ELSEVIER

Comput. Methods Appl. Mech. Engrg. 169 (1999) 297–309

Computer methods in applied mechanics and engineering

Efficient finite element solvers for the Maxwell equations in the frequency domain

Romanus Dyczij-Edlinger[a,*], Guanghua Peng[a], Jin-Fa Lee[b]

[a] *Ansoft Corp., 4 Station Sq., Suite 660, Pittsburgh, PA 15219, USA*

[b] *ECE Department, Worcester Polytechnic Inst., 100 Institute Rd., Worcester, MA 01609, USA*

Abstract

The present paper shows that certain instabilities encountered with Nedelec-type finite element implementations of the vector wave equation can be eliminated by a family of Lagrange multiplier methods. The considered approaches can be interpreted as coupled vector and scalar potential methods, including an ungauged formulation. We advocate the latter form for use with iterative solvers. We discuss an inexpensive high-frequency variant of the method and show how hierarchical finite element bases can be utilized to derive efficient, partially gauged formulations of higher order. Based on the multigrid idea, a Schwartz-type solver is constructed which can overcome a major insufficiency of conventional preconditioners.

1. Introduction

H(curl) conforming Nedelec-type [1,2] finite element methods for time-harmonic electromagnetic field problems are usually based on the vector wave equation. The problem with this formulation is that it incorporates the gauge condition for the electric field, $\nabla \cdot [\epsilon]\vec{E} = 0$, in frequency dependent form, which can lead to severe ill-conditioning of the finite element matrix. When the operating frequency or the mesh size becomes small, a whole family of eigenvalues converges towards zero. This deficiency may result in poor solution quality or even breakdown of the matrix factorization process. The most common effect seen in practise is that iterative solvers converge very poorly [3].

The problem can be overcome by means of Lagrange multipliers similar to [4]. We will investigate a family of such methods which may contain additional penalty terms. The considered approaches can be interpreted as coupled vector and scalar potential methods. Thanks to particular mapping relations between the selected finite element spaces [2], the penalty term can be chosen such that the formulation remains perfectly ungauged even after discretization.

With respect to iterative solvers, the ungauged version is very attractive, because the dimension of the solution space is at a minimum in this case. However, we advocate this formulation for the lowest order finite element scheme only. In the higher order case, we utilize hierarchical vector finite elements [5] which provide explicit basis functions for all higher order gradients to construct a more economic, partially gauged approach.

The analysis of the $\vec{E}$ field formulation also reveals that there is a second class of eigenvalues which conventional preconditioners may be unable to handle. These are related to the lowest order cavity modes of the system with resonances close to or below the operating frequency. We use multigrid ideas and, again, hierarchical finite elements to derive a robust Schwartz-type solver.

Numerical examples are given to illustrate the characteristics of the considered formulations.

* Corresponding author. E-mail: edlinger@ansoft.com

0045-7825/99/$ – see front matter
PII: S0045-7825(98)00159-5

2. Boundary value problems and finite element spaces

The time-harmonic form of the Maxwell equations in lossless regions can be written as

$$\nabla \times \vec{E} = -jk_0\eta_0[\mu_r]\vec{H}\,, \tag{1}$$

$$\nabla \times \vec{H} = jk_0/\eta_0[\epsilon_r]\vec{E}\,, \tag{2}$$

$$\nabla \cdot [\epsilon_r]\vec{E} = 0\,, \tag{3}$$

$$\nabla \cdot [\mu_r]\vec{H} = 0\,, \tag{4}$$

where $\vec{E}$ and $\vec{H}$ are phasors for the electric and magnetic fields, $[\mu_r]$ denotes the relative permeability tensor and $[\epsilon_r]$ the relative permittivity, k_0 is the free-space wavenumber and η_0 the free-space characteristic impedance. We do not consider radiation into outer space, and it is assumed that the materials are lossless and passive. Let Ω be a finite, bounded domain with boundary Γ and $\hat{n}$ the outward normal vector on Γ. We consider the solution to (1)–(4) over Ω with respect to the boundary conditions

$$\vec{H} \times \hat{n} = \overline{K}_E \quad \text{given on } \Gamma_H\,, \tag{5}$$

$$\vec{E} \times \hat{n} = \vec{0} \quad \text{given on } \Gamma_E = \Gamma - \Gamma_H\,. \tag{6}$$

For simplicity, let Γ_E, the electrically conducting part of the boundary, be non-empty and connected.

We will also consider an auxiliary Laplace problem of the type

$$\nabla \cdot [\epsilon_r]\nabla V = 0 \quad \text{in } \Omega\,, \tag{7}$$

$$\hat{n} \cdot [\epsilon_r]\nabla V = \overline{D}_n \quad \text{on } \Gamma_H\,, \tag{8}$$

$$V = 0 \quad \text{on } \Gamma_E\,. \tag{9}$$

In the finite element model, we will solve for solutions $\vec{E} \in \mathcal{W}_0(\Omega, \Gamma_E)$ and $V \in \mathcal{V}_0(\Omega, \Gamma_E)$, where

$$\mathcal{V}_0(\Omega, \Gamma_E) \subset \{V \in L^2(\Omega) | V = 0 \text{ on } \Gamma_E\}\,, \tag{10}$$

$$\mathcal{W}_0(\Omega, \Gamma_E) \subset \{\vec{E} \in \boldsymbol{L}^2(\Omega) | \nabla \times \vec{E} \in \boldsymbol{L}^2(\Omega),\ \vec{E} \times \vec{n} = 0 \text{ on } \Gamma_E\}\,, \tag{11}$$

are appropriately chosen finite element spaces defined over the same discretization. In particular, we require $(\mathcal{V}_0; \text{grad})$, $(\mathcal{W}_0; \text{curl})$ to possess the exact sequence property [2]:

$$\ker \nabla \times \vec{E}|_{\vec{E} \in \mathcal{W}_0} = \text{span}\{\nabla V \,|\, V \in \mathcal{V}_0\}\,. \tag{12}$$

Let $N_V = \dim \mathcal{V}_0$ and $N_E = \dim \mathcal{W}_0$ be the number of unconstrained scalar and vector basis functions, respectively. Then the fact that the gradient operator is a linear mapping from $\mathcal{V}_0$ into $\mathcal{W}_0$ can be represented by an $(N_E \times N_V)$ dimensional 'gradient matrix' $[G]$ such that

$$[G]: \vec{v} \to \vec{e}: \quad \vec{e} = [G]\vec{v}\,, \tag{13}$$

where $\vec{v}$ and $\vec{e}$ denote coefficient vectors associated with scalar and vector basis functions, respectively.

3. Field formulation

For *non-zero* wave numbers, $k_0 > 0$, a relevant subset of (1)–(4) is given by the vector wave equation,

$$\nabla \times [\mu_r]^{-1}\nabla \times \vec{E} - k_0^2[\epsilon_r]\vec{E} = 0\,. \tag{14}$$

However, in the static case, $k_0 = 0$, Eq. (14) no longer imposes the divergence constraint (3). In the following, we will investigate how this deficiency can affect finite element implementations. Starting from the weak form of the boundary value problem (14), (6), (5), a Galerkin approach for the finite element solution

$$\vec{E}_h = \sum e_i \vec{W}_i \tag{15}$$

with basis functions $\vec{W}_i \in \mathcal{W}_0$ leads to to an algebraic set of equations

$$([S_{EE}] - k_0^2[T_{EE}]) \cdot \vec{x}_E = k_0 \vec{r}_K \tag{16}$$

for the unknown coefficient vector $\vec{x}_E = \mathrm{col}(e_i)$. The elements of the matrices $[S_{EE}]$ and $[T_{EE}]$, and the right-hand side $\vec{r}_E$ are given by

$$[S_{EE}]_{ij} = \int_\Omega (\nabla \times \vec{W}_i) \cdot [\mu_r]^{-1} (\nabla \times \vec{W}_j) \, \mathrm{d}\Omega \,, \tag{17}$$

$$[T_{EE}]_{ij} = \int_\Omega \vec{W}_i \cdot [\epsilon_r] \vec{W}_j \, \mathrm{d}\Omega \,, \tag{18}$$

$$\vec{r}_{K,i} = -j\eta_0 \int_{\Gamma_H} \vec{W}_i \cdot \bar{\vec{K}}_E \, \mathrm{d}\Gamma \,. \tag{19}$$

From (12) and (17), it can be seen that $[S_{EE}]$ is a positive semi-definite matrix whose nullspace is formed by N_V pure gradients. Due to (18), $[T_{EE}]$ is positive definite. To characterize the numerical properties of the matrix $[S_{EE}] - k_0^2[T_{EE}]$, we consider the eigenmode problem associated with the structure, which in discretized form reads

$$([S_{EE}] - k_i^2[T_{EE}]) \cdot \vec{x}_i = 0 \,. \tag{20}$$

The solutions $(k_i^2, \vec{x}_i)$ to this generalized eigenvalue problem can be classified as follows:

Type A: *Unphysical source fields*. The eigenvectors $\vec{x}_1 \cdots \vec{x}_{N_V}$ represent gradients associated with an N_V-fold zero eigenvalue $k_1^2 = \cdots = k_{N_V}^2 = 0$. The corresponding electrical fields $\vec{E}$ do not satisfy the source-free condition (3), i.e.

$$\exists\, v_i \in \mathcal{V}_0 : \quad \int_\Omega \nabla v_i \cdot [\epsilon_r] \vec{E} \, \mathrm{d}\Omega \neq 0 \,. \tag{21}$$

Type B: *Cavity modes*. The remaining eigenvectors $\vec{x}_{N_V+1} \cdots \vec{x}_{N_E}$ and eigenvalues $0 < k_{N_V+1}^2 \leqslant \cdots \leqslant k_E^2$ are approximations to the physical cavity modes and resonance wave-numbers of the system.

We now use the eigenpairs $(k_i^2, \vec{x}_i)$ to establish bounds for the eigenvalues λ_i of the finite element matrix $[S_{EE}] - k_0^2[T_{EE}]$. Let $\tau_{\min}$ and $\tau_{\max}$ denote the minimum and maximum eigenvalues of $[T_{EE}]$. For any $\vec{x}_i$, the Rayleigh coefficient $\rho_E(\vec{x}_i)$ for $[S_{EE}] - k_0^2[T_{EE}]$ reduces to

$$\rho_E(\vec{x}_i) = \frac{\vec{x}_i^{\top}([S_{EE}] - k_0^2[T_{EE}])\vec{x}_i}{\vec{x}_i^{\top}\vec{x}_i} = (k_i^2 - k_0^2)\frac{\vec{x}_i^{\top}[T_{EE}]\vec{x}_i}{\vec{x}_i^{\top}\vec{x}_i} \,. \tag{22}$$

Now the Courant–Fischer theorem [6] yields

$$\mathrm{sign}\ \lambda_i = \mathrm{sign}(k_i^2 - k_0^2) \,, \tag{23}$$

$$|k_i^2 - k_0^2|\tau_{\min} \leqslant |\lambda_i| \leqslant |k_i^2 - k_0^2|\tau_{\max} \,. \tag{24}$$

This leads us to the following characterization of $[S_{EE}] - k_0^2[T_{EE}]$:

Type A: At any non-zero wavenumber, there exist N_V negative eigenvalues related to the set of non-physical eigenmodes of the vector wave equation. As the operating wavenumber k_0 goes towards zero, so do these eigenvalues. The reason for this behavior is that the source-free condition (3), which serves as a gauge for the electrical field, enters the vector wave equation in frequency dependent form and vanishes in the static case.

Type B: The remaining $N_E - N_V$ eigenvalues are related to the physical resonances of the considered structure. Each resonance below the operating wavenumber causes one negative eigenvalue, and each resonance above k_0

leads to a positive one. Hence, the finite element matrix will become ill-conditioned when the operating wavenumber is close to a resonance. This behavior reflects high parameter sensitivity of the underlying physical problem and can hardly be avoided. However, the number of affected modes remains very small.

The existence of *Type A* eigenvalues results in serious limitations for the practical applicability of the field method. Consider the low-frequency behavior of the formulation: For any given machine precision, we can find an operating wavenumber below which the divergence constraint (3), which is encoded as $k_0^2[G]^{\mathrm{T}}[T_{EE}]$, does not even enter the finite element matrix anymore. Even far beyond this critical wavenumber, one should always be alert that ill-conditioning of the finite element matrix can adversely affect the accuracy of the solution. Particularly (the weak form of) the divergence constraint (3) may be poorly fulfilled. We admit that this insufficiency does not necessarily disqualify the field formulation as a high-frequency approach. However, in the course of adaptive mesh refinement, we approach the static case locally in the smallest elements, i.e., we will see increasing ill-conditioning and, at some point, the same breakdown as before.

4. Lagrange multiplier and potential formulations

It has been seen that the stability problems of the field formulation are caused by the wavenumber-dependent incorporation of the source-free condition in the vector wave equation. This suggests to introduce a Lagrange multiplier $\lambda \in \mathscr{V}_0$ to impose the divergence constraint (3) directly [4]. In the following, we consider a family of Lagrange multiplier methods with or without penalty terms of the form:

$$\nabla \times [\mu_r]^{-1} \nabla \times \vec{E} - [\epsilon_r](k_0^2 \vec{E} + \beta \nabla \lambda) = 0\,, \tag{25}$$

$$\nabla \cdot [\epsilon_r](\beta \vec{E} + \gamma^2 \nabla \lambda) = 0\,, \tag{26}$$

$$\frac{j}{k_0 \eta_0}([\mu_r]^{-1} \nabla \times \vec{E}) \times \hat{n} = \bar{K}_E\,, \qquad \frac{\partial}{\partial n}([\epsilon_r]\lambda) = 0 \quad \text{on } \Gamma_H\,, \tag{27}$$

$$\vec{E} \times \hat{n} = \vec{0}\,, \qquad \lambda = 0 \qquad \text{on } \Gamma_E\,, \tag{28}$$

where the constants β and γ^2 are to be determined later on.

Let us introduce a magnetic vector potential $\vec{A}$ and an electric scalar potential V by

$$\nabla \times \vec{A} = -j\eta_0 [\mu_r] \vec{H}\,, \tag{29}$$

$$\beta \nabla \cdot [\epsilon_r] \vec{A} = -\gamma^2 \nabla \cdot [\epsilon_r] \nabla V\,, \tag{30}$$

$$\vec{E} = k_0 \vec{A} + \frac{\beta}{k_0} \nabla V\,. \tag{31}$$

By plugging (29) and (31) into the vector wave equation (14) and duplicating the gauge condition (30), we arrive at the transformed boundary value problem

$$\nabla \times [\mu_r]^{-1} \nabla \times \vec{A} - [\epsilon_r](k_0^2 \vec{A} + \beta \nabla V) = 0\,, \tag{32}$$

$$\nabla \cdot [\epsilon_r](\beta \vec{A} + \gamma^2 \nabla V) = 0\,, \tag{33}$$

$$\frac{j}{\eta_0}([\mu_r]^{-1} \nabla \times \vec{A}) \times \hat{n} = \bar{K}_E\,, \qquad \frac{\partial}{\partial n}([\epsilon_r] V) = 0 \quad \text{on } \Gamma_H\,, \tag{34}$$

$$\vec{A} \times \hat{n} = \vec{0}\,, \qquad V = 0 \qquad \text{on } \Gamma_E\,, \tag{35}$$

which is formally equivalent to (25)–(28). We prefer the potential formulation not only because it identifies the abstract Lagrange multiplier λ with an electrostatic scalar potential, but also because it will allow us later on to treat the special case $\beta^2 = k_0^2 \gamma^2$, where the solution is ambiguous, in a straightforward manner.

After discretization, the Galerkin equations for solution pairs ($\vec{A} \in \mathscr{W}_0$, $V \in \mathscr{V}_0$) read

$$\int_\Omega (\nabla \times \vec{W}_i)\cdot[\mu_r]^{-1}\nabla\times\vec{A}\,\mathrm{d}\Omega - \int_\Omega \vec{W}_i\cdot[\epsilon_r](k_0^2\vec{A}+\beta\nabla V)\,\mathrm{d}\Omega = -j\eta_0\int_{\Gamma_H}\vec{W}_i\cdot\bar{K}_t\,\mathrm{d}\Gamma \quad \forall\,\vec{W}_i\in\mathscr{W}_0\,, \tag{36}$$

$$-\int_\Omega \nabla w_k\cdot[\epsilon_r](\beta\vec{A}+\gamma^2\nabla V)\,\mathrm{d}\Omega = -j\frac{\beta\eta_0}{k_0^2}\int_{\Gamma_H}\nabla w_k\cdot\bar{K}_t\,\mathrm{d}\Gamma \quad \forall\, w_k\in\mathscr{V}_0\,. \tag{37}$$

The exact sequence property (12) implies that, even after discretization, the solution for the electric field (31) remains independent of the particular gauge imposed by (30), (37). Let us now investigate the gauge condition more specifically. When we restrict the space of weighting functions in (36) to pure gradients, we may choose $\{\nabla w_k \,|\, w_k\in\mathscr{V}_0\}$ as a basis. Then, (36) reduces to

$$-\int_\Omega \nabla w_k\cdot[\epsilon_r](k_0^2\vec{A}+\beta\nabla V)\,\mathrm{d}\Omega = -j\eta_0\int_{\Gamma_H}\nabla w_k\cdot\bar{K}_t\,\mathrm{d}\Gamma \quad \forall\, w_k\in\mathscr{V}_0 \tag{38}$$

and, by means of (37), we can give explicit constraints for the vector and scalar potentials,

$$(\gamma^2k_0^2-\beta^2)\int_\Omega \nabla w_k\cdot[\epsilon_r]\vec{A}\,\mathrm{d}\Omega = (\gamma^2k_0^2-\beta^2)j\frac{\eta_0}{k_0^2}\int_{\Gamma_H}\nabla w_k\cdot\bar{K}_t\,\mathrm{d}\Gamma \quad \forall\, w_k\in\mathscr{V}_0\,, \tag{39}$$

$$(\gamma^2k_0^2-\beta^2)\int_\Omega \nabla w_k\cdot[\epsilon_r]\nabla V\,\mathrm{d}\Omega = 0 \quad \forall\, w_k\in\mathscr{V}^0\,. \tag{40}$$

Provided that $\gamma^2k_0^2\neq\beta^2$, (40) is a discretized form of the Laplace problem (7)–(9) and leads to the trivial solution

$$V\equiv 0 \quad \text{in } \Omega\,, \tag{41}$$

i.e., in exact arithmetics, the 'Lagrange multiplier' V remains zero and independent of the penalty parameter γ^2. Eqs. (39), (30) imply that (a weak form of) the Coulomb gauge, $\nabla\cdot[\epsilon_r]\vec{A}=0$, with inhomogeneous boundary conditions on Γ_H is imposed on the vector potential.

In the singular case, $\gamma^2k_0^2=\beta^2$, (36) and (37) still guarantee that, for the resultant electrical field (31), the vector wave equation (14) and the source-free condition (3) are properly imposed. The ambiguity in the solution is restricted to an N_V dimensional subspace of gauge transformations of the form

$$\beta\vec{A}_1+\gamma^2\nabla V_1 = \beta\vec{A}_2+\gamma^2\nabla V_2\,. \tag{42}$$

According to (36), (37) the finite element equations for the potential formulation can be written as

$$\begin{bmatrix}[S_{EE}]-k_0^2[T_{EE}] & -\beta[M_{AV}]\\ -\beta[M_{VA}] & -\gamma^2[M_{VV}]\end{bmatrix}\begin{bmatrix}\vec{x}_A\\ \vec{x}_V\end{bmatrix}=\begin{bmatrix}\vec{r}_K\\ \dfrac{\beta}{k_0^2}\vec{r}_V\end{bmatrix}, \tag{43}$$

where $\vec{x}_A=\mathrm{col}(a_i)$ and $\vec{x}_V=\mathrm{col}(\varphi_k)$ are the coefficient vectors associated with $\vec{A}$ and V. The matrix and right-hand-side entries in (43) are given by

$$\vec{r}_{V,k} = -j\eta_0\int_{\Gamma_H}\nabla w_k\cdot\bar{K}_E\,\mathrm{d}\Gamma\,, \tag{44}$$

$$[M_{AV}]_{ik} = \int_\Omega \vec{W}_i\cdot[\epsilon_r]\nabla w_k\,\mathrm{d}\Omega\,, \tag{45}$$

$$[M_{VA}]_{ik} = \int_\Omega \nabla w_i\cdot[\epsilon_r]\vec{W}_k\,\mathrm{d}\Omega\,, \tag{46}$$

$$[M_{VV}]_{ik} = \int_\Omega \nabla w_i\cdot[\epsilon_r]\nabla w_k\,\mathrm{d}\Omega\,. \tag{47}$$

In the analysis below, we will use a set of identities which follow directly from the definition (12), (13) of the gradient matrix $[G]$:

$$[S_{EE}][G] \equiv 0\,, \tag{48}$$

$$[G]^{\mathrm{T}}[S_{EE}] \equiv 0\,, \tag{49}$$

$$[T_{EE}][G] \equiv [M_{AV}]\,, \tag{50}$$

$$[G]^{\mathrm{T}}[T_{EE}] \equiv [M_{VA}]\,, \tag{51}$$

$$[G]^{\mathrm{T}}[T_{EE}][G] \equiv [M_{VV}]\,. \tag{52}$$

Let us now compare the matrix properties for different choices of β and γ^2 with respect to pure gradients (which are affected by the instability of the field formulation). With the restriction

$$\begin{bmatrix} \vec{x}_A \\ \vec{x}_V \end{bmatrix} = \begin{bmatrix} [G]\vec{x}_\phi \\ \vec{x}_V \end{bmatrix} \quad \vec{x}_\phi, \vec{x}_V \in \mathbb{R}^{N_V}\,, \tag{53}$$

and the identities (48)–(52), the Rayleigh coefficient ρ_{AV} for the finite element matrix in (43) can be written as

$$\rho_{AV}(\vec{x}_\phi, \vec{x}_V) = \frac{-\begin{bmatrix} \vec{x}_\phi \\ \vec{x}_V \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} k_0^2 M_{VV} & \beta M_{VV} \\ \beta M_{VV} & \gamma^2 M_{VV} \end{bmatrix} \begin{bmatrix} \vec{x}_\phi \\ \vec{x}_V \end{bmatrix}}{\vec{x}_\phi^{\mathrm{T}} G^{\mathrm{T}} G \vec{x}_\phi + \vec{x}_V^{\mathrm{T}} \vec{x}_V}\,. \tag{54}$$

Let the eigenpairs of the matrix

$$M := \begin{bmatrix} k_0^2 I & \beta I \\ \beta\, I & \gamma^2 I \end{bmatrix} \tag{55}$$

be denoted by $(\theta_i, \hat{q}_i)$, where

$${}_1\theta_2 = \frac{\gamma^2 + k_0^2}{2} \pm \sqrt{\left(\frac{\gamma^2 + k_0^2}{2}\right)^2 + (\beta^2 - k_0^2\gamma^2)}\,, \tag{56}$$

$$\vec{q}_1 = \mathrm{col}[\beta I, (\theta_1 - k_0^2) I]\,, \qquad \hat{q}_1 = \vec{q}_1 / \|\vec{q}_1\|\,, \tag{57}$$

$$\vec{q}_2 = \mathrm{col}[(\theta_2 - \gamma^2) I, \beta n I]\,, \qquad \hat{q}_2 = \vec{q}_2 / \|\vec{q}_2\|\,. \tag{58}$$

Then, we can use the orthogonal matrix $[Q]$,

$$[Q] := [\hat{q}_1, \hat{q}_2]\,, \tag{59}$$

to rewrite the Rayleigh coefficient (54) as

$$\rho_{AV} = \frac{\begin{bmatrix} \vec{x}_\phi \\ \vec{x}_V \end{bmatrix}^{\mathrm{T}} [Q]^{\mathrm{T}} \begin{bmatrix} \theta_1 [M_{VV}] & 0 \\ 0 & \theta_2 [M_{VV}] \end{bmatrix} [Q] \begin{bmatrix} \vec{x}_\phi \\ \vec{x}_V \end{bmatrix}}{\vec{x}_\phi^{\mathrm{T}} G^{\mathrm{T}} G \vec{x}_\phi + \vec{x}_V^{\mathrm{T}} \vec{x}_V}\,. \tag{60}$$

Note that the denominator is positive definite and independent of the parameters β, γ, k_0^2. Matrix $[M_{VV}]$ is the Laplace matrix (47), which does not depend on the parameters either and $[Q]$ is orthogonal. For stability considerations, it is therefore sufficient to discuss the eigenvalues ${}_1\theta_2$. In the following, we will assume

$$\gamma^2 \geqslant 0\,. \tag{61}$$

Then, the eigenvalue θ_1 (using the positive root) is positive and bounded away from zero by

$$\theta_1(k_0^2) \geqslant \frac{\gamma^2}{2} + |\beta|\,, \tag{62}$$

provided that

$$\gamma^2 \geq \epsilon > 0 \quad \text{or} \quad |\beta| \geq \epsilon > 0\,. \tag{63}$$

For the eigenvalue θ_2 using the negative root, we have to distinguish three cases.

Case 1: $\beta^2 < k_0^2\gamma^2$. To avoid $\beta^2 \to 0$, $\theta_2 \to 0$ for $k_0^2 \to 0$, γ^2 must contain a k_0^{-2} term. But then, $\theta_1 \to \infty$ for $k_0^2 \to 0$. No stable solutions exist.

Case 2: $\beta^2 > k_0^2\gamma^2$. The eigenvalue θ_2 is negative and bounded from zero, provided that

$$\beta^2 - k_0^2\gamma^2 \geq \epsilon > 0\,. \tag{64}$$

Case 3: $\beta^2 = k_0^2\gamma^2$. This is the ungauged formulation with $\theta_2 \equiv 0$. This choice requires a solver that behaves robustly with respect to small perturbations to the right-hand-side vector and zero-eigenvectors of the matrix, such as the preconditioned conjugate gradient (PCG) method [7].

Let us shortly compare two specific variants of the formulation:

(i) *Lagrange multiplier method without penalty* ($\beta = 1$, $\gamma^2 = 0$): The method is stable. For the asymptotic behavior of the eigenvalues we get:

$$k_0^2 \ll 1\text{:} \quad {}_1\theta_2 \approx \pm 1 + \frac{k_0^2}{2}\,, \tag{65}$$

$$k_0^2 \gg 1\text{:} \quad \theta_1 \approx +k_0^2\,, \tag{66}$$

$$\theta_2 \approx -\frac{1}{k_0^2}\,. \tag{67}$$

(ii) *Ungauged potential approach with* $\beta = k_0$, $\gamma^2 = 1$: The eigenvalues are given by

$$\theta_1 = 1 + k_0^2\,, \tag{68}$$

$$\theta_2 = 0\,. \tag{69}$$

The asymptotic behavior of θ_1 is very similar in both cases. However, the eigenvalues θ_2 are quite different, which has important consequences with regard to iterative solvers. For the Lagrange multiplier method, $\theta_2 \neq 0$, so that the subspace corresponding to $\vec{q}_2$ ought to be considered in the design of the preconditioner. It follows from (66), (67) that, for large wavenumbers, the matrix condition number will grow $\propto \theta_1/\theta_2 \propto k_0^4$. In case of the ungauged formulation, the gradient subspace just implies that there are eigenvalues whose magnitudes grow $\propto k_0^2$. The eigenvector $\vec{q}_2$ describes the nullspace which does not require any preconditioning, and the fact that the dimension of the solution space has been reduced by N_V helps improve numerical convergence. For these reasons, we recommend the ungauged formulation when iterative solvers are to be used.

4.1. An incomplete Cholesky type preconditioner for the ungauged formulation

In our implementation of the PCG method, we have chosen an incomplete LU factorization (ILU) as the preconditioner. The preconditioner is derived from the static limit, where the finite element equations (43) for $\vec{A}$ and V are perfectly decoupled. We write:

$$[S_{EE}] - k_0^2[T_{EE}] \approx [S_{EE}] \approx [L_{AA}][U_{AA}]\,, \tag{70}$$

$$[M_{VV}] \approx [L_{VV}][U_{VV}]\,, \tag{71}$$

$$\begin{bmatrix} S_{EE} - k_0^2 T_{EE} & -k_0 M_{AV} \\ -k_0 M_{VA} & -M_{VV} \end{bmatrix} \approx \begin{bmatrix} [L_{AA}] & 0 \\ 0 & [L_{VV}] \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} \begin{bmatrix} [U_{AA}] & 0 \\ 0 & [U_{VV}] \end{bmatrix}, \tag{72}$$

where $[L_{AA}][U_{AA}]$ and $[L_{VV}][U_{VV}]$ are ILU decompositions of $[S_{EE}]$ and $[M_{VV}]$, respectively. Shifting techniques are applied to keep the factorization stable. In any practical finite element discretization, the vast majority of *Type B* eigenmodes has resonances far beyond the operating wavenumber. According to (22),

$$\frac{\vec{x}_i^{\,\mathrm{T}}[S_{EE}]\vec{x}_i}{\vec{x}_i^{\,\mathrm{T}}([S_{EE}] - k_0^2[T_{EE}])\vec{x}_i} = \frac{k_i^2}{k_i^2 - k_0^2} \approx 1 \quad \text{for } \vec{x}_i \text{ with } k_i^2 \gg k_0^2\,, \tag{73}$$

so (70) will be a good approximation for a large number of modes. However, the preconditioner is clearly unable to catch any negative eigenvalues related to cavity modes with resonances below the operating wavenumber. Hence, the efficiency of our solver will deteriorate when the problem domain becomes electrically large.

Regarding the applicability of PCG methods to singular systems, see [7], [8].

4.2. Simplified $\vec{A} - V$: an inexpensive formulation for high frequency applications

One disadvantage of the potential approach is that the system matrix contains almost twice as many non-zero entries as in case of the field formulation although the solution space for $\vec{E}$ remains the same. Also, the number of floating point operations per iteration is nearly doubled. The overhead is mainly caused by the coupling matrices $[M_{AV}]$ and $[M_{VA}]$ representing scalar-vector interactions. The Laplace matrix $[M_{VV}]$ is comparatively small and does not contribute much. The following implementation is designed for high-frequency applications and seeks a compromise between moderate computational costs and sufficient stability.

Let us introduce an alternative pair of potentials $(\vec{A}', V')$ by

$$\nabla \times \vec{A}' = -jk_0\eta_0[\mu_r]\vec{H}\,, \tag{74}$$

$$\vec{E} = \vec{A}' + \nabla V'\,. \tag{75}$$

We can now state an ungauged formulation as

$$\nabla \times [\mu_r]^{-1}\nabla \times \vec{A}' - [\epsilon_r](k_0^2\vec{A}' + \nabla V') = 0\,, \tag{76}$$

$$k_0^2\nabla \cdot [\epsilon_r](\vec{A}' + \nabla V') = 0\,, \tag{77}$$

$$\frac{-j}{k_0\eta_0}([\mu_r]^{-1}\nabla \times \vec{E}) \times \hat{n} = \overline{K}_E\,, \qquad \frac{\partial}{\partial n}([\epsilon_r]V') = 0 \quad \text{on } \Gamma_H\,, \tag{78}$$

$$\vec{A}' \times \hat{n} = \vec{0}\,, \qquad V' = 0 \quad \text{on } \Gamma_E\,. \tag{79}$$

The corresponding finite element equations read

$$\begin{bmatrix} [S_{EE}] - k_0^2[T_{EE}] & -k_0^2[M_{AV}] \\ -k_0^2[M_{VA}] & -k_0^2[M_{VV}] \end{bmatrix} \begin{bmatrix} \vec{x}'_A \\ \vec{x}'_V \end{bmatrix} = \begin{bmatrix} k_0\vec{r}_K \\ k_0\vec{r}_V \end{bmatrix}. \tag{80}$$

In practise, the elements of $[S_{EE}]$ and $[T_{EE}]$ are stored in a single matrix $[M_{EE}]$,

$$[M_{EE}] = [S_{EE}] - k_0^2[T_{EE}]\,. \tag{81}$$

Applying the identities (48)–(52), the equation system (80) can now be written as

$$[I \quad G]^{\mathrm{T}}[M_{EE}][I \quad G]\begin{bmatrix} \vec{x}'_A \\ \vec{x}'_V \end{bmatrix} = [I \quad G]^{\mathrm{T}}k_0\vec{r}_K\,. \tag{82}$$

Assuming that first order Lagrangian basis functions are used for the scalar potential and edge elements [2] for the vector potential, the elements of the gradient matrix $[G]$ can be determined from the following interpolation property: For a given vector basis function $\vec{W}_i$ defined along edge $\{mn\}$ from node m to node n, we have

$$\int_{edge} \vec{W}_i \cdot \mathrm{d}\vec{l} = \begin{cases} 1 & \text{along edge } \{mn\}\,, \\ 0 & \text{along all other edges}\,. \end{cases} \tag{83}$$

Therefore, the line integral of a gradient $\vec{A} \in \{\nabla V \,|\, V \in \mathcal{V}_0\}$ along edge $\{mn\}$ yields

$$\begin{aligned} \int_m^n \vec{A} \cdot \mathrm{d}\vec{l} &= \underbrace{\int_m^n a_i\vec{W}_i \cdot \mathrm{d}\vec{l}}_{} = \underbrace{\int_m^n \nabla V \cdot \mathrm{d}\vec{l}}_{} \\ &= \quad a_i \quad = V_n - V_m\,. \end{aligned} \tag{84}$$

Since the edge coefficient a_i is given by the difference of the scalar potential values at the starting and ending point, the gradient matrix has a maximum of two non-zero entries per row, namely

$$[G]_{im} = -1 \quad \text{and} \quad [G]_{in} = +1\,. \tag{85}$$

Thus, the implementation of $[G]$ requires just $2N_E$ pointers or indices. Compared to the field formulation, the overhead per matrix-times-vector multiplication has been reduced to just $2N_V$ additions plus $2N_V$ subtractions.

The preconditioner for our ILU-PCG solver is now written as

$$\begin{bmatrix} [L_{AA}] & 0 \\ 0 & [L_{VV}] \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & -k_0^2 I \end{bmatrix} \begin{bmatrix} [U_{AA}] & 0 \\ 0 & [U_{VV}] \end{bmatrix}. \tag{86}$$

We close this section with a look at stability. A comparison of (80) and (43) reveals that the present finite element matrix corresponds to

$$\beta = k_0^2\,, \tag{87}$$

$$\gamma^2 = k_0^2\,. \tag{88}$$

According to (57), the eigenvalue θ_1 will tend towards zero in the low frequency case. However, a stable form with $\beta = k_0$, $\gamma^2 = 1$, can be recovered by simple diagonal scaling. Our strategy has been to exploit the particular structure of the matrix factorization in (82) to improve efficiency and let the preconditioner do the proper rescaling. The stability limit of the simplified approach is reached, when the operating wavenumber (or the mesh size) becomes so small that the $[T_{EE}]$ matrix information incorporated in a given finite precision representation of $[M_{EE}]$ gets significantly distorted. Otherwise, the present scheme will exhibit very similar numerical properties as the basic formulation (43). In practical simulations of microwave devices, we have found the simplified method economic and reliable [9] (see Section 7.3, Fig. 4).

5. Extension to higher order elements—A partially gauged approach

The methods from Section 4 can be extended to conventional higher order elements in a straight-forward manner [10]. A better solution is to employ hierarchical basis functions [5] and make the formulation exploit their special properties to save memory and CPU time. We will demonstrate the idea for the $H_1(\text{curl})$ element space over tetrahedra [3].

Let simple, double and triple indices denote vertices, edges and facets, and ξ_p $(p = 1 \div 4)$ stand for the barycentric coordinates. Then, a hierarchical basis for $H_1(\text{curl})$ elements can be stated as follows:

$$\vec{w}_{ik} = \xi_i \nabla \xi_k - \xi_k \nabla \xi_i\,, \tag{89}$$

$$\vec{g}_{ik} = 4\nabla(\xi_i \xi_k)\,, \tag{90}$$

$$\vec{f}_{ikl} = 4\xi_i(\xi_k \nabla \xi_l - \xi_l \nabla \xi_k)\,, \quad (2 \text{ per facet})\,, \tag{91}$$

where $\vec{g}_{ik}$ represents the 6 edge basis functions of the lowest order scheme, augmented by $6 + 8$ higher order degrees of freedom, one $\vec{g}_{ik}$ function per edge mid-node and two $\vec{f}_{ikl}$ functions per face. To reflect the different types of basis functions, the $H_1(\text{curl})$ space can be written as

$$H_1(\text{curl}) = \mathcal{W}_0^e \oplus \mathcal{W}_0^{g2} \oplus \mathcal{W}_0^f\,, \tag{92}$$

$$\mathcal{W}^e = \text{span}\{\vec{w}_{ik}\}\,, \tag{93}$$

$$\mathcal{W}^{g2} = \text{span}\{\vec{g}_{ik}\}\,, \tag{94}$$

$$\mathcal{W}^f = \text{span}\{\vec{f}_{ikl}\}\,. \tag{95}$$

For the scalar potential, we will employ second order Lagrangian basis functions, i.e.,

$$W_i = \xi_i\,, \tag{96}$$

$$W_{ik} = 4\xi_i \xi_k \,, \tag{97}$$

$$\mathcal{V}^1 = \mathrm{span}\{W_i\} \,, \tag{98}$$

$$\mathcal{V}^2 = \mathrm{span}\{W_{ik}\} \,, \tag{99}$$

$$\mathcal{V} = \mathcal{V}^1 \oplus \mathcal{V}^2 \,. \tag{100}$$

Note that the mid-node basis functions $\vec{g}_{ik}$ are pure gradients. In the $\vec{A} - V$ formulation, equivalent second-order Lagrange basis functions W_{ik} are employed for the scalar potential anyway. It is therefore sufficient to have

$$\vec{A} \in \mathcal{W}_0^e \oplus \mathcal{W}_0^f \,, \tag{101}$$

i.e., the subspace $\mathcal{W}^{g2}$ can be dropped. The facial basis functions $\vec{f}_{ikl}$ are provided to enrich the range space of the curl operator; $\mathcal{W}_0^f$ does not contain any gradients. To avoid instabilities, it is therefore sufficient to take counter-measures in $\mathcal{W}_0^e$ only. We arrive at a partially gauged formulation, where ambiguities in $\vec{A}$ and ∇V are restricted to the gradients in $\mathcal{W}_0^e$. The electrical field $\vec{E}$ is now given by

$$\vec{E} = \underbrace{k_0 \vec{A}^e + \nabla V^1}_{\text{ungauged}} + k_0 \vec{A}^f + \nabla V^2 \,, \quad \vec{A}^e \in \mathcal{W}_0^e \,, \quad \vec{A}^f \in \mathcal{W}_0^f \,, \quad V^1 \in \mathcal{V}_0^1 \,, \quad V^2 \in \mathcal{V}_0^2 \,. \tag{102}$$

The finite element equations (36), (37) remain valid, and the resulting equation system takes the same form as in (43).

Again, an ILU-PCG solver is used in our computer implementation. To increase sparsity, all coupling terms between low and high order basis functions in the blocks $[S_{EE}]$ and $[M_{VV}]$ of the preconditioner (72) have been dropped. Numerical results are given in Section 7.3, Figs. 3 and 4.

6. A Schwartz-type solver for higher order elements

As said in Section 4.1, our ILU preconditioner cannot catch negative eigenvalues corresponding to physical modes with resonances below the operating wavenumber. We therefore expect the efficiency of our iterative solver to deteriorate when the number of such modes becomes significant, i.e., for electrically large structures (see [10] for some numerical results).

These negative eigenvalues depend very strongly on the resonance wave numbers of certain cavity modes which are ruled by the geometry and electromagnetic properties of the whole field domain. As a result, information describing these critical modes is spread over the entire finite element matrix rather than being localized in some small section hereof. We do not believe that conventional preconditioners such as diagonal scaling or ILU factorizations are very suitable in this case. Instead, we apply multi-grid ideas to gather critical information from the entire field domain.

We start from the observation that those cavity modes with low resonance wave numbers are also the ones with dominantly smooth field patterns, for which even the lowest order finite element scheme can yield reasonable approximations. Augmented hierarchical higher order basis functions will mainly contribute to rapidly oscillating modes, but add just minor corrections to smooth field patterns.

We suggest a similar PCG solver as before [11], but with a complete factorization replacing the ILU approximation for the lowest order scheme (edge vector basis functions and linear scalar unknowns). Obviously, the ungauged $\vec{A} - V$ formulation is unsuitable for direct solvers. For stability, either the Coulomb gauge (via a non-singular Lagrange multiplier method) or a tree-gauge [12] should be applied.

The success of the suggested method depends on whether the edge element scheme used in the complete factorization can catch the behavior of all physical modes associated with negative eigenvalues with sufficient accuracy. At present, we cannot give a theoretical limit for the maximum permissible mesh size, but numerical tests have shown that three to four elements per wavelength are sufficient, as long as the linear dimensions of the model are in the order of a few wavelengths. Early results for very large domains indicate that, even for

structures 20–30 wavelengths in size, five to six elements per wavelength are sufficient to prevent deterioration of numerical convergence (see Section 7.2 for details).

In practical applications, the Schwartz-type solver has proven to be very robust, even for electrically large problem domains. The limiting factor we see with the present implementation is that its computational complexity is approximately N^2, although with a very much smaller constant in front than in case of conventional direct solvers.

7. Numerical examples

7.1. Low frequency stability

To compare the numerical convergence characteristics of the field and potential formulation, a section of an air-filled rectangular waveguide has been investigated. (Geometrical data: length $l = 1$ m, width $w = 1$ m, height $h = 0.5\,m$.) The side walls were modeled by perfect electric conductors (PECs), one end of the waveguide was excited by the magnetic field pattern of the fundamental mode, TE_{10}, and the opposite end was terminated by a perfect magnetic conductor (PMC). A non-structured tetrahedral mesh of maximum edge length $h_{\max} = 0.13$ m was used for the discretization. For the $\vec{A} - V$ method, the preconditioner was as in (72), whereas in case the field formulation, just the $[S_{EE}]$ block was taken. Edge basis functions of lowest order were applied. The PCG iteration was terminated when the norm of the relative residual vector for the electrical field had reached the value of $\|\vec{r}_{E,\mathrm{rel}}\| \leq 10^{-6}$. Fig. 1 shows that the iteration counts for the $\vec{A} - V$ remained flat over a frequency range of 6 decades, whereas the field formulation ran into severe stability problems in the low frequency region. At the high frequency end, neither formulation was satisfactory. This problem was mainly caused by insufficient discretization and could be eliminated when the mesh was refined (not shown here).

7.2. Maximum permissible mesh size

The model of a parallel plate waveguide has been used to find out about the maximum permissible mesh size for the Schwartz type solver of Section 6. The geometry and material parameters were given by: width $w = 1$ m, height $h = 1$ m, length $l = 100$ m; permeability $\mu_r = 1$, permittivity $\epsilon_r = 1$. The waveguide was excited at one end by a plane wave. On the opposite end, a PEC was imposed. The PCG termination criterion was the same as in Section 7.1. Finite element calculations were carried out over a wide frequency range, keeping a constant mesh size of $h_{\max} = 1$ m. Fig. 2 presents the PCG iteration count as a function of the number of elements per wavelength λ. The curve has a moderate and nearly constant slope for finer discretizations, but starts growing

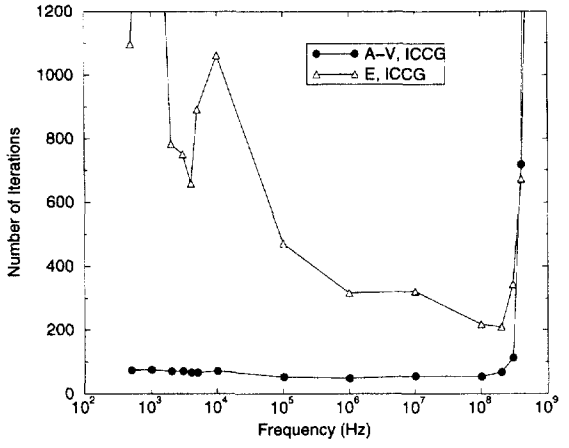

Fig. 1. PCG iteration counts for the $\vec{E}$ field and $\vec{A} - V$ potential formulation. Results for a rectangular waveguide model.

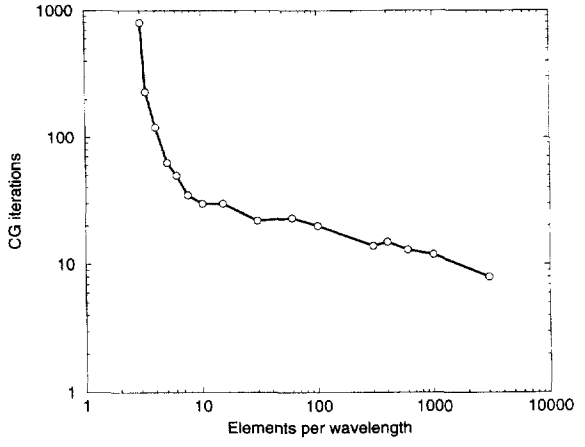

Fig. 2. PCG iteration counts versus elements per wavelength for the Schwartz-type solver. Results for a parallel plate waveguide model.

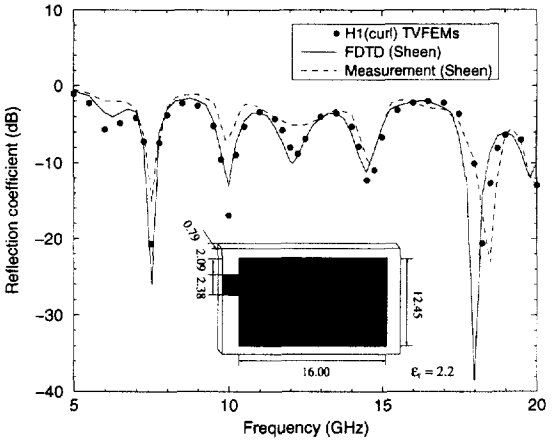

Fig. 3. A rectangular microwave patch antenna. Dimensions in mm. Reference results taken from Sheen et al. [13].

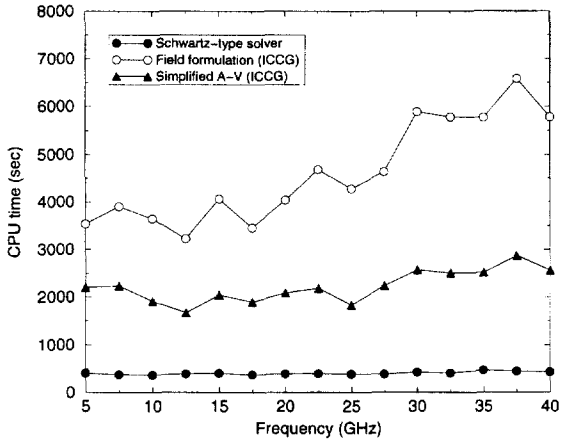

Fig. 4. CPU time versus frequency for the patch antenna. Workstation: HP 900/C100, clock speed 120 MHz.

very rapidly once the mesh density falls short of approximately 5 elements per wavelength. Note that, at this point, the electrical length of the structure is already 20λ.

7.3. A microstrip patch antenna

The rectangular microstrip patch antenna depicted in Fig. 3 has been analyzed. A comparison between the present finite element results for the return loss and solutions from the literature [13] is given in Fig. 3. $H_1(\mathrm{curl})$ element discretization of the antenna resulted in 93 738 vector unknowns and approximately 1.9 million non-zero matrix entries. The field problem was solved by the $\vec{E}$ field formulation, the partially gauged method based on the simplified $\vec{A} - V$ approach, and the Schwartz-type solver. A CPU time comparison is given in Fig. 4. For the field formulation, convergence is poor and deteriorates strongly with increasing frequency. The partially gauged $\vec{A} - V$ approach is significantly faster and exhibits a less pronounced dependence on the operating frequency. The Schwartz method is the most robust approach and, for the given number of unknowns, the fastest, too. Notice that the iteration count is almost independent of the operating frequency.

8. Conclusions

In the present paper we have shown that certain instabilities inherent in the Nedelec-type finite element implementation of the $\vec{E}$ field formulation can be overcome by a family of Lagrange multiplier methods with or without penalty term. These methods can be interpreted as coupled vector and scalar potential methods imposing the Coulomb gauge or no gauge at all. The ungauged variant has been found particularly useful for iterative solvers, because it keeps the dimension of the solution space at a minimum. There exists a simplified version of this approach, which has some limitations with regard to stability but excels by its moderate computational requirements. For higher order finite elements, the properties of hierarchical bases can be utilized to derive a partially gauged potential formulation which is faster and less memory consuming than a straight-forward extension of the ungauged approach. Due to a shortcoming in the block ILU preconditioners used with our methods, computational efficiency deteriorates when the solution domain becomes electrically large. To overcome this problem, a Schwartz-type solver has been suggested. This method takes advantage of multigrid ideas and hierarchical finite element bases and stands out for its robustness.

References

[1] J. Nedelec, Mixed finite elements in $\Re^3$, Numer. Math. 35 (1980) 315–341.
[2] A. Bossavit, Whitney forms: a class of finite elements for three-dimensional computations in electromagnetism, IEE Proc. 135(pt. A) (1988) 493–500.

[3] J.F. Lee, D.K. Sun and Z.J. Cendes, Tangential vector finite elements for electromagnetic field computation, IEEE Trans. Magn. 27 (1991) 4032–4035.
[4] H. Kanayama, H. Motoyama, K. Endo and F. Kikuchi, Three-dimensional magnetostatic analysis using Nedelec's elements, IEEE Trans. Magn. 26 (1990) 682–685.
[5] J.P. Webb and B. Forghani, Hierarchical scalar and vector tetrahedra, IEEE Trans. Magn. 29 (1993) 1495–1498.
[6] G.H. Golub and C.F. van Loan, Matrix Computations, 3rd edition (The Johns Hopkins University Press, Baltimore and London, 1996).
[7] E.F. Kaasschieter, Preconditioned conjugate gradients for solving singular systems, J. Comput. Appl. Math. 24 (1988) 265–275.
[8] Y. Notay, Incomplete factorizations of singular linear systems, BIT 29 (1989) 682–702.
[9] R. Dyczij-Edlinger, G. Peng and J.-F. Lee, A fast vector potential method using tangentially continuous vector finite elements, IEEE Trans. Microwave Theory Tech. 46(6) (1998) 863–868.
[10] R. Dyczij-Edlinger and O. Biro, A joint vector and scalar potential formulation for driven high frequency problems using hybrid edge and nodal finite elements, IEEE Trans. Microwave Theory Tech. 44(1) (1996) 15–23.
[11] G. Peng, Multigrid preconditioning in solving time-harmonic wave propagation problems using tangential vector finite elements, Ph.D. Thesis, ECE Dept., Worcester Polytechnic Institute, Worcester, MA, Nov. 1997.
[12] R. Albanese and A. Rubinacci, Solution of three dimensional eddy current problems by integral and differential methods, IEEE Trans. Magn. 24(1) (1988) 98–101.
[13] D.M. Sheen, S.M. Ali, M.D. Abouzahra, J.A. Kong, Application of the three-dimensional finite-difference time-domain method to the analysis of planar microstrip circuits, IEEE Trans. Microwave Theory Tech. 38(7) (1990) 849–857.