

Higher-Order Finite Elements for Hybrid Meshes Using New Nodal Pyramidal Elements

Morgane Bergot, Gary Cohen, Marc Duruflé

▶ **To cite this version:**

Morgane Bergot, Gary Cohen, Marc Duruflé. Higher-Order Finite Elements for Hybrid Meshes Using New Nodal Pyramidal Elements. Journal of Scientific Computing, 2010, 42 (3), pp.345–381. 10.1007/s10915-009-9334-9 . hal-00454261

HAL Id: hal-00454261
https://hal.science/hal-00454261

Submitted on 8 Feb 2010

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

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

Journal of Scientific Computing, Vol.??, No.?, 200?, 1–35.

HIGHER-ORDER FINITE ELEMENTS FOR HYBRID MESHES USING NEW NODAL PYRAMIDAL ELEMENTS*

Morgane Bergot
Projet POems, INRIA Rocquencourt, Le Chesnay, France
Email: morgane.bergot@inria.fr
Gary Cohen
Projet POems, INRIA Rocquencourt, Le Chesnay, France
Email: gary.cohen@inria.fr
Marc Duruflé
Institut Mathématique de Bordeaux, Université Bordeaux I, Bordeaux, France
Email: marc.durufle@math.u-bordeaux1.fr

Abstract

We provide a comprehensive study of arbitrarily high-order finite elements defined on pyramids. We propose a new family of high-order nodal pyramidal finite element which can be used in hybrid meshes which include hexahedra, tetrahedra, wedges and pyramids. Finite elements matrices can be evaluated through approximate integration, and we show that the order of convergence of the method is conserved. Numerical results demonstrate the efficiency of hybrid meshes compared to pure tetrahedral meshes or hexahedral meshes obtained by splitting tetrahedra into hexahedra.

Key words: pyramidal element, higher-order finite element, hybrid mesh, conformal mesh, continuous finite element, discontinuous Galerkin method, error estimates, quadrature formula.

Introduction

Highly efficient finite element methods using hexahedral meshes have been developed by Cohen [9] and his collaborators (Fauqueux [10], Pernet and Ferrières [11], [25], Duruflé [13], [14]) but currently the only systematic way to generate unstructured hexahedral meshes for a complex geometry is to generate a tetrahedral mesh, and split each tetrahedron into four hexahedra, which introduce needlessly substantial increase in the cost. However, some mesh generators are able to produce hexahedral-dominant meshes that include a minor number of tetrahedra, wedges and pyramids. The aim here is to study finite element methods on hybrid meshes in order to preserve the efficiency of the method developed for hexahedra.

Nodal finite elements are detailed in Hesthaven and Teng [20] for tetrahedra, and Cohen [9] for hexahedra. Wedge (or triangular prism) nodal finite elements are constructed as a tensor product between Legendre-Gauss-Lobatto (LGL) points on $[0,1]$ and electrostatic points on the triangle including LGL points on the edges [20]. In this work, the main effort is devoted to the construction of pyramidal finite elements, preserving conformity with the other types of elements.

Since obtaining a proper base for nodal pyramidal elements is a tricky point, two approaches have been attempted. A first approach consists in using rational functions in order to obtain nodal shape functions.

* Received ???? ??, 200? / Revised version received ???? ??, 200? /

- First works about nodal pyramidal elements have been made by Bedrosian in [2] where he noticed the impossibility of choosing polynomial shape functions if we want to preserve the conformity with other elements. As a solution, he proposes what he calls "rabbit-functions" for first-order and second-order approximations. But, the second-order approximation does not include a node at the center of the quadrilateral base, which prohibits the conformity with the second-order hexahedron.

- Zgainski *et al.* [33] perform numerical experiments with the basis functions given by Bedrosian, and propose a modified second-order set of shape functions by adding a node at the center of the quadrilateral base. However, the central basis function proposed does not satisfy the nodal condition $\varphi_i(M_j) = \delta_{ij}$, and the modification does not improve the accuracy, since the finite element space generated by this set of basis functions does not contain $\mathbb{P}_2$. The same idea is taken back by Graglia *et al.* [17] who achieve to improve the accuracy with their own second-order central basis function.

- Chatzi and Preparata [6] introduce a generalization of Bedrosian basis functions at any order for nodes regularly distributed on the pyramid. Unfortunately, these basis functions are not consistent for order greater or equal to three since polynomials are not generated by these functions.

The second approach is to split the pyramid into tetrahedra to avoid the use of rational fractions, which have the debatable reputation to make the basis functions hard to manipulate, and instead use polynomial basis functions.

- Wieners [31], Knabner and Summ [22], and Bluck and Walker [3] provide a consistent first-order set of shape functions which ensures the conformity with tetrahedra and hexahedra, by splitting a pyramid into two tetrahedra. Second-order shape functions have been proposed by Wieners, and high order shape functions by Bluck and Walker. However, the finite element space of higher order does not contain the low order finite element space, which leads to a non-consistent method in the case of non-affine pyramids. Moreover, this method requires expensive quadrature on each tetrahedron.

- Liu *et al.* [23] propose to symmetrize shape functions of Wieners, but this modification barely improves the accuracy of the method.

An other popular alternative for finite element is the *hp* approach (Szabó and Babuška [29]), e.g. with Šolín *et al.* [28] for hexahedra, tetrahedra and wedges. Several papers extend the *hp* finite element to pyramidal elements.

- Warburton [30], Sherwin [26], Sherwin *et al.* [27], and Karniadakis and Sherwin [21] provide a tensorial set of basis functions for all types of elements based on the degeneration of a cube. For tetrahedra, hexahedra and wedges, the generated finite element spaces are standards. For pyramids, the proposed generated finite element space provides an optimal convergence for affine pyramids, but not for distorted pyramids for an order greater or equal to two. Moreover, the continuous transition between pyramids and tetrahedra is not achievable for general unstructured meshes.

- Nigam and Phillips [24] propose an original finite element space by deriving pyramidal finite elements from a reference element that is the infinite pyramid. With the finite

element space they obtain, the accuracy is preserved but the dimension of this space could be reduced.

- Demkowicz *et al.* [12] and Zaglmayr [32] give the construction of partial-orthogonal basis functions for tetrahedra, hexahedra and wedges, and exploit the use of a degenerated cube for pyramidal elements to get a finite element space that preserve the optimal accuracy, with a smaller dimension than Nigam and Phillips.

In this paper, the reference element is the symmetric unit pyramid (Fig. 1.1). Let $\mathbb{P}_r$ be the polynomial space of degree r, we claim that if we choose the following finite element space

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

we are able to produce optimal error estimates in H^1 norm

$$\|u-\pi_r u\|_{1,K} \le C h^r \|u\|_{r+1,K}$$

with the notations detailed in Section 4, for continuous finite elements.

In order to evaluate integrals, we propose in Section 3 to use the same technique as Bedrosian, detailed by Hammer, Marlowe and Stroud in [18], adapted to the pyramid and which does not deteriorate the accuracy, as it will be proved in Section 4. An extension of this work is proposed for discontinuous Galerkin formulation with the same finite element space $\hat{P}_r$.

To validate this new pyramidal finite element, a dispersion analysis is carried out in the case of periodic meshes. We have observed an optimal dispersion error in $O(h^{2r})$ as obtained for other element shapes. Furthermore these elements have been tested for the Helmholtz equation with the continuous Galerkin formulation, and for the unsteady wave equation with the discontinuous Galerkin method. The numerical experiments show that they are much more efficient than purely tetrahedral elements, or hexahedral meshes generated by splitting each tetrahedron into four hexahedra.

The outline of our paper is as follows:

- In Section 1, following the classical notations of Ciarlet [7], we define two pyramidal finite elements of order r, $(\hat{K},\hat{P}_r,\hat{\Sigma})$ on the reference element, and (K,P_r,Σ) for any pyramid in the mesh ;
- A comparison to existing *hp* finite element spaces is given in Section 2, along with possible improvements of these spaces, in propositions 2.2 and 2.3 ;
- The quadrature formula used to get exact integrals, whenever it is possible, for the basis functions constructed from the finite element space $\hat{P}_r$ are presented in Section 3, thanks to a change of variable from the unit cube ;
- Section 4 is devoted to the error analysis which is performed in a classical way ;
- The case of a discontinuous Galerkin formulation is briefly treated in Section 5 ;
- Section 6 is devoted to numerical results: a dispersion analysis is performed on the wave equation in section 6.1, the stability condition (CFL) is computed on a periodic infinite mesh in section 6.2, and numerical experiments are performed in section 6.4 along with explanations about storage.

1. Arbitrary High-Order Pyramidal Element

1.1. Pyramidal Element

Definition 1.1 *A pyramid $K(x,y,z)$ is the image of the reference pyramid $\hat{K}(\hat{x},\hat{y},\hat{z})$ taken as the unit symmetrical pyramid, centered at the origin by the transformation F given by Bedrosian [2] using rational fractions, as shown in Fig. 1.1*

$$F = \sum_{1\le i\le 5} S_i\, \hat{\varphi}_i^1, \tag{1}$$

where $S_i = (x_i, y_i, z_i)$ are the vertices of the pyramid K and $\hat{\varphi}_i^1$ are the following mapping functions

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

when F is invertible.

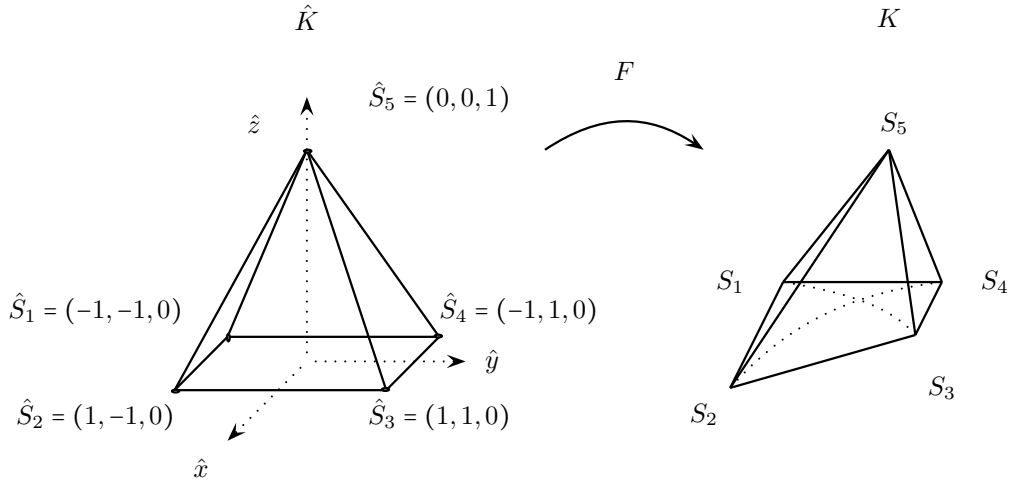


Fig. 1.1. Transformation of the reference pyramid $\hat{K}$ to the pyramid K via the transformation F

Remark 1.2 *The mapping functions $\hat{\varphi}_i^1$ are denoted with an upper index 1 as they correspond to the basis functions of order 1.*

The case of a non-invertible transformation may occur when considering a degenerated element, e.g. when the five vertices are co-planar, but the characterization of pyramids for which F is invertible remains an open question, as for hexahedra (Duruflé *et al.* [14]). In the sequel, we assume that F is always invertible.

The transformation F can be explicitly written as

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

We notice that F is affine when

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

i.e. when the base of the pyramid is a parallelogram. Furthermore, F ensures the conformity with tetrahedra and hexahedra as the shape functions becomes a two-dimensional triangular or quadrilateral shape function, since adjacent tetrahedra, wedge and hexahedra have the same property. That would not be the case if F had been chosen to be polynomial.

Remark 1.3 *The shape function of Bedrosian can be found by defining the transformation T from the unit cube $\widetilde{Q}$ to the reference pyramid $\hat{K}$*

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

For a basis function of the hexahedron $\varphi(\widetilde{x},\widetilde{y},\widetilde{z}) = (1-\widetilde{x})(1-\widetilde{y})(1-\widetilde{z})$, the transformation T gives indeed

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

Similarly, we find the other functions of Bedrosian.

1.2. A Pyramidal Finite Element Space of Order r

We place ourselves in the most restrictive case, that is continuous finite elements. The finite element space V_h on an open set Ω of $\mathbb{R}^3$ is given by

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

where P_r^F is the real space of order r for an element K of the mesh defined by

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

The finite element space $\hat{P}_r$ of order r on $\hat{K}$ is

- $\mathbb{P}_r(\hat{x},\hat{y},\hat{z}) = \left\{\hat{x}^i\hat{y}^j\hat{z}^k,\ i+j+k \le r\right\}$ when $\hat{K}$ is a tetrahedron ;
- $\mathbb{Q}_r(\hat{x},\hat{y},\hat{z}) = \left\{\hat{x}^i\hat{y}^j\hat{z}^k,\ i,j,k \le r\right\}$ when $\hat{K}$ is a hexahedron ;
- $\mathbb{P}_r(\hat{x},\hat{y}) \otimes \mathbb{P}_r(\hat{z}) = \left\{\hat{x}^i\hat{y}^j\hat{z}^k,\ i+j \le r, k \le r\right\}$ when $\hat{K}$ is a wedge ;
- and defined by identity (5) when $\hat{K}$ is a pyramid.

To use the Bramble-Hilbert's lemma and get optimal error estimates, the real space P_r^F for a pyramidal element K of the mesh must be such that

$$\mathbb{P}_r(x,y,z) \subset P_r^F. \tag{3}$$

Theorem 1.4 *When F is affine, the minimal space $\hat{P}_r$ of order r such that we have the inclusion (3) is*

$$\boxed{\hat{P}_r = \mathbb{P}_r(\hat{x},\hat{y},\hat{z}).} \tag{4}$$

When F is not affine, the minimal space $\hat{P}_r$ of order r such that we have the inclusion (3) is

$$\boxed{\hat{P}_r = \mathbb{P}_r(\hat{x},\hat{y},\hat{z}) \oplus \sum_{0\le k\le r-1} \left(\frac{\hat{x}\hat{y}}{1-\hat{z}}\right)^{r-k} \mathbb{P}_k(\hat{x},\hat{y}).} \tag{5}$$

Proof. When $F \in \mathbb{P}_1$, it is easy to see that

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

which means that taking $\hat{P}_r = \mathbb{P}_r$ when the base of the pyramid is a parallelogram is necessary and sufficient to satisfy (3).

For any base of the pyramid, we take $f \in \mathbb{P}_r$, i.e.

$$f = \sum_{\substack{0\le i,j,k\le r,\\ i+j+k\le r}} x^i\, y^j\, z^k.$$

We study the case $f = x^n$, $n \le r$. Using the transformation F, f can be written as

$$\begin{aligned} 4^n f = 4^n x^n \quad = \quad & \Big[x_1\,(1-\hat{x}-\hat{y}-\hat{z}) + x_2\,(1+\hat{x}-\hat{y}-\hat{z}) + x_3\,(1+\hat{x}+\hat{y}-\hat{z}) + x_4\,(1-\hat{x}+\hat{y}-\hat{z}) \\ & + 4\,\hat{z}x_5 + \frac{\hat{x}\hat{y}}{1-\hat{z}}(x_1+x_3-x_2-x_4)\Big]^n. \end{aligned}$$

As the part

$$\Big[\, x_1\,(1-\hat{x}-\hat{y}-\hat{z}) + x_2\,(1+\hat{x}-\hat{y}-\hat{z}) + x_3\,(1+\hat{x}+\hat{y}-\hat{z}) + x_4\,(1-\hat{x}+\hat{y}-\hat{z}) + 4\,\hat{z}x_5 \,\Big]^n$$

is in $\mathbb{P}_n(\hat{x},\hat{y},\hat{z})$, it remains to handle the terms

$$(a+b\hat{x}+c\hat{y}+d\hat{z})^k\,(\frac{\hat{x}\hat{y}}{1-\hat{z}})^{n-k} \quad k \le n-1.$$

Developing the first factor, we get terms of the form

$$\hat{z}^p\,(\alpha+\beta\hat{x}+\gamma\hat{y})^{k-p}\,(\frac{\hat{x}\hat{y}}{1-\hat{z}})^{n-k}.$$

If $p = 0$, the factor belongs to $\left(\frac{\hat{x}\hat{y}}{1-\hat{z}}\right)^{n-k} \mathbb{P}_k(\hat{x},\hat{y})$. Otherwise, we decrease the power of $\hat{z}$, by writing $\hat{z} = 1-\hat{z}+\hat{z}$

$$\hat{z}^{p-1}(\alpha+\beta\hat{x}+\gamma\hat{y})^{k-p}\,(\frac{\hat{x}\hat{y}}{1-\hat{z}})^{n-k} \;+\; \hat{z}^{p-1}(\alpha+\beta\hat{x}+\gamma\hat{y})^{k-p}\hat{x}\,\hat{y}\,(\frac{\hat{x}\hat{y}}{1-\hat{z}})^{n-k-1}.$$

Iterating this method, we erase all the powers of $\hat{z}$ to obtain a term of higher degree

$$(\alpha+\beta\hat{x}+\gamma\hat{y})^{k-p}(\hat{x}\,\hat{y})^p\,(\frac{\hat{x}\hat{y}}{1-\hat{z}})^{n-k-p}.$$

However, when $k+p \geq n$, the iterative procedure stops as we obtained the polynomial

$$\hat{z}^{p+k-n}(\alpha+\beta\hat{x}+\gamma\hat{y})^{k-p}(xy)^{n-k}, \tag{6}$$

and the degree of this polynomial is equal to $k \leq r-1$. Since $k+p \leq r-1$,

$$(\alpha+\beta\hat{x}+\gamma\hat{y})^{k-p}(\hat{x}\,\hat{y})^{p} \in \mathbb{P}_{k+p}(\hat{x},\hat{y}),$$

and the term is finally in

$$\mathbb{P}_m(\hat{x},\hat{y})\,(\frac{\hat{x}\hat{y}}{1-\hat{z}})^{n-m}$$

with $m = k+p$, $m \leq n-1$.

We let the reader convince himself that other cases can be treated similarly.

At this point, we proved that it is sufficient to take $\hat{P}_r$ as specified by Theorem 1.4 to obtain the inclusion (3). □

Corollary 1.5

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

Proof. We classically have

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

and, using the direct sums property,

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

that is

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

which provides the claimed result. □

Proposition 1.6

$$\begin{array}{rcl} \hat{P}_{r|\ \hat{x}=1-\hat{z}\ \ or\ \ \hat{x}=\hat{z}-1} & = & \mathbb{P}_r(\hat{y},\hat{z}). \\ \hat{P}_{r|\ \hat{y}=1-\hat{z}\ \ or\ \ \hat{y}=\hat{z}-1} & = & \mathbb{P}_r(\hat{x},\hat{z}). \\ \hat{P}_{r|\ \hat{z}=0} & = & \mathbb{Q}_r(\hat{x},\hat{y}). \end{array} \tag{7}$$

Proof. Any function $p \in \hat{P}_r$ can be written as

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

with $p_r \in \mathbb{P}_r(\hat{x},\hat{y},\hat{z})$ and $p_k \in \mathbb{P}_k(\hat{x},\hat{y})$.

On a triangular face, we replace $\hat{x}$ by $\pm(1-\hat{z})$ or $\hat{y}$ by $\pm(1-\hat{z})$, according to the considered face. For example for the face $\hat{x} = (1-\hat{z})$, as $p_r((1-\hat{z}),\hat{y},\hat{z})$ obviously belongs to $\mathbb{P}_r(\hat{y},\hat{z})$, rational parts become

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

As $p_k((1-\hat{z}),\hat{y}) \in \mathbb{P}_k(\hat{y},\hat{z})$, we have $p_k((1-\hat{z}),\hat{y})y^{r-k} \in \mathbb{P}_r(\hat{y},\hat{z})$, and finally $p \in \mathbb{P}_r(\hat{y},\hat{z})$. The same simplification can be done for the other faces.

On the quadrangular base, we replace $\hat{z}$ by 0: $p_r(\hat{x},\hat{y},0)$ is obviously in $\mathbb{Q}_r(\hat{x},\hat{y})$, and the rational parts become

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

As $p_k(\hat{x},\hat{y}) \in \mathbb{P}_k(\hat{x},\hat{y})$, we have $p_k(\hat{x},\hat{y})x^{r-k}y^{r-k} \in \mathbb{Q}_r(\hat{x},\hat{y})$, and finally $p \in \mathbb{Q}_r(\hat{x},\hat{y})$.

The proposition is finally proved using a dimension argument. □

Proposition 1.7

$$P_r^F(K) \subset H^1(K).$$

Proof. For $p \in P_r^F$, $p \in C^\infty(\bar{K}\backslash S_5)$ as a the rational fraction when its pole is not in the domain. The continuity in S_5 is proved by considering four pseudo-faces F_ε^i, $0 \le i \le 4$, $0 \le \varepsilon \le 1$ mapping a quarter Q_i of the pyramid. We consider the face F_ε^2 represented in Fig. 1.2 in red, Q_2 being represented in blue. F_ε^2 is such that

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

and we have

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

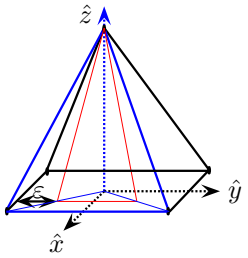


Fig. 1.2. Pseudo face F_ε^2

All cases amount to study $p = \dfrac{xy}{1-z}$ which is the difficult case.

$$p_{|F_\varepsilon^2} = \frac{(1-z)(1-\varepsilon)y}{1-z} = (1-\varepsilon)y \xrightarrow{z\to 1} 0.$$

The three other cases are similar by symmetry, and finally $p \in C^0(\bar{K})$.

As K is bounded, we get $p \in L^2(K)$. As for p, we consider ∇p on a quarter of pyramid, for example Q_2 and we consider an ε such that $M \in F_\varepsilon^2$, and

$$-(1-\varepsilon)^2 \le \frac{\partial p}{\partial z}_{|F_\varepsilon^2} \le (1-\varepsilon)^2,$$

that is $\partial_z p$ is bounded. The same technique applied for $\partial_x p$ and $\partial_y p$ leads to conclude that ∇p in bounded inside K. As K is also bounded, we have $\nabla p \in L^2(K)$. □

Proposition 1.6 ensures to a function $u \in V_h$ to be continuous across the interface between elements, whatever the type of the elements adjacent to the face, and therefore to belong to $H^1(\Omega)$ due to proposition 1.7.

Proposition 1.8 *The optimal finite element space of order r on the unit cube $\widetilde{Q}$ is*

$$\boxed{C_r = \hat{P}_r \circ T = \sum_{0 \le k \le r} \mathbb{Q}_k(\widetilde{x}, \widetilde{y})(1-\widetilde{z})^k.}$$

Proof. Using the transformation T, the polynomial part of $\hat{P}_r$ becomes

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

whereas the fractional part of $\hat{P}_r$ becomes

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

that is $\widetilde{P}_r \subset C_r$.

We now notice that

$$\dim\, C_r = \sum_{0 \le k \le r} (k+1)^2 = \frac{1}{6}(r+1)(r+2)(2r+3) = \dim\, \hat{P}_r = \dim\, \widetilde{P}_r,$$

which proves the proposition. □

1.3. Location of the Degrees of Freedom

We wish to link continuously pyramidal elements with other elements of the mesh

- hexahedra with Legendre-Gauss-Lobatto (LGL) points ;
- tetrahedra with Hesthaven "electrostatic points" constructed with LGL points on edges (Hesthaven and Teng[20]);
- wedges obtained by a tensorial product of a face of a tetrahedron of Hesthaven, that is a triangle of Hesthaven (Hesthaven [19]), with an edge with LGL points.

We place the degrees of freedom on LGL points on the quadrangular base of the pyramid, and on Hesthaven points on each triangular face. The number of degrees of freedom n_f on the faces is then

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

We add n_i degrees of freedom inside the pyramid

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

and we place them on $(r-2)$ parallel planes of k^2 degrees of freedom, as shown in the Fig. 1.3.

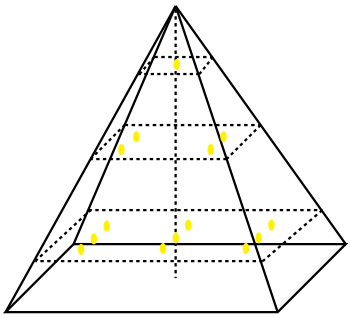


Fig. 1.3. Location of the degrees of freedom inside the pyramidal element of order 5

The total number of degrees of freedom is

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

which is precisely the dimension of $\hat{P}_r$.

Degrees of freedom can then be placed systematically on the pyramid, at any order. Each category of point is represented by a color in the Fig. 1.4 for the pyramidal elements of order two to four.

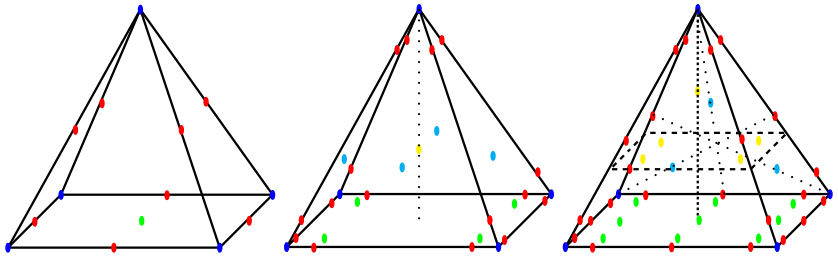


Fig. 1.4. Location of the degrees of freedom for the pyramidal elements of order 2, 3 and 4 (Color online)

1.4. Basis Functions

The basis functions on the reference pyramid $\hat{K}$ are obtained by inverting a Vandermonde system as follow.

Let $(\hat{M}_i)_{1\le i\le n_r}$ the locations of the interpolation points on the pyramid, and $(\hat{\psi}_i)_{1\le i\le n_r}$ a base of $\hat{P}_r$,

Definition 1.9 *The Vandermonde matrix* $VDM \in \mathcal{M}_{n_r}(\mathbb{R})$ *is defined by*

$$VDM_{i,j} = \hat{\psi}_i(\hat{M}_j), \quad 1 \le i, j \le n_r, \tag{8}$$

and the basis function $\hat{\varphi}_i$ *linked to the interpolation point* $\hat{M}_i$ *is then defined as*

$$\hat{\varphi}_i = \sum_{1\le j\le n_r} (VDM^{-1})_{i,j}\, \hat{\psi}_j. \tag{9}$$

Proposition 1.10 *The following set of basis functions is an orthogonal base of* $\hat{P}_r$

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

where $P_m^{i,j}(x)$ *denotes the Jacobi polynomial of order* m*, orthogonal for the weight* $(1-x)^i(1+x)^j$*.*

Proof. We note

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

We first prove that the family is orthogonal by using the transformation (2) on $\tilde{Q}_r$

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

with $0\le i,j\le r$, $0\le k\le r-max(i,j)$, and when $i=i'$ and $j=j'$

$$\int_0^1 (1-\widetilde{z})^{2max(i,j)+2} P_k^{2max(i,j)+2,0}(2\widetilde{z}-1)P_{k'}^{2max(i,j)+2,0}(2\widetilde{z}-1)\,d\widetilde{z} = C_{kk'}\,\delta_{k\,k'}.$$

We also have

$$\begin{aligned}&\left\{\hat{\psi}_{i,j,k}(\hat{x},\hat{y},\hat{z}),\ 0\le i,j\le r,\ k\le r-max(i,j)\right\}\circ T^{-1} = \\ &\left\{P_i^{0,0}(2\widetilde{x}-1)\,P_j^{0,0}(2\widetilde{y}-1)\,(1-\widetilde{z})^{max(i,j)} P_k^{2max(i,j)+2,0}(2\widetilde{z}-1),\ 0\le i,j\le r,\ k\le r-max(i,j)\right\}\subset C_r,\end{aligned}$$

that is, with an argument of dimension,

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

which proves the proposition. □

We compare the condition number of the Vandermonde matrix for monomial and orthogonal bases of $\hat{P}_r$ in the case of tetrahedral and pyramidal elements on Fig. 1.5. We notice that the condition number of the Vandermonde matrix is increasing faster for tetrahedra than for pyramids when using monomial base, whereas we observe the opposite for orthogonal base. Besides, the use of the orthogonal set of basis functions highly improves the condition number of the VDM matrix.

Remark 1.11 *The characterization of the invertibility of the Vandermonde matrix is an open question, but we observed that the VDM matrix is invertible with our choice of position for the degrees of freedom, the element is therefore unisolvent.*

2. Comparison with Existing Methods

The nodal basis functions we propose are the same as Bedrosian [2], Zgainski *et al.* [33] and Chatzi and Preparata [6] for order one. They are identical to those of Graglia *et al.* [17] for order two, and they are new for order greater or equal to three.

The finite element space C_r of order r on the unit cube $\widetilde{Q}$ defined by in proposition 1.8 is the same as the one proposed by Zaglmayr, cited in [12],

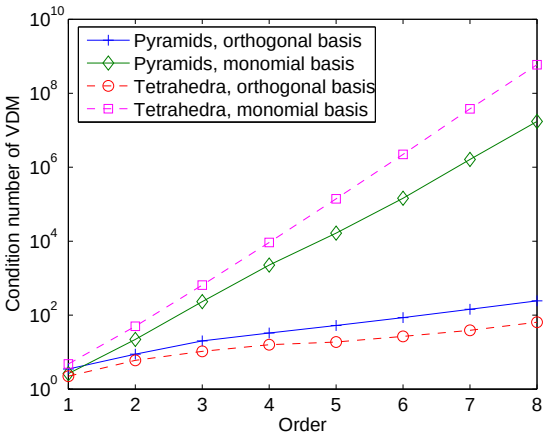


Fig. 1.5. Condition number of the VDM matrix versus the order for tetrahedral and pyramidal elements, for monomial and orthogonal basis functions

Proposition 2.1 *The subspace C_r^0 of C_r with zero trace on the boundary of $\widetilde{Q}$ is*

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

Proof. The basis functions obviously vanish on the boundary of $\widetilde{Q}$ and belongs to C_r. The dimension of the space is $\dim C_{r-3} = \dfrac{1}{6}(r-1)(r-2)(2r-3) = n_i$, which proves the proposition. □

We write transformation $\bar{T}$ from the infinite pyramid $\bar{Q}$ to the unit cube $\widetilde{Q}$.

$$\bar{T}: \begin{cases} \widetilde{x} = \bar{x} \\ \widetilde{y} = \bar{y} \\ \widetilde{z} = \dfrac{\bar{z}}{1+\bar{z}}. \end{cases} \tag{10}$$

Proposition 2.2 *The finite element space U_r proposed by Nigam and Phillips [24] on the infinite pyramid $\bar{Q}$ satisfies*

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

and contains more degrees of freedom than C_r since

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

The subspace U_r^0 of U_r whose trace is null on the boundary of the element is equal to

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

and if we replace U_0^r by $C_r^0 \circ \bar{T}$, we get the optimal space

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

Proof. Using the transformation (10), we detail the following basis functions (the others can be treated similarly by symmetry)

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

For the apex : $\dfrac{\bar{z}^r}{(1+\bar{z})^k} \circ \bar{T}^{-1} = \widetilde{z}^r \in C_r.$

For the representative vertical edge :
$\left\{\dfrac{(1-\bar{x})(1-\bar{y})\bar{z}^a}{(1+\bar{z})^r},\ 1 \le a \le r-2\right\} \circ \bar{T}^{-1} = \{(1-\widetilde{x})(1-\widetilde{y})(1-\widetilde{z})^{r-a}\widetilde{z}^a,\ 1 \le a \le r-2\} \subset C_r.$

For the representative base edge :
$\left\{\dfrac{(1-\bar{x})(1-\bar{y})\bar{x}^a}{(1+\bar{z})^r},\ 1 \le a \le r-2\right\} \circ \bar{T}^{-1} = \{(1-\widetilde{x})(1-\widetilde{y})\widetilde{x}^a(1-\widetilde{z})^r,\ 1 \le a \le r-2\} \subset C_r.$

For the representative triangular face : $\left\{\dfrac{(1-\bar{x})(1-\bar{y})\bar{x}^a\bar{z}^b}{(1+\bar{z})^r},\ a,b \ge 0, a+b \le r-1\right\} \circ \bar{T}^{-1} =$
$\{(1-\widetilde{x})(1-\widetilde{y})\widetilde{x}^a(1-\widetilde{z})^{r-b}\widetilde{z}^b,\ 0 \le a+b \le r-1\} \subset C_r.$

For the base face :
$\left\{\dfrac{(1-\bar{x})(1-\bar{y})\bar{x}^a\bar{y}^b}{(1+\bar{z})^r},\ 1 \le a,b \le r-1\right\} \circ \bar{T}^{-1} = \{(1-\widetilde{x})(1-\widetilde{y})\widetilde{x}^a\widetilde{y}^b(1-\widetilde{z})^r,\ 1 \le a,b \le r-1\} \subset C_r.$

For the interior :
$\left\{\dfrac{\bar{x}(1-\bar{x})\bar{y}(1-\bar{y})\bar{z}}{(1+\bar{z})^r}\, u(\bar{x},\bar{y},\bar{z}),\ u \in \mathbb{Q}^{r-2}(\bar{x},\bar{y},\bar{z})\right\} \circ \bar{T}^{-1} =$
$\{\widetilde{x}^{i+1}(1-\widetilde{x})\widetilde{y}^{j+1}(1-\widetilde{y})\widetilde{z}^{k+1}(1-\widetilde{z})^{r-k},\ 0 \le i,j,k \le r-2\} \supset C_r^0.$

The subspace of U_r whose trace is null on the boundary of the element is equal to

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

whose dimension is

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

Since there are $n_f = 3r^2 + 2$ basis functions associated with the boundary, we have

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

If we replace U_r^0 by $C_r^0 \circ \bar{T}$, the new finite element space $\bar{U}_r$ satisfies

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

and

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

that is we have the equality of these two spaces. □

We write transformation $\widehat{T}$ from the cube $[-1,1]^3$ to the unit cube $\widetilde{Q}$

$$\widehat{T} : \begin{cases} \widetilde{x} = \dfrac{1+a}{2} \\ \widetilde{y} = \dfrac{1+b}{2} \\ \widetilde{z} = \dfrac{1+c}{2}. \end{cases} \tag{11}$$

Proposition 2.3 *The finite element space W_r of order r introduced by Warburton [30] on the cube $[-1,1]^3$ is not optimal.*

The subspace of W_r whose trace is null on the boundary of the element is equal to

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

If we replace W_0^r by $C_r^0 \circ \widehat{T}$, and the basis functions linked to the base face by the following set of functions

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

we get the optimal space

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

Proof. Using the transformation (11), we detail the following basis functions (the others can be treated similarly by symmetry)

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

For the apex : $\left\{ \frac{1+c}{2} \right\} \circ \widehat{T}^{-1} = \widetilde{z} \ \in C_r.$

For the vertical edge : $\left\{ \left(\frac{1-a}{2}\right)\left(\frac{1-b}{2}\right)\left(\frac{1-c}{2}\right)\left(\frac{1+c}{2}\right) P_{i-1}^{1,1}(c),\ 1 \le i \le r-1 \right\} \circ \widehat{T}^{-1} =$
$\left\{ (1-\widetilde{x})(1-\widetilde{y})(1-\widetilde{z})\widetilde{z}\, P_{i-1}^{1,1}(2\widetilde{z}-1),\ 1 \le i \le r-1 \right\} \subset C_r.$

For the base edge :
$\left\{ \left(\frac{1-a}{2}\right)\left(\frac{1+a}{2}\right)\left(\frac{1-b}{2}\right)\left(\frac{1-c}{2}\right)^{i+1} P_{i-1}^{1,1}(a),\ 1 \le i \le r-1 \right\} \circ \widehat{T}^{-1} =$
$\left\{ \widetilde{x}(1-\widetilde{x})(1-\widetilde{y})(1-\widetilde{z})^{i+1}\, P_{i-1}^{1,1}(2\widetilde{x}-1),\ 1 \le i \le r-1 \right\} \subset C_r.$

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

For the triangular face :
$\left\{ \left(\frac{1-a}{2}\right)\left(\frac{1+a}{2}\right)\left(\frac{1-b}{2}\right)\left(\frac{1-c}{2}\right)^{i+1}\left(\frac{1+c}{2}\right) P_{i-1}^{1,1}(a)P_{j-1}^{2i+1,1}(c),\ i+j \le r-1, i,j \ge 1 \right\} \circ \widehat{T}^{-1} =$
$\left\{ (1-\widetilde{x})\widetilde{x}(1-\widetilde{y})(1-\widetilde{z})^{i+1}\widetilde{z}\, P_{i-1}^{1,1}(2\widetilde{x}-1)P_{j-1}^{2i+1,1}(2\widetilde{z}-1),\ i+j \le r-1, i,j \ge 1 \right\} \subset C_r.$

For the interior :
$\left\{ \left(\frac{1-a}{2}\right)\left(\frac{1+a}{2}\right)\left(\frac{1-b}{2}\right)\left(\frac{1+b}{2}\right)\left(\frac{1-c}{2}\right)^{i+j+1}\left(\frac{1+c}{2}\right) P_{i-1}^{1,1}(a)P_{j-1}^{1,1}(b)P_{k-1}^{2i+2j+1,1}(c),\ i+j+k \le r-1,\ i,j,k \ge 1 \right\} \circ$
$\widehat{T}^{-1} = \left\{ \widetilde{x}(1-\widetilde{x})\widetilde{y}(1-\widetilde{y})\widetilde{z}(1-\widetilde{z})^{i+j+1}\, P_{i-1}^{1,1}(2\widetilde{x}-1)P_{j-1}^{1,1}(2\widetilde{y}-1)P_{k-1}^{2i+2j+1,1}(2\widetilde{z}-1),\ i+j+k \le r-1,\ i,j,k \ge 1 \right\} \subset$
$C_r.$

The subspace W_r^0 of W_r whose trace is null on the boundary of the element is equal to

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

whose dimension is

$$\dim\, W_r^0 = \dim\, \mathbb{P}_{r-3} = \frac{(r-2)(r-1)r}{6}.$$

Since there are $3r^2+2$ basis functions associated with the boundary, we have

$$\dim\, W_r = \frac{(r-2)(r-1)r}{6} + 3r^2 + 2 = \frac{(r+1)(r+2)(r+3)}{6} + r^2 < \dim\, C_r.$$

If we replace the proposed set of basis functions for the base face by the following one

$$\left\{\left(\frac{1-a}{2}\right)\left(\frac{1+a}{2}\right)\left(\frac{1-b}{2}\right)\left(\frac{1+b}{2}\right)\left(\frac{1-c}{2}\right)^{max(i,j)+1} P_{i-1}^{1,1}(a)P_{j-1}^{1,1}(b),\ 1 \le i,j \le r-1\right\} \circ \widehat{T}^{-1} =$$
$$\left\{\widetilde{x}(1-\widetilde{x})\widetilde{y}(1-\widetilde{y})(1-\widetilde{z})^{max(i,j)+1}\, P_{i-1}^{1,1}(2\widetilde{x}-1)P_{j-1}^{1,1}(2\widetilde{y}-1)\ 1 \le i,j \le r-1\right\} \subset C_r,$$

and W_r^0 by $C_r^0 \circ \widehat{T}$, the new finite element space $\widehat{W}_r$ will satisfy

$$\widehat{W}_r \subset C_r \circ \widehat{T}$$

and

$$\dim\, \widehat{W}_r = \dim\, C_r,$$

that is we have the equality of the two spaces. □

Remark 2.4 *As $W_1 = \widehat{W}_1$ but $W_r \not\supset \widehat{W}_2$, using W_r as a finite element space for pyramidal elements ensures not more than a first-order convergence in H^1-norm.*

Numerical study of the dispersion error has been conducted on periodic meshes containing non-affine pyramids in order to check these theoretical results (see Fig. 6.6 in section 6.3).

3. Quadrature Formula

To evaluate integrals, we use a quadrature rule defined over the reference pyramid $\hat{K}$. A simple rule consists in taking Gauss points over the unit cube $\widetilde{Q}$ of coordinates $(\widetilde{x},\widetilde{y},\widetilde{z})$, and compute their image on the reference pyramid $\hat{K}$ of coordinates $(\hat{x},\hat{y},\hat{z})$, via the change of variable T defined by equation (2), which is a diffeomorphism from the open $\widetilde{Q}$ to the open $\hat{K}$.

For any function f, we denote

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

and the change of variable provides

$$\int_{\hat{K}} \hat{f}(\hat{x},\hat{y},\hat{z})\, d\hat{x}d\hat{y}d\hat{z} = \int_{\widetilde{Q}} 4\, \widetilde{f}(\widetilde{x},\widetilde{y},\widetilde{z})\, (1-\widetilde{z})^2\, d\widetilde{x}d\widetilde{y}d\widetilde{z}. \tag{12}$$

Definition 3.1 *Let M be the mass matrix for the pyramid K, defined by*

$$M_{i,j} = \int_K \varphi_i\varphi_j\, dxdydz = \int_{\hat{K}} |DF|\, \hat{\varphi}_i\hat{\varphi}_j\, d\hat{x}d\hat{y}d\hat{z} \tag{13}$$

and K the stiffness matrix such that

$$K_{i,j} = \int_K \nabla\varphi_i \cdot \nabla\varphi_j\, dxdydz = \int_{\hat{K}} |DF| DF^{-1} DF^{*-1} \hat{\nabla}\hat{\varphi}_i \cdot \hat{\nabla}\hat{\varphi}_j\, d\hat{x}d\hat{y}d\hat{z}. \tag{14}$$

Definition 3.2 *We define the polynomial space*

$$\mathbb{Q}_{m,n,p} = \left\{ x^i y^j z^k, 0 \le i \le m, 0 \le j \le n, 0 \le k \le p \right\}.$$

Lemma 3.3

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

Proof. Using proposition 1.8, $\widetilde{\varphi}_i(\widetilde{x}, \widetilde{y}, \widetilde{z}) \in C_r(\widetilde{x}, \widetilde{y}, \widetilde{z})$, and we obviously have $C_r \subset \mathbb{Q}_r$, which proves the lemma. □

Lemma 3.4

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

Proof. We decompose $\hat{\varphi}_i(\hat{x}, \hat{y}, \hat{z})$ into the monomial base $\hat{\psi}_j(\hat{x}, \hat{y}, \hat{z})$ of $\hat{P}_r$ and we treat the different cases.

We first consider the derivative in x, the derivative in y being treated similarly by symmetry: either $\hat{\psi}_j(\hat{x}, \hat{y}, \hat{z}) \in \mathbb{P}_r(\hat{x}, \hat{y}, \hat{z})$,

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

or

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

that is $\dfrac{\partial \widetilde{\psi}_j}{\partial \hat{x}}(\widetilde{x}, \widetilde{y}, \widetilde{z}) \in \mathbb{Q}_{r-1,r,r-1}(\widetilde{x}, \widetilde{y}, \widetilde{z})$ in both cases.

Similarly, for the derivative in z, either

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

or

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

that is $\dfrac{\partial \widetilde{\psi}_j}{\partial \hat{z}}(\widetilde{x}, \widetilde{y}, \widetilde{z}) \in \mathbb{Q}_{r,r,r-1}(\widetilde{x}, \widetilde{y}, \widetilde{z})$ in both cases. □

Lemma 3.5

$$\widetilde{DF} = \left(\frac{\partial \widetilde{F}}{\partial \hat{x}}, \frac{\partial \widetilde{F}}{\partial \hat{y}}, \frac{\partial \widetilde{F}}{\partial \hat{z}} \right) \in \mathbb{Q}^3_{0,1,0} \times \mathbb{Q}^3_{1,0,0} \times \mathbb{Q}^3_{1,1,0}\,(\widetilde{x}, \widetilde{y}, \widetilde{z}),$$

and

$$\widetilde{|DF|} \in \mathbb{Q}_{1,1,0}(\widetilde{x}, \widetilde{y}, \widetilde{z}).$$

Proof. The derivatives of F can be written as

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

The determinant is then

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

We can develop the expression in

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

which proves the lemma. □

Proposition 3.6 *When F is affine, the quadrature formula must be exact for polynomials of $(1-z)^2\mathbb{Q}_{2r}$ for the mass matrix and $(1-z)^2\mathbb{Q}_{2r,2r,2r-2}$ for the stiffness matrix, such that these matrices are exactly integrated.*

When F is not affine, for the mass matrix to be exactly integrated, the quadrature formula must be exact for polynomials of $(1-z)^2\mathbb{Q}_{2r+1,2r+1,2r}$.

Proof. In the affine case, we can factorize the mass matrix by coefficient $|DF|$ which is constant. For the mass matrix, the lemma 3.3 provides the term $\widetilde{\varphi}_i\widetilde{\varphi}_j(1-\widetilde{z})^2$ to be in $(1-\widetilde{z})^2\mathbb{Q}_{2r}$. For the stiffness matrix, thanks to lemma 3.4, the term $\frac{\partial\widetilde{\varphi}_i}{\partial\widetilde{x}}\frac{\partial\widetilde{\varphi}_j}{\partial\widetilde{x}}(1-\widetilde{z})^2$ is in $(1-\widetilde{z})^2\mathbb{Q}_{2r-2,2r,2r-2}$, the term $\frac{\partial\widetilde{\varphi}_i}{\partial\widetilde{y}}\frac{\partial\widetilde{\varphi}_j}{\partial\widetilde{y}}(1-\widetilde{z})^2$ is in $(1-\widetilde{z})^2\mathbb{Q}_{2r,2r-2,2r-2}$ and the term $\frac{\partial\widetilde{\varphi}_i}{\partial\widetilde{z}}\frac{\partial\widetilde{\varphi}_j}{\partial\widetilde{z}}(1-\widetilde{z})^2$ is in $(1-\widetilde{z})^2\mathbb{Q}_{2r,2r,2r-2}$, that is the final term is in $(1-\widetilde{z})^2\mathbb{Q}_{2r,2r,2r-2}$.

In the non-affine case, using the result of the affine case and lemma 3.5, we can easily conclude that we need a quadrature rule exact for polynomials of $(1-\widetilde{z})^2\mathbb{Q}_{2r+1,2r+1,2r}$ to exactly integrate the mass matrix. □

Remark 3.7 *Because of the rational fraction due to $\widetilde{DF}^{-1}$, the stiffness matrix can not be exactly integrated with a classical quadrature formula.*

To integrate exactly the mass matrix, we choose the following quadrature formula :

$$(\xi_i^G,\xi_j^G,\xi_k^{HM}),(\omega_i^G\omega_j^G\omega_k^{HM}),$$

where (ξ_i^G, ω_i^G) is the classical Gauss quadrature rule, exact for polynomials of $\mathbb{Q}_{2r+1}$, and $(\omega_k^{HM}, \xi_k^{HM})$ a Gauss-Jacobi rule for the evaluation of $\int (1-x)^2 f(x)$, exact for polynomials of $(1-x)^2 Q_{2r+1}$. For this last rule, quadrature points and weights have been calculated in Hammer, Marlowe and Stroud [18]. Eventually, we need $(r+1)^3$ integration points for an exact integration of the mass matrix.

4. Error Estimates

4.1. Functional Spaces and Basic Notations

Let Ω be an open set of $\mathbb{R}^3$.

Definition 4.1 *We define*

$$L^2(\Omega) = \{u \in L^2_{loc} \mid \int_\Omega |u|^2 < +\infty\}$$
$$H^m(\Omega) = \{u \in L^2(\Omega) \mid \frac{\partial^\alpha u}{\partial x^\alpha} \in L^2(\Omega), |\alpha| \le m\}$$

equipped with the usual norm in $H^m(\Omega)$

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

and the usual semi-norm in $H^m(\Omega)$

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

We obviously have the inequality

$$\forall u \in H^m(\Omega), \ \ |u|^2_{m,\Omega} \le \|u\|^2_{m,\Omega}. \tag{15}$$

An approximate integral using a Gauss-type quadrature formula is denoted $\oint^G$, and π_r will denote a projector on the polynomial space $\mathbb{P}_r$.

4.2. Presentation of the Problem

We consider a standard variational problem

$$\begin{cases} \text{Find } u \in V \text{ such that} \\ \forall v \in V,\ a(u,v) = f(v), \end{cases} \tag{16}$$

where the space $V = H^1(\Omega)$, $a(.,.)$ a continuous bilinear and coercive form, and $f(.)$ a continuous linear form. Then, given a finite-dimensional subspace V_h of the space V, the discrete problem reads

$$\begin{cases} \text{Find } u_h \in V_h \text{ such that} \\ \forall v_h \in V_h,\ a_h(u_h, v_h) = f_h(v_h), \end{cases} \tag{17}$$

where $a_h(.,.)$ is a bilinear form defined over the space V_h, uniformly V_h-elliptic, and $f_h(.)$ is a linear form defined over the space V_h.

We shall consider the simple case where $a(u,v) = \int_\Omega uv + \int_\Omega \nabla u \cdot \nabla v$.

4.3. Abstract Error Estimate

We will consider the following version of Strang's lemma

Lemma 4.2 *(Strang's lemma). If u is the solution of (16) and u_h the solution of (17), there exists a constant $C > 0$ which does not depend on the space step h such that*

$$\|u - u_h\|_1 \leq C \underbrace{\inf_{v_h \in V_h^r} \left\{ \|u - v_h\|_1 \right.}_{\text{interpolation error}} + \underbrace{\left. \sup_{w_h \in V_h^r} \frac{|a(v_h, w_h) - a_h(v_h, w_h)|}{\|w_h\|_1} \right\}}_{\text{numerical integration error}}.$$

Proof. The proof of this version of Strang's lemma is similar to the proof proposed by Ciarlet [7] for his version, noticing that, in our case, $V_h \subset V$. □

We now study separately the two terms of the right hand-side of the Strang's lemma, namely the interpolation error and the quadrature error. As $\mathbb{P}_r \subset P_r^F$, we first consider the case of $v_h = \pi_r u_h \in \mathbb{P}_r$ and get estimates in this case, and we will then take the infimum for $v_h \in P_r^F$ to get Strang's lemma. We also suppose that u is in $H^{r+1}(\Omega)$ for an Ω regular enough.

4.3.1. Interpolation error

Let Ω a set composed of n pyramids K_i

$$\Omega = \bigcup_{1 \leq i \leq n} K_i.$$

We denote

$$h_{K_i} = \text{diam } (K_i) = \sup_{x,y \in K_i} |x - y|, \quad h = \sup_{1 \leq i \leq n} h_{K_i}.$$

Let us recall the Bramble-Hilbert's lemma (Ciarlet [7])

Lemma 4.3 *(Bramble-Hilbert's lemma). Let $r \geq 1$, $m \leq r + 1$. We denote by π_r the projector to $\mathbb{P}_r$ satisfying*

$$\forall p \in \mathbb{P}_r(K),\ \pi_r p = p.$$

Then, there exists a constant $C_K > 0$ which only depends on K and r such that

$$\forall u \in H^{r+1}(K),\ \|u - \pi_r u\|_{m,K} \leq C_K h_K^{r+1-m} |u|_{r+1,K}.$$

We consider a domain made of a single pyramid K, ($\Omega = K$). For u in $H^{r+1}(\Omega)$, we apply Bramble-Hilbert lemma

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

where C_K depends on the shape of the pyramid K.

We suppose that the constants C_{K_i} for each element K_i of Ω are bounded by a constant C,

$$\forall K_i \in \Omega,\ C_{K_i} \leq C. \tag{18}$$

For $m = 1$, the sum for all the elements K_i of Ω and the inequality (15) on norms give the error estimates on Ω

$$\|u - \pi_r u\|_{1,\Omega} \leq C h^r \|u\|_{r+1,\Omega}. \tag{19}$$

Remark 4.4 *The condition $C_K \leq C$ is satisfied in the case of periodic meshes of the following sections, since the number of pyramids of different shape is finite. In a more general case, we conjecture that the boundedness of C_K is related to the existence of an upper bound for the inverse of jacobian matrix DF, as in the case of hexahedra (Girault and Raviart [16]).*

4.3.2. Quadrature Error

We now seek the minimal integration formula in order to obtain a quadrature error of order r.

Definition 4.5 *For $(v_h, w_h) \in P_r^F$ and $K \in \Omega$, we denote*

$$E_K(v_h, w_h) = \int_K v_h\, w_h\, dxdydz - \oint_K v_h w_h\, dxdydz,$$

where the approximate integral $\oint_K v_h w_h\, dxdydz$ is exact for $(1-z)^2\mathbb{Q}_{m,m,n}$.

We first get estimates for the mass term of the bilinear form a.

Lemma 4.6 *$\forall (v_h, w_h) \in P_r^F$, $K \in \Omega$ and $0 \le p, q \le m-r-1$, we have*

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

for a quadrature formula exact for polynomials of $(1-z)^2\mathbb{Q}_{m,m,m-1}$

Proof. Let us consider the integral

$$\int_K \pi_p v_h w_h\, dxdydz.$$

After change of variable, we get

$$\int_{\widetilde{Q}} 4\,(1-\widetilde{z})^2 (\pi_p \widetilde{v}_h)(\widetilde{|DF|}\widetilde{w}_h)\, d\widetilde{x}d\widetilde{y}d\widetilde{z}.$$

For a quadrature rule exact for polynomials of $(1-z)^2\mathbb{Q}_{m,m,m-1}$, as $\pi_p\widetilde{v}_h\, \widetilde{|DF|}\widetilde{w}_h \in \mathbb{Q}_{p+r+1,p+r+1,p+r}(\widetilde{x},\widetilde{y},\widetilde{z})$, we have

$$E_K(\pi_p v_h, w_h) = 0, \tag{20}$$

for $r+p+1 \le m$, that is $p \le m-r-1$, using lemma 3.3.

Similarly, we have

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

for $q \le m-r-1$, and for $p, q \le m-r-1$, we check that

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

which provides the claimed result. □

Proposition 4.7 *For $K \in \Omega$ and $v_h \in \mathbb{P}_r(K)$,*

$$\sup_{w_h \in P_r^F} \frac{E(v_h, w_h)}{\|w_h\|_{1,K}} \le C'_K h_K^r \|v_h\|_{r,K}$$

for a quadrature formula exact for polynomials of $(1-z)^2\mathbb{Q}_{m,m,m-1}$, with $m \ge 2r-1$.

Proof. We apply lemma 4.6 for $p = r-2$ and $q = 0$

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

Using the Cauchy-Schwarz inequality for both the exact and the approximate integral, which is possible since the norm defined by approximate integration is equivalent to the usual norm with a constant C_N, we get

$$|E(v_h, w_h)| \le C_N \|v_h - \pi_{r-2}v_h\|_{0,K} \|w_h - \pi_0 w_h\|_{0,K}.$$

As $v_h \in \mathbb{P}_r(K) \subset H^r(K)$ and $w_h \in H^1(K)$, Bramble-Hilbert's lemma and the inequality on the norms (15) provides

$$\|v_h - \pi_{r-2}v_h\|_{0,K} \le C_K h_K^{r-1} \|v_h\|_{r,K}.$$
$$\|w_h - \pi_0 w_h\|_{0,K} \le C_K h_K \|w_h\|_{1,K}.$$

Hence

$$|E(v_h, w_h)| \le C'_K h_K^r \|v_h\|_{r,K} \|w_h\|_{1,K}$$

with $C'_K = C_N C_K$ independent of w_h. Therefore, we obtain the claimed result. □

We then find estimates for the stiffness term of a.

Lemma 4.8

$$\left((1-\widetilde{z})^2 \widetilde{|DF|} \widetilde{DF}^{*-1} \nabla \widetilde{\varphi}_i\right)(\widetilde{x}, \widetilde{y}, \widetilde{z}) \in (1-\widetilde{z})^2 \mathbb{Q}_{r,r+1,r-1} \times \mathbb{Q}_{r+1,r,r-1} \times \mathbb{Q}_{r+1,r+1,r-1}(\widetilde{x}, \widetilde{y}, \widetilde{z}).$$

Proof. As $\widetilde{|DF|}\widetilde{DF}^{*-1}$ is the comatrix of $\widetilde{DF}$, using lemma 3.5, we get

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

that is

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

As

$$(\nabla \widetilde{\varphi}_i)(\widetilde{x}, \widetilde{y}, \widetilde{z}) \in \mathbb{Q}_{r-1,r,r-1} \times \mathbb{Q}_{r,r-1,r-1} \times \mathbb{Q}_{r,r,r-1}(\widetilde{x}, \widetilde{y}, \widetilde{z}),$$

by summing degrees, we obtain the claimed result. □

Lemma 4.9 *$\forall (v_h, w_h) \in P_r^F$, $K \in \Omega$ and $0 \le p \le m-r-1$, we have*

$$E_K(\nabla v_h, \nabla w_h) = E_K(\nabla v_h - \pi_p \nabla v_h, \nabla w_h)$$

for a quadrature formula exact for polynomials of $(1-z)^2 \mathbb{Q}_{m,m,m-2}$

Proof. As for lemma 4.6, we prove that

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

by considering the integral

$$\int_K \pi_p \nabla v_h \cdot \nabla w_h \, dxdydz.$$

After the change of variable, we get

$$\int_{\widetilde{Q}} 4\,(1-\widetilde{z})^2(\pi_p\nabla\widetilde{v}_h)\cdot(\widetilde{|DF|}\widetilde{DF}^{*-1}\nabla\widetilde{w}_h)\,d\widetilde{x}\,d\widetilde{y}\,d\widetilde{z}.$$

For a quadrature rule exact for polynomials of $(1-z)^2\mathbb{Q}_{m,m,m-2}$, using lemma 4.8, $(\pi_p\nabla\widetilde{v}_h)\cdot(\widetilde{|DF|}\widetilde{DF}^{*-1}\nabla\widetilde{w}_h) \in \mathbb{Q}_{p+r+1,p+r+1,p+r-1}(\widetilde{x},\widetilde{y},\widetilde{z})$, that is $E_K(\pi_p\nabla v_h,\nabla w_h)$ is equal to zero if $r+p+1\le m$, i.e. $p\le m-r-1$. □

Proposition 4.10 *For $K\in\Omega$ and $v_h\in\mathbb{P}_r(K)$,*

$$\forall w_h\in P_r^F,\ E_K(\nabla v_h,\nabla w_h)=0$$

for a quadrature formula exact for polynomials of $(1-z)^2\mathbb{Q}_{m,m,m-2}$, with $m\ge 2r$.

Proof. For $v_h\in\mathbb{P}_r(K)$, that is $\nabla v_h\in(\mathbb{P}_{r-1}(K))^3$, we apply lemma 4.9 with $p=r-1$ to conclude. □

We can now get the final quadrature error estimate using propositions 4.7 and 4.10, by summing for all the elements K of Ω and using (18)

$$\sup_{w_h\in V_h^r}\frac{|(a-a_h)(v_h,w_h)|}{\|w_h\|_{1,\Omega}}\le C'h^r\|v_h\|_{r+1,\Omega},\quad v_h\in\mathbb{P}_r(\Omega) \tag{21}$$

with $C'=C_NC$.

Theorem 4.11 *For a quadrature formula which is exact for polynomials of $(1-z)^2\mathbb{Q}_{m,m,m-2}$, for $m\ge 2r$, the final estimate from Strang's lemma is*

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

Proof. Summing equations (19) and (21) with $v_h=\pi_r u$, we get

$$\|u-\pi_r u\|_{1,\Omega}+\sup_{w_h\in V_h^r}\frac{|(a-a_h)(\pi_r u,w_h)|}{\|w_h\|_{1,\Omega}}\le C'h^r\|\pi_r u\|_{r,\Omega}+Ch^r\|u\|_{r+1,\Omega},\quad v_h\in\mathbb{P}_r(\Omega).$$

We then use the inequality $\|\pi_r u\|_{m,\Omega}\le\|u\|_{m,\Omega}$ to upper bound the right-hand side of the inequality with $max(C,C')$, and we take the infimum for the left-hand side thanks to inclusion (3). □

Remark 4.12 *We can take the points and weights of quadrature (ξ_k^G,ω_k^G) to get a quadrature formula exact for $\mathbb{Q}_{2r+1}$, instead of $(\xi_k^{HM},\omega_k^{HM})$ in the previous quadrature formula, without losing accuracy. Yet, this deteriorates the CFL as we shall see in section 6.2.*

5. Discontinuous Galerkin Method

For any element K of boundary ∂K, we consider the Local Discontinuous Galerkin (LDG) formulation (Castillo *et al.* [5])

$$\begin{cases} \int_K u\,\varphi + \int_K v\,\nabla\varphi - \int_{\partial K} \{v\}\cdot n\,\varphi = \int f\varphi \\ \int_K v\cdot\psi - \int_K \nabla u\cdot\psi - \int_{\partial K} \psi\cdot n\,[u] = 0 \end{cases}$$

with the following notations

$$\{v\} = \tfrac{1}{2}(v_1 + v_2)$$
$$[u] = \tfrac{1}{2}(u_2 - u_1),$$

where u_1 is the value of u restricted to K, and u_2 the value of u restricted to an element adjacent to K.

We can take the space $\hat{P}_r$ identical to the one in the continuous case. The error estimates are quite different from the ones obtained for continuous elements as the norm used for DG formulations is the L^2 norm, instead of the H^1 norm.

6. Numerical Results

6.1. Dispersion

In order to study the pyramidal elements, a dispersion analysis is performed on the wave equation, relying on the computation of the phase error on infinite periodic meshes, as in Cohen [9]. The periodic cell is a cube that can be composed of a single hexahedron; of two wedges; of two pyramids and two tetrahedra (hybrid); or of six pyramids; or of six tetrahedra as shown in Fig. 6.1. The analysis has also been carried out on periodic cells made up of distorted cubes in order to check the consistency of our method when the base of the pyramid is not a parallelogram as shown in Fig. 6.2 for the hybrid cell.

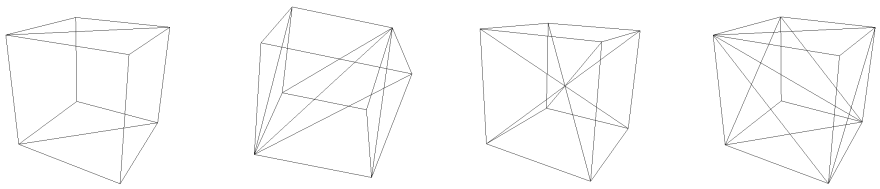


Fig. 6.1. Cells for the regular periodic mesh : wedges (top-left), hybrid(top-right), pyramids (bottom-left) and tetrahedra (bottom-right)

Considering the Helmholtz equation

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

we look for plane solutions of the form $u = e^{(\imath\vec{k}\cdot\vec{x})}$, where $\vec{k}$ is the wave vector, and $\|\vec{k}\|$ the wave number. The computational domain is reduced to a periodic cell where quasi-periodic conditions

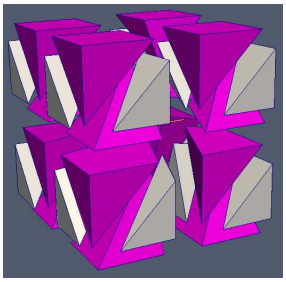


Fig. 6.2. Periodic pattern for the hybrid case, with distorted pyramids (purple) and tetrahedra (gray) (Color online)

are enforced, so that the plane wave is a solution of the continuous equations problem. The numerical eigenvalue closest to the wave number $\|\vec{k}\|$ is denoted by ω_h, so that we define q_h as

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

Since q_h should be close to 1, we can write

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

where p denotes the dispersion order of the scheme.

A dispersion order of $2r$ is obtained for both continuous and discontinuous Galerkin method, for regular as well as for distorted meshes (see Babuska and Osborn [1] for the factor 2), which coincides with the theoretical error estimates results obtained. This order of dispersion is clearly shown on the log-log curves in Fig. 6.3 for the continuous hybrid elements on a distorted mesh.

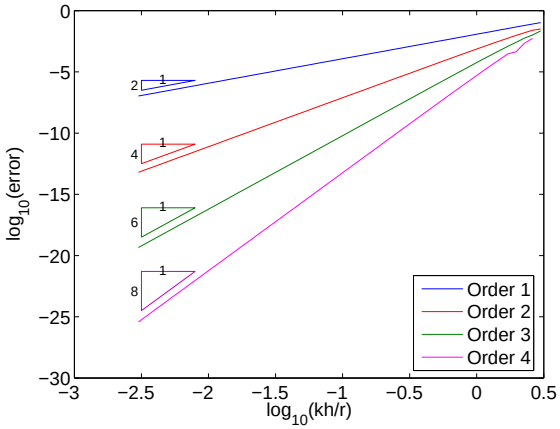


Fig. 6.3. log-log dispersion error for continuous finite elements of orders 1 to 4 for a hybrid distorted mesh

Dispersion curves are shown for regular elements of orders one to three in Fig. 6.4 for the continuous elements, and in Fig. 6.5 for the discontinuous elements. For the pyramids, exact and approximate integrations are giving very close results, and all elements exhibit similar

dispersion properties. The less dispersive element is the pyramidal one in most of the cases. The same study has been performed for distorted meshes and leads to the same conclusion.

6.2. Stability

The stability condition of the standard leap frog scheme is also computed on a periodic infinite mesh.

The CFL number, for which we have the stability condition $\Delta t \leq \text{CFL}\, h$, is defined by

$$CFL = \frac{1}{\sqrt{\max\limits_{\|\vec{k}\|\leq\pi} \lambda(M^{-1}(\vec{k})\, K(\vec{k}))}},$$

where $M(\vec{k})$ and $K(\vec{k})$ are the mass and stiffness matrices associated with the periodic cell, and $\vec{k}$ the wave vector.

For each type of element, the CFL number is given in Table 6.1 in the continuous case in Table 6.1, and in the discontinuous case in 6.2, up to order four. The stability criteria have been tested in the instationary case to check the correctness of the results.

Table 6.1: CFL for continuous finite elements with regular meshes.

	Order 1	Order 2	Order 3	Order 4
Hexahedron	0.28868	0.11785	0.06697	0.04264
Wedge	0.15794	0.07176	0.04344	0.02926
Pyramid ExactInt	0.09682	0.04803	0.03083	0.02143
Pyramid ApproxInt	0.07217	0.03335	0.01985	0.01316
Tetrahedron	0.12247	0.06253	0.03919	0.02669
Hybrid	0.15138	0.07245	0.04558	0.03191

Table 6.2: CFL for discontinuous Galerkin method with regular meshes.

	Order 1	Order 2	Order 3	Order 4
Hexahedron	0.14434	0.07144	0.04348	0.02934
Wedge	0.11471	0.04348	0.03957	0.02717
Pyramid ExactInt	0.07184	0.04058	0.02618	0.0184
Pyramid ApproxInt	0.04811	0.02544	0.01566	0.0112
Tetrahedron	0.07373	0.04467	0.03041	0.02173
Hybrid	0.09283	0.05363	0.03527	0.02448

For pyramidal elements, the CFL are computed with the quadrature formula $(\xi_k^{HM}, \omega_k^{HM})$ presented in paragraph 4.3.2 for exact integration, and with Gauss quadrature formula (ξ_k^G, ω_k^G) for approximate integration. The CFL of pyramidal elements are clearly lower for approximate integrals (60% lower on average) than for exact integrals. In this case, the CFL of pyramidal elements is close to the tetrahedral's.

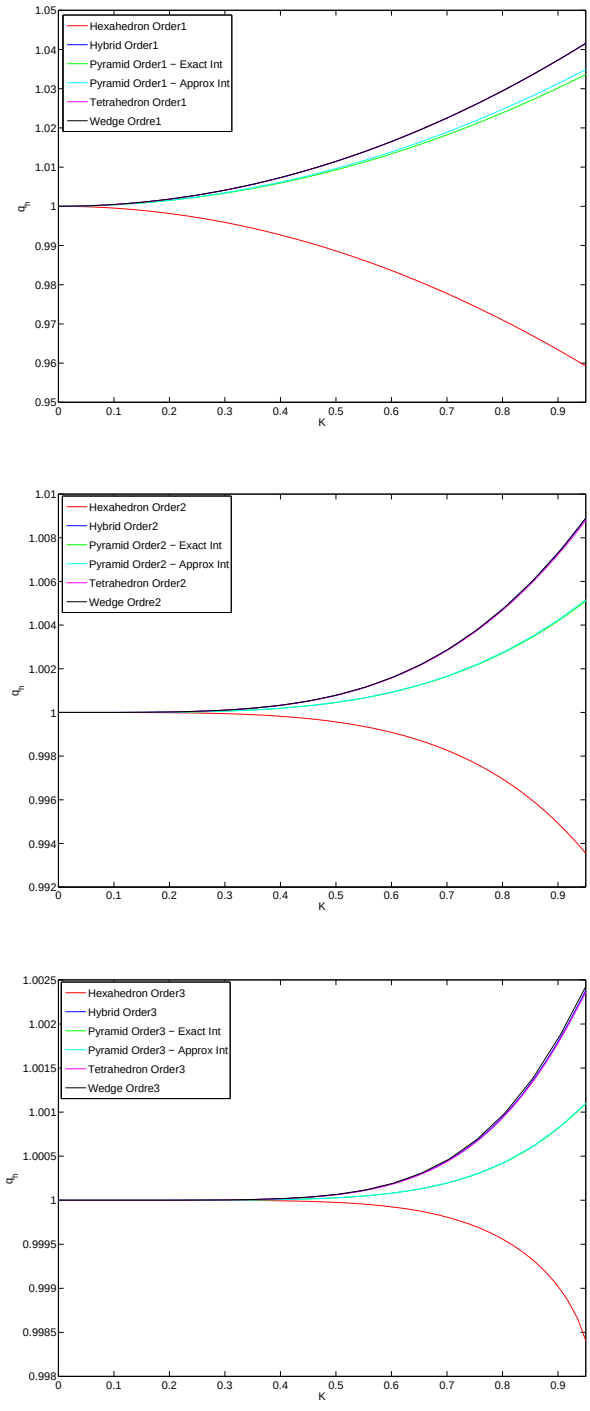


Fig. 6.4. Dispersion curves for continuous finite element of orders 1 to 3 for regular meshes ($K = \frac{6kh}{2\pi r}$)

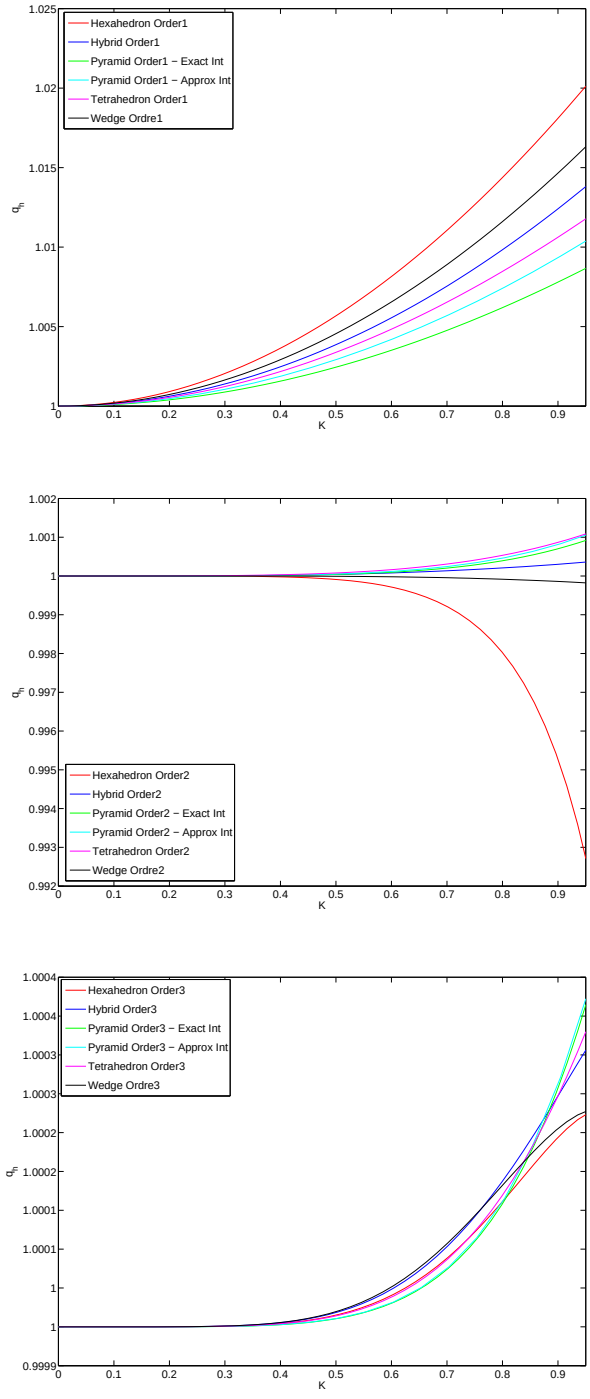


Fig. 6.5. Dispersion curves for discontinuous Galerkin method of orders 1 to 3 for regular meshes ($K = \frac{6kh}{2\pi r}$)

The same study has been performed for distorted meshes and leads to the same results, the CFL being smaller for the distorted meshes, and the CFL finally ranks as follows, in all the cases

$$CFL_{Hexa} > CFL_{Wedge} > CFL_{Tetra} > CFL_{Pyr-ExactInt} > CFL_{Pyr-ApproxInt}.$$

The CFL for a hybrid mesh is better than the tetrahedron's and pyramid's, which is a quite surprising result.

Remark 6.1 *A study of optimal location for the degrees of freedom inside the pyramid has been performed to get an optimal CFL, but this location appears to have almost no influence on the CFL.*

6.3. Numerical Comparison with Other Existing Methods

We display the dispersion obtained for the existing methods of Sherwin *et al.* [27], Nigam and Phillips [24], and Bluck and Walker [3] on Fig. 6.6, for order two and three. For order one, all these methods provide the same accuracy.

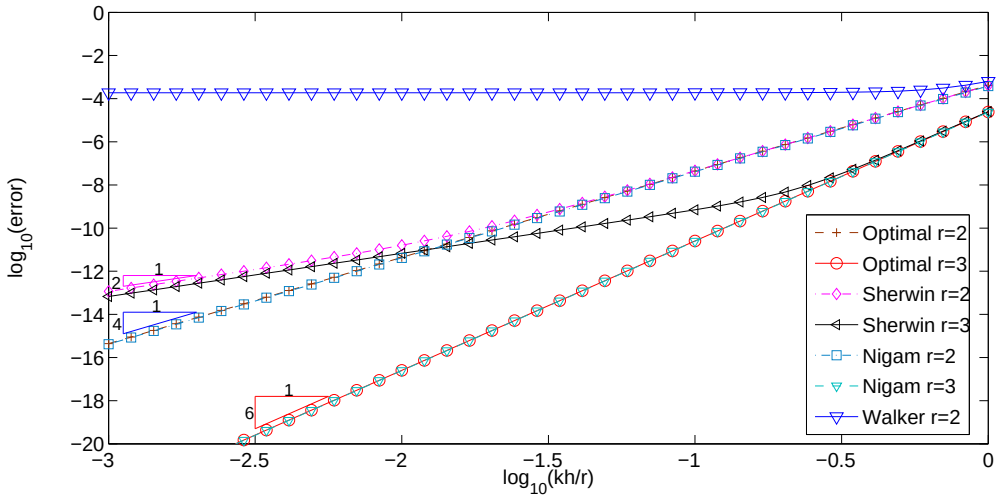


Fig. 6.6. Dispersion errors for the different types of existing elements of orders 2 and 3 with a distorted pyramidal mesh

The dispersion obtained with the space proposed by Sherwin *et al.* is of order two, whatever the order of approximation, as the basis functions for the base face and inside the pyramid are not sufficient to contain the optimal space. However, in the affine case, we check that we have a dispersion of order $2r$.

The dispersion obtained for the optimal space is equal to the space proposed by Nigam and Phillips, that is the degrees of freedom they add are not necessary as they do not increase the accuracy.

For order two, the dispersion obtained by Bluck and Walker is not consistent (dispersion of order 0) because the space proposed does not contain the space of order one. However, in a case of an affine pyramid, the dispersion obtained is in h^4 for order two.

6.4. Numerical Experiments

6.4.1. Test Case on a Cube

We first consider the Helmholtz equation on a cubic cavity with homogeneous Dirichlet boundary conditions

$$-\omega^2 u - \Delta u = f(x,y,z),$$

where

$$\Omega = [-1,1]^3 \quad \omega = 1.92\,\pi,$$

and f is a Gaussian source centered at the origin. We study convergence on a hybrid mesh with a similar pattern as for the dispersion (see Fig. 6.2).

Displaying the L^2 error obtained versus the space step h in a log-log scale in Fig. 6.7, we observe an L^2 error in $O(h^{r+1})$ as proved in the error estimates. For these experiments, we used exact integration with $r+1$ HM points in the direction $\tilde{z}$ and $r+1$ Gauss points in directions $\tilde{x}$ and $\tilde{y}$.

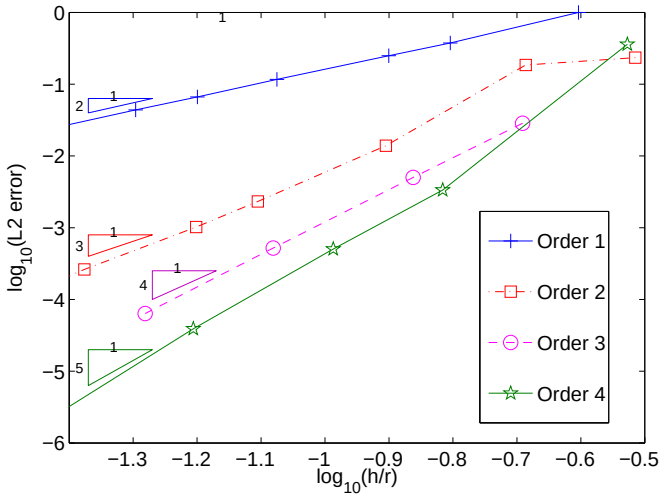


Fig. 6.7. L^2 error versus h for the cubic cavity at different order of approximations.

6.4.2. Low Storage Matrix-Vector Product

Since the involved matrices can require a huge amount of memory, in particular for high order approximation like $\mathbb{Q}_5$, we use a low storage matrix vector product (i.e. a matrix-free implementation). This is quite classical in discontinuous Galerkin methods on hexahedra (Castel *et al.* [4]) or on tetrahedra (Hesthaven and Teng [20]), and the extension to pyramid and prismatic

elements is straightforward. For continuous formulation, we use similar techniques, i.e. that we exploit factorizations of elementary mass and stiffness matrix

$$\int_{\hat{K}} |DF| \hat{\varphi}_i \, \hat{\varphi}_j = \hat{C} \, D_h \hat{C}^t$$
$$\int_{\hat{K}} |DF| DF^{-1} DF^{*-1} \hat{\nabla} \hat{\varphi}_i \cdot \hat{\nabla} \hat{\varphi}_j = \hat{R} \, B_h \hat{R}^t,$$

where matrices

$$\hat{C}_{i,j} = \hat{\varphi}_i(\xi_j) \quad \hat{R}_{i,j} = \nabla \hat{\varphi}_i(\xi_j)$$

are independent on the geometry, so they are not stored for each element, and matrices D_h and B_h

$$(D_h)_{i,j} = \omega_j |DF|(\xi_j)\delta_{i,j} \quad (B_h)_{i,j} = \omega_j (|DF| DF^{-1} DF^{*-1})(\xi_j)\delta_{i,j}$$

are respectively diagonal and block-diagonal (each block being a symmetric 3x3 matrix). Such factorization is explained by Cohen and Fauqueux in [10], and shown to be very efficient for hexahedral elements, since $\hat{R}$ is very sparse. In the case of other elements, $\hat{R}$ is dense and does not induce any gain in computational time. Yet, we still use this factorization, since the storage induced is very low as we only store matrices D_h and B_h.

6.4.3. Test Case with Curved Isoparametric Elements on a Sphere

Let us consider a sample test case of scattering by a sphere (see Fig 6.8)

$$\begin{cases} -\omega^2 u - \Delta u = 0 & \text{in } \Omega \\ \dfrac{\partial u}{\partial n} = -\dfrac{\partial u^{incident}}{\partial n} & \text{on } \Gamma \\ \dfrac{\partial u}{\partial n} - i\omega u = 0 & \text{on } \Sigma, \end{cases}$$

where Γ is a sphere of radius 3, $\omega = 2\pi$, and Σ is the boundary of the cube $[-5, 5]^3$.

To have a good approximation of the geometry, curved isoparametric elements are used. The implementation of such elements is explained in Šolín *et al.* [28], except for pyramids for which the extension is straightforward. The reference solution is computed on a refined pure hexahedral mesh with $\mathbb{Q}_7$ elements. On Fig. 6.9, three different meshes used for third order approximation are displayed.

The COCG solver (Clemens and Weiland [8]) is implemented, and can be used with or without preconditioning. In Table 6.3, the required number of degrees of freedom necessary to reach an error between one and two percent (measured in L^2 norm) are displayed for each type of mesh and at orders two, three and four. The results obtained without preconditioning and for a p-multigrid preconditioning using a damped Helmholtz equation (see Erlangga [15] for finite-difference, and Duruflé [13] for finite element) are also displayed, along with the computational time. The order of the coarsest mesh is set to 1 for orders two and three, and set to 2 for order five and $\mathbb{P}_4$.

The performance of hybrid mesh is quite similar to a purely hexahedral mesh while purely tetrahedral meshes result in much more expensive computations. The last row of this table concerns the use of a Gauss-Seidel smoother instead of the Jacobi smoother used in other rows. Jacobi smoother fails for hybrid meshes, but we don't have any explanation on this issue. In this case, a $\mathbb{Q}_5$ approximation on hexahedral mesh is much faster than for tetrahedral elements. This is the reason why it is important to use hybrid meshes with a high percentage of hexahedra.

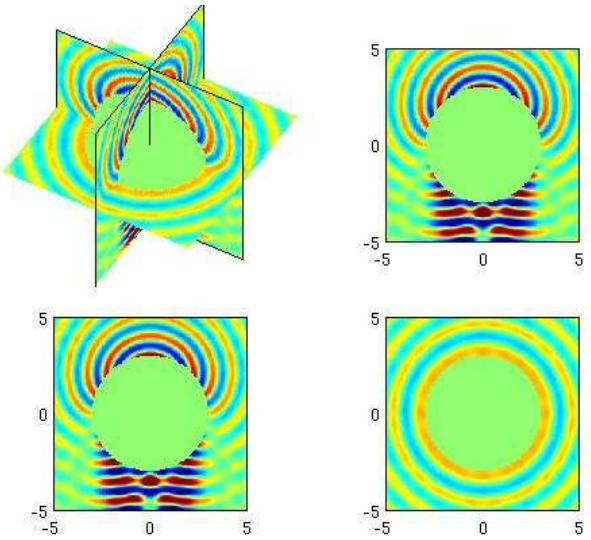


Fig. 6.8. Real part of diffracted field for a sphere of radius 3 with Neumann condition.

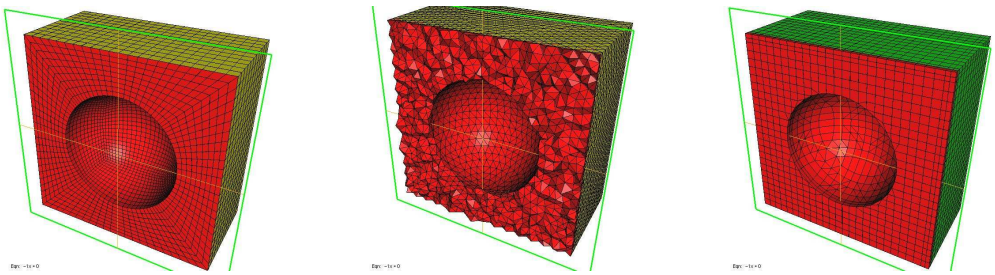


Fig. 6.9. Meshes used for third order approximation.

6.4.4. Numerical Experiment on a Piano

We now perform computations using the discontinuous Galerkin method for the transient wave equation

$$\begin{cases} \dfrac{\partial^2 u}{\partial t^2} - \Delta u = f(x,t) \quad \text{in } \Omega \\ \dfrac{\partial u}{\partial n} = 0 \quad \text{on } \Gamma \\ \dfrac{\partial u}{\partial n} + \dfrac{\partial u}{\partial t} = 0 \quad \text{on } \Sigma, \end{cases} \tag{22}$$

where Γ has the shape of the resonance cavity of a piano, and F is the surrounding parallelepiped box, as displayed in Fig. 6.10.

The source is chosen as

$$f(x,t) = \frac{1}{r_0^2} e^{-13\frac{r}{r_0}^2} e^{-4(t-t_0)^2} sin(2\pi f_0 t), \tag{23}$$

where r is the distance to the center of the source, r_0 the distribution radius of the Gaussian,

Order	2	3	5 ($\mathbb{P}_4$ for tetrahedra)
Hexahedra	964 000 dof	732 000 dof	315 000 dof
without precond.	2 762 iterations (3 410s)	2 938 iterations (2 024s)	3 467 iterations (**802**s)
preconditioned	133 iterations (637s)	127 iterations (**504**s)	130 iterations (**152s**)
Tetrahedra	1 216 000 dof	519 000 dof	339 000 dof
without precond.	2 300 iterations (12 622s)	1 656 iterations (3 490s)	1 942 iterations (17 835s)
preconditioned	58 iterations (1 019s)	51 iterations (534s)	119 iterations (587s)
Split tetrahedra	2 751 000 dof	936 000 dof	520 000 dof
without precond.	4 837 iterations (19 833s)	3 775 iterations (3 775s)	2 514 iterations (2 514s)
preconditioned	131 iterations (1 809s)	126 iterations (631s)	93 iterations (266s)
Hybrid	1 060 000 dof	455 000 dof	266 000 dof
without precond.	1 800 iterations (**2 744**s)	2 195 iterations (**1 153**s)	4 222 iterations (1 358s)
preconditioned	72 iterations (**388**s)	439 iterations (1 262s)	2 546 iterations (3 685s)
Hybrid GS precond.	69 iterations (**330s**)	76 iterations (**176s**)	128 iterations (161s)

Table 6.3: Number of degrees of freedom, number of iterations and computational time for the same accuracy.

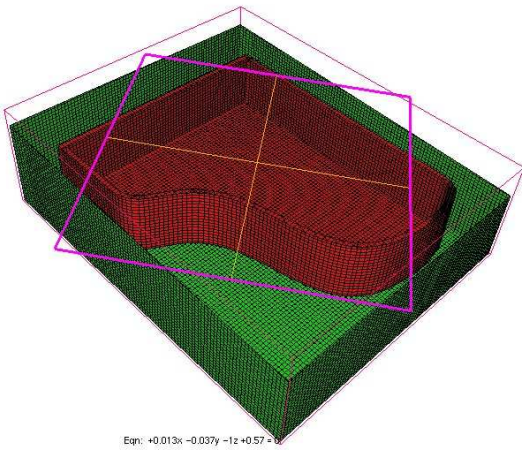


Fig. 6.10. Surface mesh of the piano-shaped cavity and surrounding box.

f_0 the frequency, and t_0 a constant. Here we have taken

$$r_0 = 0.1, \;\; f_0 = 14, \;\; t_0 = 1.858, \tag{24}$$

so that the parallelepiped box is as large as $32\lambda \times 26\lambda \times 10\lambda$ where $\lambda = \dfrac{1}{f_0}$ is the wavelength. We compute the solution from $t = 0$ until $t = 6$, and we obtain the result of Fig. 6.11.

A second-order leap frog scheme (Cohen and Fauqueux [10]) is used for the time discretization.

The reference solution is computed on a very fine mesh, and we compare two kind of meshes: a hybrid mesh and a hexahedral mesh obtained by splitting tetrahedra. Third order approximation ($\mathbb{Q}_3$) is used. The results are given in Table 6.4. We have given the computational time we would have obtained on a single processor, this time is computed by summing the

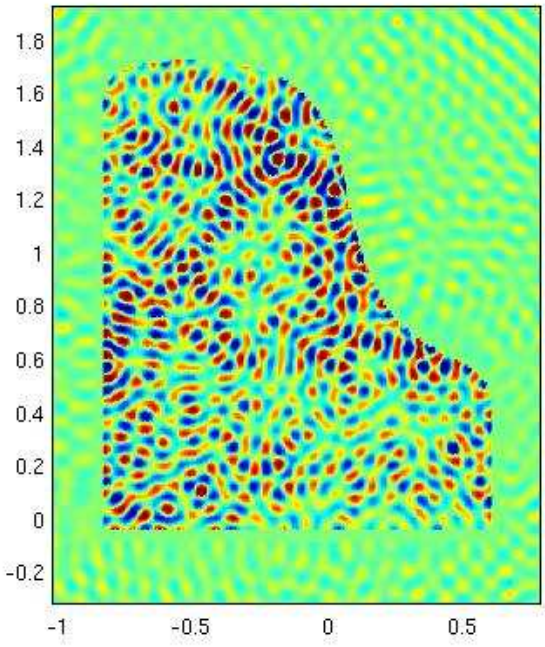


Fig. 6.11. Solution of the piano-shaped cavity on an horizontal section of the domain.

computational times for all the processors and we subtract the cost of communications. Since curved elements are not used, hexahedral meshes generated by splitting tetrahedra produce a bad approximation of the geometry, therefore they require a larger number of degrees of freedom.

Type of mesh	Split tetrahedra	Pure tetrahedra	Hybrid
Obtained accuracy	9.4 %	5.7 %	6.3 %
Degrees of freedom	49.3 millions	16.9 millions	14.88 millions
Time step	$\Delta t = 0.0002$	$\Delta t = 0.0004$	$\Delta t = 0.0005$
Computational time	**12.28 days**	**4.3** days	**1.18 day**

Table 6.4: Efficiency of different kind of meshes for the piano-shaped cavity.

Conclusion

Highly efficient pyramidal elements of any order have been constructed. Numerical experiments conducted with these elements (up to order six) exhibit a low phase error, a good CFL, and a very good behaviour in hybrid meshes.

Acknowledgments. We thank P. Ciarlet and S. Imperiale for useful suggestions and comments on the manuscript.

References

[1] I. Babuska and J. Osborn, *Eigenvalue Problems*, Handbook of Numerical Analysis Vol II, Finite Element Methods (Part 1), 1991.

[2] G. Bedrosian, *Shape Functions and Integration Formulas for Three-Dimensional Finite Element Analysis*, Int. J. Numer. Meth. Eng., **35** (1992), pp 95–108.

[3] M.J. Bluck and S.P. Walker, *Polynomial Basis Functions on Pyramidal Elements*, Comm. Numer. Meth. Engng., **24** (2008), pp 1827–1837.

[4] N. Castel, G. Cohen and M. Duruflé *Discontinuous Galerkin method for hexahedral elements and aeroacoustic*, Journal of Computational Acoustics, **17**-2 (2009), pp 175–196.

[5] P. Castillo, B. Cockburn, I. Perugia and D. Schötzau, *An a-priori Error Analysis of the Local Discontinuous Galerkin Method for Elliptic Problems*, SIAM J. Numer. Anal., **38**-5 (2000), pp 1676-1706.

[6] V. Chatzi and F.P. Preparata, *Using Pyramids in Mixed Meshes - Point Placement and Basis Functions*, Brown University, 2000.

[7] P.G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland, 1978.

[8] M. Clemens and T. Weiland, *Iterative methods for the solution of very large complex symmetric linear systems of equations in electrodynamics*, Technische Hochschule Darmstadt, Fachbereich 18 Elektrische Nachrichtentechnik, 2002.

[9] G.C. Cohen, *Higher-Order Numerical Methods for Transient Wave Equations*, Springer Verlag, 2000.

[10] G.C. Cohen and S. Fauqueux, *Mixed Finite Elements with Mass-Lumping for the Transient Wave Equation*, Journal of Computational Acoustics, **8** (2000), pp 171–188.

[11] G.C. 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**-2 (2006), pp 340–363.

[12] L. Demkowicz, J. Kurtz, D. Pardo, M. Paszynski, W. Rachowicz and A. Zdunek *Computing with hp-adaptive finite elements*, Volume 2, Chapman and Hall, 2007.

[13] M. Duruflé, *Intégration numérique et éléments finis d'ordre élevé appliqués aux équations de Maxwell en régime harmonique.*, Thèse de doctorat, Université Paris IX-Dauphine, 2006.

[14] 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**-3 (2009), pp 526–551.

[15] Y. A. Erlangga, *Some numerical aspects for solving sparse large linear systems derived from the Helmholtz equation*, Report of Delft University Technology, 2002.

[16] V. Girault and P.A. Raviart, *Finite Element Approximation of the Navier-Stokes Equations*, Berlin - Springer, 1979.

[17] R.D. Graglia, D. R. Wilton, A. F. Peterson and I-L. Gheorma, *Higher Order Interpolatory Vector Bases on Pyramidal Elements*, IEEE Trans. Ant. and Prop., **47**-5 (1999), pp 775–782.

[18] P.C. Hammer, O.J. Marlowe and A.H. Stroud, *Numerical Integration Over Simplexes and Cones*, Mathematical Tables and Other Aids to Computation, Vol. **10**-55 (1956), pp 130–137.

[19] J.S. Hesthaven, *From Electrostatics to Almost Optimal Nodal Sets for Polynomial Interpolation in a Simplex*, SIAM J. Numer. Anal., **35**-2 (1998), pp 665–676.

[20] J.S. Hesthaven and C.H. Teng, *Stable Spectral Methods on Tetrahedral Elements*, SIAM J. Numer. Anal., **21**-6 (2000), pp 2352–2380.

[21] G. Karniadakis and S. J. Sherwin, *Spectral/hp element methods for CF - Second Edition*, Oxford University Press, 2005.

[22] P. Knabner and G. Summ, *The Invertibility of the Isoparametric Mapping for Pyramidal and Prismatic Finite Elements*, Numer. Maths., **88** (2001), pp 661–681.

[23] L. Liu, K.B. Davies, K. Yuan and M. Křížek, *On symmetric pyramidal Finite Elements*, Dyn.

Contin. Discrete Impuls. Syst. Ser. B Appl. Algorithms, **11** (2004), pp 213–227.

[24] N. Nigam and J. Phillips, *Higher-Order Finite Elements on Pyramids*, 2007.

[25] 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*, Math. Comp., **76** (2007), pp 1801–1832.

[26] S.J. Sherwin, *Hierarchical hp Finite Element in Hybrid Domains*, Finite Elements in Analysis and Design, **27** (1997), pp 109–119.

[27] S.J. Sherwin, T. Warburton and G.E. Karniadakis, *Spectral/hp Methods For Elliptic Problems on Hybrid Grids*, Contemporary Mathematics, **218** (1998), pp 191–216.

[28] P. Šolín, K. Segeth and I. Dolezel, *Higher-order finite elements methods*, Studies in Advanced Mathematics, Chapman and Hall, 2003.

[29] B.A. Szabó and I. Babuška, *Finite Element Analysis*, John Wiley & Sons, 1991.

[30] T. Warburton, *Spectral/hp Methods on Polymorphic Multi-Domains: Algorithms and Applications*, Brown University, 1999.

[31] C. Wieners, *Conforming discretizations on tetrahedrons, pyramids, prisms and hexahedrons*, 1997

[32] S. Zaglmayr, *High Order Finite Elements for Electromagnetic Field Computation*, PhD thesis, Johannes Kepler University, Linz Austria, 2006.

[33] F.X. Zgainski, J.C. Coulomb, Y. Marchal, F. Claeyssen and X. Brunotte, *A New Family of Finite Elements : The Pyramidal Elements*, IEEE Transactions on Magnetics, **32**-3 (1996), pp 1393–1396.