Convergence-Optimized, Higher Order Vector Finite Elements for Microwave Simulations

Traianos V. Yioultsis and Theodoros D. Tsiboukis, *Senior Member, IEEE*

***Abstract*—We introduce a general class of higher order parameter-dependent Whitney elements, unlike previous approaches that resulted in specific element definitions. All elements of this kind provide the same solution, but their convergence properties may be significantly different. The most essential fact, though, is the introduction of an optimization procedure, which reveals the existence of an optimal, with respect to convergence, element. The produced second order elements are tested in both two-dimensional (2-D)and three-dimensional (3-D) microwave simulations.**

***Index Terms*—Finite-element methods, modeling, waveguide components.**

I. Introduction

THE FINITE-element method (FEM) is a valuable tool in computational electromagnetics and especially in microwave analysis and design. However, most simulations of this kind are still based on the well-known edge element [1], although higher orders are highly desirable for better accuracy. Generalizations of this kind do exist, [2]–[7] and actually form the basis of some widely used commercial packages. However, thanks to a recent comparative study of several second-order elements [7], it is clear that they exhibit significant differences, with respect to the convergence of the iterative method used to solve the global finite element matrix equation. In particular, the element [3] has slow convergence, compared to [2] and [7]. What we introduced in [3], though, was actually a systematic approach to construct an infinite variety of second and third order Whitney elements. We also implied that these could be related to each other by an affine transformation. In this study, we formulate the most general transformation of degrees of freedom (DOF) and shape functions (SF), thus defining an element class, depending on nine parameters. By a proper optimization scheme, based on the elemental matrix and being, therefore, efficient and easy to implement, we managed to trace the optimal parameters, resulting in less than 30% of the iterations for the reference element [3]. This is demonstrated in both two dimensional (2-D) and three-dimensional (3-D) simulations.

Manuscript received May 14, 2001; revised August 3, 2001. This work was supported in part by the Greek General Secretariat of Research and Technology under Grant 99ED325 (PENED '99). The review of this letter was arranged by Associate Editor Dr. Shigeo Kawasaki.

The authors are with the Department of Electrical and Computer Engineering, Division of Telecommunications, Aristotle University of Thessaloniki, Thessaloniki, Greece (e-mail: tsibukis@eng.auth.gr; traianos@egnatia.ee.auth.gr).

Publisher Item Identifier S 1531-1309(01)09480-6.

II. General Class of Whitney Elements

A. Basic Element and the Affine Transformation

The formation of the wide class of second-order Whitney elements starts from the basic one, derived by a justified choice of DOFs [3]. On each edge (i, j), the two DOFs are

$$F_{ij}^{i}=\int_{(i)}^{(j)}\mathbf{F}\cdot\hat{\mathbf{t}}_{ij}\zeta_i\,dl,\qquad F_{ji}^{j}=\int_{(j)}^{(i)}\mathbf{F}\cdot\hat{\mathbf{t}}_{ji}\zeta_j\,dl \tag{1}$$

while two additional DOFs are defined on each face (i, j, k)

$$\begin{aligned}F_{ijk}&=\iint_{\{i,j,k\}}\mathbf{F}\times\hat{\mathbf{n}}^{+}\cdot\nabla\zeta_j\,ds\\ F_{ikj}&=\iint_{\{i,j,k\}}\mathbf{F}\times\hat{\mathbf{n}}^{-}\cdot\nabla\zeta_k\,ds.\end{aligned} \tag{2}$$

This choice is not unique, although the kernels should follow the same patterns. A systematic procedure, based on a decoupling property and the correct geometric representation of the discrete de Rham sequence results in the following SF expressions

$$\overline{w}_{ij}^{i}=(8\zeta_i^2-4\zeta_i)\nabla\zeta_j+(-8\zeta_i\zeta_j+2\zeta_j)\nabla\zeta_i \tag{3}$$

$$\overline{w}_{ij}^{k}=-16\zeta_i\zeta_j\nabla\zeta_k+8\zeta_j\zeta_k\nabla\zeta_i+8\zeta_k\zeta_i\nabla\zeta_j \tag{4}$$

for edge and face DOFs, respectively, both in 2-D and 3-D.

The essential concept here is to define an as general as possible transformation of SFs, resulting in optimal convergence. Since it is unlikely to find this by heuristic trials, we pursue a systematic analysis by introducing transformations of DOFs and SFs,

$$\mathbf{F}'=\mathbf{M}\mathbf{F},\qquad \overline{\mathbf{W}}'=\mathbf{\Lambda}\overline{\mathbf{W}} \tag{5), (6}$$

respectively, where $\mathbf{F}$ is the column vector of DOFs for a single element, $\mathbf{F}'$ the transformed one, $\overline{\mathbf{W}}$ the column vector of vector SFs and $\overline{\mathbf{W}}'$ the new basis. Moreover, $\mathbf{M}$ and $\mathbf{\Lambda}$ are 8×8 in 2-D or 20×20 in 3-D matrices, which are related to each other, since the field expression is, in column vector form,

$$\overline{F}=\mathbf{F}'^{T}\overline{\mathbf{W}}'=(\mathbf{M}\mathbf{F})^{T}\mathbf{\Lambda}\overline{\mathbf{W}}=\mathbf{F}^{T}(\mathbf{M}^{T}\mathbf{\Lambda})\overline{\mathbf{W}} \tag{7}$$

and given that the transformation should not affect the numerical solution, it naturally comes out that

$$\mathbf{M}^{T}\mathbf{\Lambda}=\mathbf{I},\quad\text{or}\quad\mathbf{\Lambda}=(\mathbf{M}^{T})^{-1}. \tag{8}$$

The most interesting observation, though, is how the transformation affects the system matrices. To find this, we express

the transformed T-matrix in the rather unusual column vector product of the form

$$\mathbf{T}' = [T'_{ij}] = \lfloor \langle \overline{w}'_i, \overline{w}'_j \rangle \rfloor = \int_{V_e} [\overline{w}'_i \cdot \overline{w}'_j]\, dv = \int_{V_e} \overline{\mathbf{W}}' \overline{\mathbf{W}}'^T\, dv \tag{9}$$

and apply (6), which results in

$$\mathbf{T}' = \int_{V_e} \mathbf{\Lambda} \overline{\mathbf{W}}\, \overline{\mathbf{W}}^T \mathbf{\Lambda}^T\, dv = \mathbf{\Lambda T \Lambda}^T. \tag{10}$$

The same relation holds for the S-matrix. The question is what prevents a so general transformation that could even nearly diagonalize the matrix. The answer lies in the allocation of DOFs and the tangential continuity property, dictating that (5) should associate a DOF with those *defined only on the same simplex or its own subsimplices*. This means that (5) should link edge DOFs with those on the same edge only, or a face DOF with all face and edge DOFs of the same face. Hence, further links are not possible.

Therefore, in 2-D, we introduce an M-matrix of the form

$$\mathbf{M} = \left[\begin{array}{cc:cc:cc:cc} 1 & a & & & & & & \\ a & 1 & & & & & & \\ \hdashline & & 1 & a & & & & \\ & & a & 1 & & & & \\ \hdashline & & & & 1 & a & & \\ & & & & a & 1 & & \\ \hdashline e_1 & d_1 & d_2 & e_2 & f_1 & f_2 & b & c \\ f_2 & f_1 & e_2 & d_2 & d_1 & e_1 & c & b \end{array}\right] \tag{11}$$

where the face DOFs are numbered last (Fig. 1). The situation is identical in 3-D, where the 20×20 M-matrix is easily derived if the methodology shown in Fig. 1 is applied to each one of the four faces. In both cases, the element class depends on nine parameters. Although, without loss of generality, the edge-to-edge coefficient in (11) is taken equal to 1, it is very significant that b can vary, since it controls the relative magnitude of edge and face DOFs (and SFs). We stress that the simultaneous implementation of the whole class is simple and requires programming of (5) and (6) only, if numerical integration is used.

B. The Optimization Scheme

The key issue now is how to determine the transformation that provides optimal convergence. Optimization for the entire matrix would require enormous computational times. Instead, we establish a scheme, based on a criterion for the element matrix, A. It is difficult, though, to prove in mathematical terms, which criterion for the element matrix minimizes the number of iterations for the assembled system. We experimented with various criteria, but only few were successful. The condition number criterion fails, possibly due to the presence of negative eigenvalues. A much better one is proven to be the *diagonality criterion*, where the object function is

$$F(\mathbf{p}) = \frac{1}{N_f} \sum_{i=1}^{N_f} \frac{\sum_{j=1,\, j\neq i}^{N_f} |a'_{ij}(\mathbf{p})|}{|a'_{ii}(\mathbf{p})|} \tag{12}$$

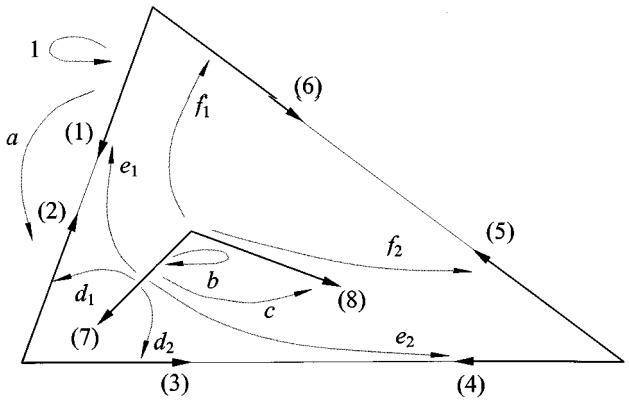

Fig. 1. Links among the transformed and the original DOFs in a 2-D or a face of a 3-D second order Whitney element [shown for DOFs (1) and (7) only].

where

$$\mathbf{p} = [a\, b\, c\, d_1\, d_2\, e_1\, e_2\, f_1\, f_2]^T$$

is the parameter vector, $a'_{ij}(\mathbf{p})$ the entries of the transformed A-matrix, and N_f the number of degrees of freedom for a single element. If (12) is minimized, the nondiagonal entries for each row will be as low as possible, compared to the diagonal element.

The best results, though, are obtained by introducing a criterion, based on the eigenvalues $\lambda_i(\mathbf{p}), i = 1, \ldots, N_f$ of the transformed A-matrix. Indeed, if we enforce that they are as close to each other as possible, the matrix will be closer to a diagonal matrix. An efficient criterion of this kind is

$$F(\mathbf{p}) = \sum_{i=1}^{N_f} |\lambda_i(\mathbf{p}) - \max \lambda_i(\mathbf{p})| \tag{13}$$

which has given the best results so far, but the quest for an even better one is, still, an open problem.

As for the optimization method, a simple line search algorithm can be chosen, since it is applied to the elemental matrix only. The algorithm searches toward the principal directions in the parameter space, $\mathbf{d}_k = \pm \mathbf{e}_i$, $\mathbf{e}_i(j) = \delta_{ij}$ and $i, j = 1, \ldots, 9$. From those 18 directions, we chose the one that minimizes the object function. Hence, the algorithm, with adaptive step, is outlined as follows

$$n = 0,\ \mathbf{p}^0 = [0\,1\,0\,0\,0\,0\,0\,0\,0]^T, \qquad \delta = \delta^0$$

do

$$\begin{aligned} \mathbf{A}' &= (\mathbf{M}(\mathbf{p}^n)^T)^{-1} \mathbf{A} \mathbf{M}(\mathbf{p}^n)^{-1} \\ \mathbf{d}_k &= \arg\min\{F(\mathbf{p}^n + \delta \mathbf{d}_i)\}, \qquad i = 1, \ldots, 18 \\ \mathbf{p}^{n+1} &= \mathbf{p}^n + \delta \mathbf{d}_k, \qquad n = n + 1 \end{aligned}$$

if

$$|\mathbf{p}^{n+1} - \mathbf{p}^{n-1}| / |\mathbf{p}^{n-1}| < e$$

then $\delta = \mu\delta$ while

$$|\mathbf{p}^{n+1} - \mathbf{p}^{n}| / |\mathbf{p}^{n}| > e.$$

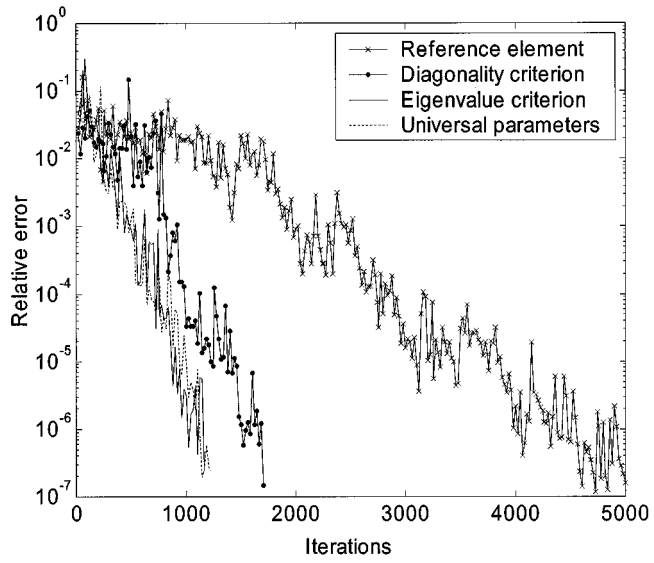

Fig. 2. Comparison of convergence rates for a 2-D problem with 22 848 degrees of freedom. The universal parameters are $a = 0.298$, $b = 2.019$, $c = -0.282$, $d_1 = 0.4$, $d_2 = -0.526$, $e_1 = 0.1$, $e_2 = 0.002$, $f_1 = 0.225$, $f_2 = -0.601$.

The if-condition is used to reduce step δ in the frequent case of stagnation. Typical values are $\delta^0 = 0.01 \div 0.1$ and $\mu = 0.5 \div 0.9$. For larger starting steps, the algorithm can be trapped to badly conditioned solutions.

III. Numerical Results

A 2-D and a 3-D application are considered and a preconditioned conjugate gradient method is used to solve the system. We have chosen a diagonal preconditioning, i.e., the preconditioner is the system matrix itself. This is a logical choice, since it requires no extra memory, while it is not guaranteed that other schemes like ICCG or controlled fill-in preconditioners respond very well to matrices generated from Whitney elements. In any case, we are interested in comparing the iterations required from a specific method, due to the improvement of the system's conditioning. The number of iterations for convergence to a relative error of 10^{-7} is investigated. In 2-D, the scattering from a dielectric cylinder is considered, via a frequency-domain E-formulation and an unsplit diagonally anisotropic PML. The frequency-domain 3-D application deals with a waveguide termination, backed by a dielectric or ferrite material. The optimizations are performed according to the most characteristic element, preferably that of worst quality factor, where the latter one is defined as three times the ratio of radii of the inscribed and the circumscribed sphere, respectively. In both cases, a set of universal parameters has been extracted from a large-scale problem, with elements of various quality factors. These sets are given for anyone that prefers to skip the optimization and are expected to work well in all cases. In Fig. 2, comparisons are performed in 2-D, whereas a comparison of different criteria for the 3-D case is given in Table I, for different problem sizes and average quality factors Q_{av}. The set of universal parameters, extracted from the largest problem in 3-D is $a = 0.367$, $b = 2.857$, $c = -0.494$, $d_1 = -0.350$, $d_2 = 0.509$, $e_1 = -0.081$, $e_2 = -0.075$, $f_1 = 0.527$, $f_2 = -0.350$. Since this set is obtained from a typical element, it is not expected to significantly depend on the problem's geometry. However, it is a matter of further investigation if it is truly universal.

TABLE I
Number of Iterations for Convergence (3-D Problem)

Total DOFs	Q_{av}	Ref. element	Diag. Criterion	Eigenvalue Criterion	Universal Parms
1830	0.55	1345	486(36%)	438(33%)	411(31%)
2328	0.73	848	307(36%)	283(33%)	251(30%)
13380	0.55	2613	982(38%)	717(27%)	747(29%)
17508	0.73	1822	849(47%)	477(26%)	450(25%)
43770	0.55	3948	1395(35%)	1065(27%)	1071(27%)
57852	0.73	2627	1263(48%)	707(27%)	678(26%)
102120	0.55	5236	1855(35%)	1402(27%)	1402(27%)

In all cases, the solution generated by different elements is identical, as predicted by (7). However, we observe a striking reduction of iterations, reaching a 25% of those for (1), (2). Surprisingly, the universal parameters give, sometimes, slightly better results. According to [7], the most efficient elements [2], [7] require about half of the iterations of (1), (2), which makes the proposed element the best one.

IV. Conclusion

We have introduced a wide class of second-order triangular and tetrahedral Whitney elements and an optimization scheme, resulting in a remarkable improvement of convergence speed. Although the parameters derived, can be widely used, we suggest the use of the optimization procedure, which is not only easy but also locates the best parameters in a few seconds. Apart from that, this study should be conceived as a general theory that provides an improvement to one of the most adverse issues in vector FEM modeling and could be applied to other elements, as well.

References

[1] A. Bossavit, "Whitney forms: A class of finite elements for three-dimensional computations in electromagnetism," *Proc. Inst. Elect. Eng. A*, vol. 135, no. 8, pp. 493–500, 1988.

[2] J. F. Lee, D. K. Sun, and Z. J. Cendes, "Tangential vector finite elements for electromagnetic field computation," *IEEE Trans. Magn.*, vol. 27, pp. 4032–4035, Sept. 1991.

[3] T. V. Yioultsis and T. D. Tsiboukis, "Development and implementation of second and third order vector finite elements in various 3-D electromagnetic field problems," *IEEE Trans. Magn.*, vol. 33, pp. 1812–1815, Mar. 1997.

[4] R. D. Graglia, D. R. Wilton, and A. F. Peterson, "Higher order interpolatory vector bases for computational electromagnetics," *IEEE Trans. Antennas Propagat.*, vol. 45, pp. 329–342, Mar. 1997.

[5] L. S. Andersen and J. L. Volakis, "Hierarchical tangential vector finite elements for tetrahedra," *IEEE Microwave Guided Wave Lett.*, vol. 8, pp. 127–129, Mar. 1998.

[6] J. P. Webb, "Hierarchal vector basis functions of arbitrary order for triangular and tetrahedral finite elements," *IEEE Trans. Antennas Propagat.*, vol. 47, pp. 1244–1253, Aug. 1999.

[7] Z. Ren and N. Ida, "Solving 3D eddy current problems using second order nodal and edge elements," *IEEE Trans. Magn.*, vol. 36, pp. 746–750, July 2000.