A Joint Vector and Scalar Potential Formulation for Driven High Frequency Problems Using Hybrid Edge and Nodal Finite Elements

Romanus Dyczij-Edlinger and Oszkar Biro

Abstract—**An advanced *A-V* method employing edge-based finite elements for the vector potential *A* and nodal shape functions for the scalar potential *V* is proposed. Both gauged and ungauged formulations are considered. The novel scheme is particularly well suited for efficient iterative solvers such as the preconditioned conjugate gradient method, since it leads to significantly faster numerical convergence rates than pure edge element schemes. In contrast to nodal finite element implementations, spurious solutions do not occur and the inherent singularities of the electromagnetic fields in the vicinity of perfectly conducting edges and corners are handled correctly. Several numerical examples are presented to verify the suggested approach.**

I. Introduction

OVER THE LAST few years, edge-based finite elements have widely been accepted as a reliable tool for the computation of electromagnetic fields. While nodal elements enforce continuity of all vector components across element interfaces, edge elements ensure tangential continuity only. Thus they can cope with material discontinuities as well as sharp metallic edges and corners in a natural way. Edge elements provide well-defined subspaces for irrotational fields and are clearly related to the classes of nodal and facet elements [1]. The most important thing, however, is that the occurrence of spurious solutions can be avoided [2]. Because of these special features, edge elements seem to be ideally suited for formulations where the unknown is the electric or magnetic field. In the case of eigenvalue problems, such methods are very powerful indeed [3]–[5]. However, in the driven case, they lead to rather ill-conditioned, indefinite matrices, which result in poor convergence rates of iterative solvers. Frequently, the widespread incomplete Cholesky conjugate gradient (ICCG) method does not converge at all [6].

Fortunately, the number of unknowns can be kept rather small for a wide range of waveguide discontinuity problems by exploiting techniques that allow the finite element mesh to be truncated very close to the discontinuity region, e.g., the boundary-marching method [7] or transfinite elements [8]. In these cases, direct solvers can be used and $\boldsymbol{E}$ or $\boldsymbol{H}$ based formulations perform very well. However, practical configurations may also comprise scatterers or discontinuities that are electrically large, highly inhomogeneous, or complicated in shape. Such structures typically lead to very large numbers of unknowns and, in consequence, to enormous computational costs when direct solvers are in use. Under these circumstances, iterative methods become very attractive or even indispensable. The slow numerical convergence of conventional formulations is a serious shortcoming in this case.

In the present work, we point out a major cause of the convergence problem and present a vector potential formulation that requires far fewer iterations than the $\boldsymbol{E}$ formulation, while retaining the same level of accuracy. Although a similar method was suggested by Albanese and Rubinacci [9] for eddy current fields as early as 1988, this technique has never been applied to wave propagation problems. While in their paper a graph theory approach is employed to ensure uniqueness, we recommend the ungauged formulation. For completeness, we also describe an inexpensive gauge procedure, similar to the Lagrange multiplier method for static fields [10]. Although the proposed method is applicable to the lossy case as well, the theory presented in this paper is primarily addressed to lossless, nonradiating structures and the proofs given below hold for real matrices only.

Improved numerical convergence is not the only benefit of the new approach. It also enabled us to solve problems for which the $\boldsymbol{E}$ method had not converged at all. (See Section V.) The trade-off is the additional scalar field, resulting in approximately twice as many nonzero matrix entries as the $\boldsymbol{E}$ formulation. We do not recommend to use the *A-V* formulation together with a direct solver, because pure edge element schemes are computationally less expensive in this case.

Several numerical examples and comparisons are presented to demonstrate the merits of the suggested approach.

II. Notations and Restrictions

In the following, a bounded contractible domain $\Omega \in \mathfrak{R}^3$ is considered. The abbreviation Γ stands for the surface of Ω and $\boldsymbol{n}$ for the outward unit normal on Γ. We write L^2 ($\boldsymbol{L}^2$) for the space of square integrable scalar (vector) functions over Ω, span ($\mathcal{B}$) for the space spanned by a basis $\mathcal{B}$, ran( ) for the range space of an operator, and ker( ) for the nullspace. As usual, the magnetic flux density is denoted by $\boldsymbol{B}$, the electric

Manuscript received August 16, 1994; revised October 2, 1995. This work was supported by the Fonds zur Foerderung der wissenschaftlichen Forschung, Austria under Grant PHY-8493.

R. Dyczij-Edlinger was with Graz University of Technology, Austria. He is now with Motorola CCRL, IL02–EA901, Schaumburg, IL 60195–1078 USA.

O. Biro is with Graz University of Technology, Kopernikusgasse 24, A-8010 Graz, Austria.

Publisher Item Identifier S 0018-9480(96)00483-8.

and magnetic fields by $\boldsymbol{E}$ and $\boldsymbol{H}$, the angular frequency by ω, the magnetic permeability by μ, the electric permittivity by ϵ, and the conductivity by σ. Throughout Ω, we assume $\mu > 0$, $\epsilon > 0$, and $\sigma \geq 0$. The considered boundary conditions are

$$\overline{\boldsymbol{E}}_t \quad \text{given on } \Gamma_E, \tag{1}$$

$$\overline{\boldsymbol{H}}_t \quad \text{given on } \Gamma_H \tag{2}$$

where $\overline{\boldsymbol{E}}_t$ and $\overline{\boldsymbol{H}}_t$ stand for the tangential components of $\boldsymbol{E}$ and $\boldsymbol{H}$, Γ_E is a nonempty connected subset of Γ, and Γ_H is the complement of Γ_E. The computational domain Ω is subdivided into finite elements equipped with both scalar and vector degrees of freedom. We denote the resulting set of nodal shape functions W by $\mathcal{N}_0$ and the set of edge shape functions $\boldsymbol{W}$ by $\mathcal{E} = \mathcal{E}_0 \cup \mathcal{E}_D$, where

$$\mathcal{N}_0 = \{W_i \in C^0(\Omega) \text{ for } i = 1 \cdots K \mid W_i = 0 \text{ on } \Gamma_E\}, \tag{3}$$

$$\mathcal{E}_0 = \{\boldsymbol{W}_i \in \boldsymbol{L}^2(\Omega) \text{ for } i = 1 \cdots N \mid \operatorname{curl} \boldsymbol{W}_i \in \boldsymbol{L}^2(\Omega), \; \boldsymbol{W}_i \times \boldsymbol{n} = \boldsymbol{0} \text{ on } \Gamma_e\}, \tag{4}$$

$$\mathcal{E}_D = \left\{\boldsymbol{W}_D \in \boldsymbol{L}^2(\Omega) \mid \operatorname{curl} \boldsymbol{W}_D \in \boldsymbol{L}^2(\Omega), \; \exists \overline{e}_D \in C : \sum_D \overline{e}_D \boldsymbol{W}_D \times \boldsymbol{n} = \overline{\boldsymbol{E}}_t \times \boldsymbol{n} \text{ on } \Gamma_E\right\}. \tag{5}$$

In (5), $\boldsymbol{C}$ stands for the set of complex numbers. Since the shape functions $\boldsymbol{W}_D \in \mathcal{E}_D$ are reserved for the incorporation of inhomogeneous Dirichlet boundary conditions, we end up with K nodal and N edge variables. For compatibility [1], the spaces $\mathcal{W}_0^0 = \operatorname{span}(\mathcal{N}_0)$ and $\mathcal{W}_0^1 = \operatorname{span}(\mathcal{E}_0)$ need to satisfy the relation

$$\boldsymbol{\nu} \in \mathcal{W}_0^1; \; \ker(\operatorname{curl} \boldsymbol{\nu}) = \{\operatorname{grad} \varphi \mid \varphi \in \mathcal{W}_0^0\}. \tag{6}$$

III. Analysis of the E-Formulation

The well-known edge element scheme comes directly from the PDE

$$\operatorname{curl}(\mu^{-1} \operatorname{curl} \boldsymbol{E}) - \omega^2 \epsilon \boldsymbol{E} = \boldsymbol{0} \tag{7}$$

and leads to the (approximate) solution $\tilde{\boldsymbol{E}} \in \operatorname{span} \mathcal{E}$ of the corresponding weak formulation

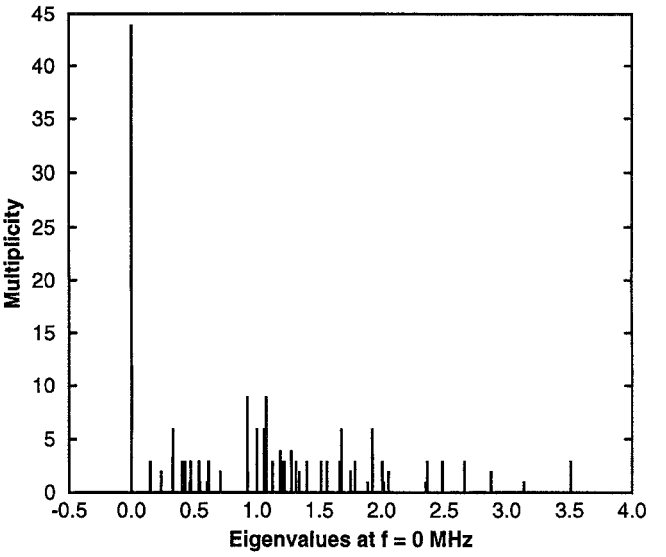
$$\forall \boldsymbol{W}_i \in \boldsymbol{W}_0^1; \quad \int_\Omega [\mu^{-1} \operatorname{curl} \tilde{\boldsymbol{E}} \cdot \operatorname{curl} \boldsymbol{W}_i - \omega^2 \epsilon \tilde{\boldsymbol{E}} \cdot \boldsymbol{W}_i] \, d\Omega = -j\omega \int_{\Gamma_H} (\overline{\boldsymbol{H}}_t \times \boldsymbol{n}) \cdot \boldsymbol{W}_i \, d\Gamma. \tag{8}$$

For the resulting system of linear equations we write

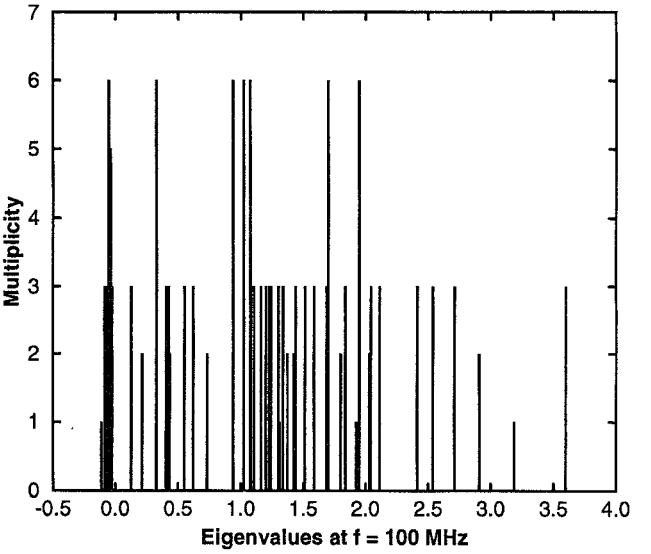
$$(\boldsymbol{F} - \omega^2 \boldsymbol{G}) \cdot \boldsymbol{e} = \boldsymbol{r}, \tag{9}$$

$$f_{ik} = \int_\Omega \mu^{-1} \operatorname{curl} \boldsymbol{W}_i \cdot \operatorname{curl} \boldsymbol{W}_k \, d\Omega, \quad \boldsymbol{W}_i, \boldsymbol{W}_k \in \mathcal{E}_0, \tag{10}$$

$$g_{ik} = \int_\Omega \epsilon \boldsymbol{W}_i \cdot \boldsymbol{W}_k \, d\Omega \tag{11}$$



(a)



(b)

Fig. 1. (a) $\boldsymbol{E}$ formulation: Spectrum of a diagonally scaled finite element matrix in the static case. (b) $\boldsymbol{E}$ formulation: Spectrum of a diagonally scaled finite element matrix at $f = 100$ MHz.

where $\boldsymbol{F}$ and $\boldsymbol{G}$ are symmetric matrices, f_{ik} and g_{ik} denote the elements of $\boldsymbol{F}$ and $\boldsymbol{G}$. $\boldsymbol{e}$ stands for the vector of unknown coefficients, and $\boldsymbol{r}$ for the right-hand side vector. Due to (6) and (10), the matrix $\boldsymbol{F}$ is positive semidefinite and its nullspace is spanned by K linearly independent vectors corresponding to all irrotational fields representable. The matrix $\boldsymbol{G}$ is positive definite. The eigenvalues of $\boldsymbol{F} - \omega^2 \boldsymbol{G}$ can be classified as follows:

Type A: In the static case, $\omega = 0$, K eigenvalues are zero. When the frequency is increased, they all move into the negative range. Thus, the matrix becomes indefinite and rather ill-conditioned. Fig. 1(a) and (b) illustrates this effect.

Type B: The remaining $N - K$ eigenvalues are related to the physical resonances of the considered structure. As long as the

lowest nonzero resonant frequency v_{K+1} of the corresponding eigenvalue problem

$$(\boldsymbol{F} - v^2\boldsymbol{G}) \cdot \boldsymbol{e} = \boldsymbol{0} \tag{12}$$

is larger than ω, they are all positive. On the other hand, each resonant frequency less than ω results in an additional negative eigenvalue of $\boldsymbol{F}-\omega^2\boldsymbol{G}$. We remark that the number of negative type B-eigenvalues is by far less than K. For a mathematical proof, see Lemma I in the Appendix.

From the fact that, for any practical finite element discretization, the main diagonal entries of $\boldsymbol{F} - \omega^2\boldsymbol{G}$ are all positive, it is clear that diagonal scaling preconditioners cannot account for any of the negative eigenvalues. In case of the widespread ICCG solvers, the large number of negative eigenvalues is also very cumbersome: Under such conditions, Cholesky-like matrix factorizations are rather sensitive to perturbations [11] and therefore the neglection of fill-in entries, which is characteristic for incomplete matrix decompositions, strongly affects the performance of the preconditioner. In our experiments, the usual level-0-factorizations have shown to be extremely inaccurate or even unstable, as indicated by the excessive growth of matrix elements and huge spectral condition numbers. More sophisticated schemes employing drop tolerances for off diagonal entries may help, but these methods are computationally expensive. From the theoretical point of view, the standard implementation of the CG method may become unstable as well [12], but in practice this seems to be less important.

IV. $\boldsymbol{A}$-V Description

Our basic idea to prevent poor numerical convergence is to avoid the displacement of the zero-eigenvalues, as described in Section III. To this aim, we propose an ungauged potential formulation employing edge elements for the vector field $\boldsymbol{A}$ and nodal shape functions for the scalar potential V. For completeness, a gauged version is also presented.

As usual, the potentials are defined by

$$\boldsymbol{A}\text{:}\quad \operatorname{curl}\boldsymbol{A} = \boldsymbol{B}, \tag{13}$$

$$V\text{:}\quad -j\omega(\operatorname{grad} V + \boldsymbol{A}) = \boldsymbol{E}. \tag{14}$$

Then, Maxwell's equations and the boundary conditions (1) and (2) lead to the transformed boundary value problem

$$\operatorname{curl}(\mu^{-1}\operatorname{curl}\boldsymbol{A}) - \omega^2\epsilon(\boldsymbol{A} + \operatorname{grad} V) = \boldsymbol{0}, \tag{15}$$

$$\boldsymbol{A} \times \boldsymbol{n} = j/\omega(\overline{\boldsymbol{E}}_t \times \boldsymbol{n}), \qquad V = 0 \quad \text{on } \Gamma_E, \tag{16}$$

$$(\mu^{-1}\operatorname{curl}\boldsymbol{A}) \times \boldsymbol{n} = \overline{\boldsymbol{H}}_t \times \boldsymbol{n} \quad \text{on } \Gamma_H. \tag{17}$$

The corresponding discrete problem consists of finding a solution ($\boldsymbol{A} \in \operatorname{span}\mathcal{E}$, $V \in \operatorname{span}\mathcal{N}$) to the weak formulation

$$\forall \boldsymbol{W}_i \in \mathcal{W}_0^1;$$
$$\int_\Omega [\mu^{-1}\operatorname{curl}\boldsymbol{A} \cdot \operatorname{curl}\boldsymbol{W}_i - \omega^2\epsilon(\boldsymbol{A} + \operatorname{grad} V) \cdot \boldsymbol{W}_i]\,\mathrm{d}\Omega = \int_{\Gamma_H} (\overline{\boldsymbol{H}}_t \times \boldsymbol{n}) \cdot \boldsymbol{W}_i\,\mathrm{d}\Gamma \tag{18}$$

that satisfies the boundary condition (16). To make the finite element matrix symmetric, we state (18) once more, for a restricted set of weighting functions $\{\operatorname{grad} W_i \mid W_i \in \mathcal{W}_0^0\} \subset \mathcal{W}_0^1$. Additionally, we introduce the parameter $\alpha \in \mathfrak{R}$ to provide a gauge. Thus, we have

$$\forall W_i \in \mathcal{W}_0^0; \qquad \int_\Omega [-\omega^2\epsilon(\boldsymbol{A} + \alpha^2 \operatorname{grad} V) \cdot \operatorname{grad} W_i]\,\mathrm{d}\Omega = \int_{\Gamma_H} (\overline{\boldsymbol{H}}_t \times \boldsymbol{n}) \cdot \operatorname{grad} W_i\,\mathrm{d}\Gamma. \tag{19}$$

When the variables assigned to the scalar potential are numbered first, the resulting system of equations is of the form

$$\boldsymbol{M}(\alpha) \cdot \begin{bmatrix} \boldsymbol{v} \\ \boldsymbol{a} \end{bmatrix} = \begin{bmatrix} \boldsymbol{r}_V \\ \boldsymbol{r}_A \end{bmatrix}, \tag{20}$$

$$\boldsymbol{M}(\alpha) = \begin{bmatrix} -\alpha^2\boldsymbol{M}_{VV} & -\boldsymbol{M}_{VA} \\ -\boldsymbol{M}_{VA}^T & \boldsymbol{F} - \omega^2\boldsymbol{G} \end{bmatrix} \tag{21}$$

where $\boldsymbol{v}$ and $\boldsymbol{a}$ denote the vectors of nodal and edge unknowns. Due to (16), (18), and (19), the elements of the submatrices $\boldsymbol{M}_{VV}$, $\boldsymbol{M}_{VA}$, and the right-hand side vectors $\boldsymbol{r}_V$, $\boldsymbol{r}_A$ are given by

$$m_{VV,ik} = \omega^2 \int_\Omega \epsilon \operatorname{grad} W_k \cdot \operatorname{grad} W_i\,\mathrm{d}\Omega, \qquad W_i,\, W_k \in \mathcal{N}_0, \tag{22}$$

$$m_{VA,ik} = \omega^2 \int_\Omega \epsilon \boldsymbol{W}_k \cdot \operatorname{grad} W_i\,\mathrm{d}\Omega, \qquad W_i \in \mathcal{N}_0,\, \boldsymbol{W}_k \in \mathcal{E}_0, \tag{23}$$

$$r_{V,i} = \int_{\Gamma_H} (\overline{\boldsymbol{H}}_t \times \boldsymbol{n}) \cdot \operatorname{grad} W_i\,\mathrm{d}\Gamma, \qquad W_i \in \mathcal{N}_0, \tag{24}$$

$$r_{A,i} = \int_{\Gamma_H} (\overline{\boldsymbol{H}}_t \times \boldsymbol{n}) \cdot \boldsymbol{W}_i \mathrm{d}\Gamma - \sum_D \overline{a}_D \int_\Omega (\mu^{-1}\operatorname{curl}\boldsymbol{W}_D \cdot \operatorname{curl}\boldsymbol{W}_i - \omega^2 \boldsymbol{W}_D \cdot \boldsymbol{W}_i)\,\mathrm{d}\Omega, \qquad \boldsymbol{W}_D \in \mathcal{E}_D,\quad \boldsymbol{W}_i \in \mathcal{E}_0. \tag{25}$$

The matrices $\boldsymbol{F}$ and $\boldsymbol{G}$ are defined as in Section III and, by analogy to (5), the constants $\overline{a}_D$ are used to impose the Dirichlet boundary conditions (16) on Γ_E. The formulation is ungauged for $\alpha^2 = 1$, because, in this case, any pair ($V_0 \in \mathcal{W}_0^0$, $\boldsymbol{A}_0 = -\operatorname{grad} V_0$) may be added to a given solution of (18) and (19). It can be shown easily that all other values of α enforce

$$\forall W_i \in \mathcal{W}_0^0; \qquad (\alpha^2 - 1)\int_\Omega \omega^2\epsilon \operatorname{grad} V \cdot \operatorname{grad} W_i\,\mathrm{d}\Omega = 0 \tag{26}$$

with the unique solution $V \equiv 0$. We remark that this method of gauging is very similar to the Lagrange multiplier method proposed in [10].

A. Factorization of the Finite Element Matrix

Let us consider the symmetric decomposition

$$\begin{aligned} \boldsymbol{M}(\alpha) &= \begin{bmatrix} -\alpha^2 \boldsymbol{M}_{VV} & -\boldsymbol{M}_{VA} \\ -\boldsymbol{M}_{VA}^T & \boldsymbol{F} - \omega^2 \boldsymbol{G} \end{bmatrix} \\ &= \begin{bmatrix} \alpha \boldsymbol{U}_{VV}^T & \boldsymbol{0} \\ \frac{1}{\alpha} \boldsymbol{U}_{VA}^T & \boldsymbol{U}_{AA}^T \end{bmatrix} \cdot \begin{bmatrix} -\boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{D}_{AA} \end{bmatrix} \\ &\quad \cdot \begin{bmatrix} \alpha \boldsymbol{U}_{VV} & \frac{1}{\alpha} \boldsymbol{U}_{VA} \\ \boldsymbol{0} & \boldsymbol{U}_{AA} \end{bmatrix} \end{aligned} \tag{27}$$

where $\boldsymbol{U_{VV}}$ and $\boldsymbol{U_{AA}}$ are upper triangular matrices, $\boldsymbol{U_{VA}}$ is a rectangular $K \times N$ matrix, and $\boldsymbol{I}$ is the unity matrix. The matrix $\boldsymbol{D_{AA}}$ is block diagonal with blocks of dimension one or two and serves to keep $\boldsymbol{U_{AA}}$ real-valued in the indefinite case [13]. From (27) we obtain

$$\boldsymbol{U}_{VV}^T \cdot \boldsymbol{U}_{VV} = \boldsymbol{M}_{VV}, \tag{28}$$

$$\boldsymbol{U}_{VV}^T \cdot \boldsymbol{U}_{VA} = \boldsymbol{M}_{VA}, \tag{29}$$

$$\begin{aligned} \tilde{\boldsymbol{M}}_{AA} &= \boldsymbol{F} - \omega^2 \boldsymbol{G} + \frac{1}{\alpha^2} \boldsymbol{U}_{VA}^T \cdot \boldsymbol{U}_{VA} \\ &= \boldsymbol{F} - \omega^2 \boldsymbol{G} + \frac{1}{\alpha^2} \boldsymbol{M}_{VA}^T \cdot \boldsymbol{M}_{VV}^{-1} \cdot \boldsymbol{M}_{VA}, \end{aligned} \tag{30}$$

$$\boldsymbol{U}_{AA}^T \cdot \boldsymbol{D}_{AA} \cdot \boldsymbol{U}_{AA} = \tilde{\boldsymbol{M}}_{AA}. \tag{31}$$

Due to (22), $\boldsymbol{M_{AA}}$ is always positive definite. Moreover, for $0 < \omega < \upsilon_{K+1}$ and $0 < \alpha^2 < 1$, even the matrix $\tilde{\boldsymbol{M}}_{\boldsymbol{AA}}$ is positive definite, or semidefinite for $\alpha^2 = 1$. A more detailed mathematical formulation can be found in Lemma II in the Appendix. Thus, the great difference to the $\boldsymbol{E}$ formulation is that, at least for frequencies below the first resonance and $0 < \alpha^2 < 1$, the factorization (27) is based on numerically stable processes, namely the Cholesky decompositions of the positive definite submatrices $\boldsymbol{M_{VV}}$ and $\tilde{\boldsymbol{M}}_{\boldsymbol{AA}}$.

B. Preconditioner

The considered incomplete Cholesky factorization is based on the matrix decomposition (27)–(31) and a positive definite approximation to $\tilde{\boldsymbol{M}}_{\boldsymbol{AA}}$. Denoting approximate matrices by apostrophes, we get

$$\boldsymbol{M}(\alpha) \approx \boldsymbol{U}'^T \cdot \boldsymbol{D}' \cdot \boldsymbol{U}', \tag{32}$$

$$\boldsymbol{D}' = \begin{bmatrix} -\boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I} \end{bmatrix}, \tag{33}$$

$$\boldsymbol{U}' = \begin{bmatrix} \alpha \boldsymbol{U}'_{VV} & \frac{1}{\alpha} \boldsymbol{U}'_{VA} \\ 0 & \boldsymbol{U}'_{AA} \end{bmatrix}, \tag{34}$$

$$\boldsymbol{U}'^T_{VV} \cdot \boldsymbol{U}'_{VV} \approx \boldsymbol{M}_{VV}, \tag{35}$$

$$\boldsymbol{U}'^T_{VV} \cdot \boldsymbol{U}'_{VA} \approx \boldsymbol{M}_{VA}, \tag{36}$$

$$\tilde{\boldsymbol{M}}'_{AA} = \boldsymbol{F} - \omega^2 \boldsymbol{G} + 1/\alpha^2 \boldsymbol{U}'^T_{VA} \cdot \boldsymbol{U}'_{VA} \approx \tilde{\boldsymbol{M}}_{AA}, \tag{37}$$

$$\boldsymbol{U}'^T_{AA} \cdot \boldsymbol{U}'_{AA} \approx \tilde{\boldsymbol{M}}''_{AA} \approx \tilde{\boldsymbol{M}}'_{AA} \approx \tilde{\boldsymbol{M}}_{AA}. \tag{38}$$

To keep the number of nonzero entries in $\boldsymbol{U}'$ low, we perform a level-0-factorization, i.e., all fill-ins are discarded. Since $\boldsymbol{M_{VV}}$ is positive definite, the calculation of $\boldsymbol{U'_{VV}}$ and $\boldsymbol{U'_{VA}}$ is unproblematic. On the other hand, the matrix $\tilde{\boldsymbol{M}}'_{\boldsymbol{AA}}$ may become indefinite (even for $0 < \alpha^2 < 1$) and cause stability problems. The reasons for this are: a) In (37), $\tilde{\boldsymbol{M}}'_{AA}$ is derived from the indefinite matrix $\boldsymbol{F} - \omega^2 \boldsymbol{G}$ by using the incomplete factor $\boldsymbol{U'_{VA}}$ instead of $\boldsymbol{U_{VA}}$. b) In the ungauged case, the matrix $\tilde{\boldsymbol{M}}_{\boldsymbol{AA}}$ to be approximated is singular. c) For $\omega > \upsilon_{K+1}$, even $\tilde{\boldsymbol{M}}_{\boldsymbol{AA}}$ is indefinite.

To overcome these difficulties, we factorize a positive definite approximation $\tilde{\boldsymbol{M}}''_{\boldsymbol{AA}}$ instead of $\tilde{\boldsymbol{M}}'_{\boldsymbol{AA}}$ itself. This is permissible as long as Ω is not too large, because then (in contrast to the $\boldsymbol{E}$ formulation) the number of negative eigenvalues of $\tilde{\boldsymbol{M}}_{\boldsymbol{AA}}$ which are improperly represented is rather small. (See Lemma II in the Appendix.)

C. Gauged Versus Ungauged Formulations

Assumption (6) guarantees that all gauge transformations of $\boldsymbol{A}$ and grad V are performed in a K-dimensional subspace of $\mathcal{W}_0^1$ [1]. For this reason, all values of α lead to the very same electric field as the $\boldsymbol{E}$ formulation discussed in Section III, which can be interpreted as a rescaled version of the $\boldsymbol{A}$-V approach with the particular gauge $V \equiv 0$. Nevertheless, from the numerical point of view, the question of how to optimally adjust the parameter α in (19) is of great importance. In particular, the reduced matrix $\tilde{\boldsymbol{M}}_{\boldsymbol{AA}}$ will become more and more ill-conditioned as α approaches 1, because K eigenvalues tend toward zero then. Also, (30) implies that very small values for α should be avoided.

However, in the case where $\alpha = 1$, $\tilde{\boldsymbol{M}}_{\boldsymbol{AA}}$ is *exactly* singular and, since the solution space ran $(\boldsymbol{M})$ is of dimension N at most, we may expect the lowest number of CG iteration steps possible with this method [14]. For this reason, the ungauged formulation is preferable for practical calculations.

For completeness, we mention an alternative way of gauging, namely by specifying a tree through the edges and eliminating the corresponding degrees of freedom [9]. We have not tried this method yet, but we are suspicious that it would lead to poor numerical convergence [15].

D. Computer Implementation

In our experimental software higher order brick elements with 20 nodes and 36 edges [16] are employed. At present, we either use no reordering strategy or we rearrange the variables in $\boldsymbol{v}$ and $\boldsymbol{a}$ (see Section IV) by applying the minimum degree algorithm [17] to the graph sets of $\boldsymbol{M_{VV}}$ and $\boldsymbol{F} - \omega^2 \boldsymbol{G}$, respectively. To form the preconditioner, we start with diagonal scaling, i.e., the magnitudes of all main diagonal entries of $\boldsymbol{M}$ are adjusted to one. After that, the incomplete matrix

TABLE I
CUBICAL CAVITY PROBLEM: PHASE ERRORS AND CONVERGENCE RATES OF THE ICCG METHOD IN CASE OF THE UNGAUGED A-V FORMULATION

frequency (MHz)	lateral length / wavelength	ICCG iterations (residual = 10^{-7}) TE_{10} mode	ICCG iterations (residual = 10^{-7}) TM_{33} mode	phase error of reflection coeff. (deg) TE_{10} mode	phase error of reflection coeff. (deg) TM_{33} mode
100	0.3336	63	54	evanescent	evanescent
200	0.6671	105	89	0.000	evanescent
300	1.0007	143	119	0.001	evanescent
400	1.3343	232	211	0.025	evanescent
500	1.6678	475	420	0.042	evanescent
600	2.0014	506	449	0.061	evanescent
700	2.3349	1127	1234	0.140	0.205
800	2.6685	958	987	0.293	0.323
900	3.0021	1882	2324	0.510	0.581
1000	3.3356	2297	2143	0.867	0.980
1100	3.6692	3150	3363	1.796	1.584
1200	4.0028	3656	3465	2.169	2.106

TABLE II
COMPARISON OF E AND UNGAUGED A-V FORMULATIONS FOR THE EXAMPLES GIVEN IN THIS PAPER

problem specification	final CG residual	fomulation	number of equations	number of non-zeros	preconditioner	CG iterations	CG iteration time (sec)	total CPU time (sec)	local error of E field*
example A: cavity resonator, irregular mesh, f = 800 MHz	10^{-4}	E	155 664	6 209 736	diag. scal.	5 012	42 256.2	43 398.3	$5.74*10^{-4}$
		ungauged A-V	155 664 + 55 255	13 500 555	diag. scal.	1 568	32 872.4	35 326.1	$5.69*10^{-4}$
					IC	601	24 330.5	26 779.1	$8.34*10^{-4}$
	10^{-7}	E	155 664	6 209 736	diag. scal.	12 269	103 335.7	104 475.3	$3.34*10^{-5}$
		ungauged A-V	155 664 + 55 255	13 500 555	diag. scal.	3 427	71 089.2	73 538.5	$4.44*10^{-7}$
					IC	958	38 703.8	41 144.4	0
example B: dielectric obstacle, course mesh, $k_0 a = 2.2$	10^{-4}	E	1120	34 244	diag. scal.	532	22.9	32.0	$1.23*10^{-4}$
		ungauged A-V	1120 + 346	71 100	diag. scal.	121	10.4	20.3	$3.46*10^{-5}$
					IC	20	3.73	17.7	$2.67*10^{-5}$
	10^{-7}	E	1120	34 244	diag. scal.	768	33.1	42.2	$1.73*10^{-6}$
		ungauged A-V	1120 + 346	71 100	diag. scal.	165	14.1	23.9	$1.40*10^{-8}$
					IC	27	5.01	21.01	0
example B: dielectric obstacle, fine mesh, $k_0 a = 2.2$	10^{-4}	E	21404	811 806	diag. scal.	1 762	2 122.0	2 219.0	$3.22*10^{-3}$
		ungauged A-V	21404 + 6909	1 754 505	diag. scal.	375	995.7	1 202.4	$3.38*10^{-4}$
					IC	57	294.9	598.4	$4.48*10^{-4}$
	10^{-7}	E	21404	811 806	diag. scal.	2 731	3 286.0	3 384.5	$5.64*10^{-6}$
		ungauged A-V	21404 + 6909	1 754 505	diag. scal.	490	1 301.1	1 507.88	$4.91*10^{-7}$
					IC	77	396.2	701.9	0
example C: hollow probe, h / l = 0.5	10^{-4}	E	20 906	744 419	diag. scal.	> 10 000	-	-	-
		ungauged A-V	20 906 + 6 608	1 607 695	diag. scal.	382	920.9	1 109.4	$2.45*10^{-4}$
					IC	75	348.76	625.3	$1.87*10^{-4}$
	10^{-7}	E	20 906	744 419	diag. scal.	> 10 000	-	-	-
		ungauged A-V	20 906 + 6 608	1 607 695	diag. scal.	659	1 584.0	1 772.2	$8.43*10^{-8}$
					IC	101	468.2	744.4	0

* See (40) for the definition.

factorization described in Section IV-B is performed. Finally, the matrix equations are solved iteratively by algorithms of the conjugate gradient type: In the complex case we employ a QMR-solver [18], while in the real case we use the standard implementation of the CG-method.

The method used to construct scalar and vector basis functions [1], [16] makes it possible to obtain the edge element approximation of $\boldsymbol{E}$ from the potentials in (14) without numerical differentiation. Because of (6), and since the unknowns correspond to the line integrals along the element edges [16], we simply obtain

$$e_{mn} = -j\omega(a_{mn} + v_n - v_m) \tag{39}$$

where e_{mn}, a_{mn} are the coefficients for $\boldsymbol{E}$ and $\boldsymbol{A}$ associated with edge $\{m \rightarrow n\}$ and v_m, v_n stand for the scalar potential in the starting and ending node of that edge.

V. NUMERICAL EXAMPLES

For convenience, we have summarized all computational data for the following examples in Table II. In the actual problem descriptions, these specifications are omitted.

A. Cubical Cavity

Lemma II in the Appendix shows that the number of negative eigenvalues of the matrix $\tilde{\boldsymbol{M}}_{\boldsymbol{AA}}$ increases with the frequency. Since, in our preconditioner, just a positive definite approximation to $\tilde{\boldsymbol{M}}_{\boldsymbol{AA}}$ is employed, the resulting factorization will become inaccurate whenever the wavelength λ becomes significantly smaller than the linear dimensions of the computational domain. To determine the practical limit, a simple cubical cavity has been investigated over a wide frequency range.

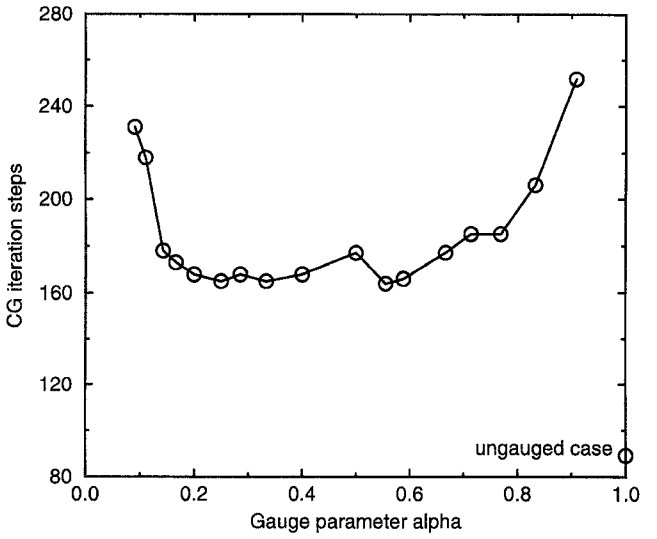


Fig. 2. A-V formulation: Number of CG iterations versus the gauge parameter α for the cubical cavity problem at $f = 200$ MHz. (CG residual norm $= 10^{-7}$, excitation: $TM_{3,3}$ mode.)

For convenience, we chose 1 m for the lateral length l, and $\epsilon_r = \mu_r = 1$, $\sigma = 0$. On one surface, the transverse components of a waveguide mode, either $TE_{1,0}$ or $TM_{3,3}$, were prescribed, while all other boundaries were assumed to be perfectly conducting. The CG method was terminated when the norm of the relative residual was less than 10^{-7}. To avoid atypically short solution times due to eigenvalues of high multiplicity, the cube was subdivided into 13 824 nonuniform finite elements. Table I shows the numerical convergence of the ICCG method and the phase error of the reflection coefficient in case of the ungauged formulation. It can be seen that the number of iteration steps grows significantly as the frequency is increased. However, up to $l \approx 2\lambda$, convergence rates are very satisfactory. Note that at the highest frequency, $f = 1.2$ GHz, the cavity is as large as $(4\lambda)^3$.

Second, the gauged scheme was tested. Fig. 2 illustrates the dependence of the convergence rate on the parameter α for $f = 200$ MHz. As predicted in Section IV-C, the number of iteration steps is markedly higher than in the ungauged case.

B. Dielectric Obstacle in Rectangular Waveguide

The configuration we considered is depicted in Fig. 3. In order to compare with previous results [19], we make use of the normalized frequency $k_0 a$, where 2a is the waveguide width and k_0 denotes the free space wavenumber. For simplicity, all waveguide walls were assumed to be perfectly conducting. First, the obstacle was lossless and its relative permittivity was set to $\epsilon_r = 6$. Second, a lossy dielectric with $\epsilon_r = 6 - j0.1$ was considered. The ungauged formulation was used in both cases. All computations were verified by our A-V nodal element software [20], but since the results are nearly the same as in the edge element case, no separate data are presented. Our solution for the magnitude of the reflection coefficient S_{11} is compared to the results from [19] in Fig. 4. The corresponding computational data given in Table II are for the course mesh.

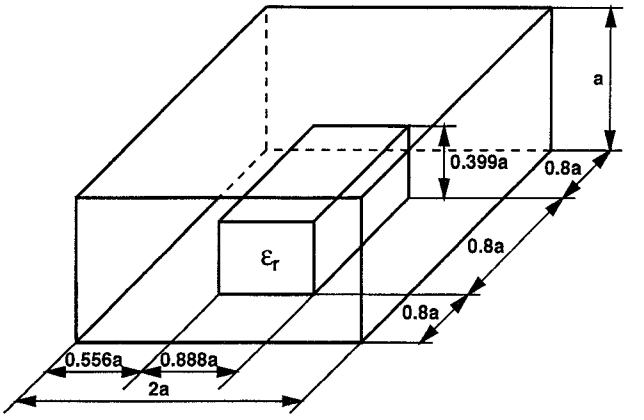


Fig. 3. Dielectric obstacle in a rectangular waveguide.

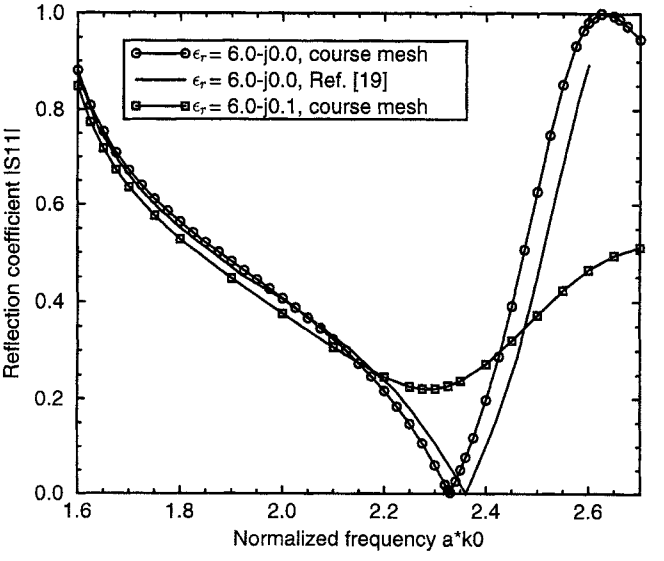


Fig. 4. Magnitude of the reflection coefficient $|S_{11}|$ of the dielectric obstacle versus the normalized frequency $k_0 a$.

C. Hollow Probe

A coaxially excited probe radiating into a rectangular waveguide has been analyzed. Fig. 5 illustrates the geometry of the considered device. At the operating frequency, $f = 1.6$ GHz, the waveguide was assumed to be perfectly matched at both ends. The walls of the hollow tube were assumed to be infinitely thin and all metallic parts perfectly conducting. Again, the ungauged method was applied. The maximum number of CG iterations for a relative residual norm of 10^{-7} was 215. As illustrated by Fig. 6, our results for the input admittance as a function of the probe height are in very good agreement with the solutions presented in [21].

D. Comparison with the E Formulation

In order to verify the efficiency of our method, we have recalculated the examples given above with both the A-V and the E method and different parameter settings. For all problems, we present solutions for CG residual norms of 10^{-4} and 10^{-7}. For the $A-V$ formulation, we applied both diagonal

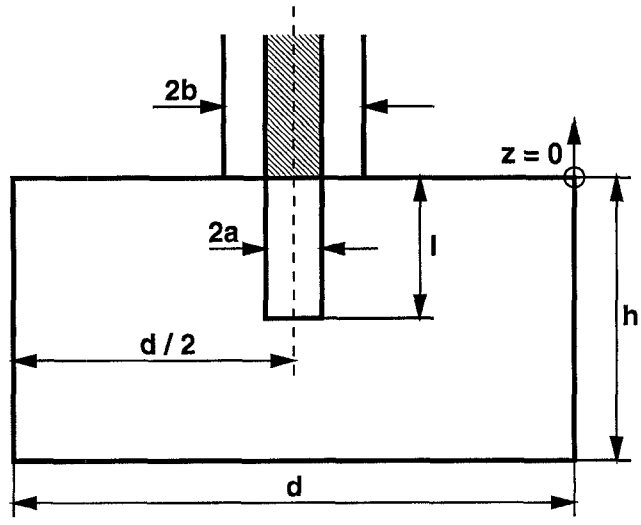


Fig. 5. Cross section of a coaxially fed hollow probe in a rectangular waveguide. Dimensions: $a = 3.1$ mm, $b = 7.13$ mm, $h = 57$ mm, and $d = 135$ mm.

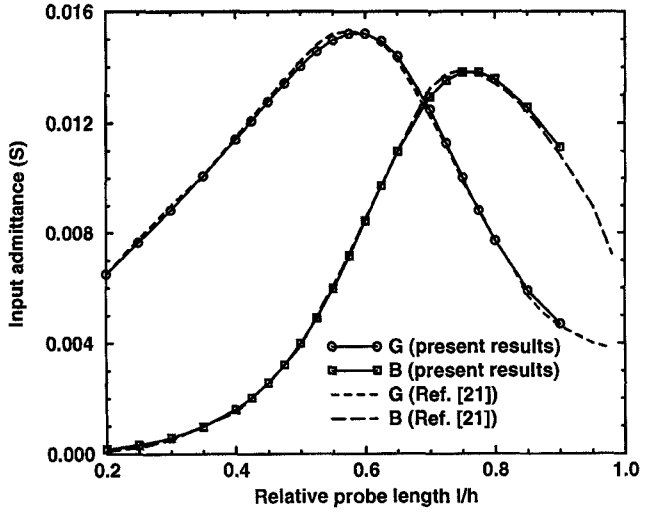


Fig. 6. Input admittance Y of the hollow probe as a function of the relative probe length l/h. Reference plane: $z = 0$.

scaling and incomplete Cholesky (IC) preconditioners. For the $\boldsymbol{E}$ scheme, however, IC never converged. To determine how well the theoretical equivalence of both formulations holds in practice, we compared the values of the electrical field strength along each finite element edge. As a measure for the largest difference observed, we used the maximum local error defined by

$$\epsilon := \max_{i=1\cdots N} \left| e_i - e_{i,AV-\text{scheme}}^{\text{residual}=10^{-7}} \right| \Big/ \max_{i=1\cdots N} \left| e_{i,AV-\text{scheme}}^{\text{residual}=10^{-7}} \right| \tag{40}$$

where e_i stands for the line integral of the electric field along edge I, which is the unknown in the $\boldsymbol{E}$ formulation [16]. Since the $\boldsymbol{E}$ method did not converge in one case, the $\boldsymbol{E}$ field solutions obtained from the $\boldsymbol{A}$-V formulation for a CG residual of 10^{-7} are used as the "true values" in (40). Table II presents a comparison of memory consumptions, CPU-times, iteration counts, and errors. It can be seen that both formulations yield almost equal results and that, especially for larger numbers of unknowns, the $\boldsymbol{A}$-V formulation is significantly faster.

VI. Conclusion

In this paper we have identified a major reason why conventional edge element formulations for driven high frequency problems are unsuitable for iterative solvers such as the preconditioned conjugate gradient method. It has been shown that significantly faster convergence rates can be obtained from an alternative, ungauged $\boldsymbol{A}$-V formulation employing edge elements for the vector potential and nodal shape functions for the scalar potential. An obvious disadvantage of the suggested method is that it results in approximately twice as many nonzero matrix entries as formulations based on the electric or magnetic field. Several numerical examples have been given to verify the suggested approach.

Appendix

Lemma I: Let the matrices $\boldsymbol{F}$ and $\boldsymbol{G}$ be defined as in Section III. Further, let $v_1^2 = \cdots = v_K^2 = 0$, $0 < v_{K+1}^2 \leq \cdots \leq v_{K+P}^2 < \omega^2$, and $\omega^2 < v_{K+P+1}^2 \leq \cdots \leq v_N^2$ be the eigenvalues of the generalized eigenvalue problem

$$(\boldsymbol{F} - v^2 \boldsymbol{G}) \cdot \boldsymbol{x} = 0. \tag{41}$$

Then, $K + P$ eigenvalues of $\boldsymbol{F} - \omega^2 \boldsymbol{G}$ are negative and $N - K - P$ eigenvalues are positive.

Proof: We denote by $\boldsymbol{x}_1 \cdots \boldsymbol{x}_N$ the eigenvectors of (41). For convenience, let them be normalized such that

$$\forall i = 1, \cdots, N; \quad \boldsymbol{x}_i^T \cdot \boldsymbol{G} \cdot \boldsymbol{x}_i = 1. \tag{42}$$

Then, with $\boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_N]$, we get from (41)

$$\boldsymbol{X}^T \cdot \boldsymbol{F} \cdot \boldsymbol{X} = \operatorname{diag}(v_i^2), \tag{43}$$

$$\boldsymbol{X}^T \cdot \boldsymbol{G} \cdot \boldsymbol{X} = \boldsymbol{I}. \tag{44}$$

For the eigenvector expansion of a vector $\boldsymbol{z} \in \Re^N$ we write

$$\boldsymbol{z} = \boldsymbol{X} \cdot \boldsymbol{s} \tag{45}$$

where $\boldsymbol{s} = \operatorname{col}[s_1, \cdots, s_N]$ is the vector of spectral coefficients. From (44), (45), it follows that

$$\boldsymbol{s}^T \cdot \boldsymbol{s} = \boldsymbol{z}^T \cdot \boldsymbol{G} \cdot \boldsymbol{z}. \tag{46}$$

Let $\boldsymbol{P} \in \Re^{N \times N}$ be a symmetric matrix and let $\pi_1 \leq \cdots \leq \pi_n$ be the eigenvalues of $\boldsymbol{P}$. Further, let Θ_i denote the set of all i-dimensional subspaces ϑ in $\Re^N$. Then, the eigenvalues of $\boldsymbol{P}$ are characterized by [11]

$$\begin{aligned} \pi_i &= \max_{\vartheta \in \Theta_{N-i+1}} \min_{\boldsymbol{z} \in \vartheta} \boldsymbol{z}^T \cdot \boldsymbol{P} \cdot \boldsymbol{z} / (\boldsymbol{z}^T \cdot \boldsymbol{z}) \\ &= \min_{\vartheta \in \Theta_i} \max_{\boldsymbol{z} \in \vartheta} \boldsymbol{z}^T \cdot \boldsymbol{P} \cdot \boldsymbol{z} / (\boldsymbol{z}^T \cdot \boldsymbol{z}). \end{aligned} \tag{47}$$

For arbitrary subspaces ϑ_i and ϑ_{N-i+1}, with $\dim \vartheta_i = i$ and $\dim \vartheta_{N-i+1} = N - i + 1$, we may therefore write

$$\min_{\boldsymbol{z} \in \vartheta_{N-i+1}} \boldsymbol{z}^T \cdot \boldsymbol{P} \cdot / (\boldsymbol{z}^T \cdot \boldsymbol{z}) \leq \pi_i \leq \max_{\boldsymbol{z} \in \vartheta_i} \boldsymbol{z}^T \cdot P \cdot \boldsymbol{z} / (\boldsymbol{z}^T \cdot \boldsymbol{z}). \tag{48}$$

In the following, we denote by $0 < \gamma_1^2 \leq \cdots \leq \gamma_N^2$ and $\kappa_1 \leq \cdots \leq \kappa_N$ the eigenvalues of $\boldsymbol{G}$ and $\boldsymbol{F} - \omega^2\boldsymbol{G}$, respectively. With that, we get from (47)

$$\forall \boldsymbol{z} \in \Re^N; \quad \gamma_1^2 \leq \boldsymbol{z}^T \cdot \boldsymbol{G} \cdot \boldsymbol{z}/(\boldsymbol{z}^T \cdot \boldsymbol{z}) \leq \gamma_N^2. \tag{49}$$

To complete the proof, we split $\Re^N$ into two linearly independent subspaces:

Case 1: Let $\vartheta_{K+P} = \text{span}\,(\boldsymbol{x}_1 \cdots \boldsymbol{x}_K \cdots \boldsymbol{x}_{K+P})$, with $v_1^2 = \cdots = v_K^2 = 0$ and $0 < v_{K+1}^2 \leq \cdots \leq v_{K+P}^2 < \omega^2$. Then, (48) and (43)–(45) lead to

$$\begin{aligned}\kappa_{K+P} &\leq \max_{\boldsymbol{z} \in \vartheta_{K+P}} \boldsymbol{z}^T \cdot (\boldsymbol{F} - \omega^2\boldsymbol{G}) \cdot \boldsymbol{z}/(\boldsymbol{z}^T \cdot \boldsymbol{z}) \\ &= \max_{\boldsymbol{z} \in \vartheta_{K+P}} \boldsymbol{s}^T \cdot \text{diag}\,(v_i^2 - \omega^2) \cdot \boldsymbol{s}/(\boldsymbol{z}^T \cdot \boldsymbol{z}). \end{aligned} \tag{50}$$

Since, for any $\boldsymbol{z} \in \vartheta_{K+P}$, the corresponding vector $\boldsymbol{s}$ is of the form $s = \text{col}\,[\boldsymbol{s}_1, \cdots \boldsymbol{s}_{K+P}; \boldsymbol{0}, \cdots, \boldsymbol{0}]$, we have

$$\begin{aligned}&\max_{\boldsymbol{z} \in \vartheta_{K+P}} \boldsymbol{s}^T \cdot \text{diag}\,(v_i^2 - \omega^2) \cdot \boldsymbol{s}/(\boldsymbol{z}^T \cdot \boldsymbol{z}) \\ &\qquad \leq (v_{k+P}^2 - \omega^2) \min_{\boldsymbol{z} \in \vartheta_{K+P}} \boldsymbol{s}^T \cdot \boldsymbol{s}/(\boldsymbol{z}^T \cdot \boldsymbol{z}) < 0. \end{aligned} \tag{51}$$

By means of (46) and (49)–(51), we arrive at

$$\begin{aligned}\kappa_{K+P} &\leq (v_{K+P}^2 - \omega^2) \min_{\boldsymbol{z} \in \vartheta_{K+P}} \boldsymbol{s}^T \cdot \boldsymbol{s}/(\boldsymbol{z}^T \cdot \boldsymbol{z}) \\ &= (v_{K+P}^2 - \omega^2) \min_{\boldsymbol{z} \in \vartheta_{K+P}} \boldsymbol{z}^T \cdot \boldsymbol{G} \cdot \boldsymbol{z}/(\boldsymbol{z}^T \cdot \boldsymbol{z}) \\ &\leq (v_{K+P}^2 - \omega^2)\gamma_1^2 < 0. \end{aligned} \tag{52}$$

Hence, $K+P$ eigenvalues of $\boldsymbol{F} - \omega^2\boldsymbol{G}$ are negative. From the special case $P = 0$, i.e., $\omega^2 < v_{K+1}^2$, it can be seen that, at any nonzero frequency, at least K eigenvalues are negative.

Case 2: Let $\vartheta_{N-K-P+1} = \text{span}\,(\boldsymbol{x}_{K+P+1} \cdots \boldsymbol{x}_N)$, with $(\omega^2 < v_{K+P+1}^2 \leq \cdots \leq v_N^2$. Now, $\boldsymbol{s} = \text{col}\,[0, \cdots 0; s_{K+P+1}, \cdots s_N]$ and, by analogy to *Case 1*, (48), (49), (43)–(46) yield

$$\begin{aligned}\kappa_{K+P+1} &\geq \min_{\boldsymbol{z} \in \vartheta_{N-K-P+1}} \boldsymbol{z}^T \cdot (\boldsymbol{F} - \omega^2\boldsymbol{G}) \cdot \boldsymbol{z}/(\boldsymbol{z}^T \cdot \boldsymbol{z}) \\ &\geq (v_{K+P+1}^2 - \omega^2) \min_{\boldsymbol{z} \in \vartheta_{N-K-P+1}} \boldsymbol{s}^T \cdot \boldsymbol{s}/(\boldsymbol{z}^T \cdot \boldsymbol{z}) \\ &\geq (v_{K+P+1}^2 - \omega^2)\gamma_1^2 > 0 \end{aligned} \tag{53}$$

So the remaining $N - K - P$ eigenvalues are all positive. □

Lemma II: Let the matrix $\tilde{\boldsymbol{M}}_{\boldsymbol{AA}}$ be defined by (30). Further, let $v_1^2 = \cdots = v_K^2 = 0$, $0 < v_{K+1}^2 \leq \cdots \leq v_{K+P}^2 < \omega^2$, and $\omega^2 < v_{K+P+1}^2 \leq \cdots \leq v_N^2$ be the eigenvalues of (41). Then, P eigenvalues of $\tilde{\boldsymbol{M}}_{\boldsymbol{AA}}$ are negative and $N - K - P$ eigenvalues are positive. For $\alpha^2 = 1$, the remaining K eigenvalues are zero and, for $0 < \alpha^2 < 1$, they are positive.

Proof: In the following, we make use of the matrices $\boldsymbol{F}$, $\boldsymbol{G}$, $\boldsymbol{M}$, $\boldsymbol{M}_{VV}$, and $\boldsymbol{M}_{VA}$ defined in Sections III and IV. The eigenvectors of (41), $\boldsymbol{x}_1 \cdots \boldsymbol{x}_N$, are assumed to be normalized by (42).

Let us first consider the ungauged case $\alpha^2 = 1$. Then, the nullspace of $\boldsymbol{M}$ is characterized by

$$-\boldsymbol{M}_{VV} \cdot \boldsymbol{v}_0 - \boldsymbol{M}_{VA} \cdot \boldsymbol{a}_0 = 0, \tag{54}$$

$$-\boldsymbol{M}_{VA}^T \cdot \boldsymbol{v}_0 + (\boldsymbol{F} - \omega^2\boldsymbol{G}) \cdot \boldsymbol{a}_0 = 0. \tag{55}$$

From its physical interpretation (see Section IV), we know that

$$\text{span}\,(\boldsymbol{a}_0) = \ker \boldsymbol{F}. \tag{56}$$

Hence,

$$\boldsymbol{F} \cdot \boldsymbol{a}_0 = 0 \tag{57}$$

and (55) reduces to

$$\boldsymbol{M}_{VA}^T \cdot \boldsymbol{v}_0 = -\omega^2\boldsymbol{G} \cdot \boldsymbol{a}_0. \tag{58}$$

Combining (54) and (58), we obtain

$$\boldsymbol{M}_{VA}^T \cdot \boldsymbol{M}_{VV}^{-1} \cdot \boldsymbol{M}_{VA} \cdot \boldsymbol{a}_0 = \omega^2\boldsymbol{G} \cdot \boldsymbol{a}_0. \tag{59}$$

Since

$$\dim \text{span}\,(\boldsymbol{a}_0) = \dim \text{span}\,(v) = K \tag{60}$$

we get from (54)

$$\forall \boldsymbol{v} \in \Re^K; \quad \boldsymbol{v} = -\boldsymbol{M}_{VV}^{-1} \cdot \boldsymbol{M}_{VA} \cdot \boldsymbol{a}_0. \tag{61}$$

In other words, any vector $\boldsymbol{v} \in \Re^K$ can be expressed in terms of $\boldsymbol{a}_0 \in \ker \boldsymbol{F}$. Thus, we may even write

$$\begin{aligned}&\forall \boldsymbol{a} \in \Re^N \; \exists \boldsymbol{a}_0 \in \ker \boldsymbol{F}: \\ &\boldsymbol{v}(\boldsymbol{a}) = -\boldsymbol{M}_{VV}^{-1} \cdot \boldsymbol{M}_{VA} \cdot \boldsymbol{a} = \boldsymbol{v}(\boldsymbol{a}_0) = -\boldsymbol{M}_{VV}^{-1} \cdot \boldsymbol{M}_{VA} \cdot \boldsymbol{a}_0. \end{aligned} \tag{62}$$

By means of (59), we arrive at

$$\begin{aligned}&\forall \boldsymbol{a} \in \Re^N \; \exists \boldsymbol{a}_0 \in \ker \boldsymbol{F}: \\ &\quad -\boldsymbol{M}_{VA}^T \cdot \boldsymbol{v}(\boldsymbol{a}) = \boldsymbol{M}_{VA}^T \cdot \boldsymbol{M}_{VV}^{-1} \cdot \boldsymbol{M}_{VA} \cdot \boldsymbol{a} \\ &\qquad\qquad = \boldsymbol{M}_{VA}^T \cdot \boldsymbol{M}_{VV}^{-1} \cdot \boldsymbol{M}_{VA} \cdot \boldsymbol{a}_0 \\ &\qquad\qquad = \omega^2\boldsymbol{G} \cdot \boldsymbol{a}_0. \end{aligned} \tag{63}$$

Next, let us investigate the expression

$$\boldsymbol{a}^T \cdot \tilde{\boldsymbol{M}}_{AA} \cdot \boldsymbol{a}. \tag{64}$$

A) Let $\boldsymbol{a}_0 \in \text{span}\,(\boldsymbol{x}_1 \cdots \boldsymbol{x}_K) = \ker \boldsymbol{F}$: With (30), (57), and (63), we get

$$\begin{aligned}&\boldsymbol{a}_0^T \cdot \tilde{\boldsymbol{M}}_{AA} \cdot \boldsymbol{a}_0 \\ &\quad = \boldsymbol{a}_0^T \cdot (-\omega^2\boldsymbol{G} + 1/\alpha^2 \boldsymbol{M}_{VA}^T \cdot \boldsymbol{M}_{VV}^{-1} \cdot \boldsymbol{M}_{VA}) \cdot \boldsymbol{a}_0 \\ &\quad = \omega^2(1 - \alpha^2)/\alpha^2 \boldsymbol{a}_0^T \cdot \boldsymbol{G} \cdot \boldsymbol{a}_0. \end{aligned} \tag{65}$$

B) Let $\boldsymbol{a}_\perp \in \text{span}\,(\boldsymbol{x}_{K+1} \cdots \boldsymbol{x}_N)$: Due to (44), we have

$$\forall \boldsymbol{a}_0 \in \ker\,(\boldsymbol{F}); \; \boldsymbol{a}_\perp^T \cdot \boldsymbol{G} \cdot \boldsymbol{a}_0 = 0. \tag{66}$$

Hence, (63) leads to

$$\begin{aligned}\boldsymbol{a}_\perp^T \cdot \tilde{\boldsymbol{M}}_{AA} \cdot \boldsymbol{a}_\perp &= \boldsymbol{a}_\perp^T \cdot (\boldsymbol{F} - \omega^2\boldsymbol{G}) \cdot \boldsymbol{a}_\perp + \omega^2/\alpha^2 \boldsymbol{a}_\perp^T \cdot \boldsymbol{G} \cdot \boldsymbol{a}_0 \\ &= \boldsymbol{a}_\perp^T \cdot (\boldsymbol{F} - \omega^2\boldsymbol{G}) \cdot \boldsymbol{a}_\perp. \end{aligned} \tag{67}$$

The rest is very similar to the proof of Lemma I, so we give a short outline only. In the following, $0 < \gamma_1^2 \leq \cdots \leq \gamma_N^2$ and $\mu_1 \leq \cdots \leq \mu_N$ denote the eigenvalues of $\boldsymbol{G}$ and $\tilde{\boldsymbol{M}}_{\boldsymbol{AA}}(\alpha)$, respectively.

Case 1: Let $\vartheta_P = \text{span}\,(\boldsymbol{x}_K + 1 \cdots \boldsymbol{x}_{K+P})$, with $0 < v_{K+1}^2 \leq \cdots \leq v_{K+P}^2 < \omega^2$. Then, (48), (49) and (67) lead to

$$\begin{aligned}\mu_P &\leq \max_{\boldsymbol{a}\in\vartheta_P} \boldsymbol{a}^T \cdot \tilde{\boldsymbol{M}}_{AA} \cdot \boldsymbol{a}/(\boldsymbol{a}^T \cdot \boldsymbol{a}) \\ &= \max_{\boldsymbol{a}\in\vartheta_P} \boldsymbol{a}^T \cdot (\boldsymbol{F} - \omega^2 \boldsymbol{G}) \cdot \boldsymbol{a}/(\boldsymbol{a}^T \cdot \boldsymbol{a}) \\ &\leq (v_{K+P}^2 - \omega^2)\gamma_1^2 < 0. \end{aligned} \tag{68}$$

Hence P eigenvalues are negative.

Case 2: Let $\vartheta_{N-P} = \text{span}\,(\boldsymbol{x}_1 \cdots \boldsymbol{x}_K;\ \boldsymbol{x}_{K+P+1} \cdots \boldsymbol{x}_N)$, with $v_1^2 = \cdots = v_K^2 = 0$ and $\omega^2 < v_{K+P+1}^2 \leq \cdots \leq v_{K+N}^2$. Due to (67), we have

$$\begin{aligned}&\forall\, \boldsymbol{z} \in \text{span}\,(\boldsymbol{x}_{K+P+1} \cdots \boldsymbol{x}_N); \\ &\quad \min \boldsymbol{a}^T \cdot \tilde{\boldsymbol{M}}_{AA} \cdot \boldsymbol{a}/(\boldsymbol{a}^T \cdot \boldsymbol{a}) \geq \gamma_1^2 (v_{K+P+1}^2 - \omega^2) > 0 \end{aligned} \tag{69}$$

and, by means of (65), we get

$$\begin{aligned}&\forall\, \boldsymbol{z} \in \text{span}\,(\boldsymbol{x}_1 \cdots \boldsymbol{x}_K); \\ &\quad \min \boldsymbol{a}^T \cdot \tilde{\boldsymbol{M}}_{AA} \cdot \boldsymbol{a}/(\boldsymbol{a}^T \cdot \boldsymbol{a}) \geq \gamma_1^2 \omega^2 (1 - \alpha^2)/\alpha^2 \geq 0. \end{aligned} \tag{70}$$

Hence

$$\begin{aligned}\mu_{P+1} &\geq \min_{\boldsymbol{a}\in\vartheta_{N-P}} \boldsymbol{a}^T \cdot \tilde{\boldsymbol{M}}_{AA} \cdot \boldsymbol{a}/(\boldsymbol{a}^T \cdot \boldsymbol{a}) \\ &\geq \min \left\{ \begin{matrix} \gamma_1^2 (v_{K+P+1}^2 - \omega^2) > 0 \\ \gamma_1^2 \omega^2 (1 - \alpha^2)/\alpha^2 \geq 0 \end{matrix} \right\} \geq 0. \end{aligned} \tag{71}$$

For $0 < \alpha^2 < 1$, we have $\mu_{P+1} > 0$. If $\alpha^2 = 1$, then $\mu_{P+1} = \cdots \mu_{P+K} = 0$ and $\mu_{P+K+1} \geq \gamma_1^2 (v_{K+P+1}^2 - \omega^2) > 0$. □

REFERENCES

[1] A. Bossavit, "A rationale for "edge elements" in 3-D fields computations," *IEEE Trans. Magn.*, vol. 24, pp. 74–79, Jan. 1988.

[2] ——, "Solving Maxwell equations in a closed cavity, and the question of 'spurious modes,'" *IEEE Trans. Magn.*, vol. 26, pp. 702–705, Mar. 1990.

[3] M. Hano, "Finite-element solution of three-dimensional resonator problems—Novel rectangular parallelepiped elements," *Trans. IEICE*, vol. J70-C, pp. 791–796, 1987.

[4] J. F. Lee, D. K. Sun, and Z. J. Cendes, "Full-Wave Analysis of Dielectric Waveguides Using Tangential Vector Finite Elements," *IEEE Trans. Microwave Theory Tech.*, vol. 39, 1262–1271, Aug. 1991.

[5] I. Bardi, O. Biro, K. Preis, G. Vrisk, and K. R. Richter, "Nodal and edge element analysis of inhomogeneously loaded 3-D cavities," *IEEE Trans. Magn.*, vol. 28, pp. 1142–1145, Mar. 1992.

[6] 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.

[7] S. L. Foo and P. P. Silvester, "Boundary-marching method for discontinuity analysis in waveguides of arbitrary cross section," *IEEE Trans. Microwave Theory Tech.*, vol. 40, pp. 1889–1893, Oct. 1992.

[8] Z. J. Cendes and J. F. Lee, "The transfinite Element Method for Modeling MMIC Devices," *IEEE Trans. Microwave Theory Tech.*, vol. 36, pp. 1639–1649, Dec. 1988.

[9] R. Albanese and A. Rubinacci, "Solution of three dimensional eddy current problems by integral and differential methods," *IEEE Trans. Magn.*, vol. 24, pp. 98–101, Jan. 1988.

[10] H. Kanayama, H. Motoyama, K. Endo, and F. Kikuchi, "Three-dimensional magnetostatic analysis using Nedelec's elements," *IEEE Trans. Magn.*, vol. 26, pp. 682–685, Mar. 1990.

[11] J. H. Wilkinson, *The Algebraic Eigenvalue Problem.* Oxford: Clarendon Press, 1965.

[12] C. C. Paige and M. A. Saunders, "Solution of sparse indefinite systems of linear equations," *SIAM J. Num. Anal.*, vol. 12, pp. 617–629, Sept. 1975.

[13] J. R. Bunch and B. N. Parlett, "Direct methods for solving symmetric indefinite systems of linear equations," *SIAM J. Numer. Anal.*, vol. 8, pp. 639–655, Dec. 1971.

[14] J. K. Reid, "On the method of conjugate gradients for the solution of large sparse systems of equations," in *Large Sparse Sets of Linear Equations*, J. K. Reid, Ed. London: Academic Press, 1971.

[15] A. Kameari, "Study on 3-D Eddy Current Analysis Using FEM," in *Proc. 4th International IGTE Symp. and European TEAM Workshop*, K. R. Richter, W. M. Rucker, O. Biro, Eds. Graz, Oct. 1990, pp. 91–99.

[16] A. Kameari, "Calculation of transient 3-D eddy current using edge-elements," *IEEE Trans. Magn.*, vol. 26, pp. 466–469, Mar. 1990.

[17] A. George and J. W. H. Liu, "The evolution of the minimum degree ordering algorithm," *SIAM Review*, vol. 31, pp. 1–19, Mar. 1989.

[18] R. W. Freund, "Conjugate gradient-type methods for linear systems with complex symmetric coefficient matrices," *SIAM J. Sci. Stat. Comput.*, vol. 13, pp. 425–448, Jan. 1992.

[19] K. Ise, K. Inoue, and M. Koshiba, "Three-dimensional finite-element solution of dielectric scattering obstacles in a rectangular waveguide," *IEEE Trans. Microwave Theory Tech.*, vol. 38, pp. 1352–1359, Sept. 1990.

[20] R. Dyczij-Edlinger, I. Bardi, O. Biro, K. Preis, and K. R. Richter, "A deterministic approach to the analysis of three-dimensional waveguide configurations by finite elements and mode matching," *IEEE Trans. Magn.*, vol. 28, pp. 1235–1238, Mar. 1992.

[21] A. G. Williamson, "Coaxially fed hollow probe in a rectangular waveguide," *Proc. Inst. Elec. Eng.*, vol. 132, pt. H, no. 5, pp. 273–285, Aug. 1985.

Romanus Dyczij-Edlinger, photograph and biography not available at the time of publication.

Oszkar Biro, photograph and biography not available at the time of publication.