Convergence of Preconditioned Conjugate Gradient Method Applied to Driven Microwave Problems

H. Igarashi, *Member, IEEE*, and T. Honma, *Member, IEEE*

***Abstract*—Driven microwave problems can be solved with the finite-element method formulated in terms of the electric field, as well as the vector and scalar potentials. It is known that the latter gives faster convergence of the preconditioned conjugate gradient method than the former. This can be understood from the following facts: namely, the preconditioned finite-element matrix of the former method can contain small negative eigenvalues which make the matrix condition worse. On the other hand, in the latter, such eigenvalues are shown to be composed of zeros and normalized ones.**

***Index Terms*—Condition number, conjugate gradient (CG) method, electromagnetic waves, finite-element method, preconditioning.**

I. INTRODUCTION

IN FINITE-ELEMENT analysis of electromagnetic fields, the preconditioned conjugate gradient (CG) method is widely used for the solution of a linear system. Since such a solution process spends a large fraction of computing time, it is of importance to improve convergence of the CG method.

In magnetostatic problems, deterioration of the CG convergence due to the tree-cotree gauging is caused by the decrease in the minimum nonzero eigenvalue which leads to worse matrix conditions [1]. In the quasi-static electromagnetic problems, superiority in the convergence of the A–V method over the A method is due to the fact that the redundant DoFs in the former method help the preconditioner eliminate small singular values, which remain after the preconditioning in the latter method [2].

The superiority of the A–V method over A method, or equivalently the E method, has also been reported in driven microwave problems [3]. In quasi-static problems, a linear system with a complex symmetric matrix is solved with the complex CG method with preconditioning by the incomplete Cholesky (IC) factorization , which is equivalent to the BiCG. In addition, the preconditioned CG method is often applied to indefinite real symmetric matrix in high-frequency problems. Although the validity of the CG method is guaranteed only for a positively definite matrix, it can also solve equations with a indefinite matrix unless there is breakdown during the CG iteration process [4]. (The CG process can be restarted from different initial values even if there is breakdown which rarely occurs.) Although there are other linear solvers such as MINRES and SYMMLQ, which are available for indefinite matrices, they will have to be tested for various points such as memory requirement and robustness for practical uses.

Manuscript received June 18, 2002.

The authors are with the Graduate School of Engineering, Hokkaido University, Kita-ku 060-8628, Japan (e-mail: iga@em-si.eng.hokudai.ac.jp).

Digital Object Identifier 10.1109/TMAG.2003.810168

In this paper, the reason for the superiority of A–V method in the driven microwave problems is discussed through spectrum analysis of the finite-element matrix.

II. FORMULATION

We consider here the electromagnetic waves in vacuum for simplicity. The Galerkin formulation for high-frequency electromagnetic fields in terms of electric field $\boldsymbol{E}$ is given by

$$\int_\Omega \nabla\times\boldsymbol{w}_e\cdot\nabla\times\boldsymbol{E}\,dv - k^2\int_\Omega \boldsymbol{w}_e\cdot\boldsymbol{E}\,dv = -j\omega\mu_0\int_S \boldsymbol{w}_e\cdot(\boldsymbol{H}\times\boldsymbol{n})dS \quad (1)$$

where Ω denotes the vacuum region with a surface S and k the wavenumber. The region Ω is discretized into finite elements with n nodes, e edges, and f faces. Moreover the corresponding basis functions are denoted by w_n, $\boldsymbol{w}_e$, and $\boldsymbol{w}_f$. (See, for example, [5] for their detailed definition.) The finite-element discretization of (1) leads to

$$[K]\{A\} = \{b\} \quad (2)$$

where

$$[K] = [C]^t\,[M_f]\,[C] - k^2\,[M_e] \quad (3)$$

and $[C]$ is the discrete counterparts of curl operator which has $f\times e$ entities. Moreover, the entities of the symmetric matrices $[M_f]$ and $[M_e]$ are given by $\int_\Omega \boldsymbol{w}_f\cdot\boldsymbol{w}_{f'}dv$ and $\int_\Omega \boldsymbol{w}_e\cdot\boldsymbol{w}_{e'}dv$, respectively. In addition, rank$([C]^t[M_f][C]) = e - n + 1$.

When electric field is represented in terms of the vector and scalar potentials, i.e., $\boldsymbol{E} = -j\omega(\boldsymbol{A}+\nabla V)$, (1) becomes

$$\int_\Omega \nabla\times\boldsymbol{w}_e\cdot\nabla\times\boldsymbol{A}\,dv - k^2\int_\Omega \boldsymbol{w}_e\cdot(\boldsymbol{A}+\nabla V)\,dv = \mu_0\int_S \boldsymbol{w}_e\cdot(\boldsymbol{H}\times\boldsymbol{n})dS. \quad (4)$$

Moreover, to make the equations closed, we additionally consider the divergence-free condition of $\boldsymbol{E}$, that is

$$-k^2\int_\Omega \nabla w_n\cdot\boldsymbol{A}\,dv - k^2\int_\Omega \nabla w_n\cdot\nabla V\,dv = 0. \quad (5)$$

The finite-element discretization of (4) and (5) yields the linear system

$$[K']\begin{Bmatrix} A \\ V \end{Bmatrix} = \begin{Bmatrix} b' \\ 0 \end{Bmatrix} \quad (6)$$

where

$$[K'] = \begin{bmatrix} [C]^t\,[M_f]\,[C] - k^2\,[M_e] & -k^2\,[M_e]\,[G] \\ -k^2[G]^t\,[M_e] & -k^2[G]^t\,[M_e]\,[G] \end{bmatrix} \quad (7)$$

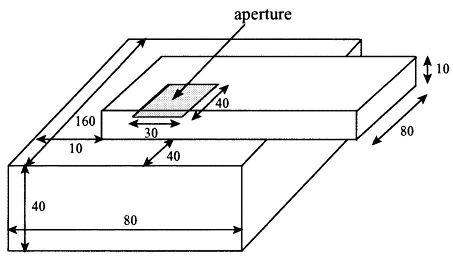

Fig. 1. Waveguide-cavity system (unit in millimeters).

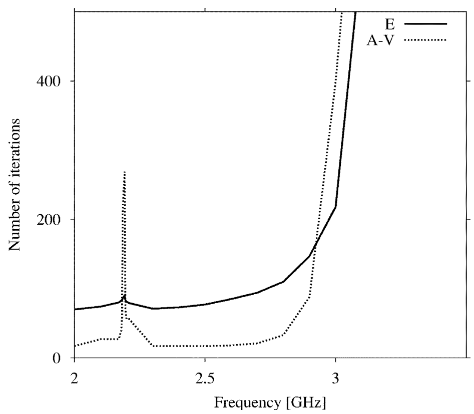

Fig. 2. Convergence of ICCG (model A1).

where $[G]$ represents the counterpart of grad which has $e \times n$ entities, and rank $([G]^t[M_e][G]) = n - 1$. Since $[G]^t[C]^t = 0$, the lower part of $[K']$ is linearly dependent on the upper part.

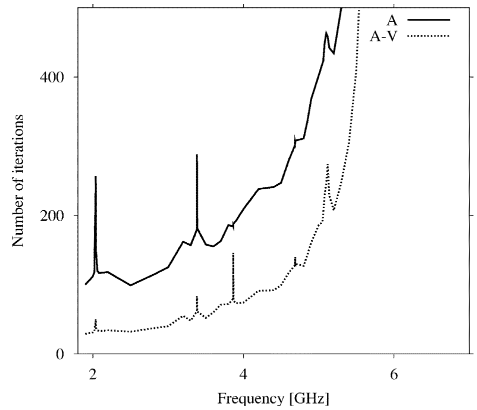

Fig. 3. Convergence of preconditioned CG (model A2).

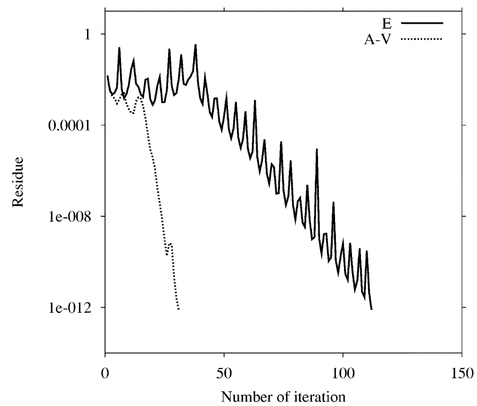

Fig. 4. Convergent process of preconditioned CG (model A2).

III. Numerical Experiment

A. Convergence of Preconditioned CG

For the first numerical experiment, the waveguide-cavity system shown in Fig. 1 is analyzed using the brick-edge elements. Electromagnetic waves are excited near the end wall of the waveguide by attributing unit value to one of the edge DoFs (i.e., hard source). There is an aperture between the waveguide and cavity, through which the electromagnetic waves are transported to the cavity.

Two numerical models are considered: model A1 contains 176 and 640 valid nodes and edges, respectively, while model A2 has 1872 and 7968. The preconditioning is performed by the IC, in which the acceleration factor γ is multiplied to the diagonal components of the finite-element matrix to make it diagonal dominant [6]. The value of γ is taken as 1.1.

Figs. 2 and 3 show the dependence of convergence of the preconditioned CG method on frequency for each model. The A–V method has overall better convergence than the E method. We see some peaks in both results, which correspond to the resonant frequencies of the cavity. Moreover, the number of iterations grows as frequency increases. When the wavelength becomes smaller than about ten times of the element size, preconditioned CG cannot provide convergent solutions.

A convergence process for model A2 is shown in Fig. 4. There are a lot of spikes, especially in the E method. These spikes are due to indefinite property of the finite-element matrix. That is, since $\{p\}^t[K]^t\{p\}$, where $\{p\}$ denotes the search vector in the CG method, is not guaranteed to be positive unlike for the positive definite matrices, nearly zero division can occur.

Next, TEM waves propagating along a parallel-plate waveguide with 50 m length and 1×1 m^2 cross section are analyzed. The waveguide is terminated by a metallic wall, whereas a TEM wave is excited at the opposite wall. This model, referred to as model B hereafter, has neither resonance nor cut-off frequency unlike model A. The model is discretized with brick-edge elements which contain valid 542 edges and 147 nodes. The convergence of model B is shown in Fig. 5, where the A–V method has again better convergence than the E method. The superiority in the convergence of A–V method over E method, which has also been reported for different numerical models [3], is discussed from a view point of the spectrum of the finite-element matrix.

B. Spectra of Diagonally Scaled Matrix

The spectra of the finite-element matrix for model A1 is considered here, whose essential property is expected to be the same as that of models A2 and B. Figs. 6 and 7 depict the absolute eigenvalues of the finite-element matrix for the E and A-V methods which has been diagonally scaled such that $\tilde{K}_{ij} = K_{ij}/\sqrt{|K_{ii}K_{jj}|}$. Note that the diagonal scaling can be regarded as the simplest preconditioning. In the E method, there are $n-1$ small clustered negative eigenvalues provided that the driven frequency is lower than the first resonance. Those small negative eigenvalues are, as can be understood from (3), relevant to

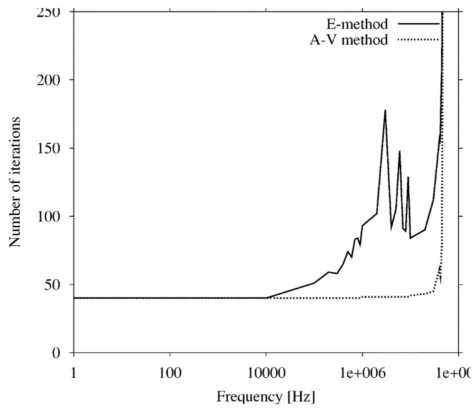

Fig. 5. Convergence of preconditioned CG (model B).

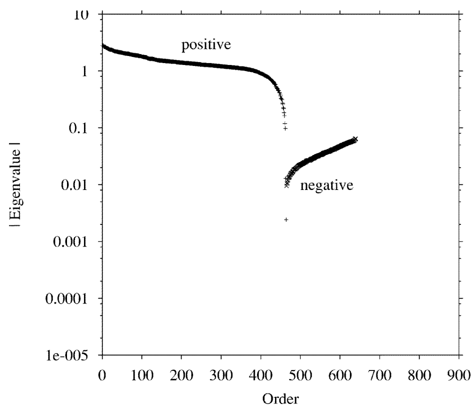

Fig. 6. Eigenvalue for E method (model A1, 2 GHz).

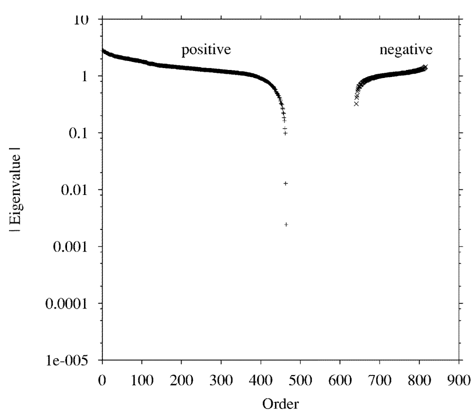

Fig. 7. Eigenvalue for A–V method (model A1, 2 GHz).

the zero eigenvalues of $[C]^t[M_f][C]$, which are shifted to negative side when frequency is increased. On the other hand, in the A–V method as shown in Fig. 7, there are the negative eigenvalues whose value is nearly minus one as well as zero eigenvalues between the negative and postive ones.

The reason for the essentially different structures in the spectra of the E and A–V method can be understood as follows. Let us consider the matrix $[\tilde{K}']$ which is obtained by diagonally scaling $[K']$. Then, it can be shown that

$$\left[\tilde{K}'\right] \simeq \begin{bmatrix} [I'] & [0] \\ [0]^t & -[I''] \end{bmatrix} \tag{8}$$

when k is sufficiently small, where $[I']$ and $[I'']$ denote $e \times e$ and $n \times n$ diagonal–dominant matrices whose $e-n+1$ and $n-1$ eigenvalues cluster near unit. Note that $[\tilde{K}']$ has strictly n zero eigenvalues because rank does not change through the scaling. Since $[\tilde{K}']$ has these $e+n$ eigenvalues in total, the small negative eigenvalues cannot exist. On the other hand, $[\tilde{K}]$ corresponding to $[I']$ which is regular for nonzero k can have small negative eigenvalues instead of zeros. That is, when k is zero, $[\tilde{K}]$ has $n-1$ zero eigenvalues, which are shifted to small negative ones as k increases. Such small eigenvalues do not exist.

One would see the property more clearly by considering a simple matrix such as

$$[k'] = \begin{bmatrix} 1-\epsilon & 2 & 2\epsilon \\ 2 & 4-\epsilon & -\epsilon \\ 2\epsilon & -\epsilon & -5\epsilon \end{bmatrix} \tag{9}$$

where $[k']$ is expected to have properties similar to $[K']$ except that the upper-left 2×2 matrix is not diagonal dominant. The parameter ϵ plays the role of k^2. The scaled matrix $[\tilde{k}']$, which is of the form of (8) for small ϵ, has the eigenvalues, for instance, 2.007, 0, and -1.007 when $\epsilon = 0.01$, whereas the upper-left 2×2 matrix of $[\tilde{k}']$, which corresponds to $[\tilde{K}]$, has 2.006 and -6.296×10^{-3} that corresponds to the small negative eigenvalue deteriorating the matrix condition.

Besides the previous important property, we can additionally observe that Figs. 6 and 7 have the nearly zero eigenvalues isolated from others. They are relevant to resonance, that is, such eigenvalues shift from positive to negative when frequency passes a resonant frequency.

The condition of an indefinite matrix can be characterized by the condition number defined as [7]

$$\kappa = \frac{\sigma_{\max}}{\sigma_{\min}^0} \tag{10}$$

where $\sigma_{\max}$ and $\sigma_{\min}^0$ are maximum and minimum nonzero singular values of the finite-element matrix (i.e., the square root of the eigenvalues of, e.g., $[K]^t[K]$).

Since a singular value becomes zero at resonance, κ becomes worse when frequency approaches one of the resonant frequencies. This is the reason why there appear peaks in the iteration number in Figs. 2 and 3.

C. Spectra of Matrix Preconditioned By IC

Figs. 8 and 9 show the singular values of the finite-element matrix preconditioned by the IC. They correspond well to the spectra in Figs. 6 and 7. The singular values of the A–V method cluster near unit while those of the E method spread down to smaller than 0.1. Since the CG method has better convergence for clustered singular values, these spectra are consistent with the observed convergence. (Note that small κ is not a sufficient but a necessary condition for good convergence.)

Although the reason for the difference in the spectra has been discussed, let us now consider the reason from a different point of view. It is known that the eigenvalues λ_i of a symmetric matrix, say $[A]$, $n \times n$, is separated by the eigenvalues μ_i of the matrix which is obtained by eliminating one column and row from $[A]$ [1], [8], that is

$$\lambda_1 \le \mu_1 \le \lambda_2 \le \mu_2 \le \cdots \le \mu_{n-1} \le \lambda_n. \tag{11}$$

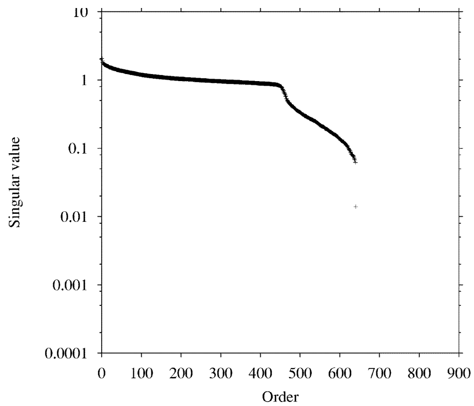

Fig. 8. Singular value for E method (model A1, 2 GHz).

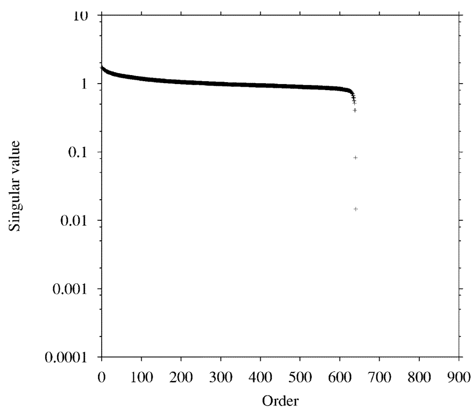

Fig. 9. Singular value for A–V method (model A1, 2 GHz).

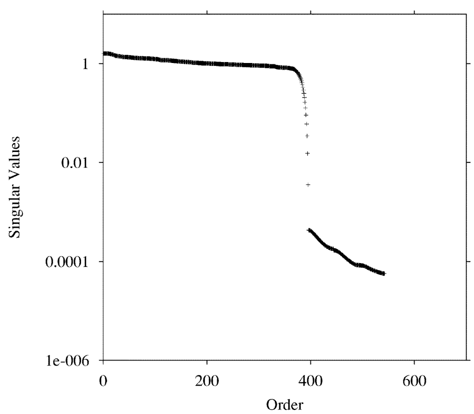

Fig. 10. Singular value for E method (model B, 1 MHz).

The matrix $[K]$ for the E method is obtained successively eliminating the redundant columns and rows from $[K']$ for the A-V method. By this elimination, the eigenvalues next to zero eigenvalues satisfy

$$\lambda \leq \mu \leq 0 \leq \mu^* \leq \lambda^*. \tag{12}$$

Equation (12) suggests that the condition of $[K]$ is worse than that of $[K']$.

The singular values of the preconditioned matrices for model B are shown in Figs. 10 and 11. We again see the singular values

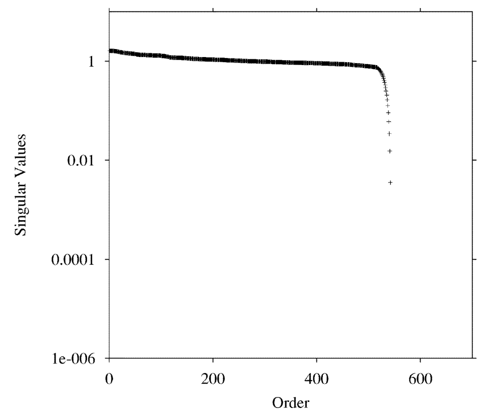

Fig. 11. Singular value for A–V method (model B, 1 MHz).

well clustered near unit for the A–V method and small singular values which spoil the convergence for the E method.

IV. Conclusion

We have discussed the reason for the superiority of the A–V method in the convergence of the preconditioned CG method over E method for the driven microwave problems. The finite-element matrix of the E method contains small negative eigenvalues that make the matrix condition worse, even after the preconditioning such as diagonal scaling and IC. On the other hand, in the A–V method, such eigenvalues are shown to be composed of zeros and normalized ones after the preconditioning. Consequently, as also has been reported for magnetostatic [1] and eddy-current [2] problems, the redundancy in the formulation improves the CG convergence.

Acknowledgment

The authors would like to acknowledge A. Kameari and A. Ahagon for helpful discussions.

References

[1] H. Igarashi, "On the property of the curl-curl matrix in finite element analysis with edge elements," *IEEE Trans. Magn.*, vol. 37, pp. 3129–3132, Sept. 2001.
[2] H. Igarashi and T. Honma, "On the convergence of ICCG applied to finite element equation for quasistatic fields," *IEEE Trans. Magn.*, vol. 38, pp. 565–568, Mar. 2002.
[3] 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.*, vol. 44, pp. 15–23, Jan. 1996.
[4] W. Joubert, "Lanczos methods for the solution of nonsymmetric systems of liner equations," *SIAM J. Matrix Anal. Appl.*, vol. 13, pp. 926–943, 1992.
[5] A. Bossavit, *Computational Electromagnetism*. New York: Academic, 1998.
[6] K. Fujiwara, T. Nakata, and H. Ohashi, "Improvement of convergence characteristics of ICCG method for the A-φ method using edge elements," *IEEE Trans. Magn.*, vol. 32, pp. 804–807, May 1996.
[7] P. C. Hansen, *Rank-Deficient and Discrete Ill-Posed Problems*. Philadelphia, PA: SIAM, 1998.
[8] G. Strang, *Linear Algebra and its Applications*. New York: Academic, 1976.