<table>
<thead>
<tr>
<th>Column 1</th>
<th>Column 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instructions for use</td>
<td></td>
</tr>
<tr>
<td>Title</td>
<td>Effect of Preconditioning in Edge-Based Finite-Element Method</td>
</tr>
<tr>
<td>Author(s)</td>
<td>Igarashi, Hajime; Yamamoto, Nobito</td>
</tr>
<tr>
<td>Citation</td>
<td>IEEE Transactions on Magnetics, 44(6), 942-945</td>
</tr>
<tr>
<td>Issue Date</td>
<td>2008-06</td>
</tr>
<tr>
<td>Doc URL</td>
<td><a href="http://hdl.handle.net/2115/38720">http://hdl.handle.net/2115/38720</a></td>
</tr>
<tr>
<td>Rights</td>
<td>© 2008 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.</td>
</tr>
<tr>
<td>Type</td>
<td>article</td>
</tr>
<tr>
<td>File Information</td>
<td>igarashi08.pdf</td>
</tr>
</tbody>
</table>

Hokkaido University Collection of Scholarly and Academic Papers: HUSCAP
Effect of Preconditioning in Edge-Based Finite-Element Method

Hajime Igarashi¹ and Nobito Yamamoto²

¹Hokkaido University, Sapporo 060-0814, Japan
²University of Electro-Communications, Tokyo 182-8585, Japan

This paper discusses mathematical properties of preconditioned finite-element matrices based on vector potential formulation (A method) and vector and scalar potential formulation (A-V method) for eddy-current problems. Numerical results show that A-V method with preconditioning is stable at all frequencies in contrast to A method. In this paper, this property is mathematically discussed by considering the diagonal scaling which is one of the simple preconditioning methods. In addition, regularization of A method is discussed.

Index Terms—A-V method, diagonal scaling, edge elements, finite-element method, preconditioning.

I. INTRODUCTION

The edge-based finite-element (FE) method has widely been used for electromagnetic field analysis. When we analyze eddy-current problems using this method, we have two ways: A method whose unknown variables are vector potential, and A-V method (or A-ϕ method) whose unknown variables are vector and scalar potentials. Also in the microwave analysis, these two formulations are available.

It is observed in numerical computations that A method gives poor convergence in the iterative solution of linear systems at relatively low frequencies. On the other hand, convergence of the A-V method is kept well even at low frequencies. One of the authors has shown that these differences between A and A-V methods mainly come from different effects of preconditioning [1], [2]. That is, a set of eigenvalues which approach zero as ω → 0 exists in A formulation for both preconditioned and original FE matrices. Due to these eigenvalues, the conditioning of the matrix becomes worse at low frequencies. When using A-V method, such eigenvalues are successfully eliminated by the preconditioning.

In this paper, the above properties are discussed from a mathematical point of view. It will be proved that A-V method with preconditioning works well without poor convergence of ICCG at all frequencies in Section IV. The diagonal scaling is used as the preconditioner for simplicity. In addition, regularization of A method will be discussed.

II. PROBLEM DEFINITION

Although we here focus on quasi-static electromagnetic fields in frequency domain, the following discussion would be valid also for quasi-static problems in time domain and microwave problems [2]. In A-V method, the following equations are solved:

\[
\nabla \times (\nu \nabla \times \mathbf{A}) + j\omega \sigma (\mathbf{A} + \nabla V) = \mathbf{J} \quad (1)
\]

\[
\nabla \cdot (\sigma \mathbf{A} + \nabla \mathbf{V}) = 0 \quad (2)
\]

where ν is the magnetic reluctivity, σ is the conductivity, A is the vector potential, V is the scalar potential, and J denotes the external current which is assumed to be divergence free. Note here that (2) is dependent on (1) because divergence of both sides of (1) yields (2).

The weak form of (1) and (2) can be written in the form

\[
\int_{\Omega} (\nu \nabla \times \mathbf{N}, \nabla \times \mathbf{A} + j\omega \sigma \mathbf{A}) \mathrm{d}v + j\omega \int_{\Omega} \sigma \nabla \mathbf{V} \cdot \mathbf{N} \mathrm{d}v = \int_{\Omega} \mathbf{N} \cdot \mathbf{J} \mathrm{d}v \quad (3)
\]

\[
j\omega \int_{\Omega} \sigma \nabla \mathbf{N} \cdot (\mathbf{A} + \nabla \mathbf{V}) \mathrm{d}v = 0 \quad (4)
\]

where \( \mathbf{N} \) and \( \mathbf{N} \) represent edge-based and scalar basis functions for approximation of \( \mathbf{A} \) and \( \mathbf{V} \), respectively. The FE discretization with edge elements provides

\[
[K] \{q\} = \{J\} \quad (5)
\]

where

\[
[K] = \begin{bmatrix}
[G] & j\omega [G] \\
[j\omega[G]^T] & j\omega [G][G]
\end{bmatrix} \quad (6)
\]

\[
[k] = [C] [\nu] [C] + j\omega [\sigma]. \quad (7)
\]

In (6) and (7), matrices \([C]\) and \([G]\), which are \( f \times e \) and \( e \times n \) matrices with entries \( \pm 1 \) and 0, represent the discrete counterparts of curl and grad, where \( n, e, \) and \( f \) denote the number of nodes, edges, and faces, respectively [3]. The matrices \([\nu]\) and \([\sigma]\) are \( f \times f \) positive definite and \( e \times e \) positive semi-definite symmetric matrices, respectively. The matrix \([C][\nu][C]^T\), whose rank is proved to be \( e-\text{n}+1 \) [4], corresponds to the FE matrix for static magnetic fields.

It is known that the relation \([G][G]^T [C] = 0\) holds, which corresponds to \( \text{div} \ \text{rot} \mathbf{A} = 0 \) in continuum systems. From this property, and with the assumption that the discrete current is divergence free, that is, \([G][J] = 0\), we can readily show that the lower \( n \) rows of matrix \([k]\) are linearly dependent on the upper \( e \) rows. Since the upper \( e \) rows are independent, the rank of \([k]\) is \( e \).

In A method, the scalar potential \( V \) in the above formulation is eliminated due to the fact that \( \text{grad} \ \mathbf{W}^0 \subset W^1 \), where \( \text{grad} \ \mathbf{W}^0 \) and \( W^1 \) represents the spaces spanned by \( \text{grad} \not V \) and \( \mathbf{N} \). This means that \( \mathbf{V} \) can be expressed in terms of \( \mathbf{A} \). Hence, the terms concerning \( \mathbf{V} \) can be eliminated from the equations. Equation (5) then reduces to the system equation of the A method

\[
[k] \{a\} = \{J\}, \quad (8)
\]
Equations (5) and (8) are usually solved using a preconditioned iterative solver such as incomplete Cholesky factorization conjugate gradient method (ICCG).

III. NUMERICAL EXAMPLE

As mentioned earlier, the convergence of ICCG applied for the finite-element matrix generated by A-V method is superior over that by the A method. This tendency comes from the fact that, in A-V method, the incomplete Cholesky decomposition for preconditioning eliminates the "floating eigenvalues" of the FE matrix which approach zero as \( \omega \to 0 \), whereas there is no such elimination in the A method. [1]. This is also valid when the diagonal scaling is used for the preconditioning. Fig. 1 shows the convergence of the CG method with the diagonal scaling for the numerical example where a metallic plate is placed above an excitation coil [1]. It is clear from this figure that the A-V method has better convergence.

It is known that the conjugate gradient methods converge rapidly when the eigenvalues of the coefficient matrix, say \([A]\), tightly cluster around away from the origin [5]. This property can be characterized using the condition number \( \frac{\sigma_{\text{max}}}{\sigma_{\text{min}}} \), where \( \sigma_{\text{max}} \) and \( \sigma_{\text{min}} \) are the maximum and minimum nonzero singular values of \([A]\) [6]. The singular values of \([A]\) are square root of the eigenvalues of \([A]^*A\), where \(^*\) denotes Hermitian conjugate.

Fig. 2 shows the distribution of singular values of the FE matrices after the diagonal scaling. In the singular values corresponding to the A method, there exists a cluster of singular values, which we call floating singular values [1], between the upper cluster including almost unit singular values and the lower cluster with zero eigenvalues. (Due to numerical errors, zero eigenvalues take nonzero values here.) The floating singular values approach zero as \( \omega \to 0 \). The conditioning is, therefore, poor for small \( \omega \to 0 \) in A method. On the other hand, there are no such floating singular values in the spectrum of the A-V method. Note that the A-V method does have floating singular values before preconditioning [1]. The main purpose of this paper is to clarify the reason why the A and A-V method have the above different convergence, or actually different spectra, from a mathematical point of view.

Because it would be easier to analyze the diagonal scaling than the incomplete Cholesky factorization, the effect of the former will be mathematically discussed in the next section. Before proceeding to the next section, we further consider the property of the diagonal scaling using a toy problem [2] to obtain an intuitive understanding of the mathematical properties. Let us consider a \( 2 \times 2 \) matrix

\[
[k'] = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} + \epsilon \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{9}
\]

The first singular matrix on the right-hand side of (9) corresponds to \([C][\mu][C]\), while the second regular matrix corresponds to \([\sigma]\). Moreover, the real parameter \( \epsilon \) stands for \( \omega \). Although the matrix in (9) is simple, it would furnish enough properties for our interest. We can see from the structure of \([k']\) that it is nearly singular, when \( \epsilon \) has a small value.

The diagonal scaling \( k'_{ij} = k_{ij}/\sqrt{k_{ij}k_{jj}} \) of \([k']\) in (9) provides

\[
\begin{bmatrix}
\frac{1}{2} & \frac{1}{2}
\end{bmatrix}
\begin{bmatrix}
\frac{1}{\sqrt{(1+\epsilon)(1+\epsilon)}} & \frac{1}{\sqrt{(1+\epsilon)(1+\epsilon)}}
\end{bmatrix}
\approx
\begin{bmatrix}
1 & 1 - \frac{5}{8}\epsilon
\end{bmatrix}
\begin{bmatrix}
1 & 1
\end{bmatrix}
\tag{10}
\]

for small \( \epsilon \). Hence, the essential structure, a singular matrix plus \( \epsilon \) times regular matrix time, does not change after the diagonal scaling. It is thus clear that the floating eigenvalue, actually \( 5\epsilon/8 \), exists for (10).

We now consider the A-V method. The corresponding toy problem is

\[
[K'] = \begin{bmatrix} 1 + \epsilon & 2 & 2\epsilon \\ 2 & 4 + \epsilon & -\epsilon \\ 2\epsilon & -\epsilon & 5\epsilon \end{bmatrix}.	ag{11}
\]
It can be seen in (11) that the last row is dependent on the upper two rows, and the upper left 2 × 2 matrix is equal to [K] in (9). These properties are similar to [K] in (5). The eigenvalues of matrix [K] are 0, 6ε, and 5 + ε. Hence, its conditioning becomes worse when ε becomes small.

The diagonal scaling of (11) leads to

\[
\tilde{K} = \begin{bmatrix}
1 & \frac{2}{\sqrt{(1+\epsilon)+(\epsilon)}} \\
\frac{\epsilon}{1+\epsilon} & 1
\end{bmatrix}_{\text{sym.}} \tag{12}
\]

The eigenvalues of (12) are shown to be 0, 1 + 2ε/5, 2 - 2ε/5. Note here that the diagonal scaling does not change the rank of the matrix. As ε → 0, nonzero eigenvalues approach 1, 2. Hence, the conditioning is kept good even for small ε in contrast to A method.

IV. PROPERTIES OF DIAGONAL SCALING

We analyze here the effect of the diagonal scaling applied for the FEM matrices of the A and A-V methods. To do so, we consider the distribution of the eigenvalues of the matrices.

In the following, regularity of [σ] is assumed to be regular, the eigenvalues of [σ]. To do so, we write [K] = (D)[K][D], where [D] is a diagonal matrix with entries D_{ii} = 1/\sqrt{\lambda_i}. Since [K] is assumed to be regular, the eigenvalues of [K] approaches of the form 0 < |\lambda_1| ≤ ... ≤ |\lambda_c|. Due to continuity of the eigenvalues with respect to \omega, the eigenvalues of [K] approaches those of [D][K][D] as \omega → 0, where [K] = [D][K][D] is the diagonal matrix corresponding to [k]. Because rank[k] = c - e + 1, there are c - 1 nonzero eigenvalues for [D][K][D]. This means that the c - 1 nonzero eigenvalues of [K] approach zero with \omega. Hence, [K] must have small eigenvalues at low frequencies, which deteriorate the convergence of the conjugate gradient methods.

We next consider the A-V method. In this case, the scaled FEM matrix can be written in the form

\[
\tilde{K} = \left[ \begin{bmatrix} D[k][D] \\
\sqrt{\lambda_i}D_i[D][\sigma][D] \\
\sqrt{\lambda_i}D_i[G][D][D] \end{bmatrix} \right] \tag{13}
\]

where [D] is the diagonal scaling matrix corresponding to [G]. We can then derive the following lemma.

Lemma 1: We assume that \omega > 0. Then the null space of \tilde{K} can be expressed in terms of linear combination of column vectors in the (e + n) × n matrix

\[
V_0 = \begin{bmatrix}
[D]^{-1}[G] \\
-\sqrt{\lambda_i}D_i^{-1}[G]
\end{bmatrix} \tag{14}
\]

Moreover, rank[\tilde{K}] = e.

Proof: It is easily shown that

\[
[\tilde{K}][V_0] = 0 \tag{15}
\]

and rank[\tilde{V}_0] = n irrespective of rank[\sigma]. To show rank[\tilde{K}] = e, let us prove that if \tilde{K}[x] = {0}, then \{x\} = \{V_0\}[z], where \{z\} ∈ C^n.

We split \{x\} into two parts as

\[
\{x\} = \begin{bmatrix}
\{x_1\} \\
\{x_2\}
\end{bmatrix} \tag{16}
\]

and we write \tilde{K}(x) = {0} in the form

\[
[D][k][D]\{x_1\} + \sqrt{\lambda_i}[D][\sigma][G][D]\{x_2\} = \{0\},
\]

\[
\sqrt{\lambda_i}D_i[G][\sigma][D]\{x_1\} + [G][k][G][D]\{x_2\} = \{0\}. \tag{17}
\]

From (17) and the assumption that [k] is regular, we obtain

\[
\{x_1\} = -\sqrt{\lambda_i}[D]^{-1}[k][G][D][x_2]. \tag{18}
\]

By considering the fact that \sqrt{\lambda_i}[\sigma][G] = [k][G], we derive from (17)

\[
\{x_1\} = -\sqrt{\lambda_i}[D]^{-1}[G][D][x_2]. \tag{19}
\]

Hence, we have

\[
\{x\} = \begin{bmatrix}
[D]^{-1}[G] \\
-\sqrt{\lambda_i}[D]^{-1}
\end{bmatrix} \begin{bmatrix}
[x_1] \\
[x_2]
\end{bmatrix}. \tag{20}
\]

Next, we consider the FEM equation after diagonal scaling

\[
\tilde{K}\{a\} = \begin{bmatrix}
[D]^{-1} \\
\sqrt{\lambda_i}D_i^{-1}
\end{bmatrix} \begin{bmatrix}
0 \\
0
\end{bmatrix}. \tag{21}
\]

We can derive the following theorem on the basis of Lemma 1.

Theorem 1: The solution of (21) can be expressed in the form

\[
\begin{bmatrix}
a \\
v^T
\end{bmatrix} = \begin{bmatrix}
[D]^{-1}[a] \\
0
\end{bmatrix} + [V_0][z]. \tag{22}
\]

where {a} is the solution of (8) and {z} ∈ C^n.

Proof: We can directly derive that \{(D)^{-1}[a], \{0\}\} is a solution of (21) by considering the facts that

\[
[D][k][a] = [D][J], \tag{23}
\]

\[
[G][J] = \{0\}. \tag{24}
\]

Moreover, on the basis of Lemma 1, we can see that this theorem holds.

Finally, we obtain the following theorem.

Theorem 2: Let us write the eigenvalues of \tilde{K} as \tilde{\lambda}_i, i = 1, 2, ..., e + n, |\tilde{\lambda}_1| ≤ |\tilde{\lambda}_2| ≤ ... ≤ |\tilde{\lambda}_{e+n}|. Then

\[
\tilde{\lambda}_1 = \tilde{\lambda}_2 = ... = \tilde{\lambda}_e = 0 \tag{25}
\]

\[
0 < |\tilde{\lambda}_{e+1}| ≤ ... ≤ |\tilde{\lambda}_{e+n}| \tag{26}
\]

at all the frequencies including the limit \omega → 0.

Proof: From Lemma 1, it is found that (25) and (26) hold when \omega > 0.

Next, we consider the case \omega → 0. To do so, we write \tilde{K} = [\tilde{P} + \tilde{Q}], where

\[
[\tilde{P}] = \begin{bmatrix}
1 & 0 \\
0 & [G][k][G][D]
\end{bmatrix} \tag{27}
\]

\[
[\tilde{Q}] = \begin{bmatrix}
\sqrt{\lambda_i}D_i[G][\sigma][D] \\
\sqrt{\lambda_i}D_i^{-1}[G][D]
\end{bmatrix}. \tag{28}
\]
Because $D_2 = 1/\sqrt{(k_0)_{ii} + j\omega \sigma[G]}_{ii}$ and $[D_0]$ is independent of frequency, we find
\[
\begin{bmatrix}
\tilde{F} & \tilde{P}_1 = \begin{bmatrix} [D_2] & [D_2] & 0 \\
0 & [D_0] & [G][D_0]
\end{bmatrix} \tilde{P}_1
\end{bmatrix}
\] (29)
\[
[\tilde{G}] \rightarrow [0]
\] (30)
as $\omega \rightarrow 0$. By multiplying $\tilde{P}_1$ by $\tilde{P}_1$, we have
\[
\tilde{P}_1 \tilde{P}_2 \cdots \tilde{P}_n = 0,
\]
where $\tilde{P}_1 \tilde{P}_2 \cdots \tilde{P}_n = 0$. The eigenvalues of $[\tilde{P}_1]$ can be written as
\[
\tilde{\lambda}_1 \tilde{\lambda}_2 \cdots \tilde{\lambda}_n = 0,
\]
where $\tilde{\lambda}_1 \tilde{\lambda}_2 \cdots \tilde{\lambda}_n = 0$. Hence, due to the continuity of eigenvalues, we have
\[
\lambda_i \rightarrow \tilde{\lambda}_i
\] (32)
as $\omega \rightarrow 0$. It is concluded from Theorem 2 that no eigenvalues of $[\tilde{K}]$ of the A-V method approach zero with $\omega$. The conditioning of $[\tilde{K}]$, therefore, keeps well even for low frequencies in contrast to $[K]$ of the A method.

V. REGULARIZED A METHOD

In the previous section, the A method is shown to have poor convergence at low frequencies in contrast to the A-V method. Here we have a question: is it possible to regularize the A method to improve its convergence? The regularized form of the A method could be written in the form
\[
[k_r] \{u_r\} = \{J\},
\]
where $p$ is a positive constant. This form has already been discussed [7]. We will here show this validity.

Lemma 2: The matrix $[k_r]$ is regular for $\omega > 0$.

Proof: We show that $[k_r]\{x\} = \{0\}$ leads to $\{x\} = \{0\}$. From $[k_r]\{x\} = \{0\}$, we have $\{x\} = \frac{\sigma[G][G]^t\{x\}}{\omega}$. Then the quadratic form $(x)^t\sigma[G][G]^t\{x\} = 0$ can be written as
\[
p[x]^t\sigma[G][G]^t\{x\} + j\omega[x]^t\sigma[G][G]^t\{x\} = 0
\] (34)
It is concluded from (34) that $\{x\} = \{0\}$.

Theorem 3: The solutions of (33) is identical to that of (8).

Proof: Since $[k_r]$ is regular, it is sufficient to show that $[k_r]\{a\} = \{J\}$ where $\{a\}$ is the solution of (8). It easy to see that
\[
[k_r]\{a\} = \{J\} + p\sigma[G][G]^t\{a\}.
\] (35)
Due to the fact that $[G]^t\sigma[G][G]^t\{a\} = [\sigma[G][G]^t]J$, we have
\[
p\sigma[G][G]^t\{a\} = \frac{p}{\omega}\sigma[G][G]^t\{J\} = 0.
\] (36)
Hence, $[k_r]\{a\} = \{J\}$.

Finally, we mention stability of the regularized A method.

Theorem 4: The eigenvalues of $[k_r]$ is set to $\mu_i$, $i = 1, 2, \ldots, e$, where $|\mu_1| \leq |\mu_2| \leq \cdots \leq |\mu_e|$. Then $|\mu_i| > 0$, $i = 1, 2, \ldots, e$ holds at all the frequencies including the limit $\omega \rightarrow 0$.

Proof: We split $[k_r]$ into two parts as follows:
\[
[k_r] = [k_{r1}] + j\omega \sigma[G][G]^t\{a\}.
\] (37)
Moreover, the eigenvalues of $[k_{r1}]$ are set to $\nu_1 \leq \nu_2 \leq \cdots \leq \nu_e$. Then we will prove that $[k_{r1}]$ is regular. We begin with the equation $[k_{r1}]\{x\} = \{0\}$, then we have $\{x\} = \{0\}$. Because $[k_0]$ and $\sigma[G][G]^t\{a\}$ are semi-positive definite, we have
\[
\{x\}^t\sigma[G][G]^t\{x\} = 0.
\] (38)
It follows from (38) that
\[
[k_0]\{x\} = 0,
\] (39)
from the first equation of (39), we can see that $\{x\} = \sigma[G]\{c\}$, where $\{c\}$ is an arbitrary $n$-dimensional vector. By multiplying $\{c\}^t$ to the second equation of (39), we have $\{x\}^t\sigma[G][G]^t\{x\} = 0$. Hence, $\{x\} = \{0\}$. Consequently, $\nu_i > 0$, $i = 1, 2, \ldots, e$. Moreover, $|\mu_i| > 0$ due to Lemma 2. It follows from these facts and continuity of eigenvalues with respect to $\omega$ that $[k_{r1}] \rightarrow [k_r]$ as $\omega \rightarrow 0$.

Although the regularized A method is stable for all frequencies like A-V method, it would be difficult to construct the regularization term by superposition. Moreover, because A-V method has more numbers of unknowns than the regularized A method, the former would have better convergence [4].

VI. CONCLUSION

The numerical results shows that the diagonal scaling in the A-V method eliminates the floating singular values which approach zero as $\omega \rightarrow 0$ in contrast to the A method. This leads to the better convergence of preconditioned CG methods for the A-V method. This fact is discussed using a toy problem. In order to show that this property is valid not in the special case, but in general, mathematical proof of this fact is given. In addition, the regularized A method is discussed.

REFERENCES


Manuscript received June 24, 2007. Corresponding author: H. Igarashi