# Post processing for the vector finite element method: accurate computing of dual field

Christian Vollaire, François Musy, Ronan Perrussel

# Post Processing for the Vector Finite Element Method: Accurate Computing of Dual Field

Christian Vollaire[1], François Musy[2], and Ronan Perrussel[1]

[1]Centre de Génie Electrique de Lyon (CEGELY—UMR CNRS 5005), Ecole Centrale de Lyon, 69134 Ecully Cedex, France
[2]Institut Camille Jordan (ICJ—UMR CNRS 5208), Ecole Centrale de Lyon, 69134 Ecully Cedex, France

**An accurate method to compute dual field in high frequency time harmonic problem is presented. From a primal field obtained by a vector finite element discretization, the dual field is obtained without numerical derivation by using a least square argument. The accuracy of the method is compared with the natural method using shape function derivatives.**

***Index Terms*—Edge elements, finite element method, post-processing.**

## I. Introduction

EDGE finite elements (FE) are useful in modeling electromagnetic phenomena because of their right physical sense. Furthermore, it has been shown that a better approximation of the solution may be obtained compared to the nodal-based FE [1]–[3]. The incomplete first-order edge FE is commonly used in high frequency time-harmonic electromagnetic modeling. The degrees of freedom are path integrals along the edges in the mesh. The first-order elements are a good compromise between the precision and the computational time, in particular the time dedicated to the resolution of the matrix system by an iterative solver. However, the knowledge of the nodal field values remains necessary for various reasons. Maximal values located at interfaces can be required to predict possible electric breakdowns. Nodal values may be necessary to achieve some additional computation: induced currents in the conductors, source term for coupled problem (magnetothermal). Postprocessors for visualization are usually based on the nodal representation of the fields. For vector finite elements, two techniques are commonly used to compute nodal values from degrees of freedom.

- Only nodal values inside an element are computed from its edge values. However, this method doesn't give a unique value on its boundary, namely on the vertices.
- An average value of the nodal field is evaluated on each node and for each region by taking into account the contribution of all the elements connected to the considered node. Previously, a post-processing technique to obtain an accurate continuous nodal representation of the field has been proposed by the authors [4]. The proposed techniques were based on a least square formulation leading to the resolution of a sparse matrix system. They were compared with success in terms of accuracy and CPU time on a FE formulation for open boundary–frequency domain problems.

The computation with a good precision of the dual field is also an important challenge in the modeling of high frequency problems because it is necessary for the calculation of additional sizes such as: power radiated by an antenna, wall Joule losses in microwave resonant metallic cavities, etc.

In this paper, the computation of the dual field from the circulation of the primal field along the edges of the mesh is investigated. These circulations are given from the edge FE code. Our method allows the computation of the dual field without numerical derivation with a very low additional computation time. Its accuracy is evaluated on the scattering of a plane wave by a perfect electrical conductor (PEC) sphere and a perfect magnetic (PM) one. The natural method using the computation of shape functions derivatives at the mesh nodes is taken as reference.

## II. Method

Let a complex vector field $E_h$ be obtained from some computations with $H(\text{curl};\Omega)$-conforming finite elements on a tetrahedral mesh $\tau_h$ of a bounded domain $\Omega$ of $\mathbb{R}^3$. We are seeking a "good representation" $B^*$ of $\nabla \times E_h$ by a continuous vector field on $\tau_h$. Let us denote $\mathbf{S}_h$ the set of faces in the mesh $\tau_h$ and for each face $S$ in $\mathbf{S}_h$ and $n_S$ a normal unit vector with an arbitrary orientation.

For $S$, a face with an oriented boundary $\partial S$ according to the orientation of $n_S$, the Stokes' theorem writes

$$\int_S \nabla \times E_h \cdot n = \int_{\partial S} E_h \cdot t. \tag{1}$$

Vector field $E_h$ being given by incomplete first-order edge elements, the flux $\alpha_S = \int_S \nabla \times E_h \cdot n_S$ through any face $S$ in $S_h$ can be deduced from the computed values $\int_e E_h \cdot t$ where $e$ are the edges on the boundary of $S$. A representation of $\nabla \times E_h$ can be deduced on the whole mesh by

$$B_h = \sum_{S \in \boldsymbol{S}_h} \alpha_S w_S \tag{2}$$

where $(w_S)_{S\in\mathbf{S}_h}$ is the basis of the $H(\text{div};\Omega)$-conforming FE space of first-order with the flux on each face in $\mathbf{S}_h$ as degrees of freedom. Then a continuous approximation of $B_h$ can be obtained by solving the following minimization problem: Find a vector field $B* \in V_h$ which minimizes for all B in $V_h$

$$\theta \left( \sum_{\mathbf{S}\in S_h} \left| \int_S B_h.n_s - \int_S B.n_s \right|^2 \right) + (1-\theta) \int_\Omega |B_h - B|^2 \tag{3}$$

Digital Object Identifier 10.1109/TMAG.2006.870937

where $V_h$ is the nodal conforming FE space of degree 1 defined on $\boldsymbol{\tau}_h$ and $\theta$ is a weigth such that $0 \leq \theta \leq 1$.

## III. Linear Problem To Solve

The unknown field $B*$ can be decomposed as $B* = \sum_{i=1}^{N_h} \xi_i v_i$ where $(v_i), i = 1$ to $N_h$ (the nodal space dimension) is the nodal vector basis of the FE space $V_h$ and the unknown vector $\xi = (\xi_i), i = 1$ to $N_h$ contains the values of each component of the field $B*$ on the vertices of the mesh $\boldsymbol{\tau}_h$.

The minimization problem (3) is a least squares problem with one solution at least. The components of the solution can be computed from a linear system with a real symmetric matrix $A_\theta$ and a right-hand side $b_\theta$, respectively, defined by

$$A_\theta = \theta A_1 + (1-\theta) A_0 \tag{4}$$

with $(A_0)_{ij} = \int_\Omega v_i.v_j i, j = 1$ to $N_h$

$$(A_1)_{ij} = \sum_{S \in \boldsymbol{S}_h} \left( \int_S v_i.n \right) \left( \int_S v_j.n \right), \quad i,j = 1 \text{ to } N_h \tag{5}$$

and

$$b_\theta = \theta b_1 + (1-\theta) b_0 \tag{6}$$

with $(b_0)_i = \int_\Omega B_h \cdot v_i$

$$(b_1)_i = \sum_{S \in \boldsymbol{S}_h} \alpha_S \left( \int_S v_i.n \right), \qquad i = 1 \text{ to } N_h. \tag{7}$$

For $0 \leq \theta < 1, A_\theta$ is a positive definite matrix and the least squares problem admits an unique solution. The limit case $\theta = 1$ is investigated in Section IV.

Numbering the unknowns according to the three components on the axes, the nodal basis can be decomposed as

$$\begin{bmatrix} \varphi_i \\ 0 \\ 0 \end{bmatrix} i = 1 \text{ to } n, \begin{bmatrix} 0 \\ \varphi_i \\ 0 \end{bmatrix} i = 1 \text{ to } n, \begin{bmatrix} 0 \\ 0 \\ \varphi_i \end{bmatrix} i = 1 \text{ to } n \tag{8}$$

where $\varphi_i, i = 1$ to n is the scalar nodal basis and n the number of vertices. It induces a partition of the matrix $A_\theta$ into a $3 \times 3$ block matrix with 9 $(n \times n)$ blocks and similarly for $b$ with 3 $(n \times 1)$ blocks

$$A_\theta = \begin{bmatrix} A_\theta^{XX} & A_\theta^{XY} & A_\theta^{XZ} \\ A_\theta^{YX} & A_\theta^{YY} & A_\theta^{YZ} \\ A_\theta^{ZX} & A_\theta^{ZY} & A_\theta^{ZZ} \end{bmatrix} b = \begin{bmatrix} b_\theta^X \\ b_\theta^Y \\ b_\theta^Z \end{bmatrix}. \tag{9}$$

From (4) and (8), it is easy to check that $A_0$ is given by

$$A_0 = \begin{pmatrix} M & 0 & 0 \\ 0 & M & 0 \\ 0 & 0 & M \end{pmatrix} \tag{10}$$

where $M$ is the usual $(n \times n)$ mass matrix defined by

$$M_{ij} = \int_\Omega \varphi_i \varphi_j, i, j = 1 \text{ to } n. \tag{11}$$

$A_1$ is not a block diagonal matrix but one sees from (5) and (8) that each block is as sparse as $M$.

In order to compute the non zero entries of $A_1$, it is convenient to introduce the sets of faces associated to each vertex $a_i$ of the mesh:

$$\mathbf{S}_h^i = \{ S \in \mathbf{S}_h / \text{vertex } a_i \text{ belongs to face } S \}.$$

Then denoting by $n_S^x, n_S^y, n_S^z$ the three coordinates of a normal unit vector $n_S$ one gets

$$\int_S \varphi_i n_S^x = \frac{n_S^x mes(S)}{3} = \sigma_S^x \quad \text{for } S \in \mathbf{S}_h^i. \tag{12}$$

Then according to (5), (8), and (12), the nonzero entries of $A_1$ are given by

$$\left(A_1^{XY}\right)_{ij} = \sum\nolimits_{S \in S_h^i \cap S_h^j} \sigma_S^x \sigma_S^y$$

and the entries of the other blocks of $A_1$ defined by (9) can be easily deduced.

In order to explicit the entries of the right-hand side $b_\theta$, we introduce the notation

$$\boldsymbol{\tau}_h^i = \{ T \in \boldsymbol{\tau}_h / \text{vertex } a_i \text{ belongs to tetrahedron } T \}.$$

From (2), (6), and (8), one gets

$$(b_0^x)_i = \sum\nolimits_{T \in \boldsymbol{\tau}_h^i} \left[ \sum\nolimits_{S \text{ face of } T} \alpha_S \left( \int_T (w_S)^X \varphi_i \right) \right]. \tag{13}$$

By denoting $x_k, k = 1$ to 4, the $X$-coordinates of the four vertices of tetrahedron $T$, from the definition of the local shape functions, it can be deduced

$$\int_T (w_S)^X \varphi_i = \frac{\varepsilon}{60} \sum\nolimits_{k=1}^{4} \eta_k (x_k - x_r) \tag{14}$$

with

$$\eta_k = \begin{cases} 1, & \text{if } k \neq m \\ 2, & \text{if } k = m \end{cases} \varepsilon = \begin{cases} +1, & \text{if normal } n_S \text{ is outgoing of } T \\ -1, & \text{if not} \end{cases}$$

where $m$ and $r$ are the respective local numbers in $T$ of $a_i$ and the vertex opposite to $S$.

The computation of $b_1$ from (7), (8), and (12) is straightforward

$$(b_1^x)_i = \sum_{S \in \mathbf{S}_h^i} \alpha_S \left( \int_S \varphi_i.n_S^x \right) = \sum_{S \in \mathbf{S}_h^i} \alpha_S \sigma_S^x. \tag{15}$$

## IV. Limit Case $\theta = 0$ and $\theta = 1$

In the particular case $\theta = 0$, the linear system reduces to $A_0 \xi = b_0$. As pointed out before in (10), matrix $A_0$ has a block diagonal structure with mass matrices on the diagonal. In order to save some computation time an explicit but approximate solution to the least square problem can be obtained by a mass-lumping technique [5]. The mass lumping allows to approximate the mass matrix by a diagonal matrix

$$\int_\Omega \varphi_i^2 \approx \frac{\beta}{4} \text{ with } \beta = \sum_{T \in \boldsymbol{\tau}_h^i} mes(T). \tag{16}$$

The components of the approximate solution of the least squares problem are obtained as $(\xi^X)_i = (4)/(\beta)(b_0^X)_i i = 1$ to $n$.

For $\theta = 1$, the minimization problem defined in (3) is a pure discrete least square problem. From (5) it can be verified that $A_1 = T^t T$ where T is a rectangular matrix $M_h \times N_h$ ($M_h$: number of faces of $\mathbf{S}_h$) with entries given by

$$T_{ki} = \int_S v_i.n \; i = 1 \text{ to } N_h \tag{17}$$

with $k$ the index of the face $S$.

Uniqueness of the solution is ensured if there exists no $\xi \neq 0$ such that $T\xi = 0$, i.e., if the columns of $T$ are linearly independent. It requires that the number of lines $M_h$ is larger than $N_h$ the number of columns, i.e., the number of faces is greater or equal to three times the number of nodes.

The uniqueness condition of the solution can be also written as: there exists no $B \in V_h - \{0\}$ such that $\int_S B.n = 0$ for any face $S$ in $\mathbf{S}_h$. It can be also translated into the following mesh characterization: there is no vector field $B \neq 0$ defined on the nodes of the mesh $\tau_h$ such that for any face $S = abc$ the vector $B(a)+B(b)+B(c)$ belongs to $S$. In the numerical experiments presented later, such a property appears to be satisfied by the meshes. In particular, the necessary condition $M_h$ greater than $N_h$ is always satisfied.

## V. Results

In the following, method 1 (using shape functions derivatives) is the reference method. Method 2 concerns the energy approximation ($\theta = 0$). Method 3 is obtained with the flux approximation ($\theta = 1$). Method 4 mixes both criteria ($\theta = h$) where h is the length average of the mesh edges. Method 5 corresponds to $\theta = 0$ (energy criterion) with the mass lumping approximation. For methods 2–4, a symmetric quasi-minimal residual solver with symmetric successive over relaxtion (SSOR) preconditioning (PQMR) is used to solve the matrix system.

The accuracy and the time required for each post processing method are tested on two examples. The numerical results are compared with analytical solutions of scattering of a plane wave by a sphere. Analytical solutions can be found in [6] and numerical formulations of scattering problems are given in [7]. A 50-mm radius sphere is meshed with 100 nodes on its surface. The frequency of the incident plane wave is 1 GHz. For the numerical computation, the absorbing boundary condition (ABC) (first-order Engquist Majda) is located at a half wavelength from the sphere (Fig. 1).

The efficiency of the different post processing strategies is evaluated on the surface of the sphere by means of a relative error estimator defined by

$$\text{Error} = \frac{\sqrt{\sum_{a\in S} |H^{\text{anal}}(a) - H^{\text{num}}(a)|^2}}{\sqrt{\sum_{a\in S} |H^{\text{anal}}(a)|^2}} \tag{18}$$

where $H^{\text{anal}}$ is the analytical solution to the magnetic formulation. The nodal values $H^{\text{num}}(a)$ are given by the post processing methods from the FE edge values obtained in the electric formulation.

In the first example, the sphere is modeled as a PEC. The problem is meshed with 4481 nodes (100 nodes on the sphere's surface) and 22 108 tetrahedral elements leading to 28 161 edges (294 edges on the sphere's surface). The total solving time including assembling and solving of the matrix system for the FE code is about 300 s on a HP J5000. Observe that the same solver as in the post-processing is used: 217 iterations of symmetric PQMR are required to solve the matrix system.

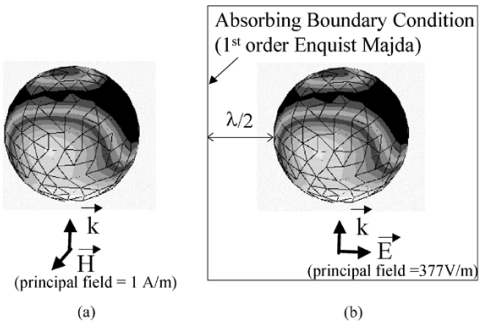


Fig. 1. (a) Analytical model and (b) numerical model of scattering by a sphere.

In the second example, the sphere consists of a perfect magnetic material with $\mu_r = 3$. A discontinuity of the magnetic field at the surface of the sphere is therefore introduced and the interior of the sphere (region 1) has to be meshed. The same mesh as for the PEC sphere (Fig. 1) is used outside of the sphere (region 2). For the whole mesh, one gets 4547 nodes and 22 783 tetrahedral elements leading to 28 805 edges (294 edges on the sphere's surface). The total time for the edge FE solving is about 436 s (537 iterations of the symmetric PQMR). The post processing is made in the only region 2 (air) according to the analytical solution [6].

For the analytical model of scattering by the sphere (PEC or magnetic), the solution in magnetic field [6] is obtained from an incident plane wave with a magnetic field $H$ ($\text{modulus} = 1$ A/m) on the $y$-axis and a propagation of the incident plane wave across the $z$-axis (Fig. 1).

For the numerical solution (edge FE code) of scattering by the sphere (PEC or magnetic), the formulation is written in term of total electrical field $E$. The incident electrical field $E$ is on the $x$-axis with a $\text{modulus} = 120\pi$ V/m and a propagation of the incident plane wave across the $z$-axis (Fig. 1). The dual field $H$ is computed with the five methods.

The proposed methods (methods 2–5) are compared to the natural method 1 which consists in computing the curl of $E_h$ on each element of the mesh from the edge values by means of the local edge basis. As explained by (19), a value of the curl of $E_h$ at a node nd is obtained as an arithmetic mean involving the $n_T$ elements containing the node. For a node at the interface between two regions, two values are obtained by averaging the field values separately in each region

$$\mathbf{B}(nd) = \frac{\sum_{T \ni nd} \mathbf{B}_T(nd)}{n_T} \quad \text{with } \mathbf{B}_T = \nabla \times \sum_{i=1}^{6} \gamma_i \mathbf{p}_i \tag{19}$$

where $p_i$ is the local edge shape function associated to edge $e_i$ in $T$ and $\gamma_i = \int_{e_i} E_h.t$

Table I presents the errors computed with (18) for the problem of scattering by the PEC and PM spheres. The number of iterations required by the symmetric PQMR and the total additional CPU time for each post-processing method are also presented.

TABLE I
ERROR ON THE SPHERES

| | method | Error | Iterations | Time (s) |
|---|---|---|---|---|
| PEC Sphere | method 1 | 0.203 | / | 3 |
| | method 2 | 0.094 | 9 | 10 |
| | method 3 | 0.115 | 20 | 22 |
| | method 4 | 0.092 | 10 | 12 |
| | method 5 | 0.209 | / | 3 |
| PM Sphere | method 1 | 0.260 | / | 3 |
| | method 2 | 0.154 | 9 | 10 |
| | method 3 | 0.138 | 19 | 19 |
| | method 4 | 0.134 | 12 | 12 |
| | method 5 | 0.273 | / | 3 |

## VI. CONCLUSION

As shown in Table I, methods 2–4 allow us to obtain a more accurate dual field. Method 2 (energy approximation) is less CPU consuming. Method 3 (flux approximation) is the most CPU time consuming. Method 4 (both criteria) gives the best precision with an acceptable additional CPU time. Method 5 (energy criterion with mass lumping approximation) is the less CPU time consuming but gives a poor precision. The additional CPU time for the post processing (all methods) remains small compared to the total edge FE solving time (300 s for the PEC sphere and 436 s for the PM sphere).

## REFERENCES

[1] C. D. Dibben and R. Metaxas, "A comparison of errors obtained with Whitney and linear edge elements," *IEEE Trans. Magn.*, vol. 33, no. 2, pp. 1524–1527, Mar. 1997.
[2] L. J. Volakis and D. B. Davdson, "Implementation issues for three dimensional vector FEM programs," *IEEE Antennas Propagat. Mag.*, vol. 42, no. 6, pp. 100–107, Jun. 2000.
[3] J. P. Webb, "Edge element and what they can do for you," *IEEE Trans. Magn.*, vol. 29, no. 2, pp. 1460–1465, Mar. 1993.
[4] C. Vollaire, F. Musy, and R. Perrussel, "Post processing for the vector finite element method: From edge to nodal values," *COMPEL*, vol. 24, no. 4, pp. 1274–1283, 2005.
[5] O. C. Zienkiewicz, *The Finite Element Method*. New York: McGraw-Hill, 1970.
[6] R. F. Harrington, *Field Computation by Moment Methods*. New York: MacMillan, 1968.
[7] Y. Bi, J. L. Bi, and J. L. Nicolas, "Vector absorbing boundary conditions for nodal or mixed finite elements," *IEEE Trans. Magn.*, vol. 32, no. 3, pp. 848–853, May 1996.

Manuscript received June 26, 2005 (e-mail: christian.vollaire@ec-lyon.fr).