 M_h + R_h$. Pour cela, on peut faire le calcul naïvement, soit un calcul en $O(r^9)$, ou essayer d'utiliser la structure des fonctions de base pour accélérer les calculs. Pour cela, on va se servir de la décomposition des matrices M_h et R_h issues du produit matrice-vecteur rapide mis au point pas Duruflé [28] pour les hexaèdres.

5.4.2 Efficacité du produit matrice-vecteur

Comme il le sera détaillé le chapitre 7 sur Galerkin discontinu, on montre que l'on a la factorisation suivante des matrices élémentaires M_h et K_h

$$M_h = \hat{C}^* A \hat{C}$$

$$K_h = \hat{S}^* B \hat{S}$$

où A et B sont des matrices respectivement diagonale et diagonale par bloc, chaque bloc étant associé à un point de quadrature

$$A_k = \omega_k \, |DF|(\hat{\xi_k})$$

$$B_k = \omega_k \, (|DF| DF^{-1} \, DF^{*-1})(\hat{\xi_k})$$

Les matrices $\hat{C}$ et $\hat{S}$ sont quant à elles des matrices indépendantes de la géométrie telles que

$$\hat{C}_{i,j} = \hat{\varphi}_j(\xi_i)$$

$$\hat{S}_{i,j} = \nabla\hat{\varphi}_j(\xi_i)$$

L'avantage de cette factorisation est avant tout un gain en stockage, puisque l'on remplace le stockage de la matrice initiale, en $O(r^6 \, n_e)$, par le stockage des matrices A et B, en $O(r^3 \, n_e)$. En outre, en fonction des cas, les matrices $\hat{C}$ et $\hat{S}$ sont plus ou moins pleines.

Dans le cas nodal, on a
- **Hexaèdres :** les fonctions de base sont tensorisées et la formule de quadrature utilisée coïncide avec les points d'interpolation, i.e.

  $$\hat{C} = I$$

  et $\hat{S}$ est creuse. La complexité du produit matrice-vecteur est alors en $O(r^4 n_e)$ et le coût de construction de la matrice est en $O(r^5 n_e)$ au lieu de $O(r^9 n_e)$.
- **Pyramides :** les points de quadrature étant tensorisés, il est possible de calculer les dérivées sur le cube unité. On a donc de manière sous-jacente la factorisation

  $$K_h = \hat{C}^* \hat{R}^* \tilde{B} \hat{R} \hat{C}$$

  avec

  $$\hat{R}_{i,j} = \tilde{\nabla}\tilde{\psi}_j(\tilde{\xi}_i)$$

  où $\tilde{\psi}_j$ sont les polynômes d'interpolation de Lagrange associés aux points ξ_j

  $$\tilde{\psi}_j(\tilde{x}) = \frac{\prod_{n \neq j} \tilde{x} - \xi_n}{\prod_{n \neq j} \xi_j - \xi_n}$$

Néanmoins, la complexité du produit avec $\hat{C}$ maintient la complexité en $O(r^6 n_e)$, ce qui rend l'algorithme plus lent qu'en ayant stocké la matrice. Le coût de construction de la matrice reste en outre en $O(r^9 n_e)$.

- **Prismes :** du fait de la tensorisation dans la direction e_z, la matrices $\hat{C}$ est creuse, ce qui conduit à une complexité en $O(r^5 n_e)$ pour le produit matrice-vecteur, donc un coût en $O(r^8 n_e)$ pour la construction de la matrice.
- **Tétraèdres :** Les matrices $\hat{C}$ et $\hat{S}$ sont pleines, mais dans le cas de tétraèdres droits, les matrices élémentaires sont précalculées, ce qui accélère les calculs. Dans le cas d'éléments courbes, on a une complexité en $O(r^6 n_e)$ pour le produit matrice-vecteur et un coût en $O(r^9 n_e)$ pour la construction de la matrice.

Concernant les fonctions hiérarchiques, les fonctions ont été construites de sorte qu'elles s'écrivent sous forme tensorisées sur le cube après passage par la transformation T pour tous les éléments. Comme il sera détaillé dans le chapitre 7, on a alors une factorisation de la matrice $\hat{C}$

$$\hat{C} = \hat{C}_1 \hat{C}_2 \hat{C}_3$$

où les matrices $\hat{C}_1, \hat{C}_2$ et $\hat{C}_3$ sont creuses. La complexité du produit matrice-vecteur pour les fonctions de base hiérarchiques est alors en $O(r^4 n_e)$ et le coût de construction de la matrice globale est en $O(r^7 n_e)$, ce qui les rend plus attractives pour des ordres élevés.

En faisant la factorisation, on a cependant remplacé un produit-matrice vecteur avec une matrice $-\omega^2 M_h + K_h$ par un produit avec quatre matrices, la matrice $\hat{S}$ étant de taille $3n_r \times n_r$, donc a priori trois fois plus volumineuse que C. Le coût de calcul du produit matrice-vecteur en utilisant cette factorisation est donc a priori au moins huit fois plus lent qu'en ayant stocké la matrice $-\omega^2 M_h + K_h$. En pratique, puisque des valeurs sont ajoutées au moment de l'assemblage, si bien que la matrice globale contient moins d'entrées que la somme des entrées des matrices élémentaires. Ce gain de stockage induit un gain de temps supplémentaire.

Les tableaux 5.1, 5.2, 5.3 et 5.4 présentent les résultats comparatifs obtenus pour les tétraèdres, les pyramides, les prismes et les hexaèdres respectivement. Les temps de calcul obtenus pour 100 itérations du COCG sur un maillage contenant un million de ddls pour des éléments non-affines sont indiqués pour les éléments nodaux et hiérarchiques utilisant la factorisation, et pour la méthode utilisant la matrice non factorisée. La taille de la matrice est indiquée entre parenthèse. Pour un million de ddls, 7 vecteurs représentent 56 Mo (réel double précision), et on remarque bien que le gain de stockage est réalisé pour un ordre supérieur ou égal à 2.

Sur les tableaux 5.1 et 5.2, on observe un temps de calcul environ 15 fois plus lent pour les tétraèdres, et 12 fois plus lent pour les pyramides en utilisant la factorisation. On constate néanmoins qu'il est nettement plus efficace d'utiliser cette factorisation sur les pyramides nodales. Concernant les prismes, ainsi que le montre le tableau 5.3, la factorisation est plus efficace à partir de l'ordre 5.

Pour récapituler sur le temps de calcul, on a comparé les approches suivantes

- **Matrice stockée :** Bien que le coût de stockage soit prohibitif sur des ordres élevés, le temps de calcul reste souvent compétitif
- **Fonctions de base nodales :** elles sont intéressantes pour $r \geq 5$ sur les hexaèdres et les prismes. Il faut préférer stocker la matrice pour les tétraèdres, et pour les pyramides, il est préférable de calculer les dérivées via le cube.
- **Fonctions de base hiérarchiques :** elles fournissent un algorithme rapide en $O(r^4)$, mais en pratique le gain est intéressant pour $r \geq 6$ (sauf pour les tétras).
- Pour les hexaèdres, il est toujours préférable d'utiliser les fonctions de base nodales

On notera aussi que le cas de l'équation de Helmholtz est le plus « pénalisant » car sur d'autres équations où le nombre d'inconnues scalaires n_s est plus élevé, le gain obtenu pour le produit matrice-vecteur en utilisant la factorisation sera plus important. En effet, le stockage de la matrice est proportionnel à n_s^2 alors que les termes prépondérants du coût du produit matrice-vecteur sont proportionnels à n_s.

TAB. 5.1 – Temps de calcul pour 100 itérations du COCG sur un maillage contenant un million de ddls et uniquement des tétraèdres

	r = 2	r = 3	r = 4	r = 5	r = 6	r = 7	r = 8	r = 9	r = 10
Nodal	291s	251s	246s	525s	810s	967s	1281s	2058s	3516s
Hiérarchique	502s	368s	313s	316s	310s	321s	346s	344s	365s
Matrice stockée	18.65s	27.36s	43.02s	53.84s	71.83s	93.2s	119.9s	152s	185.6s

TAB. 5.2 – Temps de calcul pour 100 itérations du COCG sur un maillage contenant un million de ddls et uniquement des pyramides non-affines

	r = 2	r = 3	r = 4	r = 5	r = 6	r = 7	r = 8	r = 9	r = 10
Nodal	327s	388s	499s	725s	1021s	1487s	1918s	2789s	4345s
Nodal/Cube	263s	212s	247s	268s	336s	453s	529s	721s	1120s
Hiérarchique	285.6s	199.9s	182.5s	173.7s	183.7s	202.2s	193.9s	208.3s	238.3s
Matrice stockée	26.2s	37.5s	54.9s	79.3s	112.9s	171.7s	233.9s	274.9s	358.8s
	(276 Mo)	(487 Mo)	(781 Mo)	(1175 Mo)	(1684 Mo)	(2314 Mo)	(3086Mo)	(4025Mo)	(5131Mo)

TAB. 5.3 – Temps de calcul pour 100 itérations du COCG sur un maillage contenant un million de ddls et uniquement des prismes non-affines

	r = 2	r = 3	r = 4	r = 5	r = 6	r = 7	r = 8	r = 9	r = 10
Nodal	171.9s	141.4s	108.3s	116.7s	124s	198.3s	143s	173.2s	193s
Hiérarchique	203.4s	140.2s	128s	115.2s	116.6s	120.5s	120.1s	131.9s	129s
Matrice stockée	29.86s	44.6s	70.9s	97.9s	138.8s	187.3s	266.8s	331.4s	448.5s
	(327Mo)	(593 Mo)	(970 Mo)	(1480 Mo)	(2149 Mo)	(2971 Mo)	(3985 Mo)	(5239 Mo)	(6716 Mo)

TAB. 5.4 – Temps de calcul pour 100 itérations du COCG sur un maillage contenant un million de ddls et uniquement des hexaèdres non-affines

	r = 2	r = 3	r = 4	r = 5	r = 6	r = 7	r = 8	r = 9	r = 10
Nodal	77.3s	53.7s	48.7s	42.5s	44.6s	41.8s	42.3s	43.3s	46s
Hiérarchique	98.5s	73s	63.9s	61.5s	61.8s	62.7s	77s	66.2s	67.9s
Matrice stockée	22.2s	32.3s	45.4s	60.9s	78.6s	98.5s	119.7s	147.6s	171.1s
	(266 Mo)	(431 Mo)	(636 Mo)	(881 Mo)	(1170 Mo)	(1490 Mo)	(1852 Mo)	(2266 Mo)	(2717 Mo)

5.4.3 Conditionnement des matrices

Une problématique concerne le conditionnement des matrices, afin d'avoir des algorithmes itératifs performants. Ce problème n'est bien sûr pas crucial dans le sens où il est en général nécessaire d'utiliser un préconditionneur afin d'obtenir des performances raisonnables et, souvent, atténuer fortement le mauvais conditionnement initial.

Sur la figure 5.3, on montre le conditionnement de la matrice de masse pour les différents éléments avec les fonctions de base nodales et hiérarchiques. On remarque que les fonctions de base nodales fournissent un conditionnement nettement meilleur que les fonctions de base hiérarchiques pour les matrices de masse et de rigidité.

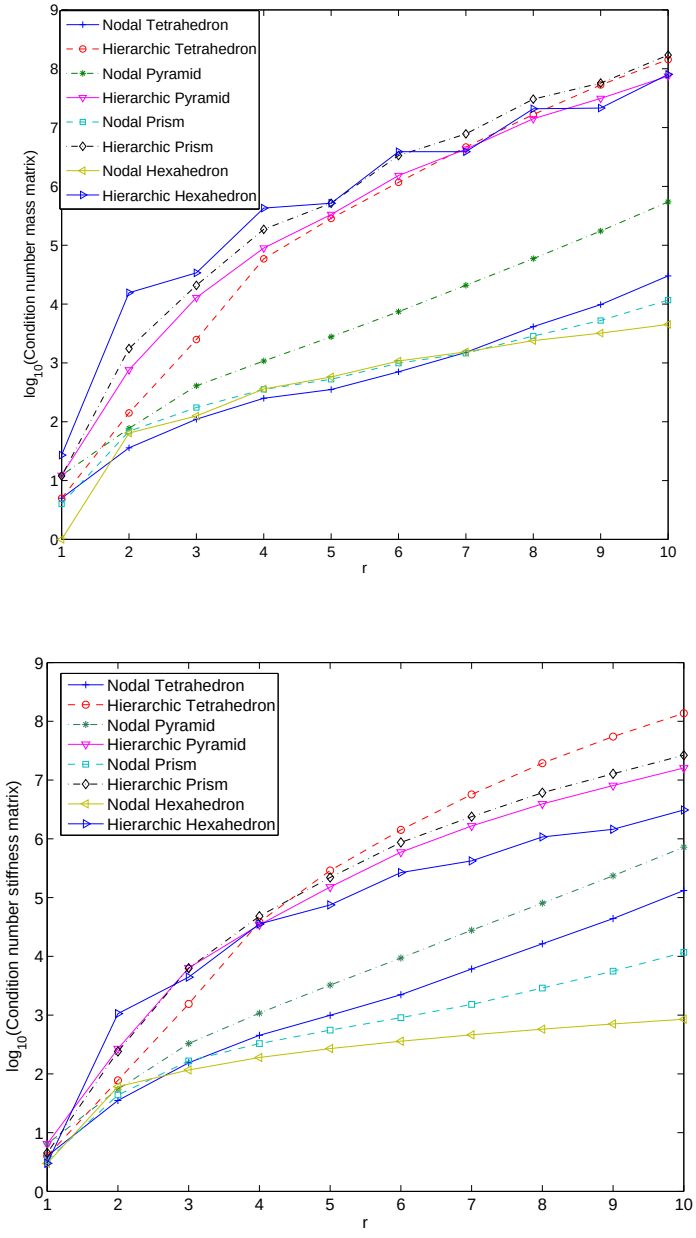

FIG. 5.3 – Conditionnement en échelle logarithmique pour la matrice de masse (en haut) et la matrice de rigidité (en bas)

On étudie le cas de la diffraction d'une sphère de rayon 3 (voir section 14.1, figure 14.1) et on compare le nombre d'itérations nécessaire pour avoir un résidu relatif inférieur à 10^{-6} pour une maillage hybride. Le tableau 5.5 résume les résultats observés. On observe sans surprise que le nombre d'itérations augmente avec l'ordre, et qu'il est beaucoup plus élevé pour les fonctions de base hiérarchiques. Ce n'est cependant pas en soit problématique, puisque souvent ces disparités peuvent être gommées par le préconditionneur employé.

TAB. 5.5 – Nombre d'itérations pour BICGCR sans préconditionnement pour la diffraction d'une sphère. Maillage hybride contenant 500 000 degrés de liberté

	r = 2	r= 3	r=4	r=5	r=6	r=7
Nodal	1 162	2 000	3 078	5 676	10 076	17 952
Hiérarchique	2 144	5 939	17 290	50 742	> 107 620	> 200 000

Troisième partie

Éléments finis orthogonaux pour une formulation discontinue

Chapitre 6

Éléments finis orthogonaux d'ordre arbitrairement élevé

Dans cette partie, nous nous intéressons à des éléments adaptés aux méthodes de Galerkin discontinues. Nous définissons tout d'abord des éléments permettant d'obtenir une matrice de masse la plus creuse possible. Nous présentons ensuite un procédé utilisant les propriétés des fonctions de base des éléments et permettant de diminuer le coût de calcul de la matrice de masse en réduisant le nombre d'évaluations d'intégrales par formules de quadrature.

Sommaire

6.1 Problématique

L'espace d'approximation V_h d'ordre r est défini par

$$V_h = \left\{ u \in (L^2(\Omega))^{n_s} \text{ tel que } u_{|K} \circ F \in \hat{P}_r(\hat{x}, \hat{y}, \hat{z}) \right\}$$

où F est la transformation permettant de passer d'un élément K du maillage à un élément de référence $\hat{K}$ et les espaces $\hat{P}_r$ sont les espaces d'approximations sur $\hat{K}$.

Comme dans le cas H^1, on prend la transformation F et les éléments de référence $\hat{K}$ donnés par la définition 2.1.1. Pour les mêmes raisons que pour H^1, on a besoin d'avoir $\mathbb{P}_r \subset P_r^F$ pour avoir des estimations d'erreur en $O(h^r)$ en norme L^2 : on choisit donc de prendre les $\hat{P}_r$ comme donnés par le théorème 2.2.3.

Remarque 6.1.1 *Comme le rappelle Hartmann [42], si l'adjoint est consistant, ce qui est le cas si le système est symétrique, on a même une convergence en $O(h^{r+1})$ en norme L^2.*

Puisque la CFL ne dépend pas du choix de la base de l'espace d'approximation (voir Cohen [17]), on peut considérer plusieurs types de bases de $\hat{P}_r$ pour les éléments. Par un choix judicieux des fonctions de base, on souhaite minimiser le stockage, le coût de calcul des matrices de masse élémentaires et du produit matrice-vecteur. Les fonctions nodales ont cependant l'avantage de restreindre l'évaluation des intégrales de bord aux degrés de liberté associés à la frontière des éléments, mais la matrice de masse obtenue avec ces fonctions est pleine, tout comme celle obtenue avec des fonctions de base monomiales. La structure des polynômes orthogonaux utilisés par Kirby *et al.* [48] et des fonctions de la proposition 2.3.3 permet de creuser la matrice de masse, qui est même égale à l'identité dans le cas d'éléments affines, et l'utilisation de fonctions de base tensorisées, ou même semi-tensorisées induit également un produit-matrice vecteur rapide. On va donc proposer des fonctions orthogonales semi-tensorisées creusant au maximum la matrice de masse.

6.2 Fonctions de base orthogonales

6.2.1 Base pour éléments non affines

On rappelle que, d'après le théorème 2.2.3, l'espace d'approximation optimal est différent selon que l'élément est affine ou non. Dans un premier temps, on traite le cas des éléments non-affines. On définit une base orthogonale de $\hat{P}_r$ dans l'esprit de celle de la proposition 2.3.3 pour les éléments continus, mais adaptée aux éléments discontinus.

Proposition 6.2.1 *Les fonctions de base suivantes forment une base orthogonale de l'espace $\hat{P}_r$ dans le cas d'éléments non-affines*

- ***Hexaèdre***

$$\hat{\varphi}^G_{i_1}(\hat{x})\,\hat{\varphi}^G_{i_2}(\hat{y})\,\hat{\varphi}^G_{i_3}(\hat{z}), \quad 0 \leqslant i_1, i_2, i_3 \leqslant r,$$

  où

$$\hat{\varphi}^G_i(\hat{x}) = \frac{\prod_{j \neq i} \hat{x} - \xi^G_j}{\prod_{j \neq i}, \xi^G_i - \xi^G_j},$$

- ***Prisme***

$$P^{0,0}_{i_1}\left(\frac{2\hat{x}}{1-\hat{y}} - 1\right)(1-\hat{y})^{i_1} P^{2i_1+1,0}_{i_2}(2\hat{y}-1)\,\varphi^G_{i_3}(\hat{z}), \quad 0 \leqslant i_1 + i_2, i_3 \leqslant r,$$

- ***Pyramide***

$$P^{0,0}_{i_1}\left(\frac{\hat{x}}{1-\hat{z}}\right) P^{0,0}_{i_2}\left(\frac{\hat{y}}{1-\hat{z}}\right)(1-\hat{z})^{max(i_1,i_2)} P^{2max(i_1,i_2)+2,0}_{i_3}(2\hat{z}-1),$$
$$0 \leqslant i_1, i_2 \leqslant r,\ 0 \leqslant i_3 \leqslant r - max(i_1, i_2),$$

Les $P^{i,j}_m(x)$ sont les polynômes de Jacobi orthonormalisés d'ordre m, orthogonaux pour les poids $(1-x)^i(1+x)^j$, et les ξ^G_j sont les points de Gauss-Legendre sur $[0,1]$ (cf Hammer, Marlowe et Stroud [41]).

Preuve. La preuve est similaire à celle de la proposition 2.3.3.

Remarque 6.2.2 *La différence avec la base orthogonale de la proposition 2.3.3 est que l'on utilise les fonctions d'interpolation de Lagrange avec points de Gauss, et non plus de Gauss-Lobatto sur les hexaèdres et les prismes.*

6.2.2 Base pour éléments affines

D'après le théorème 2.2.3), lorsque les éléments sont affines, on peut utiliser $\hat{P}_r = \mathbb{P}_r$.

Proposition 6.2.3 *Les fonctions de base suivantes forment une base orthogonales de l'espace* $\mathbb{P}_r$

$$\textit{Pour } 0 \leqslant i_1 + i_2 + i_3 \leq r,$$

- ***Hexaèdre***

$$P_{i_1}^{0,0}(2\hat{x}-1)P_{i_2}^{0,0}(2\hat{y}-1)P_{i_3}^{0,0}(2\hat{z}-1),$$

- ***Prisme***

$$P_{i_1}^{0,0}\left(\frac{2\hat{x}}{1-\hat{y}}-1\right)(1-\hat{y})^{i_1}P_{i_2}^{2i_1+1,0}(2\hat{y}-1)\,P_{i_3}^{0,0}(2\hat{z}-1),$$

- ***Pyramide***

$$P_{i_1}^{0,0}\left(\frac{\hat{x}}{1-\hat{z}}\right)P_{i_2}^{0,0}\left(\frac{\hat{y}}{1-\hat{z}}\right)(1-\hat{z})^{i_1+i_2}P_{i_3}^{2(i_1+i_2)+2,0}(2\hat{z}-1),$$

- ***Tétraèdre***

$$P_{i_1}^{0,0}\left(\frac{2\hat{x}}{1-\hat{y}-\hat{z}}-1\right)P_{i_2}^{2i_1+1,0}\left(\frac{2\hat{y}}{1-\hat{z}}-1\right)(1-\hat{y}-\hat{z})^{i_1}P_{i_3}^{2(i_1+i_2)+2,0}(2\hat{z}-1)\,(1-\hat{z})^{i_2}.$$

Preuve. La preuve est similaire à celle de la proposition 2.3.3.

6.3 Construction de la matrice de masse

6.3.1 Hexaèdres et éléments affines

Dans le cas des hexaèdres, de par la structure des fonctions de base, on a condensation de masse, donc une matrice de masse diagonale. Concernant les tétraèdres et les éléments affines, le jacobien étant constant, il est clair qu'en utilisant les fonctions orthogonales de la proposition 6.2.3, la matrice est également diagonale.

6.3.2 Algorithme rapide pour les pyramides

On utilise les fonctions de base de la proposition 6.2.1. Pour faciliter les calculs, on écrit les intégrales sur le cube unité $\widetilde{Q}$ grâce à la transformation T définie par l'équation 2.1.2. On rappelle que la matrice de masse s'écrit alors (cf équation 2.2.9)

$$(M_h)_{i,j} = 4\int_{\widetilde{Q}} M\,\widetilde{|DF|}\,\widetilde{\varphi}_i\,\widetilde{\varphi}_j\,(1-\widetilde{z})^2\,d\widetilde{x}\,d\widetilde{y}\,d\widetilde{z},$$

On traitera ici le cas où M est constante. Pour simplifier les calculs, on prendra $M = I$.

En rappelle en outre que, d'après le lemme 3.1.1, $\widetilde{|DF|}$ peut s'écrire

$$\widetilde{|DF|} = A + B_1(2\widetilde{x}-1) + B_2(2\widetilde{y}-1) + C(2\widetilde{x}-1)(2\widetilde{y}-1).$$

On rappelle la propriété des polynômes de Jacobi suivante (voir Szegö [69])

Propriété 6.3.1

$$tP_k^{i,j}(t) = \gamma_k^{i,j}P_{k+1}^{i,j}(t) + \alpha_k^{i,j}P_k^{i,j}(t) + \beta_k^{i,j}P_{k-1}^{i,j}(t), \tag{6.3.1}$$

où

$$\alpha_k^{i,j} = -a_{i,j,k}\frac{(2k+i+j)(i^2-j^2)}{2k+i+j},$$

$$\beta_k^{i,j} = b_{i,j,k}\frac{2(k+i)(k+j)(2k+i+j+2)}{2k+i+j},$$

$$\gamma_k^{i,j} = c_{i,j,k}2(k+1)(k+i+j+1),$$

et où $a_{i,j,k}$, $b_{i,j,k}$ *et* $c_{i,j,k}$ *sont les coefficients d'orthonormalisation des polynômes de Jacobi* $P_k^{i,j}$.

En utilisant la propriété 6.3.1, on décompose alors la matrice de masse comme suit, pour tout $i = (i_1, i_2, i_3)$, $1 \leqslant i \leqslant n_r$ et pour tout j_3, $0 \leqslant j_3 \leqslant r$ dépendants de la numérotation

– Terme en A

$$4\int_{\widetilde{Q}}\widetilde{\varphi}_i\widetilde{\varphi}_j\,(1-\widetilde{z})^2\,d\widetilde{x}\,d\widetilde{y}\,d\widetilde{z} = \underbrace{\int_0^1 P_{i_1}^{0,0}(2\widetilde{x}-1)P_{j_1}^{0,0}(2\widetilde{x}-1)\,d\widetilde{x}}_{\delta_{i_1j_1}} \underbrace{\int_0^1 P_{i_2}^{0,0}(2\widetilde{y}-1)P_{j_2}^{0,0}(2\widetilde{y}-1)\,d\widetilde{y}}_{\delta_{i_2j_2}} \underbrace{4\int_0^1(1-\hat{z})^{max(i_1,i_2)+max(j_1,j_2)+2}P_{i_3}^{2max(i_1,i_2)+2,0}(2\widetilde{z}-1)P_{j_3}^{2max(j_1,j_2)+2,0}(2\widetilde{z}-1)\,d\widetilde{z}}_{\delta_{i_3j_3}};$$

– Terme en B_1

$$4\int_{\widetilde{Q}}\widetilde{\varphi}_i\widetilde{\varphi}_j\,(2\widetilde{x}-1)(1-\widetilde{z})^2\,d\widetilde{x}\,d\widetilde{y}\,d\widetilde{z} = \underbrace{\int_0^1 (2\widetilde{x}-1)P_{i_1}^{0,0}(2\widetilde{x}-1)P_{j_1}^{0,0}(2\widetilde{x}-1)\,d\widetilde{x}}_{\gamma_{i_1}^{0,0}\delta_{i_1+1j_1}+\beta_{i_1}^{0,0}\delta_{i_1-1j_1}} \underbrace{\int_0^1 P_{i_2}^{0,0}(2\widetilde{y}-1)P_{j_2}^{0,0}(2\widetilde{y}-1)\,d\widetilde{y}}_{\delta_{i_2j_2}} \underbrace{4\int_0^1(1-\hat{z})^{max(i_1,i_2)+max(j_1,j_2)+2}P_{i_3}^{2max(i_1,i_2)+2,0}(2\widetilde{z}-1)P_{j_3}^{2max(j_1,j_2)+2,0}(2\widetilde{z}-1)\,d\widetilde{z}}_{\mathfrak{C}_{i_1,j_1}^{i_2,j_2}(i_3,j_3)};$$

– Terme en B_2

$$4\int_{\widetilde{Q}}\widetilde{\varphi}_i\widetilde{\varphi}_j\,(2\widetilde{y}-1)(1-\widetilde{z})^2\,d\widetilde{x}\,d\widetilde{y}\,d\widetilde{z} = \underbrace{\int_0^1 P_{i_1}^{0,0}(2\widetilde{x}-1)P_{j_1}^{0,0}(2\widetilde{x}-1)\,d\widetilde{x}}_{\delta_{i_1j_1}} \underbrace{\int_0^1 (2\widetilde{y}-1)P_{i_2}^{0,0}(2\widetilde{y}-1)P_{j_2}^{0,0}(2\widetilde{y}-1)\,d\widetilde{y}}_{\gamma_{i_2}^{0,0}\delta_{i_2+1j_2}+\beta_{i_2}^{0,0}\delta_{i_2-1j_2}} \underbrace{4\int_0^1(1-\hat{z})^{max(i_1,i_2)+max(j_1,j_2)+2}P_{i_3}^{2max(i_1,i_2)+2,0}(2\widetilde{z}-1)P_{j_3}^{2max(j_1,j_2)+2,0}(2\widetilde{z}-1)\,d\widetilde{z}}_{\mathfrak{C}_{i_1,j_1}^{i_2,j_2}(i_3,j_3)};$$

– Terme en C

$$4\int_{\widetilde{Q}}\widetilde{\varphi}_i\widetilde{\varphi}_j\,(2\widetilde{x}-1)(2\widetilde{y}-1)(1-\widetilde{z})^2\,d\widetilde{x}\,d\widetilde{y}\,d\widetilde{z} = \underbrace{\int_0^1 (2\widetilde{x}-1)P_{i_1}^{0,0}(2\widetilde{x}-1)P_{j_1}^{0,0}(2\widetilde{x}-1)\,d\widetilde{x}}_{\gamma_{i_1}^{0,0}\delta_{i_1+1j_1}+\beta_{i_1}^{0,0}\delta_{i_1-1j_1}} \underbrace{\int_0^1 (2\widetilde{y}-1)P_{i_2}^{0,0}(2\widetilde{y}-1)P_{j_2}^{0,0}(2\widetilde{y}-1)\,d\widetilde{y}}_{\gamma_{i_2}^{0,0}\delta_{i_2+1j_2}+\beta_{i_2}^{0,0}\delta_{i_2-1j_2}} \underbrace{4\int_0^1(1-\hat{z})^{max(i_1,i_2)+max(j_1,j_2)+2}P_{i_3}^{2max(i_1,i_2)+2,0}(2\widetilde{z}-1)P_{j_3}^{2max(j_1,j_2)+2,0}(2\widetilde{z}-1)\,d\widetilde{z}}_{\mathfrak{C}_{i_1,j_1}^{i_2,j_2}(i_3,j_3)}.$$

La matrice de masse peut donc se calculer comme suit, pour tout $i = (i_1, i_2, i_3)$, $1 \leqslant i \leqslant n_r$ et pour tout j_3, $0 \leqslant j_3 \leqslant r$

- $M_h[i,i] = 4A$
- $M_h[i,(i_1+1,i_2,j_3)] = 4B_1\,\gamma_{i_1}^{0,0}\,\mathcal{C}_{i_1,i_1+1}^{i_2,i_2}(i_3,j_3)$
- $M_h[i,(i_1-1,i_2,j_3)] = 4B_1\,\beta_{i_1}^{0,0}\,\mathcal{C}_{i_1,i_1-1}^{i_2,i_2}(i_3,j_3)$
- $M_h[i,(i_1,i_2+1,j_3)] = 4B_2\,\gamma_{i_2}^{0,0}\mathcal{C}_{i_1,i_1}^{i_2,i_2+1}(i_3,j_3)$
- $M_h[i,(i_1,i_2-1,j_3)] = 4B_2\,\beta_{i_2}^{0,0}\mathcal{C}_{i_1,i_1}^{i_2,i_2-1}(i_3,j_3)$
- $M_h[i,(i_1+1,i_2+1,j_3)] = 4C\,\gamma_{i_1}^{0,0}\,\mathcal{C}_{i_1,i_1+1}^{i_2,i_2}(i_3,j_3)\,\gamma_{i_2}^{0,0}\,\mathcal{C}_{i_1,i_1}^{i_2,i_2+1}(i_3,j_3)$
- $M_h[i,(i_1+1,i_2-1,j_3)] = 4C\,\gamma_{i_1}^{0,0}\,\mathcal{C}_{i_1,i_1+1}^{i_2,i_2}(i_3,j_3)\,\beta_{i_2}^{0,0}\,\mathcal{C}_{i_1,i_1}^{i_2,i_2-1}(i_3,j_3)$
- $M_h[i,(i_1-1,i_2+1,j_3)] = 4C\,\beta_{i_1}^{0,0}\,\mathcal{C}_{i_1,i_1-1}^{i_2,i_2}(i_3,j_3)\,\gamma_{i_2}^{0,0}\,\mathcal{C}_{i_1,i_1}^{i_2,i_2+1}(i_3,j_3)$
- $M_h[i,(i_1-1,i_2-1,j_3)] = 4C\,\beta_{i_1}^{0,0}\,\mathcal{C}_{i_1,i_1-1}^{i_2,i_2}(i_3,j_3)\,\beta_{i_2}^{0,0}\,\mathcal{C}_{i_1,i_1}^{i_2,i_2-1}(i_3,j_3)$

La matrice de masse contient donc $O(r^4)$ valeurs non nulles au lieu de $O(r^6)$ dans le cas d'une base nodale, et les intégrales $\mathcal{C}_{i_1,j_1}^{i_2,j_2}(i_3,j_3)$ peuvent être précalculées. Le coût de calcul de la matrice est donc en $O(r^4)$.

Pour inverser la matrice, on utilise une factorisation de Cholesky LL^*. Pour améliorer le profil de la matrice de masse, et ainsi diminuer le stockage de la factorisation LL^*, on peut utiliser un algorithme de renumérotation. Par exemple, à partir de notre premier choix de numérotation, l'algorithme Symmetric Approximate Minimum Degree permutations (symamd) développé par Amestoy *et al.* [1] nous donne les résultats de la figure 6.1. Dans ce cas, pour l'ordre 2, le profil de la matrice est de 21% plus petit, 34% à l'ordre 3, 54% à l'ordre 5 et 70% à l'ordre 8.

6.3.3 Algorithme rapide pour les prismes

En ce qui concerne les prismes, d'après l'équation 2.2.8, la matrice de masse sur le cube unité $\widetilde{Q}$ s'écrit

$$(M_h)_{i,j} = \int_{\widetilde{Q}} (1-\widetilde{y})\,\widetilde{|DF|}\widetilde{\varphi}_i\widetilde{\varphi}_j\,d\widetilde{x}d\widetilde{y}d\widetilde{z}$$

et, d'après le lemme 3.1.1, le jacobien s'écrit sous la forme

$$\widetilde{|DF|}(\widetilde{x},\widetilde{y},\widetilde{z}) = (A + B_3\widetilde{z} + D\widetilde{z}^2) + (B_1 + C_1\widetilde{z})\,\widetilde{x}(1-\widetilde{y}) + (B_2 + C_2\widetilde{z})\,\widetilde{y}.$$

Comme pour la pyramide, on utilise les bases orthogonales de la proposition 6.2.1 et la propriété 6.3.1 pour décomposer le calcul de la matrice de masse selon les termes du jacobien. Ainsi, pour tout $i = (i_1,i_2,i_3)$ et $j = (j_1,j_2,j_3)$ dépendants de la numérotation, on a

- Terme en $A + B_3\widetilde{z} + D\widetilde{z}^2$:

$$\int_{\widetilde{Q}} (A + B_3\widetilde{z} + D\widetilde{z}^2)\,\widetilde{\varphi}_i\widetilde{\varphi}_j\,(1-\widetilde{y})\,d\widetilde{x}d\widetilde{y}d\widetilde{z} = \underbrace{\int_0^1 P_{i_1}^{0,0}(2\widetilde{x}-1)P_{j_1}^{0,0}(2\widetilde{x}-1)\,d\widetilde{x}}_{\delta_{i_1j_1}} \underbrace{\int_0^1 (1-\widetilde{y})^{i_1+j_1+1}P_{i_2}^{2i_1+1,0}(2\widetilde{y}-1)P_{j_2}^{2j_1+1,0}(2\widetilde{y}-1)\,d\widetilde{y}}_{\delta_{i_2j_2}} \underbrace{\int_0^1 (A + B_3\widetilde{z} + D\widetilde{z}^2)\,\varphi_{i_3}^G(\widetilde{z})\varphi_{j_3}^G(\widetilde{z})\,d\widetilde{z}}_{\omega_{i_3}(A+B_3\xi_{i_3}+D\xi_{i_3}^2)\delta_{i_3j_3}\text{ avec une formule de quadrature de Gauss}}$$

- Terme en $B_1 + C_1\widetilde{z}$:

$$\int_{\widetilde{Q}} (B_1 + C_1\widetilde{z})\,\widetilde{x}\,(1-\widetilde{y})\,\widetilde{\varphi}_i\widetilde{\varphi}_j\,d\widetilde{x}d\widetilde{y}d\widetilde{z} = \underbrace{\int_0^1 \widetilde{x}\,P_{i_1}^{0,0}(2\widetilde{x}-1)P_{j_1}^{0,0}(2\widetilde{x}-1)\,d\widetilde{x}}_{\frac{1}{2}\left(\gamma_{i_1}^{0,0}\delta_{i_1+1j_1} + \delta_{i_1j_1} + \beta_{i_1}^{0,0}\delta_{i_1-1j_1}\right)} \underbrace{\int_0^1 (1-\widetilde{y})^{i_1+j_1+2}P_{i_2}^{2i_1+1,0}(2\widetilde{y}-1)P_{j_2}^{2j_1+1,0}(2\widetilde{y}-1)\,d\widetilde{y}}_{\mathcal{C}_{i_1,j_1}^{i_2,j_2}} \underbrace{\int_0^1 (B_1 + C_1\widetilde{z})\,\varphi_{i_3}^G(\widetilde{z})\varphi_{j_3}^G(\widetilde{z})\,d\widetilde{z}}_{\omega_{i_3}(B_1+C_1\xi_{i_3})\delta_{i_3j_3}\text{ avec une formule de quadrature de Gauss}}$$

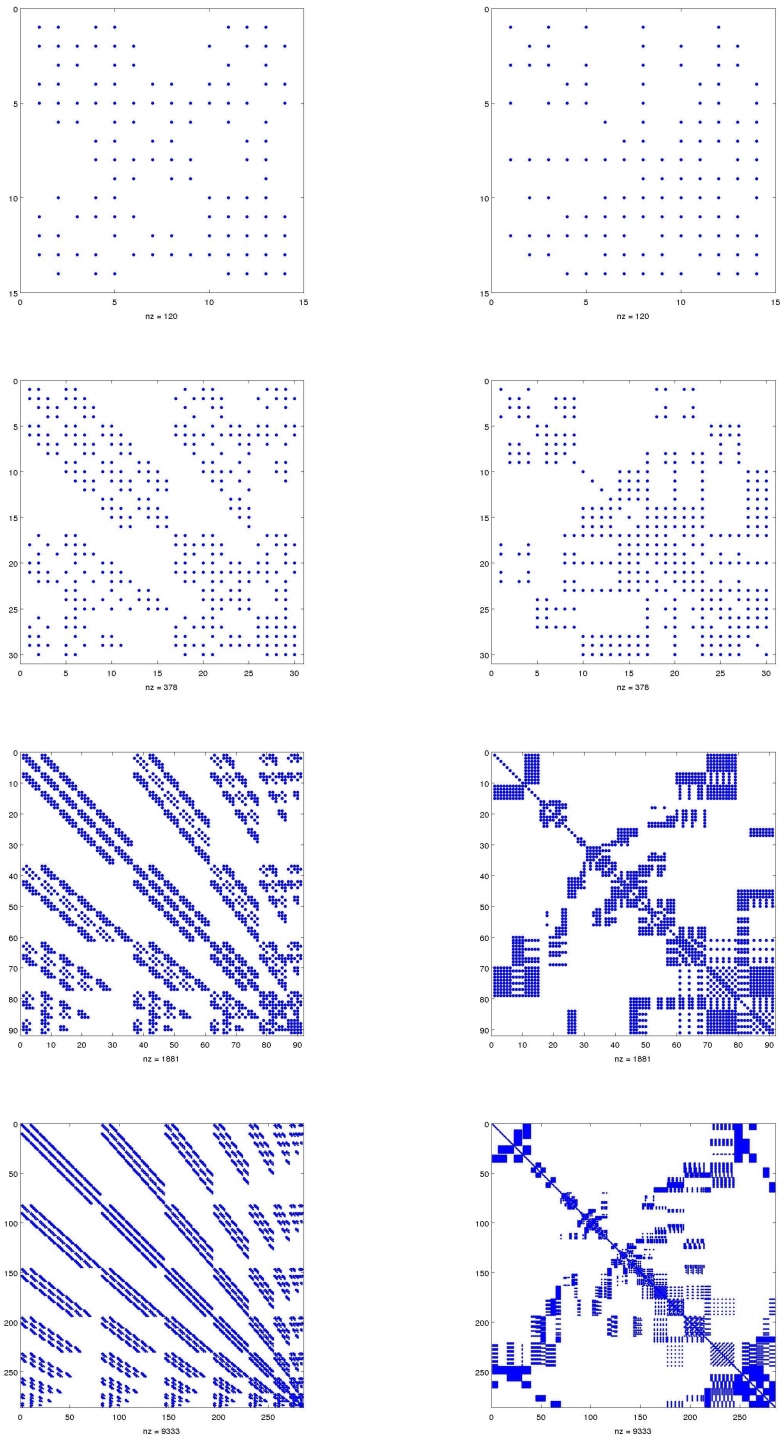

FIG. 6.1 – Profil de la matrice de masse avec des éléments pyramidaux d'ordre 2, 3, 5 et 8 avant (gauche) et après (droite) renumérotation par un algorithme symamd

– Terme en $B_2 + C_2\widetilde{z}$:

$$\int_{\widetilde{Q}} (B_2 + C_2\widetilde{z})\,\widetilde{y}\,(1-\widetilde{y})\,\widetilde{\varphi}_i\widetilde{\varphi}_j\,d\widetilde{x}d\widetilde{y}d\widetilde{z} = \underbrace{\int_0^1 P_{i_1}^{0,0}(2\widetilde{x}-1)P_{j_1}^{0,0}(2\widetilde{x}-1)\,d\widetilde{x}}_{\delta_{i_1 j_1}} \underbrace{\int_0^1 \widetilde{y}\,(1-\widetilde{y})^{i_1+j_1+1}\,P_{i_2}^{2i_1+1,0}(2\widetilde{y}-1)P_{j_2}^{2j_1+1,0}(2\widetilde{y}-1)\,d\widetilde{y}}_{\frac{1}{2}\left(\gamma_{i_2}^{2i_1+1,0}\delta_{i_2+1 j_2} + (\alpha_{i_2}^{2i_1+1,0}+1)\delta_{i_2 j_2} + \beta_{i_2}^{2i_1+1,0}\delta_{i_2-1 j_2}\right)} \underbrace{\int_0^1 (B_2 + C_2\widetilde{z})\,\varphi_{i_3}^G(\widetilde{z})\varphi_{j_3}^G(\widetilde{z})\,d\widetilde{z}}_{\omega_{i_3}(B_2+C_2\xi_{i_3})\delta_{i_3 j_3} \text{ avec une formule de quadrature de Gauss}}$$

La matrice de masse se remplit alors comme suit, pour tout $i = (i_1, i_2, i_3)$

– $M[i,i] = \omega_{i_3}\left(A + \dfrac{(B_2+\xi_{i_3}C_2)}{2}(\alpha_{i_2}^{2i_1+1,0}+1) + \xi_{i_3}B_3 + \xi_{i_3}^2 D\right)$

– $M[i,(i_1,i_2+1,i_3)] = \omega_{i_3}\dfrac{(B_2+\xi_{i_3}C_2)}{2}\gamma_{i_2}^{2i_1+1,0}$

– $M[i,(i_1,i_2-1,i_3)] = \omega_{i_3}\dfrac{(B_2+\xi_{i_3}C_2)}{2}\beta_{i_2}^{2i_1+1,0}$

et pour tout j_2

– $M[i,(i_1,j_2,i_3)] = \omega_{i_3}\dfrac{(B_1+\xi_{i_3}C_1)}{2}\mathcal{C}_{i_1,i_1}^{i_2,j_2}(i_3,i_3)$

– $M[i,(i_1+1,j_2,i_3)] = \omega_{i_3}\dfrac{(B_1+\xi_{i_3}C_1)}{2}\gamma_{i_1}^{0,0}\mathcal{C}_{i_1,i_1+1}^{i_2,j_2}(i_3,i_3)$

– $M[i,(i_1-1,j_2,i_3)] = \omega_{i_3}\dfrac{(B_1+\xi_{i_3}C_1)}{2}\beta_{i_1}^{0,0}\mathcal{C}_{i_1,i_1-1}^{i_2,j_2}(i_3,i_3)$

Comme pour la pyramide, les intégrales 1D peuvent être précalculées, si bien que le coût final de la construction de la matrice de masse est en $O(r^4)$.

Du fait de la tensorisation et de la condensation de masse en z, la matrice de masse est diagonale par blocs pour chaque élément, y compris dans le cas d'éléments nodaux pour lesquels le nombre de valeurs non nulles dans ma matrice est en $O(r^5)$. En utilisant les fonctions de base orthogonales, chaque bloc peut devenir plus creux, mais le gain est moins spectaculaire que pour les éléments pyramidaux.

Étant donné la structure diagonale par blocs de la matrice de masse, une renumérotation n'est pas nécessaire pour réduire le profil de la matrice. En effet, à partir de notre premier choix de numérotation, l'algorithme symamd nous donne les résultats de la figure 6.2, et pour l'ordre 3, le profil de la matrice est de 3% plus petit, 6% à l'ordre 5, 9% à l'ordre 6 et 12% à l'ordre 8.

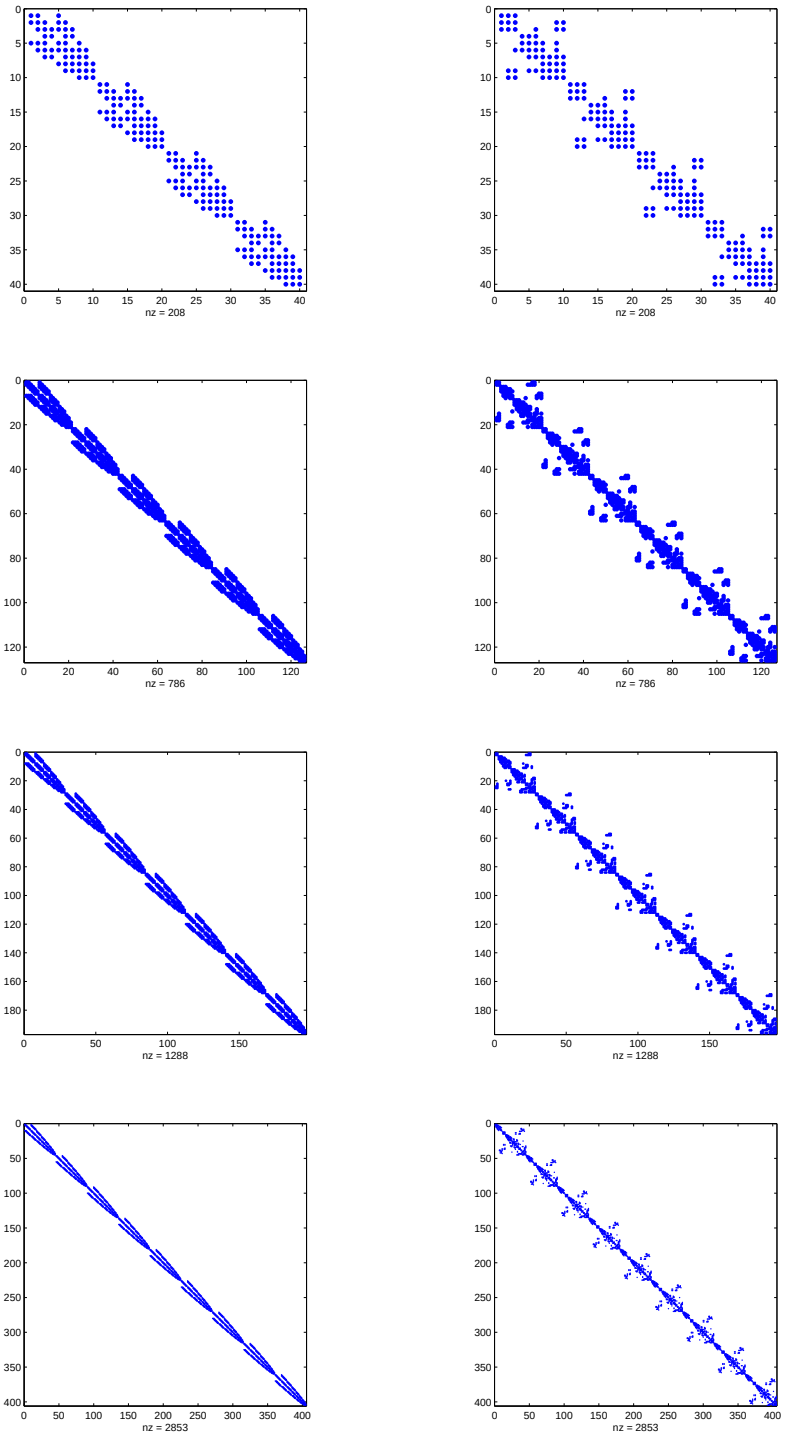

FIG. 6.2 – Profil de la matrice de masse avec des éléments prismatiques d'ordre 3, 5, 6 et 8 avant (gauche) et après (droite) renumérotation par un algorithme symamd

Chapitre 7

Produit matrice-vecteur rapide

On détaille ici un algorithme de produit matrice-vecteur rapide adapté à la structure tensorisée des fonctions de base utilisées pour la formulation de Galerkin discontinue.

Sommaire

7.1 Introduction

Dans la discrétisation en temps (voir section 1.2.3 dans le chapitre 1), l'une des étapes consiste à effectuer le calcul du produit matrice-vecteur suivant

$$y^n = (R_h - S_h)\, U^n$$

Stocker la matrice creuse $R_h - S_h$ et calculer le produit matrice-vecteur standard serait une solution onéreuse car la quantité de mémoire nécessaire au stockage d'une si grosse matrice peut être très importante, surtout lorsque l'on utilise de l'ordre élevé.

Puisqu'à l'intérieur de chaque élément (excepté l'hexaèdre), les degrés de liberté interagissent entre eux, c'est à dire

$$\int_K \frac{\partial \varphi_k}{\partial x_i} \varphi_j \neq 0, \quad \forall j,k,$$

le nombre de valeurs non nulles de la matrice $R_h - S_h$ est en $O(n_e r^6)$, où n_e est le nombre d'éléments du maillage, et r l'ordre d'approximation. Le temps de calcul nécessaire au produit matrice-vecteur serait donc également en $O(n_e r^6)$ si la matrice $R_h - S_h$ était stockée.

Une autre solution, bien connue pour les tétraèdres (Hesthaven [46]) consiste à calculer le produit matrice-vecteur sans stocker la matrice. Pour les fonctions de base nodales, le temps de calcul serait toujours en $O(n_e r^6)$, mais en utilisant les fonctions orthogonales, la tensorisation des fonctions de base induit un algorithme en $O(n_e r^4)$, comme le remarque Warburton dans sa thèse [72].

7.2 Méthode générale

On considère le produit matrice-vecteur suivant

$$y_j = \int_K \sum_{1 \leqslant i \leqslant d} \left(A_i\, U \cdot \frac{\partial \varphi_j}{\partial x_i} - B_i \frac{\partial U}{\partial x_i} \cdot \varphi_j \right) dx - \int_{\partial K} \left(N_1 \{U\} + N_2 [U] \right) \cdot \varphi_j \, ds$$

On utilise F^{-1} pour transformer un élément K du maillage en l'élément de référence $\hat{K}$, et le changement de variable T de la définition 2.2.7 pour transformer l'élément de référence $\hat{K}$ en cube unité $\widetilde{Q}$. Définissons $G = F \circ T$, la transformation de l'élément K en le cube unité $\widetilde{Q}$. On utilise également une transformation g^{-1} d'une face ∂K vers la face de référence $\partial \widetilde{Q}$.

Sur $\widetilde{Q}$, on a

$$\widetilde{u} = \sum_k \widetilde{u}_k \, \widetilde{\varphi}_k,$$

on peut écrire y_j comme suit

$$\begin{aligned} y_j = \quad & \int_{\widetilde{Q}} \widetilde{|DG|} \sum_{1 \leqslant i \leqslant d} \sum_{1 \leqslant l \leqslant d} \left(A_i \widetilde{u} \cdot (\widetilde{DG^{-1}})_{l,i} \frac{\partial \widetilde{\varphi}_j}{\partial \widetilde{x}_l} - B_i \, (\widetilde{DG}^{-1})_{l,i} \frac{\partial \widetilde{u}}{\partial \widetilde{x}_l} \cdot \widetilde{\varphi}_j \right) d\widetilde{x}\, d\widetilde{y}\, d\widetilde{z} \\ & - \int_{\partial \widetilde{Q}} \widetilde{|Dg|} \left(N_1 \{\widetilde{u}\} + N_2 [\widetilde{u}] \right) \cdot \widetilde{\varphi}_j \, d\widetilde{s}. \end{aligned}$$

Les intégrales de volume sont calculées grâce à une formule de quadrature (ω_m, ξ_m) adaptée au cube $\widetilde{Q}$, tandis que les intégrales de surface sont évaluées avec une formule de quadrature (ω'_n, ξ'_n) adaptée aux faces $\partial \widetilde{Q}$ du cube.

On définit

$$\bar{U} = -\left(N_1 \left\{ \widetilde{u}(\xi'_n) \right\} + N_2 \left[\widetilde{u}(\xi'_n) \right] \right).$$

On écrit finalement

$$\begin{aligned} y_j = \quad & \sum_m \omega_m \widetilde{|DG|}(\xi_m) \sum_{1 \leqslant i \leqslant d} \sum_{1 \leqslant l \leqslant d} \left((\widetilde{DG}^{-1})_{l,i} \, A_i \widetilde{u} \cdot \frac{\partial \widetilde{\varphi}_j}{\partial \widetilde{x}_l} - B_i \, (\widetilde{DG}^{-1})_{l,i} \frac{\partial \widetilde{u}}{\partial \widetilde{x}_l} \cdot \widetilde{\varphi}_j \right)(\xi_m) \\ & + \sum_n \omega'_n \widetilde{|Dg|}(\xi'_n)\, \bar{U} \cdot \widetilde{\varphi}_j(\xi'_n). \end{aligned}$$

On décompose le produit matrice-vecteur en plusieurs étapes

Pour les intégrales de volume :

1. Calcul de

$$v_m = \widetilde{u}(\xi_m) = \sum_k \widetilde{u}_k \, \widetilde{\varphi}_k(\xi_m)$$

$$dv_m^l = \frac{\partial \widetilde{u}}{\partial \widetilde{x}_l}(\xi_m) = \sum_k \widetilde{u}_k \frac{\partial \widetilde{\varphi}_k}{\partial \widetilde{x}_l}(\xi_m)$$

2. Application de la géométrie et des coefficients physiques

$$v_m^{1,l} = \sum_{1 \leqslant i \leqslant d} \omega_m \, \widetilde{|DG|}(\xi_m) \, \widetilde{DG}_{l,i}^{-1} A_i v_m$$
$$v_m^2 = - \sum_{1 \leqslant i,l \leqslant d} \omega_m \, \widetilde{|DG|}(\xi_m) \, B_i \, \widetilde{DG}_{l,i}^{-1} \, dv_m^l$$

3. Calcul de

$$w_j^1 = \sum_m \sum_{1 \leqslant l \leqslant d} v_m^{1,l} \cdot \frac{\partial \widetilde{\varphi}_j}{\partial \widetilde{x}_l}(\xi_m)$$
$$w_j^2 = \sum_m v_m^2 \cdot \widetilde{\varphi}_j(\xi_m)$$

Pour les intégrales de surface :

1. Calculs de

$$s_n = \sum_k u_k \, \widetilde{\varphi}_k(\xi'_n)$$

2. Application de la géométrie et des coefficients physiques

$$\bar{s}_n = - \left(N_1 \{ s_n \} + N_2 [s_n] \right)$$
$$s_n^1 = \omega'_n \, \widetilde{|Dg|}(\xi'_n) \, \bar{s}_n$$

3. Calcul de

$$w_j^3 = \sum_n s_n^1 \cdot \widetilde{\varphi}_j(\xi'_n)$$

Vecteur final :

$$y_j = w_j^1 + w_j^2 + w_j^3$$

On considère les points de quadrature tensorisés suivants

$$\xi_m = (\xi_{m_1}, \xi_{m_2}, \xi_{m_3})$$

et les fonctions de base tensorisées suivantes

$$\varphi_j(x,y,z) = \widetilde{\varphi}_{j_1}(\widetilde{x}) \, \widetilde{\varphi}_{j_2}^{j_1}(\widetilde{y}) \, \widetilde{\varphi}_{j_3}^{j_1,j_2}(\widetilde{z})$$

On détaille à présent comment les différentes étapes de l'algorithme comportant des sommes peuvent se décomposer grâce à la tensorisation. En effectuant les sommes le long de chaque coordonnée $\widetilde{x}$, $\widetilde{y}$ ou $\widetilde{z}$, on peut réduire les sommes à $r+1$ termes, au lieu de $(r+1)^3$ termes si les fonctions de base et les points de quadratures n'étaient pas tensorisés.

7.3 Calcul des intégrales

7.3.1 Intégrales de volume

1. Pour $m = (m_1, m_2, m_3)$, on souhaite calculer

$$v_m = \sum_{k_1,k_2,k_3} \widetilde{\varphi}_{k_1}(\xi_{m_1}) \, \widetilde{\varphi}_{k_2}^{k_1}(\xi_{m_2}) \, \widetilde{\varphi}_{k_3}^{k_1,k_2}(\xi_{m_3}) \, u_{k_1,k_2,k_3}.$$

On note $\widetilde{\varphi}_{j_3}^{j_1,j_2}$ lorsque la fonction de base dépend de j_1 et j_2. La triple somme se scinde alors en trois sommes simples

$$\begin{aligned} u_{k_1,k_2,m_3}^1 &= \sum_{k_3} \widetilde{\varphi}_{k_3}^{k_1,k_2}(\xi_{m_3}) \, u_{k_1,k_2,k_3} \\ u_{k_1,m_2,m_3}^2 &= \sum_{k_2} \widetilde{\varphi}_{k_2}^{k_1}(\xi_{m_2}) \, u_{k_1,k_2,m_3}^1 \\ v_{m_1,m_2,m_3} &= \sum_{k_1} \widetilde{\varphi}_{k_1}(\xi_{m_1}) \, u_{k_1,m_2,m_3}^2 \end{aligned}$$

Remarque 7.3.1 *À ce stade, on remarque que la dépendance entre φ_j et ξ_m doit être « opposée », sinon on ne pourrait exploiter la structure tensorisée des points de quadrature et des fonctions de base pour avoir un algorithme rapide. Ainsi, des points de quadratures semi-tensorisés du type*

$$\xi_m = (\xi_{m_1}^{m_2,m_3}, \xi_{m_2}^{m_3}, \xi_{m_3})$$

pourraient également être utilisés.

Les trois sommes font intervenir $O(r)$ termes et sont calculées pour $O(r^3)$ valeurs, ce qui signifie que l'on a un coût en $O(r^4)$. Chaque somme pouvant être interprétée comme un produit matrice-vecteur, on a

$$\begin{aligned} U^1 &= C_1\, U \\ U^2 &= C_2\, U^1 \\ V &= C_3\, U^2, \end{aligned}$$

c'est à dire que l'on a une factorisation de la matrice C

$$C_{m,k} = \widetilde{\varphi}_k(\xi_m)$$

qui est

$$C = C_3\, C_2\, C_1$$

Alors que la matrice C est dense, les matrices C_1, C_2 et C_3 sont creuses. Comme elles sont également indépendantes de la géométries, les matrices C_1, C_2 et C_3 sont précalculées pour chaque type d'élément de référence.
Ainsi, on a

$$V = C\, U$$

Pour $n = n_1, n_2, n_3$, on veut maintenant calculer

$$dv_n^l(\xi_n) = \sum_k \widetilde{u}_k \frac{\partial \widetilde{\varphi}_k}{\partial \widetilde{x}_l}(\xi_n)$$

Comme, pour chaque type d'élément, on a l'inclusion $C_r \subset \mathbb{Q}_r$ (voir remarque 2.2.9), on peut écrire

$$\widetilde{u}(x,y,z) = \sum_k \widetilde{u}_k\, \widetilde{\varphi}_k(\widetilde{x},\widetilde{y},\widetilde{z}) = \sum_{m_1,m_2,m_3} v_{m_1,m_2,m_3}\, \psi_{m_1}(\widetilde{x})\, \psi_{m_2}(\widetilde{y})\, \psi_{m_3}(\widetilde{z})$$

où ψ_{m_1}, ψ_{m_2} , ψ_{m_3} sont les polynômes d'interpolation de Lagrange associés respectivement aux points ξ_{m_1}, ξ_{m_2} et ξ_{m_3}

$$\begin{aligned} \psi_{m_1}(\widetilde{x}) &= \frac{\displaystyle\prod_{n_1 \neq m_1} \widetilde{x} - \xi_{n_1}}{\displaystyle\prod_{n_1 \neq m_1} \xi_{m_1} - \xi_{n_1}} \\ \psi_{m_2}(\widetilde{y}) &= \frac{\displaystyle\prod_{n_2 \neq m_2} \widetilde{y} - \xi_{n_2}}{\displaystyle\prod_{n_2 \neq m_2} \xi_{m_2} - \xi_{n_2}} \\ \psi_{m_3}(\widetilde{z}) &= \frac{\displaystyle\prod_{n_3 \neq m_3} \widetilde{z} - \xi_{n_3}}{\displaystyle\prod_{n_3 \neq m_3} \xi_{m_3} - \xi_{n_3}} \end{aligned}$$

on a

$$\widetilde{\nabla} \psi_{m_1}(x)\, \psi_{m_2}(y)\, \psi_{m_3}(z) = \begin{vmatrix} \displaystyle\sum_{m_1,m_2,m_3} v_{m_1,m_2,m_3} \frac{d\psi_{m_1}}{x}(\widetilde{x})\, \psi_{m_2}(\widetilde{y})\, \psi_{m_3}(\widetilde{z}) \\ \displaystyle\sum_{m_1,m_2,m_3} v_{m_1,m_2,m_3}\, \psi_{m_1}(\widetilde{x}) \frac{d\psi_{m_2}}{d\widetilde{y}}(\widetilde{y})\, \psi_{m_3}(\widetilde{z}) \\ \displaystyle\sum_{m_1,m_2,m_3} v_{m_1,m_2,m_3}\, \psi_{m_1}(\widetilde{x})\, \psi_{m_2}(\widetilde{y}) \frac{d\psi_{m_3}}{d\widetilde{z}}(\widetilde{z}) \end{vmatrix}$$

Puisque

$$\psi_{m_1}(\xi^{n_1}) = \delta_{m_1,n_1}, \quad \psi_{m_2}(\xi^{n_2}) = \delta_{m_2,n_2}, \quad \psi_{m_3}(\xi^{n_3}) = \delta_{m_3,n_3},$$

les triples sommes sur m_1, m_2, m_3 se réduisent à de simples sommes

$$\begin{aligned}\left(\frac{\partial v}{\partial \widetilde{x}}\right)_{n_1,n_2,n_3} &= \sum_{m_1} \frac{d\psi^{m_1}}{d\widetilde{x}}(\xi_{n_1})\, v_{m_1,n_2,n_3}\\ \left(\frac{\partial v}{\partial \widetilde{y}}\right)_{n_1,n_2,n_3} &= \sum_{m_2} \frac{d\psi^{m_2}}{d\widetilde{y}}(\xi_{n_2})\, v_{n_1,m_2,n_3}\\ \left(\frac{\partial v}{\partial \widetilde{z}}\right)_{n_1,n_2,n_3} &= \sum_{m_3} \frac{d\psi^{m_3}}{d\widetilde{z}}(\xi_{n_3})\, v_{n_1,n_2,m_3}\end{aligned}$$

ces opérations peuvent être vues comme des produits matrice-vecteur

$$dV = RV$$

où la matrice R est creuse et indépendante de la géométrie. Elle peut donc être précalculée pour chaque type d'élément.

Finalement, on a la factorisation suivante

$$dV = R\,C\,U.$$

2. On calcule

$$\begin{aligned}v_m^{1,l} &= \sum_{1\leqslant i\leqslant d} \omega_m\, \widetilde{|DG|}(\xi_m)\, \widetilde{DG}_{l,i}^{-1} A_i v_m\\ v_m^2 &= -\sum_{1\leqslant i\leqslant d} \omega_m\, \widetilde{|DG|}(\xi_m)\, B_i \,(\sum_{1\leqslant l\leqslant d} \widetilde{DG}_{l,i}^{-1} dv_m^l).\end{aligned}$$

La complexité de cette opération est en $O(r^3)$ et peut être vue comme un produit matrice-vecteur

$$\begin{aligned}V^1 &= A\,V\\ V^2 &= B\,dV\end{aligned}$$

où les matrices A et B sont diagonales par bloc, chaque bloc étant lié à une formule de quadrature, et dépendant de la géométrie.

3. On s'intéresse à l'étape

$$w_j^1 = \sum_{m,l} v_m^{1,l} \cdot \frac{\partial \widetilde{\varphi}_j}{\partial \widetilde{x}_l}(\xi_m)$$

qui est l'opération transposée du calcul des dérivées des fonctions de base sur les points de quadrature. Ainsi, on a

$$W^1 = C^*\,R^*\,V^1.$$

On s'intéresse à présent à l'étape

$$w_j^2 = \sum_m v_m^2 \cdot \widetilde{\varphi}_j(\xi_m)$$

qui peut être interprétée comme

$$W^2 = C^*\,V^2,$$

c'est à dire

$$W^2 = C_1^*\,C_2^*\,C_3^*\,V^2.$$

Remarque 7.3.2 *Pour les hexaèdres, en utilisant les fonctions de base de la définition 6.2.1, la matrice C utilisée pour le calcul de u sur les points de quadrature est égale à l'identité*

$$C = I.$$

C'est ce qui fait que le produit matrice-vecteur est si rapide pour ces éléments, et la raison pour laquelle on souhaite avoir un maximum d'hexaèdres dans les maillages.

7.3.2 Intégrales de surface

1. On s'intéresse à l'étape
$$s_p = \widetilde{u}(\xi'_p) = \sum_k \widetilde{u}_k \, \widetilde{\varphi}_k(\xi'_p).$$
Puisque l'on considère les faces du cube unité, on a trois familles de points de quadrature
$$\begin{array}{c}(\delta, \xi'_{p_2}, \xi'_{p_3}) \\ (\xi'_{p_1}, \delta, \xi'_{p_3}) \\ (\xi'_{p_1}, \xi'_{p_2}, \delta)\end{array}$$
où δ vaut 0 ou 1. Le point de départ consiste à considérer le développement de u sur les fonctions de base du cube, c'est à dire
$$\widetilde{u}(\widetilde{x}, \widetilde{y}, \widetilde{z}) = \sum_{m_1, m_2, m_3} v_{m_1, m_2, m_3} \, \psi_{m_1}(\widetilde{x}) \, \psi_{m_2}(\widetilde{y}) \, \psi_{m_3}(\widetilde{z})$$
On calcule ensuite u sur les familles de points
$$\begin{array}{rcl} u^1_{m_2, m_3} & = & \widetilde{u}(\delta, \xi_{m_2}, \xi_{m_3}) \\ u^2_{m_1, m_3} & = & \widetilde{u}(\xi_{m_1}, \delta, \xi_{m_3}) \\ u^3_{m_1, m_2} & = & \widetilde{u}(\xi_{m_1}, \xi_{m_2}, \delta) \end{array}$$
Ces opérations sont de simples sommes
$$\begin{array}{rcl} u^1_{m_2, m_3} & = & \displaystyle\sum_{m_1} \psi_{m_1}(\delta) \, v_{m_1, m_2, m_3} \\ u^2_{m_1, m_3} & = & \displaystyle\sum_{m_2} \psi_{m_2}(\delta) \, v_{m_1, m_2, m_3} \\ u^3_{m_1, m_2} & = & \displaystyle\sum_{m_3} \psi_{m_3}(\delta) \, v_{m_1, m_2, m_3} \end{array}$$
qui peuvent être interprétée comme des produits matrice-vecteur avec des matrices creuses P_1, P_2 et P_3
$$U^1 = P_1 V, \quad U^2 = P_2 V, \quad U^3 = P_3 V$$
Puis, si la face considérée est une face quadrangulaire, on peut séparer le calcul
$$s^1_{p_2, p_3} = \sum_{m_2, m_3} \psi_{m_2}(\xi'_{p_2}) \, \psi_{m_3}(\xi'_{p_3}) \, u^1_{m_2, m_3}$$
en deux étapes
$$\begin{array}{rcl} z_{m_2, p_3} & = & \displaystyle\sum_{m_3} \psi_{m_3}(\xi'_{p_3}) \, u^1_{m_2, m_3} \\ s^1_{p_2, p_3} & = & \displaystyle\sum_{m_2} \psi_{m_2}(\xi'_{p_2}) \, z_{m_2, p_3} \end{array}$$
ce qui peut là encore être interprété comme un produit matrice-vecteur avec des matrices creuses
$$S^1 = T_2 \, T_1 \, U^1$$
Si la face considérée est une face triangulaire, en utilisant des points de quadrature symétriques (voir Dunavant [27]) qui ne sont pas tensorisés, on a seulement
$$S^1 = T \, U^1$$
où la matrice T est dense, mais restreinte à la face. La complexité du calcul de s_p est alors en $O(r^3)$ si l'élément est un hexaèdre, ne comportant donc que des faces quadrangulaires, et en $O(r^4)$ pour les autres éléments à cause des faces triangulaires. Pour certains éléments, certaines faces quadrangulaires ne sont bien évidemment pas traitées puisqu'elles se réduisent à un seul point sur l'élément K réel. Par exemple, pour les pyramides la face $z = 1$ n'est pas traitée.
2. On calcule
$$\bar{s}_n = -\left(N_1 \{s_n\} + N_2 [s_n]\right)$$
$$s^1_n = \omega'_n \, \widetilde{|Dg|}(\xi'_n) \, \bar{s}_n$$
3. On s'intéresse ensuite au calcul de
$$w^3_j = \sum_n s^1_n \cdot \widetilde{\varphi}_j(\xi'_n)$$
Cette étape est l'opération transposée du calcul de s_n, c'est à dire que le calcul de w^3 est fait en utilisant la transposée des matrices P_1, P_2, P_3, T_1 et T_2 définies précédemment.

7.4 Coût final

En faisant la somme finale, on obtient donc
- Calcul de w_j^1 et w_j^2 : coût en $O(r^4)$ pour les étapes 1 et 3, coût en $O(r^3)$ pour l'étape 2,
- Calcul de w_j^3 : coût en $O(r^3)$ pour les hexaèdres, en $O(r^4)$ pour les tétraèdres pour les étapes 1 et 3, coût en $O(r^2)$ pour l'étape 2,

soit un coût final en $O(r^4)$ au lieu de $O(r^6)$ pour un produit matrice-vecteur classique.

Chapitre 8

Étude numérique des éléments discontinus

Afin de dégager des propriétés numériques des éléments construits dans le chapitre 6, on effectue une analyse de dispersion et une étude de stabilité des schémas obtenus à partir de ces éléments. Un cas test numérique vient finalement confirmer le bon comportement des éléments au sein d'un maillage hybride.

Sommaire

8.1 Propriétés numériques des éléments

8.1.1 Analyse de dispersion

Comme pour les éléments continus (cf section 4.1 du chapitre 4), on effectue une analyse de dispersion pour les éléments discontinus construits précédemment, avec l'équation de Helmholtz ou avec les équations de Maxwell. On rappelle que l'étude est faite sur des cellules périodiques prises comme un cube découpé en un unique hexaèdre, deux prismes, deux pyramides et deux tétraèdres (hybride), six pyramides ou six tétraèdres comme le montre la figure 4.1. Afin de vérifier la consistance de nos méthodes lorsque les éléments ne sont pas affines, l'analyse a également été faite sur des cellules périodiques faites à partir de cubes déformés (voir figure 4.2).

Pour l'équation de Helmholtz et les équations de Maxwell sans pénalisation, c'est à dire $\alpha = 0$ dans l'équation 1.2.4, on obtient un ordre de dispersion de $2r$ avec l'espace optimal, pour les maillages avec éléments affines comme pour les maillages déformés. Avec pénalisation, c'est à dire $\alpha < 0$ dans l'équation 1.2.4, on obtient une erreur de dispersion en $O(h^{2r+1})$ pour l'espace optimal, tandis que l'utilisation de l'espace $\mathbb{P}_r$ conduit à un taux de convergence plus faible, comme le montre la figure 8.1 pour les équations de Maxwell sur un maillage périodique composé de pyramides non-affines.

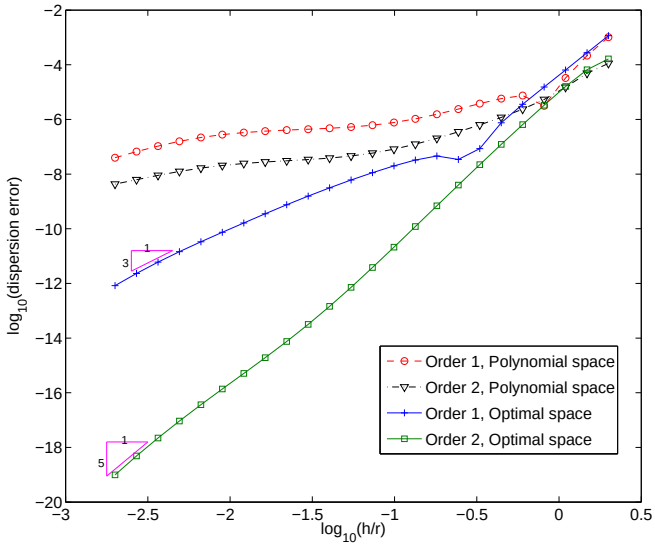

FIG. 8.1 – Erreur de dispersion pour un maillage périodique composé de pyramides non-affines pour les équations de Maxwell

Les courbes de dispersion pour l'équation de Helmholtz avec les éléments réguliers d'ordre 1 à 3 sont présentées sur la figure 8.2. Comme dans le cas continu, tous les types d'éléments présentent les mêmes propriétés de dispersion. L'élément le moins dispersif est l'élément pyramidal dans la plupart des cas. La même étude a été faite pour les éléments déformés, comme présenté sur la figure 8.3 pour les ordres 1 et 2, et mène aux mêmes conclusions. Dans les deux cas, la dispersion pour tous les éléments diminue lorsque l'on monte en ordre.

8.1.2 Étude de stabilité

Les méthodes de Galerkin discontinues étant utilisées de manière préférentielles pour la résolution de cas instationnaires, l'étude de la CFL pour les élément discontinus est très importante. Pour chaque type d'élément, le tableau 8.1 donne la CFL obtenue jusqu'à l'ordre 4 sur des maillages régulier, et jusqu'à l'ordre 2 sur des maillages irrégulier. Le critère de stabilité dans le cas instationnaire a également été recherché afin de vérifier la validité de ces résultats.

Comme dans le cas continu, l'intégration exacte donne une CFL plus élevée qu'avec une l'intégration minimale pour les éléments pyramidaux, et les conditions CFL pour tous les éléments se classent comme suit

$$CFL_{Hexa} > CFL_{Wedge} > CFL_{Tetra} > CFL_{Pyr-ExactInt} > CFL_{Pyr-ApproxInt}.$$

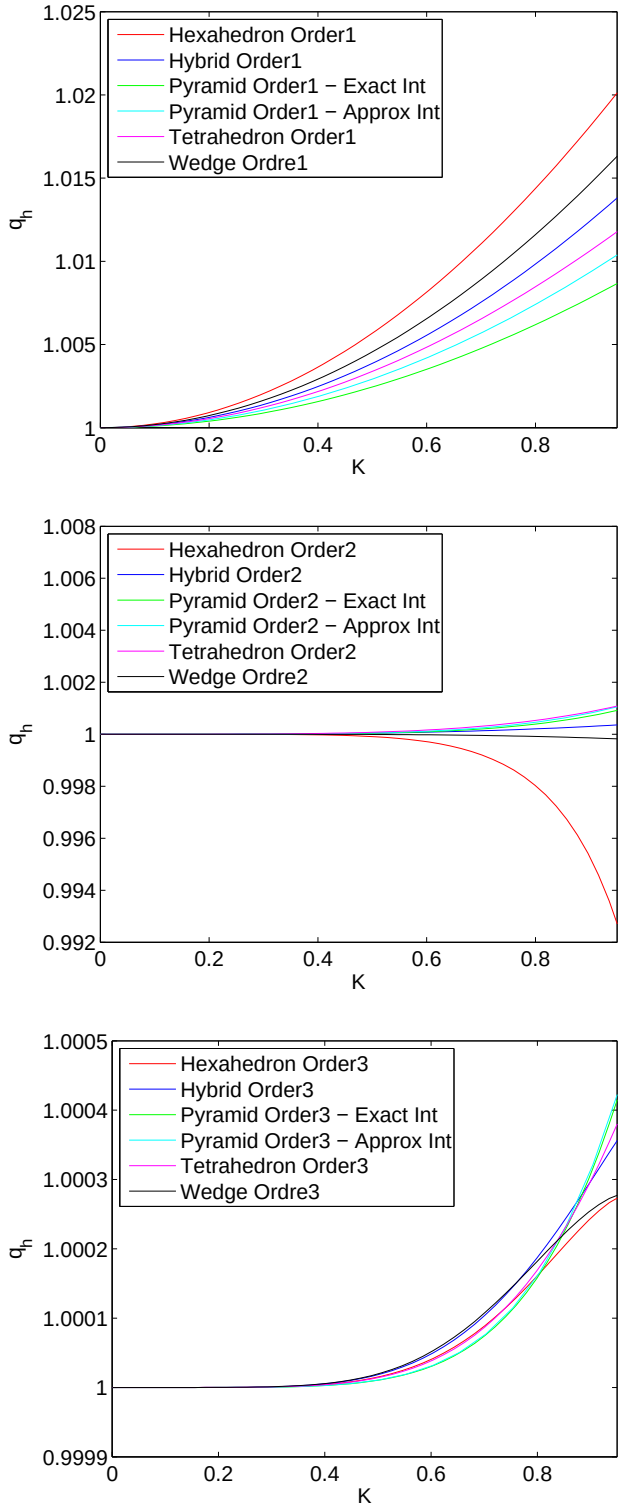

FIG. 8.2 – Courbes de dispersion pour une méthode de Galerkin discontinue pour les ordres 1 à 3 sur un maillage régulier ($K = \frac{6kh}{2\pi r}$)

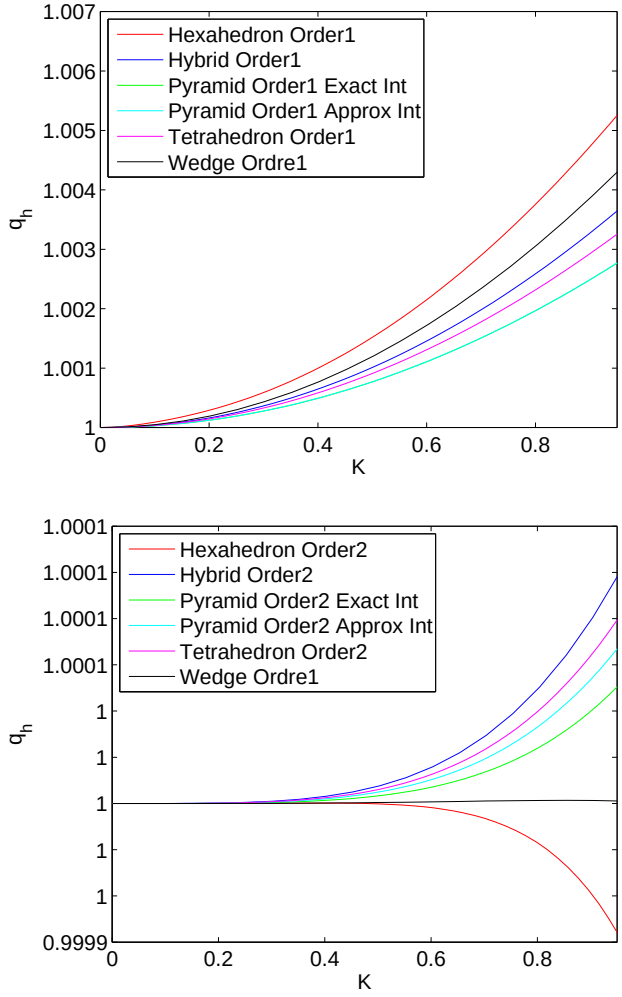

FIG. 8.3 – Courbes de dispersion pour une méthode de Galerkin discontinue pour les ordres 1 et 2 sur un maillage déformé ($K = \frac{6kh}{2\pi r}$)

TAB. 8.1 – Stabilité des éléments continus pour un maillage régulier et un maillage déformé avec les éléments finis discontinus

Élément	Maillage régulier				Maillage déformé	
	Ordre 1	Ordre 2	Ordre 3	Ordre 4	Ordre 1	Ordre 2
Hexaèdre	0.14434	0.07144	0.04348	0.02934	0.139306	0.06712
Prisme	0.11471	0.04348	0.03957	0.02717	0.102650	0.05186
Pyramide IntExacte	0.07184	0.04058	0.02618	0.0184	0.047548	0.02512
Pyramide IntApprox	0.04811	0.02544	0.01566	0.0112	-	-
Hybride	0.09283	0.05363	0.03527	0.02448	0.071584	0.03758
Tétraèdre	0.07373	0.04467	0.03041	0.02173	0.041028	0.02380

Remarque 8.1.1 *Du fait que l'on a beaucoup plus de degrés de liberté en discontinu et que la cellule de base du motif est huit fois plus grosse pour les maillages irréguliers, les temps de calcul sont beaucoup plus longs dans ce cas. C'est pourquoi les calculs de dispersion et de CFL sur maillage déformé pour les éléments discontinu n'ont pas été menés au delà de l'ordre 2.*

8.2 Convergence

On souhaite vérifier l'ordre de convergence obtenu par l'étude de dispersion. On considère les équations de Maxwell en domaine temporel sur une cavité cubique $[-5,5]^3$ avec une source gaussienne

$$f = x\, e^{-\alpha r^2}\, e_x$$

avec un maillage hybride non-affine composé de cellules hybrides (voir figure 4.2).

On trace l'erreur obtenue en norme L^2 par rapport au pas du maillage h en échelle logarithmique sur la figure 8.4. On observe que l'erreur en norme L^2 est en $O(h^{r+1})$ comme on le souhaitait. Pour ce cas test, l'intégration utilisée pour les pyramides est celle avec $r+1$ points de Gauss-Jacobi dans la direction $\tilde{z}$ et $r+1$ points de Gauss dans les directions $\tilde{x}$ et $\tilde{y}$.

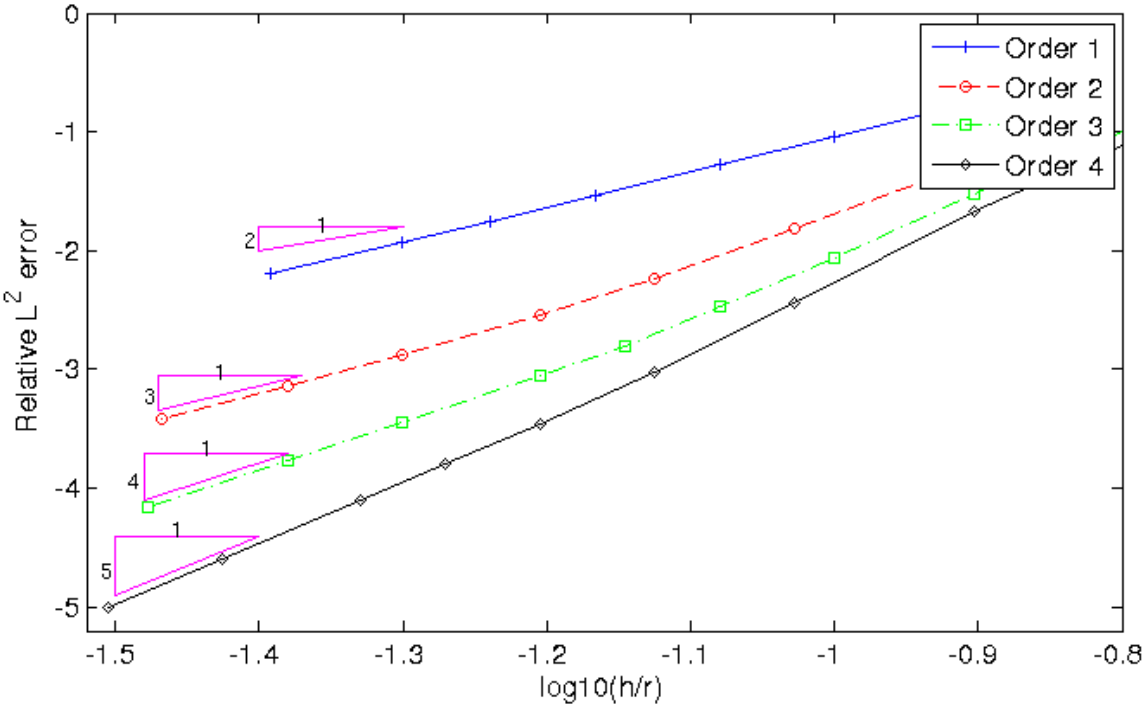

FIG. 8.4 – Erreur relative en norme L^2 sur un maillage hybride déformé (tétraèdres + pyramides) pour un cube de taille 10λ avec un schéma de Runge-Kutta d'ordre 4.

8.3 Équations de Maxwell sur un cone-sphère

On étudie les équations de Maxwell en régime temporel sur le cas-test d'un cone-sphère de bord Γ placé dans une boite parallélépipédique $[-3.5,3.5]\times[-5,10]\times[-3.5,3.5]$

On met une condition de conducteur parfait sur Γ, et la condition de Sommerfeld est approchée en utilisant des couches PML entourant le domaine de calcul. Pour la source f, on prend une gaussienne en espace et un Ricker en temps

$$f(x,t) = \frac{1}{r_0^2}\, e^{-13\frac{r}{r_0}^2}\, \pi^2\, (f_0 t - 1)^2\, e^{-\pi^2 (f_0\, t-1)^2}$$

où r est la distance au centre de la source, r_0 le rayon de distribution de la gaussienne, f_0 la fréquence centrale du Ricker. Pour cette expérience, on a pris $f_0 = 1.5$ et $r_0 = 0.9$.

On étudie le cas d'un maillage hybride et d'un maillage hexaédrique obtenu en découpant chaque tétraèdre d'un maillage tétraédrique en 4 hexaèdres. Le maillage hybride et le maillage tétraédrique utilisé pour obtenir le maillage hexaédrique sont représentés sur la figure 8.5.

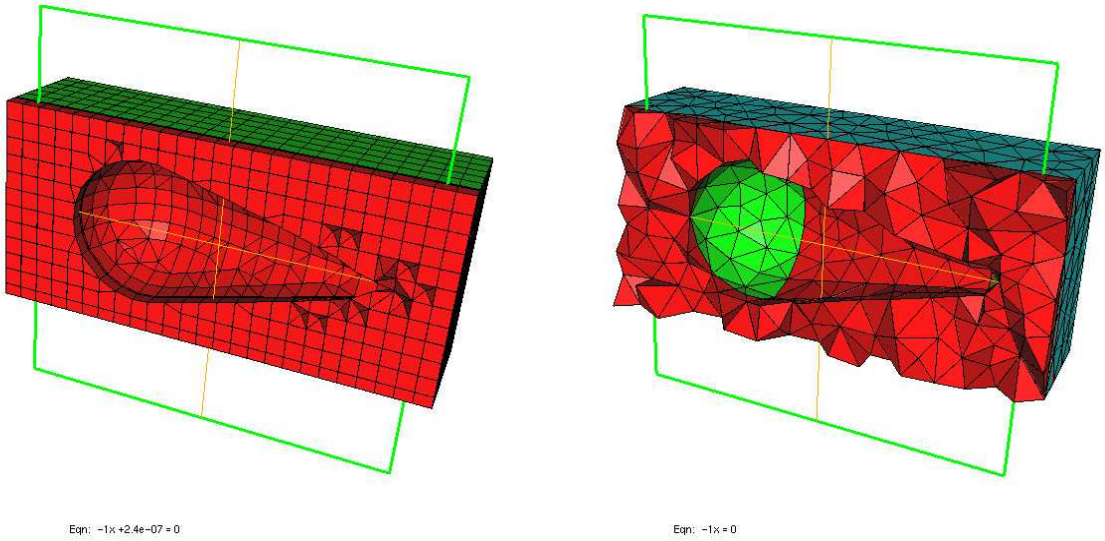

FIG. 8.5 – Maillages utilisés : maillage hybride (à gauche) et maillage tétraédrique (à droite) à la base du maillage hexaédrique utilisé

On utilise un schéma de Runge-Kutta d'ordre 4 pour la discrétisation en temps et des éléments d'ordre 4 pour la discrétisation spatiale. Les instantanés à $T = 2$, $T = 4$ et $T = 13$ sont donnés sur la figure 8.6 sur le maillage hybride.

On compare les solutions obtenues sur chaque type de maillage avec une solution de référence calculée sur une maillage hybride plus fin ($h = 0.2$) et des éléments d'ordre 5. Les résultats obtenus sont résumés dans le tableau 8.2. Le cas des tétraèdres découpés courbes n'a pu être traité à cause de la difficulté d'obtenir des éléments courbes valides à partir d'un maillage tétraédrique.

TAB. 8.2 – Erreur, nombre de degrés de liberté pour E_x, pas de temps et temps de calcul pour les différents types de maillages

Type de maillage	Hexaédrique	Hybride	
		Droit	Courbe
Données	erreur de 9,58% 24,95 millions ddls $\Delta t = 0.003$ $h = 0.5$ (tétra)	erreur de 4,15% 8,93 millions ddls $\Delta t = 0.006$ $h = 0.25$	erreur de 0,15%
Temps de calcul à $T = 20$	7d 13h 41min	1d 9h 21min	

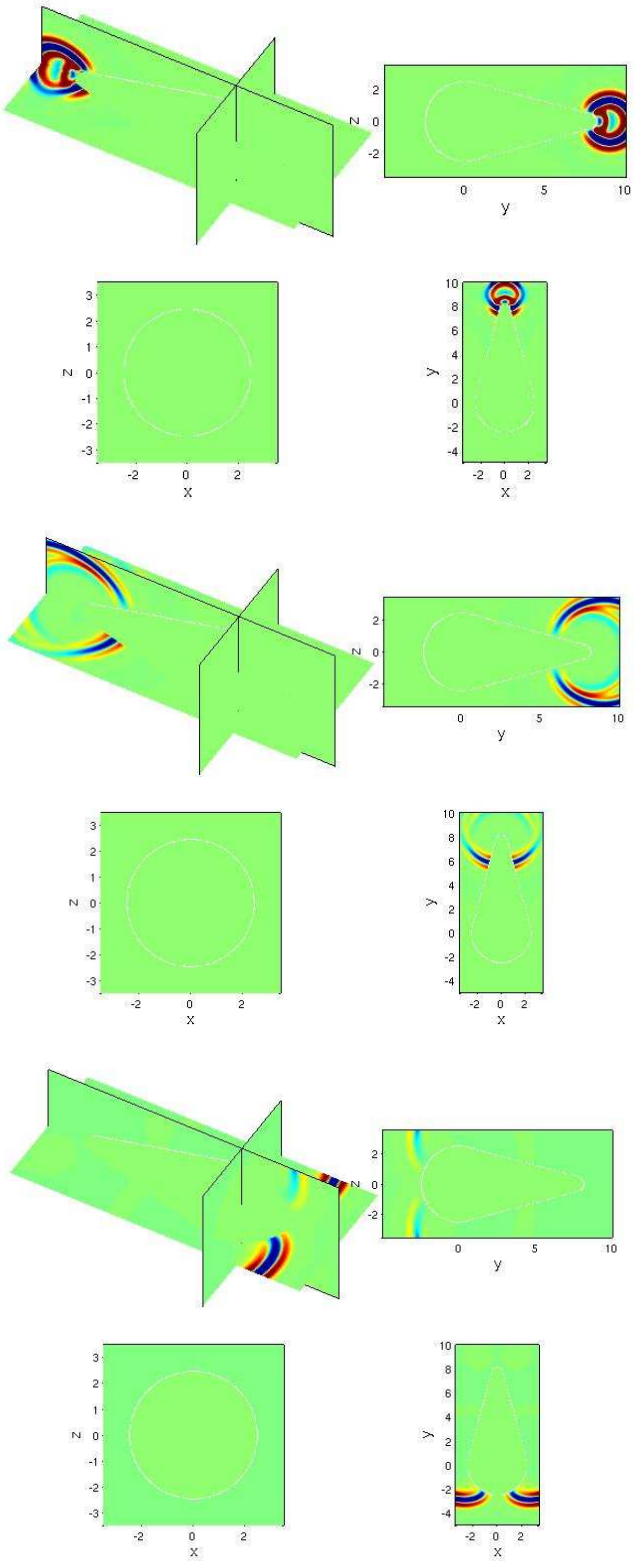

FIG. 8.6 – Instantanés aux temps $T = 2s$ (haut), $T = 4s$ (milieu) et $T = 13s$ (bas) pour le maillage hybride

Chapitre 9

Comparaison avec d'autres méthodes

Nous présentons ici quelques éléments qui peuvent être utilisés avec des méthodes discontinues, et nous effectuons des comparaisons numériques sur les équations de Maxwell.

Sommaire

9.1 Présentation d'autres types d'éléments

9.1.1 Hexaèdre dégénéré

Il est possible de considérer les pyramides comme des éléments hexaédriques dégénérés, obtenus par la « transformation de Duffy » (T^{-1} avec la définition 2.2.7) présentée sur la figure 9.1. On peut également construire des éléments finis sur ce principe en plaçant des degrés de liberté sur des points de type Gauss et les fonctions de base associées sur le cube, et en appliquant la transformation. Le nombre de degrés de liberté est alors bien plus important que dans le cas de l'élément « optimal » mais on a la condensation de masse.

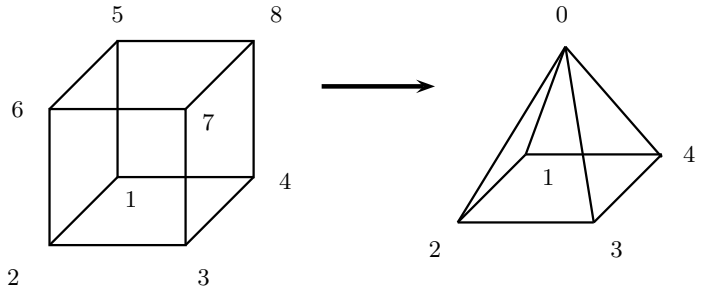

FIG. 9.1 – Transformation de Duffy.

Bien que cette transformation conduise à des fonctions de base interpolant correctement la solution à l'intérieur de la pyramide, de nombreux problèmes apparaissent (Bedrosian [5]). Cet élément n'est évidemment pas utilisable pour des éléments continus puisque la position des points sur les faces triangulaires est fixée dès le départ et ne coïncide pas avec les degrés de liberté sur les faces d'un tétraèdre classique. De plus, on a pris l'espace $\mathbb{Q}_r$ au lieu de C_r comme espace d'approximation, ce qui signifie que les fonctions de base obtenues par cette transformation ne respectent donc pas les conditions de restriction sur les faces triangulaires. Il faut alors utiliser des points de quadrature tensorisés sur chaque face triangulaire pour intégrer $\mathbb{Q}_r$.

Le tableau 9.1 présente les résultats de la CFL pour les éléments pyramidaux obtenus par la transformation de Duffy pour différentes configurations de degrés de liberté obtenus à partir d'un produit tensoriel de formules de quadrature. L'espace polynomial pour lequel la formule de quadrature est exacte est indiquée, ainsi que l'ordre de dispersion obtenu.

Il est clair que, malgré l'atout que constitue la condensation de masse, l'utilisation de ces éléments est vivement déconseillée car la CFL est très pénalisante.

9.1.2 Éléments nodaux et monomiaux

Évidemment, il est possible d'utiliser les éléments nodaux définis pour la formulation continue dans le chapitre 2, ou les fonctions de base monomiales de $\hat{P}_r$ comme fonctions de base. Cependant, ces fonctions n'ont pas de propriété particulière : la matrice de masse est pleine et mal conditionnée, le nombre de valeurs non nulles dans la matrice de masse est en $O(r^6)$. De plus, prendre les fonctions monomiales implique une matrice mal conditionnée. En revanche, l'utilisation des fonctions nodales a l'avantage de localiser le calcul des flux aux degrés de liberté de la face.

9.1.3 Astuce de Warburton

Dans le cas discontinu, l'inversion de la matrice de masse peut être évitée en considérant la transformation non conforme H^1 proposée par Warburton (CANUM 2010)

$$\hat{\varphi}_i = \frac{1}{\sqrt{|DF|}}\, \varphi_i \circ F^{-1}. \tag{9.1.1}$$

Avec cette transformation, la matrice de masse s'écrit alors

$$(M_h)_{i,j} = \int_K \varphi_i \cdot \varphi_j dx = \int_{\hat{K}} |DF| \frac{\hat{\varphi}_i}{\sqrt{|DF|}} \cdot \frac{\hat{\varphi}_j}{\sqrt{|DF|}} dx = \int_{\hat{K}} \hat{\varphi}_i \cdot \hat{\varphi}_j dx,$$

c'est à dire que la matrice de masse est indépendante de la géométrie. Si l'on utilise en outre des fonctions orthonormales, la matrice est égale à l'identité.

TAB. 9.1 – Stabilité et ordre de dispersion des différents éléments pyramidaux avec une méthode de Galerkin discontinue

Élément (base - z)	Quadrature	Maillage régulier $r=1$	$r=2$	$r=3$	$r=4$	$r=5$	Disp
Pyramide exacte	$(1-z)^2\,\mathbb{Q}_{2r+1,2r+1,2r}$	0.07300	0.04109	0.02644	0.01843	0.01351	2r
Pyramide G - G	$\mathbb{Q}_{2r+1,2r+1,2r+1}$	0.04870	0.02567	0.01577	0.01065	0.00766	2r
Rapport CFL		1.5	1.6	1.7	1.7	1.8	
Pyramide L - RJ	$(1-z)^2\,\mathbb{Q}_{2r-1,2r-1,2r}$	0.07879	0.04521	0.02883	0.01986	0.01357	2r
Rapport CFL		0.9	0.9	0.9	0.9	1.0	
Hexaèdre G - G	$\mathbb{Q}_{2r+1,2r+1,2r+1}$	0.02349	0.00651	0.00243	0.00100	0.00056	2r
Rapport CFL		3.1	6.3	10.9	16.8	23.9	
Hexaèdre G - R	$\mathbb{Q}_{2r+1,2r+1,2r}$	0.04505	0.00963	0.00323	0.00137	0.00068	2r
Rapport CFL		1.6	4.3	8.2	13.4	19.9	
Hexaèdre G - J	$(1-z)\,\mathbb{Q}_{2r+1,2r+1,2r+1}$	0.04210	0.01274	0.00498	0.00233	0.00123	2r
Rapport CFL		1.7	3.2	5.3	7.9	11.0	
Hexaèdre G - RJ	$(1-z)\,\mathbb{Q}_{2r+1,2r+1,2r}$	0.06373	0.01654	0.00613	0.00277	0.00143	2r
Rapport CFL		1.1	2.5	4.3	6.7	9.5	
Hexaèdre L - G	$\mathbb{Q}_{2r-1,2r-1,2r+1}$	0.02192	0.00705	0.00284	0.00134	0.00071	2 (r-1)
Rapport CFL		3.3	5.8	9.3	13.8	19.1	
Hexaèdre L - R	$\mathbb{Q}_{2r-1,2r-1,2r}$	0.03448	0.00975	0.00365	0.00164	0.00084	2 (r-1)
Rapport CFL		2.1	4.2	7.2	11.2	16.1	
Hexaèdre L - J	$(1-z)\,\mathbb{Q}_{2r-1,2r-1,2r+1}$	0.03594	0.01301	0.00563	0.00278	0.00152	2 (r-1)
Rapport CFL		2.0	3.2	4.7	6.6	8.9	
Hexaèdre L - RJ	$(1-z)\,\mathbb{Q}_{2r-1,2r-1,2r}$	0.04913	0.01661	0.00687	0.00329	0.00175	2 (r-1)
Rapport CFL		1.5	2.5	3.8	5.6	7.7	

G = Gauss, R = Radau, L = Lobatto, J= Jacobi, RJ = Radau-Jacobi

La matrice de rigidité R_h s'écrit quant à elle

$$\begin{aligned}(R_h)_{j,k} &= \int_{\hat{K}} |DF| \sum_{1\leqslant i\leqslant d} A_i \frac{\partial(\frac{\hat{\varphi}_k}{\sqrt{|DF|}})}{\partial x_i} \cdot \frac{\hat{\varphi}_j}{\sqrt{|DF|}} d\hat{x} - \int_{\hat{K}} |DF| \sum_{1\leqslant i\leqslant d} B_i \frac{\hat{\varphi}_k}{\sqrt{|DF|}} \cdot \frac{\partial(\frac{\hat{\varphi}_j}{\sqrt{|DF|}})}{\partial x_i} d\hat{x} \\ &= \int_{\hat{K}} \sum_{1\leqslant i\leqslant d} A_i \frac{\partial \hat{\varphi}_k}{\partial x_i} \cdot \hat{\varphi}_j \, d\hat{x} - \int_{\hat{K}} \sum_{1\leqslant i\leqslant d} B_i \hat{\varphi}_k \cdot \frac{\partial \hat{\varphi}_j}{\partial x_i} \, d\hat{x} \\ &\quad + \frac{1}{2} \int_{\hat{K}} \sum_{1\leqslant i\leqslant d} (B_i - A_i) \frac{1}{|DF|} \frac{\partial |DF|}{\partial x_i} \hat{\varphi}_k \cdot \hat{\varphi}_j \, d\hat{x}\end{aligned}$$

On peut utiliser le produit matrice-vecteur détaillé dans le chapitre 7 sans surcoût important. De plus, puisque la matrice de masse est égale à l'identité, le calcul est plus rapide, comme en témoigne le tableau 9.12. Cependant, la matrice de rigidité fait intervenir les dérivées du jacobien qui doivent être manipulées avec attention dans le cas des pyramides. En effet, dans le cas de pyramides non-affines, le jacobien est singulier

$$\frac{\partial |DF|}{\partial \hat{x}} = B_1 \frac{1}{1-\hat{z}} + C \frac{\hat{y}}{(1-\hat{z})^2} = B_1 \frac{1}{1-\widetilde{z}} + C \frac{2\widetilde{y}-1}{1-\widetilde{z}}$$

On ne doit évidemment pas prendre de point de quadrature sur le point de singularité. En outre, à cause de la fraction rationnelle en $(1-\hat{z})$ qui apparait en plus lors du calcul des intégrales de volume, on doit utiliser des points de quadrature Gauss-Jacobi induisant une formule de quadrature exacte pour les polynômes de $(1-\widetilde{z})\mathbb{Q}_r$ au lieu de $(1-\widetilde{z})^2\mathbb{Q}_r$ comme choisi habituellement pour les pyramides.

On remarque que les constantes ne sont pas approchées par cette technique qui ne pourra donc être en $O(h)$. L'ajout d'une fonction de base est proposée par Warburton pour permettre aux constantes d'être incluses dans l'espace d'approximation, mais elle dépend de la géométrie, si bien que la technique perd de son intérêt. Par ailleurs, les fonctions d'ordre 1 ne sont pas générées, à moins de rajouter des fonctions qui dépendent elles aussi de la géométrie.

9.2 Comparaison numérique des éléments

9.2.1 Astuce de Warburton

On considère les équations de Maxwell en domaine temporel sur une cavité cubique $[-5,5]^3$ avec une source gaussienne

$$f = x\, e^{-\alpha r^2}\, e_x$$

avec un maillage hybride non-affine composé de cellules hybrides (voir figure 4.2).

Les solutions pour $T = 5$ et $T = 50$ sont données sur la figure 9.2. Une analyse de l'erreur est réalisée pour la solution obtenue au temps $T = 50s$ pour les ordres 3 et 5 en utilisant un solveur creux pour calculer la matrice de masse avec l'astuce de Warburton. Comme le montre la figure 9.3, comme on s'y attendait, cette astuce ne permet pas d'obtenir un schéma h-convergent, alors que la convergence en h est assurée par notre méthode.

Cependant, on peut conjecturer que l'astuce de Warburton permet d'obtenir un schéma p-convergent puisque l'erreur de consistance du schéma décroit en passant de l'ordre 3 à l'ordre 5. De plus, l'erreur de consistance est faible (moins de 2%), et même acceptable dans la plupart des simulations.

9.2.2 Hexaèdres

Les fonctions de base orthogonales de la proposition 6.2.1 permettent d'obtenir une matrice de masse diagonale, y compris pour les hexaèdres non-affines, alors que les fonctions de base obtenues par le produit tensoriel de polynômes de Legendre, proposées par Kirby *et al.* [48] (avec $P_1 = P_2 = P_3 = r$) et Warburton [72] ne permettent d'obtenir la condensation de masse que pour les éléments affines.

Lorsque l'hexaèdre est affine, on peut utiliser les fonctions de base de la proposition 6.2.3. On écrit les dérivées des polynomes de Legendre de la manière suivante

$$\frac{dP_i^{0,0}}{dx}(x) = \sum_{j=0}^{i-1} \eta_j^i P_j^{0,0}(x).$$

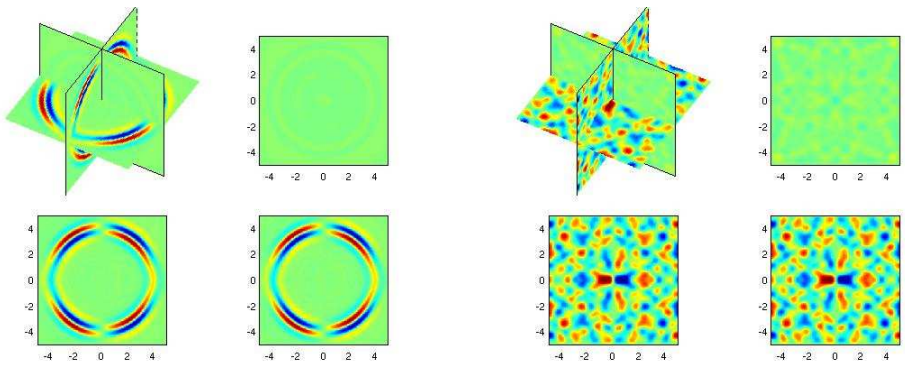

FIG. 9.2 – Solution pour $T = 5$ (gauche) et $T = 50$ (droite) pour une cavité cubique de taille 10λ

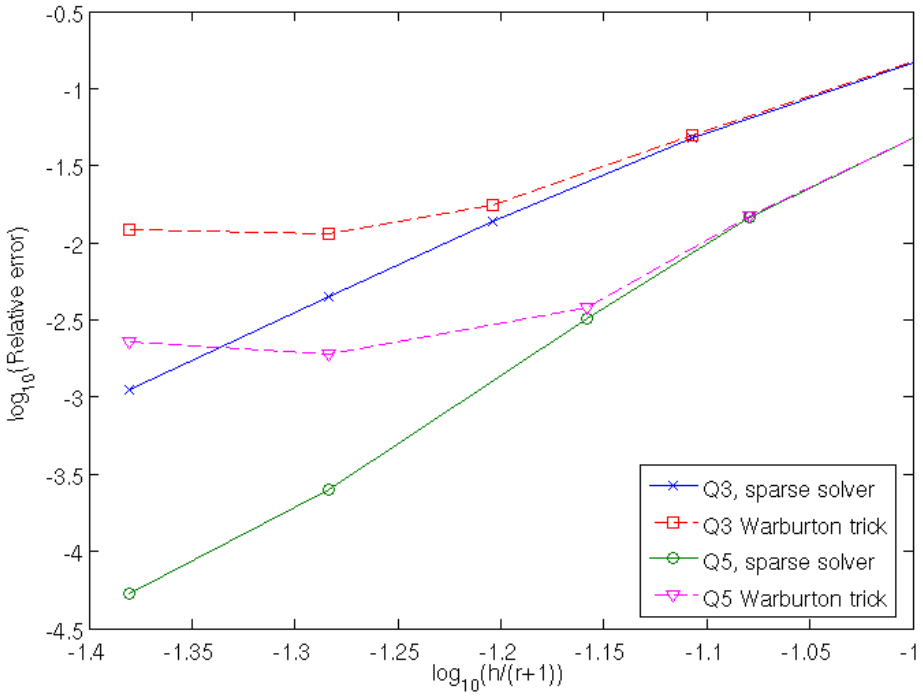

FIG. 9.3 – Erreur relative en norme L^2 sur un maillage hybride déformé (tétraèdre + pyramides) pour un cube de taille 10λ avec un schéma de Runge-Kutta d'ordre 4

Grâce à l'orthogonalité, pour un hexaèdre affine, on a

$$\begin{aligned}\int_{\hat{K}} \frac{\partial \varphi_{i_1,i_2,i_3}}{\partial \hat{x}} \varphi_{j_1,j_2,j_3} dx &= \eta^{i_1}_{j_1} \delta_{i_2,j_2} \delta_{i_3,j_3} \\ \int_{\hat{K}} \frac{\partial \varphi_{i_1,i_2,i_3}}{\partial \hat{y}} \varphi_{j_1,j_2,j_3} dx &= \eta^{i_2}_{j_2} \delta_{i_1,j_1} \delta_{i_3,j_3} \\ \int_{\hat{K}} \frac{\partial \varphi_{i_1,i_2,i_3}}{\partial \hat{z}} \varphi_{j_1,j_2,j_3} dx &= \eta^{i_3}_{j_3} \delta_{i_1,j_1} \delta_{i_2,j_2}\end{aligned}$$

La matrice de rigidité est donc plus creuse dans ce cas que dans le cas où l'on utilise les fonctions de la proposition 6.2.1. Le coût asymptotique du produit avec la matrice de rigidité est en r^4 tandis qu'il est en $12r^4$ lorsque l'on utilise la base de $\mathbb{Q}_r$. Cependant, le coût de calcul de la matrice de flux reste élevé, même s'il n'est qu'en $O(r^3)$. En effet, pour la face $x = 0$ par exemple, on a

$$u(0,\hat{y},\hat{z}) = \sum_{i_1,i_2,i_3} u_{i_1,i_2,i_3} P^{0,0}_{i_1}(-1)$$

On peut donc écrire les valeurs de s sur la face en utilisant le développement par polynômes de Legendre

$$s(\hat{y},\hat{z}) = \sum_{i_2,i_3} s_{i_2,i_3} P^{0,0}_{i_2}(\hat{y}) P^{0,0}_{i_3}(\hat{z})$$

On obtient s sous forme de vecteur avec le produit matrice-vecteur suivant

$$S = PU$$

où

$$P_{(i_2,i_3),(j_1,j_2,j_3)} = P^{0,0}_{j_1}(-1)\, \delta_{i_2,j_2}\, \delta_{i_2,j_2}$$

Puisque l'on peut écrire s avec des polynômes de Legendre qui forment une base orthonormée sur chaque face du cube, il n'est plus nécessaire d'utiliser une formule de quadrature pour évaluer les intégrales de surface dans le produit matrice-vecteur rapide.

On s'attend à un calcul beaucoup plus efficace en utilisant cette base plutôt que celle de $\mathbb{Q}_r$, mais bien qu'il y ait moins de degrés de liberté pour $\mathbb{P}_r$ que pour $\mathbb{Q}_r$, il faudra utiliser un maillage beaucoup plus fin avec la base de $\mathbb{P}_r$ pour avoir la même précision. La comparaison des temps de calcul pour les deux bases est donnée dans le tableau 9.2 pour un nombre égal de degrés de liberté (un million). L'utilisation de $\mathbb{P}_r$ semble intéressante à partir de l'ordre 6 et reste relativement constante à ordre élevé. Cependant, la différence de coût entre les deux bases reste peu importante.

TAB. 9.2 – Temps de calcul pour 100 itérations avec un maillage d'hexaèdres affines comportant un million de degrés de liberté pour l'inconnue E_x.

Ordre	2	3	4	5	6	7	8	10	12	15
$\mathbb{Q}_r$	136s	111s	107s	100s	101s	99s	109s	105s	113s	140s
$\mathbb{P}_r$	214s	157s	135s	117s	111s	105s	106s	101s	97s	99s

Pour une comparaison plus fine, on étudie la norme L^2 obtenue pour une cavité cubique maillée avec des petits cubes. Grâce à la symétrie, on peut se contenter d'étudier le cube $[0,5\lambda]^2$. Les tableaux 9.3 et 9.4 présentent les temps de calculs minimaux et le nombre minimal de degrés de liberté nécessaires à l'obtention d'une erreur inférieure à 1% pour les bases de $\mathbb{P}_r$ et $\mathbb{Q}_r$.

TAB. 9.3 – Pas de temps, nombre de degrés de liberté et temps de calcul nécessaires à l'obtention d'une erreur inférieure à 1% pour une cavité cubique (maillage avec des hexaèdres affines) avec base orthogonale de $\mathbb{P}_r$.

Ordre	3	4	5	6	8	10	14	18	26
Δt	0.035	0.0405	0.0394	0.0395	0.0348	0.0325	0.0272	0.0235	0.0179
Ddls	540 000	240 065	153 664	111 804	84 480	61 776	43 520	35 910	29 232
Temps CPU	11 818s	2865s	1387s	856s	695s	500s	444s	376s	620s

L'utilisation de la base de $\mathbb{P}_r$ semble donner de meilleurs résultats qu'en utilisant la base de $\mathbb{Q}_r$ au dessus de l'ordre 6, mais la différence est là encore minime aussi nous préconisons d'utiliser la base de $\mathbb{Q}_r$ dans tous les cas, de manière à n'avoir qu'une seule méthode valable aussi bien pour le cas affine que non-affine.

TAB. 9.4 – Pas de temps, nombre de degrés de liberté et temps de calcul nécessaires à l'obtention d'une erreur inférieure à 1% pour une cavité cubique (maillage avec des hexaèdres affines) avec base orthogonale de $\mathbb{Q}_r$.

Ordre	3	4	5	6	7	9	11	16	30
Δt	0.032	0.036	0.0353	0.0344	0.0381	0.0315	0.0301	0.023	0.0156
Ddls	512 000	216 000	157 464	117 649	64 000	64 000	46 656	39 304	29 791
Temps CPU	5 575s	1800s	1180s	1013s	417s	580s	426s	1134s	3520s

9.2.3 Prismes

Comme pour les hexaèdres, on compare le temps de calcul pour construire et inverser la matrice de masse avec l'algorithme décrit dans la section 6.3.3 dans le cas de fonctions de base nodales et orthogonales pour un maillage composé de prismes non-affines et comportant un million de degrés de liberté pour l'inconnue E_x. Les résultats sont données dans le tableau 9.5. Le temps de calcul pour les fonctions de base nodales croît très vite avec l'ordre utilisé, ce qui rend les fonctions orthogonales très attractives.

TAB. 9.5 – Temps de calcul de la matrice de masse avec un maillage de prismes non-affines comportant un million de degrés de liberté pour l'inconnue E_x.

Ordre	2	3	4	5	6	7	8	9	10
Nodal	0.5s	0.935s	1.66s	2.99s	5.27s	8.02s	12.4s	20s	30.1s
Orthogonal	0.524s	0.766s	1.19s	1.83s	2.9s	4.04s	5.74s	8.06s	11.4s

Les fonctions orthogonales présentée dans la proposition 6.2.1 sont semblables à celles proposées par Kirby *et al.* [48] et Warburton [72], à la différence que nous utilisons des fonctions d'interpolation de Lagrange avec points de Gauss-Legendre à la place d'un polynôme de Legendre pour la partie en z. Pour les fonctions nodales, une tensorisation en z existe déjà, et donne un produit matrice-vecteur en $O(r^5)$ au lieu de $O(r^6)$, mais on peut obtenir un coût de $O(r^4)$ en utilisant des fonctions orthogonales tensorisées. Dans le tableau 9.6, le temps de calcul est donné pour la base nodale et la base orthogonale. Les résultats donnent un léger avantage à la base orthogonale à partir de l'ordre 5.

TAB. 9.6 – Temps de calcul pour 100 itérations avec un maillage de prismes non-affines comportant un million de degrés de liberté pour l'inconnue E_x.

Ordre	2	3	4	5	6	7	8
Nodale	296	312	239	342	343	353	400
Orthogonal	338	326	280	315	292	317	361

Comme pour les hexaèdres, on compare les bases de $\mathbb{P}_r$ et $\mathbb{W}_r$ dans le cas d'éléments affines. Dans le cas de $\mathbb{P}_r$, on prend les bases de la proposition 6.2.3, et celles de la proposition 6.2.1 pour $\mathbb{W}_r$. Le nombre de degrés de liberté et le temps de calcul nécessaire à l'obtention d'une erreur inférieure à 1% pour une cavité cubique dans les deux cas sont données dans les tableaux 9.7 et 9.8.

TAB. 9.7 – Pas de temps, nombre de degrés de liberté et temps de calcul nécessaires à l'obtention d'une erreur inférieure à 1% pour une cavité cubique (maillage avec des prismes affines) avec base orthogonale de $\mathbb{P}_r$.

Ordre	3	4	5	6	7	8	9
Δt	0.0268	0.0296	0.0309	0.029	0.0285	0.0268	0.0258
Ddls	351 520	172 955	96 768	84 000	61 440	56 595	47 520
Temps CPU	30 330s	12 920s	7 145s	7028s	5 560s	6 306s	6 602s

Dans ce cas, l'utilisation des bases de l'espace $\mathbb{W}_r$ semble bien meilleure, même si $\mathbb{P}_r$ comporte moins de degrés de liberté.

TAB. 9.8 – Pas de temps, nombre de degrés de liberté et temps de calcul nécessaires à l'obtention d'une erreur inférieure à 1% pour une cavité cubique (maillage avec des prismes affines) avec base orthogonale de $\mathbb{W}_r$.

Ordre	3	4	5	6	7	8	9
Δt	0.0277	0.0306	0.031	0.0298	0.0268	0.025	0.025
Ddls	425 920	205 800	126 000	100 352	98 784	87 480	68 750
Temps CPU	22 020s	6 834s	3 904s	3 340s	3 930s	4 397s	3 727s

9.2.4 Pyramides

Comme pour les hexaèdres et les prismes, on compare le temps de calcul pour construire et inverser la matrice de masse avec l'algorithme décrit dans la section 6.3.2 dans le cas de fonctions de base nodales et orthogonales pour un maillage composé de pyramides non-affines et comportant un million de degrés de liberté pour l'inconnue E_x. Les résultats sont données dans le tableau 9.9. Le temps de calcul pour les fonctions de base nodales croît très vite avec l'ordre utilisé, ce qui rend les fonctions orthogonales très attractives.

TAB. 9.9 – Temps de calcul de la matrice de masse avec un maillage de pyramides non-affines comportant un million de degrés de liberté pour l'inconnue E_x.

Ordre	2	3	4	5	6	7	8
Nodale	1.91s	7.98s	26.2s	73.6s	177s	391s	786s
Orthogonale	0.577s	0.98s	1.32s	2.27s	4.08s	5.17s	7.62s

Dans le tableau 9.10, on donne le gain de stockage de la matrice de masse et de sa factorisation de Cholesky lorsque l'on utilise la base orthogonale, par rapport au cas où l'on utilise une base nodale.

TAB. 9.10 – Gain de stockage pour la matrice de masse et sa factorisation de Cholesky en utilisant une base orthogonale à la place d'une base nodale pour des pyramides non-affines

Ordre	2	3	4	5	6	7	8
Gain pour la matrice	1.81	2.94	4.46	6.38	8.7	11.4	14.6
Gain pour la factorisation	1.52	1.91	2.39	2.83	3.38	3.85	4.8

On compare à présent différents solveurs permettant d'inverser la matrice de masse : un solveur plein, un solver creux et un solveur itératif. Comme le montre le tableau 9.11, le solveur creux surpasse les autres solveurs, notamment parce que le solveur itératif nécessite 10 itérations pour obtenir un résidu plus petit que 10^{-12}.

Dans le tableau 9.12, on compare ensuite le temps de calcul pour 100 itérations avec un maillage comportant un million de degrés de liberté pour l'inconnue E_x pour la base nodale, la base orthogonale et l'astuce de Warburton. L'astuce de warburton donne un résultat constant (entre 400 et 500s dans le cas présent) en fonction de l'ordre d'approximation. Les fonctions orthogonales donnent un coût de calcul inférieur à celui des fonctions nodales à partir de l'ordre 3, et est acceptable pour les ordres élevés. En revanche, le coût pour les fonctions nodales rend rapidement la méthode inutilisable pour les ordres élevés.

Comme pour les autres éléments, on compare les bases orthogonales de $\mathbb{P}_r$ et celles de $\mathbb{B}_r$ dans le cas d'éléments affines. Les résultats sont présentés dans les tableaux 9.13 et 9.14. On note que la base orthogonale de $\mathbb{P}_r$ dans la proposition 6.2.3 est la même que celle proposée par Kirby *et al.* [48].

Au vu des résultats, il semble préférable d'utiliser les fonctions de base de $\mathbb{B}_r$ plutôt que celles de $\mathbb{P}_r$ pour les pyramides affines.

9.2.5 Tétraèdres

Les tétraèdres étant affines, on utilise généralement des matrices de rigidité précalculées sur l'élément de référence. On compare donc ici les fonctions de base nodales pour lesquelles on effectue le produit matrice-vecteur en utilisant des matrices précalculées par une méthode décrite par Hesthaven et Warburton dans [46], et les fonctions de base orthogonales pour lesquelles on passe par des formules de quadrature (avec $(r+1)^3$ points). Le tableau 9.15 donne les temps de calcul pour ces deux méthodes à différents ordres d'approximation.

TAB. 9.11 – Comparaison de différents solveurs pour l'inversion de la matrice de masse avec des pyramides non-affines

Ordre	2	3	4	5	6	7	8
Solveur plein	91s	144s	183s	295s	340s	503s	652s
Solveur creux	74s	101s	121s	148s	178s	227s	235s
Solveur itératif	150s	150s	149s	331s	366s	403s	439s

TAB. 9.12 – Temps de calcul pour 100 itérations avec un maillage de pyramides non-affines comportant un million de degrés de liberté pour l'inconnue E_x.

Ordre	2	3	4	5	6	7	8
Nodale	378s	532s	702s	1135s	2425s	7618s	15350s
Orthogonale	523s	505s	508s	569s	619s	692s	766s
Astuce de Warburton	460s	432s	411s	451s	471s	494s	548s

TAB. 9.13 – Pas de temps, nombre de degrés de liberté et temps de calcul nécessaires à l'obtention d'une erreur inférieure à 1% pour une cavité cubique (maillage avec des pyramides affines) avec base orthogonale de $\mathbb{P}_r$.

Order	3	4	5	6	7	8	9
Δt	0.0238	0.0258	0.0248	0.024	0.0216	0.0205	0.0203
Dofs	960 000	461 370	336 000	258 048	246 960	213 840	165 000
Computation Time	60 449s	27 564s	22 422s	19 111s	21 433s	22 074s	19 056s

TAB. 9.14 – Pas de temps, nombre de degrés de liberté et temps de calcul nécessaires à l'obtention d'une erreur inférieure à 1% pour une cavité cubique (maillage avec des pyramides affines) avec base orthogonale de $\mathbb{B}_r$.

Ordre	3	4	5	6	7	8
Δt	0.0276	0.0307	0.028	0.0285	0.0268	0.0268
Ddls	737 280	330 000	279 552	181 440	153 000	109 440
Temps CPU	28 267s	10 898s	10 800s	7 246s	6 933s	5 204s

En outre, il est bien connu que l'utilisation de fonctions de base orthogonales n'est pas très attractive pour les ordres bas (voir Warburton [72]). Comme Warburton, on vérifie ainsi que les fonctions nodales sont plus efficaces lorsque $r < 10$.

TAB. 9.15 – Temps de calcul pour 100 itérations avec un maillage de tétraèdres comportant un million de degrés de liberté pour l'inconnue E_x.

Ordre	2	3	4	5	6	7	8	9	10
Nodale	379s	358s	327s	343s	428s	541s	680s	1023s	1751s
Orthogonale	527s	601s	559s	595s	679s	722s	798s	992s	1074s

Quatrième partie

Éléments finis d'arête pour une formulation $H(rot)$

Chapitre 10

Éléments finis d'ordre arbitrairement élevé

Comme dans le cas d'une formulation H^1, on construit des éléments finis (K, P_r^F, Σ_r) d'ordre r pour l'espace fonctionnel $H(rot)$. Nous cherchons là encore à obtenir un espace « optimal » au sens de la convergence en norme $H(rot)$, et des degrés de liberté permettant de conserver la continuité des composantes tangetielles. Nous donnons en particulier des fonctions de base hiérarchiques vérifiant la continuité des composantes tangentielles aux interfaces des éléments du maillage. L'accent sera porté plus particulièrement sur les pyramides, bien que l'espace d'approximation pour les hexaèdres et les prismes soit relativement nouveau.

Sommaire

10.1 Définition des éléments

Pour chaque type d'élément, on reprend la définition 2.1.1 du chapitre 2 pour les éléments K, $\hat{K}$ et la transformation F.

10.2 Espace d'approximation d'ordre r

10.2.1 Espace d'approximation optimal sur l'élément de référence

L'espace d'approximation V_h sur un ouvert Ω de $\mathbb{R}^3$ est donné par

$$V_h = \{u \in H(rot,\Omega) \mid u_{|K} \in P_r^F(K)\},$$

où P_r^F est l'espace d'approximation réel d'ordre r pour un élément K quelconque du maillage défini par

$$P_r^F(K) = \left\{u \mid DF^* u \circ F \in \left(\hat{P}_r(\hat{K})\right)^{n_s}\right\}$$

où DF est le jacobien de la transformation F.

L'objectif est de construire un espace d'approximation $\hat{P}_r$ permettant d'avoir une convergence optimale en norme $H(rot)$.

Définition 10.2.1 *On définit les espaces polynomiaux suivants (Nédélec [56])*

- *En 2D*

$$\begin{aligned} \mathcal{S}_r(x,y) &= \left\{u = (u_1,u_2) \in \left(\tilde{\mathbb{P}}_r\right)^2,\ u_1\, x + u_2\, y = 0\right\} \\ \mathcal{R}_r(x,y) &= \left(\mathbb{P}_{r-1}(x,y)\right)^2 \oplus \mathcal{S}_r(x,y) \end{aligned}$$

- *En 3D*

$$\begin{aligned} \mathcal{S}_r(x,y,z) &= \left\{u = (u_1,u_2,u_3) \in \left(\tilde{\mathbb{P}}_r\right)^3,\ u_1\, x + u_2\, y + u_3\, z = 0\right\} \\ \mathcal{R}_r(x,y,z) &= \left(\mathbb{P}_{r-1}(x,y,z)\right)^3 \oplus \mathcal{S}_r(x,y,z) \end{aligned}$$

et l'espace suivant

$$\mathbb{W}_{p,q}(x,y,z) = \mathbb{P}_p(x,y) \otimes \mathbb{P}_q(z)$$

On donne à présent une caractérisation et la dimension de l'espace $\mathcal{S}_r$

Propriété 10.2.2 *L'espace $\mathcal{S}_r$ est engendré par les familles suivantes*

- *En 2D*

$$\begin{bmatrix} \hat{x}^i \hat{y}^{j+1} \\ -\hat{x}^{i+1}\hat{y}^j \end{bmatrix}, \qquad \begin{matrix} 0 \leqslant i,j \leqslant r-1 \\ i+j = r-1 \end{matrix}$$

et

$$dim\ \mathcal{S}_r = r$$

- *En 3D*

$$\begin{bmatrix} x^{i-1} y^j z^{r-i-j+1} \\ 0 \\ -x^i y^j z^{r-i-j} \end{bmatrix}, \qquad \begin{matrix} i+j \leqslant r \\ 1 \leqslant i \leqslant r \\ 0 \leqslant j \leqslant r-1 \end{matrix}$$

$$\begin{bmatrix} 0 \\ x^i y^{j-1} z^{r-i-j+1} \\ -x^i y^j z^{r-i-j} \end{bmatrix}, \qquad \begin{matrix} i+j \leqslant r \\ 0 \leqslant i \leqslant r-1 \\ 1 \leqslant j \leqslant r \end{matrix}$$

$$\begin{bmatrix} x^{i-1} y^j \\ -x^i y^{j-1} \\ 0 \end{bmatrix}, \qquad \begin{matrix} i+j = r+1 \\ 1 \leqslant i,j \leqslant r \end{matrix}$$

et

$$dim\ \mathcal{S}_r = r(r+2)$$

Preuve.

- En 2D : Une base de $\tilde{\mathbb{P}}_r$ est $\left\{\hat{x}^i\hat{y}^j,\ i+j=r\right\}$. En prenant $u_1 = \hat{x}^i\hat{y}^j, i+j=r$, on a $u_2 = -\hat{x}^{i+1}\hat{y}^{j-1}$, soit, pour que $u_2 \in \tilde{\mathbb{P}}_r$, $j \neq 0$ et $i \neq r$. Réciproquement, en prenant $u_2 = \hat{x}^i\hat{y}^j, i+j=r$, on a $u_1 = -\hat{x}^{i-1}\hat{y}^{j+1}$, soit, pour que $u_1 \in \tilde{\mathbb{P}}_r$, $i \neq 0$ et $j \neq r$. Finalement, on peut écrire $u_1 = \hat{x}^i\hat{y}^{j+1}, i+j=r-1$ et $u_2 = -\hat{x}^{i+1}\hat{y}^j, i+j=r-1$.
  Le résultat sur la dimension est immédiat.

– En 3D : voir Bergot et Lacoste [6]. □

Définition 10.2.3 *Pour un élément K d'un maillage d'arête de longueur moyenneh, l'espace d'approximation P_r^F* ***optimal*** *est l'espace d'approximation de dimension minimale tel que $\mathcal{R}_r \subset P_r^F$.*

Theorème 10.2.4 *L'espace d'approximation P_r^F optimal pour un élément K du maillage permet, pour une solution suffisamment régulière, d'obtenir une erreur d'interpolation sur l'élément en $O(h^r)$ pour la norme $H(rot)$.*

Preuve. Voir Chapitre 11 sur les estimations d'erreur.

On cherche donc, pour chaque élément, l'espace optimal $\hat{P}_r$ sur l'élément de référence tel que l'on ait $\mathcal{R}_r \subset P_r^F$ sur l'élément du maillage, via la transformation F. Pour cette étude, on se base sur les travaux fondateurs de Nédélec [56] pour éléments finis $H(rot)$-conformes de la première famille.

Theorème 10.2.5 *L'espace d'approximation optimal $\hat{P}_r$ d'ordre r tel que l'on a $\mathcal{R}_r \subset P_r^F$ est*

– ***Tétraèdre et transformation F affine :***

$$\boxed{\hat{P}_r \;=\; \mathcal{R}_r(\hat{x},\hat{y},\hat{z})} \tag{10.2.1}$$

dont la dimension est

$$dim\ \mathcal{R}_r = \frac{r(r+2)(r+3)}{2}$$

– ***Hexaèdres :***

$$\boxed{\hat{P}_r = \mathbb{Q}_{r-1,r+1,r+1}(\hat{x},\hat{y},\hat{z}) \times \mathbb{Q}_{r+1,r-1,r+1}(\hat{x},\hat{y},\hat{z}) \times \mathbb{Q}_{r+1,r+1,r-1}(\hat{x},\hat{y},\hat{z})} \tag{10.2.2}$$

dont la dimension est

$$dim\ \mathbb{Q}_r(x,y,z) = 3r(r+2)^2$$

– ***Prismes :***

$$\boxed{\hat{P}_r = [\mathcal{R}_r(\hat{x},\hat{y}) \otimes \mathbb{P}_{r+1}(\hat{z})] \times \mathbb{W}_{r+1,r-1}(\hat{x},\hat{y},\hat{z})} \tag{10.2.3}$$

dont la dimension est

$$dim\ P_r(x,y,z) = \frac{r(r+2)(3r+7)}{2}$$

– ***Pyramides :***

$$\begin{aligned}
\hat{P}_r \;=\; & \mathbb{B}_{r-1}(\hat{x},\hat{y},\hat{z})^3 \;\oplus\; \left\{ \begin{bmatrix} \dfrac{\hat{x}^p\,\hat{y}^{p+1}}{(1-\hat{z})^{p+1}} \\ \dfrac{\hat{x}^{p+1}\,\hat{y}^p}{(1-\hat{z})^{p+1}} \\ \dfrac{\hat{x}^{p+1}\,\hat{y}^{p+1}}{(1-\hat{z})^{p+2}} \end{bmatrix},\ 0 \leqslant p \leqslant r-1 \right\} \\
& \oplus \left\{ \begin{bmatrix} \dfrac{\hat{x}^m\,\hat{y}^{n+2}}{(1-\hat{z})^{m+1}} \\ 0 \\ \dfrac{\hat{x}^{m+1}\,\hat{y}^{n+2}}{(1-\hat{z})^{m+2}} \end{bmatrix} \oplus \begin{bmatrix} 0 \\ \dfrac{\hat{x}^{n+2}\,\hat{y}^m}{(1-\hat{z})^{m+1}} \\ \dfrac{\hat{x}^{n+2}\,\hat{y}^{m+1}}{(1-\hat{z})^{m+2}} \end{bmatrix},\ 0 \leqslant m \leqslant n \leqslant r-2 \right\} \\
& \oplus \left\{ \begin{bmatrix} \dfrac{\hat{x}^p\,\hat{y}^q}{(1-\hat{z})^{p+q-r}} \\ 0 \\ \dfrac{\hat{x}^{p+1}\,\hat{y}^q}{(1-\hat{z})^{p+q+1-r}} \end{bmatrix} \oplus \begin{bmatrix} 0 \\ \dfrac{\hat{x}^q\,\hat{y}^p}{(1-\hat{z})^{p+q-r}} \\ \dfrac{\hat{x}^q\,\hat{y}^{p+1}}{(1-\hat{z})^{p+q+1-r}} \end{bmatrix},\ \begin{matrix} 0 \leqslant p \leqslant r-1 \\ 0 \leqslant q \leqslant r+1 \end{matrix} \right\}
\end{aligned} \tag{10.2.4}$$

dont la dimension est

$$dim\ \hat{P}_r = \frac{r(r+3)(2r+3)}{2}$$

Avant de prouver ce théorème, définissons les notations pour les espaces $\hat{P}_r$ des différents types d'éléments

Définition 10.2.6 *On note*
- ***Tétraèdres*** *:* $\hat{P}_r = \mathcal{R}_r(\hat{x},\hat{y},\hat{z})$
- ***Hexaèdres*** *:* $\hat{P}_r = \mathcal{Q}_r(\hat{x},\hat{y},\hat{z})$
- ***Prisme*** *:* $\hat{P}_r = \mathcal{W}_r(\hat{x},\hat{y},\hat{z})$
- ***Pyramide*** *:* $\hat{P}_r = \mathcal{B}_r(\hat{x},\hat{y},\hat{z})$

Preuve.
- Lorsque F est affine, il est immédiat que
$$\hat{P}_r(\hat{K}) = \mathcal{R}_r(\hat{K}) \Longleftrightarrow P_r^F(K) = \mathcal{R}_r(K).$$
- Lorsque l'élément n'est pas affine, voire la preuve de la proposition 10.2.9.

En ce qui concerne les dimensions,
- **Tétraèdres** : Voir Bergot et Lacoste [6].
- **Hexaèdres** : Le résultat est immédiat en utilisant le résultat classique
$$\dim \mathbb{Q}_{p,q,s}(x,y,z) = (p+1)(q+1)(s+1),$$
- **Prisme** : D'après la propriété 10.2.2
$$\dim \mathcal{S}_r(\hat{x},\hat{y}) \otimes \mathbb{P}_{r+1}(\hat{z}) = r(r+2)^2.$$
Comme
$$\dim \mathbb{W}_{p,q}(x,y,z) = \dim \mathbb{P}_p(x,y) \dim \mathbb{P}_q(z) = (q+1)\,\frac{(p+1)(p+2)}{2},$$
on a
$$\dim \mathcal{W}_r(x,y,z) = r(r+2)^2 + r\,\frac{(r+2)(r+3)}{2}$$
d'où le résultat.
- **Pyramide** : La dimension des cinq familles est $2\dfrac{r(r-1)}{2} + 2r(r+2) + r$, d'où
$$\dim \mathcal{B}_r(x,y,z) = 3\,\frac{r(r+1)(2r+1)}{6} + 2\frac{r(r-1)}{2} + 2r(r+2) + r = \frac{r(r+3)(2r+3)}{2}$$
ce qui achève la démonstration. □

10.2.2 Espace d'approximation optimal sur le cube symétrique

Comme dans le cas continu, on peut considérer les différents éléments comme un cube dégénéré grâce à la transformation T de la définition 2.2.7. En ce qui concerne les pyramides, pour encore plus de simplicité, on considère la transformation $\breve{T}$ suivante

Définition 10.2.7 *La transformation permettant de passer du cube symétrique* $\breve{Q}(\breve{x},\breve{y},\breve{z}) = [-1,1]^2 \times [0,1]$ *à la pyramide de référence* $\hat{K}$ *est*
$$\breve{T} : \begin{cases} \hat{x} = \breve{x}\,(1-\breve{z}) \\ \hat{y} = \breve{y}\,(1-\breve{z}) \\ \hat{z} = \breve{z}. \end{cases} \tag{10.2.5}$$

Ce changement de variable définit un difféomorphisme de l'ouvert $\breve{Q}$ vers l'ouvert $\hat{K}$, et pour toute fonction f, on note
$$\breve{f}(\breve{x},\breve{y},\breve{z}) = \hat{f}(\hat{x},\hat{y},\hat{z}),$$

Remarque 10.2.8 *Tout ce qui est valable pour* T *est valable pour* $\breve{T}$*, en particulier*
$$C_r = \hat{P}_r \circ T = \hat{P}_r \circ \breve{T}$$

On remarque qu'il s'agit bien ici d'un changement de variable et non d'une transformation $H(rot)$-conforme puisque l'on n'ajoute pas le DT^* dans le changement de variable. Cela signifie notamment que l'on est virtuellement resté sur l'élément en gardant notamment l'équation des faces et les directions de dérivation sur les éléments, ce qui est important pour vérifier les restrictions sur les faces.

On peut ainsi écrire les espaces d'approximation $\hat{P}_r$ sur le cube unité après transformation par T, ou sur le cube symétrique $\breve{Q}$ après transformation par $\breve{T}$.

Proposition 10.2.9 *L'espace optimal d'approximation C_r d'ordre r sur le cube unité $\widetilde{Q}$ ou sur le cube symétrique $\breve{Q}$ est*

- ***Tétraèdre et transformation F affine :***

$$C_r = \mathcal{R}_r(\hat{x},\hat{y},\hat{z}) \circ T = \mathcal{R}_r(\widetilde{x}\,(1-\widetilde{y})\,(1-\widetilde{z}), \widetilde{y}\,(1-\widetilde{z}), \widetilde{z})$$

- ***Hexaèdres :***

$$C_r = \mathcal{Q}_r(\widetilde{x},\widetilde{y},\widetilde{z})$$

- ***Prismes :***

$$C_r = \mathcal{W}_r(\hat{x},\hat{y},\hat{z}) \circ T = \mathcal{W}_r(\widetilde{x}\,(1-\widetilde{y}), \widetilde{y}, \widetilde{z})$$

- ***Pyramides :***

$$\begin{aligned}
C_r &= \mathcal{B}_r(\hat{x},\hat{y},\hat{z}) \circ \breve{T} \\
&= \left(\mathbb{B}_{r-1} \circ \breve{T}(\breve{x},\breve{y},\breve{z})\right)^3 \oplus \left\{ \begin{bmatrix} \breve{x}^p \breve{y}^{p+1}(1-\breve{z})^p \\ \breve{x}^{p+1}\breve{y}^p(1-\breve{z})^p \\ \breve{x}^{p+1}\breve{y}^{p+1}(1-\breve{z})^p \end{bmatrix}, \ 0 \leqslant p \leqslant r-1 \right\} \\
&\oplus \left\{ \begin{bmatrix} \breve{x}^m \breve{y}^{n+2}(1-\breve{z})^{n+1} \\ 0 \\ \breve{x}^{m+1}\breve{y}^{n+2}(1-\breve{z})^{n+1} \end{bmatrix} \oplus \begin{bmatrix} 0 \\ \breve{x}^{n+2}\breve{y}^m(1-\breve{z})^{n+1} \\ \breve{x}^{n+2}\breve{y}^{m+1}(1-\breve{z})^{n+1} \end{bmatrix}, \ 0 \leqslant m \leqslant n \leqslant r-2 \right\} \\
&\oplus \left\{ \begin{bmatrix} \breve{x}^p \breve{y}^q(1-z)^r \\ 0 \\ \breve{x}^{p+1}\breve{y}^q(1-\breve{z})^r \end{bmatrix} \oplus \begin{bmatrix} 0 \\ \breve{x}^q \breve{y}^p(1-\breve{z})^r \\ \breve{x}^q \breve{y}^{p+1}(1-\breve{z})^r \end{bmatrix}, \ \begin{matrix} 0 \leqslant p \leqslant r-1 \\ 0 \leqslant q \leqslant r+1 \end{matrix} \right\}
\end{aligned}$$

Remarque 10.2.10 *L'espace $\mathbb{B}_{r-1} \circ \breve{T}(\breve{x},\breve{y},\breve{z})$ est l'espace C_{r-1} du cas H^1.*

Preuve. L'égalité entre $\hat{P}_r(\hat{x},\hat{y},\hat{z}) \circ T$ et l'espace C_r correspondant proposé pour chaque type d'élément est immédiate par l'application de la transformation.

Reste à prouver l'optimalité de espaces C_r, et donc des espaces $\hat{P}_r$ lorsque l'élément n'est pas affine. La preuve dans le cas des hexaèdres d'ordre 1 est donnée par Falk, Gatto et Monk dans [32]. À l'ordre quelconque, on détaille le cas des pyramides, les autres éléments étant un peu plus simples à traiter. On considère pour cela un monôme $p(x,y,z)$ de $\mathcal{R}_r$, on calcule $\hat{p} = p \circ F$, on applique ensuite DF^* à $\hat{p}$ et on regarde enfin l'ensemble des monômes obtenus.

D'après la proposition 2.1.4, la transformation F de la pyramide peut s'écrire sous la forme

$$F(\hat{x},\hat{y},\hat{z}) = A_0 + A_1\hat{x} + A_2\hat{y} + A_3\hat{z} + \frac{\hat{x}\hat{y}}{1-\hat{z}}C,$$

où A_1, A_2, A_3 et C appartiennent à $\mathbb{R}^3$ et dépendent de la géométrie. Sur le cube symétrique $\breve{Q}$, la transformation F s'écrit alors

$$F(\breve{x},\breve{y},\breve{z}) = A_0 + A_1\breve{x}(1-\breve{z}) + A_2\breve{y}(1-\breve{z}) + A_3\breve{z} + \breve{x}\breve{y}(1-\breve{z})C$$

Les dérivées de F valent

$$\begin{aligned}
\frac{\partial F}{\partial \hat{x}} &= A_1 + \frac{\hat{y}}{1-\hat{z}}C = A_1 + C\breve{y} \\
\frac{\partial F}{\partial \hat{y}} &= A_2 + \frac{\hat{x}}{1-\hat{z}}C = A_2 + C\breve{x} \\
\frac{\partial F}{\partial \hat{z}} &= A_3 + \frac{\hat{x}\hat{y}}{(1-\hat{z})^2}C = A_3 + C\breve{x}\breve{y}
\end{aligned}$$

Lorsqu'on applique DF^* à un champ $\breve{p} \in \mathbb{R}^3$, on calcule

$$D\breve{F}^*\breve{p} = \begin{vmatrix} (A_1 + C\breve{y}) \cdot \breve{p} \\ (A_2 + C\breve{x}) \cdot \breve{p} \\ (A_3 + C\breve{x}\breve{y}) \cdot \breve{p} \end{vmatrix}$$

Pour $p \in \mathbb{P}^3_{r-1}(K)$, d'après le théorème 2.2.3 et la proposition 2.2.8, il est équivalent de dire que $\breve{p} \in C^3_{r-1}(\breve{Q})$. On considère un monôme $\breve{p}$ de C^3_{r-1}

$$\breve{p} = \breve{x}^i \breve{y}^j (1-\breve{z})^k E_0$$

avec $0 \leqslant i,j \leqslant k \leqslant r-1$ et E_0 vecteur constant de $\mathbb{R}^3$. On a donc

$$D\breve{F}^{*}\breve{p} = \left| \begin{array}{l} A_1 \cdot E_0\, \breve{x}^i \breve{y}^j (1-\breve{z})^k + C \cdot E_0\, \breve{x}^i \breve{y}^{j+1} (1-\breve{z})^k \\ A_2 \cdot E_0\, \breve{x}^i \breve{y}^j (1-\breve{z})^k + C \cdot E_0\, \breve{x}^{i+1} \breve{y}^j (1-\breve{z})^k \\ A_3 \cdot E_0\, \breve{x}^i \breve{y}^j (1-\breve{z})^k + C \cdot E_0\, \breve{x}^{i+1} \breve{y}^{j+1} (1-\breve{z})^k \end{array} \right.$$

On a

$$\left\{ \left| \begin{array}{l} A_1 \cdot E_0\, \breve{x}^i \breve{y}^j (1-\breve{z})^k \\ A_2 \cdot E_0\, \breve{x}^i \breve{y}^j (1-\breve{z})^k \\ A_3 \cdot E_0\, \breve{x}^i \breve{y}^j (1-\breve{z})^k \end{array} \right. , \quad 0 \leqslant i,j \leqslant k \leqslant r-1 \right\} = C^3_{r-1}$$

ce qui signifie qu'on a besoin de toutes les fonctions de C^3_{r-1}.

Concernant la partie en C, on distingue quatre cas

- $i = j = k = p-1$, avec $1 \leqslant p \leqslant r$, on obtient alors

$$\left| \begin{array}{l} \breve{x}^{p-1} \breve{y}^p (1-\breve{z})^{p-1} \\ \breve{x}^p \breve{y}^{p-1} (1-\breve{z})^{p-1} \\ \breve{x}^p \breve{y}^p (1-\breve{z})^{p-1} \end{array} \right. \in \mathbb{B}_r$$

- $i = k = m+1$ et $j < k$, donc $j \leqslant m$ et $m \leqslant r-2$. On a :

$$\underbrace{\left| \begin{array}{l} 0 \\ \breve{x}^{m+2} \breve{y}^j (1-\breve{z})^{m+1} \\ \breve{x}^{m+2} \breve{y}^{j+1} (1-\breve{z})^{m+1} \end{array} \right.}_{\in \mathbb{B}_r} + \underbrace{\left| \begin{array}{l} \breve{x}^{m+1} \breve{y}^{j+1} (1-\breve{z})^{m+1} \\ 0 \\ 0 \end{array} \right.}_{\in C^3_{r-1}}$$

- $j = k = m+1$ et $i = p < k$, alors $p \leqslant m$ et $m \leqslant r-2$. On a

$$\underbrace{\left| \begin{array}{l} \breve{x}^p \breve{y}^{m+2} (1-\breve{z})^{m+1} \\ 0 \\ \breve{x}^{p+1} \breve{y}^{m+2} (1-\breve{z})^{m+1} \end{array} \right.}_{\in \mathbb{B}_r} + \underbrace{\left| \begin{array}{l} 0 \\ \breve{x}^{p+1} \breve{y}^{m+1} (1-\breve{z})^{m+1} \\ 0 \end{array} \right.}_{\in C^3_{r-1}}$$

- $i,j < k$, dans ce cas $DF^{*}E \in C^3_{r-1}$

On remarque donc que pour générer $\mathbb{P}^3_{r-1}$, on a de manière nécessaire et suffisante l'espace suivant

$$C^3_{r-1} \oplus \left| \begin{array}{l} \breve{x}^j \breve{y}^{k+2} (1-\breve{z})^{k+1} \\ 0 \\ \breve{x}^{j+1} \breve{y}^{k+2} (1-\breve{z})^{k+1} \end{array} \right. \oplus \left| \begin{array}{l} 0 \\ \breve{x}^{k+2} \breve{y}^j (1-\breve{z})^{k+1} \\ \breve{x}^{k+2} \breve{y}^{j+1} (1-\breve{z})^{k+1} \end{array} \right. \oplus \left| \begin{array}{l} \breve{x}^{p-1} \breve{y}^p (1-\breve{z})^{p-1} \\ \breve{x}^p \breve{y}^{p-1} (1-\breve{z})^{p-1} \\ \breve{x}^p \breve{y}^p (1-\breve{z})^{p-1}, \end{array} \right. \quad j \leqslant k \leqslant r-2,\ 1 \leqslant p \leqslant r$$

Cet espace est à considérer si l'on veut construire des éléments finis de la seconde famille de Nédélec ($\mathbb{P}^3_r$ pour les tétraèdres).

On considère à présent une fonction de $\mathcal{S}_r$

$$\left| \begin{array}{l} 0 \\ x^i y^j z^{k+1} \\ -x^i y^{j+1} z^k \end{array} \right.$$

avec $i+j \leqslant r-1$ et $k = r-1-i-j$.

Le degré de $x^i y^j z^k$ est exactement $r-1$, donc ce monôme s'écrit comme une combinaison linéaire de $\breve{x}^i \breve{y}^j (1-\breve{z})^{r-1}$ avec $i,j \leq r-1$. Pour simplifier, les calculs, il est équivalent de considérer une fonction de S_r de la forme

$$\breve{x}^i \breve{y}^j (1-\breve{z})^{r-1} \left| \begin{array}{l} 0 \\ z \\ -y \end{array} \right.$$

pour $i,j \leqslant r-1$.

On pose

$$E = \left| \begin{array}{l} 0 \\ z \\ -y \end{array} \right. = \left| \begin{array}{l} 0 \\ (A^z_0 - A^z_3) + \breve{x}(1-\breve{z}) A^z_1 + \breve{y}(1-\breve{z}) A^z_2 - (1-\breve{z}) A^z_3 + \breve{x}\breve{y}(1-\breve{z}) C^z \\ -[(A^y_0 - A^y_3) + \breve{x}(1-\breve{z}) A^y_1 + \breve{y}(1-\breve{z}) A^y_2 - (1-\breve{z}) A^y_3 + \breve{x}\breve{y}(1-\breve{z}) C^y] \end{array} \right.$$

avec

$$\begin{aligned} A_0 &= (A_0^x, A_0^y, A_0^z) \\ A_1 &= (A_1^x, A_1^y, A_1^z) \\ A_2 &= (A_2^x, A_2^y, A_2^z) \\ A_3 &= (A_3^x, A_3^y, A_3^z) \\ C &= (C^x, C^y, C^z) \end{aligned}$$

La première composante de DF^*E s'écrit alors

$$(DF^*E)_x = (A_1 + C\breve{y}) \cdot E = \begin{vmatrix} A_1^x + C^x\breve{y} \\ A_1^y + C^y\breve{y} \\ A_1^z + C^z\breve{y} \end{vmatrix} \cdot \begin{vmatrix} 0 \\ (A_0^z - A_3^z) + \breve{y}(1-\breve{z})A_2^z - (1-\breve{z})A_3^z + \breve{x}(1-\breve{z})(A_1^z + C^z\breve{y}) \\ -[(A_0^y - A_3^y) + \breve{y}(1-\breve{z})A_2^y - (1-\breve{z})A_3^y + \breve{x}(1-\breve{z})(A_1^y + C^y\breve{y})] \end{vmatrix}$$

soit

$$\begin{aligned} (DF^*E)_x &= [A_1^y(A_0^z - A_3^z) - A_1^z(A_0^y - A_3^y)] + [C^y(A_0^z - A_3^z) - C^z(A_0^y - A_3^y)]\breve{y} + [A_1^yA_2^z - A_1^zA_2^y]\breve{y}(1-\breve{z}) \\ &\quad + [C^yA_2^z - C^zA_2^y]\breve{y}^2(1-\breve{z}) + [A_1^zA_3^y - A_1^yA_3^z](1-\breve{z}) + [A_3^yC^z - A_3^zC^y]\breve{y}(1-\breve{z}) \\ &= b_0 + b_1\breve{y} + (b_2+b_3)\breve{y}(1-\breve{z}) + b_4(1-\breve{z}) + b_5\breve{y}^2(1-\breve{z}) \end{aligned}$$

où

$$\begin{aligned} b_0 &= A_1^y(A_0^z - A_3^z) - A_1^z(A_0^y - A_3^y) \\ b_1 &= C^y(A_0^z - A_3^z) - C^z(A_0^y - A_3^y) \\ b_2 &= A_1^yA_2^z - A_1^zA_2^y \\ b_3 &= A_3^yC^z - A_3^zC^y \\ b_4 &= A_1^zA_3^y - A_1^yA_3^z \\ b_5 &= C^yA_2^z - C^zA_2^y \end{aligned}$$

La deuxième composante de DF^*E s'écrit

$$(DF^*E)_y = (A_2 + C\breve{x}) \cdot E = \begin{vmatrix} A_2^x + C^x\breve{x} \\ A_2^y + C^y\breve{x} \\ A_2^z + C^z\breve{x} \end{vmatrix} \cdot \begin{vmatrix} 0 \\ (A_0^z - A_3^z) + \breve{x}(1-\breve{z})A_1^z - (1-\breve{z})A_3^z + \breve{y}(1-\breve{z})(A_2^z + C^z\breve{x}) \\ -[(A_0^y - A_3^y) + \breve{x}(1-\breve{z})A_1^y - (1-\breve{z})A_3^y + \breve{y}(1-\breve{z})(A_2^y + C^y\breve{x})] \end{vmatrix}$$

i.e.

$$\begin{aligned} (DF^*E)_y &= [A_2^y(A_0^z - A_3^z) - A_2^z(A_0^y - A_3^y)] + [C^y(A_0^z - A_3^z) - C^z(A_0^y - A_3^y)]\breve{x} + [A_2^yA_1^z - A_2^zA_1^y]\breve{x}(1-\breve{z}) \\ &\quad + [C^yA_1^z - C^zA_1^y]\breve{x}^2(1-\breve{z} + [A_3^yA_2^z - A_3^zA_2^y](1-\breve{z}) + [A_3^yC^z - A_3^zC^y]\breve{x}(1-\breve{z}) \\ &= b_6 + b_1\breve{x} + (b_3 - b_2)\breve{x}(1-\breve{z}) + b_8(1-\breve{z}) + b_7\breve{x}^2(1-\breve{z}) \end{aligned}$$

où

$$\begin{aligned} b_6 &= A_2^y(A_0^z - A_3^z) - A_2^z(A_0^y - A_3^y) \\ b_7 &= C^yA_1^z - C^zA_1^y \\ b_8 &= A_3^yA_2^z - A_3^zA_2^y \end{aligned}$$

On calcule la dernière composante

$$(DF^*E)_z = (A_3 + C\breve{x}\breve{y}) \cdot E = \begin{vmatrix} A_3^x + C^x\breve{x}\breve{y} \\ A_3^y + C^y\breve{x}\breve{y} \\ A_3^z + C^z\breve{x}\breve{y} \end{vmatrix} \cdot \begin{vmatrix} 0 \\ (A_0^z - A_3^z) + \breve{x}(1-\breve{z})A_1^z + \breve{y}(1-\breve{z})A_2^z - (1-\breve{z})A_3^z + \breve{x}\breve{y}(1-\breve{z})C^z \\ -[(A_0^y - A_3^y) + \breve{x}(1-\breve{z})A_1^y + \breve{y}(1-\breve{z})A_2^y - (1-\breve{z})A_3^y + \breve{x}\breve{y}(1-\breve{z})C^y] \end{vmatrix}$$

soit

$$\begin{aligned} (DF^*E)_z &= [A_3^yA_0^z - A_3^zA_0^y] + [C^y(A_0^z - A_3^z) - C^z(A_0^y - A_3^y)]\breve{x}\breve{y} + [A_1^zA_3^y - A_1^yA_3^z]\breve{x}(1-\breve{z}) \\ &\quad + [C^yA_1^z - C^zA_1^y]\breve{x}^2\breve{y}(1-\breve{z}) + [A_3^yA_2^z - A_3^zA_2^y]\breve{y}(1-\breve{z}) + [C^yA_2^z - C^zA_2^y]\breve{x}\breve{y}^2(1-\breve{z}) \\ &\quad + 2[A_3^yC^z - A_3^zC^y]\breve{x}\breve{y}(1-\breve{z}) \\ &= b_9 + b_1\breve{x}\breve{y} + b_4\breve{x}(1-\breve{z}) + b_7\breve{x}^2(1-\breve{z}) + b_5\breve{x}\breve{y}^2(1-\breve{z}) \\ &\quad + (b_3 - b_2)\breve{x}\breve{y}(1-\breve{z}) + (b_3 + b_2)\breve{x}\breve{y}(1-\breve{z}) + b_8\breve{y}(1-\breve{z}) \end{aligned}$$

avec

$$b_9 = A_3^yA_0^z - A_3^zA_0^y$$

On a donc

$$DF^*E = \begin{vmatrix} b_0 \\ b_6 \\ b_9 \end{vmatrix} + b_1 \begin{vmatrix} \check{y} \\ \check{x} \\ \check{x}\check{y} \end{vmatrix} + (b_2+b_3) \begin{vmatrix} \check{y}(1-\check{z}) \\ 0 \\ \check{x}\check{y}(1-\check{z}) \end{vmatrix} + (b_3-b_2) \begin{vmatrix} 0 \\ \check{x}(1-\check{z}) \\ \check{x}\check{y}(1-\check{z}) \end{vmatrix} + b_4 \begin{vmatrix} 1-\check{z} \\ 0 \\ \check{x}(1-\check{z}) \end{vmatrix}$$
$$+b_5 \begin{vmatrix} \check{y}^2(1-\check{z}) \\ 0 \\ \check{x}\check{y}^2(1-\check{z}) \end{vmatrix} + b_7 \begin{vmatrix} 0 \\ \check{x}^2(1-\check{z}) \\ \check{x}^2\check{y}(1-\check{z}) \end{vmatrix} + b_8 \begin{vmatrix} 0 \\ (1-\check{z}) \\ \check{y}(1-\check{z}) \end{vmatrix}$$

La fonction

$$\check{x}^i\check{y}^j(1-\check{z})^{r-1} \begin{vmatrix} b_0 \\ b_6 \\ b_9 \end{vmatrix}$$

appartient à C^3_{r-1}.

Les coefficients $b_1, (b_2+b_3), (b_3-b_2), b_4, b_5, b_7$ et b_8, vus comme des fonctions de A_0, A_1, A_2, A_3, C, sont linéairement indépendants, on a donc de manière nécessaire et suffisante les fonctions suivantes

$$\check{x}^i\check{y}^j(1-\check{z})^{r-1} \begin{vmatrix} \check{y} \\ \check{x} \\ \check{x}\check{y} \end{vmatrix}, \check{x}^i\check{y}^j(1-\check{z})^{r-1} \begin{vmatrix} \check{y}(1-\check{z}) \\ 0 \\ \check{x}\check{y}(1-\check{z}) \end{vmatrix}, \check{x}^i\check{y}^j(1-\check{z})^{r-1} \begin{vmatrix} 1-\check{z} \\ 0 \\ \check{x}(1-\check{z}) \end{vmatrix}, \check{x}^i\check{y}^j(1-\check{z})^{r-1} \begin{vmatrix} \check{y}^2(1-\check{z}) \\ 0 \\ \check{x}\check{y}^2(1-\check{z}) \end{vmatrix}$$
$$\check{x}^i\check{y}^j(1-\check{z})^{r-1} \begin{vmatrix} 0 \\ \check{x}(1-\check{z}) \\ \check{x}\check{y}(1-\check{z}) \end{vmatrix}, \check{x}^i\check{y}^j(1-\check{z})^{r-1} \begin{vmatrix} 0 \\ \check{x}^2(1-\check{z}) \\ \check{x}^2\check{y}(1-\check{z}) \end{vmatrix}, \check{x}^i\check{y}^j(1-\check{z})^{r-1} \begin{vmatrix} 0 \\ (1-\check{z}) \\ \check{y}(1-\check{z}) \end{vmatrix}$$

que l'on peut regrouper suivant trois groupes

- Un premier groupe

$$\begin{vmatrix} \check{x}^i\check{y}^{j+1}(1-\check{z})^{r-1} \\ \check{x}^{i+1}\check{y}^j(1-\check{z})^{r-1} \\ \check{x}^{i+1}\check{y}^{j+1}(1-\check{z})^{r-1} \end{vmatrix}, 0 \leqslant i,j \leq r-1$$

  qui est le même que celui rencontré lors du traitement de $\mathbb{P}^3_{r-1}$ (avec $k = r-1$)

- Deux autres groupes

$$\begin{vmatrix} \check{x}^i\check{y}^j(1-\check{z})^r \\ 0 \\ \check{x}^{i+1}\check{y}^j(1-\check{z})^r \end{vmatrix}, \quad \begin{vmatrix} 0 \\ \check{x}^j\check{y}^i(1-\check{z})^r \\ \check{x}^j\check{y}^{i+1}(1-\check{z})^r \end{vmatrix}, \quad i \leq r-1, j \leq r+1$$

  qui sont les deux dernières familles de $\mathcal{B}_r$.

Lorsqu'on effectue les calculs pour les deux autres familles de fonctions de S_r

$$\begin{vmatrix} x^i y^j z^{k+1} \\ 0 \\ -x^{i+1}y^j z^k \end{vmatrix}, \quad \begin{vmatrix} x^i y^{j+1} z^k \\ -x^{i+1}y^j z^k \\ 0 \end{vmatrix}, \quad k = r-1-i-j,$$

on obtient alors exactement la même forme pour DF^*E, seule l'expression des coefficients $b_0, b_1, ..., b_9$ étant différente.

On a donc montré que pour générer $\mathcal{R}_r$, toutes les fonctions de $\mathcal{B}_r$ étaient nécessaires et suffisantes. □

Remarque 10.2.11 *Comme dans le cas de l'espace optimal H^1, nous avons identifié les coefficients indépendants pour les espaces des trois premiers ordres avant d'établir une conjecture sur la forme de l'espace pour tout ordre.*

La preuve pour l'espace des hexaèdres d'ordre 1 est détaillée dans [32].

10.3 Degrés de liberté et fonctions de base

10.3.1 Éléments finis nodaux

10.3.1.1 Localisation des degrés de liberté

Comme dans le cas H^1, on souhaite lier les éléments entre eux de sorte que les composantes tangentielles de chaque fonction soient continues à l'interface. Bien que peu utilisée pour les éléments finis d'arête, on peut considérer une approche nodale comme l'ont étudié Cohen et Monk [21] sur les hexaèdres.

On utilisera les degrés de liberté suivants.

Hexaèdre : Comme le montre la figure 10.1 pour l'hexaèdre d'ordre 1, on place dans chaque direction e_i, un degré de liberté sur chaque point issu du produit tensoriel entre les points de Gauss-Lobatto intérieurs d'ordre $r+1$ selon e_i et les points de Gauss-Lobatto d'ordre $r+1$ selon les deux autres directions.
On a au total $3r(r+2)^2$ degrés de liberté, ce qui est précisément la dimension de $\mathcal{Q}_r$.

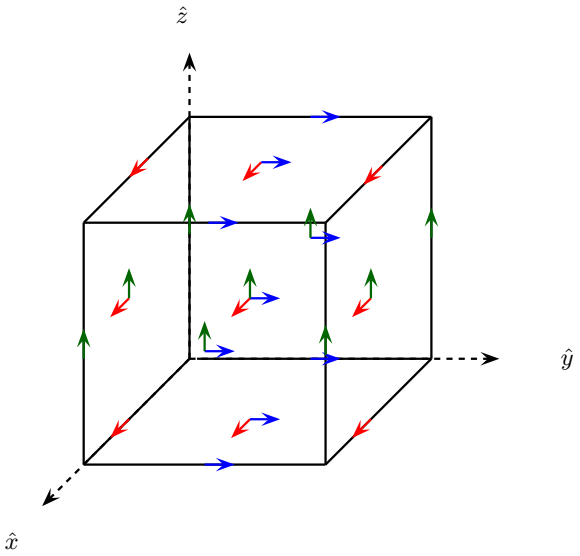

FIG. 10.1 – Localisation des degrés de liberté pour l'élément hexaédrique d'ordre 1

Tétraèdres : Comme le montre la figure 10.2 pour le tétraèdre d'ordre 1, on place les degrés de liberté comme suit

- Pour les arêtes, on place un degré de liberté sur chaque point de Gauss-Lobatto intérieur d'ordre $r+1$
- Pour les faces, on place deux degrés de liberté orientés selon une base de la face, sur les points intérieurs du triangle d'Hesthaven (Hesthaven [43]) d'ordre $r+1$
- Pour le volume, on prend les points intérieurs d'un tétraèdre de Hesthaven (Hesthaven et Teng [44]) d'ordre $r+1$, et on place trois degrés de liberté suivant une base de l'espace.

Il y a finalement $6r + 4r(r-1) + \dfrac{r(r-1)(r-2)}{2}$ degrés de liberté, qui est la dimension de $\mathcal{R}_r$.

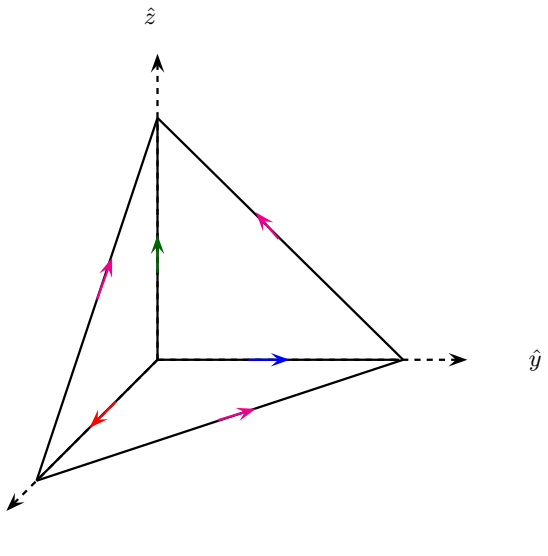

FIG. 10.2 – Localisation des degrés de liberté pour l'élément tétraédrique d'ordre 1

Prismes : Comme le montre la figure 10.3 pour le prisme d'ordre 1, on place les degrés de liberté selon $(\hat{x}, \hat{y})$ en prenant le produit tensoriel d'une face triangulaire identique à celle d'un tétraèdre, et une arête de Gauss-Lobatto d'ordre $r+1$. Pour les degrés de liberté selon $\hat{z}$, ils sont placés sur les points issus du produit tensoriel d'un triangle H^1 d'ordre $r+1$ par une arête avec points de Gauss-Lobatto d'ordre $r+1$ intérieurs. On retrouve ainsi les mêmes faces quadrangulaires que sur l'hexaèdre.

On obtient donc $r(r+2)(r+2) + r\dfrac{(r+2)(r+3)}{2}$ degrés de liberté, soit autant que la dimension de $\mathcal{W}_r$

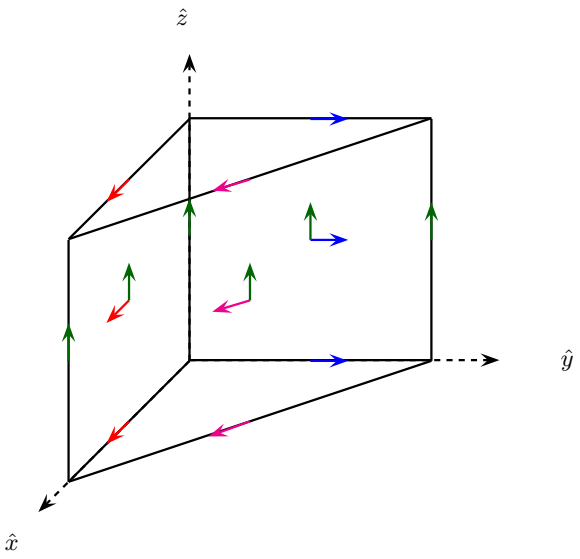

FIG. 10.3 – Localisation des degrés de liberté pour l'élément prismatique d'ordre 1

Pyramides : Comme le montre la figure 10.4 pour la pyramide d'ordre 1, on place les degrés de liberté comme pour les faces de tétraèdre sur les faces triangulaires, comme pour les faces de l'hexaèdre sur la base, et on place 3 degrés de liberté par point intérieur, placés comme les points intérieurs d'une pyramide H^1 d'ordre $r+1$.

Il a y donc $4r(r+2) + 2r(r-2) + \dfrac{r(r-1)(2r-1)}{2}$ degrés de liberté, qui est bien la dimension de $\mathcal{B}_r$.

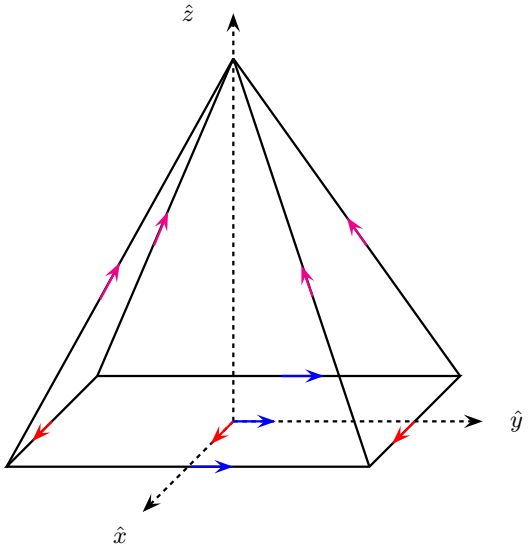

FIG. 10.4 – Localisation des degrés de liberté pour l'élément pyramidal d'ordre 1

10.3.1.2 Fonctions de base

Comme pour le cas continu, on cherche la base nodale de chaque espace $\hat{P}_r$ sur les éléments de référence $\hat{K}$, les fonctions réelles sur K étant déduites de celles-ci à l'aide de la transformation $H(rot)$-conforme

$$\hat{p} = DF^* p \circ F. \tag{10.3.1}$$

Les fonctions de base $(\hat{\varphi}_i)$ sur l'élément de référence $\hat{K}$ sont obtenues comme suit.

Définition 10.3.1 *Soit $(\hat{M}_i)_{1\leqslant i\leqslant n_r}$ les coordonnées des points d'interpolation sur l'élément $\hat{K}$ **munis d'une orientation selon un vecteur** $\hat{t}_i$, et $(\hat{\psi}_i)_{1\leqslant i\leqslant n_r}$ une base de $\hat{P}_r$. La matrice de Vandermonde $VDM \in \mathcal{M}_{n_r}(\mathbb{R})$ est définie par*

$$VDM_{i,j} = \hat{\psi}_i(\hat{M}_j)\cdot\hat{t}_j,\ \ 1\leqslant i,j\leqslant n_r,$$

et la fonction de base $\hat{\varphi}_i$ liée au point d'interpolation $\hat{M}_i$ est alors définie par

$$\hat{\varphi}_i = \sum_{1\leqslant j\leqslant n_r} (VDM^{-1})_{i,j}\,\hat{\psi}_j. \tag{10.3.2}$$

Remarque 10.3.2 *Les points $\hat{M}_i$ sont munis d'une orientation, mais peuvent évidemment correspondre à un même point géométrique.*

Remarque 10.3.3 *Comme dans le cas H^1, la caractérisation de l'inversibilité de la matrice de Vandermonde est une question ouverte. Cependant on observe qu'avec notre choix pour le positionnement des degrés de liberté, la matrice de Vandermonde est inversible, c'est à dire que l'élément est unisolvant.*

Le désavantage principal de l'utilisation de fonctions nodales est que la matrice de Vandermonde peut être mal conditionnée, ce qui conduit à des erreurs numériques importantes lors du calcul de la base. La figure 2.8 présente ainsi la comparaison entre le conditionnement de la matrice de Vandermonde pour la première famille optimale (« Gauss-Lobatto ») et pour la première famille (« Gauss », voir chapitre 13, section 13.1.3) dans le cas d'éléments tétraédriques, pyramidaux et prismatiques nodaux. On remarque que le conditionnement ne change pas beaucoup suivant les deux familles, et que les fonctions de base ψ_i choisies pour les pyramides souffrent d'un mauvais conditionnement qui handicape la méthode dès que $r \geqslant 5$. Pour les tétraèdres et le prisme, le conditionnement reste bon.

FIG. 10.5 – Conditionnement de la matrice de Vandermonde en fonction de l'ordre pour les éléments

Nous allons donc chercher une base de fonctions hiérarchiques de $\hat{P}_r$ pour éviter d'avoir à utiliser la méthode nodale, un autre avantage des fonctions hiérarchiques étant d'être tensorisées (sur le cube).

10.3.2 Éléments finis d'arête

Comme dans le cas H^1, on cherche une base de chaque espace $\hat{P}_r$ sur les éléments de référence $\hat{K}$, adaptées à la structure de l'espace fonctionnel $H(rot)$, et qui permette la conformité $H(rot)$ avec les autres éléments. Pour cela, on doit vérifier que chaque fonction de base vérifie la continuité tangentielle, qui est la contrainte délicate lors de la construction de la base.

Proposition 10.3.4 *Les fonctions suivantes forment une base hiérarchique $H(rot)$-conforme de $\hat{P}_r$ et sur tout le maillage*

- ***Hexaèdre :*** *On considère les paramètres suivants*

$$\begin{cases} \lambda_1 = \hat{x} \\ \lambda_2 = \hat{y} \\ \lambda_3 = \hat{z} \\ \lambda_4 = 1 - \hat{x} \\ \lambda_5 = 1 - \hat{y} \\ \lambda_6 = 1 - \hat{z} \end{cases}$$

FONCTIONS $H(rot)$ HIÉRARCHIQUES POUR L'HEXAÈDRE

Pour une arête a *: soient* a_1 *et* a_2 *les faces ne contenant aucun sommet de* a $(a_1 < a_2)$

Si a *est orientée selon* e_x

$$\begin{bmatrix} \lambda_{a_1}\lambda_{a_2} \\ 0 \\ 0 \end{bmatrix} P_i^{0,0}(2\hat{x}-1), \quad 0 \leqslant i \leqslant r-1, \quad 1 \leqslant a \leqslant 4$$

Si a *est orientée selon* e_y

$$\begin{bmatrix} 0 \\ \lambda_{a_1}\lambda_{a_2} \\ 0 \end{bmatrix} P_i^{0,0}(2\hat{y}-1), \quad 0 \leqslant i \leqslant r-1, \quad 1 \leqslant a \leqslant 4$$

Si a *est orientée selon* e_z

$$\begin{bmatrix} 0 \\ 0 \\ \lambda_{a_1}\lambda_{a_2} \end{bmatrix} P_i^{0,0}(2\hat{z}-1), \quad 0 \leqslant i \leqslant r-1, \quad 1 \leqslant a \leqslant 4$$

Pour une face f *: soit* f_1 *la face directement opposée à* f

Si f *est dans le plan* (e_x, e_y)

$$\begin{bmatrix} \lambda_2\lambda_5\lambda_{f_1} \\ 0 \\ 0 \end{bmatrix} P_i^{0,0}(2\hat{x}-1)P_j^{1,1}(2\hat{y}-1)$$
$$\begin{bmatrix} 0 \\ \lambda_1\lambda_4\lambda_{f_1} \\ 0 \end{bmatrix} P_j^{1,1}(2\hat{x}-1)P_i^{0,0}(2\hat{y}-1) \qquad 0 \leqslant i,j \leqslant r-1, \quad 1 \leqslant f \leqslant 2$$

Si f *est dans le plan* (e_y, e_z)

$$\begin{bmatrix} 0 \\ \lambda_3\lambda_6\lambda_{f_1} \\ 0 \end{bmatrix} P_i^{0,0}(2\hat{y}-1)P_j^{1,1}(2\hat{z}-1)$$
$$\begin{bmatrix} 0 \\ 0 \\ \lambda_2\lambda_5\lambda_{f_1} \end{bmatrix} P_j^{1,1}(2\hat{y}-1)P_i^{0,0}(2\hat{z}-1) \qquad 0 \leqslant i,j \leqslant r-1, \quad 1 \leqslant f \leqslant 2$$

Si f *est dans le plan* (e_x, e_z)

$$\begin{bmatrix} \lambda_3\lambda_6\lambda_{f_1} \\ 0 \\ 0 \end{bmatrix} P_i^{0,0}(2\hat{x}-1)P_j^{1,1}(2\hat{z}-1)$$
$$\begin{bmatrix} 0 \\ 0 \\ \lambda_1\lambda_4\lambda_{f_1} \end{bmatrix} P_j^{1,1}(2\hat{x}-1)P_i^{0,0}(2\hat{z}-1) \qquad 0 \leqslant i,j \leqslant r-1, \quad 1 \leqslant f \leqslant 2$$

Pour les fonctions intérieures *:*

$$\begin{bmatrix} \lambda_2\lambda_3\lambda_5\lambda_6 \\ 0 \\ 0 \end{bmatrix} P_k^{0,0}(2\hat{x}-1)P_i^{1,1}(2\hat{y}-1)P_j^{1,1}(2\hat{z}-1)$$
$$\begin{bmatrix} 0 \\ \lambda_1\lambda_3\lambda_4\lambda_6 \\ 0 \end{bmatrix} P_i^{1,1}(2\hat{x}-1)P_k^{0,0}(2\hat{y}-1)P_j^{1,1}(2\hat{z}-1) \qquad 0 \leqslant i,j,k \leqslant r-1$$
$$\begin{bmatrix} 0 \\ 0 \\ \lambda_1\lambda_2\lambda_4\lambda_5 \end{bmatrix} P_i^{1,1}(2\hat{x}-1)P_j^{1,1}(2\hat{y}-1)P_k^{0,0}(2\hat{z}-1)$$

– ***Prisme :*** *On considère les paramètres suivants*

$$\begin{cases} \lambda_1 = \lambda_4 = 1 - \hat{x} - \hat{y} \\ \lambda_2 = \lambda_5 = \hat{x} \\ \lambda_3 = \lambda_6 = \hat{y} \end{cases} \qquad \begin{cases} \beta_1 = 1 - \hat{z} \\ \beta_2 = \hat{z} \end{cases}$$

FONCTIONS $H(rot)$ HIÉRARCHIQUES POUR LE PRISME

Pour une arête horizontale a *: l'arête est dirigée d'un sommet* a_1 *vers* a_2*, et* f' *est la face horizontale opposée*

$$(\lambda_{a_1}\nabla\lambda_{a_2} - \lambda_{a_2}\nabla\lambda_{a_1})\,\beta_{f'}\,P_i^{0,0}(\lambda_{a_2} - \lambda_{a_1}), \quad 0 \leqslant i \leqslant r-1, \quad 1 \leqslant a \leqslant 6$$

Pour une arête verticale a *: soit* a_1 *la face ne contenant aucun sommet de* a

$$\begin{bmatrix} 0 \\ 0 \\ \lambda_{a_1} \end{bmatrix} P_i^{0,0}(\beta_2 - \beta_1), \quad 0 \leqslant i \leqslant r-1, \quad 1 \leqslant a \leqslant 3$$

Pour une face quadrangulaire f *: soit* $[a_1, a_2]$ *une arête en commun avec une face triangulaire* f'*, et* f_1 *et* f_2 *les deux autres faces quadrangulaires*

$$(\lambda_{a_1}\nabla\lambda_{a_2} - \lambda_{a_2}\nabla\lambda_{a_1})\,\beta_1\,\beta_2\,P_i^{0,0}(\lambda_{a_2} - \lambda_{a_1})P_j^{1,1}(2\hat{z}-1)$$

$$\begin{bmatrix} 0 \\ 0 \\ \lambda_{a_1}\,\lambda_{a_2} \end{bmatrix} P_j^{1,1}(\lambda_{a_2} - \lambda_{a_1})P_i^{0,0}(2\hat{z}-1) \qquad \begin{matrix} 0 \leqslant i \leqslant r-1 \\ 0 \leqslant j \leqslant r-1 \end{matrix} \quad 1 \leqslant f \leqslant 3$$

Pour une face triangulaire f *: soient* $[a_1, a_2]$ *et* $[a_1, a_3]$ *deux arêtes en commun avec deux faces faces quadrangulaires* f_1 *et* f_2 *respectivement, et* f' *la face horizontale opposée*

$$\begin{matrix} (\lambda_{a_1}\nabla\lambda_{a_2} - \lambda_{a_2}\nabla\lambda_{a_1})\,\lambda_{f_1}\,\beta_{f'}\,P_i^{0,0}(\lambda_{a_2} - \lambda_{a_1})P_j^{0,0}(\lambda_{a_3} - \lambda_{a_1}) \\ (\lambda_{a_1}\nabla\lambda_{a_3} - \lambda_{a_3}\nabla\lambda_{a_1})\,\lambda_{f_2}\,\beta_{f'}\,P_i^{0,0}(\lambda_{a_2} - \lambda_{a_1})P_j^{0,0}(\lambda_{a_3} - \lambda_{a_1}) \end{matrix} \quad 0 \leqslant i+j \leqslant r-2, \quad 1 \leqslant f \leqslant 2$$

Pour les fonctions intérieures *:*

$$(\lambda_2\nabla\lambda_3 - \lambda_3\nabla\lambda_2)\,\lambda_1\,\beta_1\,\beta_2\,P_{ijk}(\hat{x},\hat{y},\hat{z})$$

$$(\lambda_1\nabla\lambda_3 - \lambda_3\nabla\lambda_1)\,\lambda_2\,\beta_1\,\beta_2\,P_{ijk}(\hat{x},\hat{y},\hat{z})$$

$$\begin{bmatrix} 0 \\ 0 \\ \lambda_1\,\lambda_2\,\lambda_3 \end{bmatrix} P_{ijk}(\hat{x},\hat{y},\hat{z}) \qquad \begin{matrix} 0 \leqslant i+j \leqslant r-2 \\ 0 \leqslant k \leqslant r-1 \end{matrix}$$

avec

$$P_{ijk}(\hat{x},\hat{y},\hat{z}) = P_i^{0,0}\left(\frac{2\hat{x}}{1-\hat{y}} - 1\right)(1-\hat{y})^i P_j^{2i+1,0}(2\hat{y}-1)P_k^{0,0}(2\hat{z}-1)$$

– ***Pyramide :*** *On considère les paramètres suivants*

$$\begin{cases} \beta_1 = \dfrac{1-\hat{x}-\hat{z}}{2} \\ \beta_2 = \dfrac{1-\hat{y}-\hat{z}}{2} \\ \beta_3 = \dfrac{1+\hat{x}-\hat{z}}{2} \\ \beta_4 = \dfrac{1+\hat{y}-\hat{z}}{2} \end{cases} \qquad \begin{cases} \lambda_1 = \dfrac{\beta_1\,\beta_2}{1-\hat{z}} \\ \lambda_2 = \dfrac{\beta_2\,\beta_3}{1-\hat{z}} \\ \lambda_3 = \dfrac{\beta_3\,\beta_4}{1-\hat{z}} \\ \lambda_4 = \dfrac{\beta_4\,\beta_1}{1-\hat{z}} \\ \lambda_5 = \hat{z} \end{cases} \qquad \begin{cases} \gamma_1 = \dfrac{2\hat{z}+\hat{x}+\hat{y}}{2} \\ \gamma_2 = \dfrac{2\hat{z}-\hat{x}+\hat{y}}{2} \\ \gamma_3 = \dfrac{2\hat{z}-\hat{x}-\hat{y}}{2} \\ \gamma_4 = \dfrac{2\hat{z}+\hat{x}-\hat{y}}{2} \end{cases} \qquad \begin{cases} \delta_1 = \delta_3 = \hat{x} \\ \delta_2 = \delta_4 = \hat{y} \end{cases}$$

FONCTIONS $H(rot)$ HIÉRARCHIQUES POUR LA PYRAMIDE

Pour une arête horizontale a *: l'arête est dirigée d'un sommet* a_1 *vers* a_2*, et les arêtes horizontales adjacentes sont* $[a_1, a_4]$ *et* $[a_2, a_3]$

$$(\lambda_{a_1}\nabla(\lambda_{a_2}+\lambda_{a_3}) - \lambda_{a_2}\nabla(\lambda_{a_1}+\lambda_{a_4}))\,P_i^{0,0}(\delta_a), \quad 0 \leqslant i \leqslant r-1, \quad 1 \leqslant a \leqslant 4$$

Pour une arête verticale a *: soit* s *le sommet de* a *appartenant à la base*

$$(\lambda_s\nabla\lambda_5 - \lambda_5\nabla\lambda_s)\,P_i^{0,0}(\gamma_s), \quad 0 \leqslant i \leqslant r-1, \quad 1 \leqslant a \leqslant 4$$

Pour la base *:*

$$\begin{aligned} &(\lambda_1\nabla(\lambda_2+\lambda_3) - \lambda_2\nabla(\lambda_1+\lambda_4))\;\beta_4\,P_i^{0,0}\Big(\frac{\beta_3-\beta_1}{1-\hat{z}}\Big)P_j^{1,1}\Big(\frac{\beta_4-\beta_2}{1-\hat{z}}\Big)(1-\hat{z})^{\max(i,j)-1} \\ &(\lambda_1\nabla(\lambda_3+\lambda_4) - \lambda_4\nabla(\lambda_2+\lambda_1))\;\beta_3\,P_j^{1,1}\Big(\frac{\beta_3-\beta_1}{1-\hat{z}}\Big)P_i^{0,0}\Big(\frac{\beta_4-\beta_2}{1-\hat{z}}\Big)(1-\hat{z})^{\max(i,j)-1} \end{aligned} \qquad 0 \leqslant i,j \leqslant r-1$$

Pour une face triangulaire f *: soit* $[a_1, a_2]$ *l'arête verticale d'arêtes adjacentes* $[a_1, a_4]$ *et* $[a_2, a_3]$*, et* f_1 *la face triangulaire de base* $[a_1, a_4]$

$$\begin{aligned} &(\lambda_{a_2}\nabla(\lambda_{a_1}+\lambda_{a_4}) - \lambda_{a_1}\nabla(\lambda_{a_2}+\lambda_{a_3}))\;\lambda_5\,P_i^{0,0}(\delta_f)P_j^{0,0}(\gamma_{a_1}) \\ &(\lambda_{a_1}\nabla\lambda_5 - \lambda_5\nabla\lambda_{a_1})\;\beta_{f_1}\,P_i^{0,0}(\delta_f)P_j^{0,0}(\gamma_{a_1}) \end{aligned} \qquad 0 \leqslant i+j \leqslant r-2$$

Pour les fonctions intérieures *:*

$$\begin{aligned} &(\lambda_1\nabla(\lambda_2+\lambda_3) - \lambda_2\nabla(\lambda_1+\lambda_4))\;\beta_4\,\lambda_5\,P_{ijk}(\hat{x},\hat{y},\hat{z}) \\ &(\lambda_1\nabla(\lambda_3+\lambda_4) - \lambda_4\nabla(\lambda_2+\lambda_1))\;\beta_3\,\lambda_5\,P_{ijk}(\hat{x},\hat{y},\hat{z}) \\ &(\lambda_1\nabla\lambda_5 - \lambda_5\nabla\lambda_1)\;\beta_3\,\beta_4\,P_{ijk}(\hat{x},\hat{y},\hat{z}) \end{aligned} \qquad \begin{matrix} 0 \leqslant i,j \leqslant r-2, \\ 0 \leqslant k \leqslant r-2-max(i,j) \end{matrix}$$

avec

$$P_{ijk}(\hat{x},\hat{y},\hat{z}) = P_i^{0,0}\Big(\frac{\beta_3-\beta_1}{1-\hat{z}}\Big)P_j^{0,0}\Big(\frac{\beta_4-\beta_2}{1-\hat{z}}\Big)P_k^{2\max(i,j)+2,0}(2\hat{z}-1)(1-\hat{z})^{\max(i,j)-1}$$

– ***Tétraèdre :*** *On considère les paramètres suivants*

$$\begin{cases} \lambda_1 = 1 - \hat{x} - \hat{y} - \hat{z} \\ \lambda_2 = \hat{x} \\ \lambda_3 = \hat{y} \\ \lambda_4 = \hat{z} \end{cases}$$

FONCTIONS H^1 HIÉRARCHIQUES POUR LE TÉTRAÈDRE

Pour une arête a *: l'arête est dirigée d'un sommet* a_1 *vers* a_2

$$(\lambda_{a_1}\nabla\lambda_{a_2} - \lambda_{a_2}\nabla\lambda_{a_1})\ P_i^{0,0}(\lambda_{a_2} - \lambda_{a_1}), \quad 0 \leqslant i \leqslant r-1, \quad 1 \leqslant a \leqslant 6$$

Pour une face triangulaire f *: soient* a_1, a_2 *et* a_3 *les sommets de* f, f_1 *l'autre face contenant* $[a_1, a_2]$ *et* f_2 *l'autre face contenant* $[a_1, a_3]$

$$\begin{array}{l} (\lambda_{a_1}\nabla\lambda_{a_2} - \lambda_{a_2}\nabla\lambda_{a_1})\ \lambda_{f_1}\ P_{ij}(\hat{x},\hat{y},\hat{z}) \\ \qquad\qquad\qquad\qquad 0 \leqslant i+j \leqslant r-2, \quad 1 \leqslant f \leqslant 4 \\ (\lambda_{a_1}\nabla\lambda_{a_3} - \lambda_{a_3}\nabla\lambda_{a_1})\ \lambda_{f_2}\ P_{ij}(\hat{x},\hat{y},\hat{z}) \end{array}$$

avec

$$P_{ij}(\hat{x},\hat{y},\hat{z}) = P_i^{0,0}(\lambda_{a_2} - \lambda_{a_1}) P_j^{0,0}(\lambda_{a_3} - \lambda_{a_1})$$

Pour les fonctions intérieures *:*

$$(\lambda_1\nabla\lambda_4 - \lambda_4\nabla\lambda_1)\ \lambda_2\lambda_3\ P_{ijk}(\hat{x},\hat{y},\hat{z})$$

$$(\lambda_1\nabla\lambda_2 - \lambda_2\nabla\lambda_1)\ \lambda_3\lambda_4\ P_{ijk}(\hat{x},\hat{y},\hat{z}) \qquad 0 \leqslant i+j+k \leqslant r-3$$

$$(\lambda_1\nabla\lambda_3 - \lambda_3\nabla\lambda_1)\ \lambda_2\lambda_4\ P_{ijk}(\hat{x},\hat{y},\hat{z})$$

avec

$$P_{ijk}(\hat{x},\hat{y},\hat{z}) = P_i^{0,0}\left(\frac{2\hat{x}}{1-\hat{y}-\hat{z}} - 1\right)(1-\hat{y}-\hat{z})^i P_j^{2i+1,0}\left(\frac{2\hat{y}}{1-\hat{z}} - 1\right)(1-\hat{z})^j P_k^{2(i+j+1),0}(2\hat{z}-1)$$

Preuve. Par construction, l'ensemble des fonctions forme une base de $\hat{P}_r$ pour chaque élément et les restrictions aux arêtes, aux faces triangulaires et quadrangulaires sont égales (à une rotation près pour les faces triangulaires, à prendre en compte lors de la programmation) pour tous les éléments.

La démonstration du fait que l'ensemble des fonctions forme une base de $\hat{P}_r$ pour chaque type d'élément ne sera pas détaillée.

10.4 Conformité

On rappelle le théorème concernant les conditions de conformité $H(rot)$.

Theorème 10.4.1

$$\begin{cases} \forall K \in \Omega,\ P_r^F(K) \subset H(rot)(K) \\ V_h \wedge n \subset \mathcal{C}^0(\bar{\Omega}),\ \textit{pour } n \textit{ normale sortante à chaque face} \end{cases} \Longrightarrow V_h \subset H(rot,\Omega)$$

Preuve. Voir par exemple Monk [55].

Theorème 10.4.2 *Avec les espaces* P_r^F *construits dans la section 10.2 et le choix de degré de liberté de la section 10.3, on a*

$$V_h(\Omega) \subset H(rot,\Omega).$$

Preuve.

– Vérifions tout d'abord le premier point du théorème 10.4.1. Soit $p \in P_r^F$. Lorsque p est polynomiale, ce qui est le cas pour les tétraèdres, les prismes et les hexaèdres, on a de manière évidente $p \in \mathcal{C}^0(K)^3$ et $rot\, p \in \mathcal{C}^0(K)$. Puisque K est borné, on a alors immédiatement $p \in L^2(K)^3$ et $rot\, p \in L^2(K)$.

Concernant les pyramides, on a deux cas
– Si $p \in \mathbb{B}^3_{r-1}$, en réutilisant les résultat obtenus pour H^1, on a $p \in L^2(K)^3$ et $\nabla p_1, \nabla p_2, \nabla p_3 \in L^2(K)^3$, donc $rot\, p \in L^2(K)^3$.
– Sinon, $p \in A_r^3$ où $A_r = \left\{ \dfrac{\hat{x}^i \hat{y}^j}{(1-\hat{z})^{i+j-k}}, 0 \leqslant i,j \leqslant k+1 \leqslant r, 1 \leqslant\leqslant r-1 \right\}$. Comme dans le cas H^1, on prouve la continuité en S_5 en considérant la pseudo-face F_ε^2, représentée en rouge sur la figure 10.6, telle que

$$\begin{cases} \hat{x} = (1-\hat{z})(1-\varepsilon) \\ -(1-\hat{z})(1-\varepsilon) \leqslant \hat{y} \leqslant (1-\hat{z})(1-\varepsilon) \\ 0 \leqslant \hat{z} \leqslant 1, \end{cases}$$

et

$$\forall M = (\hat{x}, \hat{y}, \hat{z}) \in Q_2,\ \exists \varepsilon \in [0,1],\ M \in F_\varepsilon^2.$$

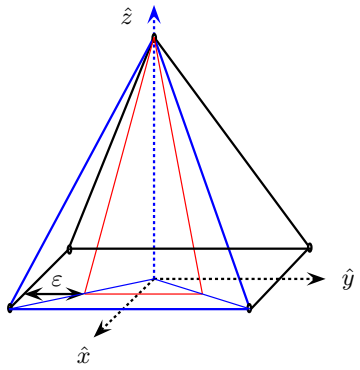

FIG. 10.6 – Pseudo face F_ε^2

Pour un point M de Q_2 attaché à une face F_ε^2, on a

$$-(1-\varepsilon)^{i+j}(1-\hat{z})^k \leqslant p(M) \leqslant (1-\varepsilon)^{i+j}(1-\hat{z})^k$$

c'est à dire

$$\begin{array}{l} p(M) \xrightarrow{\hat{z}\to 1} 0 \text{ si } k > 0 \\ p(M) \text{ borné si k=0} \end{array}$$

Puisque K est borné, on a $p \in L^2(K)$.

Pour $rot\, p$, on distingue là encore deux cas
– Soit

$$p = \begin{bmatrix} \dfrac{\hat{x}^p\, \hat{y}^{p+1}}{(1-\hat{z})^{p+1}} \\ \dfrac{\hat{x}^{p+1}\, \hat{y}^p}{(1-\hat{z})^{p+1}} \\ \dfrac{\hat{x}^{p+1}\, \hat{y}^{p+1}}{(1-\hat{z})^{p+2}} \end{bmatrix} = \frac{1}{p+1} grad \left(\frac{\hat{x}^{p+1}\, \hat{y}^{p+1}}{(1-\hat{z})^{p+1}} \right), 0 \leqslant p \leqslant r-1 \quad \Longrightarrow \quad rot\, p = 0$$

– Soit p est de la forme

$$p = \begin{bmatrix} \dfrac{\hat{x}^i\, \hat{y}^j}{(1-\hat{z})^{i+j-k}} \\ 0 \\ \dfrac{\hat{x}^{i+1}\, \hat{y}^j}{(1-\hat{z})^{i+j-k+1}} \end{bmatrix}, \quad \begin{array}{l} 0 \leqslant i,j \leqslant k+1 \\ 1 \leqslant k \leqslant r+1 \end{array} \quad \Longrightarrow \quad rot\, p = \begin{bmatrix} \dfrac{j\, \hat{x}^{i+1}\, \hat{y}^{j-1}}{(1-\hat{z})^{i+j-k+1}} \\ \dfrac{(j-k+1)\, \hat{x}^i \hat{y}^j}{(1-\hat{z})^{i+j-k+1}} \\ -\dfrac{\hat{x}^i\, \hat{y}^{j-1}}{(1-\hat{z})^{i+j-k}} \end{bmatrix}$$

ou son symétrique en y. Dans ce cas, en un point M de Q_2 attaché à une face F_ε^2, on a

$$-(1-\varepsilon)^{i+j}(1-\hat{z})^{k-1} \leqslant \frac{\hat{x}^{i+1}\hat{y}^{j-1}}{(1-\hat{z})^{i+j-k+1}} \leqslant (1-\varepsilon)^{i+j}(1-\hat{z})^{k-1}$$

$$-(1-\varepsilon)^{i+j}(1-\hat{z})^{k-1} \leqslant \frac{\hat{x}^{i}\hat{y}^{j}}{(1-\hat{z})^{i+j-k+1}} \leqslant (1-\varepsilon)^{i+j}(1-\hat{z})^{k-1}$$

$$-(1-\varepsilon)^{i+j-1}(1-\hat{z})^{k-1} \leqslant \frac{\hat{x}^{i}\hat{y}^{j-1}}{(1-\hat{z})^{i+j-k}} \leqslant (1-\varepsilon)^{i+j-1}(1-\hat{z})^{k-1}$$

On a donc

$$(1-\varepsilon)^{i+j-1}(1-\hat{z})^{k-1} \xrightarrow{z\to 1} 0 \text{ pour k>1}$$
$$(1-\varepsilon)^{i+j-1}(1-\hat{z})^{k-1} \text{ borné pour } k=1$$

c'est à dire que $rot\, p$ est borné. Comme K est borné, on a donc $rot\, p \in L^2(K)$.

- Concernant le deuxième point du théorème 10.4.1, il faut vérifier que les restrictions de la composante tangentielle des espaces P_r^F sur chaque type d'élément sur les faces de même type sont identiques. Par construction, la transformation de Piolat 10.3.1 respecte la conformité $H(rot)$ (voir Monk [55]). Il reste donc à vérifier que les restrictions des espaces $\hat{P}_r$ à chaque type de face sont les mêmes pour tous les éléments.

  Pour un paramétrage (η,ξ) de la face considérée, il est immédiat que la restriction de $\mathcal{R}_r(\hat{x},\hat{y},\hat{z})$ à toute face triangulaire du tétraèdre est dans $\mathcal{R}_r(\eta,\xi)$, et que la restriction de $\mathcal{Q}_r(\hat{x},\hat{y},\hat{z})$ à toute face quadrangulaire de l'hexaèdre est dans $\mathbb{Q}_{r-1,r+1}(\eta,\xi)\times\mathbb{Q}_{r+1,r-1}(\eta,\xi)$. Or retrouve aisément ces résultats sur le prismes, mais nous allons détailler le cas de la pyramide.
  - Sur la base quadrangulaire, $\hat{z}=0$ et $n=[0,0,-1]$. Pour toutes les familles qui composent $\mathcal{B}_r$, on a de manière quasi immédiate $(p\wedge n)\wedge n_{|\hat{z}=0} = [-p_2(\hat{x},\hat{y},0), p_1(\hat{x},\hat{y},0), 0] \in \mathbb{Q}_{r+1,r-1}(\hat{x},\hat{y})\times\mathbb{Q}_{r-1,r+1}(\hat{x},\hat{y})$
  - Sur une face triangulaire, on considère par exemple la face $\hat{x}=(1-\hat{z})$, avec $n=[1,0,1]$. La partie difficile va être d'identifier les monômes de $\mathcal{S}_r(\hat{y},\hat{z})$. On écrit donc tout d'abord $\mathcal{S}_r(\hat{y},\hat{z})$ sur la face.

    On considère la transformation f du triangle de référence $\hat{T}(\eta,\xi)$ à la face $\hat{x}=1-\hat{z}$ de $\hat{K}(\hat{x},\hat{y},\hat{z})$

$$f = \left|\begin{array}{l} \hat{x} = 1-\xi \\ \hat{y} = 2\eta+\xi-1 \\ \hat{z} = \xi \end{array}\right.$$

    La transformation f est un difféomorphisme de $\hat{T}$ vers la face $\hat{x}=1-\hat{z}$ de $\hat{K}$. On a

$$Df = \begin{bmatrix} 0 & -1 \\ 2 & 1 \\ 0 & 1 \end{bmatrix} \qquad Df^{-*} = \frac{1}{4}\begin{bmatrix} 1 & -2 \\ 2 & 0 \\ -1 & 2 \end{bmatrix}$$

    où Df est le jacobien de la transformation f et Df^{-*} la transposée du pseudo-inverse de Df. Pour $\varphi \in \mathcal{S}_r(\eta,\xi)$, d'après la proposition 10.2.2, φ s'écrit

$$\varphi(\eta,\xi) = \begin{bmatrix} \eta^i\xi^{j+1} \\ -\eta^{i+1}\xi^j \end{bmatrix}$$

    pour $i+j=r-1$. Sur la face $\hat{x}=1-\hat{z}$ de $\hat{K}$, φ s'écrit donc

$$\varphi(\hat{x},\hat{y},\hat{z}) = Df^{-*}\varphi(\eta,\xi)\circ f^{-1} = \left(\frac{\hat{y}+1-\hat{z}}{2}\right)^i \hat{z}^j \begin{bmatrix} \hat{y}+1 \\ 2\hat{z} \\ -\hat{y}-1 \end{bmatrix}$$

    On peut ajouter autant de fonctions de $\mathbb{P}_{r-1}^2$ que l'on veut sur la face sans modifier l'espace engendré : on rajoute $(t_1-2t_2)\left(\frac{\hat{y}+1-\hat{z}}{2}\right)^i \hat{z}^j$ avec

$$t_1 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \qquad t_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}.$$

On obtient ainsi

$$\mathcal{S}_r(\hat{y},\hat{z}) = \left\{ \begin{bmatrix} \hat{y}^{j+1}(1-\hat{z})^i \\ -2\hat{y}^j(1-\hat{z})^{i+1} \\ -\hat{y}^{j+1}(1-\hat{z})^i \end{bmatrix}, \quad i+j = r-1 \right\}$$

Identifions maintenant les monômes sur la face. On calcule $p_n(\hat{y},\hat{z}) = (p \wedge n) \wedge n_{|\hat{x}=(1-\hat{z})}$

– Si $p \in \mathbb{B}^3_{r-1}$, on a de manière immédiate $p_n \in \mathbb{P}^2_{r-1}(\hat{y},\hat{z})$.

– Sinon

$$0 \leqslant i \leqslant r-1, \qquad p = \begin{bmatrix} \dfrac{\hat{x}^i\,\hat{y}^{i+1}}{(1-\hat{z})^{i+1}} \\ \dfrac{\hat{x}^{i+1}\,\hat{y}^i}{(1-\hat{z})^{i+1}} \\ \dfrac{\hat{x}^{i+1}\,\hat{y}^{i+1}}{(1-\hat{z})^{i+2}} \end{bmatrix} \quad \Longrightarrow \quad p_n = \begin{bmatrix} 0 \\ \hat{y}^i \\ 0 \end{bmatrix} \in \mathbb{P}^2_{r-1}(\hat{y},\hat{z})$$

$$0 \leqslant m \leqslant n \leqslant r-2, \quad p = \begin{bmatrix} \dfrac{\hat{x}^m\,\hat{y}^{n+2}}{(1-\hat{z})^{m+1}} \\ 0 \\ \dfrac{\hat{x}^{m+1}\,\hat{y}^{n+2}}{(1-\hat{z})^{m+2}} \end{bmatrix} \quad \Longrightarrow \quad p_n = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$0 \leqslant m \leqslant n \leqslant r-2, \quad p = \begin{bmatrix} 0 \\ \dfrac{\hat{x}^{n+2}\,\hat{y}^m}{(1-\hat{z})^{m+1}} \\ \dfrac{\hat{x}^{n+2}\,\hat{y}^{m+1}}{(1-\hat{z})^{m+2}} \end{bmatrix} \quad \Longrightarrow \quad p_n = \begin{bmatrix} \hat{y}^{m+1}(1-\hat{z})^{n-m} \\ -2\hat{y}^m(1-\hat{z})^{n-m+1} \\ -\hat{y}^{m+1}(1-\hat{z})^{n-m} \end{bmatrix} \in \mathbb{P}^2_{r-1}(\hat{y},\hat{z})$$

$$\begin{matrix} 0 \leqslant p \leqslant r-1 \\ 0 \leqslant q \leqslant r+1 \end{matrix}, \qquad p = \begin{bmatrix} \dfrac{\hat{x}^p\,\hat{y}^q}{(1-\hat{z})^{p+q-r}} \\ 0 \\ \dfrac{\hat{x}^{p+1}\,\hat{y}^q}{(1-\hat{z})^{p+q+1-r}} \end{bmatrix} \quad \Longrightarrow \quad p_n = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{matrix} 0 \leqslant p \leqslant r-1 \\ 0 \leqslant q \leqslant r+1 \end{matrix}, \qquad p = \begin{bmatrix} 0 \\ \dfrac{\hat{x}^q\,\hat{y}^p}{(1-\hat{z})^{p+q-r}} \\ \dfrac{\hat{x}^q\,\hat{y}^{p+1}}{(1-\hat{z})^{p+q+1-r}} \end{bmatrix} \quad \Longrightarrow \quad p_n = \begin{bmatrix} \hat{y}^{p+1}(1-\hat{z})^{r-p-1} \\ -2\hat{y}^p(1-\hat{z})^{r-p} \\ -\hat{y}^{p+1}(1-\hat{z})^{r-p-1} \end{bmatrix} \in \mathcal{S}_r(\hat{y},\hat{z})$$

Le même résultat peut être obtenu de manière symétrique sur les autres faces.

En utilisant un argument de dimension, on en déduit que

$$\begin{array}{rcl} \hat{P}_{r|\,\hat{x}=1-\hat{z} \text{ ou } \hat{x}=\hat{z}-1} & = & \mathcal{R}_r(\hat{y},\hat{z}) \\ \hat{P}_{r|\,\hat{y}=1-\hat{z} \text{ ou } \hat{y}=\hat{z}-1} & = & \mathcal{R}_r(\hat{x},\hat{z}) \\ \hat{P}_{r|\,\hat{z}=0} & = & \mathbb{Q}_{r-1,r+1}(\hat{x},\hat{y}) \times \mathbb{Q}_{r+1,r-1}(\hat{x},\hat{y}), \end{array} \tag{10.4.1}$$

ce qui correspond bien aux mêmes restrictions d'espaces sur les faces des autres éléments. □

Chapitre 11

Formule de quadrature et estimations d'erreur

On évoque ici brièvement les formules de quadrature utilisées pour effectuer les calculs d'intégration et on tente de justifier le choix de l'espace d'approximation en effectuant des estimations d'erreur.

Sommaire

11.1 Intégration par formule de quadrature

11.1.1 Intégration exacte

On rappelle l'expression des matrices de masse M_h et de rigidité R_h sur un élément K du maillage

$$\begin{aligned} M_{hi,j} &= \int_K \varphi_i \cdot \varphi_j \, dx\, dy\, dz = \int_{\hat{K}} |DF| \, DF^{-1} DF^{*-1} \hat{\varphi}_i \cdot \hat{\varphi}_j \, d\hat{x}\, d\hat{y}\, d\hat{z} \\ R_{hi,j} &= \int_K rot\, \varphi_i \cdot rot\, \varphi_j \, dx\, dy\, dz = \int_{\hat{K}} \frac{1}{|DF|} DF^* DF rot\, \hat{\varphi}_i \cdot rot\, \hat{\varphi}_j \, d\hat{x}\, d\hat{y}\, d\hat{z}, \end{aligned} \tag{11.1.1}$$

où DF et $|DF|$ désignent respectivement la jacobienne et le jacobien de F.

En utilisant la définition 2.2.7 de la transformation T permettant de passer de l'élément de référence $\hat{K}$ au cube unité $\widetilde{Q}$, on a

$$\begin{aligned} M_{hi,j} &= \int_{\widetilde{Q}} \widetilde{|DF|} \, \widetilde{|T|} \, \widetilde{DF^{-1}} \, \widetilde{DF^{*-1}} \, \widetilde{\varphi}_i \cdot \widetilde{\varphi}_j \, d\widetilde{x}\, d\widetilde{y}\, d\widetilde{z} \\ R_{hi,j} &= \int_{\widetilde{Q}} \frac{1}{\widetilde{|DF|}} \, \widetilde{|T|} \, \widetilde{DF^*} \, \widetilde{DF} \, rot\, \widetilde{\varphi}_i \cdot rot\, \widetilde{\varphi}_j \, d\widetilde{x}\, d\widetilde{y}\, d\widetilde{z}, \end{aligned} \tag{11.1.2}$$

où $|T|$ désigne le déterminant du changement de variable T. On rappelle une nouvelle fois que, d'après la définition 2.2.7, on a

Hexaèdre :	$\widetilde{|T|} = 1$
Prisme :	$\widetilde{|T|} = (1-\widetilde{y})$
Pyramide :	$\widetilde{|T|} = 4(1-\widetilde{z})^2$
Tétraèdre :	$\widetilde{|T|} = (1-\widetilde{y})(1-\widetilde{z})^2$

D'après la structure des deux matrices, qui font intervenir DF^{-1} et DF^{*-1}, ou $|DF|^{-1}$, il est clair que l'intégration exacte des matrices n'est possible que lorsque F est affine. Dans ce cas, à une constante près, les matrices sont de la forme

$$\begin{aligned} M_{hi,j} &= \int_{\widetilde{Q}} A \, \widetilde{|T|} \, \widetilde{\varphi}_i \cdot \widetilde{\varphi}_j \, d\widetilde{x}\, d\widetilde{y}\, d\widetilde{z} \\ R_{hi,j} &= \int_{\widetilde{Q}} B \, \widetilde{|T|} \, rot\, \widetilde{\varphi}_i \cdot rot\, \widetilde{\varphi}_j \, d\widetilde{x}\, d\widetilde{y}\, d\widetilde{z}, \end{aligned} \tag{11.1.3}$$

où A et B sont deux matrices constantes.

On cherche à intégrer les matrices de masse et de rigidité exactement dans le cas affine. Pour cela, on cherche l'espace de polynômes auquel appartient le contenu des intégrales et pour lequel il faudra avoir une intégration exacte. Comme l'on va généralement utiliser des formules de quadratures issues de la tensorisation de formules de 1D, nous cherchons les inclusions dans des espaces de type $\mathbb{Q}_{m,n,p}$.

Lemme 11.1.1 *Pour chaque type d'élément, on a*

- ***Hexaèdre :***
$$\forall i \in [\![1, n_r]\!], \;\; \widetilde{\varphi}_i \in \mathbb{Q}_{r-1,r+1,r+1} \times \mathbb{Q}_{r+1,r-1,r+1} \times \mathbb{Q}_{r+1,r+1,r-1} \, (\widetilde{x}, \widetilde{y}, \widetilde{z})$$
- ***Prisme :***
$$\forall i \in [\![1, n_r]\!], \;\; \widetilde{\varphi}_i \in \mathbb{Q}_{r-1,r,r+1} \times \mathbb{Q}_{r,r,r+1} \times \mathbb{Q}_{r+1,r+1,r-1} \, (\widetilde{x}, \widetilde{y}, \widetilde{z})$$
- ***Pyramide :***
$$\forall i \in [\![1, n_r]\!], \;\; \widetilde{\varphi}_i \in \mathbb{Q}_{r-1,r+1,r} \times \mathbb{Q}_{r+1,r-1,r} \times \mathbb{Q}_{r+1,r+1,r} \, (\widetilde{x}, \widetilde{y}, \widetilde{z})$$
- ***Tétraèdre :***
$$\forall i \in [\![1, n_r]\!], \;\; \widetilde{\varphi}_i \in \mathbb{Q}_{r-1,r,r} \times \mathbb{Q}_r \times \mathbb{Q}_r \, (\widetilde{x}, \widetilde{y}, \widetilde{z})$$

Preuve.

- **Hexaèdre :** Il s'agit de la définition de l'espace optimal.
- **Prisme :** Soit $\hat{p} \in \mathcal{R}_r(\hat{x}, \hat{y})$

  Si $\hat{p} \in \mathbb{P}^2_{r-1}(\hat{x}, \hat{y})$, en appliquant en utilisant la définition 2.2.7 de la transformation T, on a
  $$\widetilde{p} \in \mathbb{Q}_{r-1,r-1} \times \mathbb{Q}_{r-1,r-1}(\widetilde{x}, \widetilde{y})$$

  Si $\hat{p} \in \mathcal{S}_r(\hat{x}, \hat{y})$, en utilisant la propriété 10.2.2, on a
  $$\begin{bmatrix} \widetilde{x}^i (1-\widetilde{y})^i \widetilde{y}^{j+1} \\ -\widetilde{x}^{i+1} (1-\widetilde{y})^{i+1} \widetilde{y}^j \end{bmatrix} \; i+j = r-1 \quad \Longrightarrow \; \widetilde{p}(\widetilde{x}, \widetilde{y}) \in \mathbb{Q}_{r-1,r} \times \mathbb{Q}_r(\widetilde{x}, \widetilde{y})$$

  Puisque $\widetilde{\varphi}_i \in (\mathcal{R}_r(\widetilde{x}(1-\widetilde{y}), \widetilde{y}) \otimes \mathbb{P}_{r+1}(\widetilde{z})) \times \mathbb{W}_{r+1,r-1}(\widetilde{x}(1-\widetilde{y}), \widetilde{y}, \widetilde{z})$, on obtient bien le résultat annoncé.

- **Pyramide :** Immédiat d'après l'expression de l'espace C_r de la propriété 10.2.9.
- **Tétraèdre :** On utilise la décomposition de l'espace $\mathcal{R}_r(\hat{x},\hat{y},\hat{z})$.
  Si $\hat{\varphi}_i \in \mathbb{P}^3_{r-1}(\hat{x},\hat{y},\hat{z})$, en utilisant la définition 2.2.7 de la transformation T, il vient

$$\widetilde{\varphi}_i \in \mathbb{Q}^3_{r-1,r-1,r-1}(\widetilde{x},\widetilde{y},\widetilde{z})$$

  Si $\hat{\varphi}_i \in \mathcal{S}^3_r(\hat{x},\hat{y},\hat{z})$, en passant sur $\widetilde{Q}$, en utilisant la propriété 10.2.2 on a

$$\begin{bmatrix} \widetilde{x}^{i-1}(1-\widetilde{y})^{i-1}\widetilde{y}^{j}(1-\widetilde{z})^{i+j-1}\widetilde{z}^{r-i-j+1} \\ 0 \\ -\widetilde{x}^{i}(1-\widetilde{y})^{i}\widetilde{y}^{j}(1-\widetilde{z})^{i+j}\widetilde{z}^{r-i-j} \end{bmatrix} \quad \begin{matrix} i+j\leqslant r \\ 1\leqslant i\leqslant r \\ 0\leqslant j\leqslant r-1 \end{matrix} \quad \Longrightarrow \widetilde{\varphi} \in \mathbb{Q}_{r-1,r-1,r}\times\{0\}\times\mathbb{Q}_r(\widetilde{x},\widetilde{y},\widetilde{z})$$

$$\begin{bmatrix} 0 \\ \widetilde{x}^{i}(1-\widetilde{y})^{i}\widetilde{y}^{j-1}(1-\widetilde{z})^{i+j-1}\widetilde{z}^{r-i-j+1} \\ -\widetilde{x}^{i}(1-\widetilde{y})^{i}\widetilde{y}^{j}(1-\widetilde{z})^{i+j}\widetilde{z}^{r-i-i} \end{bmatrix} \quad \begin{matrix} i+j\leqslant r \\ 0\leqslant i\leqslant r-1 \\ 1\leqslant j\leqslant r \end{matrix} \quad \Longrightarrow \widetilde{\varphi} \in \{0\}\times\mathbb{Q}_{r-1,r-1,r}\times\mathbb{Q}_{r-1,r,r}(\widetilde{x},\widetilde{y},\widetilde{z})$$

$$\begin{bmatrix} \widetilde{x}^{i-1}(1-\widetilde{y})^{i-1}\widetilde{y}^{j}(1-\widetilde{z})^{i+j-1} \\ -\widetilde{x}^{i}(1-\widetilde{y})^{i}\widetilde{y}^{j-1}(1-\widetilde{z})^{i+j-1} \\ 0 \end{bmatrix} \quad \begin{matrix} i+j=r+1, \\ 1\leqslant i,j\leqslant r \end{matrix} \quad \Longrightarrow \widetilde{\varphi}_i \in \mathbb{Q}_{r-1,r,r}\times\mathbb{Q}_r\times\{0\}(\widetilde{x},\widetilde{y},\widetilde{z})$$

  d'où le résultat. □

Proposition 11.1.2 *Pour avoir l'intégration exacte de la matrice de masse dans le cas affine, la formule de quadrature utilisée doit être exacte pour les polynômes de :*

- ***Hexaèdre*** *:* $\mathbb{Q}_{2r+2}(\widetilde{x},\widetilde{y},\widetilde{z})$ *;*
- ***Prisme*** *:* $(1-\widetilde{y})\,\mathbb{Q}_{2r+2}(\widetilde{x},\widetilde{y},\widetilde{z})$ *;*
- ***Pyramide*** *:* $(1-\widetilde{z})^2\,\mathbb{Q}_{2r+2,2r+2,2r}(\widetilde{x},\widetilde{y},\widetilde{z})$ *;*
- ***Tétraèdre*** *:* $(1-\widetilde{y})(1-\widetilde{z})^2\,\mathbb{Q}_{2r}(\widetilde{x},\widetilde{y},\widetilde{z})$.

Preuve. En utilisant le lemme 11.1.1, via l'écriture de la matrice de masse sur le cube unité $\widetilde{Q}$ donnée par l'équation 11.1.3, le résultat est obtenu en additionnant les puissances. □

Lemme 11.1.3 *Pour chaque type d'élément, on a*

- ***Hexaèdre :***
$$\forall i \in [\![1,n_r]\!],\ \ rot\,\widetilde{\varphi}_i \in \mathbb{Q}_{r+1,r,r}\times\mathbb{Q}_{r,r+1,r}\times\mathbb{Q}_{r,r,r+1}(\widetilde{x},\widetilde{y},\widetilde{z})$$
- ***Prisme :***
$$\forall i \in [\![1,n_r]\!],\ \ rot\,\widetilde{\varphi}_i \in \mathbb{Q}_{r+1,r+1,r}\times\mathbb{Q}_r\times\mathbb{Q}_{r-1,r-1,r+1}(\widetilde{x},\widetilde{y},\widetilde{z})$$
- ***Pyramide :***
$$\forall i \in [\![1,n_r]\!],\ \ rot\,\widetilde{\varphi}_i \in \mathbb{Q}_{r+1,r,r+1}\times\mathbb{Q}_{r,r+1,r+1}\times\mathbb{Q}_{r,r,r-1}(\widetilde{x},\widetilde{y},\widetilde{z})$$
- ***Tétraèdre :***
$$\forall i \in [\![1,n_r]\!],\ \ rot\,\widetilde{\varphi}_i \in \mathbb{Q}_{r,r-1,r-1}\times\mathbb{Q}_{r-1}\times\mathbb{Q}_{r-1,r-2,r-1}(\widetilde{x},\widetilde{y},\widetilde{z})$$

Preuve.

- **Hexaèdre :** Le rotationnel des fonctions $\hat{\varphi}_i$ sur $\widetilde{Q}$ s'écrit

$$rot\,\widetilde{\varphi}_i = \begin{bmatrix} j_3\,\widetilde{x}^{i_3}\,\widetilde{y}^{j_3-1}\,\widetilde{z}^{k_3} - k_2\,\widetilde{x}^{i_2}\,\widetilde{y}^{j_2}\,\widetilde{z}^{k_2-1} \\ k_1\,\widetilde{x}^{i_1}\,\widetilde{y}^{j_1}\,\widetilde{z}^{k_1-1} - i_3\,\widetilde{x}^{i_3-1}\,\widetilde{y}^{j_3}\,\widetilde{z}^{k_3} \\ i_2\,\widetilde{x}^{i_2-1}\,\widetilde{y}^{j_2}\,\widetilde{z}^{k_2} - j_1\,\widetilde{x}^{i_1}\,\widetilde{y}^{j_1-1}\,\widetilde{z}^{k_1} \end{bmatrix} \quad \begin{matrix} i_1,j_2,k_3\leqslant r-1 \\ i_2,i_3,j_1,j_3,k_1,k_2\leqslant r+1 \end{matrix}$$

  d'où le résultat.
- **Prisme :** Le rotationnel des fonctions $\hat{\varphi}_i$ sur $\widetilde{Q}$ est

$$\begin{aligned} rot\,\widetilde{\varphi}_i \quad = \quad & \begin{bmatrix} j_3\,\widetilde{x}^{i_3}(1-\widetilde{y})^{i_3}\widetilde{y}^{j_3-1}\,\widetilde{z}^{k_3} - k_2\,\widetilde{x}^{i_2}(1-\widetilde{y})^{i_2}\widetilde{y}^{j_2}\,\widetilde{z}^{k_2-1} \\ k_1\widetilde{x}^{i_1}(1-\widetilde{y})^{i_1}\widetilde{y}^{j_1}\widetilde{z}^{k_1-1} - i_3\,\widetilde{x}^{i_3-1}(1-\widetilde{y})^{i_3-1}\widetilde{y}^{j_3}\,\widetilde{z}^{k_3} \\ i_2\,\widetilde{x}^{i_2-1}(1-\widetilde{y})^{i_2-1}\widetilde{y}^{j_2}\,\widetilde{z}^{k_2} - j_1\,\widetilde{x}^{i_1}(1-\widetilde{y})^{i_1}\widetilde{y}^{j_1-1}\,\widetilde{z}^{k_1} \end{bmatrix} \quad \begin{matrix} i_1+j_1,i_2+j_2,k_3\leqslant r-1 \\ i_3+j_3,k_1,k_2,k\leqslant r+1 \end{matrix} \\ & + \begin{bmatrix} k\,\widetilde{x}^{i+1}(1-\widetilde{y})^{i+1}\widetilde{y}^{j}\,\widetilde{z}^{k-1} \\ k\,\widetilde{x}^{i}(1-\widetilde{y})^{i}\widetilde{y}^{j+1}\,\widetilde{z}^{k-1} \\ -(i-j-1)\widetilde{x}^{i}(1-\widetilde{y})^{i}\widetilde{y}^{j}\,\widetilde{z}^{k} \end{bmatrix} \quad \begin{matrix} i+j=r-1 \\ 0\leqslant i,j\leqslant r-1 \\ 0\leqslant k\leqslant r+1 \end{matrix} \end{aligned}$$

  ce qui permet d'obtenir le résultat annoncé.

– **Pyramide :** Le rotationnel des fonctions $\hat{\varphi}_i$ sur $\breve{Q}$ est

$$
\begin{aligned}
rot\,\widetilde{p} \;=\; & \begin{bmatrix} j_3\,\widetilde{x}^{i_3}\,\widetilde{y}^{j_3-1}\,(1-\widetilde{z})^{i_3+j_3-1}\widetilde{z}^{k_3} - k_2\,\widetilde{x}^{i_2}\,\widetilde{y}^{j_2}\,(1-\widetilde{z})^{i_2+j_2}\widetilde{z}^{k_2-1} \\ k_1\,\widetilde{x}^{i_1}\,\widetilde{y}^{j_1}\,(1-\widetilde{z})^{i_1+j_1}\widetilde{z}^{k_1-1} - i_3\,\widetilde{x}^{i_3-1}\,\widetilde{y}^{j_3}\,(1-\widetilde{z})^{i_3+j_3-1}\widetilde{z}^{k_3} \\ i_2\,\widetilde{x}^{i_2-1}\,\widetilde{y}^{j_2}\,(1-\widetilde{z})^{i_2+j_2-1}\widetilde{z}^{k_2} - j_1\,\widetilde{x}^{i_1}\,\widetilde{y}^{j_1-1}\,(1-\widetilde{z})^{i_1+j_1-1}\widetilde{z}^{k_1} \end{bmatrix} \quad \begin{matrix} i_1+j_1+k_1 \leqslant r-1 \\ i_2+j_2+k_2 \leqslant r-1 \\ i_3+j_3+k_3 \leqslant r-1 \end{matrix} \\
& + \begin{bmatrix} (r+j_3-k_3)\,\widetilde{x}^{r+i_3-k_3}\,\widetilde{y}^{r+j_3-k_3-1}\,(1-\widetilde{z})^{r+i_3+j_3-k_3-1} \\ (r-k_1)\,\widetilde{x}^{r+i_1-k_1}\,\widetilde{y}^{r+j_1-k_1}\,(1-\widetilde{z})^{r+i_1+j_1-k_1-1} \\ (r+i_2-k_2)\,\widetilde{x}^{r+i_2-k_2-1}\,\widetilde{y}^{r+j_2-k_2}\,(1-\widetilde{z})^{r+i_2+j_2-k_2-1} \end{bmatrix} \quad \begin{matrix} i_1+j_1 \leqslant k_1 \\ i_2+j_2 \leqslant k_2 \\ i_3+j_3 \leqslant k_3 \\ 0 \leqslant k_1,k_2,k_3 \leqslant r-1 \end{matrix} \\
& - \begin{bmatrix} (r-k_2)\,\widetilde{x}^{r+i_2-k_2}\,\widetilde{y}^{r+j_2-k_2}\,(1-\widetilde{z})^{r+i_2+j_2-k_2-1} \\ (r+i_3-k_3)\,\widetilde{x}^{r+i_3-k_3-1}\,\widetilde{y}^{r+j_3-k_3}\,(1-\widetilde{z})^{r+i_3+j_3-k_3-1} \\ (r+j_1-k_1)\,\widetilde{x}^{r+i_1-k_1}\,\widetilde{y}^{r+j_1-k_1-1}\,(1-\widetilde{z})^{r+i_1+j_1-k_1-1} \end{bmatrix} \\
& + \begin{bmatrix} -(p+1)\,\widetilde{x}^{p+1}\,\widetilde{y}^{p}\,\widetilde{z}(1-\widetilde{z})^{p-1} \\ (p+1)\,\widetilde{x}^{p}\,\widetilde{y}^{p+1}\,\widetilde{z}(1-\widetilde{z})^{p-1} \\ 0 \end{bmatrix} \quad 0 \leqslant p \leqslant r-1 \\
& + \begin{bmatrix} (n+2)\,\widetilde{x}^{m+1}\,\widetilde{y}^{n+1}\,(1-\widetilde{z})^{m} \\ 0 \\ -(n+2)\,\widetilde{x}^{m}\,\widetilde{y}^{n+1}\,(1-\widetilde{z})^{m} \end{bmatrix} \quad 0 \leqslant m \leqslant n \leqslant r-2 \\
& + \begin{bmatrix} 0 \\ -(n+2)\,\widetilde{x}^{n+1}\,\widetilde{y}^{m+1}\,(1-\widetilde{z})^{n} \\ (n+2)\,\widetilde{x}^{n+1}\,\widetilde{y}^{m}\,(1-\widetilde{z})^{n} \end{bmatrix} \quad 0 \leqslant m \leqslant n \leqslant r-2 \\
& + \begin{bmatrix} q\,\widetilde{x}^{p+1}\,\widetilde{y}^{q-1}\,(1-\widetilde{z})^{r+1} \\ \widetilde{x}^{p}\,\widetilde{y}^{q}\,(1-\widetilde{z})^{r-1}\,((p+q-r)-(p+1)(1-\widetilde{z})^{2} \\ -q\,\widetilde{x}^{p}\,\widetilde{y}^{q-1}\,(1-\widetilde{z})^{r-1} \end{bmatrix} \quad \begin{matrix} 0 \leqslant p \leqslant r-1 \\ 0 \leqslant q \leqslant r+1 \end{matrix} \\
& + \begin{bmatrix} -\widetilde{x}^{q}\,\widetilde{y}^{p}\,(1-\widetilde{z})^{r-1}((p+q-r)-(1-\widetilde{z})^{2}p+1) \\ -q\,\widetilde{x}^{q-1}\,\widetilde{y}^{p+1}\,(1-\widetilde{z})^{r+1} \\ q\,\widetilde{x}^{q-1}\,\widetilde{y}^{p}\,(1-\widetilde{z})^{r-1} \end{bmatrix} \quad \begin{matrix} 0 \leqslant p \leqslant r-1 \\ 0 \leqslant q \leqslant r+1 \end{matrix}
\end{aligned}
$$

d'où le résultat.

– **Tétraèdre :** Le rotationnel de $\hat{\varphi}_i$ sur $\widetilde{Q}$ est

$$
\begin{aligned}
\widetilde{p} \;=\; & \begin{bmatrix} j_3\,\widetilde{x}^{i_3}\,(1-\widetilde{y})^{i_3}\widetilde{y}^{j_3-1}\,(1-\widetilde{z})^{i_3+j_3-1}\widetilde{z}^{k_3} - k_2\,\widetilde{x}^{i_2}\,(1-\widetilde{y})^{i_2}\widetilde{y}^{j_2}\,(1-\widetilde{z})^{i_2+j_2}\widetilde{z}^{k_2-1} \\ k_1\,\widetilde{x}^{i_1}\,(1-\widetilde{y})^{i_1}\widetilde{y}^{j_1}\,(1-\widetilde{z})^{i_1+j_1}\widetilde{z}^{k_1-1} - i_3\,\widetilde{x}^{i_3-1}\,(1-\widetilde{y})^{i_3-1}\widetilde{y}^{j_3}\,(1-\widetilde{z})^{i_3+j_3-1}\widetilde{z}^{k_3} \\ i_2\,\widetilde{x}^{i_2-1}\,(1-\widetilde{y})^{i_2-1}\widetilde{y}^{j_2}\,(1-\widetilde{z})^{i_2+j_2-1}\widetilde{z}^{k_2} - j_1\,\widetilde{x}^{i_1}\,(1-\widetilde{y})^{i_1}\widetilde{y}^{j_1-1}\,(1-\widetilde{z})^{i_1+j_1-1}\widetilde{z}^{k_1} \end{bmatrix} \quad \begin{matrix} i_1+j_1+k_1 \leqslant r-1 \\ i_2+j_2+k_2 \leqslant r-1 \\ i_3+j_3+k_3 \leqslant r-1 \end{matrix} \\
& + \begin{bmatrix} -j\,\widetilde{x}^{i}\,(1-\widetilde{y})^{i}\widetilde{y}^{j-1}\,(1-\widetilde{z})^{i+j-1}\widetilde{z}^{r-i-j} \\ (r-j+1)\,\widetilde{x}^{i-1}\,(1-\widetilde{y})^{i-1}\widetilde{y}^{j}\,(1-\widetilde{z})^{i+j-1}\widetilde{z}^{r-i-j} \\ -j\,\widetilde{x}^{i-1}\,(1-\widetilde{y})^{i-1}\widetilde{y}^{j-1}\,(1-\widetilde{z})^{i+j-2}\widetilde{z}^{r-i-j+1} \end{bmatrix} \quad \begin{matrix} i+j \leqslant r \\ 1 \leqslant i \leqslant r \\ 0 \leqslant j \leqslant r-1 \end{matrix} \\
& + \begin{bmatrix} -(r-i+1)\,\widetilde{x}^{i}\,(1-\widetilde{y})^{i}\widetilde{y}^{j-1}\,(1-\widetilde{z})^{i+j-1}\widetilde{z}^{r-i-j} \\ i\,\widetilde{x}^{i-1}\,(1-\widetilde{y})^{i-1}\widetilde{y}^{j}\,(1-\widetilde{z})^{i+j-1}\widetilde{z}^{r-i-j} \\ i\,\widetilde{x}^{i-1}\,(1-\widetilde{y})^{i-1}\widetilde{y}^{j-1}\,(1-\widetilde{z})^{i+j-2}\widetilde{z}^{r-i-j+1} \end{bmatrix} \quad \begin{matrix} i+j \leqslant r \\ 0 \leqslant i \leqslant r-1 \\ 1 \leqslant j \leqslant r \end{matrix} \\
& + \begin{bmatrix} 0 \\ 0 \\ -(i+j)\,\widetilde{x}^{i-1}\,(1-\widetilde{y})^{i-1}\widetilde{y}^{j-1}\,(1-\widetilde{z})^{i+j-2} \end{bmatrix} \quad \begin{matrix} i+j=r+1 \\ 1 \leqslant i,j \leqslant r \end{matrix}
\end{aligned}
$$

ce qui achève la démonstration. □

Proposition 11.1.4 *Pour avoir l'intégration exacte de la matrice de rigidité dans le cas affine, la formule de quadrature utilisée doit être exacte pour les polynômes de :*

– ***Hexaèdre*** *:* $\mathbb{Q}_{2r+2}(\widetilde{x},\widetilde{y},\widetilde{z})$ *;*
– ***Prisme*** *:* $(1-\widetilde{y})\,\mathbb{Q}_{2r+2}(\widetilde{x},\widetilde{y},\widetilde{z})$ *;*
– ***Pyramide*** *:* $(1-\widetilde{z})^2\,\mathbb{Q}_{2r+2}(\widetilde{x},\widetilde{y},\widetilde{z})$ *;*
– ***Tetraèdre*** *:* $(1-\widetilde{y})(1-\widetilde{z})^2\,\mathbb{Q}_{2r,2r-2,2r-2}(\widetilde{x},\widetilde{y},\widetilde{z})$.

Preuve. En utilisant le lemme 11.1.3, via l'écriture de la matrice de masse sur le cube unité $\widetilde{Q}$ donnée par l'équation 11.1.3, le résultat est obtenu en additionnant les puissances. □

11.1.2 Formule de quadrature

Puisque l'on ne peut pas intégrer les matrices de masse et de rigidité dans le cas non affine, on doit se contenter d'une intégration approchée. On choisit d'utiliser dans le cas non affine les mêmes formules de quadrature qui permettent d'intégrer exactement les matrices dans le cas affine.

En pratique, on utilise les formules de quadrature suivantes pour chacun des types d'élément

- **Hexaèdre :**
  - *hp* :
  $$(\xi^G_{r+1},\xi^G_{r+1},\xi^G_{r+1}),(\omega^G_{r+1},\omega^G_{r+1},\omega^G_{r+1})$$
  - nodal :
  $$(\xi^{GL}_{r+1},\xi^{GL}_{r+1},\xi^{GL}_{r+1}),(\omega^{GL}_{r+1},\omega^{GL}_{r+1},\omega^{GL}_{r+1})$$
- **Prisme :**
  - *hp* :
  $$(\xi^{tri}_{r+1},\xi^G_{r+1}),(\omega^{tri}_{r+1},\omega^G_{r+1})$$
  - nodal :
  $$(\xi^{tri}_{r+1},\xi^{GL}_{r+1}),(\omega^{tri}_{r+1},\omega^{GL}_{r+1})$$
- **Pyramide :**
  $$(\xi^G_{r+1},\xi^G_{r+1},\xi^{GJ2}_{r+1}),(\omega^G_{r+1}\omega^G_{r+1}\omega^{GJ2}_{r+1}),$$
- **Tétraèdre :**
  $$(\xi^{tetra}_r,\omega^{tetra}_r)$$

où

- $(\xi^G_{r+1},\omega^G_{r+1})$ est la formule de quadrature de Gauss d'ordre $r+1$, exacte pour les polynômes de $\mathbb{Q}_{2r+3}$,
- $(\xi^{GL}_{r+1},\omega^{GL}_{r+1})$ est la formule de quadrature de Gauss-Lobatto d'ordre $r+1$, exacte pour les polynômes de $\mathbb{Q}_{2r+1}$,
- $(\omega^{GJa}_{r+1},\xi^{GJa}_{r+1})$ est la formule de quadrature de Gauss-Jacobi d'ordre $r+1$, exacte pour les polynômes de $(1-x)^a\mathbb{Q}_{2r+3}$.
- $(\xi^{tri}_{r+1},\omega^{tri}_{r+1})$ est une formule de quadrature pour le triangle d'ordre $r+1$, exacte pour les polynômes de $\mathbb{P}_{2r+2}$, comme par exemple celle décrite par Dunavant [27],
- $(\xi^{tetra}_r,\omega^{tetra}_r)$ est une formule de quadrature d'ordre r pour le tétraèdre d'ordre, exacte pour les polynômes de $\mathbb{P}_{2r}$, comme par exemple celle décrite par Šolín *et al.* [71].

Finalement, on utilise au maximum $(r+2)^3$ points d'intégration pour tous les types d'élément, au lieu de $(r+1)^3$ pour la première famille (voir chapitre 13, section 13.1.3)

11.2 Estimation d'erreur abstraite

11.2.1 Présentation du problème

On considère le problème variationnel standard suivant

$$\left\{\begin{array}{l}\text{Trouver } u\in V \text{ tel que}\\ \forall v\in V,\ a(u,v)=f(v),\end{array}\right. \tag{11.2.1}$$

où $V=H(rot,\Omega)$, et où $a(.,.)$ désigne une forme bilinéaire continue et coercive, et $f(.)$ un forme linéaire continue. Pour un sous-espace de dimension finie V_h de l'espace V, le problème discret s'écrit alors

$$\left\{\begin{array}{l}\text{Trouver } u_h\in V_h \text{ tel que}\\ \forall v_h\in V_h,\ a_h(u_h,v_h)=f_h(v_h),\end{array}\right. \tag{11.2.2}$$

où $a_h(.,.)$ désigne une forme bilinéaire définie sur l'espace V_h, uniformément V_h-elliptique, et $f_h(.)$ une forme linéaire définie sur l'espace V_h.

On considèrera le cas simple suivant

$$a(u,v)=\int_\Omega u\cdot v+\int_\Omega rot\,u\cdot rot\,v.$$

Comme dans le cas H^1, on considère le lemme de Strang 3.2.1 et on étudie séparément l'erreur d'interpolation et l'erreur de quadrature. On supposera également que u est dans $H^{r+1}(\Omega)$ pour Ω suffisamment régulier.

11.2.2 Lemme de Strang

On considère la version suivante du lemme de Strang

Lemme 11.2.1 *(Lemme de Strang). Si u est solution de (11.2.1) et u_h est solution de (11.2.2), il existe une constante $C > 0$ ne dépendant pas du pas d'espace h telle que*

$$\|u - u_h\|_{rot} \leqslant C \underbrace{\inf_{v_h \in V_h} \left\{ \|u - v_h\|_{rot,\Omega} \right.}_{erreur\ d'interpolation} + \underbrace{\left. \sup_{w_h \in V_h} \frac{|a(v_h, w_h) - a_h(v_h, w_h)|}{\|w_h\|_{rot,\Omega}} \right\}}_{erreur\ d'intégration\ numérique}.$$

Preuve. En remarquant que $V_h \subset V$, la preuve de cette version du lemme de Strang est similaire à la preuve proposée par Ciarlet [14]. □

On étudie à présent séparément les deux termes du membre de droite de l'inégalité du lemme de Strang, c'est à dire l'erreur d'interpolation et l'erreur de quadrature. On supposera aussi que u est dans $H^{r+1}(\Omega)$ pour Ω suffisamment régulier.

11.2.3 Erreur d'interpolation

Soit Ω un ouvert lipschitzien de $\mathbb{R}^3$ composé de n_e éléments K

$$\Omega = \bigcup_K K.$$

On note

$$h_K = \text{diam } (K) = \sup_{(x,y)\in K} |x - y|, \quad h = \max_K h_K$$
$$\rho_K = \sup_B \{\text{diam } (B),\ B \text{ boule incluse dans } K\}$$

et on suppose que le maillage est tel qu'il existe $\sigma > 0$ tel que

$$\frac{h_K}{\rho_K} \leqslant \sigma$$

Pour estimer l'erreur d'interpolation, on a besoin d'un résultat d'interversion entre l'opérateur rot et un opérateur de projection. On définit au préalable l'espace suivant

Définition 11.2.2 *L'espace d'approximation pour l'approximation $H(div)$ sur les tétraèdres est*

$$\mathcal{D}_r = \mathbb{P}^3_{r-1} \oplus \tilde{\mathbb{P}}_{r-1} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Lemme 11.2.3 *Il existe $\pi_{\mathcal{D}_r}$ interpolant sur $\mathcal{D}_r$ et $\pi_{\mathcal{R}_r}$ interpolant sur $\mathcal{R}_r$ tels que*

$$\forall u \in H(rot),\ \pi_{\mathcal{D}_r}(rot\ u) = rot\ (\pi_{\mathcal{R}_r} u)$$

Preuve. Voir Monk [55].

Définition 11.2.4 *Soit $I_h \in \mathcal{L}(H^r(rot,\Omega), H(rot,\Omega))$ un opérateur tel que*

$$\forall u_h \in V_h, I_h u_h = u_h$$

et

$$\forall u \in H(rot,\Omega),\ ||I_h u||_{rot,\Omega} \leqslant C_{rot,\Omega} ||u||_{rot,\Omega}$$

On notera I_h^K la restriction de I_h à un élément K.

On suppose que l'interpolant I_h est borné dans $H(rot)$, et que les projecteurs $\pi_{\mathcal{R}_r}$ et $\pi_{\mathcal{D}_r}$ sont bornés dans $L^2(K)$. On notera

$$\begin{array}{ll} \forall u \in L^2(\Omega), & ||I_h u||_{0,\Omega} \leqslant C_{0,\Omega}||u||_{0,\Omega} \\ \forall u \in H(rot,\Omega), & ||I_h u||_{rot,\Omega} \leqslant C_{rot,\Omega}||u||_{rot,\Omega} \end{array}$$

et puisque $\pi_{\mathcal{R}_r}$ et $\pi_{\mathcal{D}_r}$ sont les projecteurs orthogonaux,

$$\forall u \in L^2(K), \quad \begin{array}{l} ||\pi_{\mathcal{R}_r} u||_{0,K} \leqslant C_{\mathcal{R}_r}||u||_{0,K} \\ ||\pi_{\mathcal{D}_r} u||_{0,K} \leqslant C_{\mathcal{D}_r}||u||_{0,K} \end{array}$$

Proposition 11.2.5 *Pour $u \in H^r(rot,\Omega)$, il existe une constante $C_\Omega > 0$ ne dépendant que de r telle que*

$$||u - I_h u||_{rot,\Omega} \leqslant C_\Omega\, h^r\, ||u||_{r,rot,\Omega}.$$

où h désigne la longueur caractéristique maximale sur tous les éléments K du maillage.

Preuve. Pour $u \in H^r(rot,\Omega)$, on utilise l'inégalité suivante pour se ramener à des estimations locales

$$||u - I_h u||^2_{rot,\Omega} = \sum_K ||u - I_h u||^2_{rot,K}.$$

L'inégalité triangulaire donne

$$||u - I_h u||_{rot,K} \leqslant ||u - \pi_{\mathcal{R}_r} u||_{rot,K} + ||\pi_{\mathcal{R}_r} u - I_h u||_{rot,K}$$

Puisque $\mathcal{R}_r \subset P_r^F$, on a $I_h \pi_{\mathcal{R}_r} u = \pi_{\mathcal{R}_r} u$, d'où

$$||u - I_h u||_{rot,K} \leqslant (1 + ||I_h||_{H(rot)})\, ||u - \pi_{\mathcal{R}_r} u||_{rot,K}$$

Or I_h est borné dans $H(rot)$, en prenant $C'_{rot,K} = 1 + C_{rot,K}$, on a

$$||u - I_h u||_{rot,K} \leqslant C'_{rot,K}\, ||u - \pi_{\mathcal{R}_r} u||_{rot,K}$$

En utilisant le lemme d'interversion 11.2.3, on a

$$||rot\ (u - \pi_{\mathcal{R}_r})u||_{0,K} = ||rot\ u - \pi_{\mathcal{D}_r} rot\ u||_{0,K}$$

Comme $\mathbb{P}^3_{r-1} \in \mathcal{R}_r$ et $\mathbb{P}^3_{r-1} \in \mathcal{D}_r$ et que les projecteurs $\pi_{\mathcal{R}_r}$ et $\pi_{\mathcal{D}_r}$ sont bornés dans L^2, on a

$$\begin{array}{lcl} ||u - \pi_{\mathcal{R}_r} u||_{0,K} & \leqslant & (1 + C_{\mathcal{R}_r})\, ||u - \pi_{r-1} u||_{0,K} \\ ||rot\ u - \pi_{\mathcal{D}_r} rot\ u||_{0,K} & \leqslant & (1 + C_{\mathcal{D}_r})\, ||rot\ u - \pi_{r-1} rot\ u||_{0,K} \end{array}$$

Or, d'après le lemme de Bramble-Hilbert sur un élément K pour $m = 0$, $s = r - 1$ et $\Pi = \pi_{r-1}$, on a les estimations locales suivantes

$$\begin{array}{lcl} ||u - \pi_{r-1} u||_{0,K} & \leqslant & C_{0,K}\, h_K^r\, |u|_{r,K} \\ ||rot\ u - \pi_{r-1} rot\ u||_{0,K} & \leqslant & C_{rot,K}\, h_K^r\, |rot\ u|_{r,K} \end{array}$$

Ainsi, en prenant $C_K = C_{0,K}(1 + C_{\mathcal{R}_r}) + (1 + C_{\mathcal{D}_r})C_{rot,K}$ et en utilisant l'inégalité des normes 1.1.1, on obtient finalement

$$||u - I_h u||^2_{rot,K} \leqslant C_K^2 h_K^{2r} ||u||^2_{r,rot,K}$$

En prenant $C_\Omega = \max_K C_K$ et $h = \max_K h_K$, on obtient ainsi

$$||u - I_h u||^2_{rot,\Omega} \leqslant C_\Omega^2\, h^{2r} \sum_K ||u||^2_{r,rot,K}.$$

soit, en réutilisant l'inégalité de Cauchy-Schwarz discrète et en prenant la racine,

$$||u - I_h u||_{rot,\Omega} \leqslant C_\Omega\, h^r\, ||u||_{r,rot,\Omega}$$

qui est le résultat annoncé. □

Remarque 11.2.6 *La condition $\max_K C_K$ borné est vérifiée dans le cas d'un maillage périodique dans les études numériques qui suivront puisque le nombre de formes de chaque type d'élément utilisé dans le maillage est fini. Dans un cas plus général, on conjecture que l'aspect borné de C_K est lié à l'existence d'une borne supérieure pour l'inverse de la matrice jacobienne DF, comme dans le cas des hexaèdres (Girault et Raviart [35]).*

Remarque 11.2.7 *Au vu de la preuve, l'espace de dimension minimale permettant d'obtenir une erreur d'interpolation sur l'élément en $O(h^r)$ pour la norme $H(rot)$ pour une solution suffisamment régulière est conditionné par l'espace de dimension minimale permettant d'obtenir une erreur d'interpolation sur l'élément en $O(h^r)$ pour la norme $H(div)$. En effet, en notant ces deux espaces A et B respectivement, une condition nécessaire pour avoir une convergence en $O(h^r)$ pour la norme $H(rot)$ est la suivante*

$$\begin{gathered}\mathbb{P}^3_{r-1} \subset A \\ \mathbb{P}^3_{r-1} \subset B \\ rot\, B = Ker(div)A\end{gathered}$$

Il serait possible de construire des projecteurs π_A et π_B tels que

$$\forall u \in H(rot,\Omega),\ \pi_B rot\, u = rot\, \pi_A u$$

Prendre $A \subset P_r^F$ avec $A = \mathcal{R}_r$ et $B = \mathcal{D}_r$ convient, mais il peut exister des espaces A et B de dimension plus petite adaptés à chaque élément.

11.2.4 Erreur de quadrature

Comme dans le cas H^1, on pourrait donc également chercher la formule de quadrature minimale nous permettant d'obtenir une erreur de quadrature en $O(h^r)$.

11.2.4.1 Cas affine

Lorsque F est affine, la matrice de masse est semblable à celle de la formulation H^1, ce qui signifie que l'on peut avoir les mêmes résultats que dans ce cas (voir section 3.2.4.1).

D'après le lemme 11.1.1, on remarque que pour tous les éléments

$$\forall i \in [\![1, n_r]\!],\ \ \widetilde{\varphi}_i \in (\mathbb{Q}_{r+s_1,r+s_1,r+s_2})^3(\widetilde{x},\widetilde{y},\widetilde{z})$$

où

Hexaèdre :	$s_1 = 1,\ s_2 = 1$
Prisme :	$s_1 = 1,\ s_2 = 1$
Pyramide :	$s_1 = 1,\ s_2 = 0$
Tétraèdre :	$s_1 = 0,\ s_2 = 0$

On a le résultat suivant

Lemme 11.2.8 *Pour une formule de quadrature exacte pour les polynômes de $\widetilde{|T|}\mathbb{Q}_{m,m,n}$, on a*

$$\forall (v_h, w_h) \in P_r^F,\ E_K(v_h, w_h) = E_K(v_h - \pi_p v_h, w_h - \pi_q w_h)$$

pour $m \geqslant r + \max(p,q) + s_1$ et $n \geqslant r + \max(p,q) + s_2$ avec $p + q \leqslant r$.

Preuve. On a

$$\forall (v_h, w_h) \in P_r^F,\ E_K(v_h - \pi_p v_h, w_h - \pi_q w_h) = E_K(v_h, w_h) - E_K(v_h, \pi_q w_h) - E_K(\pi_p v_h, w_h) + E_K(\pi_p v_h, \pi_q w_h)$$

On considère tout d'abord l'intégrale

$$\int_K \pi_p v_h\, w_h\, dx\, dy\, dz.$$

Après changement de variable, on obtient

$$\int_{\widetilde{Q}} \widetilde{|T|} \cdot \widetilde{(\pi_p v_h)}\ \widetilde{w}_h\ d\widetilde{x}\, d\widetilde{y}\, d\widetilde{z}.$$

D'après le lemme 11.1.1, on a $\widetilde{(\pi_p v_h)} \cdot \widetilde{w}_h \in \mathbb{Q}_{p+r+s_1,p+r+s_1,p+r+s_2}(\widetilde{x},\widetilde{y},\widetilde{z})$, si bien que pour une formule de quadrature exacte pour les polynômes de $\widetilde{|T|}\mathbb{Q}_{m,m,n}$, on a donc

$$E_K(\pi_p v_h, w_h) = 0, \tag{11.2.3}$$

dès que $m \geqslant r + p + s_1$ et $n \geqslant r + p + s_2$.

De la même manière, pour $m \geqslant r+q+s_1$ et $n \geqslant r+q+s_2$, on a

$$E_K(v_h, \pi_q w_h) = 0$$

Lorsque $m \geqslant r+\max(p,q)+s_1$ et $n \geqslant r+\max(p,q)+s_2$ avec $p+q \leqslant r$, on a donc

$$E_K(\pi_p v_h, \pi_q w_h) = 0,$$

ce qui prouve le résultat avancé. □

Cependant, il est difficile d'obtenir un résultat semblable à celui de la proposition 3.2.10. En effet, le lemme de Bramble-Hilbert ne donne d'estimation qu'avec des normes entières, ce qui pose problème lorsque l'on veut estimer $\|w_h - \pi_0 w_h\|_{0,K}$ en fonction de $\|w_h\|_{rot,K}$.

Le cas de la matrice de rigidité reste quant à lui délicat. En effet, bien que le fait que F soit affine simplifie la structure de la matrice, plusieurs points délicats rencontrés dans le cas H^1 subsistent (voir section 3.2.4.2) notamment le fait que l'on a peu d'informations sur le lien entre $\nabla I_h u$ et $I_h \nabla u$.

11.2.4.2 Cas non-affine

Lorsque F n'est pas affine, aucune partie ni de la matrice de masse ni de la matrice de rigidité ne peut être intégrée exactement du fait de la présence de $|DF|^{-1}$, DF^{-1}, DF^{*-1} qui sont rationnels. Appliquer la même méthode que pour le cas affine est donc compromis.

Nous avons cependant constaté numériquement que la formule de quadrature présentée dans la section 11.1.2 permettait d'obtenir une erreur globale d'ordre r.

Chapitre 12

Étude numérique des éléments d'arête

On effectue une analyse de dispersion et une étude de stabilité des schémas obtenus à partir des éléments finis d'arête construits dans le chapitre 10 pour en dégager les propriétés numériques. Un cas test numérique vient confirmer le bon comportement des éléments au sein d'un maillage hybride.

Sommaire

12.1 Propriétés numériques

12.1.1 Analyse de dispersion

Comme pour les éléments continus et discontinus (voir les sections 4.1 et 8.1.1), on effectue une analyse de dispersion pour les éléments finis d'arête construits précédemment, avec les équations de Maxwell. L'étude est faite sur des cellules périodiques prises comme un cube découpé en un unique hexaèdre, deux prismes, deux pyramides et deux tétraèdres (hybride), six pyramides ou six tétraèdres, comme le montre la figure 4.1. Afin de vérifier la consistance de nos méthodes lorsque les éléments ne sont pas affines, l'analyse a également été faite sur des cellules périodiques faites à partir de cubes déformés (voir figure 4.2).

Les courbes de dispersion pour les équations de Maxwell avec les éléments réguliers d'ordre 1 à 3 sont présentées sur la figure 8.2. La même étude a été faite pour les éléments déformés, comme présenté sur la figure 8.3 pour les ordres 1 à 3. Dans les deux cas, les hexaèdres et les prismes présentent une dispersion très proche, tandis que les tétraèdres, les pyramides et la cellule hybride sont plus dispersifs, et la dispersion pour tous les éléments diminue lorsque l'on monte en ordre.

12.1.2 Étude de stabilité

Comme dans le cas continu, on souhaite étudier la condition de stabilité de ces éléments pour l'équation des ondes avec une discrétisation en temps par un schéma centré d'ordre deux. On considère un maillage périodique infini comme pour l'analyse de dispersion.

Pour chaque type d'élément, le tableau 12.1 donne la CFL obtenue jusqu'à l'ordre 4 sur des cellules régulières. La formule de quadrature considérée pour chaque type d'élément est celle présentée dans la section 11.1.2

TAB. 12.1 – Stabilité des éléments d'arête pour un maillage régulier

	Maillage régulier			
Élément	Ordre 1	Ordre 2	Ordre 3	Ordre 4
Hexaèdre	0.18257	0.10334	0.06611	0.04597
Prisme	0.171290	0.10236	0.06610	0.04602
Pyramide	0.17233	0.09209	0.05749	0.04109
Hybride	0.19891	0.11056	0.07131	0.04967
Tétraèdre	0.23983	0.12759	0.08023	0.05611

De manière générale, la condition CFL des différents types d'éléments se classe comme suit

$$CFL_{Tetra} > CFL_{Hybride} > CFL_{Hexa} > CFL_{Prisme} > CFL_{Pyramide}$$

La CFL sur les maillages hybrides est meilleure que celle obtenue sur les pyramides, ce qui constitue un résultat assez surprenant, mais qui peut s'expliquer par le nombre de degrés de liberté, moins élevé dans le cas des maillages obtenus à partir d'une cellule hybride.

Remarque 12.1.1 *Comme pour H^1, les CFL ont été vérifiées dans le cas instationnaire.*

12.2 Convergence

On souhaite vérifier l'ordre de convergence obtenu pour l'étude de dispersion. On considère les équations de Maxwell en régime harmonique (voir équation 1.3.9) sur une cavité cubique $[-1,1]^3$ avec conditions de conducteur parfait au bord. On prend $\omega = 2\pi$ et une source gaussienne polarisée selon e_x centrée à l'origine.

On étudie la convergence sur un maillage hybride avec des motifs similaires à ceux utilisés dans l'étude de dispersion et de stabilité.

On trace l'erreur obtenue en norme $H(rot)$ par rapport au pas du maillage h en échelle log-log sur la figure 12.3. On observe que l'erreur en norme $H(rot)$ est en $O(h^r)$ comme nous l'avons démontré lors des estimations d'erreur.

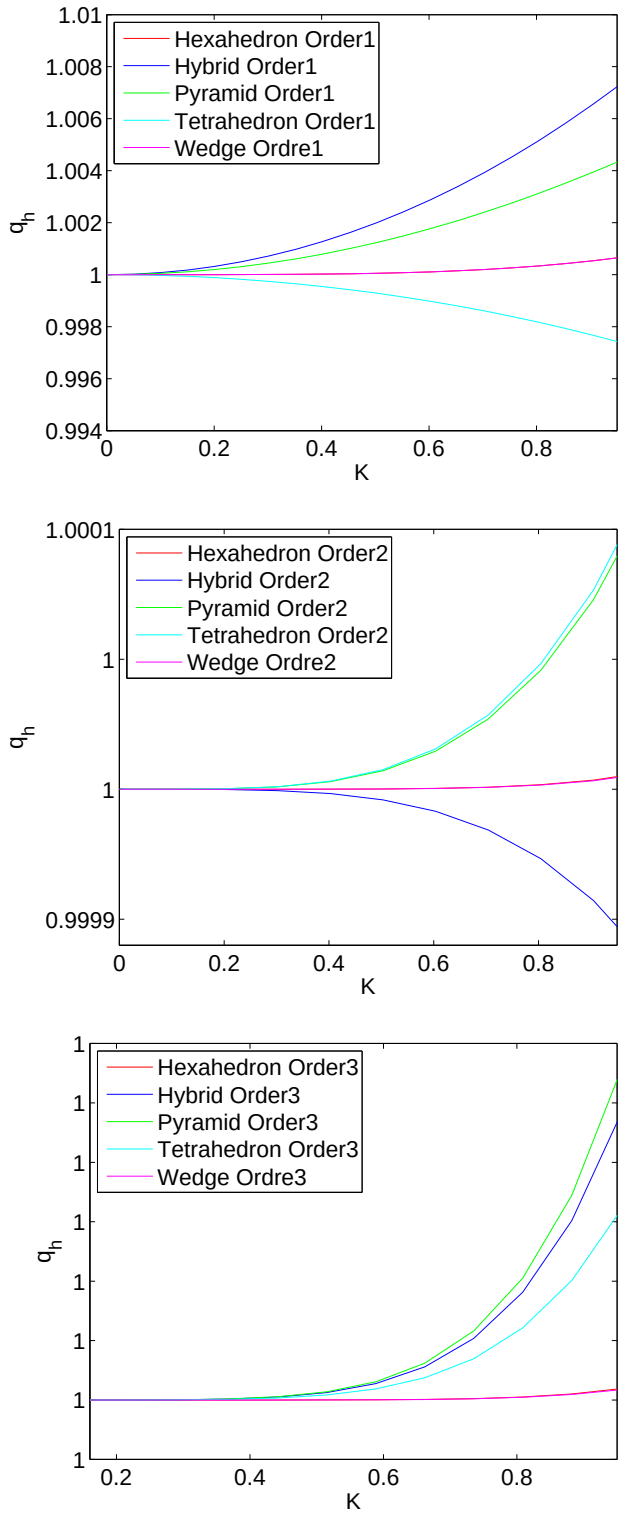

FIG. 12.1 – Courbes de dispersion pour les éléments finis d'arête d'ordre 1 à 3 sur un maillage régulier ($K = \frac{6kh}{2\pi r}$)

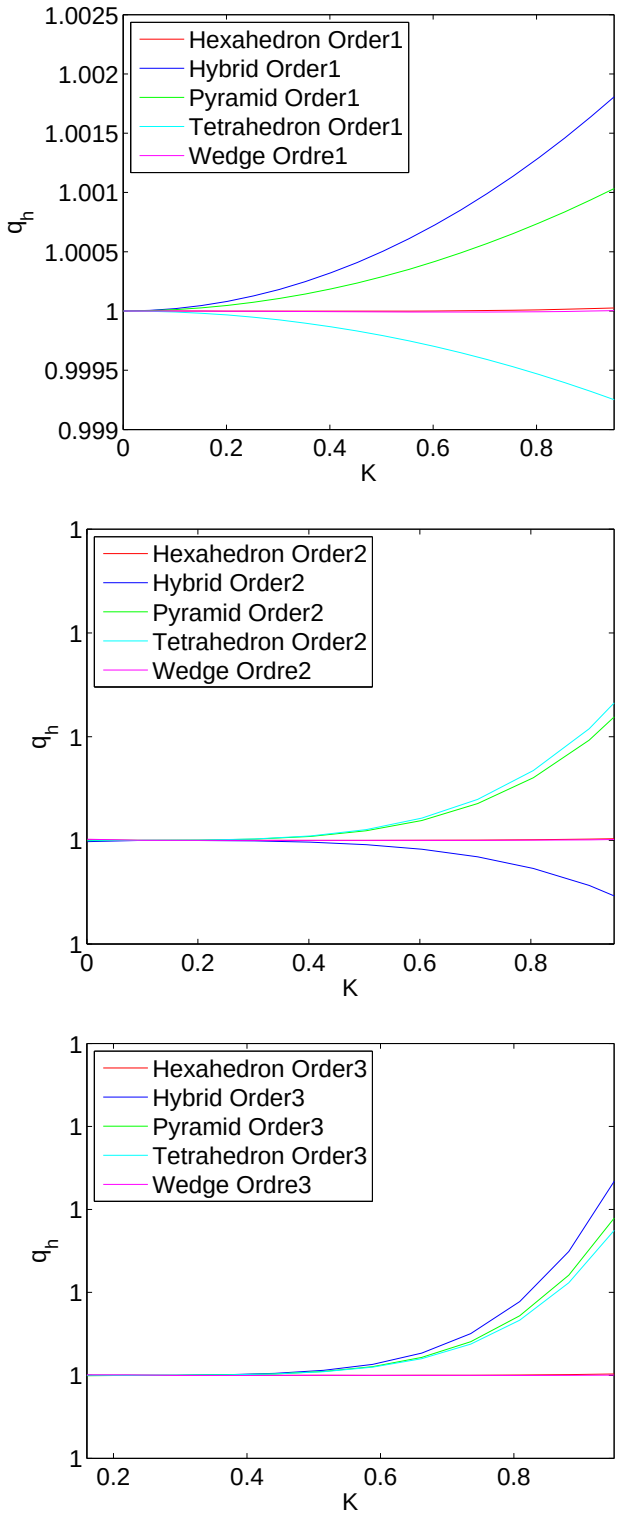

FIG. 12.2 – Courbes de dispersion pour les éléments finis d'arête d'ordre 1 à 3 sur un maillage déformé ($K = \frac{6kh}{2\pi r}$)

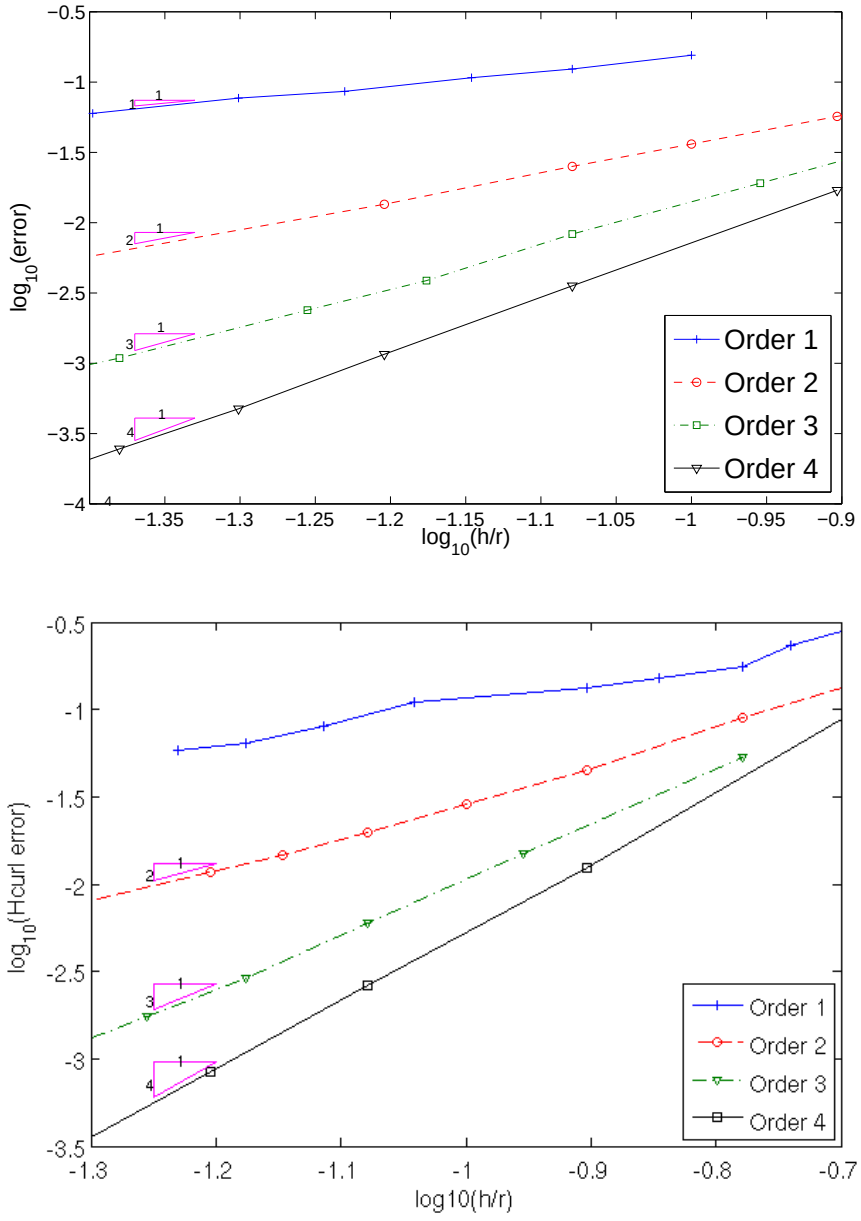

FIG. 12.3 – Erreur en norme L^2 (en haut) et $H(rot)$ (en bas) par rapport au pas du maillage h pour une cavité cubique maillée avec des cellules hybrides pour différents ordres d'approximation.

Remarque 12.2.1 *Si l'on reprend les travaux de Monk [55] qui prouve que, pour les éléments de la première famille de Nédélec [56], on a une une convergence en $O(h^r)$ pour les normes $H(rot)$ et L^2, tandis que les éléments de la seconde famille de Nédélec [57] convergent en $O(h^r)$ en norme $H(rot)$ et en $O(h^{r+1})$ en norme L^2, on peut considérer que l'on a construit les* ***éléments optimaux de la première famille****.*

12.3 Modes propres parasites

Un enjeu important lorsque l'on construit des éléments finis d'arête est de savoir s'ils sont à l'origine ou non de modes propres parasites (voir par exemple Boffi *et al.* [8], Costabel et Dauge [52]).

On vérifie ainsi que les éléments optimaux construits précédemment ne génèrent pas de valeurs propres parasites. Pour cela, on calcule les modes propres d'une cavité parallélépipédique dont on connaît les fréquences propres analytiques avec des conditions de conducteur parfait sur le bord de la cavité.

On rappelle que pour un parallélépipède (a,b,c), les fréquences propres, i.e. les valeurs propres associées à l'opérateur $rot\, rot$, sont données par l'expression suivante

$$f(n,m,p) = \frac{c_0}{2}\sqrt{\left(\frac{n}{a}\right)^2 + \left(\frac{m}{b}\right)^2 + \left(\frac{p}{c}\right)^2},\ m,n,p \geqslant 0,\ \min(m+n, m+p, n+p) > 0$$

où $c_0 = \frac{1}{\sqrt{\varepsilon\mu}}$, avec ε et μ constants. On rappelle que les modes en double 0 ne correspondent pas à des modes physiques.

Pour les ordres 1 à 4, on calcule les fréquences propres numériques obtenues dans le cas de la cavité$[-1,1]^3$ maillée à partir d'une cellule déformée hexaédrique, prismatique ou hybride pour $c_0 = 1$. On vérifie à chaque fois que les 25 premières valeurs propres sont dans le spectre théorique avec une erreur de moins de $0,1\%$ et que la multiplicité des valeurs propres est respectée.

Remarque 12.3.1 *Pour les hexaèdres de la seconde famille, Duruflé [28], [18] remarque qu'avec un maillage droit, les valeurs propres parasites apparaissent sur le spectre théorique, c'est à dire que leur valeur correspond à un mode physique, mais la multiplicité n'est plus respectée. Or le fait d'utiliser un maillage déformé permet de séparer les valeurs propres parasites du spectre physique. C'est pourquoi, dans notre cas, nous avons utilisé des cellules déformées pour mettre en valeur le fait qu'il n'y a pas de valeurs propres parasites.*

La figure 12.4 présente ainsi les valeurs propres numériques avec leur multiplicité et les valeurs théoriques pour des éléments d'ordre 3 avec les différentes cellules. La multiplicité de certaines valeurs propres est plus élevée pour le cube pour des raisons de symétrie : par exemple, la valeur propre correspondant au mode $(1,1,1)$ est de multiplicité 2 sur le cube, alors qu'elle est de multiplicité 1 sur un parallélépipède quelconque. Sur la figure 12.5, on a représenté les modes propres 12, 17 et 25 pour tous les types de cellule, à l'ordre 3.

12.4 Équations de Maxwell sur un cone-sphère

Afin de vérifier le bon fonctionnement des éléments dans un maillage hybride, on présentera un cas-test standard sur les équations de Maxwell en régime harmonique pour lequel les résultats sont bien connus.

On considère la diffraction par un cone-sphère de bord Γ placé dans une boite parallélépipédique Σ, avec des conditions de conducteur parfait sur Γ et des conditions absorbantes sur Σ. On considère le cas où le champ incident est une onde plane telle que l'onde arrive sur le cone-sphère par la pointe. On choisit ω tel que l'on ait 10 points par longueur d'onde.

La solution numérique obtenue pour un maillage hybride contenant environ un million de degrés de liberté et utilisant des éléments d'ordre 2 est donnée par la figure 12.6. Sur un maillage hybride plus fin contenant 2.2 millions de degrés de liberté, on observe une erreur de 1.8% en norme $H(rot)$ par rapport à cette solution de référence.

On lance la même simulation sur deux maillages différents (voir figure 12.7) avec des éléments d'ordre 2. Comme pour l'expérience H^1, on utilise volontairement des éléments droits et non des éléments courbes afin d'éviter la difficulté des éléments dégénérés qui peuvent apparaitre sur des maillages grossiers. En revanche, pour avoir un pas de maillage équivalent pour les deux types de maillage, il faut prendre un maillage tétraédrique très grossier, ce qui signifie que l'on approche la surface de manière très grossière.

Le nombre de degrés de liberté utilisés et l'erreur en norme L^2 et $H(rot)$ par rapport à la solution de référence sont présentés dans le tableau 12.2. Comme on pouvait s'y attendre du fait de la mauvaise approximation de la géométrie, les éléments hexaédriques donnent de piètres résultats puisque l'on obtient une précision deux fois supérieure en utilisant trois fois moins de degrés de liberté avec un maillage hybride.

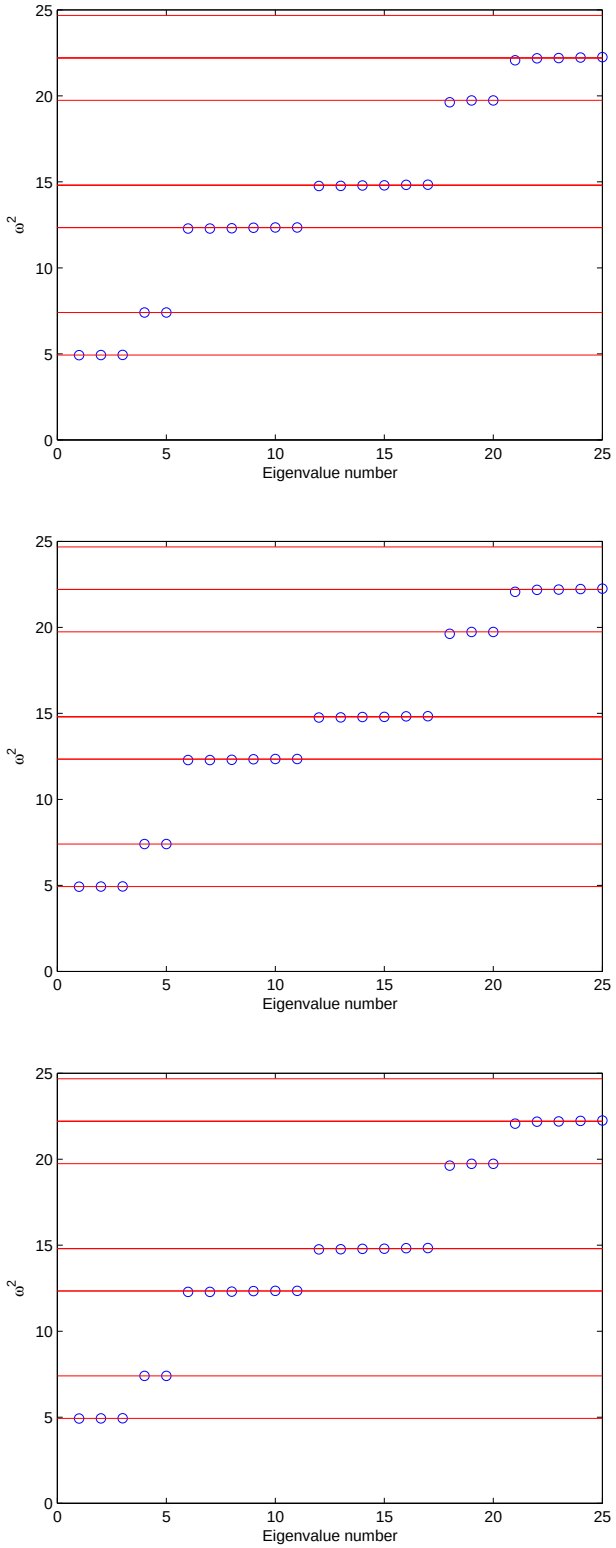

FIG. 12.4 – Valeurs propres numériques (en bleu) et valeurs propres théorique (en rouge) avec, de haut en bas, un maillage prismatique, hybride et hexaédrique pour des éléments d'ordre 3

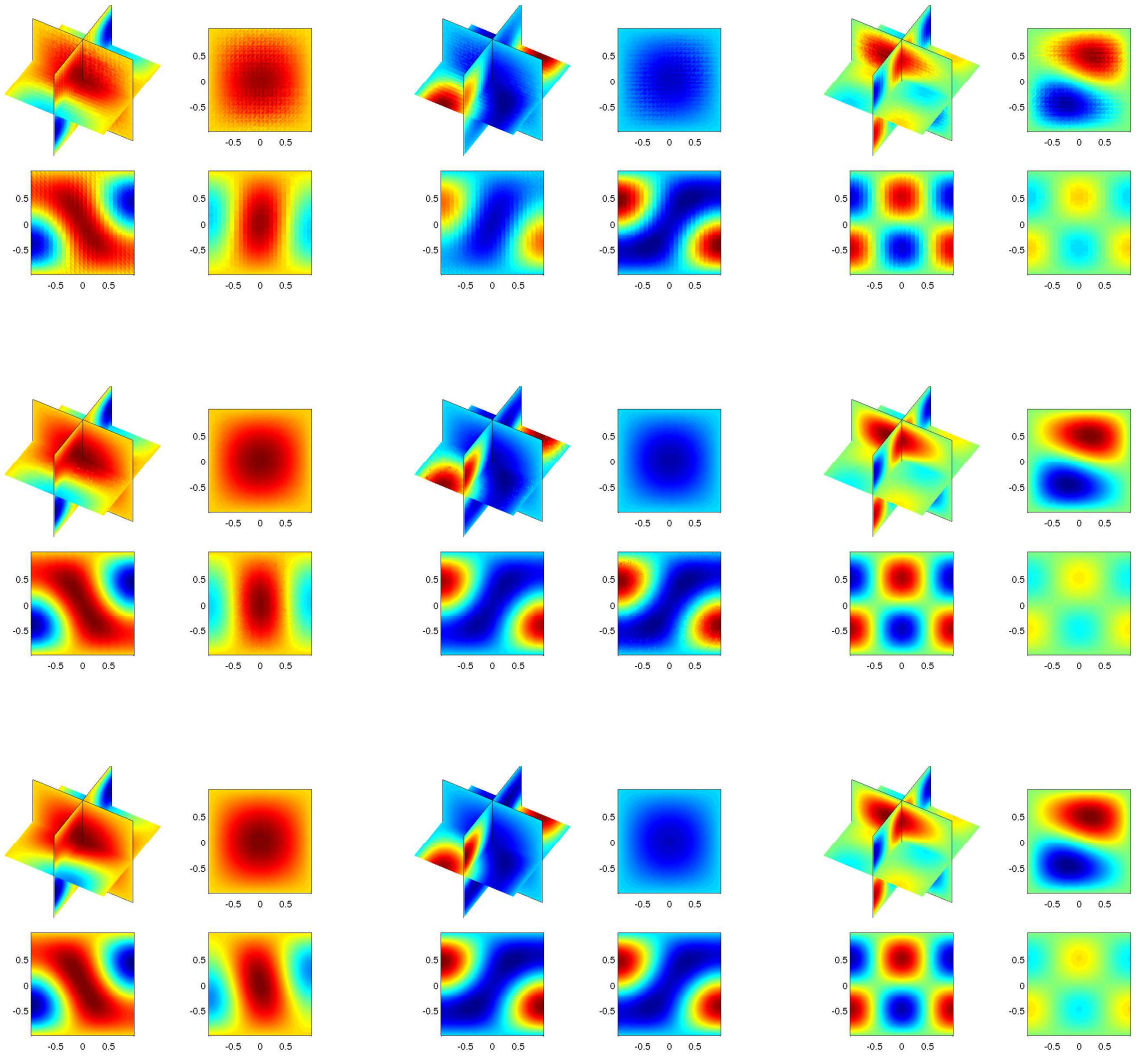

FIG. 12.5 – Modes propres 12, 17 et 25 (de gauche à droite) pour des éléments d'ordre 1 à 3 (de haut en bas) avec des maillages hybrides déformés de pas h/r constant.

TAB. 12.2 – Nombre de degrés de liberté et erreurs en norme L^2 et en norme $H(rot)$ par rapport à une solution de référence pour deux types de maillages

Type de maillage	Nombre de ddl	Erreur L^2	Erreur $H(rot)$
Tétras découpés	905 878 ddls	19.8%	19.5%
Hybride	296 469 ddls	8.87%	8.49%

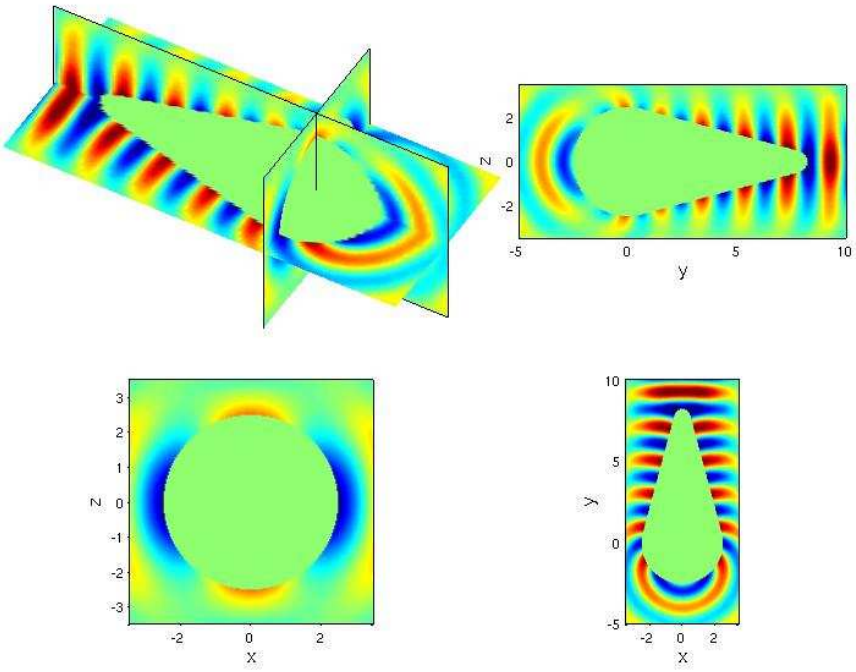

FIG. 12.6 – Partie réelle de la composante E_x du champ diffracté par le cone-sphère sur un maillage hybride pour des éléments d'ordre 3

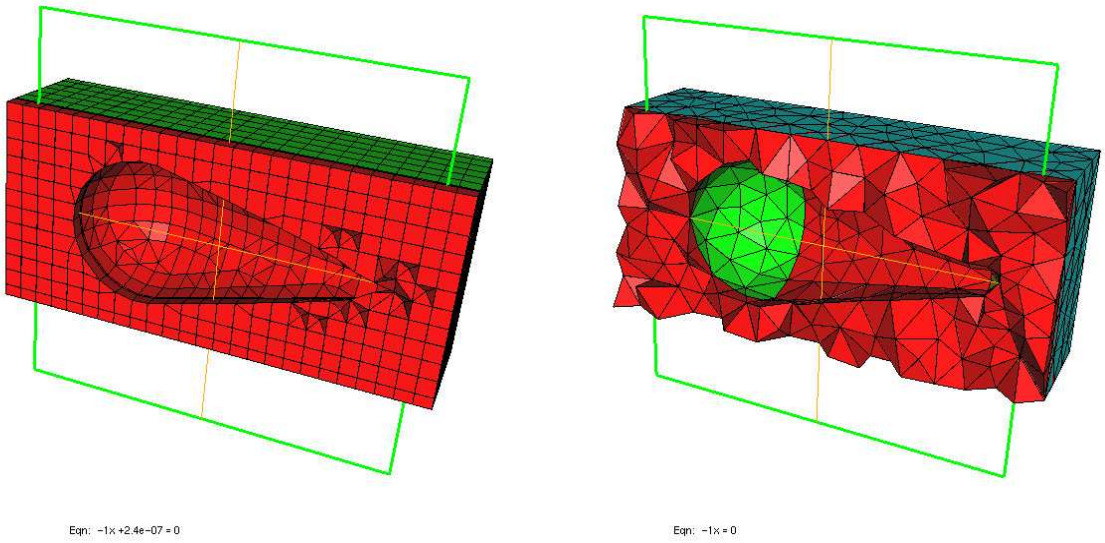

FIG. 12.7 – Maillages utilisés : maillage hybride (à gauche) et maillage tétraédrique (à droite) à la base du maillage hexaédrique utilisé

Chapitre 13

Comparaison entre différentes méthodes

Nous comparons ici les éléments optimaux obtenus dans le chapitre 10 avec ceux que l'on peut trouver dans la littérature. Des éléments finis de la première famille sont également construits sur tous les types d'éléments et comparés avec les éléments de la littérature. Une comparaison numérique est effectuée pour les deux types d'éléments ainsi construits.

Sommaire

13.1 Éléments finis d'arête

13.1.1 Introduction

Les éléments finis dits « d'arête » sont introduits par Nédélec qui construit successivement deux familles différentes pour différents types d'éléments

- Une première famille, que l'on notera $(\hat{K}, \hat{P}_r^1, \hat{\Sigma}_r^1)$ est construite pour les tétraèdres et les hexaèdres dans [56]
- Une seconde famille, désignée par $(\hat{K}, \hat{P}_r^2, \hat{\Sigma}_r^2)$ dans ce qui suit, est mise au point pour les tétraèdres, les hexaèdres et les prismes dans [57].

Ces deux familles présentent différentes propriétés qui ne seront pas discutées ici.

13.1.2 Éléments de la littérature

13.1.2.1 Tétraèdres

Pour la première famille des éléments tétraédriques, Nédélec propose $\hat{P}_r^1 = \mathcal{R}_r(\hat{x}, \hat{y}, \hat{z})$, c'est à dire

$$\hat{P}_r^1 = \hat{P}_r.$$

Ces éléments permettent effectivement d'obtenir une convergence en $O(h^r)$ pour la norme $H(rot)$ comme l'a montré Monk [55] par des estimations d'erreur.

Concernant la seconde famille de tétraèdres, l'espace d'approximation proposé est $\hat{P}_r^2 = (\mathbb{P}_r(\hat{x}, \hat{y}, \hat{z}))^3$. Or,

$$\hat{P}_r \subset \hat{P}_r^2 \subset \hat{P}_{r+1},$$

c'est à dire que la convergence en norme $H(rot)$ de ces éléments est en en $O(h^r)$, ce qu'a également montré Monk [55]. Cependant, d'après les estimations d'erreur effectuées par Monk [55], la convergence en norme L^2 est en $O(h^{r+1})$, alors qu'elle est en $O(h^r)$ pour la première famille.

13.1.2.2 Hexaèdres

Concernant les éléments hexaédriques, l'espace proposé par Nédélec pour la première famille est $\hat{P}_r^1 = \mathbb{Q}_{r-1,r,r}(\hat{x}, \hat{y}, \hat{z}) \times \mathbb{Q}_{r,r-1,r}(\hat{x}, \hat{y}, \hat{z}) \times \mathbb{Q}_{r,r,r-1}(\hat{x}, \hat{y}, \hat{z})$. On a donc

$$\hat{P}_{r-1} \subset \hat{P}_r^1 \subset \hat{P}_r,$$

ce qui signifie que les éléments hexaédriques de la première famille ne permettent d'obtenir qu'une erreur de convergence en norme $H(rot)$ en $O(h^{r-1})$ pour des éléments non affines, comme l'avait constaté numériquement Duruflé [28]). La convergence est cependant bien d'ordre r pour des éléments affines.

Pour la seconde famille, Nédélec propose de prendre $\hat{P}_r^2 = (\mathbb{Q}_r(\hat{x}, \hat{y}, \hat{z}))^3$. On a ainsi également

$$\hat{P}_{r-1} \subset \hat{P}_r^2 \subset \hat{P}_{r+1}.$$

On s'attend donc à obtenir une erreur en norme $H(rot)$ en $O(h^{r-1})$. Cependant, les résultats sur la seconde famille sont difficiles à obtenir sur les hexaèdres à cause des modes parasites (voir Duruflé [28] pour les quadrangles, pour lesquels l'ordre de convergence est très difficile à déterminer).

À la suite des travaux de Arnold, Boffi et Falk [3] sur les quadrangles pour $H(div)$, qui est le fondement des présents travaux, Falk, Gatto et Monk [32] cherchent l'espace optimal d'ordre 1 et retrouvent bien $\hat{P}_1$. Une preuve d'optimalité et des estimations d'erreur sont également proposées pour l'espace d'ordre 1.

13.1.2.3 Prismes

La première famille de Nédélec est étendue aux prismes par Monk [55] qui prend comme espace d'approximation $\hat{P}_r^1 = (\mathcal{R}_r(\hat{x}, \hat{y}) \otimes \mathbb{P}_r(\hat{z})) \times \mathbb{W}_{r,r-1}(\hat{x}, \hat{y}, \hat{z}))$. La seconde famille est proposée par Nédélec [57] qui propose $\hat{P}_r^2 = (\mathbb{W}_{r,r}(\hat{x}, \hat{y}, \hat{z}))^3$.

On a donc de manière immédiate

$$\hat{P}_{r-1} \subset \hat{P}_r^1 \subset \hat{P}_r$$
$$\hat{P}_{r-1} \subset \hat{P}_r^2 \subset \hat{P}_{r+1}$$

Dans le cas des deux familles, on s'attend donc à n'avoir une convergence qu'en $O(h^{r-1})$ avec ces éléments.

13.1.2.4 Pyramides

En ce qui concerne les pyramides, l'approche la plus générale consiste à tenter de construire les éléments de la première famille de Nédélec, plus rarement la seconde famille. Beaucoup d'auteurs utilisent pour cela les formes de Whitney.

Certains auteurs ont construit des éléments finis d'arête pour les pyramides pour les ordres 1 ou 2, souvent suffisants pour la plupart des expériences numériques.

- Coulomb, Zgainski et Maréchal [22] construisent une première famille pour les ordres 1 et 2 grâce à des fonctions de base liées aux arêtes exprimées en fonctions des bases nodales H^1. L'espace d'ordre 1 ne contient pas $\hat{P}_1$. À l'ordre 2, l'espace généré par les fonctions de base proposées contient $\hat{P}_1$ mais les fonctions de base ajoutées sur les faces s'annulent sur les autres faces, alors que seule la composante tangentielle devrait s'annuler, si bien que l'espace généré est inclus dans $\hat{P}_3$ sans l'être dans $\hat{P}_2$.
  Les auteurs construisent également une seconde famille pour l'ordre 1 à partir de la première famille. Les fonctions ne sont placées que sur les arêtes et l'espace généré est inclus dans $\hat{P}_2$ sans contenir $\hat{P}_1$.
- Gradinaru et Hiptmair [36] utilisent les formes de Whitney pour construire des éléments pyramidaux en partant d'un élément de référence cubique. À l'ordre 1, ils proposent des fonctions de base sur les arêtes dont certaines ne vérifient malheureusement pas la continuité de la composante tangentielle sur les faces. En corrigeant certains signes, on retrouve les fonctions de base de Coulomb, Zgainski et Maréchal d'ordre 1.
- Doucet *et al.* [25] présentent une généralisation des formes de Whitney pour tous les éléments et obtiennent le même espace d'ordre 1 que Coulomb, Zgainski et Maréchal et Gradinaru et Hiptmair.
- Graglia et Gheorma [38] poursuivent l'étude de Graglia *et al.* [37] sur les tétraèdres et les hexaèdres en construisant des fonctions de base nodales sur les pyramides à partir de fonctions d'arête d'ordre 1 et de points d'interpolation régulièrement répartis. À l'ordre 1, ils retrouvent également l'espace d'ordre 1 des auteurs précédents. Cependant, les fonctions de base des ordres supérieurs sont obtenues en multipliant les fonctions de base d'ordre 1 par des polynômes : les ordres supérieurs ne vont donc jamais générer $\hat{P}_1$.

Quelques auteurs construisent des éléments d'ordre quelconque sur les pyramides.

- Zaglmayr citée par Demkowicz [23] donne une expression de l'espace d'approximation d'ordre quelconque sur les pyramides pour la deuxième famille en utilisant le cube unité comme élément de référence. L'espace proposé est assorti de restrictions qui doivent permettre l'élimination de certains termes de l'espace pour obtenir la conformité $H(rot)$. L'espace d'ordre 1 comporte 4 degrés de liberté, l'espace d'ordre 2 en comporte 21, ce qui nous parait problématique pour la discrétisation. Une publication à venir de Zaglmayr permettra probablement de clarifier les propriétés de cet espace et de disposer de fonctions de base.
- Dans un premier article, Nigam et Phillips [58] présentent des fonctions de base $H(rot)$-conformes pour un ordre quelconque en utilisant une pyramide unité infinie comme élément de référence. Cependant, la composante tangentielle des fonctions liées aux arêtes verticales n'est pas polynomiale sur les deux faces adjacentes à l'arête. Après une légère modification de ces fonctions de base, l'espace obtenu d'ordre r contient $\hat{P}_{r-1}$ mais pas $\hat{P}_r$.
- Dans un second article, Nigam et Phillips [59] présentent un espace de dimension plus petite que le premier, toujours sur un élément de référence pris comme la pyramide unité infinie. Cependant, pour un ordre r quelconque, l'espace proposé ne contient pas $\hat{P}_1$, si bien que le schéma n'est pas consistant sur des pyramides non-affines.

D'autres approches existent pour traiter les éléments pyramidaux

- Des travaux très théoriques ont été réalisés par Bossavit [10] qui construit des éléments finis sur tous les types d'éléments en utilisant deux opérations simples sur les formes de Whitney [73]. Le formalisme ainsi développé est élégant, mais aucune base ni élément pratique de construction des éléments n'est donné par l'auteur.
- L'équivalent de Wieners [74], Knabner et Summ [49], et Bluck et Walker [7] pour les éléments $H(rot)$ est proposé par Marais et Davidson [53] : afin d'éviter l'utilisation des éléments pyramidaux, ils proposent de découper les pyramides en deux tétraèdres. Cependant, comme la méthode n'est pas consistante au delà de l'ordre 1 en H^1, cette approche n'a pas été testée en $H(rot)$.

13.1.3 Première famille

13.1.3.1 Espace d'approximation

Nous présentons ici une nouvelle famille d'éléments finis d'arête pour les éléments pyramidaux permettant de faire la transition entre les tétraèdres, les hexaèdres et les prismes de la première famille construits par Nédélec [56] et Monk [55].

En effet, les hexaèdres de la première famille sont relativement populaires, même si la convergence n'est pas optimale dans le cas de maillages non-réguliers. Ils restent aussi intéressants à utiliser lorsque le maillage hexaédrique est de bonne qualité, les maillages que l'on considère contenant par exemple souvent des cubes.

Définition 13.1.1 *Pour les différents types d'éléments, on prendra les espaces $\hat{P}_r^1$ suivants*

- ***Tétraèdres et transformation affine*** *:*

$$\boxed{\hat{P}_r^1 = \mathcal{R}_r(\hat{x},\hat{y},\hat{z})}$$

- ***Hexaèdres*** *:*

$$\boxed{\hat{P}_r^1 = \mathbb{Q}_{r-1,r,r}(\hat{x},\hat{y},\hat{z}) \times \mathbb{Q}_{r,r-1,r}(\hat{x},\hat{y},\hat{z}) \times \mathbb{Q}_{r,r,r-1}(\hat{x},\hat{y},\hat{z})}$$

- ***Prisme*** *:*

$$\boxed{\hat{P}_r^1 = (\mathcal{R}_r(\hat{x},\hat{y}) \otimes \mathbb{P}_r(\hat{z})) \times \mathbb{W}_{r,r-1}(\hat{x},\hat{y},\hat{z}))}$$

- ***Pyramide*** *:*

$$\begin{aligned}\hat{P}_r^1 = \mathbb{B}_{r-1}(\hat{x},\hat{y},\hat{z})^3 &\oplus \left\{ \begin{bmatrix} \frac{\hat{x}^p\,\hat{y}^{p+1}}{(1-\hat{z})^{p+1}} \\ \frac{\hat{x}^{p+1}\,\hat{y}^p}{(1-\hat{z})^{p+1}} \\ \frac{\hat{x}^{p+1}\,\hat{y}^{p+1}}{(1-\hat{z})^{p+2}} \end{bmatrix}, \quad 0 \leqslant p \leqslant r-1 \right\} \\ \oplus \left\{ \begin{bmatrix} \frac{\hat{x}^m\,\hat{y}^{n+2}}{(1-\hat{z})^{m+1}} \\ 0 \\ \frac{\hat{x}^{m+1}\,\hat{y}^{n+2}}{(1-\hat{z})^{m+2}} \end{bmatrix} \oplus \begin{bmatrix} 0 \\ \frac{\hat{x}^{n+2}\,\hat{y}^m}{(1-\hat{z})^{m+1}} \\ \frac{\hat{x}^{n+2}\,\hat{y}^{m+1}}{(1-\hat{z})^{m+2}} \end{bmatrix}, \quad 0 \leqslant m \leqslant n \leqslant r-2 \right\} \\ \oplus \left\{ \begin{bmatrix} \frac{\hat{x}^p\,\hat{y}^q}{(1-\hat{z})^{p+q-r}} \\ 0 \\ \frac{\hat{x}^{p+1}\,\hat{y}^q}{(1-\hat{z})^{p+q+1-r}} \end{bmatrix} \oplus \begin{bmatrix} 0 \\ \frac{\hat{x}^q\,\hat{y}^p}{(1-\hat{z})^{p+q-r}} \\ \frac{\hat{x}^q\,\hat{y}^{p+1}}{(1-\hat{z})^{p+q+1-r}} \end{bmatrix}, \quad \begin{matrix} 0 \leqslant p \leqslant r-1 \\ 0 \leqslant q \leqslant r \end{matrix} \right\}\end{aligned}$$

Les espaces pour le tétraèdre, le prisme et l'hexaèdre sont ceux de Nédélec [56], tandis que celui des pyramides est nouveau. Les différences entre $\hat{P}_r$ et $\hat{P}_r^1$ sont notées en rouge.

Définition 13.1.2 *On notera les espaces $\hat{P}_r^1$ des différents types d'éléments*

- ***Tétraèdres*** *: $\hat{P}_r^1 = \mathcal{R}_r(\hat{x},\hat{y},\hat{z})$*
- ***Hexaèdres*** *: $\hat{P}_r^1 = \mathcal{Q}_r^1(\hat{x},\hat{y},\hat{z})$*
- ***Prisme*** *: $\hat{P}_r^1 = \mathcal{W}_r^1(\hat{x},\hat{y},\hat{z})$*
- ***Pyramide*** *: $\hat{P}_r^1 = \mathcal{B}_r^1(\hat{x},\hat{y},\hat{z})$*

13.1.3.2 Fonctions de base

Comme dans le cas H^1, on propose des fonctions de base hiérarchiques pour les espaces $\hat{P}_r^1$ définis précédemment. Les différences entre les bases proposées pour $\hat{P}_r$ et celles proposées pour $\hat{P}_r^1$ sont là encore notées en rouge.

Proposition 13.1.3 *Les fonctions suivantes forment une base hiérarchique $H(rot)$-conforme de $\hat{P}_r^1$ et sur tout le maillage*

- ***Hexaèdre*** *: On considère les paramètres suivants*

$$\begin{cases} \lambda_1 = \hat{x} \\ \lambda_2 = \hat{y} \\ \lambda_3 = \hat{z} \\ \lambda_4 = 1-\hat{x} \\ \lambda_5 = 1-\hat{y} \\ \lambda_6 = 1-\hat{z} \end{cases}$$

FONCTIONS $H(rot)$ HIÉRARCHIQUES POUR L'HEXAÈDRE

Pour une arête a *: soient* a_1 *et* a_2 *les faces ne contenant aucun sommet de* a $(a_1 < a_2)$

Si a *est orientée selon* e_x

$$\begin{bmatrix} \lambda_{a_1}\lambda_{a_2} \\ 0 \\ 0 \end{bmatrix} P_i^{0,0}(2\hat{x}-1), \quad 0 \leqslant i \leqslant r-1, \quad 1 \leqslant a \leqslant 4$$

Si a *est orientée selon* e_y

$$\begin{bmatrix} 0 \\ \lambda_{a_1}\lambda_{a_2} \\ 0 \end{bmatrix} P_i^{0,0}(2\hat{y}-1), \quad 0 \leqslant i \leqslant r-1, \quad 1 \leqslant a \leqslant 4$$

Si a *est orientée selon* e_z

$$\begin{bmatrix} 0 \\ 0 \\ \lambda_{a_1}\lambda_{a_2} \end{bmatrix} P_i^{0,0}(2\hat{z}-1), \quad 0 \leqslant i \leqslant r-1, \quad 1 \leqslant a \leqslant 4$$

Pour une face f *: soit* f_1 *la face directement opposée à* f

Si f *est dans le plan* (e_x, e_y)

$$\begin{bmatrix} \lambda_2\,\lambda_5\,\lambda_{f_1} \\ 0 \\ 0 \end{bmatrix} P_i^{0,0}(2\hat{x}-1)P_j^{1,1}(2\hat{y}-1)$$

$$\begin{bmatrix} 0 \\ \lambda_1\,\lambda_4\,\lambda_{f_1} \\ 0 \end{bmatrix} P_j^{1,1}(2\hat{x}-1)P_i^{0,0}(2\hat{y}-1)$$

$$\begin{matrix} 0 \leqslant i \leqslant r-1 \\ 0 \leqslant j \leqslant r-2 \end{matrix}, \quad 1 \leqslant f \leqslant 2$$

Si f *est dans le plan* (e_y, e_z)

$$\begin{bmatrix} 0 \\ \lambda_3\,\lambda_6\,\lambda_{f_1} \\ 0 \end{bmatrix} P_i^{0,0}(2\hat{y}-1)P_j^{1,1}(2\hat{z}-1)$$

$$\begin{bmatrix} 0 \\ 0 \\ \lambda_2\,\lambda_5\,\lambda_{f_1} \end{bmatrix} P_j^{1,1}(2\hat{y}-1)P_i^{0,0}(2\hat{z}-1)$$

$$\begin{matrix} 0 \leqslant i \leqslant r-1 \\ 0 \leqslant j \leqslant r-2 \end{matrix}, \quad 1 \leqslant f \leqslant 2$$

Si f *est dans le plan* (e_x, e_z)

$$\begin{bmatrix} \lambda_3\,\lambda_6\,\lambda_{f_1} \\ 0 \\ 0 \end{bmatrix} P_i^{0,0}(2\hat{x}-1)P_j^{1,1}(2\hat{z}-1)$$

$$\begin{bmatrix} 0 \\ 0 \\ \lambda_1\,\lambda_4\,\lambda_{f_1} \end{bmatrix} P_j^{1,1}(2\hat{x}-1)P_i^{0,0}(2\hat{z}-1)$$

$$\begin{matrix} 0 \leqslant i \leqslant r-1 \\ 0 \leqslant j \leqslant r-2 \end{matrix}, \quad 1 \leqslant f \leqslant 2$$

Pour les fonctions intérieures *:*

$$\begin{bmatrix} \lambda_2\,\lambda_3\,\lambda_5\,\lambda_6 \\ 0 \\ 0 \end{bmatrix} P_i^{0,0}(2\hat{x}-1)P_j^{1,1}(2\hat{y}-1)P_k^{1,1}(2\hat{z}-1)$$

$$\begin{bmatrix} 0 \\ \lambda_1\,\lambda_3\,\lambda_4\,\lambda_6 \\ 0 \end{bmatrix} P_k^{1,1}(2\hat{x}-1)P_i^{0,0}(2\hat{y}-1)P_j^{1,1}(2\hat{z}-1) \qquad \begin{matrix} 0 \leqslant i \leqslant r-1 \\ 0 \leqslant j,k \leqslant r-2 \end{matrix}$$

$$\begin{bmatrix} 0 \\ 0 \\ \lambda_1\,\lambda_2\,\lambda_4\,\lambda_5 \end{bmatrix} P_j^{1,1}(2\hat{x}-1)P_k^{1,1}(2\hat{y}-1)P_i^{0,0}(2\hat{z}-1)$$

– ***Prisme :*** *On considère les paramètres suivants*

$$\begin{cases} \lambda_1 = \lambda_4 = 1 - \hat{x} - \hat{y} \\ \lambda_2 = \lambda_5 = \hat{x} \\ \lambda_3 = \lambda_6 = \hat{y} \end{cases} \qquad \begin{cases} \beta_1 = 1 - \hat{z} \\ \beta_2 = \hat{z} \end{cases}$$

FONCTIONS $H(rot)$ HIÉRARCHIQUES POUR LE PRISME

Pour une arête horizontale a *: l'arête est dirigée d'un sommet* a_1 *vers* a_2*, et* f' *est la face horizontale opposée*

$$(\lambda_{a_1}\nabla\lambda_{a_2} - \lambda_{a_2}\nabla\lambda_{a_1})\ \beta_{f'}\ P_i^{0,0}(\lambda_{a_2} - \lambda_{a_1}), \quad 0 \leqslant i \leqslant r-1, \quad 1 \leqslant a \leqslant 6$$

Pour une arête verticale a *: soit* a_1 *la face ne contenant aucun sommet de* a

$$\begin{bmatrix} 0 \\ 0 \\ \lambda_{a_1} \end{bmatrix} P_i^{0,0}(\beta_2 - \beta_1), \quad 0 \leqslant i \leqslant r-1, \quad 1 \leqslant a \leqslant 3$$

Pour une face quadrangulaire f *: soit* $[a_1, a_2]$ *une arête en commun avec une face triangulaire* f'*, et* f_1 *et* f_2 *les deux autres faces quadrangulaires*

$$(\lambda_{a_1}\nabla\lambda_{a_2} - \lambda_{a_2}\nabla\lambda_{a_1})\ \beta_1\,\beta_2\,P_i^{0,0}(\lambda_{a_2} - \lambda_{a_1})P_j^{1,1}(2\hat{z}-1)$$

$$\begin{bmatrix} 0 \\ 0 \\ \lambda_{a_1}\,\lambda_{a_2} \end{bmatrix} P_j^{1,1}(\lambda_{a_2} - \lambda_{a_1})P_i^{0,0}(2\hat{z}-1) \qquad \begin{matrix} 0 \leqslant j \leqslant r-2 \\ 0 \leqslant i \leqslant r-1 \end{matrix} \qquad 1 \leqslant f \leqslant 3$$

Pour une face triangulaire f *: soient* $[a_1, a_2]$ *et* $[a_1, a_3]$ *deux arêtes en commun avec deux faces faces quadrangulaires* f_1 *et* f_2 *respectivement, et* f' *la face horizontale opposée*

$$\begin{matrix} (\lambda_{a_1}\nabla\lambda_{a_2} - \lambda_{a_2}\nabla\lambda_{a_1})\ \lambda_{f_1}\,\beta_{f'}\,P_i^{0,0}(\lambda_{a_2} - \lambda_{a_1})P_j^{0,0}(\lambda_{a_3} - \lambda_{a_1}) \\ (\lambda_{a_1}\nabla\lambda_{a_3} - \lambda_{a_3}\nabla\lambda_{a_1})\ \lambda_{f_2}\,\beta_{f'}\,P_i^{0,0}(\lambda_{a_2} - \lambda_{a_1})P_j^{0,0}(\lambda_{a_3} - \lambda_{a_1}) \end{matrix} \qquad 0 \leqslant i+j \leqslant r-2, \quad 1 \leqslant f \leqslant 2$$

Pour les fonctions intérieures *:*

$$\begin{matrix} (\lambda_2\nabla\lambda_3 - \lambda_3\nabla\lambda_2)\ \lambda_1\,\beta_1\,\beta_2\,P_{ijk}(\hat{x},\hat{y},\hat{z}) \\ (\lambda_1\nabla\lambda_3 - \lambda_3\nabla\lambda_1)\ \lambda_2\,\beta_1\,\beta_2\,P_{ijk}(\hat{x},\hat{y},\hat{z}) \end{matrix} \qquad \begin{matrix} 0 \leqslant i+j \leqslant r-2 \\ 0 \leqslant k \leqslant r-2 \end{matrix}$$

$$\begin{bmatrix} 0 \\ 0 \\ \lambda_1\,\lambda_2\,\lambda_3 \end{bmatrix} P_{ijk}(\hat{x},\hat{y},\hat{z}) \qquad \begin{matrix} 0 \leqslant i+j \leqslant r-3 \\ 0 \leqslant k \leqslant r-1 \end{matrix}$$

avec

$$P_{ijk}(\hat{x},\hat{y},\hat{z}) = P_i^{0,0}\left(\frac{2\hat{x}}{1-\hat{y}} - 1\right)(1-\hat{y})^i P_j^{2i+1,0}(2\hat{y}-1)P_k^{0,0}(2\hat{z}-1)$$

– ***Pyramide :*** *On considère les paramètres suivants*

$$\begin{cases} \beta_1 = \dfrac{1-\hat{x}-\hat{z}}{2} \\ \beta_2 = \dfrac{1-\hat{y}-\hat{z}}{2} \\ \beta_3 = \dfrac{1+\hat{x}-\hat{z}}{2} \\ \beta_4 = \dfrac{1+\hat{y}-\hat{z}}{2} \end{cases} \qquad \begin{cases} \lambda_1 = \dfrac{\beta_1\,\beta_2}{1-\hat{z}} \\ \lambda_2 = \dfrac{\beta_2\,\beta_3}{1-\hat{z}} \\ \lambda_3 = \dfrac{\beta_3\,\beta_4}{1-\hat{z}} \\ \lambda_4 = \dfrac{\beta_4\,\beta_1}{1-\hat{z}} \\ \lambda_5 = \hat{z} \end{cases} \qquad \begin{cases} \gamma_1 = \dfrac{2\hat{z}+\hat{x}+\hat{y}}{2} \\ \gamma_2 = \dfrac{2\hat{z}-\hat{x}+\hat{y}}{2} \\ \gamma_3 = \dfrac{2\hat{z}-\hat{x}-\hat{y}}{2} \\ \gamma_4 = \dfrac{2\hat{z}+\hat{x}-\hat{y}}{2} \end{cases} \qquad \begin{cases} \delta_1 = \delta_3 = \hat{x} \\ \delta_2 = \delta_4 = \hat{y} \end{cases}$$

FONCTIONS $H(rot)$ HIÉRARCHIQUES POUR LA PYRAMIDE

Pour une arête horizontale a *: l'arête est dirigée d'un sommet* a_1 *vers* a_2*, et les arêtes horizontales adjacentes sont* $[a_1,a_4]$ *et* $[a_2,a_3]$

$$(\lambda_{a_1}\nabla(\lambda_{a_2}+\lambda_{a_3}) - \lambda_{a_2}\nabla(\lambda_{a_1}+\lambda_{a_4}))\, P_i^{0,0}(\delta_a), \quad 0 \leqslant i \leqslant r-1, \quad 1 \leqslant a \leqslant 4$$

Pour une arête verticale a *: soit* s *le sommet de* a *appartenant à la base*

$$(\lambda_s\nabla\lambda_5 - \lambda_5\nabla\lambda_s)\, P_i^{0,0}(\gamma_s), \quad 0 \leqslant i \leqslant r-1, \quad 1 \leqslant a \leqslant 4$$

Pour la base *:*

$$\begin{aligned} &(\lambda_1\nabla(\lambda_2+\lambda_3) - \lambda_2\nabla(\lambda_1+\lambda_4))\ \beta_4\, P_i^{0,0}\Big(\frac{\beta_3-\beta_1}{1-\hat{z}}\Big)P_j^{1,1}\Big(\frac{\beta_4-\beta_2}{1-\hat{z}}\Big)(1-\hat{z})^{\max(i,j)-1} \\ &(\lambda_1\nabla(\lambda_3+\lambda_4) - \lambda_4\nabla(\lambda_2+\lambda_1))\ \beta_3\, P_j^{1,1}\Big(\frac{\beta_3-\beta_1}{1-\hat{z}}\Big)P_i^{0,0}\Big(\frac{\beta_4-\beta_2}{1-\hat{z}}\Big)(1-\hat{z})^{\max(i,j)-1} \end{aligned} \qquad \begin{matrix} 0 \leqslant i \leqslant r-1 \\ 0 \leqslant j \leqslant r-2 \end{matrix}$$

Pour une face triangulaire f *: soit* $[a_1,a_2]$ *l'arête verticale d'arêtes adjacentes* $[a_1,a_4]$ *et* $[a_2,a_3]$*, et* f_1 *la face triangulaire de base* $[a_1,a_4]$

$$\begin{aligned} &(\lambda_{a_2}\nabla(\lambda_{a_1}+\lambda_{a_4}) - \lambda_{a_1}\nabla(\lambda_{a_2}+\lambda_{a_3}))\ \lambda_5\, P_i^{0,0}(\delta_f)P_j^{0,0}(\gamma_{a_1}) \\ &(\lambda_{a_1}\nabla\lambda_5 - \lambda_5\nabla\lambda_{a_1})\ \beta_{f_1}\, P_i^{0,0}(\delta_f)P_j^{0,0}(\gamma_{a_1}) \end{aligned} \qquad 0 \leqslant i+j \leqslant r-2$$

Pour les fonctions intérieures *:*

$$\begin{aligned} &(\lambda_1\nabla(\lambda_2+\lambda_3) - \lambda_2\nabla(\lambda_1+\lambda_4))\ \beta_4\,\lambda_5\, P_{ijk}(\hat{x},\hat{y},\hat{z}) \\ &(\lambda_1\nabla(\lambda_3+\lambda_4) - \lambda_4\nabla(\lambda_2+\lambda_1))\ \beta_3\,\lambda_5\, P_{ijk}(\hat{x},\hat{y},\hat{z}) \\ &(\lambda_1\nabla\lambda_5 - \lambda_5\nabla\lambda_1)\ \beta_3\,\beta_4\, P_{ijk}(\hat{x},\hat{y},\hat{z}) \end{aligned} \qquad \begin{matrix} 0 \leqslant i,j \leqslant r-2, \\ 0 \leqslant k \leqslant r-2-max(i,j) \end{matrix}$$

avec

$$P_{ijk}(\hat{x},\hat{y},\hat{z}) = P_i^{0,0}\Big(\frac{\beta_3-\beta_1}{1-\hat{z}}\Big)P_j^{0,0}\Big(\frac{\beta_4-\beta_2}{1-\hat{z}}\Big)P_k^{2\max(i,j)+2,0}(2\hat{z}-1)(1-\hat{z})^{\max(i,j)-1}$$

– ***Tétraèdre :*** *Voir proposition 10.3.4*

Preuve. Comme les fonctions de base générant $\hat{P}_r$, les fonctions hiérarchiques ci-dessus sont construites de sorte à assurer la continuité des composantes tangentielles.

En considérant les différences avec $\hat{P}_r$, le fait que l'ensemble des fonctions ci-dessus forme une base de $\hat{P}_r^1$ pour chaque type d'élément apparait aisément, aussi la démonstration ne sera-t-elle pas détaillée.

13.1.3.3 Propriétés

Propriété 13.1.4 *Pour tous les types d'éléments, les espaces* $\hat{P}_r^1$ *vérifient les inclusions suivantes*

$$\hat{P}_{r-1} \subset \hat{P}_r^1 \subset \hat{P}_r$$

Preuve. Le résultat est immédiat en considérant les différences entre $\hat{P}_r$ et $\hat{P}_r^1$. □

On vérifie en outre que, pour tous les types d'élément, la dimension de l'espace $\hat{P}_1^1$ est égale au nombre d'arêtes de l'élément. Historiquement, il s'agit en effet de la façon dont ont été construits les éléments finis d'arête, et de la raison pour laquelle ils portent ce nom.

Concernant les tétraèdres, les prismes et les hexaèdres, le fait est acquis. Pour les pyramides, la dimension de l'espace $\mathcal{B}_r^1$ est

$$\dim \mathcal{B}_r^1 = \frac{r(2r^2+9r+5)}{2}$$

ce qui donne 8 degrés de liberté à l'ordre 1, qui correspondent bien aux circulations sur les 8 arêtes de la pyramide.

13.2 Comparaison d'éléments pyramidaux

13.2.1 Comparaison théorique

On effectue ici la comparaison de notre espace $\hat{P}_r^1$ avec les espaces trouvés dans la littérature.

- À l'ordre 1, on retrouve les espaces de Coulomb, Zgainski et Maréchal [22], Doucet *et al.* [25] et Graglia et Gheorma [38].
- On retrouve l'espace d'ordre 1 de Gradinaru et Hiptmair [36] en faisant les modifications suivante dans les fonctions de base γ_6 et γ_7 proposées

$$\gamma_6 = \begin{bmatrix} -z + \dfrac{yz}{1-z} \\ \dfrac{xz}{1-z} \\ x - \dfrac{xy}{1-z} + \dfrac{xyz}{(1-z)^2} \end{bmatrix}, \quad \gamma_7 = \begin{bmatrix} \dfrac{yz}{1-z} \\ -z + \dfrac{xz}{1-z} \\ y - \dfrac{xy}{1-z} + \dfrac{xyz}{(1-z)^2} \end{bmatrix}$$

- On fait les modifications suivantes sur les fonctions de base liées aux arêtes verticales proposées par Nigam et Phillips [58]

$$\tilde{F}_{e_1} = \left\{ \frac{1}{(1+z)^{k+1-\gamma}} \begin{bmatrix} -z(y-1) \\ -z(x-1) \\ (x-1)(y-1) \end{bmatrix}, 0 \leqslant \gamma \leqslant k-1 \right\}$$

Ainsi, la composante tangentielle de ces fonctions devient polynomiale sur les deux faces adjacentes à l'arête.

En utilisant les transformations $\bar{T}$ et T afin de passer de la pyramide infinie à la pyramide de référence, on remarque que l'espace engendré par les fonctions de base liées aux arêtes et aux faces ainsi obtenues est égal à celui généré par les fonctions de base liées aux arêtes et aux faces de la proposition 13.1.3.

L'espace généré par toutes les fonctions contient $\hat{P}_r^1$, mais compte $3r(r-1)^2$ fonctions de base à l'intérieur, soit $\dfrac{r(r-1)(4r-5)}{2}$ de plus qu'à l'intérieur de $\hat{P}_r^1$ pour obtenir le même ordre de convergence. En mettant les fonctions de base intérieures de la proposition 13.1.3, après transformation grâce à $\bar{T}$ et T, on retrouve ainsi le même espace $\hat{P}_r^1$.

- Le second espace proposé par Nigam et Phillips [59] est de dimension $\dfrac{r(2r^2+7r+7)}{2}$, soit $r(r-1)$ de moins que la dimension de $\hat{P}_r^1$. Cependant, il n'y a pas de relation d'inclusion avec $\hat{P}_r^1$ puisque $\hat{P}_1$ n'est pas engendré par l'espace proposé. En effet, les fonctions suivantes de $\hat{P}_1$ ne sont pas dans l'espace proposé par les auteurs

$$\begin{bmatrix} \dfrac{\hat{y}^2}{(1-\hat{z})} \\ 0 \\ \dfrac{\hat{x}\,\hat{y}^2}{(1-\hat{z})^2} \end{bmatrix} \quad \begin{bmatrix} 0 \\ \dfrac{\hat{x}^2}{(1-\hat{z})} \\ \dfrac{\hat{x}^2\,\hat{y}}{(1-\hat{z})^2} \end{bmatrix}$$

En revanche, l'espace contient $\mathcal{R}_r$, la convergence est donc optimale pour les pyramides affines.

- Zaglmayr citée par Demkowicz [23] proposent un espace présenté comme étant un espace d'approximation pour la seconde famille, avec des restrictions pour éliminer certains termes. Cependant, une expression pratique de l'espace est difficile à obtenir, aussi n'a-t-il pas été étudié plus précisément, ni codé.

TAB. 13.1 – Inclusion des espaces pour différents éléments de la littérature

Espace V_r	Inclusion $\hat{P}_r^1$	Inclusion $\hat{P}_r$
Coulomb, Zgainski et Maréchal [22] $r = 2$	$\hat{P}_1^1 \subset V_2$	$\hat{P}_1 \subset V_2 \subset \hat{P}_3$
Graglia et Gheorma [38] $r = 2$	$\hat{P}_1^1 \subset V_2$	$\hat{P}_1 \not\subset V_2$
Nigam et Phillips [58]	$\hat{P}_r^1 \subset V_r$	$\hat{P}_{r-1} \subset V_r$
Nigam et Phillips [59]	$\hat{P}_1^1 \subset V_r$	$\hat{P}_1 \not\subset V_r$

13.2.2 Comparaison numérique

On vérifie tout d'abord numériquement les différents résultats d'inclusions de la section 13.1.2.4. Les résultats sont résumés dans le tableau 13.1.

On compare à présent l'ordre de dispersion obtenu pour différents espaces proposés dans la littérature avec l'espace optimal. La figure 13.2.2 indique ainsi l'erreur de dispersion obtenue pour un maillage constitué uniquement de pyramides affines ou d'un mélange de pyramides dont la moitié est non-affine pour différents espaces. On constate que l'on retrouve les résultats d'inclusion dans l'ordre de dispersion, pour les pyramides affines comme pour dans le cas non-affine, sauf pour l'espace de Graglia et Gheorma [38] d'ordre 2. En effet, dans le cas non-affine, on observe une dispersion d'ordre 2 alors que l'on s'attendait à avoir une dispersion d'ordre 0, l'espace proposé ne contenant pas $\hat{P}_1$. Cependant, si l'ordre de dispersion donne une bonne idée de ce que sera l'ordre de convergence, il peut arrive que l'ordre de dispersion soit plus élevé que l'ordre de convergence.

On souhaite obtenir confirmation des résultats d'ordre de dispersion en vérifiant la convergence en norme $H(rot)$ sur un cas simple. On considère une cavité cubique $[-1,1]^3$ maillée avec des pyramides, la moitié d'entre elles étant non-affine et on place une source gaussienne au centre de la cavité. La figure 13.2 donne les résultats obtenus pour les éléments d'ordre 1, 2 et 3. On retrouve globalement les ordres de convergence obtenus pour la dispersion, bien que les résultats soient moins lisibles. En revanche, les éléments de Graglia et Gheorma [38] et Coulomb, Zgainski et Maréchal [22] sont peu robustes, ce qui perturbe les résultats de convergence. La présence de parasites avait déjà été signalé par Marais et Davidson [53].

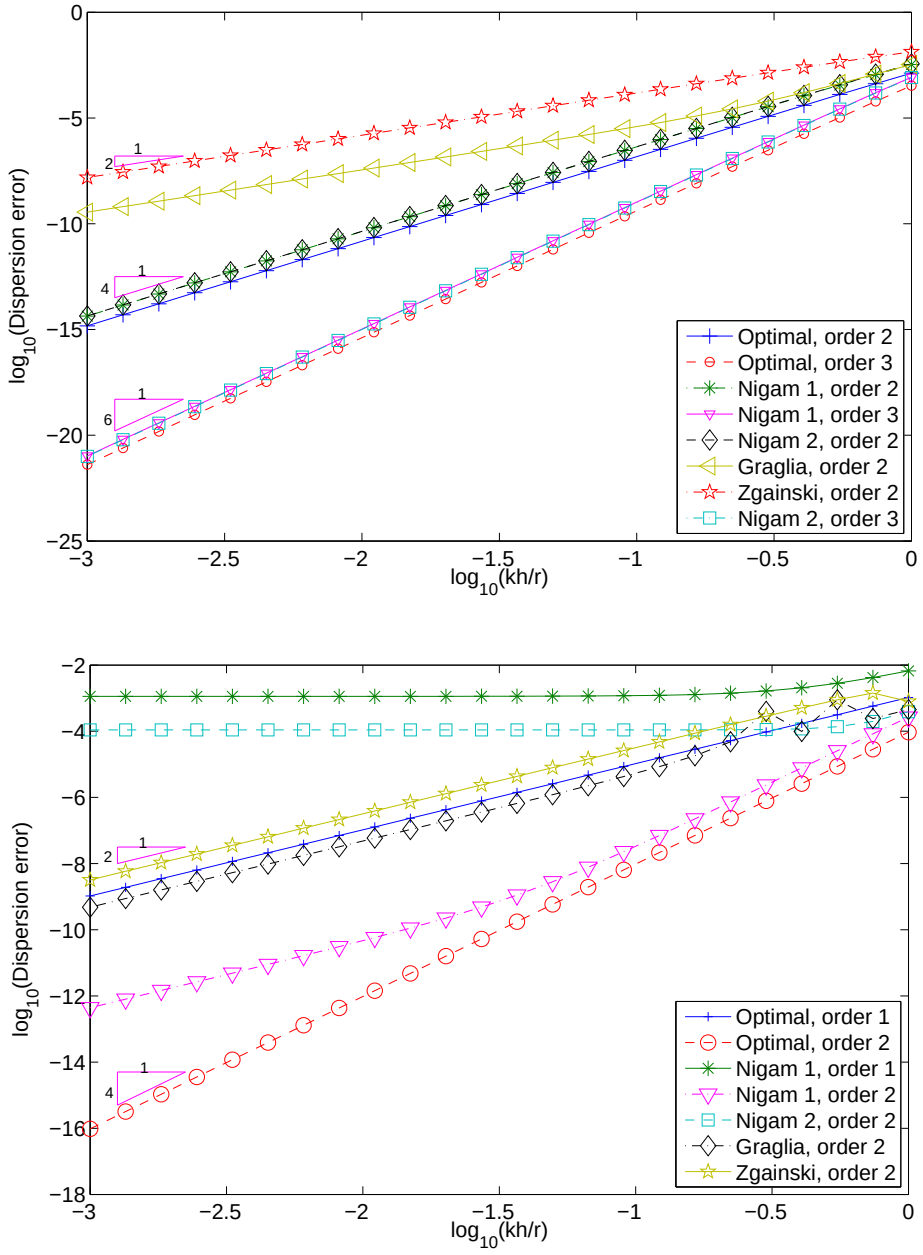

FIG. 13.1 – Erreur de dispersion en échelle logarithmique pour un maillage ne contenant que des pyramides affines (en haut) et un mélange de pyramides affines et non-affines (en bas) pour différents éléments finis d'arête pyramidaux.

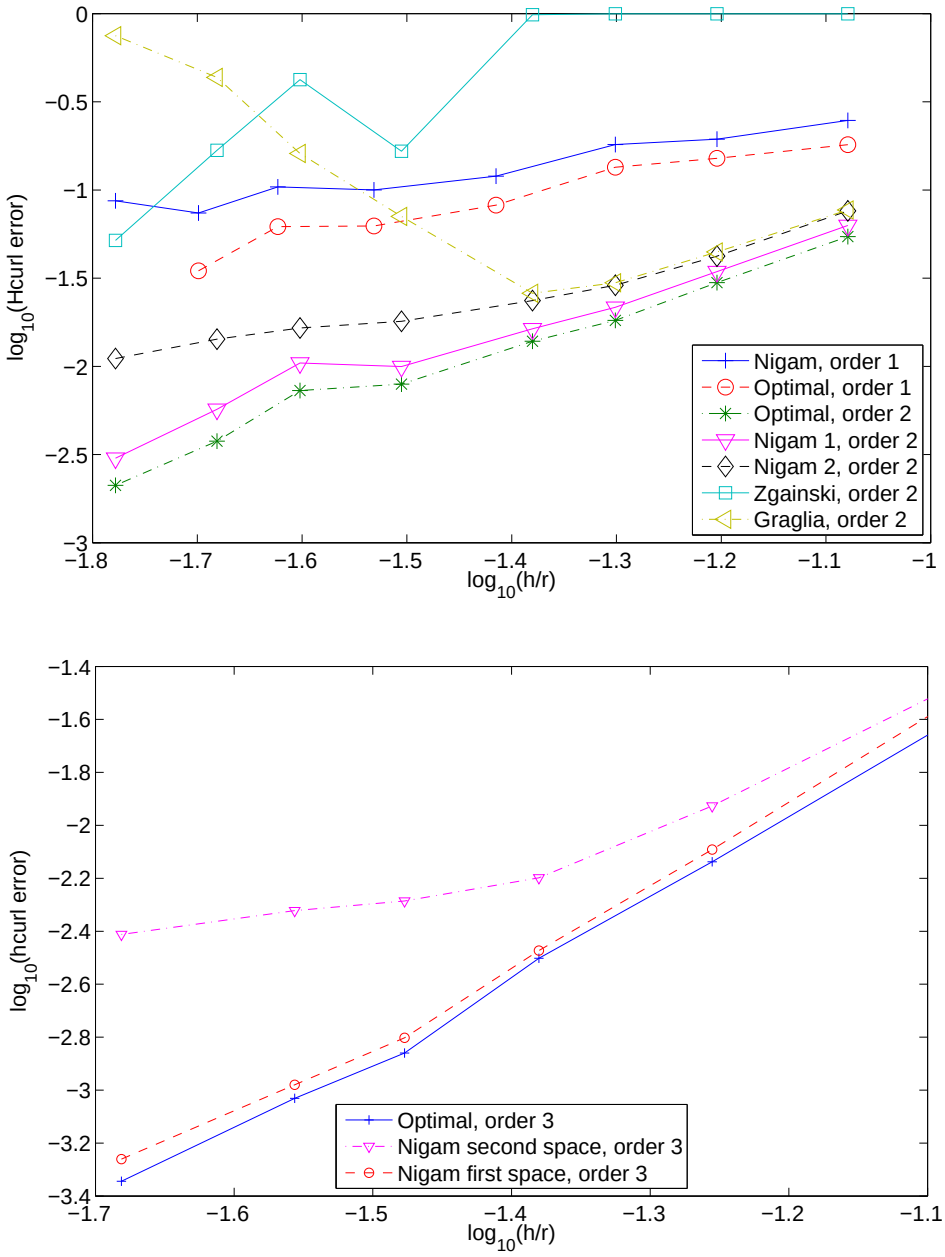

FIG. 13.2 – Erreur en norme H(rot) en échelle logarithmique pour un maillage ne contenant que des pyramides, certaines étant non-affines. Comparaison de divers éléments finis d'arête pyramidaux d'ordre un et deux (en haut) et d'ordre 3 (en bas).

Afin de mieux illustrer ce phénomène de parasites, on effectue la même étude que dans la section 12.3, c'est à dire que l'on calcule les valeurs propres dans une cavité cubique pour laquelle on connait les valeurs propres théoriques. La cavité est maillée avec des pyramides. D'après la figure 13.3, on obtient effectivement un grand nombre de parasites pour les espaces de Graglia et Gheorma [38] et Coulomb, Zgainski et Maréchal [22]. Un exemple de mode physique et de mode parasite calculé avec les éléments de Graglia et Gheorma est illustré sur la figure 13.4.

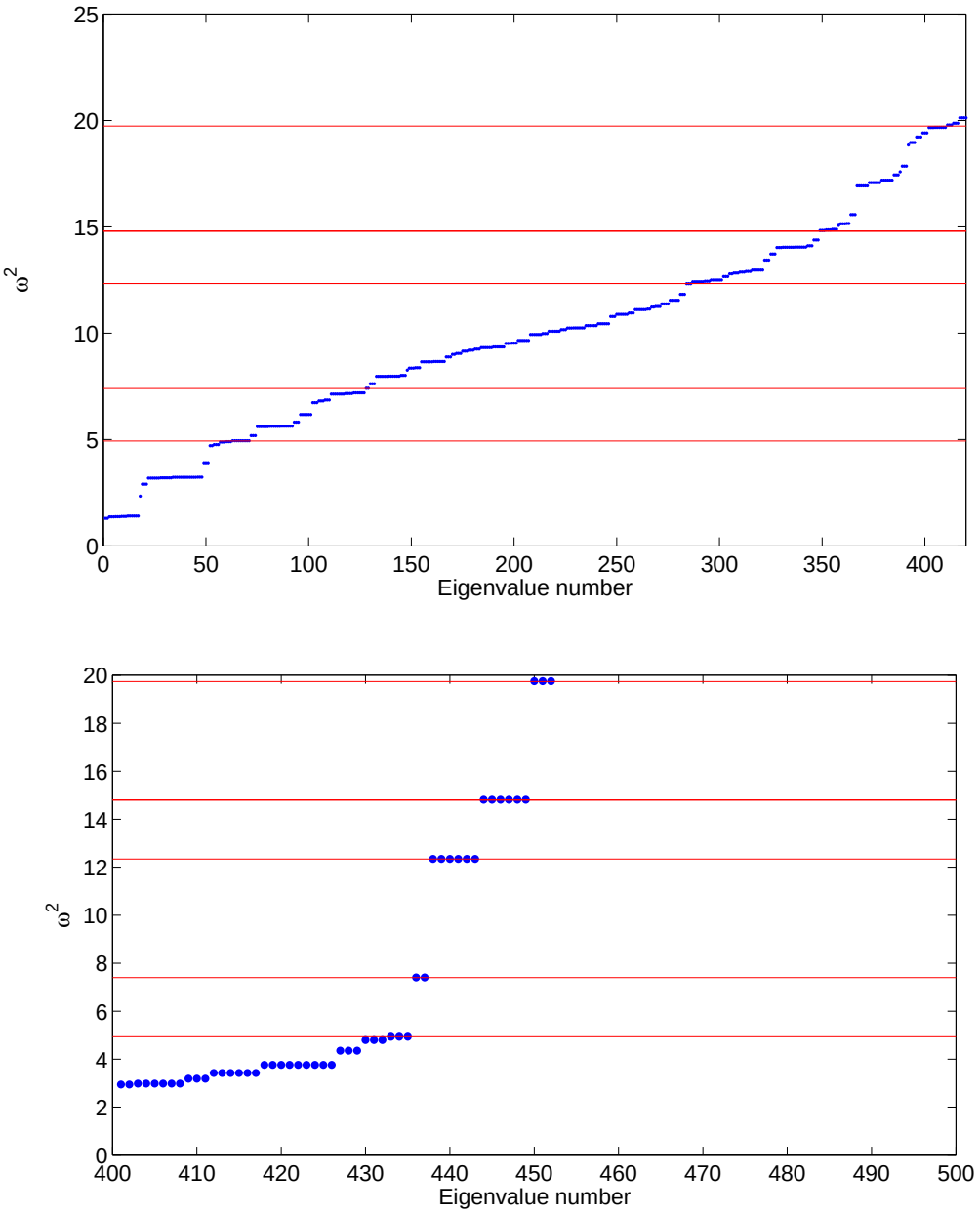

FIG. 13.3 – Distribution des valeurs propres pour un maillage pyramidal contenant 15 000 ddls avec les espaces de Coulomb, Zgainski et Maréchal (en haut) et Graglia et Gheorma (en bas). Les valeurs propres analytiques sont représentés par des lignes rouges, les valeurs propres numériques par des points bleus.

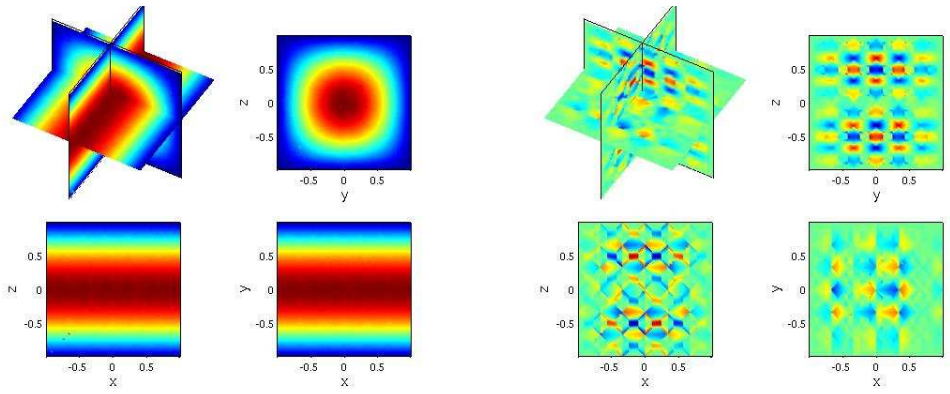

FIG. 13.4 – Exemple de deux modes propres obtenus sur une cavité cubique maillée avec des pyramides non-affines avec les éléments finis d'arête de Graglia et Gheorma d'ordre 2.

Pour terminer, on récapitule les différentes propriétés des éléments finis d'arête pyramidaux dans le tableau 13.2.

TAB. 13.2 – Propriété des différents espaces

	Zgainski $r=2$	Graglia $r=2$	Nigam&Phillips 1	Nigam&Phillips 2	Optimal
Convergence en affine	$O(h)$	$O(h)$	$O(h^r)$	$O(h^r)$	$O(h^r)$
Convergence en non-affine	$O(h)$	$O(1)$	$O(h^{r-1})$	$O(1)$	$O(h^r)$
Parasites	oui	oui	non	non	non
Compatibilité	non	oui	oui	oui	oui

Remarque 13.2.1 *Il est à noter que les éléments de Graglia d'ordre 2 sont bien compatibles avec les tétraèdres et hexaèdres de la première famille, alors que les éléments de Zgainski d'ordre 2 ne vérifient pas cette compatibilité. En effet, la restriction des fonctions intérieures à la base quadrilatère est*

$$Span\left\{\begin{vmatrix}0\\x(1-x)(1-y)^2\end{vmatrix},\begin{vmatrix}0\\x(1-x)y^2\end{vmatrix},\begin{vmatrix}(1-y)y(1-x)^2\\0\end{vmatrix},\begin{vmatrix}(1-y)y(1-x)^2\\0\end{vmatrix}\right\}$$

et la restriction sur les faces triangulaires

$$Span\left\{\begin{vmatrix}(1-x-y)y\\0\end{vmatrix},\begin{vmatrix}xy\\xy\end{vmatrix}\right\}$$

Or ces fonctions n'appartiennent pas à la première famille des quadrangles et des triangles.

13.3 Diagramme de De Rham

La plupart des auteurs s'intéressant aux éléments finis d'arête tentent de les construire sur différents types d'éléments et pour les formulations H^1, $H(rot)$, $H(div)$ et L^2 de sorte à pouvoir les incorporer dans un formalisme global utilisant des formes différentielles discrètes introduites par Whitney [73].

Le but est en fait d'obtenir une suite d'espaces d'approximation pour chacune des formulation vérifiant le diagramme suivant

$$\begin{array}{ccccccc} H^1 & \xrightarrow{grad} & H(rot) & \xrightarrow{rot} & H(div) & \xrightarrow{div} & L^2 \\ \cup & & \cup & & \cup & & \cup \\ W^1_{r+1} & \xrightarrow{grad} & W^{rot}_r & \xrightarrow{rot} & W^{div}_{r-1} & \xrightarrow{div} & W^2_{r-2} \end{array} \tag{13.3.1}$$

Ce diagramme, directement lié à la décomposition de Helmholtz, est appelé **diagramme de De Rham** (voir Monk [55]).

Les auteurs s'attachent en particulier à vérifier les inclusions suivantes

$$\begin{aligned} grad\, W^1_{r+1} &\subset W^{rot}_r \\ rot\, W^{rot}_r &\subset W^{div}_{r-1} \\ div\, W^{div}_{r-1} &\subset W^2_{r-2} \end{aligned} \tag{13.3.2}$$

Plus particulièrement, lorsque les espaces d'approximations vérifient

$$\begin{aligned} \text{Im}\, grad\, W^1_{r+1} &= \text{Ker}\, W^{rot}_r = \left\{u \in W^{rot}_r \,|\, rot\, u = 0\right\} \\ \text{Im}\, rot\, W^{rot}_r &= \text{Ker}\, W^{div}_{r-1} = \left\{u \in W^{div}_{r-1} \,|\, div\, u = 0\right\} \\ \text{Im}\, div\, W^{div}_{r-1} &= \text{Ker}\, W^2_{r-2} = W^2_{r-2} \end{aligned} \tag{13.3.3}$$

on dit que la séquence est **exacte** (voir Demkowicz [23]).

Dular *et al.* [26] ont proposé un formalisme pour construire les tétraèdres, les hexaèdres et les prismes de façon à respecter le diagramme de De Rham. Pour les pyramides, Nigam et Phillips [58] et [59] ainsi que Zaglamayr citée dans [23] construisent leurs espaces d'approximation d'ordre r de telle sorte que tous les espaces construits vérifient la séquence du diagramme.

Bien que nous ne nous soyons pas préoccupés du respect du diagramme lors de la construction de nos éléments, vérifions que nous avons la propriété d'inclusion 13.3.2 pour les espaces d'approximation H^1 et $H(rot)$.

Proposition 13.3.1 *Pour tout ordre r, on a les inclusions suivantes*

$$\begin{aligned} grad\, \hat{P}^{H^1}_r &\subset \hat{P}^{H(rot)}_r \\ grad\, \hat{P}^{H^1}_r &\subset \hat{P}^1_r \end{aligned}$$

Preuve. En remarquant que

$$\begin{aligned} grad\, \mathbb{P}_r(\hat{x},\hat{y},\hat{z}) &= \left(\mathbb{P}_{r-1}(\hat{x},\hat{y},\hat{z})\right)^3 \\ grad\, \mathbb{Q}_{m,n,p}(\hat{x},\hat{y},\hat{z}) &= \mathbb{Q}_{m-1,n,p}(\hat{x},\hat{y},\hat{z}) \times \mathbb{Q}_{m,n-1,p}(\hat{x},\hat{y},\hat{z}) \times \mathbb{Q}_{m,n,p-1}(\hat{x},\hat{y},\hat{z}) \end{aligned}$$

la preuve est immédiate pour les hexaèdres, les prismes et les tétraèdres.

Concernant les pyramides, séparons $\mathbb{B}_r(\hat{x},\hat{y},\hat{z})$ selon sa partie polynomiale et sa partie rationnelle

- Concernant la partie polynomiale, on a de manière immédiate $grad\, \mathbb{P}_r(\hat{x},\hat{y},\hat{z}) \in \hat{P}_r$.
- Concernant la partie rationnelle, on considère $\hat{p} = \dfrac{\hat{x}^i \hat{y}^j}{(1-\hat{z})^{i+j-k}}$ avec $0 \leqslant i+j \leqslant k \leqslant r-1$. On a

$$grad\, \hat{p} = \begin{bmatrix} (r+i-k)\dfrac{\hat{x}^{r+i-k-1}\hat{y}^{r+j-k}}{(1-z)^{r-k}} \\ (r+j-k)\dfrac{\hat{x}^{r+i-k}\hat{y}^{r+j-k-1}}{(1-z)^{r-k}} \\ (r-k)\dfrac{\hat{x}^{r+i-k}\hat{y}^{r+j-k}}{(1-z)^{r-k+1}} \end{bmatrix} \quad 0 \leqslant i+j \leqslant k \leqslant r-1$$

Pour $1 \leqslant i+j \leqslant k \leqslant r-1$, on a $grad\, \hat{p} \in \left(\mathbb{B}_{r-1}\right)^3$. Reste à traiter le cas de $i=j=0$ pour $0 \leqslant k \leqslant r-1$. On obtient ainsi

$$grad\, \hat{p} = (r-k)\begin{bmatrix} \dfrac{\hat{x}^{r-k-1}\hat{y}^{r-k}}{(1-z)^{r-k}} \\ \dfrac{\hat{x}^{r-k}\hat{y}^{r-k-1}}{(1-z)^{r-k}} \\ \dfrac{\hat{x}^{r-k}\hat{y}^{r-k}}{(1-z)^{r-k+1}} \end{bmatrix} \in \left\{ \begin{bmatrix} \dfrac{\hat{x}^p\,\hat{y}^{p+1}}{(1-\hat{z})^{p+1}} \\ \dfrac{\hat{x}^{p+1}\,\hat{y}^p}{(1-\hat{z})^{p+1}} \\ \dfrac{\hat{x}^{p+1}\,\hat{y}^{p+1}}{(1-\hat{z})^{p+2}} \end{bmatrix}, \quad 0 \leqslant p \leqslant r-1 \right\}$$

ce qui achève la démonstration pour les deux espaces. □

On montre numériquement que l'on a même la séquence exacte

$$\text{Im } grad\, \hat{P}_r^{H^1} = \text{Ker } \hat{P}_r^{H(rot)}$$

$$\text{Im } grad\, \hat{P}_r^{H^1} = \text{Ker } \hat{P}_r^{1}$$

en comparant la dimension de l'espace $grad\, \hat{P}_r^{H^1}$ et celle du noyau de la matrice de rigidité. Compte tenu des inclusions de la proposition 13.3.1, les deux dimensions étant égales, on a égalité des espaces.

Cinquième partie

Étude numérique

Chapitre 14

Expériences numériques en régime harmonique

On réalise à présent des expériences numériques sur des cas réels avec les éléments étudiés précédemment. On considère ici différents types d'équations issus des problèmes de propagation d'onde en régime harmonique, ce chapitre concerne donc les éléments finis pour formulation continue.

Sommaire

14.1 Sphère avec éléments isoparamétriques

On souhaite tester les éléments isoparamétriques sur un cas simple. On considère la diffraction par une sphère Γ de rayon $r = 3$, placée dans le cube $[-5,5]^3$ de frontière Σ

$$\begin{cases} -\omega^2 u - \Delta u = 0 & \text{dans } \Omega \\ \dfrac{\partial u}{\partial n} = -\dfrac{\partial u^{incident}}{\partial n} & \text{sur } \Gamma \\ \dfrac{\partial u}{\partial n} - i\,\omega\, u = 0 & \text{sur } \Sigma, \end{cases}$$

On prend $\omega = 2\pi$.

Le maillage de l'ensemble est présenté sur la figure 14.2) pour les différents types de maillage utilisés. Pour obtenir une bonne approximation de la géométrie, on considère des éléments courbes isoparamétriques dont a construction est expliquée dans la section 2.1.2. La solution de référence, présentée sur la figure 14.2, est calculée sur un maillage hexaédrique très fin avec des éléments d'ordre 7.

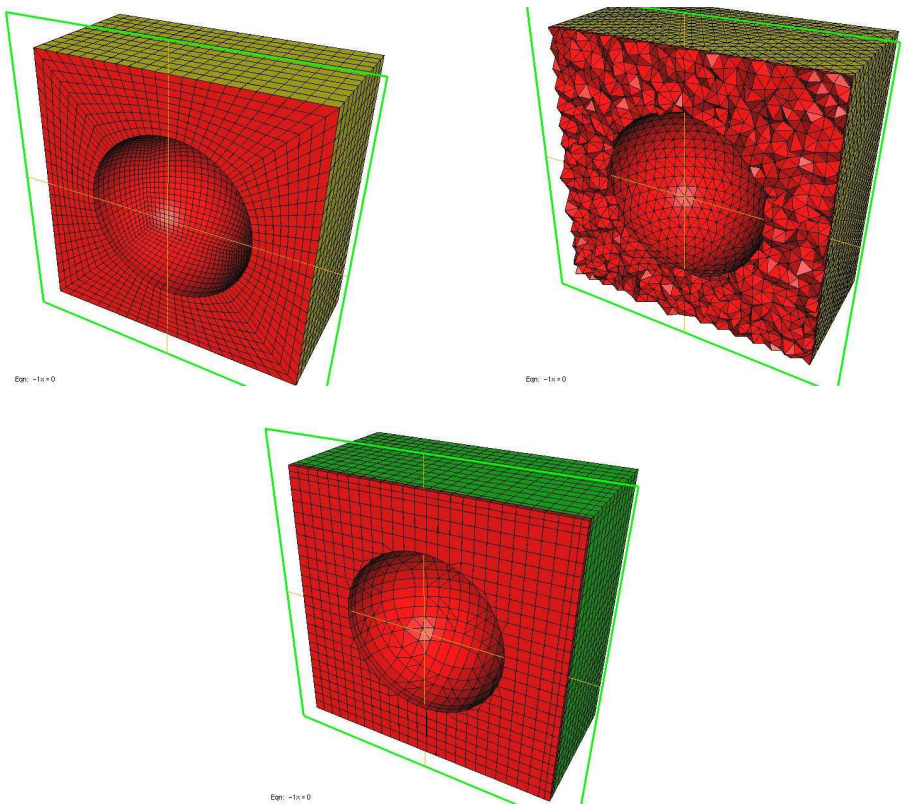

FIG. 14.1 – Maillages utilisés pour l'approximation d'ordre 3.

Pour résoudre le système linéaire, on utilise le solveur COCG de Clemens et Weiland [15]) auquel on peut adjoindre une étape de préconditionnement par une itération p-multigrille en utilisant l'équation de Helmholtz avec terme d'amortissement (voir Erlangga [31] pour les différences finies, et Duruflé [28] pour les éléments finis). On utilise un algorithme de Jacobi comme lisseur, mais on peut également utiliser l'algorithme de Gauss-Seidel.

Dans le tableau 14.1, on indique le nombre de degrés de liberté nécessaires à l'obtention d'une erreur entre 1% et 2% en norme L^2 pour chaque type de maillage, aux ordres 2, 3 et 4. On donne également les résultats obtenus avec ou sans étape de préconditionnement, ainsi que les temps de calcul. On prend un maillage grossier pour l'ordre $5 - \mathbb{P}_4$, et des maillages un peu plus fins pour les ordres 2 et 3.

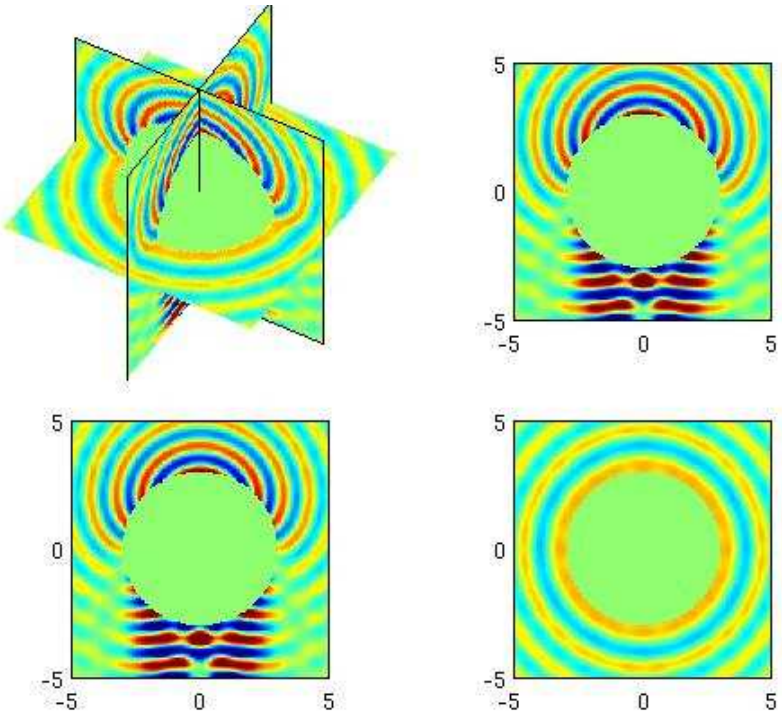

FIG. 14.2 – Partie réelle du champ diffracté par une sphère de rayon $r = 3$ avec des conditions de Neumann.

Ordre	2	3	5 ($\mathbb{P}_4$ pour les tétraèdres)
Hexaèdre	964 000 ddl	732 000 ddl	315 000 ddl
sans precond.	2 762 itérations (3 410s)	2 938 itérations (2 024s)	3 467 itérations (**802**s)
avec precond. J	133 itérations (637s)	127 itérations (**504**s)	130 itérations (**152s**)
Tetraèdre	1 216 000 ddl	519 000 ddl	339 000 ddl
sans precond.	2 300 itérations (12 622s)	1 656 itérations (3 490s)	1 942 itérations (17 835s)
avec precond. J	58 itérations (1 019s)	51 itérations (534s)	119 itérations (587s)
Tetraèdre découpé	2 751 000 ddl	936 000 ddl	520 000 ddl
sans precond.	4 837 itérations (19 833s)	3 775 itérations (3 775s)	2 514 itérations (2 514s)
avec precond. J	131 itérations (1 809s)	126 itérations (631s)	93 itérations (266s)
Hybride	1 060 000 ddl	455 000 ddl	266 000 ddl
sans precond.	1 800 itérations (**2 744**s)	2 195 itérations (**1 153**s)	4 222 itérations (1 358s)
avec precond.	72 itérations (**388**s)	439 itérations (1 262s)	2 546 itérations (3 685s)
avec precond. GS	69 itérations (**330s**)	76 itérations (**176s**)	128 itérations (161s)

TAB. 14.1 – Nombre de degrés de liberté, nombre d'itérations et temps de calcul pour une précision équivalente.

14.2 Avion

14.2.1 Géométrie et maillage

On étudie la diffraction d'une onde plane par un avion « simplifié », c'est à dire sans réacteur. Sur cette géométrie, un maillage hybride nous a été gracieusement fourni par la société produisant HyperMesh. Comme on peut le voir sur la figure 14.3, le maillage est constitué de tétraèdres près de l'objet, de parallélépipèdes à l'extérieur, et les pyramides sont utilisées pour assurer la transition. Le maillage volumique contient 83 832 hexaèdres, 47 041 tétraèdres, 3876 pyramides et aucun prisme. Les hexaèdres et les pyramides sont ici tous affines.

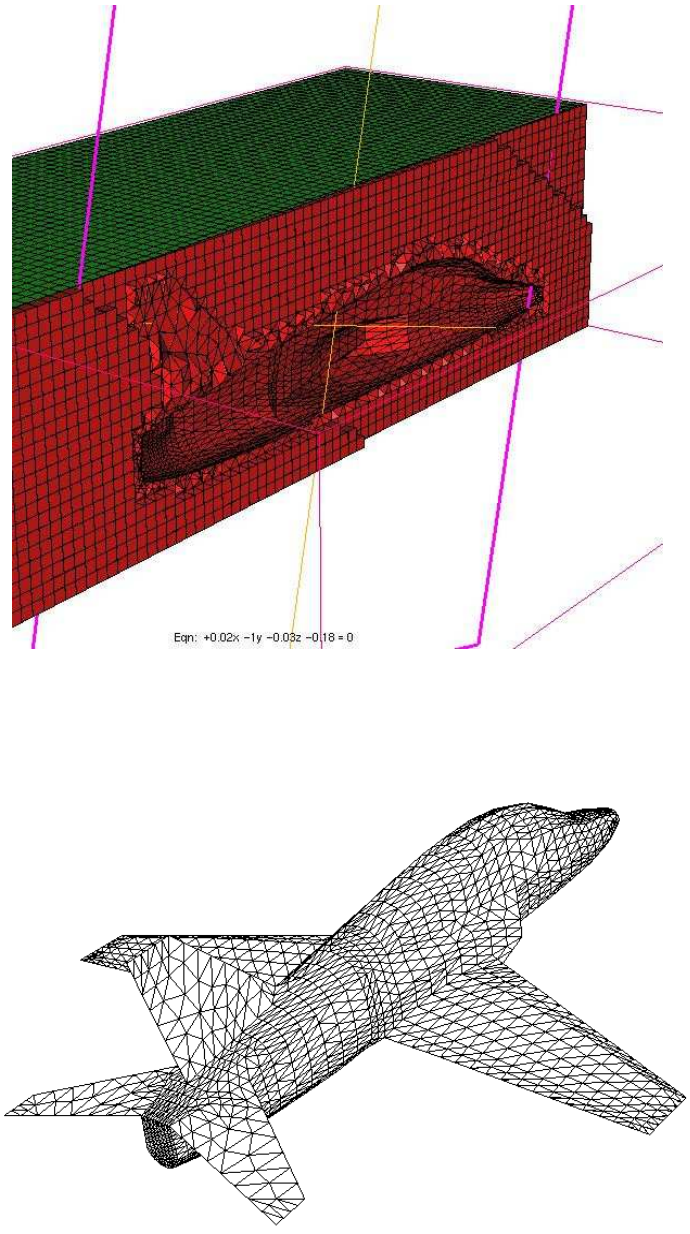

FIG. 14.3 – Maillage hybride autour d'un avion (en haut) et maillage de surface de l'avion extrait a partir du maillage hybride (en bas)

Le domaine de calcul est environ de la taille d'un parallélépipède $30\lambda \times 22\lambda \times 8\lambda$, où λ est la longueur d'onde, et on prend ici pour vecteur d'onde

$$k = \begin{vmatrix} \omega \sin\theta \cos\phi \\ \omega \sin\theta \sin\phi \\ \omega \cos\theta \end{vmatrix}$$

On a choisi $\theta = 90°$ et $\phi = 60°$.

14.2.2 Équation de Helmholtz

On prend pour champ incident

$$u^{\text{inc}} = \exp^{ik\cdot x}$$

et on prend une une fréquence adimensionnelle de 2, soit 600 Mhz.

La partie réelle du champ diffracté obtenu est affichée sur la figure 14.4.

Afin d'éviter de mesurer des erreurs géométriques, puisque l'on ne dispose pas d'éléments courbes sur cette géométrie, on compare la solution numérique obtenue pour les ordres 4 et 5 sur le même maillage. Lorsqu'on compare la solution entre le maillage hybride et le maillage tétraèdres découpés, on a observé une erreur de 14.3%.

Pour des éléments d'ordre 4, on obtient les performances du tableau 14.2 sur un maillage hybride, un maillage tétraédriques et un maillage hexaédrique obtenu par découpage d'un maillage tétraédrique.

Type de maillage	Hybride	Hexaédrique (tétras découpés)	Tétraédrique
Nombre ddls	6.08 millions	13.2 millions	5.39 millions
Erreur L^2	1.05 %	3.1 %	1.14 %
Nombre d'itérations COCG	13 113	94 500	24 325
Temps de calcul	24 253s	981 139s	80 274s
Nombre d'itérations avec préconditionneur	193	781	268
Temps de calcul avec préconditionneur	2 870s	68 354s	9 117s
Stockage de la matrice	1Go	9,16Go	3.15Go

TAB. 14.2 – Performances obtenues pour la diffraction d'un avion avec des éléments d'ordre 4 pour l'équation de Helmholtz.

On a fait tourner le cas sans préconditionneur sur 128 processeurs, on note ici le temps de calcul total obtenu en sommant les temps sur les différents processeurs, communications comprises. Pour le cas avec préconditionneur, on utilise une itération multigrille avec amortissement, comme détaillé dans la thèse de Marc Duruflé [28], et le cas est lancé sur un seul processeur.

Pour le maillage hybride, on a stocké la matrice : les hexaèdres du maillage étant des cubes, la matrice de rigidité contient beaucoup plus de zéros que dans le cas quelconque, et il est alors plus avantageux de stocker la matrice quel que soit r. Sur ce cas, la taille de la matrice pour les deux types de maillage explique en partie la grande différence entre les temps de calcul. Il est clair qu'ici les maillages hybrides fournissent des gains de performances appréciables.

14.2.3 Équations de Maxwell

On reprend le même cas pour les équations de Maxwell, mais avec une longueur d'onde deux fois plus élevée, c'est à dire une fréquence adimensionnelle de 1 (soit 300 Mhz). Le domaine est donc de taille $15\lambda \times 11\lambda \times 4\lambda$. L'angle d'incidence est le même, et l'onde est polarisée suivant e_z.

La solution numérique a été calculée avec la première famille d'ordre 2, qui est effectivement d'ordre 2 puisque tous les éléments sont affines. Le maillage contient 2,43 millions de degrés de liberté.

On observe alors le champ diffracté de la figure 14.5 sur un maillage contenant 2.43 millions de degrés de liberté. Le champ diffracté est principalement polarisé suivant e_z comme on le voit par exemple sur la composante suivant x qui est relativement petite en comparaison avec la composante suivant z.

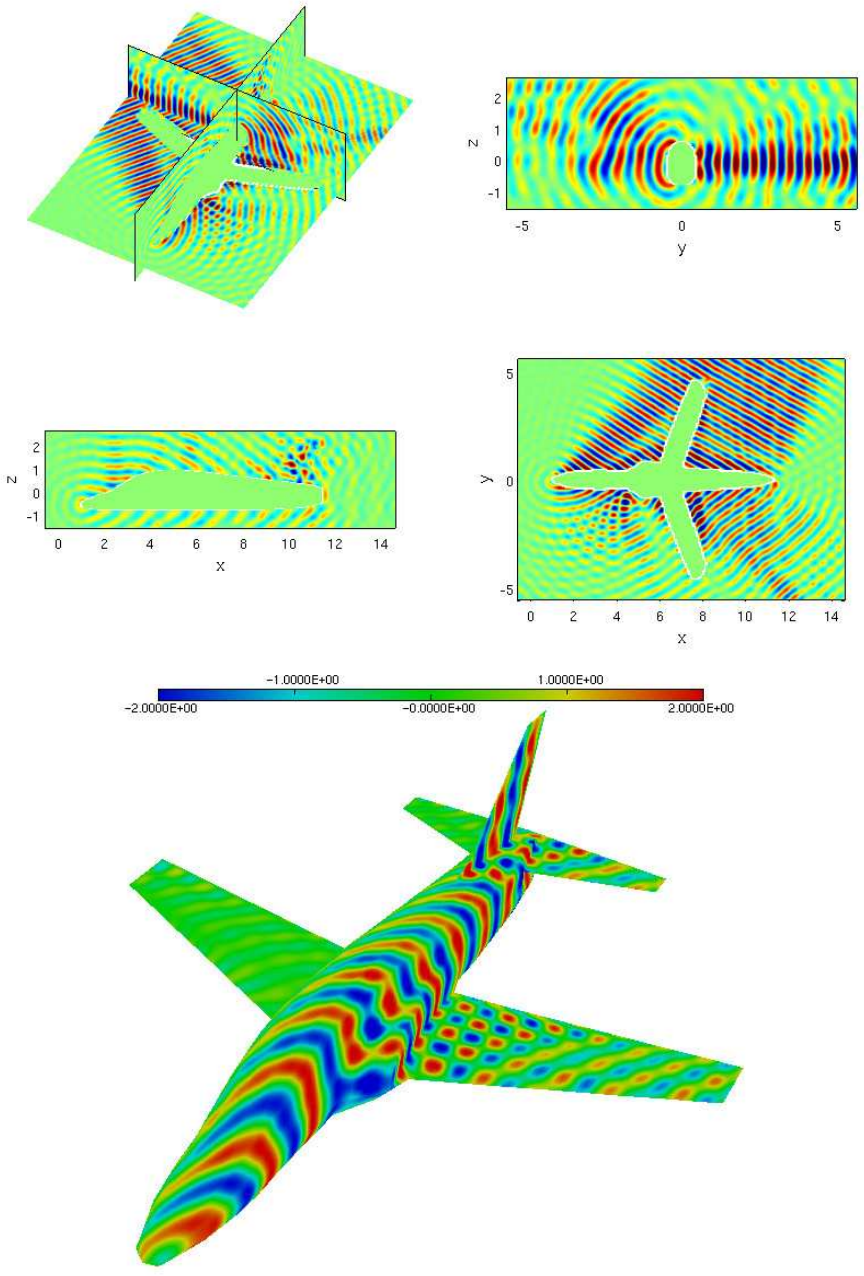

FIG. 14.4 – Partie réelle du champ diffracté.

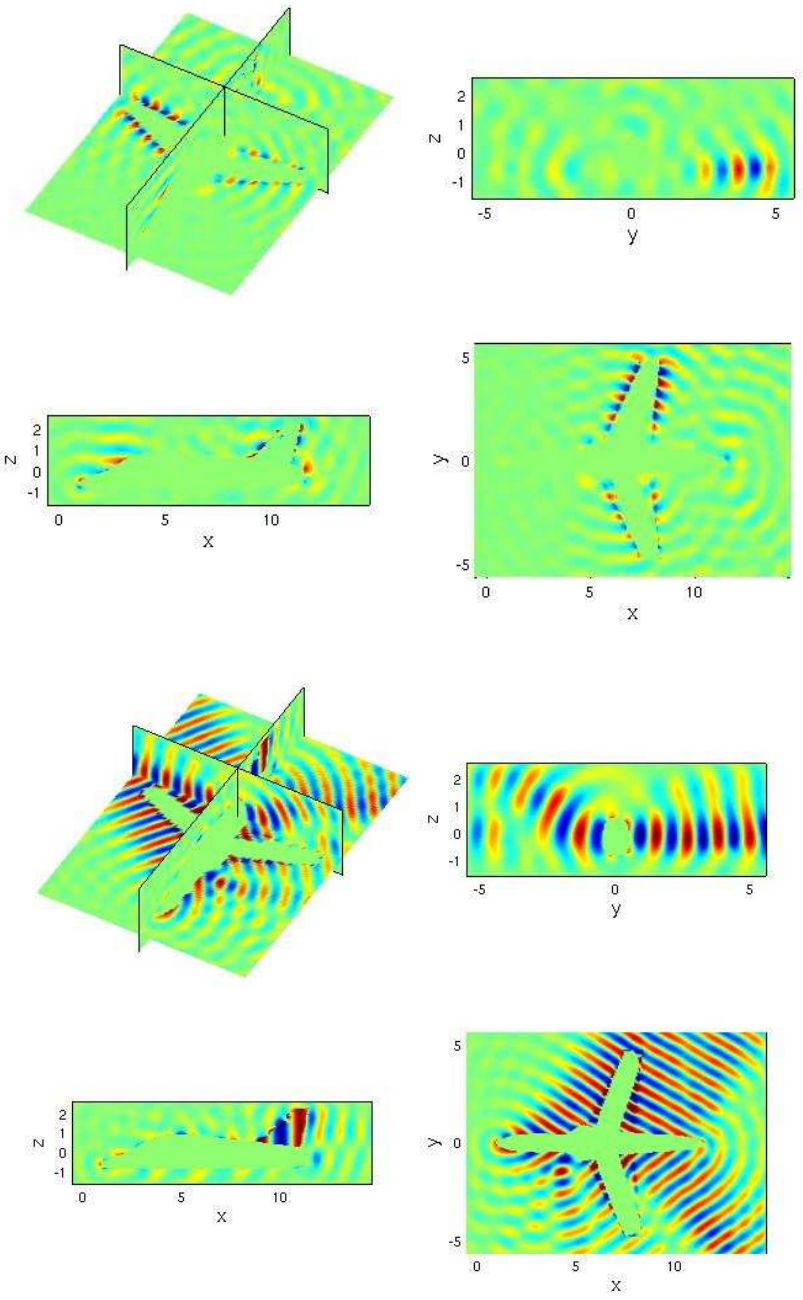

FIG. 14.5 – Partie réelle du champ diffracté par l'avion suivant e_x (en haut) et suivant e_z (en bas).

Chapitre 15

Expériences numériques en régime temporel

On présente ici les résultats des expériences numériques effectuées sur des géométries complexes en régime temporel. On utilise donc les éléments finis pour formulation discontinue avec différents types d'équations issus des problèmes de propagation d'onde.

Sommaire

15.1 Équation des ondes

15.1.1 Piano

On considère l'équation des ondes avec $c = 1$ sur la cavité résonante d'un piano Γ placée dans une boite parallélépipédique Σ.

$$\begin{cases} \dfrac{\partial^2 u}{\partial t^2} - \Delta u = f(x,t) \quad \text{in } \Omega \\ \dfrac{\partial u}{\partial n} = 0 \quad \text{on } \Gamma \\ \dfrac{\partial u}{\partial n} + \dfrac{\partial u}{\partial t} = 0 \quad \text{on } \Sigma, \end{cases} \tag{15.1.1}$$

L'ensemble du maillage est présenté sur la figure 15.1.

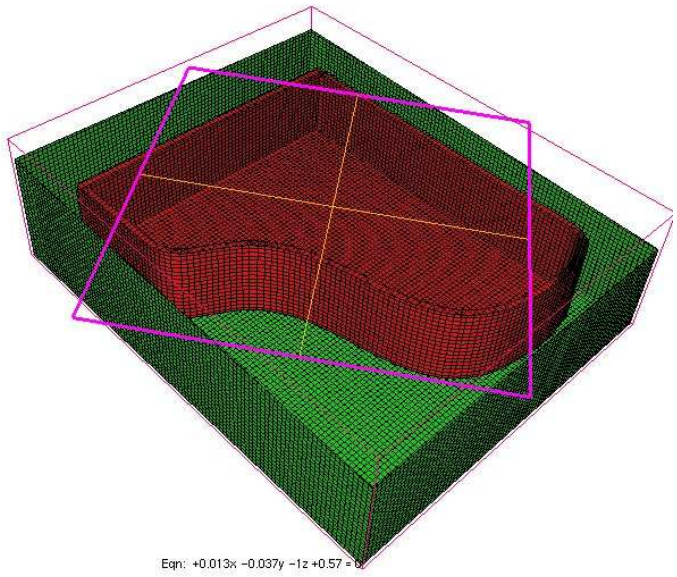

FIG. 15.1 – Maillage surfacique de la cavité en forme de piano et de la boite qui l'entoure.

La source est prise comme

$$f(x,t) = \frac{1}{r_0^2} e^{-13\frac{r}{r_0}^2} e^{-4(t-t_0)^2} sin(2\pi f_0 t), \tag{15.1.2}$$

où r est la distance du centre à la source, r_0 est la distribution radiale de la gaussienne, f_0 est la fréquence et t_0 une constante. On a pris

$$r_0 = 0.1, \;\; f_0 = 14, \;\; t_0 = 1.858, \tag{15.1.3}$$

de sorte que la taille de la boite de calcul soit $32\lambda \times 26\lambda \times 10\lambda$, où $\lambda = \dfrac{1}{f_0}$ est la longueur d'onde. Pour la discrétisation en temps, on utilise un schéma de saute-mouton d'ordre 2 (Cohen et Fauqueux [19]). On calcule la solution de $t = 0$ à $t = 6$. La solution à $t = 6$ est présentée sur la figure 15.2.

La solution de référence est calculée sur un maillage très fin, et on compare deux types de maillages : un maillage hybride et un maillage hexaédrique obtenu à partir d'un maillage tétraédrique dont on a découpé chaque élément en 4 hexaèdres. On utilise une approximation d'ordre 3 pour la discrétisation spatiale. Les résultats de l'expérience sont données dans le tableau 15.1 qui précise le temps de calcul que l'on aurait obtenu sur un seul processeur en additionnant les temps de calcul sur chaque processeur et en soustrayant les temps de communication.

Comme on utilise des éléments droits, pour éviter d'avoir une mauvaise approximation de la géométrie, on utilise un maillage tétraédrique assez fin avant de découper chaque tétraèdre en hexaèdre, si bien que le nombre de degrés de liberté dans le cas du maillage hexaédrique est très élevé.

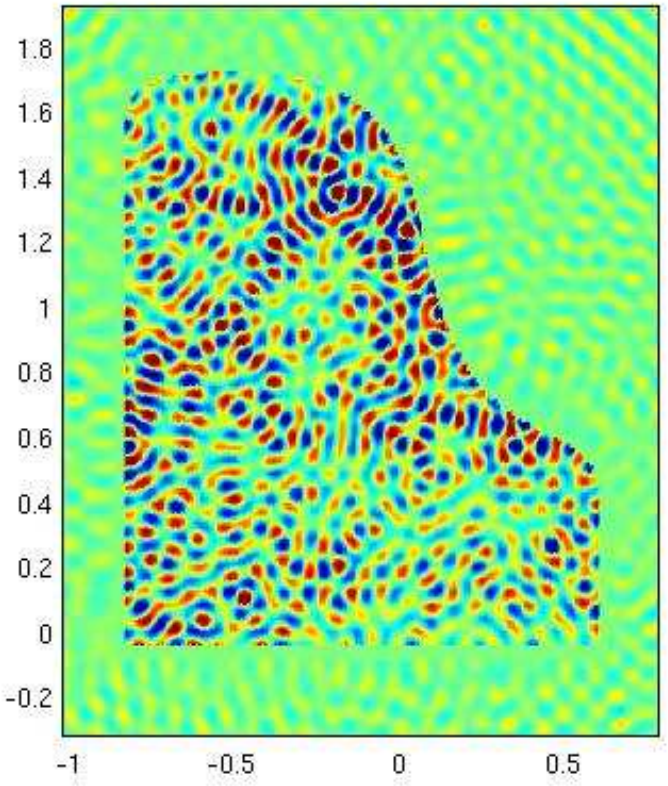

FIG. 15.2 – Solution pour la cavité en forme de piano sur une section horizontale du domaine à $t = 6$.

Type de maillage	Tétraèdres découpés	Tétraèdres	Hybride
Précision obtenue	9.4 %	5.7 %	6.3 %
Nombre de ddl	49.3 millions	16.9 millions	14.88 millions
Pas de temps	$\Delta t = 0.0002$	$\Delta t = 0.0004$	$\Delta t = 0.0005$
Temps de calcul	**12.28 jours**	**4.3 jours**	**1.18 jours**

TAB. 15.1 – Efficacité de différents types de maillages pour le piano.

15.2 Équations de Maxwell

15.2.1 Cas-test de la sphère

On considère les équations de Maxwell sur une sphère parfaitement conductrice de diamètre 10λ placée dans un cube de côté 16λ. On considère que $\varepsilon = \mu = 1$ et une source égale à

$$f(r,t) = e^{-13\,r^2}\,(t-1)\,e^{-\pi^2(t-1)^2}$$

On s'intéressera au cas où la fréquence centrale de la source est égale à 1.

Pour un niveau de précision similaire, on compare différents types de maillages ,présentés sur la figure 15.3, pour un ordre d'approximation égal à 3.

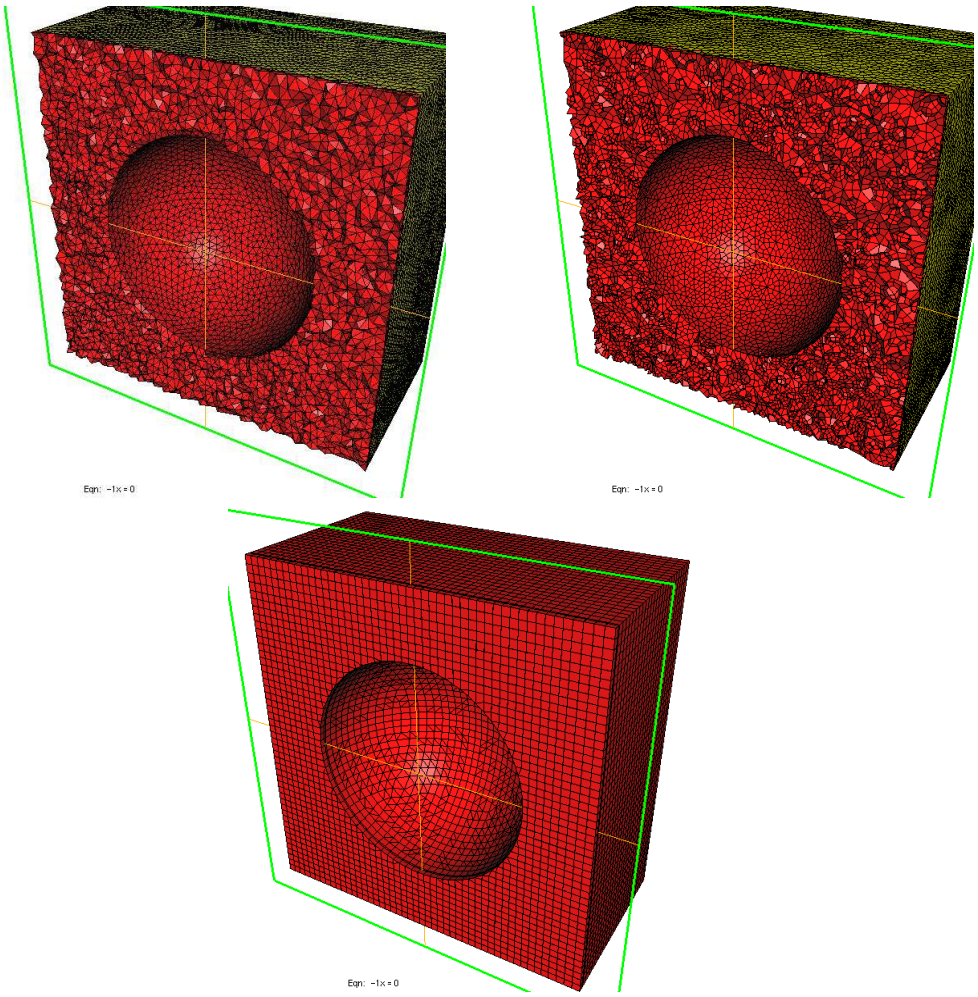

FIG. 15.3 – Maillage purement tétraédrique (à gauche), purement hexaédrique obtenu en découpant des tétraèdres en 4 (à droite), maillage mixte (en bas)

La figure 15.4 présente la réflexion de l'onde sur la sphère, tandis que le tableau 15.2 résume les résultats obtenus pour les différents types de maillage. Le nombre de degrés liberté est compté pour une seule inconnue (par exemple E_x), il faut donc multiplier par 3 pour obtenir le nombre d'inconnues totales pour modéliser le champ E. Le temps de calcul est pris pour T=10s, en additionnant le temps de calcul de chaque processeur et en retranchant le coût des communications. Ce temps est donc équivalent au temps qu'aurait pris la simulation sur un seul processeur.

Au vu des résultats, il est clair que le maillage mixte permet d'obtenir un gain important en temps de calcul.

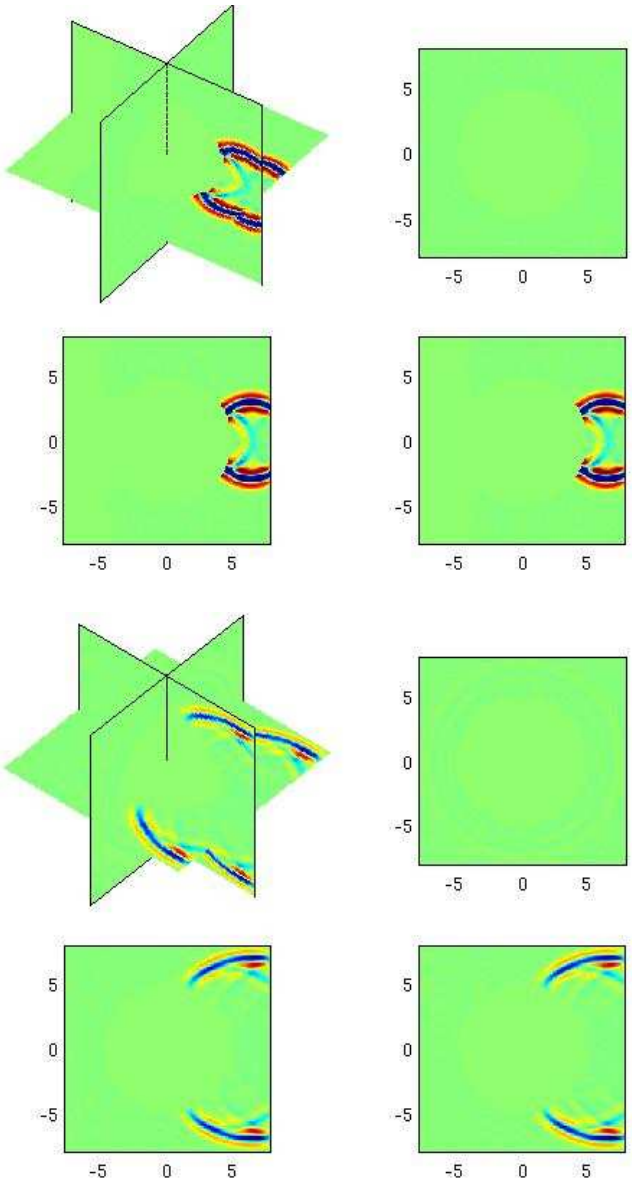

FIG. 15.4 – Composante E_x pour $t = 4$, et $t = 8$

TAB. 15.2 – Erreur, nombre de degrés de liberté, pas de temps et temps de calcul pour les différents types de maillages

Type de maillage	Tétraédrique	Hexaédrique	Hybride
Données	erreur de 7.7% 13.3 millions ddls $\Delta t = 0.01$	erreur de 6.6% 27.8 millions ddls $\Delta t = 0.0035$	erreur de 3.2% 6.3 millions ddls $\Delta t = 0.01$
Temps de calcul	**12h 43min**	**1j 21h 7min**	**2h 9min**

15.2.2 Montgolfière

On considère la diffraction par une montgolfière placée dans une boite parallélépipédique de taille $[-250, 50] \times [-130, 180] \times [90, 490]$, avec la source suivante

$$f(x,t) = e^{-13.8(\frac{r}{r_0})^2} e^{-0.001(t-t_0)^2} sin(2\pi f_0 t)$$

où $r_0 = 15$, $f_0 = 0.08$. On place une condition de conducteur parfait sur le bord de la montgolfière, et des conditions de Silver-Müller sur la boite.

Le maillage hybride utilisé pour ce calcul, présenté sur la figure 15.5, a été choisi pour donner une erreur en norme L^2 inférieure à 1% avec une approximation d'ordre 5,

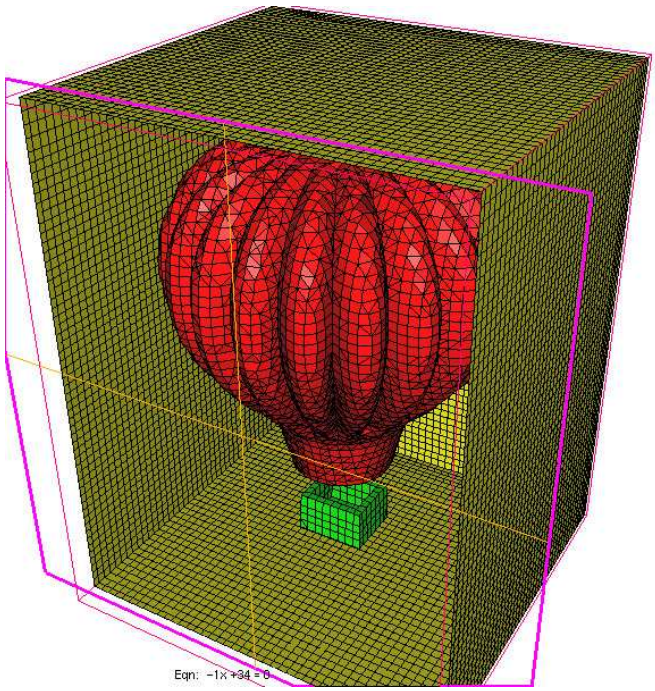

FIG. 15.5 – Maillage hybride de la montgolfière.

Pour cette expérience, on utilise le schéma saute-mouton classique pour la discrétisation en temps. La figure 15.6 montre la composante E_x de la solution numérique obtenue pour $t = 288$ et $t = 432$.

On compare chacune des solutions à une solution de référence calculée sur le même maillage mais en utilisant une approximation d'ordre $r+1$ au lieu de r. Le tableau 15.3 détaille les temps de calcul nécessaires à l'obtention d'une erreur en norme L^2 inférieure à 1%. On compare également l'utilisation de fonctions nodales et orthogonales (voir chapitre 9) dans le cas du maillage hybride.

TAB. 15.3 – Erreur, nombre de degrés de liberté, pas de temps et temps de calcul pour les différents types de maillages

Type de maillage	Tétraédrique	Hybride	
		Nodal	Ortho
Données	erreur de 11% 37.9 millions ddls $\Delta t = 0.046$	erreur de 9.3% 22.4 millions ddls $\Delta t = 0.032$	
Temps de calcul	**11j 10h 19min**	**4j 2h 36min**	**3j 17h 2min**

Cette expérience numérique a été réalisée sur 256 processeurs, le temps de calcul indiqué étant la somme des temps CPU obtenus sur chaque processeur, en retirant le temps de communication.

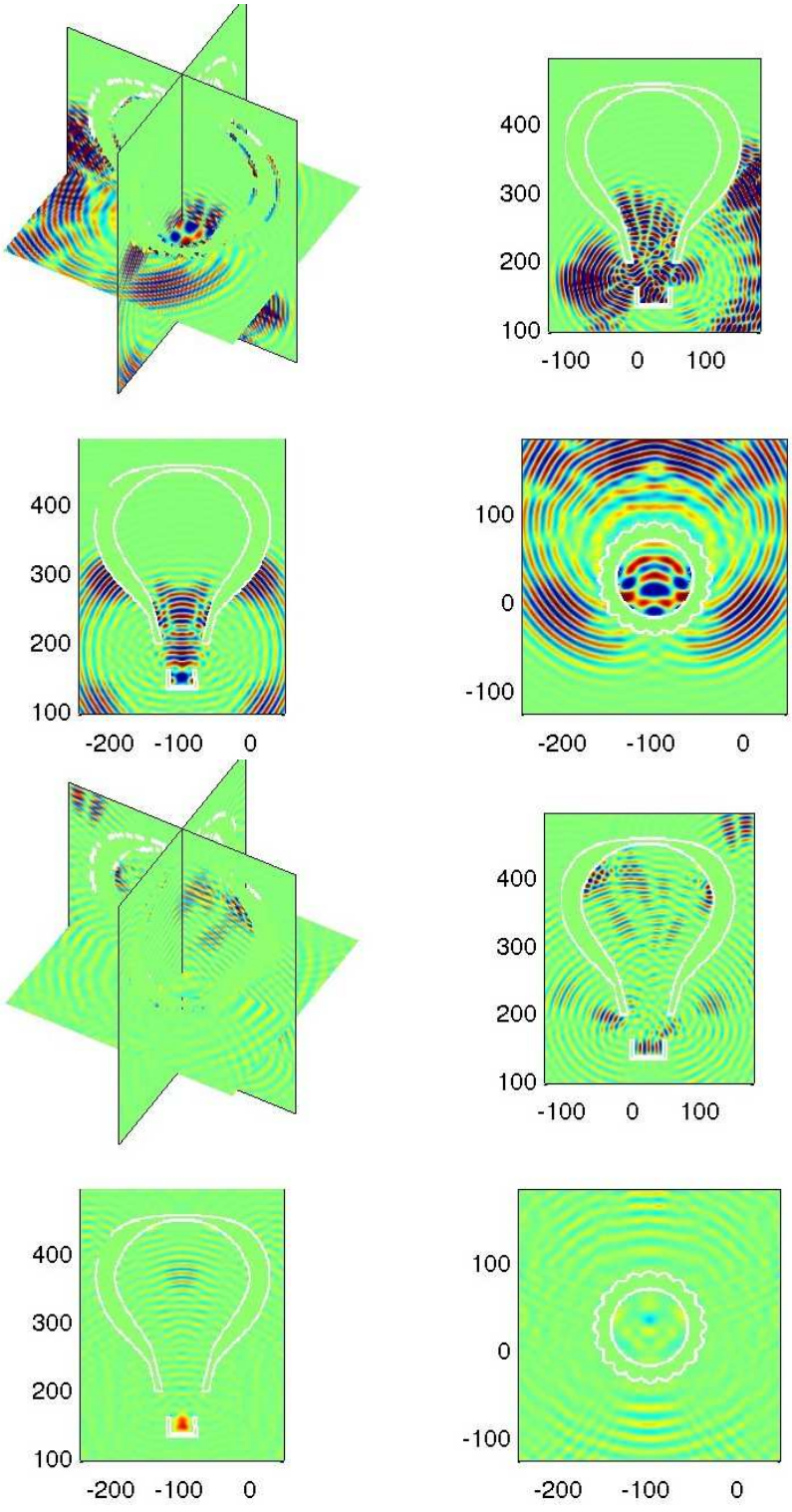

FIG. 15.6 – Solution obtenue à $t = 288$ (haut) et $t = 432$ (bas).

15.2.3 Avion

On considère à présent les équations de Maxwell en régime instationnaire sur le cas de l'avion présenté dans la section 14.2. On considère une source gaussienne en espace et sinusoïdale en temps, des conditions de Dirichlet (conducteur parfait) sur le bord de l'avion et des conditions absorbantes sur le bord de la boite de calcul.

On utilise une méthode de Galerkin discontinue avec des fonctions de base nodales pour les tétraèdres, des fonctions de base orthogonales pour la pyramide, et les points de Gauss pour les hexaèdres. On choisit une fréquence adimensionnelle de 3 (soit 900 Mhz), le centre de la gaussienne est placé en $(13, 4, 1.5)$ et le rayon de distribution est de 0.4.

Pour des éléments d'ordre 4, on mesure le temps de calcul obtenu pour $T = 120\,t_0$, où t_0 est la période de la source sinusoïdale, ce qui revient à prendre un temps final physique de $1.33e-7$. Les résultats obtenus sont indiqués dans le tableau 15.4. La solution obtenue à $t = 15\,t_0$ et à $t = 52.5\,t_0$ est présentée sur la figure 15.7. Les calculs ont été effectués sur 256 processeurs en sommant les temps de chaque processeur et en retranchant le coût des communications. On a observé une efficacité parallèle supérieure à 80%.

Type de maillage	Hybride	Tétraèdres découpés
Nombre ddls	12.3 millions	22.9 millions
Erreur L^2	3.84 %	4.55 %
Pas de temps Δt	0.0014	0.00056
Temps de calcul	**5j 18h 6min**	**13j 12h 48min**

TAB. 15.4 – Performances de l'avion pour les équations de Maxwell en régime temporel

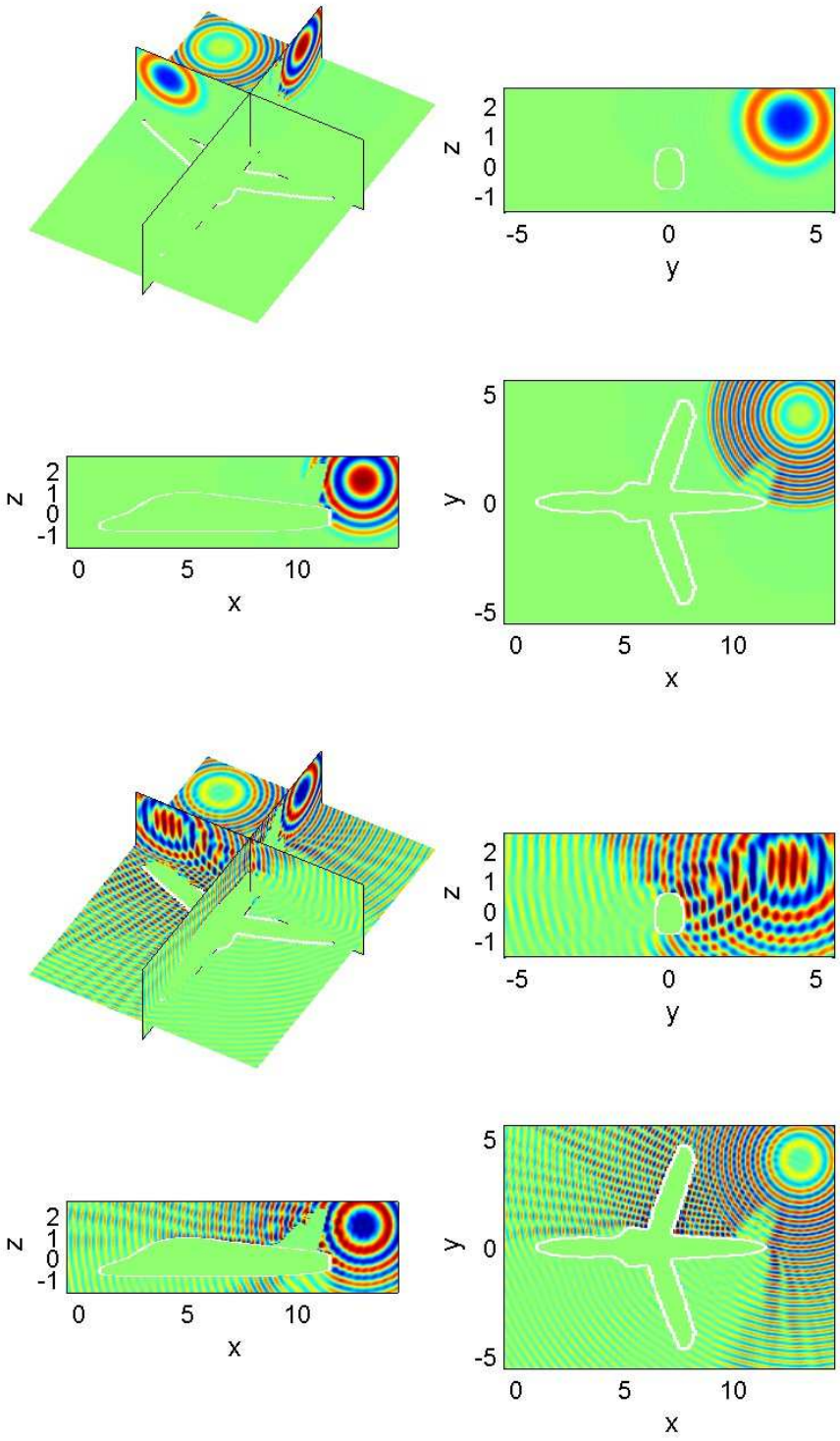

FIG. 15.7 – Solution à $t = 15t_0$ et à $t = 52.5t_0$

15.3 Amélioration de la CFL

15.3.1 Ordre variable

Afin d'utiliser un nombre plus réduit de degrés de liberté pour l'expérience de la section 15.2.3, on va adapter l'ordre de telle sorte à approcher au mieux la règle des dix points par longueur d'onde.

Soit h_i la longueur moyenne des arêtes de l'élément i, on a utilisé la règle suivante pour déterminer l'ordre pour chaque élément.

- Si $h_i < 0.014$, ordre 1,
- Si $h_i < 0.105$, ordre 2,
- Si $h_i < 0.175$, ordre 3,
- Si $h_i < 0.287$, ordre 4,
- Si $h_i < 0.378$, ordre 5

Avec cette règle, sachant que le pas de maillage des cubes est de 0.2, la plupart des éléments sont d'ordre 4 comme pour l'ordre constant. Il n'y a pas d'élément d'ordre 1,et l'ordre maximal est 5 sur tout le maillage. Le maillage de l'avion avec l'ordre variable est présenté sur la figure 15.8.

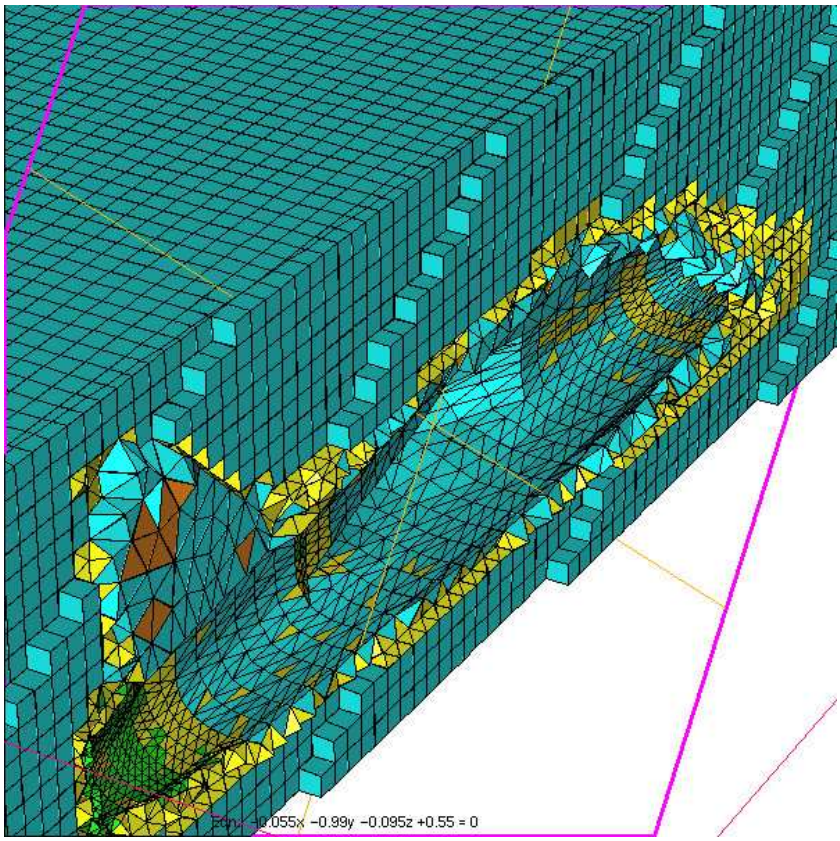

FIG. 15.8 – Maillage du jet avec ordre variable : ordre 2 en vert, ordre 3 en jaune, ordre 4 en cyan, ordre 5 en orange

15.3.2 Pas de temps local

Pour réduire encore le temps de calcul, il est avantageux de considérer un pas de temps local. En effet, la CFL est restreinte par le plus petit élément du maillage. Pour la stratégie de pas de temps local, nous avons utilisé le schéma symplectique de Piperno [63]. Une première étape consiste à calculer le pas de temps associé à chaque élément. Pour ce faire, on évalue la plus grande valeur propre de la matrice associée à un petit maillage contenant l'élément et ses voisins, comme représenté sur la figure 15.9.

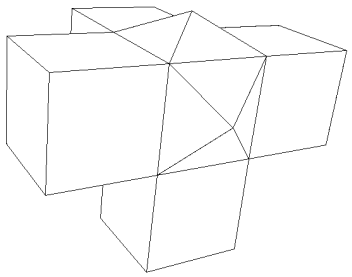

FIG. 15.9 – Petit maillage utilisé pour évaluer la CFL de l'élément central

On définit alors le pas de temps Δt_e de l'élément

$$\Delta t_e = \frac{c_r}{\lambda_{max}}$$

où c_r est un coefficient de sécurité calculé sur un maillage régulier, qui dépend donc de l'ordre d'approximation r. On choisit pour valeur de c_r les valeurs suivantes

$$c_r = \begin{cases} 0.89 \text{ si } r = 1 \\ 0.95 \text{ si } r = 2 \\ 0.98 \text{ si } r = 3 \\ 0.99 \text{ si } r = 4 \\ 0.992 \text{ si } r \geq 5 \end{cases} .$$

Cette approche a été validée en 2D : que ce soit sur des maillages réguliers ou sur des maillages quelconques, la CFL ainsi obtenue était toujours inférieure à la CFL exacte, et le taux d'erreur entre les deux CFL ne dépassait pas le pourcent.

Une fois les pas de temps locaux optimaux calculés, on se fixe un pas de temps nominal $\Delta t_{\text{nominal}}$ strictement inférieur au pas de temps maximal et on affecte à chaque élément un niveau ℓ. Dans l'approche de Piperno, un élément de niveau ℓ a pour pas de temps $\dfrac{\Delta t}{2^\ell}$ et on dit que si $\dfrac{\Delta t_{\text{nominal}}}{2^\ell} \leq c\,\Delta t_e < \dfrac{\Delta t_{\text{nominal}}}{2^{\ell-1}}$, l'élément e est de niveau ℓ, c étant un coefficient de sécurité. En pratique, on a pris $c = 0.99$.

Le niveau de chaque élément du maillage de l'avion est représenté sur la figure 15.10 et la répartition des différents niveaux ℓ est présentée sur le tableau 15.5.

Niveau	Hybride	Tétraèdres découpés
0	83282	47715
1	48375	133568
2	1290	3691
3	61	11

TAB. 15.5 – Nombre d'éléments par niveau pour les deux maillages.

Une fois les niveaux ℓ déterminés, il est parfois nécessaire de prendre un pas de temps Δt plus petit que le pas de temps nominal pour des raisons de stabilité. En effet le schéma de Piperno ne permet pas de contrôler de manière très précise la CFL globale du schéma en fonction des CFL locales, si bien qu'un ajustement est parfois nécessaire. Dans le cas des tétraèdres découpés, on a dû prendre $\Delta t = 0.004$ alors que $\Delta t_{\text{nominal}} = 0.005$. En revanche, pour le maillage hybride, aucune instabilité n'a été observée avec ce pas de temps.

Les résultats obtenus sont indiqués dans le tableau 15.6 sur lequel figure le ratio entre le pas de temps maximal et le pas de temps minimal obtenu sur les deux maillages. On voit que pour le maillage hybride, certains éléments sont assez contraignants, puisque le ratio dépasse 9. Les temps de calculs obtenus ont été mesurés sur un seul processeur puisque nous n'avons pas disposé de suffisamment de temps pour paralléliser de manière satisfaisante la stratégie de pas de temps local.

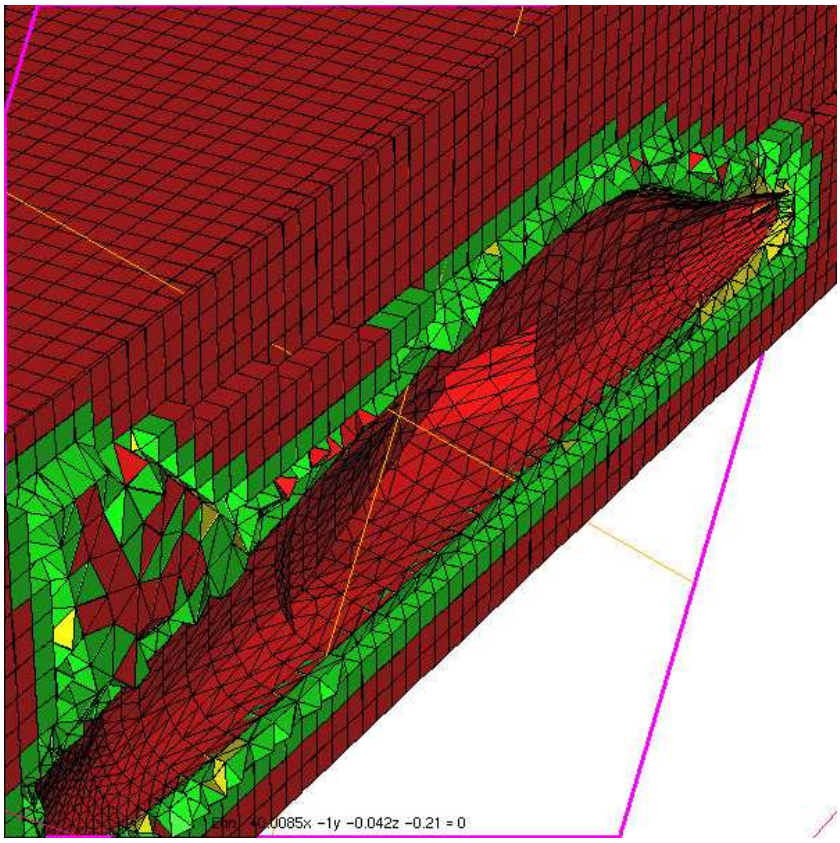

FIG. 15.10 – Maillage de l'avion avec pas de temps local : niveau 0 en rouge, niveau 1 en vert, niveau 2 en jaune et niveau 3 en cyan

Type de maillage	Hybride	Tétraèdres découpés
Nombre ddls	11.7 millions	20.7 millions
Erreur L^2	4.07 %	5.15 %
Ratio $\frac{\Delta t_{max}}{\Delta t_{min}}$	9.04	6.85
Pas de temps	0.005	0.004
Temps de calcul	**1j 2h**	**4j 11h**

TAB. 15.6 – Performances de l'avion pour les équations de Maxwell en régime temporel avec pas de temps local

Conclusion

Conclusions générales

Le but de ces travaux était de construire des éléments finis d'ordre élevé compatibles avec l'utilisation de maillages hybrides conformes pour la résolution de systèmes linéaires hyperboliques en régimes temporel et harmonique. L'objectif était plus particulièrement de mettre au point des éléments finis pour des formulations H^1, $H(rot)$ et LDG qui soient optimaux au sens de la convergence pour la norme de l'espace considéré. L'implémentation efficace des différents éléments construits, l'étude de leurs propriétés numériques, leur comparaison à différents éléments pouvant être trouvés dans la littérature et leur applications à des cas concrets ont été des points cruciaux au cours de cette thèse.

Dans les parties II et IV qui traitent respectivement des formulation H^1 et $H(rot)$, nous nous sommes efforcés de construire de manière systématique des éléments finis d'ordre élevé pour tous les types d'éléments (hexaèdres, prismes, tétraèdres et pyramides) de telle sorte qu'ils puissent être intégrés à un maillage hybride conforme. Les résultats concernant ces deux formulations sont les suivants

- la construction d'éléments finis nodaux et hiérarchiques, optimaux au sens de la convergence en norme H^1 ou $H(rot)$, pour un ordre quelconque
- une étude théorique des formules de quadrature à utiliser pour l'intégration exacte des matrices de masse et de rigidité dès que cela est possible
- un calcul d'estimations d'erreur sur des maillages hybrides : l'estimation de l'erreur d'interpolation a été obtenue dans le cas des formulations H^1 et $H(rot)$, tandis que l'erreur de quadrature sur la matrice de masse a été obtenue dans le cas de la formulation H^1
- des résultats numériques sur la dispersion et la stabilité des éléments construits, l'absence de modes parasites pour les éléments finis $H(rot)$
- la vérification numérique de l'optimalité des éléments
- le bon comportement des éléments au sein d'un maillage hybride
- la comparaison théorique et numérique de nos éléments avec des éléments trouvés dans la littérature, en particulier pour les éléments pyramidaux.
- la construction d'une première famille d'éléments finis d'arête pyramidaux pour la formulation $H(rot)$ permettant leur utilisation avec les éléments finis d'arête de la première famille classiques

On vérifie également que les deux espaces d'approximation construits vérifient le diagramme de De Rham.

En ce qui concerne les éléments discontinus dont il est question dans la partie III, nous sommes parvenus à

- construire des éléments finis tensorisés optimaux au sens de la convergence en norme L^2, pour un ordre quelconque
- mettre au point un algorithme rapide de construction de la matrice de masse pour les éléments pyramidaux et prismatique, ainsi qu'un algorithme de produit matrice-vecteur rapide pour tous les éléments, en utilisant les propriétés des fonctions de base semi-orthogonales.
- obtenir des résultats de dispersion et de stabilité pour les éléments construits
- vérifier l'optimalité de la convergence pour tous les éléments

Nous avons ensuite utilisé ces éléments pour réaliser des expériences numériques sur des cas réels. Dans la partie V, nous montrons ainsi les avantages de l'utilisation de maillages hybrides par rapport aux maillages purement tétraédriques ou purement hexaédriques obtenus en découpant chaque tétraèdre d'un maillage tétraédrique en hexaèdre. Les résultats obtenus sont concluants, bien que les outils de maillage hybride ne soient pas encore très au point.

Des résultats numériques ont en outre pu être obtenu sur des maillages utilisant des éléments droits ou courbes. L'ordre variable et une stratégie de pas de temps local ont également été implémentés et utilisés dans un cas concret pour réduire la restriction sur le pas de temps due à la CFL.

Perspectives

Plusieurs travaux restent à entreprendre à l'issue de cette thèse.

Du point de vue théorique, ils portent majoritairement sur les erreurs de quadrature pour la matrice de masse dans le cas $H(rot)$ et pour la matrice de rigidité dans les cas H^1 et $H(rot)$

Concernant l'aspect numérique, les points suivants peuvent être étudiés pour améliorer les résultats obtenus dans les présents travaux

- Trouver des formules de quadrature permettant d'augmenter légèrement la CFL des éléments $H(rot)$
- Implémenter une méthode de pas de temps local plus performante pour accélérer les calculs.

On peut également envisager la construction des éléments finis d'arête optimaux de la seconde famille pour la formulation $H(rot)$, bien que l'intérêt de la construction d'une telle famille soit discutable lorsque l'on dispose d'éléments finis permettant d'avoir une convergence optimale pour la norme $H(rot)$. Mais la principale perspective est la construction d'éléments finis optimaux d'ordre élevé pour une formulation $H(div)$ afin de compléter le diagramme de De Rham. La procédure mise au point pour H^1 et $H(rot)$ étant systématique, elle s'étend aisément au cas $H(div)$.

Bibliographie

[1] P. Amestoy, T. A. Davis, and I. S. Duff. Algorithm 837 - amd, an approximate minimum degree ordering algorithm. *ACM Transactions on Mathematical Software*, 30(3) :381–388, 2004.

[2] D. Arnold, D. Boffi, and R. Falk. Approximation by quadrilateral finite elements. *Mathematics of Computation*, 71(239) :909–922, 2002.

[3] D. N. Arnold, D. Boffi, and R. S. Falk. Quadrilateral $h(div)$ finite elements. *SIAM J. Numer. Anal.*, 42(6) : 2429–2451, 2005.

[4] I. Babuska and J. Osborn. *Eigenvalue Problems*. 1991.

[5] G. Bedrosian. Shape functions and integration formulas for three-dimensional finite element analysis. *International Journal of Numerical Methods in Engineering*, 35 :95–108, 1992.

[6] M. Bergot and P. Lacoste. Generation of higher-order polynomial bases of nédélec $h(curl)$ finite elements for Maxwell's equations. *Journal of Computational and Applied Mathematics*, 234(6) :1937–1944, 2010.

[7] M. Bluck and S. Walker. Polynomial basis functions on pyramidal elements. *Comm. Numer. Meth. Engng.*, 24 :1827–1837, 2008.

[8] D. Boffi, P. Fernandes, L. Gastaldi, and I. Perugia. Computational models of electromagnetic resonators : analysis of edge element approximation. *SIAM Journal on Numerical Analysis*, 36 :1264–1290, 1999.

[9] A.-S. Bonnet-BenDhia, E. Lunéville, and P. Ciarlet. La méthode des éléments finis. Technical report, Cours de MASTER MA201, 2008.

[10] A. Bossavit. A uniform rationale for whitney forms on various supporting shapes. *Mathematics and Computers in Simulation*, 80 :1567–1577, 2009.

[11] H. Carpenter and C. A. Kennedy. Fourth-order 2n-storage runge-kutta schemes. Technical report, NASA Langley Research Center, 1994.

[12] N. Castel, G. Cohen, and M. Duruflé. Discontinuous Galerkin method for hexahedral elements and aeroacoustic. *Journal of Computational Acoustics*, 17(2) :175–196, 2009.

[13] V. Chatzi and F. Preparata. Using pyramids in mixed meshes - point placement and basis functions. Technical report, Brown University, 2000.

[14] P. Ciarlet. *The Finite Element Method for Elliptic Problems*. North-Holland, 1978.

[15] M. Clemens and T. Weiland. Iterative methods for the solution of very large complex symmetric linear systems of equations in electrodynamics. Fachbereich 18 elektrische nachrichtentechnik, Technische Hochschule Darmstadt, 2002.

[16] B. Cockburn and C.-W. Shu. The local discontinuous Galerkin method for time-dependent convection-diffusion systems. *SIAM Journal on Numerical Analysis*, 35 :2440–2463, 1998.

[17] G. Cohen. *Higher-Order Numerical Methods for Transient Wave Equations*. Springer Verlag, 2002.

[18] G. Cohen and M. Duruflé. Non spurious spectral-like element methods for Maxwell's equations. *Journal of Computational Mathematics*, 25 :282–304, 2007.

[19] G. Cohen and S. Fauqueux. Mixed finite elements with mass-lumping for the transient wave equation. *Journal of Computational Acoustics*, 8 :171–188, 2000.

[20] G. Cohen, X. Ferrieres, and S. Pernet. A spatial high-order hexahedral discontinuous Galerkin method to solve Maxwell equations in time domain. *Journal of Computational Physics*, 217 :340–363, 2006.

[21] G. Cohen and P. Monk. Gauss point mass-lumping schemes for Maxwell's equations. *NMPDE Journal*, 14 (1) :63–88, 1998.

[22] J. C. Coulomb, F. X. Zgainski, and Y. Maréchal. Apyramidal element to link hexahedral, prismatic and tetrahedral edge finite elements. *IEEE Transactions on Magnetics*, 33(2) :1362–1365, 1997.

[23] L. Demkowicz, J. Kurtz, D. Pardo, M. Paszynski, W. Rachowicz, and A. Zdunek. *Computing With hp-Adaptive Finite Element, Volume II.* Chapman and Hal, 2007.

[24] C. Doucet. *Approximation des champs électromagnétiques sur les maillages éléments finis hybrides conformes.* PhD thesis, Université Joseph Fourier, 2008.

[25] C. Doucet, I. Charpentier, J.-L. Coulomb, and C. Guérin. Extraction of finite element basis functions from the cellular topology of meshes. *IIIE Transactions on Magnetics*, 44(6) :726–729, 2008.

[26] P. Dular, J.-Y. Hody, A. Nicolet, A. Genon, and W. Legros. Mixed finite elements associated with a collection of tetrahedra, hexahedra and prisms. *IIIE Transactions on Magnetics*, 30(5) :2980–2983, 1994.

[27] D. Dunavant. High degree efficient symmetrical gaussian quadrature rules for the triangle. 21 :1129–1148, 1985.

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

[29] M. Duruflé, P. Grob, and P. Joly. Influence of the gauss and gauss-lobatto quadrature rules on the accuracy of a quadrilateral finite element method in the time domain. *Numerical Methods for Partial Differential Equations*, 25 :526–551, 2009.

[30] P. R. E. Godlewski. *Hyperbolic Systems of Conservation Laws.* Ellipses, 1991.

[31] Y. A. Erlangga. *A robust and efficient iterative method for the numerical solution of the Helmholtz equation.* PhD thesis, University of Delft, 2005.

[32] R. Falk, P. Gatto, and P. Monk. Hexahedral $h(div)$ and $h(curl)$ finite elements. *ESAIM : M2AN*, 2010.

[33] G. Gassner, F. Lörcher, C. Munz, and J. Hesthaven. Polymorphic nodal elements and their application in discontinuous Galerkin methods. *Journal of Computational Physics*, 228(5) :1573–1590, 2009.

[34] W. Gautschi. Generalized gauss-radau and gauss-lobatto formulae. *BIT*, 44(4) :711–720, 2004.

[35] V. Girault and P. Raviart. *Finite Element Approximation of the Navier-Stokes Equations.* Berlin - Springer, 1979.

[36] V. Gradinaru and R. Hiptmair. Whitney elements on pyramids. *Electronic Transactions on Numerical Analysis*, 8 :154–168, 1999.

[37] R. D. Graglia, D. R. Wilton, and A. F. Peterson. Higher order interpolatory vector bases for computational electromagnetics. *IEEE Transactions on Antennas and Propagation*, 45(3) :329–342, 1997.

[38] R. D. Graglia, D. R. Wilton, A. F. Peterson, and I.-L. Gheorma. Higher order interpolatory vector bases on pyramidal elements. *IEEE Transactions on Antennas and Propagation*, 47(5) :775–782, 1999.

[39] W. Hackbusch. *Multigrid methods and applications.* Springer-Verlag, 1985.

[40] W. Hackbusch. *Iterative solution of large sparse systems of equations.* Springer Verlag, 1994.

[41] P. Hammer, O. Marlowe, and A. Stroud. Numerical integration over simplexes and cones. *Mathematical Tables and Other Aids to Computation*, 10(55) :130–137, 1956.

[42] R. Hartmann. Adjoint consistency analysis of discontinuous Galerkin discretizations. *SIAM J. Numer. Anal.*, 45(6) :2671–2696, 2007.

[43] J. Hesthaven. From electrostatics to almost optimal nodal sets for polynomial interpolation in a simplex. *SIAM Journal on Numerical Analysis*, 35(2) :665–676, 1998.

[44] J. Hesthaven and C. Teng. Stable spectral methods on tetrahedral elements. *SIAM Journal on Numerical Analysis*, 21(6) :2352–2380, 2000.

[45] J. Hesthaven and T. Warburton. High-order nodal methods on unstructured grids. i. time-domain solution of Maxwell's equations. *J. Comput. Phys.*, 181(1) :186–221, 2002.

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

[47] G. Karniadakis and S. J. Sherwin. *Spectral/hp element methods for CF - Second Edition.* Oxford University Press, 2005.

[48] R. Kirby, T. Warburton, I. Lomtev, and G. Karniadakis. A discontinuous Galerkin spectral/hp method on hybrid grids. *Applied Numerical Mathematics*, 33 :393–405, 2000.

[49] P. Knabner and G. Summ. The invertibility of the isoparametric mapping for pyramidal and prismatic finite element. *Numer. Math.*, 88 :661–681, 2001.

[50] C. Lee, S. Wong, and S. Lie. On increasing the order and density of 3d finite element meshes. *Comm. in Num. Meth. in Engin.*, 17 :55–68, 2001.

[51] L. Liu, K. Davies, K. Yuan, and M. Křížek. On symmetric pyramidal finite elements. *Dyn. Contin. Discrete Impuls. Syst. Ser. B Appl. Algorithms*, 11 :213–227, 2004.

[52] M. M. Costabel. Computation of resonance frequencies for Maxwell equations in non smooth domains. volume 31, 2003.

[53] N. Marais and D. Davidson. Conforming arbitrary order hexahedral/tetrahedral hybrid discretisation. *Electronics Letters*, 44(24), 2008.

[54] M.J.S.Chin-Joe-Kong, W. Mulder, and M. V. Veldhuizen. Higher-order triangular and tetrahedral finite element with mass lumping for solving the wave equation. *Journal of Engineering Mathematics*, 35 :405–426, 1999.

[55] P. Monk. *Finite elements methods for Maxwell's equations.* Oxford Science Publication, 2002, 2002.

[56] J. C. Nédélec. Mixed finite elements in $\mathbb{R}^3$. *Numer. Math.*, 35(3) :315–341, 1980.

[57] J. C. Nédélec. A new family of mixed finite elements in $\mathbb{R}^3$. *Numer. Math.*, 51(1) :57–81, 1986.

[58] N. Nigam and J. Phillips. Higher-order finite elements on pyramids. *submitted to Num. Math.*, 2007.

[59] N. Nigam and J. Phillips. Numerical integration for high order pyramidal finite elements. *submitted to Comp. Meth. in App. Mech. and Eng.*, 2010.

[60] S. Owen and S. Saigal. Formation of pyramid elements for hexahedra to tetrahedra transitions. *Comput. Methods Appl. Mech. Eng.*, 190 :4505–4518, 2001.

[61] S. Pernet. *Étude de méthodes d'ordre élevé pour résoudre les équations de Maxwell dans le domaine temporel. Application á la détection et á la compatibilité électromagnétique.* PhD thesis, Université de Paris IX Dauphine, 2004.

[62] S. Pernet and X. Ferrieres. hp a-priori error estimates for a non-dissipative spectral discontinuous Galerkin method to solve the Maxwell equations in the time domain. *Mathematics of Computation*, 76 :1801–1832, 2007.

[63] S. Piperno. Symplectic local time-stepping in non-dissipative DGTD methods applied to wave propagation problems. *ESAIM : M2AN*, 40(5) :815–841, 2006.

[64] Y. Saad. *Iterative methods for sparse linear systems.* Series in Computer Science, 1996.

[65] S. Sherwin. Hierarchical hp finite element in hybrid domains. *Finite Elements in Analysis and Design*, 27 : 109–117, 1998.

[66] S. Sherwin, T. Warburton, and G. Karniadakis. Spectral/hp methods for elliptic problems on hybrid grids. *Contemporary Mathematics*, 218 :191–216, 1998.

[67] A. Stroud. *Approximate calculation of multiple integrals.* Prentice-Hall Series in Automatic Computation. Prentice-Hall Inc., Englewood Cliffs, N.J., 1971.

[68] B. Szabó and I. Babuška. *Finite Element Analysis.* John Wiley & Sons, 1991.

[69] G. Szegö. *Orthogonal Polynomials - Ch. 4 Jacobi Polynomials.* 4th ed Amer. Math. Soc., Providence, RI, 1975.

[70] C. Tapp. *Anisotrope Gitter - Generierung und Verfeinerung.* PhD thesis, University of Erlangen, 1999.

[71] P. Šolín, K. Segeth, and I. Dolezel. *Higher-Order Finite Elements Methods.* Studies in Advanced Mathematics, Chapman and Hall, 2003.

[72] T. Warburton. *Spectral/hp Methods on Polymorphic Multi-Domains : Algorithms and Applications.* PhD thesis, Brown University, 1999.

[73] H. Whitney. *Geometric Integration Theory.* Princeton University Press, 1957.

[74] C. Wieners. *Conforming discretizations on tetrahedrons, pyramids, prisms and hexahedrons.* 1957.

[75] S. Zaglmayr. *High Order Finite Elements for Electromagnetic Field Computation.* PhD thesis, Johannes Kepler University, Linz Austria, 2006.

[76] F. X. Zgainski, J. C. Coulomb, and Y. Maréchal. A new family of finite elements the pyramidal elements. *IEEE Transactions on Magnetics*, 32(3) :1393–1396, 1996.

[77] S. Zhang. Invertible jacobian fot hexahedral elements - part 1 - bijectivity. *submitted to Numer. Math.*, 2005.

[78] S. Zhang. Invertible jacobian fot hexahedral elements - part 2 - global positivity. *submitted to Numer. Math.*, 2005.

[79] S. Zhang. Invertible jacobian fot hexahedral elements - part 3 - algorithm and examples. *in preparation*, 2005.

Vu : le Président

M. .....

Vu : les suffrageants

MM. .....

Vu et permis d'imprimer :

Le Vice-Président du Conseil Scientifique Chargé de la Recherche de l'Université PARIS IX DAUPHINE.

Résumé

Dans cette thèse, nous nous intéressons à la construction d'éléments finis d'ordre élevé adaptés aux maillages hybrides, pour la résolution de systèmes hyperboliques linéaires en régimes harmonique et temporel. L'accent est plus particulièrement porté sur la construction d'éléments pyramidaux.

On étudie trois formulations pour lesquelles on cherche des éléments finis « optimaux » au sens de la convergence dans la norme de l'espace considéré pour la formulation. Pour les formulations H^1 et $H(rot)$, on construit des éléments finis « optimaux » nodaux et *hp*. Les matrices élémentaires sont évaluées grâce à des formules de quadrature adaptées et des estimations d'erreur sont effectuées pour vérifier la convergence des éléments optimaux construits. Pour la formulation discontinue LDG (Local Discontinuous Galerkin), on présente des éléments utilisant des fonctions de base orthogonales permettant de mettre au point une construction de la matrice de masse et un produit matrice-vecteur rapides. Dans le cas des trois formulations, on étudie les propriétés numériques des éléments construits, on vérifie que l'on retrouve bien numériquement la convergence théorique et on compare nos éléments avec d'autres éléments trouvés dans la littérature.

Finalement, on présente des expériences numériques en 3D avec l'équation des ondes ou de Helmholtz, et les équations de Maxwell dans le cas des régimes temporels et harmoniques. On montre ainsi l'efficacité des maillages hybrides par rapport aux maillages purement tétraédriques ou aux maillages hexaédriques obtenus en découpant chaque tétraèdre d'un maillage purement tétraédrique en quatre hexaèdres.

Mots clés : maillage hybride conforme, éléments finis d'ordre élevé, méthodes de Galerkin continue et discontinue, éléments finis nodaux, *hp* et d'arête, formule de quadrature, estimations d'erreur, équations des ondes et de Helmholtz, équations de Maxwell.

Abstract

In this thesis, we are interested in the construction of high-order finite elements adapted to hybrid meshes for the resolution of time-dependent and time-harmonic linear hyperbolic systems. We paid a special attention to the construction of pyramidal elements.

We search « optimal » finite elements for three different formulations, the optimality being in the sense of the convergence in the norm of the space considered for the formulation. For H^1 and $H(curl)$ formulations, optimal nodal and *hp* finite elements are constructed. The elementary matrices are evaluated with appropriate quadrature formula, and error estimates are performed to check the convergence of the constructed optimal elements. For the LDG (Local Discontinuous Galerkin) formulation, we present finite elements using orthogonal basis functions that allow us to design fast construction of the mass matrix and matrix-vector product. In the three cases, we present numerical properties of the elements, we check numerically that we get the theoretical convergence, and we compare our elements with other elements found in the literature.

Finally, numerical experiments in 3D are conducted with time-dependent and time-harmonic equations (wave or Helmholtz equation, and Maxwell's equations). We show the efficiency of hybrid meshes compared to pure tetrahedral meshes or hexahedral meshes obtained by splitting tetrahedra into four hexahedra.

Key words : conformal hybrid mesh, higher-order finite element, continuous and discontinuous Galerkin methods, nodal, hp and edge finite elements, quadrature formula, error estimates, Helmholtz and wave equations, Maxwell's equations.