# NON SPURIOUS SPECTRAL-LIKE ELEMENT METHODS FOR MAXWELL'S EQUATIONS* 

Gary Cohen and Marc Duruflé<br>(INRIA, Domaine de Voluceau, Rocquencourt - BP 105, 78153 Le Chesnay Cedex, France<br>Email: gary.cohen@inria.fr, Marc.Durufle@inria.fr)


#### Abstract

In this paper, we give the state of the art for the so called "mixed spectral elements" for Maxwell's equations. Several families of elements, such as edge elements and discontinuous Galerkin methods (DGM) are presented and discussed. In particular, we show the need of introducing some numerical dissipation terms to avoid spurious modes in these methods. Such terms are classical for DGM but their use for edge element methods is a novel approach described in this paper. Finally, numerical experiments show the fast and low-cost character of these elements.

Mathematics subject classification: 65M60, 65M70, 35Q60, 34L16, 65N25. Key words: Maxwell's equations, Discontinuous Galerkin methods, Edge elements, Mass lumping.


## 1. Introduction

For a long time, Maxwell's equations were mainly solved in the time-harmonic domain. The evolution of radar techniques showed the limit of this formulation which can only treat monochromatic sources. Both engineers and researchers were then motivated to use equations in the time domain which can take into account large frequency sources in one resolution. The first and most popular approximation of Maxwell's equations in the time domain was provided by the Yee's scheme [29], commonly called FDTD (Finite Difference in the Time Domain) by engineers, which is basically a centered second order finite difference approximation of Maxwell's equation.

Although easy to implement, FDTD has some difficulties to treat complex geometries. In fact, the staircase approximation of curved boundary can produce spurious reflections which can substantially pollute the solution. On the other hand, finite element methods (FETD) have the major drawback of producing a $n$-diagonal ( $n$ can grow up to several tens in 3D) mass matrix which must be inverted at each time-step, which is a serious handicap for FETD versus FDTD whose mass matrix is the identity matrix. This mass matrix does not present any difficulty to time harmonic problems, for which even the stiffness matrix must be inverted. For this reason, industry was reluctant to use FETD for a long time and FDTD remains the reference for Maxwell's equations in the time domain for 40 years!

The mass lumping technique is an efficient alternative to mass matrix inversion. However, this technique was well known for lower order continuous (or $H^{1}$ ) elements but not obvious for higher-order approximations. A first step towards a general mass lumping technique was made by Hennart et al. [19,20] and independently by Young [30] which proposed to use GaussLobatto quadrature formulas to get mass lumping for continuous quadrilateral or hexahedral

[^0]finite elements. Besides mass lumping, these formulas ensure to keep the order induced by the finite element approximation [3]. First introduced for ODE or parabolic problems, this technique was extended to the wave equation by Cohen et al. [11] and renamed spectral element methods [22]. The non trivial extension of this technique to triangular and tetrahedral elements was later realized by Cohen et al. [10] for triangles up to third order and Mulder et al. [23] for higher-order triangles and tetrahedra.

The problem of the mass matrix inversion was solved for the wave equation and remained a challenging problem for Maxwell's equations. A first try was done by Haugazeau et al. [18] but this approach remains restricted to first-order approximation. A second and natural step was to extend spectral element techniques to edge (or $H($ curl $)$ ) elements. This was done by Cohen et al. for orthogonal meshes for the first family [12] and for any mesh for the second family of edge elements of any order [13]. The extension to triangular and tetrahedral meshes was realized by Elmkies et al. [16, 17] but lead to efficient approximations up to the second-order elements.

Due to the storage of the stiffness matrix, even by using mass lumping techniques, FETD remained much more expensive than FDTD in terms of storage and, to a lesser extent, in computational time. This ultimate problem was solved by using a mixed $H$ (curl) - $L^{2}$ formulation of Maxwell's equations based on $H($ curl $)$-conform definition of the curl operator in both spaces [13]. This technique provides a local definition of the stiffness matrices which induces a substantial gain of storage. It was later extended to acoustics [6] and linear elastodynamics [7]. A detailed presentation of all these techniques can be found in [5].

Unfortunately, although $H^{1}$ spectral elements and $H^{1}$ and $H$ (curl) triangular and tetrahedral elements behave quite well for any mesh, $H$ (curl) spectral elements present important parasitic waves for very distorted meshes, which are often used in industrial problems. For this reason, discontinuous Galerkin methods appeared as an efficient alternative for Maxwell's equations. First introduced by Hesthaven [21] for tetrahedra, this approach was adapted by Cohen et al. [9] to the spectral element point of view, which provided a low-storage as well as fast method to solve Maxwell's equations. This approach seemed to deal better with parasitic waves but eigenvalues considerations showed that such waves were however present in this method. All these remarks motivated us to discuss the numerical dissipation terms which can attenuate parasitic waves. This discussion is the new part of our survey.

Our paper is divided into four parts. In a first section, we present the continuous formulations of the Maxwell's equations and different approaches for its approximation. In the second section, we discuss the parasitic modes through an eigenvalue analysis. In the third section, we introduce dissipative jump terms to get rid of parasitic modes. The last section is devoted to the approximation of the time-harmonic problem by the methods described in the first section. Finally, some numerical experiments are presented.

## 2. Different Approximations of the Problem

### 2.1. Formulations of the continuous problem

In this paper, we are interested in solving the so-called lossy Maxwell's equations in anisotropic media which read

Find $(\vec{E}, \vec{H}): \Omega \times] 0, T\left[\rightarrow \mathbb{R}^{3}\right.$ such that

$$
\begin{align*}
& \underline{\varepsilon} \frac{\partial \vec{E}}{\partial t}-\nabla \times \vec{H}+\underline{\underline{\sigma}} \vec{E}=-\vec{J},  \tag{2.1a}\\
& \underline{\underline{\mu}} \frac{\partial \vec{H}}{\partial t}+\nabla \times \vec{E}=0, \tag{2.1b}
\end{align*}
$$

where $\vec{J}$ is a given function of time and space which can model a pulse or a plane wave. $\underline{\underline{\varepsilon}}, \underline{\underline{\sigma}}$ and $\underline{\underline{\mu}}$ are the permittivity, conductivity and permeability which can take into account heterogeneous $\overline{\text { and }}$ anisotropic media.

To these equations, we add homogeneous initial conditions on $(\vec{E}, \vec{H})$ and the perfectly conducting condition $\vec{E} \times \vec{n}=0$ on $\partial \Omega$.

By plugging Eq. (2.1b) in (2.1a), this system can be expressed as the following unique equation in $\vec{E}$ :

$$
\begin{equation*}
\underline{\underline{\varepsilon}} \frac{\partial^{2} \vec{E}}{\partial t^{2}}+\nabla \times\left(\underline{\underline{\mu}}^{-1}(\nabla \times \vec{E})\right)+\underline{\underline{\sigma}} \frac{\partial \vec{E}}{\partial t}=-\vec{J}^{\prime}, \tag{2.2}
\end{equation*}
$$

where $\vec{J}^{\prime}=\partial \vec{J} / \partial t$.

### 2.2. General issues

### 2.2.1. Some general principles

In this section, we introduce some basic principles which will be used for the different approximations of the Maxwell's problem described below.

1. Let us first define the fundamental space for Maxwell's equations

$$
\begin{equation*}
H_{0}(\text { curl }, \Omega)=\left\{\vec{v}_{h} \in\left[L^{2}(\Omega)\right]^{3} \text { such that } \nabla \times \vec{v}_{h} \in\left[L^{2}(\Omega)\right]^{3}\right\} . \tag{2.3}
\end{equation*}
$$

Functions of this functional space are characterized by the continuity of their tangential component.
2. For any quadrilateral or hexahedral mesh, basis functions are defined on the reference element $\widehat{K}$ which is the unit square in 2 D and the unit cube in 3 D . These functions (generally called hat functions because indicated by a hat) are functions of $\widehat{\vec{x}} \in \widehat{K}$. Then, a basis functions in any element $K_{j}$ of a mesh $\mathcal{Q}$ is obtained by a transform based on $\vec{F}_{j}$ such that $\vec{F}_{j}(\widehat{K})=K_{j}$. Obviously, this basis function depends on $\vec{x}=\vec{F}_{j}(\widehat{\vec{x}})$.
In our different approximations, we use the $H($ curl $)$-conform transform which maps a hat vector basis function $\widehat{\vec{\varphi}}$ onto a basis function $\vec{\varphi}$ on $K_{j}$ as follows:

$$
\vec{\varphi} \circ \vec{F}_{j}=D F_{i}^{*-1} \widehat{\vec{\varphi}},
$$

where $D F_{i}^{*-1}$ is the inverse of the transposed Jacobian matrix of $\vec{F}_{j}$.
This transform is called $H$ (curl)-conform because it ensures the continuity of the tangential component of $\vec{\varphi}$ all over the mesh (cf. the appendix of [5]).
3. Another important point in the following is the use of mass lumping. This property is obtained by defining interpolation points for the basis functions (which are basically Lagrange interpolation polynomials) as quadrature points and by computing mass integrals by the corresponding quadrature formula. In the case of Maxwell's equations, this technique provides block-diagonal mass matrices.

### 2.2.2. A general framework for edge elements

The edge element methods are based on the following variational formulation of problem (2.2):
Find $\vec{E} \in L^{2}\left(0, T ; H_{0}(c \overrightarrow{u r l}, \Omega)\right)$ such that

$$
\begin{align*}
& \frac{\mathrm{d}^{2}}{\mathrm{~d} t^{2}} \int_{\Omega} \underline{\underline{\varepsilon}} \vec{E} \cdot \vec{\varphi} \mathrm{~d} \vec{x}+\int_{\Omega} \underline{\underline{\mu}}^{-1}(\nabla \times \vec{E}) \cdot \nabla \times \vec{\varphi} \mathrm{d} \vec{x}+\frac{\mathrm{d}}{\mathrm{~d} t} \int_{\Omega} \underline{\underline{\sigma}} \vec{E} \cdot \vec{\varphi} \mathrm{~d} \vec{x} \\
= & -\int_{\Omega} \overrightarrow{J^{\prime}} \cdot \vec{\varphi} \mathrm{d} \vec{x}, \quad \forall \vec{\varphi} \in H_{0}(c \vec{r} l, \Omega) . \tag{2.4}
\end{align*}
$$

Then, we construct an edge element approximation based on this variational formulation which reads

Find $\vec{E}_{h} \in L^{2}\left(0, T ; \vec{V}_{h}^{r}\right)$ such that

$$
\begin{align*}
& \frac{\mathrm{d}^{2}}{\mathrm{~d} t^{2}} \int_{\Omega} \underline{\underline{\varepsilon}} \vec{E}_{h} \cdot \vec{\varphi}_{h} \mathrm{~d} \vec{x}+\int_{\Omega} \underline{\underline{\mu}}^{-1}\left(\nabla \times \vec{E}_{h}\right) \cdot \nabla \times \vec{\varphi}_{h} \mathrm{~d} \vec{x}+\frac{\mathrm{d}}{\mathrm{~d} t} \int_{\Omega} \underline{\underline{\sigma}} \vec{E}_{h} \cdot \vec{\varphi}_{h} \mathrm{~d} \vec{x} \\
= & -\int_{\Omega} \vec{J}^{\prime} \cdot \vec{\varphi}_{h} \mathrm{~d} \vec{x}, \quad \forall \vec{\varphi}_{h} \in \vec{V}_{h}^{r} . \tag{2.5}
\end{align*}
$$

### 2.3. A first family of edge elements

For the first family of edge elements (defined in [24]), we have,

$$
\begin{align*}
\vec{V}_{h}^{r}=\{ & \vec{v}_{h} \in H_{0}(c \overrightarrow{u r l}, \Omega) \text { such that } \forall K_{j} \in \mathcal{Q}, \\
& \left.D F_{i}^{*} \vec{v}_{\left.\right|_{K_{j}}} \circ \vec{F}_{j} \in Q_{r-1, r, r} \times Q_{r, r-1, r} \times Q_{r, r, r-1}\right\}, \tag{2.6}
\end{align*}
$$

where

$$
\begin{equation*}
Q_{r_{1}, r_{2}, r_{3}}=\left\{p(\vec{x})=\sum_{\ell=0}^{r_{1}} \sum_{m=0}^{r_{2}} \sum_{n=0}^{r_{3}} a_{\ell, m, n} x_{1}^{\ell} x_{2}^{m} x_{3}^{n}, a_{\ell, m, n} \in \mathbb{R}\right\} . \tag{2.7}
\end{equation*}
$$

For spectral edge elements, the basis functions on $\widehat{K}$ for this family are constructed from 1D Lagrange interpolation polynomials based on Gauss and Gauss-Lobatto quadrature points (denoted ( $\widehat{\xi}_{\ell}^{G}, \ell=1, \cdots, r$ ) and ( $\widehat{\xi}_{\ell}^{G L}, \ell=1, \cdots, r+1$ ) respectively). For each set of points, the corresponding polynomials are denoted $\widehat{g}_{\ell}$ and $\widehat{l}_{\ell}$ respectively. For instance, the function of $Q_{r-1, r, r} \times Q_{r, r-1, r} \times Q_{r, r, r-1}$ corresponding to the point $\left(\widehat{\xi}_{\ell}^{G}, \widehat{\xi}_{m}^{G L}, \widehat{\xi}_{n}^{G L}\right) \in \widehat{K}$ reads

$$
\widehat{\vec{\varphi}}_{\ell, m, n}\left(\widehat{x}_{1}, \widehat{x}_{2}, \widehat{x}_{3}\right)=\widehat{g}_{\ell}\left(\widehat{x}_{1}\right) \widehat{l}_{m}\left(\widehat{x}_{2}\right) \widehat{l}_{n}\left(\widehat{x}_{3}\right) \vec{e}_{1},
$$

where $\vec{e}_{1}$ is the first basis vector of $\mathbb{R}^{3}$. In the same way, functions corresponding to points with a Gauss coordinate as second (resp. third) coordinate, are supported by $\vec{e}_{2}$ (resp. $\vec{e}_{3}$ ).

For any $\widehat{g}_{\ell}$ and any $\widehat{\xi}_{m}^{G}$ and for any $\widehat{l}_{\ell}$ and any $\widehat{\xi}_{m}^{G L}$, we have $\widehat{g}_{\ell}\left(\widehat{\xi}_{m}^{G}\right)=\delta_{\ell m}$ and $\widehat{l}_{\ell}\left(\widehat{\xi}_{m}^{G L}\right)=\delta_{\ell m}$, where $\delta_{\ell m}$ is the Kronecker symbol.

The corresponding degrees of freedom on faces are given in Fig. 2.1.

### 2.4. A second family of edge elements

The definition of $\vec{V}_{h}^{r}$ for the second family of edge elements is much simpler. It reads

$$
\begin{equation*}
\vec{V}_{h}^{r}=\left\{\vec{v}_{h} \in H_{0}(c \overrightarrow{u r l}, \Omega) \text { such that } \forall K_{j} \in \mathcal{Q}, D F_{i}^{*} \vec{v}_{h_{\left.\right|_{K}}} \circ \vec{F}_{j} \in\left(Q_{r}\right)^{3}\right\}, \tag{2.8}
\end{equation*}
$$



Fig. 2.1. The degrees of freedom of the first family of edge elements for $\vec{E}$ on faces of cubes when $r=2$.
where $Q_{r}=Q_{r, r, r}, Q_{r, r, r}$ defined as in (2.7).
For spectral edge elements, the basis functions on $\widehat{K}$ for this family are constructed from 1D Lagrange interpolation polynomials based on Gauss-Lobatto quadrature points. For any point $\ell$, the corresponding polynomial is denoted $\widehat{\varphi}_{\ell}$. For instance, the function of $\left(Q_{r}\right)^{3}$ corresponding to the point $\left(\widehat{\xi}_{\ell}^{G L}, \widehat{\xi}_{m}^{G L}, \widehat{\xi}_{n}^{G L}\right) \in \widehat{K}$ in the $\vec{e}_{p}$-direction reads

$$
\widehat{\widehat{\varphi}}_{\ell, m, n}^{p}\left(\widehat{x}_{1}, \widehat{x}_{2}, \widehat{x}_{3}\right)=\widehat{\varphi}_{\ell}\left(\widehat{x}_{1}\right) \widehat{l}_{\varphi}\left(\widehat{x}_{2}\right) \widehat{l}_{\varphi}\left(\widehat{x}_{3}\right) \vec{e}_{p}
$$

Moreover, we have

$$
\widehat{\widehat{\varphi}}_{\ell, m, n}^{p}\left(\widehat{\xi}_{i}^{G L}, \widehat{\xi}_{j}^{G L}, \widehat{\xi}_{k}^{G L}\right)=\delta_{\ell i} \delta_{m j} \delta_{n k} \vec{e}_{p}
$$

The corresponding degrees of freedom are given in Fig. 2.2.
This method provides a block-diagonal mass matrix on any mesh. This structure is obtained by numbering sequentially the $n_{i}$ degrees of freedom around a point $i$ and by using a GaussLobatto quadrature formula for computing the mass (and also stiffness) integrals. Therefore, the size of a block is equal to $n_{i}$.

### 2.5. Mixed formulation

Although the above method solves the problem of mass lumping, which is important in terms of computational time, due to the $H(c u r l)$-conform transform, the storage of the stiffness matrix on non-orthogonal meshes remains substantial. This storage is still a drawback of FETD, even mass-lumped, versus FDTD. A way to overcome this problem is to use a mixed formulation.

In order to write this new approach, we use equations (2.1a)-(2.1b) with the following variational formulation:

$$
\begin{align*}
& \frac{\mathrm{d}}{\mathrm{~d} t} \int_{\Omega} \underline{\underline{\varepsilon}} \vec{E} \cdot \vec{\varphi} \mathrm{~d} \vec{x}-\int_{\Omega} \vec{H} \cdot \nabla \times \vec{\varphi} \mathrm{d} \vec{x}+\int_{\Omega} \underline{\underline{\sigma}} \vec{E} \cdot \vec{\varphi} \mathrm{~d} \vec{x}=-\int_{\Omega} \vec{J} \cdot \vec{\varphi} \mathrm{~d} \vec{x}, \forall \vec{\varphi} \in H_{0}(c \overrightarrow{r r l}, \Omega)  \tag{2.9a}\\
& \frac{\mathrm{d}}{\mathrm{~d} t} \int_{\Omega} \underline{\underline{\mu}} \vec{H} \cdot \vec{\psi} \mathrm{~d} \vec{x}+\int_{\Omega} \nabla \times \vec{E} \cdot \vec{\psi} \mathrm{~d} \vec{x}=0, \forall \vec{\psi} \in\left[L^{2}(\Omega)\right]^{3} \tag{2.9b}
\end{align*}
$$



Fig. 2.2. The degrees of freedom of the second family of edge elements for $\vec{E}$ on faces of cubes when $r=2$.
which provides the approximate problem

$$
\begin{align*}
& \frac{\mathrm{d}}{\mathrm{~d} t} \int_{\Omega} \underline{\underline{\varepsilon}} \vec{E}_{h} \cdot \vec{\varphi}_{h} \mathrm{~d} \vec{x}-\int_{\Omega} \vec{H}_{h} \cdot \nabla \times \vec{\varphi}_{h} \mathrm{~d} \vec{x}+\int_{\Omega} \underline{\underline{\sigma}} \vec{E}_{h} \cdot \vec{\varphi}_{h} \mathrm{~d} \vec{x}=-\int_{\Omega} \vec{J} \cdot \vec{\varphi}_{h} \mathrm{~d} \vec{x}, \forall \vec{\varphi}_{h} \in \vec{V}_{h}^{r} \\
& \frac{\mathrm{~d}}{\mathrm{~d} t} \int_{\Omega} \underline{\underline{\mu}} \vec{H}_{h} \cdot \vec{\psi}_{h} \mathrm{~d} \vec{x}+\int_{\Omega} \nabla \times \vec{E}_{h} \cdot \vec{\psi}_{h} \mathrm{~d} \vec{x}=0, \forall \vec{\psi}_{h} \in \vec{W}_{h}^{r} \tag{2.10b}
\end{align*}
$$

where

$$
\begin{equation*}
\vec{W}_{h}^{r}=\left\{\vec{v}_{h} \in\left[L^{2}(\Omega)\right]^{3} \text { such that } \forall K_{j} \in \mathcal{Q}, D F_{i}^{*} \vec{v}_{\left.\right|_{K_{j}}} \circ \vec{F}_{j} \in\left(Q_{r}\right)^{3}\right\} \tag{2.11}
\end{equation*}
$$

Basis functions of $\vec{W}_{h}^{r}$ are constructed exactly in the same way as in Section 2.4 but they are discontinuous. In fact, this discontinuous character makes useless the use of the $H$ (curl)conform transform which ensures the tangential continuity. Its presence in $\vec{W}_{h}^{r}$ is an artifact to provide an important gain of storage based on the following identity:

$$
\begin{equation*}
\forall K_{j} \in \mathcal{Q}, \forall \vec{\varphi}_{h} \in \vec{V}_{h}^{r}, \forall \vec{\psi}_{h} \in \vec{W}_{h}^{r}, \int_{K_{j}} \nabla \times \vec{\varphi}_{h} \cdot \vec{\psi}_{h} \mathrm{~d} \vec{x}=\operatorname{sign}\left(J_{j}\right) \int_{\widehat{K}} \widehat{\nabla} \times \widehat{\vec{\varphi}}_{h} \cdot \widehat{\vec{\psi}}_{h} \mathrm{~d} \widehat{\vec{x}}, \tag{2.12}
\end{equation*}
$$

where $\widehat{\nabla}$ is the $\nabla$ operator in $\widehat{\vec{x}}$ coordinates, $J_{j}=\operatorname{det} D F_{i}$ and $\widehat{\vec{\varphi}}_{h}=\left.D F_{i}^{*} \vec{\varphi}_{h}\right|_{K_{j}} \circ \vec{F}_{j}$.
Identity (2.12) means that the knowledge of the stiffness matrix on $\widehat{K}$ implies its knowledge on any element $K_{j}$ (modulo one sign per element). In other words, if one assemblies the stiffness matrix at each time-step (which is a low-cost operation, in particular for higher-order approximations), the storage of the stiffness matrix is reduced to its storage over $\widehat{K}$ (which is peanuts!), the storage of one sign per element and the storage of the mass matrix of (2.10b) which is a $3 \times 3$ block-diagonal matrix. Of course, this induces a huge gain of storage, as indicated by the left curve of Fig. 2.3. Besides this gain of storage, due to orthogonality properties of the basis functions over $\widehat{K}$, this approach also induces a substantial gain of CPU time, as one can see in the right curve of Fig. 2.3.

By using a leapfrog scheme in time, (2.10a)-(2.10b) can be written as follows in its discrete


Fig. 2.3. Storage (left) and CPU time (right) ratios for spectral elements and mixed spectral elements for the same number of degrees of freedom.
form

$$
\begin{align*}
& B_{\varepsilon} \frac{\mathbf{E}^{\mathbf{n + 1}}-\mathbf{E}^{\mathbf{n}-\mathbf{1}}}{\Delta t}-R_{h} \mathbf{H}^{\mathbf{n}+\mathbf{1} / \mathbf{2}}+B_{\sigma} \frac{\mathbf{E}^{\mathbf{n}+\mathbf{1}}+\mathbf{E}^{\mathbf{n}}}{2}=0  \tag{2.13a}\\
& B_{\mu} \frac{\mathbf{H}^{\mathbf{n}+\mathbf{1} / \mathbf{2}}-\mathbf{H}^{\mathbf{n}-\mathbf{1} / \mathbf{2}}}{\Delta t}+R_{h}^{*} \mathbf{E}^{\mathbf{n}}=0 \tag{2.13b}
\end{align*}
$$

where $B_{\varepsilon}, B_{\sigma}$ are $n_{i} \times n_{i}$ block-diagonal mass matrices, $B_{\mu}$ is a $3 \times 3$ block-diagonal mass matrix and $R_{h}$ is a stiffness matrix which needs almost no storage.

A detailed description of the above methods can be found in [5].

## Remarks:

1. The mixed formulation was also applied to the acoustics equation [6] and the linear elastodynamics system [7].
2. A proof of the equivalence between the approximations defined in (2.4) and in (2.10a)(2.10b) is given in Pernet's thesis [26]. This proof is a powerful generalization of the equivalence theorem proven in [6].
3. This mixed formulation can also be applied to the first family of edge elements. Although inefficient in terms of mass lumping on non-orthogonal meshes for this family, this formulation provides a fast algorithm of matrix-vector product very useful for the timeharmonic problem, as discussed in Section 4. More generally, mixed spectral elements enable to isolate the geometry from the differential operators. This property provides a fast matrix-product technique which can be efficiently use for a large class of hyperbolic problems.
4. Another approximation of (2.9a)-(2.9b) for the first family of edge elements, in which $\left[L^{2}(\Omega)\right]^{3}$ is replaced by $H(\operatorname{div}, \Omega)$, is given in [12]. This approach leads to a variational generalization of the Yee's scheme at any order.

### 2.6. Discontinuous Galerkin methods

### 2.6.1. Description of the method

In a discontinuous Galerkin method (DGM), the solution is sought in $\left[L^{2}(\Omega)\right]^{3}$. So, one cannot express a continuous variational formulation. The approximate problem reads as follows.

For each $K_{j} \in \mathcal{Q}$, we set

$$
\begin{align*}
\int_{K_{j}} \varepsilon \frac{\partial \mathbf{E}}{\partial t} \cdot \psi d x= & -\int_{K_{j}} \sigma \mathbf{E} \cdot \psi d x+\int_{K_{j}} \nabla \times \mathbf{H} \cdot \psi d x-\int_{K_{j}} \vec{J} \cdot \psi d x \\
& +\alpha \int_{\partial K_{j}} \llbracket \mathbf{H} \times \mathbf{n} \rrbracket_{\partial K_{j}}^{K_{j}} \cdot \psi d s-\beta \int_{\partial K_{j}} \llbracket \mathbf{n} \times \mathbf{E} \rrbracket{ }_{\partial K_{j}}^{K_{j}} \cdot(\mathbf{n} \times \psi) d s  \tag{2.14}\\
\int_{K_{j}} \mu \frac{\partial \mathbf{H}}{\partial t} \cdot \phi d x= & -\int_{K_{j}} \nabla \times \mathbf{E} \cdot \phi d x+\gamma \int_{\partial K_{j}} \llbracket \mathbf{E} \times \mathbf{n} \rrbracket_{\partial K_{j}}^{K_{j}} \cdot \phi d s \\
& -\delta \int_{\partial K_{j}} \llbracket \mathbf{n} \times \mathbf{H} \rrbracket_{\partial K_{j}}^{K_{j}} \cdot(\mathbf{n} \times \phi) d s \tag{2.15}
\end{align*}
$$

where $\alpha, \beta, \gamma, \delta$ are real constants, $\phi \in \vec{W}_{h}^{r}, \psi \in \vec{W}_{h}^{r}$ and $\llbracket \mathbf{v} \rrbracket_{\partial K_{j}}^{K_{j}}$ is the jump of a vector-valued function $\mathbf{v}$ across the boundary of $K_{j}$.

Basis functions of $\vec{W}_{h}^{r}$ are constructed exactly in the same way as in Section 2.5. Now, the discontinuous character of these basis functions allows us to use either Gauss-Lobatto or Gauss points as interpolation and quadrature points. A dispersion analysis of the schemes obtain by using each set of points shows a better accuracy for Gauss points which actually provide an exact value of the integrals. However, since some of the points are located on the faces of the unit cube, Gauss-Lobatto points give an immediate definition of the jumps whereas Gauss points require an extrapolation. However, due to orthogonality properties, this extrapolation is 1 D and induces a reasonable additional cost $[9,26]$.

### 2.6.2. Discrete energy analysis

If we compute the discrete energy of (2.14)-(2.15) (when $\sigma=0$ and $\vec{J}=0$ )

$$
\mathcal{E}(t)=\sum_{K_{i} \subset \Omega}\left\{\int_{K_{i}}\left(\underline{\underline{\epsilon}} \vec{E}_{K_{i}}\right) \cdot \vec{E}_{K_{i}} d x+\int_{K_{i}}\left(\underline{\mu} \vec{H}_{K_{i}}\right) \cdot \vec{H}_{K_{i}} d x\right\}
$$

one can show that, for $-\alpha=\gamma=\frac{1}{2}, \alpha \geq 0$ and $\delta \geq 0$, we get

$$
\begin{equation*}
\frac{\mathrm{d} \mathcal{E}}{\mathrm{~d} t}(t)=\sum_{K_{i} \cap K_{j}}\left\{-\beta\left\|\llbracket \vec{n} \times \vec{E} \rrbracket_{K_{i} \cap K_{j}}\right\|_{\Gamma}^{2}-\delta\left\|\llbracket \vec{n} \times \vec{H} \rrbracket_{K_{i} \cap K_{j}}\right\|_{\Gamma}^{2}\right\}, \tag{2.16}
\end{equation*}
$$

which provides a decreasing energy when $\beta>0$ and $\delta>0$ and so, a dissipative approximation of the problem. On the other hand, if we set $\beta=0$ and $\delta=0$, we get a conservative approximation.

Now, the discrete form of this approximation with a leapfrog scheme in time reads

$$
\begin{align*}
& B_{\varepsilon} \frac{\vec{E}^{n+1}-\vec{E}^{n}}{\Delta t}+B_{\sigma} \frac{\vec{E}^{n+1}+\vec{E}^{n}}{2}-R_{h} \vec{H}^{n+1 / 2}+\alpha S_{h} \vec{H}^{n+1 / 2}-\beta D_{h} \vec{E}^{n}+\vec{J}^{n+1 / 2}=0,  \tag{2.17}\\
& B_{\mu} \frac{\vec{H}^{n+1 / 2}-\vec{H}^{n-1 / 2}}{\Delta t}+R_{h} \vec{E}^{n}+\gamma S_{h}^{*} \vec{E}^{n}-\delta D_{h}^{*} \vec{H}^{n-1 / 2}=0 \tag{2.18}
\end{align*}
$$

where $B_{\varepsilon}, B_{\sigma}$ and $B_{\mu}$ are $3 \times 3$ block diagonal mass matrices, $R_{h}$ is a stiffness matrix which needs to be stored on the unit element, $S_{h}$ is a jump matrix which needs to be stored on the


Fig. 2.4. Splitting a tetrahedron into four hexahedra.
unit element and $D_{h}$ is a jump matrix which needs to be stored (its storage is equivalent to that of a mass matrix). Moreover, the jump terms in which $D_{h}$ appear need to be uncentered in order to keep the scheme explicit, which introduces some numerical dissipation. However, this is not too much troublesome since these two terms are themselves dissipative. So, in order to reduce the storage and to avoid numerical dissipation, it seems a priori better to set $\beta=\delta=0$. This was our first choice, as described in details in [9, 26].

## Remarks:

1. Plane wave analysis and error estimates can be found in [26]. Error analysis shows that, due to the presence of $D F_{i}$ which is bilinear on non-regular meshes, a $Q_{r}$ approximation is of order $r-1$. In particular, this implies that the first-order approximation is not consistent on non-regular meshes. This theoretical result was confirmed by numerical experiments.
2. For Gauss-Lobatto rules, one can center the dissipative terms and remain explicit. In particular, this property ensures to keep the same stability condition for any value of $\beta$ and $\delta$. For Gauss rules, the uncentered character of these terms induces a loss of stability ( $\simeq 10 \%$ for $\beta=0.1$ ).

### 2.7. Numerical issues

Both DGM and the second family of edge elements on slightly non-regular meshes provided performant results in 2D and 3D in terms of storage and CPU time. In particular comparisons with FDTD showed its efficiency [8]. Such meshes are easy to construct in 2D but not always possible to get in 3D. In fact, for complex geometries, it is difficult to obtain a purely hexahedral mesh. A palliative to this difficulty is to produce a hexahedral mesh from a tetrahedral mesh by splitting each tetrahedron into four hexahedra, as shown in Fig. 2.4. Unfortunately, this technique leads to very distorted meshes which produces very important parasitic waves.

## 3. Eigenvalue Analysis

Since one can write a spectral decomposition the solution of Maxwell's equations, the study of the eigenvalues problem related to the Maxwell's equations in their continuous or discrete form can exhibit possible parasitic modes.

In [14], the problem of spurious modes for nodal elements is dealt. The author suggests a weighted regularization to avoid these spurious modes. In [1], the second family is shown as spectrally non-correct in the case of rectangular meshes. The authors proved that "hp" elements based on the first family for rectangular meshes provide a spurious-free method. Buffa and Perugia [2] treat the Maxwell eigenvalue for several Discontinuous Galerkin formulations: interior penalty IP, non-symmetric interior penalty NIP and Local Discontinuous Galerkin (LDG), which is the formulation we use. The authors conclude that the formulation is spurious free for first and second family on tetrahedra and for first family on hexahedra. Nevertheless, they think that the method has spurious modes for the second family. Our study will focus on the second family, in order to give a new enlightenment.


Fig. 3.1. Test meshed cavities in 2D and 3D.

Our eigenvalue analysis is developed on two test cases which are unit square and cubic cavities meshed by using split triangles or tetrahedra as shown in Fig. 3.1 with a $Q_{3}$ approximation. For these two cavities, we compute the eigenvalues in the range of frequencies contained in the solution.

### 3.1. Discontinuous Galerkin methods

For DGM without numerical dissipation (i.e. $\beta=\delta=0$ ), the behaviour of the eigenvalues is somehow strange. In 2 D , we only get one parasitic eigenvalue but a huge amount of such eigenvalues appear in 3D (Fig. 3.2). If we compare with the second family of edge elements (Fig. 3.6), we get about 25 times more parasitic eigenvalues in 3D!

An efficient way to get rid of these parasitic eigenvalues is to introduce some numerical dissipation (i.e. to set $\beta>0$ or/and $\delta>0$ ). In this way, the parasitic eigenvalues become all complex with a negative imaginary part. This last property implies that all the parasitic waves become evanescent. In Fig. 3.3, we show the eigenvalues obtained by setting $\beta=0.1$


Fig. 3.2. Eigenvalues for the 2D (left) and 3D (right) cavities computed with DGM. Physical eigenvalues are indicated by red lines.


Fig. 3.3. Eigenvalues in the complex plane for $\beta=0.1$ (left) and $\beta=0.5$ (right) computed with DGM.
then $=0.5$ and $\delta=0$ in 3D. One can notice that for both values of the parameter, a lot of parasitic eigenvalues are not too much affected by the change of parameter.

The effect of numerical dissipation on the solution can be seen in Fig. 3.4, in which we show snapshots of the solution (on the $x-y$ plane) in a unit cubic cavity after a wave travel of 30 wavelengths. An interesting question is "how the overall accuracy on the solution depends on the penalization parameter $\beta$ ?". In order to answer to this question, we compute the solution, in time-harmonic domain, in a cubic cavity meshed by tetrahedra split into hexahedra, for different values of $\beta$. The results are displayed in Fig. 3.5. As the reader can see, the error is large when $\beta=0$, and for $\beta>0.1$, the error is quite small. An "optimal value" can be seen around $\beta=2$, but the accuracy is not very sensitive to the variations of $\beta$.

### 3.2. Second family of edge elements

In the 2 D and the 3D case, edge elements present a substantial number of parasitic eigenvalues, even more important in 3D, as shown in Fig. 3.4. These parasitic eigenvalues confirm the


Fig. 3.4. Snapshots of the solution (on the $x-y$ plane) in unit cubic cavity after a wave travel of 30 wavelengths for $\beta=0$ (left) and $\beta=0.1$ (right) computed with DGM.
presence of parasitic waves in the solution. A priori, these parasitic eigenvalues seem to make this family of edge elements inefficient. However, this problem can be overcome in a similar way as DGM, by adding some numerical dissipation to this approximation. These dissipative terms can be introduced in (2.10a)-(2.10b) as follows

$$
\begin{align*}
\sum_{K_{j} \in \mathcal{Q}}\left\{\frac{\mathrm{~d}}{\mathrm{~d} t} \int_{K_{j}} \underline{\varepsilon}_{\underline{\varepsilon}}^{\underline{E}} \cdot \vec{\varphi}_{h} \mathrm{~d} \vec{x}-\int_{K_{j}} \vec{H}_{h} \cdot \nabla \times \vec{\varphi}_{h} \mathrm{~d} \vec{x}+\int_{K_{j}} \underline{\underline{\sigma}} \vec{E}_{h} \cdot \vec{\varphi}_{h} \mathrm{~d} \vec{x}\right. \\
\left.\quad-\beta^{\prime} \int_{\partial K_{j}} \llbracket \mathbf{E} \cdot \mathbf{n} \rrbracket_{\partial K_{j}}^{K_{j}}(\psi \cdot \mathbf{n}) d s+\int_{K_{j}} \vec{J} \cdot \vec{\varphi}_{h} \mathrm{~d} \vec{x}\right\}=0, \quad \forall \vec{\varphi}_{h} \in \vec{V}_{h}^{r}  \tag{3.1a}\\
\sum_{K_{j} \in \mathcal{Q}}\left\{\frac{\mathrm{~d}}{\mathrm{~d} t} \int_{K_{j}}{ }_{=}^{\mu \vec{H}_{h} \cdot \vec{\psi}_{h} \mathrm{~d} \vec{x}+\int_{K_{j}} \nabla \times \vec{E}_{h} \cdot \vec{\psi}_{h} \mathrm{~d} \vec{x}}\right. \\
\left.\quad-\delta^{\prime} \int_{\partial K_{j}} \llbracket \mathbf{n} \times \mathbf{H} \rrbracket_{\partial K_{j}}^{K_{j}} \cdot\left(\mathbf{n} \times \vec{\psi}_{h}\right) d s\right\}=0, \quad \forall \vec{\psi}_{h} \in \vec{W}_{h}^{r} \tag{3.1b}
\end{align*}
$$

As for (2.14)-(2.15), one can write the following energy identity:

$$
\begin{equation*}
\frac{\mathrm{d} \mathcal{E}}{\mathrm{~d} t}(t)=\sum_{K_{i} \cap K_{j}}\left\{-\beta^{\prime}\left\|\llbracket \vec{n} \cdot \vec{E} \rrbracket_{K_{i} \cap K_{j}}\right\|_{\Gamma}^{2}-\delta^{\prime}\left\|\llbracket \vec{n} \times \vec{H} \rrbracket_{K_{i} \cap K_{j}}\right\|_{\Gamma}^{2}\right\} \tag{3.2}
\end{equation*}
$$

The use of the above dissipative terms provides the same effect as for DGM on the eigenvalues as shown in Figs. 3.7 and 3.8.

## Remarks:

1. The dissipative term in (3.1a) is not equal to zero because only the tangent components of $\vec{E}$ are continuous.
2. The dissipative term in (3.1b) is similar to this of DGM but was not yet tested.
3. Both for DGM and edge elements, the dissipation was tested in cavities for very long times (even for time-harmonic problems) for $Q_{3}$ and higher-order approximations. For all these experiments, we did not notice substantial effect on the amplitudes of the solutions.


Fig. 3.5. Relative error made on the numerical solution according to the parameter $\beta$, for a Q3 approximation (left) and a Q5 approximation (right).


Fig. 3.6. Eigenvalues for the 2D (left) and 3D (right) cavities computed with edge elements. Physical eigenvalues are indicated by red lines.

### 3.3. The case of the first family

The first family of edge elements does not provide mass-lumping on non-orthogonal meshes. However, this family is very useful for time-harmonic problem, as we shall see in the next section. Its efficiency is based on a fast matrix-vector product but also on the fact that, for some quadrature rules, this family is free of spurious modes

In fact, due to the structure of its degrees of freedom, both Gauss and Gauss-Lobatto quadrature rules are suitable to this family. For the mass matrices, Gauss points lead to non invertible mass matrices whereas Gauss-Lobatto quadrature rules behave well. The stiffness matrices can also be computed by both quadrature rules. However, as shown in Fig. 3.9, the Gauss-Lobatto rule leads to a spurious-free approximation.

## 4. Time-Harmonic Problems

In this section, we only study the first family, which is spurious free and which needs less storage than other methods to reach the same accuracy.


Fig. 3.7. Eigenvalues in the complex plane for $\beta^{\prime}=0.1$ (left) and $\beta^{\prime}=0.5$ (right) and $\delta^{\prime}=0$ computed with the second family of edge elements.


Fig. 3.8. Snapshots of the solution (on the $x-y$ plane) in unit cubic cavity after a wave travel of 30 wavelengths for $\beta^{\prime}=0$ (left) and $\beta^{\prime}=0.1$ (right) and $\delta^{\prime}=0$ computed with the second family of edge elements.

### 4.1. Fast matrix-vector product

After applying the Fourier transform in time and discretizing the variational formulation (2.5), we get the following linear system:

$$
\begin{equation*}
-\omega^{2} M_{h} E_{h}+K_{h} E_{h}=F_{h} \tag{4.1}
\end{equation*}
$$

The mass matrix $M_{h}$ and stiffness matrix $K_{h}$ can be written as

$$
\begin{aligned}
\left(M_{h}\right)_{j, k} & =\int_{\Omega} \varepsilon \varphi_{j} \cdot \varphi_{k}, \\
\left(K_{h}\right)_{j, k} & =\int_{\Omega} \mu^{-1} \nabla \times \varphi_{j} \cdot \nabla \times \varphi_{k}
\end{aligned}
$$



Fig. 3.9. Eigenvalues for the 3D first family of edge elements with Gauss (left) and Gauss-Lobatto (right) quadrature rules for the stiffness matrix. Physical eigenvalues are indicated by red lines.

After a change of variables, the integrands are computed over the unit cube $\hat{K}$ as follows

$$
\begin{aligned}
\left(M_{h}\right)_{j, k} & =\int_{\hat{K}} J_{i} D F_{i}^{-1} \varepsilon D F_{i}^{-t} \hat{\varphi}_{j} \cdot \hat{\varphi}_{k}, \\
\left(K_{h}\right)_{j, k} & =\int_{\hat{K}} \frac{1}{J_{i}} D F_{i}^{*} \mu^{-1} D F_{i} \hat{\nabla} \times \hat{\varphi}_{j} \cdot \hat{\nabla} \times \hat{\varphi}_{k} .
\end{aligned}
$$

We have supposed here that Jacobian $J_{i}$ is always positive. This can be obtained by a permutation on the local numbers of the vertices in each hexahedron.

These two matrices are computed by using Gauss-Lobatto quadrature rules. We need then to compute the following $3 \times 3$ symmetric matrices $\left(A_{h}\right)_{k, k}$ and $\left(B_{h}\right)_{k, k}$ at each quadrature point $\xi_{k}^{G L}$ :

$$
\begin{aligned}
\left(A_{h}\right)_{k, k} & =\omega_{k}^{G L} J_{i} D F_{i}^{-1} \varepsilon D F_{i}^{-t}\left(\hat{\xi}_{k}^{G L}\right), \\
\left(B_{h}\right)_{k, k} & =\frac{\omega_{k}^{G L}}{J_{i}} D F_{i}^{*} \mu^{-1} D F_{i}\left(\hat{\xi}_{k}^{G L}\right) .
\end{aligned}
$$

Matrices $A_{h}$ and $B_{h}$ are block-diagonal matrices, with $3 \times 3$ symmetric blocks. An "interpolation" operator $\hat{C}$ is defined by

$$
\hat{C}_{j, k}=\hat{\varphi}_{j}\left(\hat{\xi}_{k}^{G L}\right)
$$

In the same way, a discrete "curl" $\hat{R}$ operator is defined by

$$
\hat{R}_{j, k}=\hat{\nabla} \times \hat{\varphi}_{j}^{G L}\left(\hat{\xi}_{k}^{G L}\right)
$$

where functions $\hat{\varphi}_{j}^{G L}$ are basis functions associated to Gauss-Lobatto points. The reader can notice that matrix $\hat{R}$ is the same as the stiffness matrix introduced for the mixed formulation.

Theorem 4.1. With the previous notations, we get the following factorizations:

$$
M_{h}=\hat{C} A_{h} \hat{C}^{*}, \quad K_{h}=\hat{C} \hat{R} B_{h} \hat{R}^{*} \hat{C}^{*}
$$

Proof. Let us write the elementary mass matrix by using quadrature rules

$$
\left(M_{h}\right)_{j, k}=\sum_{m=1}^{(r+1)^{3}}\left(A_{h}\right)_{m, m} \hat{\varphi}_{j}\left(\hat{\xi}_{m}^{G L}\right) \cdot \hat{\varphi}_{k}\left(\hat{\xi}_{m}^{G L}\right)
$$

If we denote $F_{h}$ the matrix such that

$$
\left(F_{h}\right)_{j, m}=\left(A_{h}\right)_{m, m} \hat{\varphi}_{j}\left(\hat{\xi}_{m}^{G L}\right),
$$

we have a classical matrix-matrix product

$$
M_{h}=F_{h} \hat{C}^{*}
$$

Matrix $F_{h}$ can be easily decomposed as

$$
F_{h}=\hat{C} A_{h} .
$$

The factorization of the elementary mass matrix is proven. In the same way, we write the elementary stiffness matrix

$$
\left(K_{h}\right)_{j, k}=\sum_{m=1}^{(r+1)^{3}}\left(B_{h}\right)_{m, m} \hat{\nabla} \times \hat{\varphi}_{j}\left(\hat{\xi}_{m}^{G L}\right) \cdot \hat{\nabla} \times \hat{\varphi}_{k}\left(\hat{\xi}_{m}^{G L}\right) .
$$

As for the mass matrix, we have the following factorization

$$
K_{h}=\hat{S} B_{h} \hat{S}^{*}
$$

where matrix $\hat{S}$ is defined by

$$
\hat{S}=\nabla \times \hat{\varphi}_{j}\left(\hat{\xi}_{k}^{G L}\right)
$$

$\hat{S}$ is different from $\hat{R}$, because we take basis functions of the first family instead of basis functions associated to Gauss-Lobatto points. But we claim the identity (4.2):

$$
\begin{equation*}
\hat{S}=\hat{C} \hat{R} \tag{4.2}
\end{equation*}
$$

Because of polynomial inclusion $\left(Q_{r-1, r, r} \times Q_{r, r-1, r} \times Q_{r, r, r-1} \subset Q_{r}^{3}\right)$, we have

$$
\hat{\varphi}_{j}(\hat{x})=\sum_{m=1}^{3(r+1)^{3}} \hat{C}_{j, m} \hat{\varphi}_{m}^{G L}(\hat{x})
$$

By taking the curl of this expression, and for all $\hat{x}$ equal to a Gauss-Lobatto point, we get

$$
\hat{\nabla} \times \hat{\varphi}_{j}\left(\hat{\xi}_{k}^{G L}\right)=\sum_{m=1}^{3(r+1)^{3}} \hat{C}_{j, m} \hat{\nabla} \times \hat{\varphi}_{m}^{G L}\left(\hat{\xi}_{k}^{G L}\right) .
$$

By definition of the operators, this can be rewritten as

$$
\hat{S}_{j, k}=\sum_{m=1}^{3(r+1)^{3}} \hat{C}_{j, m} \hat{R}_{m, k} .
$$

Identity (4.2) is proven, as well as the factorization of the elementary stiffness matrix.
Complexity of the matrix-vector product $\left(-\omega^{2} M_{h}+K_{h}\right) X_{h}$ is in $\mathcal{O}\left(r^{4}\right)$. Moreover, the number of operations is equal to : $\left((r+1)^{3}[24(r+1)+12 r+30]\right) N_{e}$, where $N_{e}$ represents the number of hexahedra of the mesh. The required storage is equal to $12(r+1)^{3} N_{e}$ coefficients,


Fig. 4.1. At left, computational time according to the order of approximation, at right storage required. A standard matrix-vector product (by storing the full matrix) and the fast matrix-vector are compared.
which is equivalent to four vectors if $r$ is large enough. In order to compare the different orders of approximation, we compute the two quantities:

$$
\frac{\text { Computational cost }}{\text { Number of d.o.f.s }}, \quad \frac{\text { Storage cost }}{\text { Number of d.o.f.s }}
$$

These quantities are displayed in Fig. 4.1 As we can see in this figure, the fast matrix-vector product is more efficient than standard matrix-vector product for order greater or equal to 3 , and also leads to a lower storage than the standard method for order greater or equal to 2 .

Remark: The factorization is also true if an almost-"exact" integration is used ( $k+1$ Gauss points instead of $k+1$ Gauss-Lobatto points). In this case, $\hat{R}$ has the same sparsity pattern, but $\hat{C}$ is a full matrix. Fortunately, thanks to tensorization of basis functions on the hexahedron, the triple sum induced by the matrix-vector product $\hat{C} X$ can be decomposed into three single sums. By doing that, we use an implicit factorization: $\hat{C}=\hat{C}_{1} \hat{C}_{2} \hat{C}_{3}$, where the intermediaries matrices $\hat{C}_{1}, \hat{C}_{2}$ and $\hat{C}_{3}$ are sparse. A fast matrix-vector is obtained, but it is slower than the fast matrix-vector product obtained with approximate integration. More precisely, it is 67 $\%$ slower, if we use exact integration with no improvement in accuracy. This difference was confirmed numerically.

### 4.2. Preconditioning technique

The iterative solver used is an iterative solver which only works for complex symmetric matrices. It is called BICGCR (BIConjugate Gradient Conjugate Residual method) and is described in [4]. Numerical experiments show the superiority of this solver over classical COCG, GMRES or BICGSTAB [15]. Nevertheless, the number of iterations can be very large for high frequency problems with heterogeneities, and for fine meshes. In order to be efficient for the linear system, this basic iterative solver needs to be accelerated by a preconditioning technique.

The main idea of the preconditioning technique used in this article, is to consider damped Maxwell equations by taking $\varepsilon^{\prime}$ complex of the form

$$
\varepsilon^{\prime}=\varepsilon(\theta+\mathrm{i} \eta)
$$

Table 4.1: Number of iterations of BICGCR for a perfectly conducting sphere, and memory used for factorization. The preconditioning used is incomplete factorization on the finite element matrix $Q_{1}$.

| Threshold | $1 e-4$ | $1 e-3$ | 0.01 | 0.05 | 0.08 | 0.1 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $\theta=1, \eta=0$ | $30 / 370 \mathrm{Mo}$ | $\infty / 350 \mathrm{Mo}$ | $\infty / 340 \mathrm{Mo}$ | $\infty / 326 \mathrm{Mo}$ | $\infty / 318 \mathrm{Mo}$ | $\infty / 314 \mathrm{Mo}$ |
| $\theta=1, \eta=0.5$ | $55 / 299 \mathrm{Mo}$ | $55 / 242 \mathrm{Mo}$ | $55 / 149 \mathrm{Mo}$ | $82 / 74 \mathrm{Mo}$ | $116 / 55 \mathrm{Mo}$ | $145 / 47 \mathrm{Mo}$ |
| $\theta=1, \eta=1$ | $97 / 244 \mathrm{Mo}$ | $97 / 197 \mathrm{Mo}$ | $99 / 108 \mathrm{Mo}$ | $110 / 53 \mathrm{Mo}$ | $133 / 40 \mathrm{Mo}$ | $155 / 34 \mathrm{Mo}$ |



Fig. 4.2. Subdivided $Q_{1}$ mesh of a elementary quadrilateral $Q_{2}$ transformation $F_{i}$ to go from unit square to a small square of the subdivided mesh.


Fig. 4.3. At left, time for 1000 iterations of COCG on a test-case of 200000 d.o.f.s. At right, storage required before iterative solving (iterations vectors are not included).

Damping is obtained under the assumption $\eta>0$. The preconditioner is constructed on the damped Maxwell equations, while non-damped Maxwell equations are solved.

The effect of damping is to "localize" the inverse of the matrix so that the incomplete factorization works fine. In Table 4.1, we put the number of iterations of BICGCR preconditioned by ILUT, according to $\eta$ and different values of the threshold. The incomplete factorization is called ILUT [27], but we only store the $U$ factor, considering that the $L$ factor can be recovered by symmetry. We can see that, if you don't use damping, the incomplete factorization crashes very fast, and there is no gain in storage. The gain in storage increases with damping, but also the number of iterations. The parameters chosen will be equal to $(1,1)$ for all other numerical experiments of the paper.

In order to have a low storage, we use a subdivided a $Q 1$ mesh to compute a $Q 1$ matrix. The procedure can be summarized by Fig. 4.2. There is an exact matching between degrees


Fig. 5.1. The dielectric spherical torus.
of freedom (d.o.f.s) of the high-order mesh and d.o.f.s of low-order provided a diagonal scaling. This diagonal scaling is due to the transform $D F_{i}^{*-1}$ incorporated in the definition of the basis functions.

### 4.3. Cost of the matrix-vector product. Comparison with tetrahedra

In this section, we numerically compare hexahedral and tetrahedral elements for the first family. The results are summarized in Fig. 4.3. The mesh and the matrix are main components of the storage requirement. For $Q_{1}$ and $R_{1}$, the mesh represents the main part. As we can see, the storage needed decreases for hexahedra, whereas it dramatically increases for tetrahedra. The hexahedral elements provide a matrix-vector product which becomes faster than for tetrahedral elements for an order of approximation greater or equal to 3 . In practice, the use of high order for tetrahedra of the first family is quite difficult because of this high storage.

## 5. Numerical Results

### 5.1. A dielectric spherical torus

In this experiment, we treat a metallic sphere coated by a dielectric layer (Fig. 5.1) in the time domain. The radius of the sphere is equal to 0.5 m whereas the width of the dielectric is 0.25 m . The relative permittivity $\varepsilon_{r}$ in the dielectric is equal to 10 . This sphere is illuminated by a Gaussian plane wave of frequency 0.5 GHz .

In Fig. 5.2, we give the comparison of the value of the electric field at the point $(-0.76,0)$ computed by FDTD (10 and 20 points per wavelength) and our method with a mesh composed of 16984 hexahedra (obtained by split tetrahedra adapted to the wavelength) for $Q_{3}$ and $Q_{5}$ approximations. The FDTD with 20 points/ $\lambda$ requires 1100 s of CPU versus 300 s for the $Q_{3}$ DG approximation and 10 times more storage for a very bad accuracy in long times.

### 5.2. The Cobra cavity

The scattering of a cobra cavity in the time-harmonic domain is studied (cf. Fig. 5.3). Its length is 3 m and a plane wave of frequency 1 GHz illuminates its open end. If we want to

Table 5.1: Efficiency of first family on cobra cavity.

| Order | Number d.o.f.s | Error on the RCS | BICGCR | ILUT(0.05) |
| :---: | :---: | :---: | :---: | :---: |
| $Q_{4}$ | 412000 | 0.45 dB | 14039 s | 2247 |
|  |  |  | s |  |
|  |  |  | $3400(47 \mathrm{Mo})$ | $1900(391 \quad \mathrm{Mo})$ |
| $Q_{6}$ | 187000 | 0.4 dB | 12096 s | 846 |
|  |  |  | $41500(22 \mathrm{Mo})$ | $1700(161 \quad \mathrm{Mo})$ |



Fig. 5.2. $E_{y}$ component of the electric field at a point of the domain after propagation across 10 wavelengths (left) and 120 wavelengths (right).


Eqn: $+0.013 x-1 y-0.072 z+0.72=0$


Fig. 5.3. At left mesh of the cavity, at right real part of $x$-component of electric field.
obtain an error less than 0.5 dB for the RCS (see Fig. 5.4), we get results of Table 5.1. For a $Q_{6}$ approximation, we need 187000 d.o.f.s, the computational time is 846 s, if we use the ILUT preconditioner. The iterative solver converged in 1700 iterations and needs 161 Mo of storage.


Fig. 5.4. Radar cross section for the cobra cavity.


Fig. 5.5. The surfacic mesh of the plane (before splitting into hexahedra) (left) and the final snapshot (right).

In this case, we used curved hexahedral elements, in order to have a good approximation of the geometry. For the definition of a curved $F_{i}$, we refer to the book of Solin et al. [28].

### 5.3. Scattering by an airplane

This last experiment illustrates the possibilities of the methods in terms of storage and CPU time. We consider an airplane of length 11 m in the time domain, illuminated by a timeharmonic plane wave of frequency 0.75 GHz . The whole mesh (obtained by split tetrahedra) contains about 78000 elements and a $Q_{4}$ DG approximation was used, which leads to about 30000000 of degrees of freedom. A local time-stepping technique was used and the dissipation coefficient $\beta$ defined in (2.14) is equal to 0.1 . This experiment requires a storage of 0.8 Go and the wave used 25 hours of CPU on a monoprocessor computer ( 2.4 Go and 3.2 GHz ) to reach the end of the plane. In Fig. 5.5, we show the triangular surfacic mesh (before splitting) and a snapshot at the final time. We validated the accuracy of the solution by a comparison with a $Q_{5}$ approximation on the same mesh.

## 6. Conclusion

Mixed spectral approximations provide fast and low-storage algorithms both in the time domain and the time-harmonic domain, provided some numerical dissipation. Their advantage over tetrahedral methods is obvious. On the other hand, time domain approximation can efficiently compete with FDTD methods. However, the use of split tetrahedra for meshing remains a handicap for the performance of the method. Some alternative meshing strategies are under study.

## References

[1] D. Boffi, M. Costabel, M. Dauge and L. Demkowicz, Discrete compactness for the hp version of rectangular edge finite elements, SIAM J. Numer. Anal., 44:3 (2006), 979-1004.
[2] A. Buffa and I. Perugia, Discontinuous Galerkin approximation of the Maxwell eigenproblem, SIAM J. Numer. Anal., to appear.
[3] P.G. Ciarlet, The Finite Element Method for Elliptic Problems, North-Holland, 1987.
[4] M. Clemens and T. Weiland, Iterative methods for the solution of very large complex symmetric linear systems of equations in electrodynamics, Technische Hochschule Darmstadt, 2002.
[5] G. Cohen, Higher Order Numerical Methods for Transient Wave Equations, Springer, 2001.
[6] G. Cohen and S. Fauqueux, Mixed finite elements with mass-lumping for the transient wave equation, J. Comput. Acoust., 8:1 (2000), 171-188.
[7] G. Cohen and S. Fauqueux, Mixed spectral finite elements for the linear elasticity system in unbounded domains, SIAM J. Sci. Comput., 26:3 (2005), 864-884.
[8] G. Cohen, X. Ferrieres and S Pernet, High spatial order finite element method to solve Maxwell's equations in time domain, IEEE T. Antenn. Propag., 53:9 (2005), 2889-2899.
[9] G. Cohen, X. Ferrieres and S. Pernet, A spatial high-order hexahedral discontinuous Galerkin methods to solve Maxwell's equations in time domain, J. Comput. Phys., 217 (2006), 340-363.
[10] G. Cohen, P. Joly, J.E. Roberts and N. Tordjman, Higher order triangular finite elements with mass lumping for the wave equation, SIAM J. Numer. Anal., 38:6 (2001), 2047-2078.
[11] G. Cohen, P. Joly and N. Tordjman, Higher-order finite elements with mass lumping for the 1-D wave equation, Finite Elem. Anal. Des., 17:3-4 (1994), 329-336.
[12] G. Cohen and P. Monk, Gauss point mass lumping schemes for Maxwell's equations, Numer. Meth. Part. Diff. Eq., 14:1 (1998), 63-88.
[13] G. Cohen and P. Monk, Mur-Nédélec finite element schemes for Maxwell's equations, Comput. Method. Appl. Mech. Engrg., 169:3-4 (1999), 197-217.
[14] M. Costabel and M. Dauge, Computation of Resonance Frequencies for Maxwell Equations in Non Smooth Domains, Lecture Notes in Computational Science and Engineering, Vol. 31, Springer, 2003.
[15] M. Duruflé, Integration numérique et éléments finis d'ordre élevé appliqués aux équations de Maxwell en régime harmonique, PhD thesis, U. de Paris-Dauphine, 2006.
[16] A. Elmkies and P. Joly, Éléments finis d'arête et condensation de masse pour les équations de Maxwell: le cas 2D, C.R.A.S., Math., 324, série I, 1997, 1287-1293.
[17] A. Elmkies and P. Joly, Éléments finis d'arête et condensation de masse pour les équations de Maxwell: le cas de la dimension 3, C.R.A.S., Math., 325, série I, 1997, 1217-1222.
[18] Y. Haugazeau and P. Lacoste, Condensation de la matrice masse pour les éléments finis mixtes de H(rot), C.R. Acad. Sci. Paris Sér. I-Math., 316:5 (1993), 509-512.
[19] J.-P. Hennart, E. Sainz and M. Villegas, On the efficient use of the finite element method in static neutron diffusion calculations, Comput. Method. Nuclear Engrg., 1 (1979), 3-87.
[20] J.-P. Hennart, Topics in Finite Element Discretization of Parabolic Evolution Problems, Lecture Notes in Mathematics, Springer-Verlag, 909, 1982, pp. 185-199.
[21] J.S. Hesthaven and T. Warburton, High-order nodal methods on unstructured grids. I. Timedomain solution of Maxwell's equations, J. Comput. Phys., 181 (2002), 1-34.
[22] Y. Maday and A.T. Patera, Spectral element methods for the incompressible Navier-Stokes equations, in State of the Art Survey in Computational Mechanics, A.K. Noor (Ed.), 1989, pp. 71-143.
[23] W.A. Mulder, Higher-order mass-lumped finite elements for the wave equation, J. Comput. Acoust., 9:2 (2001), 671-680.
[24] J.-C. Nédélec, Mixed finite elements in $\mathbb{R}^{3}$, Numer. Math., 35:3 (1980), 315-341.
[25] J.-C. Nédélec, A new family of mixed finite elements in $\mathbb{R}^{3}$, Numer. Math., 50:1 (1986), 57-81.
[26] S. Pernet, Étude de méthodes d'ordre élevé pour résoudre les équations de Maxwell dans le domaine temporel. Application à la détection et à la compatibilité électromagnétique, PhD thesis, U. de Paris-Dauphine, 2004.
[27] Y. Saad, Iterative Methods for Sparse Linear Systems, Series in Computer Science, SIAM, Philadelphia, 1996.
[28] P. Solin, K. Segeth and I. Dolezel, Higher-Order Finite Elements Methods, Studies in Advanced Mathematics, Chapman and Hall, 2003.
[29] K. Yee, Numerical solutions of initial boundary value problems involving Maxwell's equations in isotropic media, IEEE T. Antenn. Propag., 16 (1966), 302-307.
[30] L.C. Young, An efficient finite element method for reservoir simulation, Proc. of the 53rd Annual Fall Technical Conference and Exhibition of the Society of Petroleum Engineers of AIME, Houston, Texas Oct. 1-3, 1978.


[^0]:    ${ }^{*}$ Received November 2, 2006; Final revised February 7, 2007.

