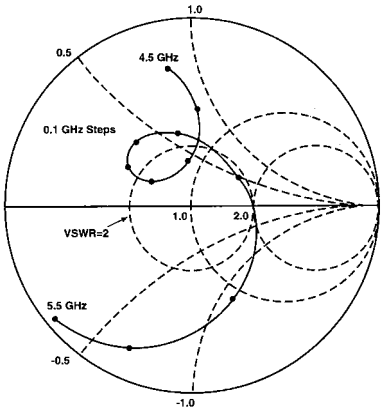
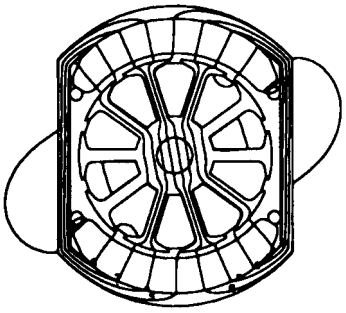
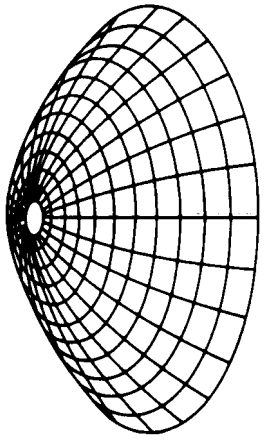


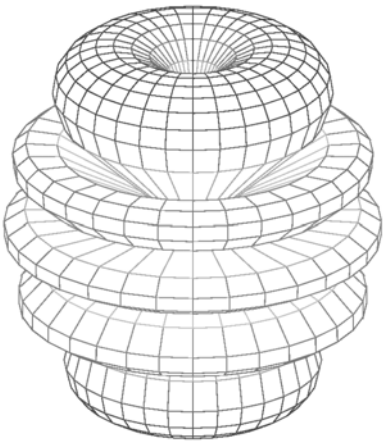


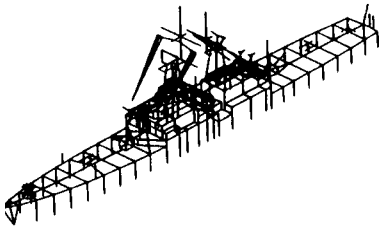
Applied Computational Electromagnetics Society Journal

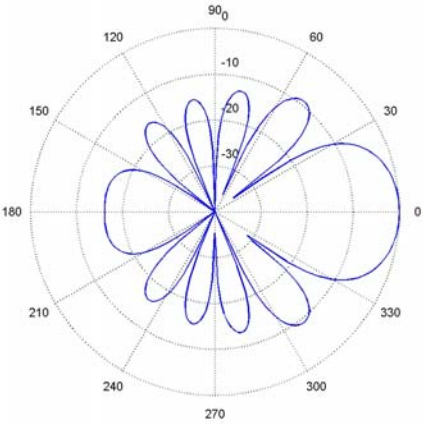


Editor-in-Chief
Atef Z. Elsherbeni

July 2005
Vol. 20 No. 2
ISSN 1054-4887







GENERAL PURPOSE AND SCOPE: The Applied Computational Electromagnetics Society ***(ACES)*** Journal hereinafter known as the ***ACES Journal*** is devoted to the exchange of information in computational electromagnetics, to the advancement of the state-of-the art, and the promotion of related technical activities. A primary objective of the information exchange is the elimination of the need to "re-invent the wheel" to solve a previously-solved computational problem in electrical engineering, physics, or related fields of study. The technical activities promoted by this publication include code validation, performance analysis, and input/output standardization; code or technique optimization and error minimization; innovations in solution technique or in data input/output; identification of new applications for electromagnetics modeling codes and techniques; integration of computational electromagnetics techniques with new computer architectures; and correlation of computational parameters with physical mechanisms.

SUBMISSIONS: The ***ACES Journal*** welcomes original, previously unpublished papers, relating to applied computational electromagnetics. Typical papers will represent the computational electromagnetics aspects of research in electrical engineering, physics, or related disciplines. However, papers which represent research in applied computational electromagnetics itself are equally acceptable.

Manuscripts are to be submitted through the upload system of ***ACES*** web site http://aces.ee.olemiss.edu See "Information for Authors" on inside of back cover and at ***ACES*** web site. For additional information contact the Editor-in-Chief:

Dr. Atef Elsherbeni
Department of Electrical Engineering
The University of Mississippi
University, MS 386377 USA
Phone: 662-915-5382 Fax: 662-915-7231
Email: atef@olemiss.edu

SUBSCRIPTIONS: All members of the Applied Computational Electromagnetics Society who have paid their subscription fees are entitled to receive the ***ACES Journal*** with a minimum of three issues per calendar year and are entitled to download any published journal article available at http://aces.ee.olemiss.edu.

Back issues, when available, are $15 each. Subscriptions to ***ACES*** are through the web site. Orders for back issues of the ***ACES Journal*** and changes of addresses should be sent directly to ***ACES*** Executive Officer:

Dr. Richard W. Adler
ECE Department, Code ECAB
Naval Postgraduate School
833 Dyer Road, Room 437
Monterey, CA 93943-5121 USA
Fax: 831-649-0300
Email: *rwa@attglobal.net*

Allow four week's advance notice for change of address. Claims for missing issues will not be honored because of insufficient notice or address change or loss in mail unless the Executive Officer is notified within 60 days for USA and Canadian subscribers or 90 days for subscribers in other countries, from the last day of the month of publication. For information regarding reprints of individual papers or other materials, see "Information for Authors".

LIABILITY. Neither ***ACES***, nor the ***ACES Journal*** editors, are responsible for any consequence of misinformation or claims, express or implied, in any published material in an ***ACES Journal*** issue. This also applies to advertising, for which only camera-ready copies are accepted. Authors are responsible for information contained in their papers. If any material submitted for publication includes material that has already been published elsewhere, it is the author's responsibility to obtain written permission to reproduce such material.

APPLIED COMPUTATIONAL ELECTROMAGNETICS SOCIETY JOURNAL

Editor-in-Chief
Atef Z. Elsherbeni

July 2005
Vol. 20 No. 2

ISSN 1054-4887

The ACES Journal is abstracted in INSPEC, in Engineering Index, and in DTIC.

The first, fourth, and sixth illustrations on the front cover have been obtained from the Department of Electrical Engineering at the University of Mississippi.

The third and fifth illustrations on the front cover have been obtained from Lawrence Livermore National Laboratory.

The second illustration on the front cover has been obtained from FLUX2D software, CEDRAT S.S. France, MAGSOFT Corporation, New York.

THE APPLIED COMPUTATIONAL ELECTROMAGNETICS SOCIETY

JOURNAL

Vol. 20 No. 2 July 2005

TABLE OF CONTENTS

Effective Preconditioners for the Solution of Hybrid FEM/MoM Matrix Equations using Combined Formulations

Chunlei Guo and Todd H. Hubing
Department of Electrical and Computer Engineering
University of Missouri-Rolla
Rolla, MO 65409

ABSTRACT

Hybrid FEM/MoM modeling codes generate large systems of equations that are generally solved using inward-looking, outward-looking or combined formulations. For many types of problems, the combined formulation is preferred because it does not require a direct inversion of the coefficient matrices and can be solved using iterative solution techniques. An effective preconditioner is a crucial part of the solution process in order to guarantee convergence. However, it can be difficult to generate effective, memory-efficient preconditioners for large problems. This paper investigates preconditioners that use the FEM solution and an absorbing boundary condition (ABC). Various techniques are explored to reduce the memory required by the preconditioner while maintaining effectiveness. Practical problems are presented to evaluate the effectiveness of these preconditioners in various situations.

I. INTRODUCTION

The hybrid finite-element-method/method-of-moments (FEM/MoM) combines the finite element method (FEM) and the method of moments (MoM) and has been used to analyze signal integrity (SI) [1], electromagnetic scattering, and radiation problems [2–6]. FEM is used to model structures with geometrical complexity and inhomogeneous materials. MoM is used to model larger metallic structures and to provide an exact radiation boundary condition (RBC) to terminate the FEM mesh. These two methods are coupled by enforcing field continuity on the boundary separating the FEM and MoM regions.

There are three ways of formulating hybrid FEM/MoM methods [7–9]. The outward-looking formulation constructs an RBC from MoM and incorporates it into the FEM equations. This formulation has been used by Ji *et al*. [9], Jin and Volakis [10], and Ramahi and Mittra [11]. The inward-looking formulation incorporates an RBC constructed from FEM into the MoM equations. This formulation has been utilized by Jin and Liepa [12], Yuan *et al.* [13], and Sheng *et al.* [14]. These two formulations usually involve direct or indirect inversion of the FEM or MoM matrices, so they can be computationally expensive. The combined formulation, on the other hand, combines the FEM and MoM equations and solves for all unknowns at the same time using an iterative solver without requiring a direct matrix inversion. This formulation has been employed by Sheng *et al.* [18]. Techniques to reduce the complexity of the matrix-vector multiplication associated with the MoM part, such as the fast multipole method (FMM) [15, 28], multilevel fast multipole algorithm (MLFMA) [16], and adaptive integral method (AIM) [17], can be readily incorporated into a combined formulation.

The matrix equation generated using the combined formulation is partly full and partly sparse. This matrix is usually ill-conditioned, and the iterative solver may converge very slowly or not at all without an effective preconditioner. An effective preconditioner can reduce the necessary iterations dramatically, resulting in a significant reduction in the overall simulation time. Thus, a preconditioner is a crucial part of the iterative solution. Generally a more accurate approximation of the system results in a more effective preconditioner.

A major feature of a preconditioner is its memory efficiency. In most cases, a preconditioner which utilizes less memory can be developed using a less accurate approximation of the original system. However, this may cause the iterative solver to require more steps to converge; so there is usually a tradeoff between the speed and the memory-efficiency of a preconditioner.

For a system with a small number of unknowns (e.g. 10^3), it is relatively easy to use the complete or incomplete LU (ILU) decomposition of the hybrid

matrix to make an effective preconditioner without exceeding the memory available on typical personal computers. For a system with a large number of unknowns (e.g. 10^6), FMM or similar techniques can be used to reduce the memory requirement. Diagonal, block-diagonal, or near-neighbor matrices are often used to build preconditioners [17], but such preconditioners do not usually yield the most efficient solution. In [19], Liu and Jin proposed a preconditioner using a sparse matrix generated by FEM and an absorbing boundary condition (ABC). This preconditioner was shown to improve the convergence of iterative solvers greatly. This paper further investigates the FEM/ABC preconditioner and proposes a modified preconditioner for geometries with large metallic surfaces.

Section II of this paper presents the necessary formulations. Representative examples are introduced in Section III. Reordering techniques to reduce the number of fill-in elements in ILU decompositions are discussed in Section IV and a modified preconditioner that further reduces the memory requirement is proposed. Finally, conclusions from the work presented here are drawn in Section V.

II. FORMULATIONS

The Hybrid FEM/MoM Using the Combined Formulation

In the hybrid FEM/MoM, an electromagnetic problem is divided into an interior equivalent part and an exterior equivalent part. The interior part is modeled using the FEM and the exterior part is modeled using a surface integral equation method. The two parts are coupled by enforcing the continuity of tangential fields on the FEM and MoM boundary. FEM can be used to analyze the interior equivalent part and generates a sparse matrix equation of the form,

$$\begin{bmatrix} A_{ii} & A_{is} \\ A_{si} & A_{ss} \end{bmatrix} \begin{bmatrix} E_i \\ E_s \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & B_{ss} \end{bmatrix} \begin{bmatrix} 0 \\ J_s \end{bmatrix} + \begin{bmatrix} g_i \\ g_s \end{bmatrix}. \quad (1)$$

A detailed explanation of Equation (1) can be found in [9].

The exterior equivalent problem can be analyzed by using an electric field integral equation (EFIE), magnetic field integral equation (MFIE), or a combined field integral equation (CFIE), which is a linear combination of the EFIE and MFIE. Regardless of the choice of testing functions or integral equations, the MoM matrix equation has the following form [9],

$$[C]\ [J_s] = [D]\ [E_s] - [F] \quad (2)$$

where J_s is a set of unknown complex scalar coefficients, C and D are dense coefficient matrices, and F is the combined source term specifically given by [20],

$$F_m = \int_{S_m} \mathbf{f}_m(\mathbf{r}) \bullet \left[\mathbf{E}^{\text{inc}}(\mathbf{r}) + \eta_0 \hat{n} \times \mathbf{H}^{\text{inc}}(\mathbf{r}) \right] dS \quad (3)$$

where $\boldsymbol{f}_m(\boldsymbol{r})$ is the set of basis functions on the surface and $\hat{n}$ is a unit normal vector pointing outward from the surface S.

Neither Equation (1) nor Equation (2) can be solved independently. These two equations form a coupled and determined system. The combined formulation is obtained from Equations (1) and (2) as,

$$\begin{bmatrix} A_{ii} & A_{is} & 0 \\ A_{si} & A_{ss} & -B_{ss} \\ 0 & D & -C \end{bmatrix} \begin{bmatrix} E_i \\ E_s \\ J_s \end{bmatrix} = \begin{bmatrix} g_i \\ g_s \\ F \end{bmatrix}. \quad (4)$$

Unlike the inward-looking formulation or the outward-looking formulation, the combined formulation doesn't require an explicit inversion of any matrix and solves for all the unknowns simultaneously. For many configurations, with proper preconditioning, the combined formulation is the most computationally efficient of the three formulations.

Equation (4) can be written in the form,

$$Mx = b \quad (5)$$

where

$$M = \begin{bmatrix} A_{ii} & A_{is} & 0 \\ A_{si} & A_{ss} & -B_{ss} \\ 0 & D & -C \end{bmatrix} \quad (6)$$

$x = \begin{bmatrix} E_i & E_s & J_s \end{bmatrix}^T$, and $b = \begin{bmatrix} g_i & g_s & F \end{bmatrix}^T$. Notice that the matrix M is a hybrid matrix, which is partly full and partly sparse. The convergence rate of an iterative solver for Equation (5) is highly dependent on the condition number of M.

The matrix M usually has a very large condition number (e.g. on the order of 10^6 or higher), which results in poor convergence or non-convergence of the iterative solution. However, Equation (5) can be transformed into another linear equation with the same solution,

$$PMx = Pb \tag{7}$$

where $P \approx M^{-1}$ is a preconditioner matrix. With the proper choice of P, the matrix PM has better spectral properties than M and the number of iterations required to converge is greatly reduced. Ideally, the construction and application of a good preconditioner should be fast without requiring a lot of memory.

Application of FEM and ABC as a Preconditioner

Liu and Jin proposed a preconditioner that applies an absorbing boundary condition on the truncation surface to approximate the MoM boundary condition [19]. A sparse preconditioning matrix can be formed by replacing the EFIE or MFIE equations with first-order ABCs [8] on the truncation surface S,

$$\hat{n} \times \nabla \times \boldsymbol{E}^s(\boldsymbol{r}) = jk_0 \boldsymbol{E}^s(\boldsymbol{r}),\ \boldsymbol{r} \in S \tag{8}$$

where $\boldsymbol{E}^s = \boldsymbol{E} - \boldsymbol{E}^{inc}$ is the scattered electric field. From Equation (8) we can derive,

$$\begin{aligned} &\boldsymbol{E}(\boldsymbol{r}) + \eta_0 \hat{n} \times \boldsymbol{H}(\boldsymbol{r}) = \\ &\qquad \boldsymbol{E}^{inc}(\boldsymbol{r}) + \eta_0 \hat{n} \times \boldsymbol{H}^{inc}(\boldsymbol{r}), \quad r \in S. \end{aligned} \tag{9}$$

The same basis functions on the surface S used in the hybrid FEM/MoM method can be applied to approximate the **E** and **H** fields in Equation (9), and the basis functions $\boldsymbol{f}_m(\boldsymbol{r})$ can be used to test Equation (9), resulting in a matrix equation of the form,

$$\left[B_{ss}^T\right]\left[E_s\right] + \left[H_{ss}\right]\left[J_s\right] = \left[L\right] \tag{10}$$

where $\left[B_{ss}^T\right]$ is the transpose of the matrix B_{ss} in Equation (1), H_{ss} is given by

$$\left[H_{ss}\right]_{mn} = \int_{S_m} \eta_0 \boldsymbol{f}_m(\boldsymbol{r}) \bullet \boldsymbol{f}_n(\boldsymbol{r}) dS \tag{11}$$

and L is the source term given by,

$$\begin{aligned} L_m = &\int_{S_m} \boldsymbol{f}_m(\boldsymbol{r}) \bullet \boldsymbol{E}^{inc}(\boldsymbol{r})\, dS \\ &+ \eta_0 \int_{S_m} \boldsymbol{f}_m(\boldsymbol{r}) \bullet \hat{n} \times \boldsymbol{H}^{inc}(\boldsymbol{r})\, dS. \end{aligned} \tag{12}$$

The ABC approximates the MoM boundary condition described using the TENH form of the CFIE.

Combining Equation (10) with Equation (1), we have,

$$\begin{bmatrix} A_{ii} & A_{is} & 0 \\ A_{si} & A_{ss} & -B_{ss} \\ 0 & B_{ss}^T & H_{ss} \end{bmatrix} \begin{bmatrix} E_i \\ E_s \\ J_s \end{bmatrix} = \begin{bmatrix} g_i \\ g_s \\ K \end{bmatrix} \tag{13}$$

and we define

$$Q = \begin{bmatrix} A_{ii} & A_{is} & 0 \\ A_{si} & A_{ss} & -B_{ss} \\ 0 & B_{ss}^T & H_{ss} \end{bmatrix}. \tag{14}$$

Notice that the matrix Q is very sparse. Now we have two systems: the hybrid FEM/MoM system described by Equation (4) and the FEM/ABC system described by Equation (13).

Since Equation (8) describes the behavior of the electric field in free space far from the sources, the ABC truncation surface should not be too close to the scatterer. When the ABC truncation surface is far enough from the scatterer's surface (e.g. $\lambda/10$), the FEM/ABC system described by Equation (13) can be a good physical approximation of the hybrid system described by Equation (4).

If the same computational domain is analyzed using both Equations (4) and (13), the matrix M is of the same order as the matrix Q. The matrix Q is a physical approximation of M, and Q^{-1} can be used as a preconditioner to improve the iterative solution of Equation (6). Q is highly sparse, and doesn't require much additional memory. Q^{-1} can be generated much faster and more efficiently than M^{-1}, particularly for problems with a lot of MoM boundary elements.

Since the preconditioning technique requires the FEM/MoM boundary to be located far from the surface of the scatterer, more elements may be required increasing the order of the system of equations. However, in many situations the amount of additional computational resources required by these extra elements is small compared to the resources saved by using this preconditioning technique.

III. SAMPLE PROBLEMS

Four sample problems were used to evaluate the preconditioning techniques discussed in later sections. The first problem is a perfectly conducting (PEC) sphere, which does not require any FEM elements to model. The second problem is a dielectric-coated sphere, where the coating is thin relative to the radius of the sphere. This structure requires both FEM and MOM elements to model. The third problem is a solid dielectric sphere

requiring many more FEM elements. These spherical structures are convenient because they can be modeled analytically. The FEM part of the hybrid system becomes more dominant from the first problem to the third problem. The remaining problem is a printed circuit board (PCB) power bus structure, which is a structure of particular interest to EMC and signal integrity engineers. Each sample configuration is modeled at 3 GHz.

Problem 1: Perfectly Conducting Sphere

The first sample configuration is a perfectly conducting sphere. The radius of the sphere is 8 cm, as shown in Fig. 1. The incident wave travels along the z-axis, and the polarization of the E field is along x-axis. The goal is to model the scattered far fields. The most convenient way to model this structure is to use MoM on the surface of the sphere, so FEM is not required.

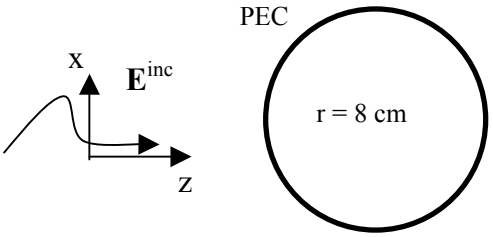


Figure1. Scattering from a PEC sphere.

Problem 2: Dielectric-Coated PEC Sphere

The second configuration is a dielectric-coated PEC sphere. The radius of this sphere is also 8 cm. The coated dielectric material has a thickness of 5 mm (0.05λ at 3 GHz) and relative dielectric constant of 4.0-j1.0, as shown in Fig. 2. The same incident wave as Problem 1 is applied. The field in the interior of the dielectric material is analyzed using FEM, and the equivalent current on the truncation surface is modeled using MoM.

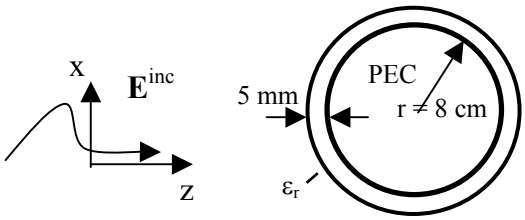


Figure 2. Scattering from a dielectric-coated PEC sphere.

Problem 3: Dielectric Sphere

The third sample configuration is a dielectric sphere. The radius of this sphere is again 8 cm and the relative dielectric constant of the sphere material is 4.0, as indicated in Fig. 3. The same incident wave as Problem 1 is applied. The field in the interior of the dielectric sphere is analyzed using FEM, and the equivalent current on the truncation surface is modeled using MoM.

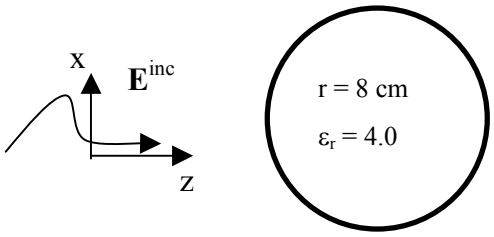


Figure 3. Scattering from a dielectric sphere.

Problem 4: Power Bus Structure

The fourth problem is to model the input impedance of a PCB power bus structure. As shown in Fig. 4, the board dimensions are 30 mm × 20 mm × 2 mm. The top and bottom planes are PECs. The relative dielectric constant of the material between the planes is 4.2. An ideal current source is located at ($x_i = 10$ mm, $y_i = 5$ mm) to excite the structure. Such a structure usually requires a large number of FEM elements between the planes in order to control the aspect ratio of the tetrahedra. This results in a lot of triangular MoM boundary elements if the FEM/MoM boundary is located on the surface of the metal planes.

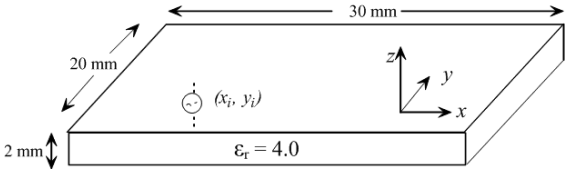


Figure 4. A PCB power bus structure.

Discretization of Sample Structures

Since the MoM provides an exact RBC on the truncation surface, it doesn't matter how far the truncation surface is from the scatterer. However, it is usually convenient to choose the truncation surface to coincide with the physical boundary of the scatterer to minimize the computational domain. Defining the distance between the truncation surface and physical boundary of the scatterer as d, Table 1(a) summarizes the discretization of the sample problems when $d = 0$ (i.e., the truncation surface coincides with the physical boundary of the spheres or PCB). The mesh density on this truncation surface is about 10 elements/wavelength. The total number of unknowns is given by the sum of the number of E_i, E_s, and J_s elements.

Table 1. Summary of discretization of the sample problems

(a) when $d = 0$

	# of tetrahedra	# of E_i	# of E_s	# of J_s	Total # of unknowns
Problem 1	0	0	0	2589	2589
Problem 2	5491	3660	2910	2910	9480
Problem 3	6541	6379	2589	2589	11557
Problem 4	456	274	68	552	832

(b) when $d = 5$ mm

	# of tetrahedra	# of E_i	# of E_s	# of J_s	Total # of unknowns
Problem 1	5491	3660	2910	2910	9480
Problem 2	11923	10851	3549	3549	17949
Problem 3	12032	12628	2910	2910	18448
Problem 4	3353	2724	1464	1464	5652

In order for the FEM-ABC preconditioner to be effective, the truncation surface has to be moved away from the physical boundary of the scatterer. In this study, a truncation surface with $d = \lambda/20 = 5$ mm in each direction was used when the FEM/ABC preconditioner was employed. This choice of d should be sufficient to provide a good preconditioner [19]. Table 1(b) summarizes the discretization when d = 5 mm. The mesh density on this truncation surface is also about 10 elements/wavelength.

As we can see from Table 1, applying the FEM-ABC preconditioner increases the total number of unknowns roughly by a factor of 2 to 7 for the structures studied in this paper. For the PEC sphere, no FEM elements are necessary when d = 0. Only the coefficient matrix C needs to be saved and only MoM is applied in this case. When the truncation boundary is moved away from the surface of the sphere, tetrahedral finite elements are added and the memory requirement at least doubles since the matrix D (which is as dense as the C matrix) also needs to be saved.

For the dielectric-coated sphere, the number of tetrahedral elements roughly doubles when d is increased from 0 to 5 mm. The number of E_i elements roughly triples, and the number of E_s and J_s elements also increases due to the larger surface area.

For the dielectric sphere, there are a large number of tetrahedral elements even when the truncation boundary coincides with the physical boundary of the sphere. However, the number of unknowns still increases significantly when the truncation surface is extended beyond the physical boundary of the sphere.

For the power bus structure, the thickness of the board requires a fine tetrahedral mesh in the dielectric in order to ensure that the tetrahedra have a reasonable aspect ratio. When d is increased from 0 to 5 mm, many more tetrahedral elements must be used to discretize the computational domain between the physical boundary of the board and the truncation boundary, resulting in a large increase in the total number of unknowns. At lower frequencies (longer wavelengths), the boundary would need to be located even further from the scattering surfaces.

IV. PRECONDITIONING TECHNIQUES

The inverse of the matrix Q in Equation (14) can be used as a preconditioner for iterative solutions of Equation (4). However, it is usually very expensive to derive an explicit inverse of this matrix. An incomplete LU factorization of the matrix Q will result in $Q \approx LU$, where L is a sparse, lower-triangular matrix, and U is a sparse, upper-triangular matrix. The preconditioner, P, is then given by $P = Q^{-1} \approx (LU)^{-1}$, where the inversion is actually replaced by forward and back substitution at each iteration.

There are two popular ILU schemes, one based on the structure of the matrix being factored, and the other based on the numerical values of the

elements in L or U generated during factorization [21]. In the first scheme, an element in L or U is dropped if the element in the corresponding position of the original matrix is zero, no matter how large this element is. In the second scheme, an element in L or U is discarded only if its magnitude is smaller than a specified *drop tolerance*. The second scheme often yields more accurate factorizations than the first scheme. Variations of each scheme and hybridizations of these schemes are also described in the literature [21].

In this study, the LU factorization based on drop tolerance in MATLAB was used [25]. The drop tolerance was set to 1.0×10^{-6} for the results presented here. A smaller drop tolerance yields a more accurate factorization, but produces more fill-in elements. Fill-in elements refer to matrix entries that are zero in the original matrix Q and are nonzero in the L and U matrices. In order to reduce the number of fill-in elements (and the memory required to store these elements), the matrix Q was reordered before the factorization.

Reducing the Number of Fill-ins During ILU by Reordering

There are various reordering algorithms, including variable band, nested dissection and minimum degree [22, 23]. A good variable band algorithm is the reverse Cuthill-McKee algorithm to minimize the bandwidth of a matrix [24]. The minimum degree algorithm is based on graph theory and reduces fill-in elements during Gaussian elimination [25, 26]. In [9], it is shown that the symmetric reverse Cuthill-McKee algorithm (SYMRCM) and symmetric minimum degree algorithm (SYMMMD) effectively reduce the fill-ins during a complete LU on a sparse matrix generated using FEM. In this work, besides the SYMRCM and SYMMMD algorithms, another minimum degree algorithm, the symmetric approximate minimum degree reordering technique (SYMAMD) was also investigated [25]. This algorithm is usually faster than the symmetric minimum degree algorithm and yields a better ordering.

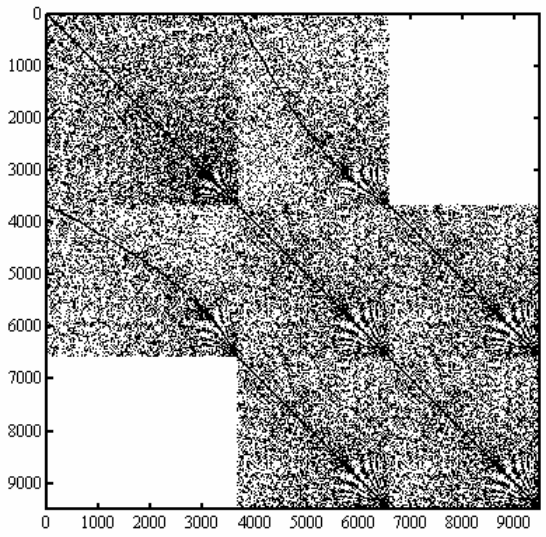
The sparsity pattern of the matrix Q generated using FEM and ABC for Problem 1 (d = 5 mm) is shown in Fig. 4(a). The average number of nonzero elements is about 12 elements per row in this case, which indicates that Q is highly sparse. The sparsity pattern using various reordering algorithms is also shown in Fig. 5.

Table 2 lists the number of nonzero elements in the L and U matrices after an ILU factorization of matrix Q using a drop tolerance of 1.0×10^{-6}. Problems 2 and 3 could not be factored within the available memory without reordering. For L and U using sparse complex values with double precision, the required memory (in bytes) is roughly given by the number of nonzero elements times 20. The memory required to store the L and U matrices is also listed in Table 2. As we can see, the number of fill-in elements during ILU is greatly reduced by reordering the matrix. It is also much faster to perform ILU factorizations when the reordering schemes are applied. On average, the SYMRCM algorithm performed a little better than the other algorithms. This is probably due to the asymmetric nature of the matrix Q.

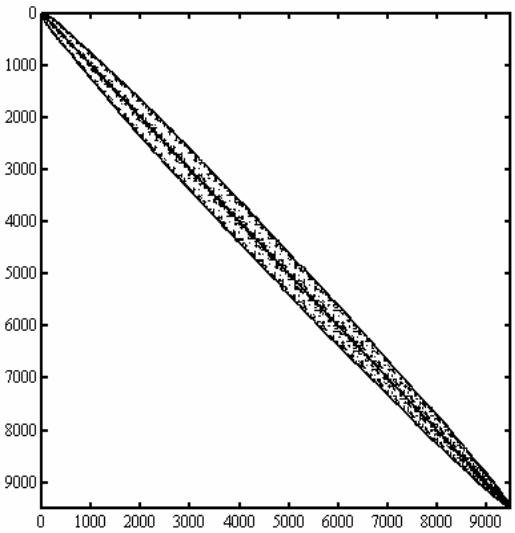
Iterative Solver Behavior

After L and U are generated, they can be applied to the iterative solver at each iteration and do not have to be explicitly inverted. In this study, a bi-conjugate gradient stabilized (BICGSTAB) solver was utilized [21, 27]. Table 3 summarizes the number of iterations required to achieve a solution with a convergence factor of 1.0×10^{-3}. The convergence factor is the maximum value for the normalized residual norm, $\|Mx - b\| / \|b\|$. In other words, the BICGSTAB solver has converged once $\|Mx - b\| / \|b\| \le 1.0 \times 10^{-3}$ is achieved. The maximum number of iterations investigated in this study was 500.

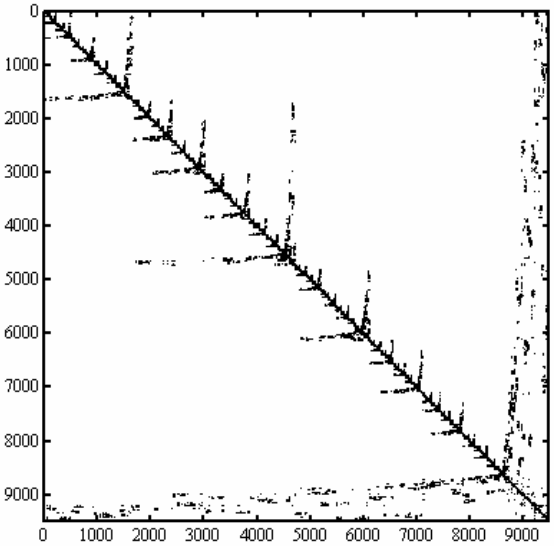
The general behavior of the BICGSTAB solver is described as being divergent, convergent, or stagnant in Table 3. For divergent behavior, the normalized residual norm bounces between certain values above the required tolerance as the number of iterations increases. For convergent behavior, the normalized residual goes below the tolerance in less than 500 iterations. For stagnant behavior, the normalized residual norm remains the same for two consecutive iterations. The BICGSTAB solver stops the solution process before reaching the maximum number of iterations if stagnant behavior occurs.



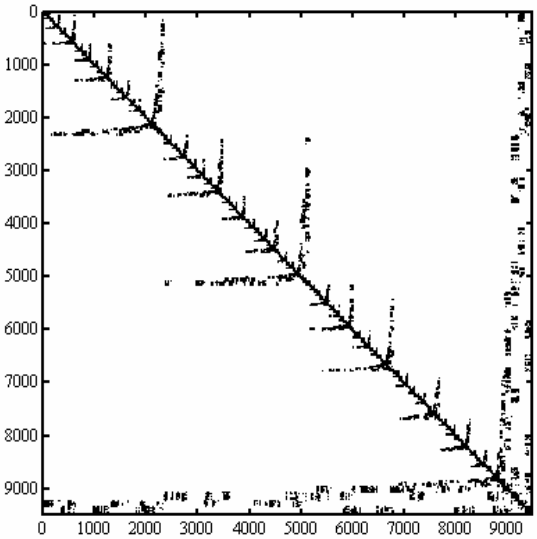
(a) Sparsity pattern of the original matrix Q.



(b) Sparsity pattern of the matrix Q after SYMRCM ordering.



(c) Sparsity pattern of the matrix Q after SYMMMD ordering.



(d) Sparsity pattern of the matrix after SYMAMD ordering.

Figure 5. Sparsity pattern for Problem 1 matrix generated using FEM and ABC.

Table 2. The number of nonzero elements in L and U after ILU

	No ordering (Mbytes)	SYMRCM (Mbytes)	SYMMMD (Mbytes)	SYMAMD (Mbytes)
Problem 1	21151493 (423)	2507653 (50)	2701435 (54)	2710165 (54)
Problem 2	Out of memory	9534816 (190)	10000965 (200)	10415806 (208)
Problem 3	Out of memory	27943414 (559)	33754661 (675)	Out of memory
Problem 4	8875433 (176)	2319723 (46)	2504835 (50)	2437372 (49)

Table 3. Iterations required using un-preconditioned and preconditioned BICGSTAB

	Un-preconditioned BICGSTAB			Preconditioned BICGSTAB		
	# of Iterations	Converged (Yes/No)	General behavior	# of Iterations	Converged (Yes/No)	General behavior
Problem 1	500	No	Divergent	14	Yes	Convergent
Problem 2	500	No	Divergent	14	Yes	Convergent
Problem 3	500	No	Divergent	35	Yes	Convergent
Problem 4	47	No	Stagnant	27	Yes	Convergent

Table 4. Number of nonzero elements in L and U and iterations required to converge when the coupling between FEM and ABC was discarded, and SYMMMD was used

	Problem 1	Problem 2	Problem 3	Problem 4	
				Radiation	Scattering
# of nonzero element in *L* and *U* (MBytes)	377965 (8)	2515255 (50)	17715266 (354)	597535 (12)	
# of iterations for convergence	24	26	500 (Did not converge)	103	40

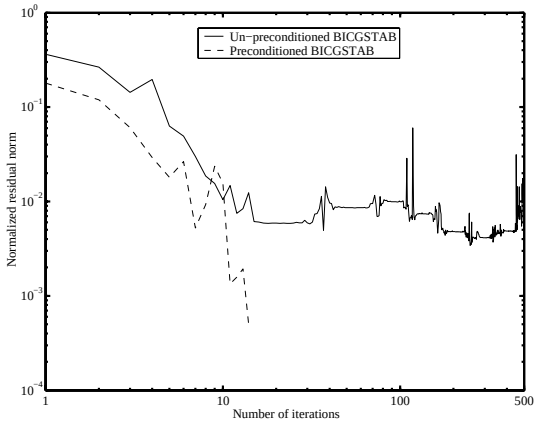


Figure 6. Convergence/divergence curve for Problem 1.

Figure 6 plots the normalized residual norm as a function of the number of iterations for Problem 1 using the un-preconditioned and preconditioned BICGSTAB solver. As we can see, the FEM-ABC preconditioner greatly reduces the number of iterations required for convergence.

Reducing the Number of Fill-ins by Decoupling FEM and ABC

In this study, a modified FEM/ABC preconditioner requiring less memory was also evaluated. In Equation (13), the coupling between FEM and ABC is achieved through the B_{ss} and B_{ss}^T coefficient matrices. Although the elements in these matrices are bigger than those on the same row in the diagonal entries of the matrix Q, we discarded the B_{ss} and B_{ss}^T matrices and used the resulting sparse matrix, Q', to construct preconditioners. For scattering problems like Problems 1, 2 and 3, discarding the coupling between the FEM and ABC is effectively the same as imposing a PEC boundary condition on the truncation surface. For radiation problems like Problem 4, discarding the coupling between the FEM and ABC effectively imposes a perfectly magnetically conducting (PMC) boundary condition on the truncation surface.

Discarding the elements corresponding to the coupling between FEM and ABC dramatically reduces the number of fill-ins during ILU factorization. Table 4 lists the number of nonzero elements and the number of iterations required for convergence. The memory required to store the L and U matrices is given in parentheses.

Comparing the results in Table 4 to the results in Tables 2 and 3, we observe that this preconditioner works reasonably well for PEC and dielectric-coated PEC spheres. The number of iterations required to converge is higher, but the memory required is significantly reduced. Since the ABC truncation surface is close to the PEC sphere in both cases, discarding the coupling between the

FEM and ABC (implicitly applying a PEC boundary condition) is a reasonable approximation of the FEM and ABC. Figures 7 and 8 show the calculated radar cross section (RCS) for the PEC sphere in Problem 1 and the dielectric-coated sphere in Problem 2 using the decoupled FEM-ABC as preconditioner, respectively. Analytical results for the RCS of this geometry obtained using the Mie series [29] are also provided. The results obtained using the two methods agree with each other very well.

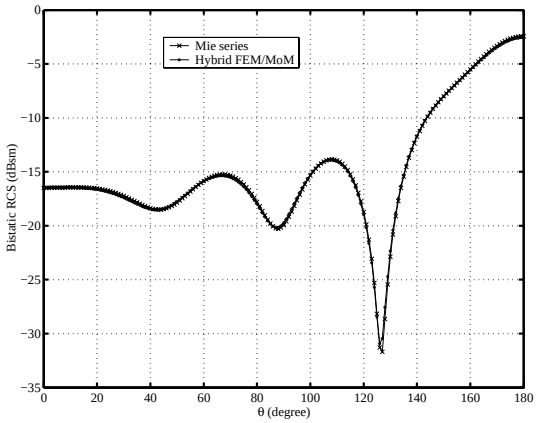


Figure 7. Calculated RCS for Problem 1.

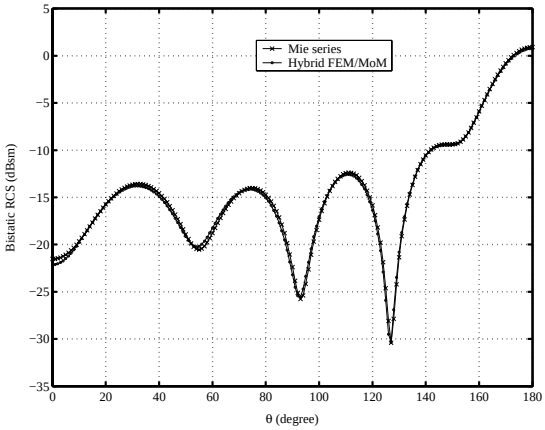


Figure 8. Calculated RCS for Problem 2.

For the PCB power bus structure, the memory was reduced by a factor of 4, however, the number of iterations increased by a factor of 4, as shown in the first sub-column in Table 4 for Problem 4. In this case, the excitation is located inside the FEM region. The preconditioner implicitly applies a PMC boundary condition. To demonstrate the different behavior of the proposed preconditioner for scattering problems vs. radiation problems, the same configuration was modeled with the incident wave from Problem 1 instead of the internal current source. This change only affects the terms on the right hand side of Equation (4). The memory required by the preconditioners in the scattering case is the same as that in the radiation case. Using the original preconditioner with FEM and ABC coupled together, 17 iterations are required to reach a 1.0×10^{-3} convergence factor. Applying the new preconditioner with FEM and ABC decoupled (i.e. implicitly applying a PEC boundary condition), the memory is still reduced by a factor of 4 and 40 iterations are required to converge.

For the dielectric sphere, the iterative solver did not converge for Problem 3. The normalized residual norm oscillated around 3.0×10^{-3}. Therefore, in this case, discarding the coupling between the FEM and ABC elements resulted in a poorer preconditioner.

Eliminating the coupling terms between the FEM and ABC portions of the preconditioning matrix appears to work pretty well for scattering problems from structures with large metallic surfaces. However, it does not work as well for radiation problems or for modeling structures without large metal surfaces.

V. CONCLUSIONS

In this paper, four sample problems were used to investigate the application of preconditioning techniques to the iterative solution of matrix equations resulting from the hybrid FEM/MoM method employing a combined formulation. These techniques were based on the FEM/ABC, which yields a physical approximation of the geometry being evaluated. An ILU factorization employing a drop tolerance was used to construct the preconditioner. Reordering algorithms reduced the number of nonzero elements in the L and U matrices by a factor of 4 to 8, depending on the geometry and reordering scheme applied.

When FEM/ABC preconditioners were applied to the solution of the hybrid FEM/MoM system of equations, the convergence rate of the iterative solution improved significantly. These preconditioners work very well for calculating the scattering from PEC spheres (with few FEM elements), dielectric coated spheres, and dielectric spheres (with many FEM elements). They also worked well for modeling radiation and scattering from a PCB power bus structure (with many FEM elements).

The memory required by the preconditioner can limit the size of the problems that can be modeled. This memory requirement can be reduced significantly by discarding the coupling between the FEM and ABC elements in the preconditioner matrix. In our examples, memory was reduced by a factor of 2 to 7. This modified FEM/ABC preconditioning technique worked very well for analyzing the scattering from PEC and dielectric-coated metal spheres. It also worked well for analyzing the scattering from a PCB. However, it did not work well for modeling the radiation from the PCB or for modeling a dielectric sphere. Since this approach implicitly simulates a PEC boundary for scattering problems (or a PMC boundary for radiation problems), it is generally expected to work well for the analysis of scattering from geometries that have a PEC surface near the FEM/MoM boundary.

ACKNOWLEDGEMENT

The authors wish to express their thanks to Professor W. C. Chew of University of Illinois at Urbana-Champaign for providing the Mie series code.

REFERENCES

[1] Y. Ji and T. H. Hubing, "On the modeling of a gapped power-bus structure using a hybrid FEM/MoM approach", *IEEE Trans. Electromag. Compat.,* vol. 44, no. 4, pp. 566–569, Nov. 2002.

[2] X. Yuan, "Three dimensional electromagnetic scattering from inhomogeneous objects by the hybrid moment and finite element method," *IEEE Trans. Microwave Theory Tech.*, vol. 38, pp. 1053–1058, Aug. 1990.

[3] J.-M. Jin and J. L. Volakis, "Electromagnetic scattering by and transmission through a three-dimensional slot in a thick conducting plane," *IEEE Trans. Antennas Propagat.*, vol. 39, pp. 543–550, Apr. 1991.

[4] J.-M. Jin and J. L. Volakis, "A hybrid finite element method for scattering and radiation by microstrip patch antennas and arrays residing in a cavity," *IEEE Trans. Antennas Propagat.*, vol. 29, pp. 1598–1604, Nov. 1991.

[5] M. W. Ali, T. H. Hubing, and J. L. Drewniak, "A hybrid FEM/MOM technique for electromagnetic scattering and radiation from dielectric objects with attached wires," *IEEE Trans. Electromag. Compat.*, vol. 39, pp. 304–314, Nov. 1997.

[6] J. Angélini, C. Soize, and P. Soudais, "Hybrid numerical method for harmonic 3D Maxwell equations: Scattering by a mixed conducting and inhomogeneous anisotropic dielectric medium," *IEEE Trans. Antennas Propagat.*, vol. 41, pp. 66–76, Jan. 1993.

[7] L. W. Pearson, A. F. Peterson, L. J. Bahrmasel, and R. A. Whitaker, "Inward-looking and outward-looking formulations for scattering from penetrable objects," *IEEE Trans. Antennas Propagat.*, vol. 40, pp. 714–720, Jun. 1992.

[8] A. F. Peterson, S. L. Ray, and R. Mittra, *Computational Methods for Electromagnetics*, New York: IEEE Press and Oxford University Press, 1997.

[9] Y. Ji, H. Wang and T. H. Hubing, "A novel preconditioning technique and comparison of three formulations for the hybrid FEM/MoM method", *Appl. Computat. Electromagn. Soc. (ACES) J.*, vol. 15, pp. 103–114, Jul. 2000.

[10] J.-M. Jin and J. L. Volakis, "A finite element-boundary integral formulation for scattering by three-dimensional cavity-backed apertures," *IEEE Trans. Antennas Propagat.*, vol. 39, pp. 97–104, Jan. 1991.

[11] O. M. Ramahi and R. Mittra, "Finite element solution for a class of unbounded geometries," *IEEE Trans. Antennas Propagat.*, vol. 39, pp. 244–250, Feb. 1991.

[12] J.-M. Jin and V. V. Liepa, "A note on hybrid finite element method for solving scattering problems," *IEEE Trans. Antennas Propagat.*, vol. 36, pp. 1486–1490, Oct. 1988.

[13] X. Yuan, D. R. Lynch, and J .W. Strohbehn, "Coupling of finite element and moment methods for electromagnetic scattering from inhomogeneous objects," *IEEE Trans. Antennas Propagat.*, vol. 38, pp. 386–393, Mar. 1990.

[14] X. Q. Sheng and E. K. Yung, "Implementation and experiments of a hybrid algorithm of the MLFMA–Enhanced FE-BI method for open-region inhomogeneous electromagnetic problems", *IEEE Trans. Antennas. Propagat.*, vol. 50, no. 2, pp. 163–167, Feb. 2002.

[15] R. Coifman, V. Roklin, and S. Wandzura, "The fast multipole method for the wave equation: a pedestrian prescription," *IEEE Antennas Propag. Mag.*, vol. 35, pp. 7–12, June 1993.

[16] J. M. Song and W. C. Chew, "MLFMA for electromagnetic scattering from large complex objects," *IEEE Trans. Antennas Propag.*, vol. 45, pp. 1488–1493, Oct. 1997.

[17] E. Bleszynski, M. Bleszynski, and T. Jaroszewicz, "A fast integral-equation solver for electromagnetic scattering problems", *IEEE AP-S Int. Symp. Dig.*, Seattle, WA, pp. 416–419, June 1994.

[18] X.-Q. Sheng, J.-M. Jin, J. Song, C.-C. Lu, and W.C. Chew, "On the formulation of hybrid finite-element and boundary-integral methods for 3-D scattering," *IEEE Trans. Antennas Propagat.*, vol. 46, pp. 303–311, Mar. 1998.

[19] J. Liu and J.-M. Jin, "A highly effective preconditioner for solving the finite element-boundary integral matrix equation of 3-D scattering", *IEEE Trans. Antennas Propagat.*, vol. 50, no. 9, pp. 1212–1221, Sep. 2002.

[20] Y. Ji, H. Wang, and T. H. Hubing, "A numerical investigation of interior resonances in the hybrid FEM/MoM method", IEEE *Trans. Antennas Propagat.*, vol. 51, no. 2, pp. 347–349, Feb. 2003.

[21] Y. Saad, *Iterative methods for Sparse Linear Systems*, Boston: PWS Publishing Company, 1996.

[22] I. S. Duff, A. M. Erisman and J. K. Reid, *Direct Methods for Sparse Matrices*, Claerndon Press, Oxford, 1986.

[23] J. L. Volakis, A. Chatterjee, and L. C. Kempel, *Finite Element Method for Electromagnetics*, New York: IEEE Press and Oxford University Press, 1998.

[24] E. Cuthill and J. McKee, "Reducing the bandwidth of sparse symmetric matrices," *Proceedings of the 24th National Conference of the ACM*, New Jersey: Brandon Systems Press, 1969.

[25] MATLAB manuals, MATLAB is a product of Mathworks, Inc, U.S.A.

[26] A. George and J. W. H. Liu, *Computer Solution of Large Sparse Positive Definite Systems*, New Jersey: Prentice-Hall, 1981.

[27] R. Barrett, M. Berry, T. F. Chan, F. Demmel, J. M. Donato, J. Dongarra, V. Eijkhout, R. Pozo, C. Romine, and H. Van der Vorst, *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods*, Philadelphia: SIAM, 1994.

[28] C. Guo and T. H. Hubing, "Development and application of a fast multipole method in a hybrid FEM/MoM field solver", *Journal of the Applied Computational Electromagnetics Society*, vol. 19, no. 3, pp. 126-134, Nov. 2004.

[29] W. C. Chew, *Waves and Fields in Inhomogeous Media*, New York: IEEE Press, 1995.

Chunlie Guo earned his BSEE and MSEE degrees from Tsinghua University in 1998 and 2000, respectively. He is currently pursuing a Ph.D. degree in Electrical Engineering at the University of Missouri-Rolla. His graduate research is focused on the development of numerical modeling techniques for analyzing signal integrity and electromagnetic compatibility problems.

Todd Hubing received his BSEE degree from MIT in 1980, his MSEE degree from Purdue University in 1982, and his Ph.D. in Electrical Engineering from North Carolina State University in 1988.

From 1982 to 1989, he was an electromagnetic compatibility engineer for IBM in Research Triangle Park, NC. He is currently a Professor of Electrical and Computer Engineering at the University of Missouri-Rolla. He serves on the Board of Directors and is a Past President of the IEEE EMC Society.

A 2D Pseudo-Spectral Approach of Photonic Crystal Slabs

K. Varis[(a)], and A. R. Baghai-Wadji[(b)]
[(a)] Helsinki University of Technology, Optoelectronics Laboratory,
P.O. Box 3000, FIN-02015 HUT, Finland, EU
karri.varis@hut.fi
[(b)] Vienna University of Technology, E366,
Gusshausstrasse 27-29, A-1040 Vienna, Austria, EU
alireza.baghai-wadji@tuwien.ac.at

Abstract— **We consider an L_1-periodic dielectric slab which is characterized by the dielectric function $\varepsilon(x+L_1,z)=\varepsilon(x,z)$ as a 2D model for photonic crystals. We assume that there is no variation in $y-$direction, with fields varying time-harmonically according to $\exp(-j\omega t)$. In order to solve electromagnetic wave propagation in such structures, we diagonalize the Maxwell's equations with respect to the $z-$ coordinate. As demonstrated in this paper, diagonalized forms greatly facilitate the implementation of the finite difference method. The L_1-periodicity of the fields suggests expansions in terms of spatially harmonic functions. However, contrary to the commonly-used Bloch inhomogeneous plane waves, we utilize expansions of the form $\psi(x,z)=\sum_{n=0}^{N-1}\psi_n(z)\exp[j(k_n+K)x]$. For the determination of the coefficient functions $\psi_n(z)$ we employ a sophisticated, yet, easy-to-apply implementation of a finite difference discretization scheme in the $z-$direction which permits virtually arbitrary L_1-periodic $\varepsilon(x,z)$ profile functions. It will be demonstrated that the proposed hybridization of the plane-wave decomposition and the finite difference method leads to a robust and flexible method of analysis with a wide range of applications. As an example, we consider TE-polarized electromagnetic waves which propagate in the assumed dielectric slab along the $x-$axis.**

I. INTRODUCTION

We consider an L_1-periodic dielectric slab characterized by the permittivity function $\varepsilon(x+L_1,z)=\varepsilon(x,z)$ as a 2D model for photonic crystals. We assume that there is no variation in $y-$direction, with fields varying time-harmonically according to $\exp(-j\omega t)$. We present a general scheme for the diagonalization of Maxwell's equations with respect to the $z-$ coordinate and consider TE-polarized electromagnetic waves propagating in the assumed dielectric slab along the $x-$axis. The L_1-periodicity of the fields suggests expansions in terms of spatially harmonic functions. However, contrary to the commonly-used Bloch inhomogeneous plane waves, we utilize expansions of the form $\psi(x,z)=\sum_{n=0}^{N-1}\psi_n(z)\exp[j(k_n+K)x]$. For the determination of the coefficient functions $\psi_n(z)$ we employ a sophisticated, yet, easy-to-apply implementation of a finite difference discretization scheme in $z-$direction which permits virtually arbitrary permittivity profile functions $\varepsilon(x,z)$. It will be demonstrated that the proposed hybridization of the plane-wave decomposition and the finite difference method leads to a robust and flexible method of analysis with a wide range of applications.

Contrary to the standard finite difference implementations which include the entirety of the $\mathbf{E}-$ and $\mathbf{H}-$components, in our technique, we use an FD discretization, which only involves an ''optimized subset'' of the field components [1,2]. Stated more precisely, only those field components are involved in our formalism, which enter the interface- or boundary conditions on $z=$ const planes: It turns out that these ''transversal'' field components are the only unknowns in our problems; the remaining ''normal'' components can be uniquely determined once the transversal fields have been calculated.

This paper is organized as follows: In Section II. we briefly comment on the diagonalization procedure. The electromagnetic wave propagating in a photonic crystal, as specified above, decouples into a TE- and a TM-polarized mode. In this paper we focus on TE-modes. In Section III. we discuss the discretization and approximation of the fields. Section IV. is devoted to formulating appropriate boundary conditions for our problem. Thereby, assuming free space below and above our structure, we formulate discrete boundary conditions in a matrix form. Section V. is devoted to the specifics of our numerical calculations. We discuss tools and measures which we have developed to enhance the speed, and at the same time, the accuracy of our computations. In Section VI. we discuss a glimpse of the numerical results which we have obtained. We compare our results with available data. Section VII. concludes our discussion.

Notation: In the following we exploit the L_1-periodicity property and assume that the real-valued variable K (to be specified soon) varies in the interval $\left[-\pi/L_1,\pi/L_1\right]$. Furthermore, we have defined

$$k_n=\begin{cases}\frac{2\pi}{L_1}n & 0\le n\le\frac{N}{2}\\ \frac{2\pi}{L_1}(n-N) & \frac{N}{2}<n\le N-1\end{cases}. \quad (1)$$

II. Diagonalization

Consider the Maxwell's curl equations under the following two assumptions: (i) no variation in $y-$direction ($\partial y \equiv 0$), and (ii) isotropic materials specified by a constant permeability μ, and an L_1- periodic inhomogeneous permittivity function $\varepsilon(x,z)$: $\varepsilon(x+L_1,z) = \varepsilon(x,z)$. Assumptions (i) and (ii) imply the constitutive equations in the forms $\mathbf{B} = \mu\mathbf{H}$ and $\mathbf{D} = \varepsilon(x,z)\mathbf{E}$. Consequently, Maxwell's equations are:

$$-\partial_x \begin{bmatrix} 1 \\ 0 \end{bmatrix} H_3 + \partial_z \begin{pmatrix} H_1 \\ H_2 \end{pmatrix} = -j\omega\varepsilon(x,z) \begin{pmatrix} E_2 \\ -E_1 \end{pmatrix} \tag{2a}$$

$$-\partial_x \begin{bmatrix} 1 \\ 0 \end{bmatrix} E_3 + \partial_z \begin{pmatrix} E_1 \\ E_2 \end{pmatrix} = j\omega\mu \begin{pmatrix} H_2 \\ -H_1 \end{pmatrix} \tag{2b}$$

$$\partial_x H_2 = -j\omega\varepsilon(x,z)E_3 \tag{2c}$$

$$\partial_x E_2 = j\omega\mu H_3. \tag{2d}$$

Our goal is the diagonalization of (2) with respect to the $z-$axis. In other words, we are aiming at an equation of the form: $\mathscr{L}(x,z)\vec{\psi} = \partial_z\vec{\psi}$. Equations (2a) and (2b) can be written in the form:

$$\begin{pmatrix} 0 & 0 & 0 & j\omega\mu \\ 0 & 0 & -j\omega\mu & 0 \\ 0 & -j\omega\varepsilon(x,z) & 0 & 0 \\ j\omega\varepsilon(x,z) & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} E_1 \\ E_2 \\ H_1 \\ H_2 \end{pmatrix} + \begin{pmatrix} \partial_x & 0 \\ 0 & 0 \\ 0 & \partial_x \\ 0 & 0 \end{pmatrix} \begin{pmatrix} E_3 \\ H_3 \end{pmatrix} = \partial_z \begin{pmatrix} E_1 \\ E_2 \\ H_1 \\ H_2 \end{pmatrix}, \tag{3}$$

while (2c) and (2d) give:

$$\begin{pmatrix} E_3 \\ H_3 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & \frac{\partial_x}{j\omega\mu} \\ 0 & 0 \\ -\frac{\partial_x}{j\omega\varepsilon(x,z)} & 0 \end{pmatrix}^T \begin{pmatrix} E_1 \\ E_2 \\ H_1 \\ H_2 \end{pmatrix}. \tag{4}$$

Substituting (4) into (3) we obtain the desired diagonalized form, which can be written in the form $\mathscr{L}(x,z)\vec{\psi} = \partial_z\vec{\psi}$. Note that the $z-$dependence in $\mathscr{L}(x,z)$ is due to the $z-$dependence in the $\varepsilon(x,z)$ function. The differential $\mathscr{L}(x,z)-$operator is devoid of $z-$derivatives. The $z-$diagonalized form can be interpreted as follows: Evaluate $\mathscr{L}(x,z)$ at a certain point z, say, z_0, to obtain $\mathscr{L}(x,z_0)$. Determine the expressions for $\mathscr{L}(x,z_0)\vec{\psi}$. This gives the rate of change of $\vec{\psi}$ in the $z-$direction at $z = z_0$, i.e. $\partial_z\vec{\psi}$ at $z = z_0$. In the present case the system of equations $\mathscr{L}(x,z)\vec{\psi} = \partial_z\vec{\psi}$ decouples into the following subsets:

1) $z-$diagonalized transversal electric fields: It is straightforward to show that the $z-$diagonalized transversal electric fields satisfy equation (5),

$$\begin{pmatrix} 0 & -j\omega\mu \\ -j\omega\varepsilon(x,z) + \frac{1}{j\omega\mu}\partial_x^2 & 0 \end{pmatrix} \begin{pmatrix} E_2 \\ H_1 \end{pmatrix} = \partial_z \begin{pmatrix} E_2 \\ H_1 \end{pmatrix} \tag{5a}$$

$$H_3 = \frac{1}{j\omega\mu}\partial_x E_2. \tag{5b}$$

2) $z-$diagonalized transversal magnetic fields: The $z-$diagonalized equation for the transversal magnetic fields is

$$\begin{pmatrix} 0 & j\omega\mu - \partial_x\frac{1}{j\omega\varepsilon(x,z)}\partial_x \\ j\omega\varepsilon(x,z) & 0 \end{pmatrix} \begin{pmatrix} E_1 \\ H_2 \end{pmatrix} = \partial_z \begin{pmatrix} E_1 \\ H_2 \end{pmatrix} \tag{6a}$$

$$E_3 = -\frac{1}{j\omega\varepsilon(x,z)}\partial_x H_2. \tag{6b}$$

III. Analysis of TE-Polarized Waves in Periodic Dielectrics

In this paper we will focus on the TE-polarized waves, i.e. we consider (5), which we write in the form:

$$\left(-j\omega\varepsilon(x,z) + \frac{1}{j\omega\mu}\partial_x^2\right)E_2 = \partial_z H_1 \tag{7a}$$

$$-j\omega\mu H_1 = \partial_z E_2. \tag{7b}$$

A. Discretization of the Fields

As will be clear in the sequel we discretize the fields in $x-$ and $z-$directions differently. We exploit the periodicity conditions in $x-$direction and decompose the fields in spatial harmonics in this direction. However, we choose a finite difference discretization in $z-$direction. The following sections are devoted to the procedural details.

1) Treatment of the $x-$dependence: L_1-periodicity suggests the following approximation for the fields:

$$E_2(x,z) = \sum_{n=0}^{N-1} e_n(z)e^{j(k_n+K)x} \tag{8a}$$

$$H_1(x,z) = \sum_{n=0}^{N-1} h_n(z)e^{j(k_n+K)x}. \tag{8b}$$

For $z = z_0 = \text{const}$ the finite sums at the RHS of (8) are infinitely differentiable with respect to x. Substitute (8) into (7). Define the linear functional D_m

$$D_m(f) = \frac{1}{L_1}\int_0^{L_1} dx f(x,\cdot)e^{-j(k_m+K)x} \tag{9}$$

for $m \in [0, N-1]$. Apply D_m to both sides of (7). Use the fact that

$$\delta[m-n] = \frac{1}{L_1}\int_0^{L_1} dx e^{-j(k_m-k_n)x} \tag{10}$$

with $\delta[n]$ being the Kronecker symbol, and

$$\varepsilon_{m-n}(z) = \frac{1}{L_1}\int_0^{L_1} dx \varepsilon(x,z)e^{-j(k_m-k_n)x} \tag{11}$$

to obtain

$$\frac{j}{\omega\mu}(k_m+K)^2 e_m(z) - j\omega\sum_{n=0}^{N-1} e_n(z)\varepsilon_{m-n}(z) = \partial_z h_m(z) \tag{12a}$$

$$-j\omega\mu h_m(z) = \partial_z e_m(z). \tag{12b}$$

Consider the integral representation for $\varepsilon_{m-n}(z)$, Eq. (11), with $m,n \in [0, N-1]$. Using $k_m - k_n = \frac{2\pi}{L_1}(m-n)$ we obtain,

$$\varepsilon_{m-n}(z) = \frac{1}{L_1}\int_0^{L_1} dx \varepsilon(x,z)e^{-j\frac{2\pi}{L_1}(m-n)x}. \tag{13}$$

Observe that $\min(m-n) = 0-(N-1) = -N+1$ and $\max(m-n) = (N-1)-0 = N-1$. Therefore, for every fixed value of z we need to evaluate Fourier integrals of the form

$$\varepsilon_l(z) = \frac{1}{L_1}\int_0^{L_1} dx \varepsilon(x,z)e^{-j\frac{2\pi}{L_1}lx} \tag{14}$$

at $2N-1$ discrete ''frequency'' values in the range $l \in [-N+1, N-1]$. In our simulations we have evaluated integrals (14) by utilizing Fast Fourier Transform (FFT), and requiring $2N-1$ sampling points $\varepsilon(x_i, z_0)$; $i \in [1, 2N-1]$ of the function $\varepsilon(x, z_0)$ for every fix value z_0 of z.

B. Discretization in z−direction

As pointed out earlier we use a finite difference discretization in $z-$direction. However, in contrast to the standard techniques, our implementation of the finite difference technique involves Fourier coefficients rather than the fields in real space. In what follows we demonstrate the way how we discretize the $z-$ dependent part of the coefficients. To this end it is advantageous to adopt the abbreviation $f_m^i = f_m(i\Delta_z)$. Thereby, f_m^i means the m^{th} Fourier coefficient of the function $f(x,z)$ sampled at $z = i\Delta_z$. Using this notation we obtain:

$$\frac{j\Delta_z}{\omega\mu}(k_m+K)^2 e_m^{i-\frac{1}{2}} - j\omega\Delta_z\sum_{n=0}^{N-1} e_n^{i-\frac{1}{2}}\varepsilon_{m-n}^{i-\frac{1}{2}} + h_m^{i-1} = h_m^i \tag{15a}$$

$$-j\omega\mu\Delta_z h_m^i + e_m^{i-\frac{1}{2}} = e_m^{i+\frac{1}{2}}. \tag{15b}$$

These equations in the matrix form read:

$$\mathbf{A}^{i-\frac{1}{2}}\mathbf{e}^{i-\frac{1}{2}} + \mathbf{h}^{i-1} = \mathbf{h}^i \tag{16a}$$

$$\mathbf{B}^i\mathbf{h}^i + \mathbf{e}^{i-\frac{1}{2}} = \mathbf{e}^{i+\frac{1}{2}}. \tag{16b}$$

The coefficient matrices $\mathbf{A}$ and $\mathbf{B}$ in (16) have the following structure:

$$\mathbf{A} = -j\Delta_z\left\{\omega\mathbf{P} - \frac{1}{\omega\mu}(\mathbf{Q}+K\mathbf{I})^2\right\} \tag{17}$$

with

$$\mathbf{P} = \begin{bmatrix} \varepsilon_0 & \varepsilon_{-1} & \varepsilon_{-2} & \cdot & \cdot & \cdot & \varepsilon_{-N+1} \\ \varepsilon_1 & \varepsilon_0 & \varepsilon_{-1} & \cdot & \cdot & \cdot & \varepsilon_{-N+2} \\ \varepsilon_2 & \varepsilon_1 & \varepsilon_0 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \varepsilon_{N-2} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \varepsilon_{N-1} & \varepsilon_{N-2} & \cdot & \cdot & \cdot & \cdot & \varepsilon_0 \end{bmatrix} \tag{18}$$

and

$$\mathbf{Q} = \begin{bmatrix} k_0 & 0 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & k_1 & 0 & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & \cdot & \cdot & k_{N-1} \end{bmatrix} \tag{19}$$

and

$$\mathbf{B} = -j\omega\mu\Delta_z\mathbf{I}. \tag{20}$$

C. Boundedness property of $\mathbf{P}_\infty$

For analyzing the stability of the system of equations, which we will construct in the next section, it is

imperative to investigate the properties of the involved matrices. The finite dimensional matrix $\mathbf{P}(=\mathbf{P_N})$ in (18) is the truncated version (projection) of an infinite dimensional matrix that we denote by $\mathbf{P}_\infty$. It is recognized that we are concerned with Toeplitz matrices, which considered as kernels, have rich analytical properties. It is easily seen that $\mathbf{P}_\infty$ is uniquely determined by a two-sided infinite sequence $\{\varepsilon_l | l = 0, \pm 1, \pm 2, \cdots\}$ with $\mathbf{P}_\infty(\mathbf{m},\mathbf{n}) = \varepsilon_{\mathbf{m}-\mathbf{n}}$ $(m,n = 0,1,2,\cdots)$.

Furthermore, we understand that for any physically realizable dielectric medium the function $\varepsilon(x,z_i)$ is in $L^\infty(-\pi/2,\pi/2)$, the space of all essentially bounded functions $f(x)$ defined on $[-\pi/2,\pi/2]$, which are finite in the norm $||f||_\infty$. Here $||f||_\infty$ denotes the essential supremum of $|f(x)|$, with $x \in [-\pi/2,\pi/2]$.

In addition, we remember that we obtained ε_l as Fourier coefficients of $\varepsilon(x,z)$ (evaluated at a fixed point z_0) with respect to the functions $e_l(x) = \exp(j2\pi lx)$, $(l = 0,\pm 1,\pm 2,\cdots)$. Using standard theorems in the theory of Toeplitz matrices, we can show that based on these properties $\mathbf{P}_\infty$ $(m,n) = \varepsilon_{m-n}$, and thus $\mathbf{P_N}$ are bounded.

1) Creating the global system of equations: The fields along the $z-$ direction can be determined recursively by using (16) which we write in the form:

$$\mathbf{A}^{i-\frac{1}{2}}\mathbf{e}^{i-\frac{1}{2}} + \mathbf{h}^{i-1} - \mathbf{h}^i = 0 \tag{21a}$$

$$\mathbf{B}^i\mathbf{h}^i + \mathbf{e}^{i-\frac{1}{2}} - \mathbf{e}^{i+\frac{1}{2}} = 0. \tag{21b}$$

We terminate our computational grid at two bounding $z = \text{const}$ levels: The lower bound being $z = -\Delta_z/2$ and the upper bound being $z = (n_z + 1/2)\Delta_z$. Therefore, the simulation domain occupies the region $z \in [0,n_z]\Delta_z$. As is evident from (21) the magnetic fields are evaluated at even multiples of $\Delta_z/2$ while the electric fields are evaluated at odd multiples of $\Delta_z/2$. Consequently, we need to calculate the electric fields at $z = -\Delta_z/2$ and $z = (n_z + 1/2)\Delta_z$, i.e., $\exp(-\Delta_z/2)$ and $\exp((n_z + 1/2)\Delta_z)$. As will be shown in the next section these electric field values will be calculated from the field distributions in the adjacent media immediately below and above the corrugated slab. In view of (21), and with i running through the interval $[0,n_z]$, we can assemble the desired global system. The interlaced algorithm, which we have utilized for constructing the global system, is sketched in Fig. 1.

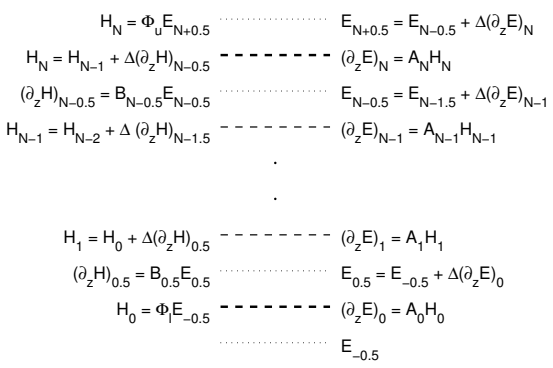


Fig. 1. A representation of how electric and magnetic fields are defined on interlaced layers. The text at the LHS should describe where and how the $h-$field has been computed, while the text on the RHS should provide the same information for the $e-$field. The slab is confined to the layers 0 and N. The $e-$fields outside the slab on the outermost layers are used to formulate the boundary conditions in terms of matrices Φ_u and Φ_l.

IV. Boundary Conditions

A. Open Boundary Problems

Assume free space for $z < 0$ and $z > n_z\Delta_z$. In free space our diagonalized equations take a particularly simple form: For the field expansion coefficients $e_n(z)$ and $h_n(z)$ appearing in (8), we can use $e_n\exp(\lambda_n z)$ and $h_n\exp(\lambda_n z)$, respectively. Thus we have:

$$E_2(x,z) = \sum_{n=0}^{N-1} e_n e^{j(k_n+K)x} e^{\lambda_n z} \tag{22a}$$

$$H_1(x,z) = \sum_{n=0}^{N-1} h_n e^{j(k_n+K)x} e^{\lambda_n z}. \tag{22b}$$

We next substitute (22) into (7), and apply the functional D_m to the terms involved. Noting that ε is a constant, we have $\varepsilon_{m-n}(z) = \delta[m-n]$. Therefore, we obtain:

$$\frac{j}{\omega\mu}(k_n+K)^2 e_n - j\omega\varepsilon e_n = \lambda_n h_n \tag{23a}$$

$$-j\omega\mu h_n = \lambda_n e_n. \tag{23b}$$

Solving for h_n from (23b), and substituting the result into (23a) we arrive at

$$\frac{j}{\omega\mu}(k_n+K)^2 e_n - j\omega\varepsilon e_n = -\frac{\lambda_n^2}{j\omega\mu} e_n. \tag{24}$$

For nontrivial solutions we obtain $(k_n+K)^2 - \omega^2\mu\varepsilon = \lambda_n^2$. Or, equivalently,

$$\lambda_n = \pm\sqrt{(k_n+K)^2 - \omega^2\mu\varepsilon} = \pm w_n. \tag{25}$$

As the next step we establish a relationship between $e^{-1/2}$ and h^0 to formulate the boundary condition we are looking for. We use the fact that $e_n^0 = e_n^{-1/2} e^{\lambda_n\Delta_z/2}$. In free space Sommerfeld's radiation condition permits

only those fields which decay at infinity. This condition is met by taking the branch $\lambda_n = w_n$, as defined in (25). Then, we obtain $e_n^0 = e_n^{-1/2}\, e^{w_n \Delta_z/2}$. From (23b) we obtain: $h_n^0 = -(w_n / j\omega\mu) e_n^0$. Using the last two equations we arrive at

$$h_n^0 + \frac{w_n}{j\omega\mu} e_n^{-\frac{1}{2}} e^{w_n \frac{1}{2}\Delta_z} = 0. \tag{26}$$

This equation defines the desired ''boundary'' matrix $\mathbf{C^{-1/2}}$ which relates $e^{-1/2}$ to h^0. Analogously we obtain the ''boundary'' matrix $\mathbf{C^{n_z+1/2}}$ which relates $h_n^{n_z}$ to $e_n^{n_z+1/2}$. Proceeding similarly we obtain

$$h_n^{n_z} - \frac{w_n}{j\omega\mu} e_n^{n_z+\frac{1}{2}} e^{\frac{1}{2}\Delta_z w_n} = 0 \tag{27}$$

which yields the desired boundary condition. Note that the resulting equality $\mathbf{C^{-1/2}} = \mathbf{C^{n_z+1/2}}$ is an implication of $\varepsilon^{-1/2} = \varepsilon^{n_z+1/2}$.

B. Inhomogeneous boundary problems

The scheme presented above is not restricted to homogeneous boundaries. Assume that the dielectric function outside the slab satisfies the following two conditions: for $z < 0$ or $z > h_z$ (i) $\partial_z \varepsilon(x,z) \equiv 0$, and (ii) $\varepsilon(x + L_1) = \varepsilon(x)$, where L_1 is the periodicity length in the slab. For the fields we have the expansions:

$$E_2(x,z) = \sum_{m=0}^{N-1} f_m \sum_{n=0}^{N-1} e_{m,n} e^{j(k_n+K)x} e^{\lambda_m z} \tag{28a}$$

$$H_1(x,z) = \sum_{m=0}^{N-1} f_m \sum_{n=0}^{N-1} h_{m,n} e^{j(k_n+K)x} e^{\lambda_m z}. \tag{28b}$$

The constituent terms in these equations are built from the eigenvectors and the corresponding eigenvalues λ_m (the propagation constants in $z-$direction). In order to compute the desired eigenpairs, we substitute

$$E_2(x,z) = \sum_{n=0}^{N-1} e_n e^{j(k_n+K)x} e^{\lambda z} \tag{29a}$$

$$H_1(x,z) = \sum_{n=0}^{N-1} h_n e^{j(k_n+K)x} e^{\lambda z} \tag{29b}$$

into (7) and process the LHS of the equation as described in section III.A. This leads to an algebraic eigenvalue equation of dimension $2N$ from which the desired eigenvalues and vectors can be solved numerically. Only eigenvalues which lead to decaying fields should be considered as explained above. This procedure is a generalization of the homogeneous boundary. In homogeneous case, the ''boundary'' matrices are diagonal because each eigenvector has only one non-zero element. In inhomogeneous case, the eigenvectors have in general N non-zero elements and the resulting boundary matrices are full.

V. Solving linear systems of the equations

In the preceding sections we explained the theoretical basis of our method. In the following we will describe how this theory can be applied to eigenmode and excitation problems. In addition we will explain how to solve the involved equation systems efficiently using iterative solvers.

A. Excitation problems

In this section we consider an elementary excitation problem, which can be formulated in terms of the following interface condition for the magnetic field:

$$\lim_{\delta\to 0} h_1\left(z^0 + \frac{\delta}{2}\right) - h_1\left(z^0 - \frac{\delta}{2}\right) = \rho_2(z^0). \tag{30}$$

Here $h_1(z)$ denotes the $x-$directional magnetic field component and ρ_2 represents a $y-$directional elementary current element.

In order to discretize (30), consider a three-point central difference scheme:

$$h_1(k\Delta) + \Delta \left\{ \frac{\partial h_1}{\partial z} \right\}\Bigg|_{z=\left(k+\frac{1}{2}\right)\Delta} = h_1\left((k+1)\Delta\right). \tag{31}$$

Next add one point to the system, at position $((k+1)\Delta - \delta)$ where δ is an infinitesimally small but finite number, and insert the dipole source ρ_2 at location $\left((k+1)\Delta - \frac{\delta}{2}\right)$. In view of (30) we can write (31) as

$$h_1(k\Delta) + \Delta \left\{ \frac{\partial h_1}{\partial z} \right\}\Bigg|_{z=\left(k+\frac{1}{2}\right)\Delta} + \rho_2 = h_1\left((k+1)\Delta\right). \tag{32}$$

As a generalization, we can substitute the $x-$directional Fourier expansion in place of the scalar variables above. Evaluating the derivative as in (12a) and writing the terms using the notation of (21) we obtain

$$h_k - h_{k+1} + B'_{k+\frac{1}{2}} e_{k+\frac{1}{2}} = -\rho_{k+\frac{1}{2}}. \tag{33}$$

Obviously the mere difference between this equation and (21b) is the excitation term at the RHS. In (33) the excitation has been indexed by $\left(k+\frac{1}{2}\right)$ due to the finite resolution in our discretized system: We cannot specify the position of the dipole source more precisely than stating that it is located somewhere between the layers k and $(k+1)$.

From the above discussion it can be concluded that for solving excitation problems we merely need to replace the zero vector at the RHS of (21b) by the Fourier transform of the current distribution. A similar procedure can be conducted *mutatis mutandis* for the determination of electric fields due to magnetic currents.

The discretized inhomogeneous system of equations describing the present excitation problem has the form

$$\mathbf{A}(\omega,\mathbf{K})\mathbf{x} = \mathbf{b}(\omega,\mathbf{K}) \tag{34}$$

where $\mathbf{x}$ represents the unknown coefficient vector and $\mathbf{b}$ is the excitation vector. The solution to this problem can be obtained by using standard techniques, e.g. LU-decomposition or Gauss elimination, unless matrix $\mathbf{A}$ is singular, in which case we recommend to resort to one of the techniques described below.

Once the solution to (34) has been obtained, we use (8) to construct the fields in real domain. As pointed out earlier, the expansions of the fields are simply Fourier transforms with respect to x on various layers $z = i\Delta$. Therefore, we can synthesis the fields from the coefficients merely by applying the inverse Fourier transform.

B. Eigenstate problems

1) Theorem: A homogeneous linear system

$$\mathbf{Ax} = \mathbf{0} \tag{35}$$

always has the trivial solution $\mathbf{x} = \mathbf{0}$. Nontrivial solution exist if and only if $rank(\mathbf{A}) < \mathbf{n}$. If $rank(\mathbf{A}) = \mathbf{r} < \mathbf{n}$, these solutions, together with $\mathbf{x} = \mathbf{0}$, form a vector space of dimension $n - r$.

One way to determine the rank of a matrix is to count all its eigenvalues which are equal to zero. The corresponding eigenvectors expand the null space of $\mathbf{A}$ among which we can construct all the solutions.

This information can also be obtained through singular value decomposition. If n singular values are zero, then the null space of $\mathbf{A}$ has dimension n and the equation system has a solution of degeneracy n.

A more efficient way to investigate the singularity of a matrix is to perform an LU-decomposition and calculate its determinant by multiplying the diagonal elements of the upper diagonal matrix. The determinant itself is generally inappropriate for determining the singularity of a matrix due to the lack of a uniform scale: The determinant can be very large even if the matrix is nearly singular or vice versa. In our case it does not really matter much since we are aiming at ratios. We construct the system matrix for consecutive ω values and compare the associated determinants. For ω's near a singularity, the value of the determinant drops sharply and we can iterate towards the resonant frequency. In the next session we address the details of the iterative scheme used.

C. Iterative solver

For large problems, iterative rather than direct solvers should be considered since they often significantly speed up the computations. Our choice for iterative solver has been the Transpose Free Quasi Minimal Residual method (TFQMR), which is a Krylov-subspace method for non-Hermitian matrices [3]. It is efficient, tolerant against breakdowns, and handles singular matrices well.

TFQMR (as many other iterative solvers) only requires products by the matrix to be solved and, thus, the matrix never actually has to be constructed. All is needed is a routine that returns the product of the matrix by a given trial vector.

1) Generating matrix products: The operator in (7a) has two parts: the spatial derivatives and a multiplication by a function ε in the spatial domain. Derivatives are trivially simple to compute in the Fourier domain as they reduce to algebraic multiplication by the respective Fourier expansion term. On the other hand, multiplication by ε leads to a convolution - or - in discretized version, to a multiplication by a convolution matrix (18). It is known from the theory of Fourier transforms that convolution in Fourier domain corresponds to a multiplication in real domain and vice versa. Therefore, the convolution can be evaluated by inverse Fourier transforming the coefficient matrix, multiplying by ε in the real domain, and finally Fourier transforming back. This approach is justifiable because multiplication by the (full) convolution matrix is an $O(N^3)$ operation for N coefficients, but in real domain, we multiply spatial fields with the corresponding ε, requiring only $O(N)$ operations. Dominating factor, $O(N\ln(N))$, comes from the FFT.

This approach can not be used in constructing the system matrix A, it is only amenable to evaluating matrix vector products.

2) Preconditioning: The convergence rate of iterative methods decisively depends on the matrix they are applied to. Occasionally, they may completely fail to converge. However, instead of solving $\mathbf{Ax} = \mathbf{b}$ we can solve the equivalent form

$$\mathbf{M_1^{-1}AM_2^{-1}}(\mathbf{M_2x}) = \mathbf{M_1^{-1}b} \tag{36}$$

for the new unknown vector $\mathbf{y} = \mathbf{M_2x}$ and the RHS $\mathbf{c} = \mathbf{M_1^{-1}b}$. Our expectation is that the solver may converge faster for the new matrix $\mathbf{M_1^{-1}AM_2^{-1}}$. The primary objective is then to find suitable preconditioner matrices $\mathbf{M_1}$ and $\mathbf{M_2}$ such that their inverse can be computed with a reasonable effort, and that they would transform the matrix into a nearly diagonal one. To this end various standard techniques have been developed, e.g. partial LU-decomposition, but we decided to use a problem-specific strategy in which relevant information from $\mathbf{A}$ is used. In our case we use only one sided preconditioning and set $\mathbf{M_1} = \mathbf{I}$. The right preconditioner matrix $\mathbf{M_2}$ is constructed from three diagonals of $\mathbf{A}$. In effect, $\mathbf{M_2}$ corresponds to a system matrix of a modified problem in which $\partial_x\varepsilon(x,z) \equiv 0$. The discretized dielectric function for the ''reduced''

problem can be obtained by averaging the original dielectric function on each $z-$layer, $\varepsilon_k = ave(\varepsilon_k)$, over one period.

We do not invert $\mathbf{M_2}$ explicitly. Instead, we solve $\mathbf{M_2 z} = \hat{\mathbf{y}} \Rightarrow \mathbf{z} = \mathbf{M_2^{-1}}\hat{\mathbf{y}}$, where $\hat{\mathbf{y}}$ is a trial vector given by TFQMR. After this, $\mathbf{z}$ is multiplied by $\mathbf{A}$ as explained above and the result is returned TFQMR. The linear system concerning $\mathbf{M_2}$ can be solved efficiently using Gauss elimination because there are non-zero elements only on three diagonals. The elements of $\mathbf{M_2}$ are computed once for each different $\mathbf{A}$ and then reused in the subsequent iterations.

3) Iterative solution of Solving eigenstates: Eigenstates can also be solved iteratively as will be demonstrated in section VI. This technique does not involve matrix factorization directly nor the computation of its determinant. Consider the following system

$$\mathbf{A}(\omega, \mathbf{K})\mathbf{x} = \mathbf{y} \tag{37}$$

where $\mathbf{y}$ is some non-zero vector, while the remaining terms are defined as before. This is an inhomogeneous system and can be solved using the iterative technique described above. If $\mathbf{A}(\omega, \mathbf{K})$ approaches a singular point, then $\mathbf{x}$ approaches an eigenvector of $\mathbf{A}(\omega, \mathbf{K})$ corresponding to an eigenvalue 0, nearly independently of $\mathbf{y}$. Furthermore, the norm of the solution $\mathbf{x}$ approaches infinity. The proof of this statement and limitations on the choice of $\mathbf{y}$ are provided in [4]. This property can be used as a measure for matrix singularity in place of, e.g. the determinant.

In many instances iterative solutions are preferred since they require less computer resources than the direct factorization of the matrix, and, at the same time, give the field solution.

VI. Numerical Results

We solve dispersion diagrams for two problems and compare results with those obtained by a planewave method (PWM) [5]. In addition, we present field solutions for a third problem, which consists of a slab with Gaussian dielectric profile function. We use TFQMR to find the field solution due to a single dipole excitation. Our objective is to find the singular points of the system matrix where the solution norm approaches infinity. Because it is easier to search zeros rather than infinities we use an ''inverted'' form instead. Then our objective function becomes $F(\omega) = \left(\left\|\mathbf{A^{-1}}(\omega)\mathbf{b}\right\|_2\right)^{-1/2}$ where $\mathbf{A}(\omega)$ is the system matrix and $\mathbf{b}$ is the excitation vector. The minimization takes place in two steps: first we bracket the minimum between two points and then we decrease the interval to the desired accuracy. When bracketing the minimum, we fit a second order polynomial to three points of $F(\omega)$ in order to estimate the location of a fourth point. This allows us to adapt the step size according to the derivative and absolute value of $F(\omega)$ and therefore take great leaps away from the minimum and small steps in its vicinity. When the minimum has been bracketed, we switch to golden section search in order to decrease the bracketed interval.

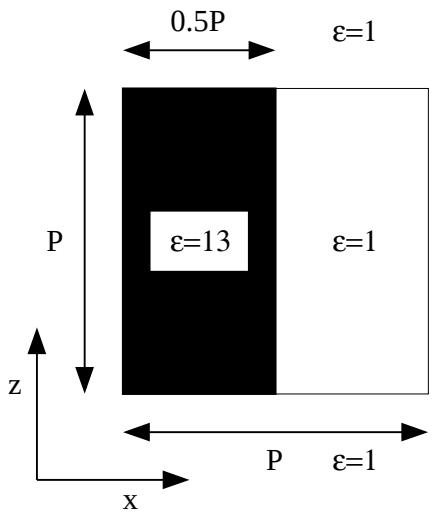


Fig. 2. One unit cell of the test structure 1. The structure is periodic in $x-$ direction with period P. Above and below the slab, as well as the space between the corrugation is free space.

In the planewave method, the problem is assumed to be periodic in all directions. Therefore, in order to apply this technique, we need to periodize the structure artificially by adding sufficiently large free space in the $z-$direction. The resulting enlarged unit cell, i.e. a supercell, is then periodized. This approach is justifiable if the modes are confined around the slab in $z-$direction, such that immediate neighboring supercells have negligible interaction. This can be verified by repetitively solving the problem with larger and larger supercell until the results converge.

A. A slab with rectangular corrugation

Our first test case is a slab with a periodic and rectangular corrugation. The period is $L_1 = P$, the height of the slab $h = P$ and the pitch-to-mark ratio is 0.5 ($l = 0.5P$). The relative dielectric constant of the slab is $\varepsilon = 13$. Above and below the slab, as well as, the space between the corrugation is free space. A schematic picture of the structure is presented in Fig. 2.

In order to apply the planewave method we have used a supercell with dimensions $L_1 = P$ and $L_z = 12P$. The computed dispersion diagram is shown in Fig. 3. Red curves with circular markers are obtained using our method, while blue lines with cross markers have been computed using the planewave method. The thick black line indicates the lightline, above which modes are radiating. Note that due to the artificial periodiza-

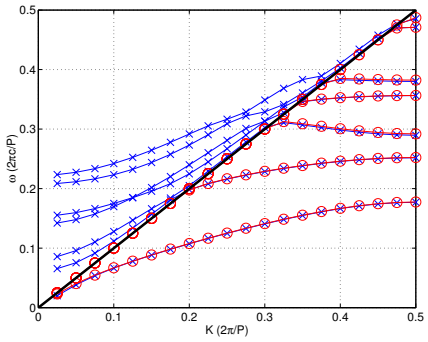


Fig. 3. Dispersion diagram for a slab with rectangular corrugation. Red lines with circular markers are computed using our method, while blue lines with cross markers have been computed using planewave method. In our method we used 64 planewaves in $x-$direction and 64 finite differences in $z-$direction. For the planewave method we used 64 planewaves in a unitcell; i.e. 64x(12x64) planewaves all together.

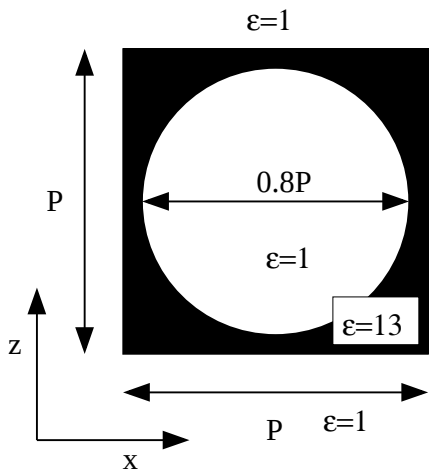


Fig. 4. One unit cell of the test structure 2. The structure is periodic in $x-$direction with period $L_1 = P$. Above and below the slab in $z-$direction, we assume free space.

tion in the planewave method, even unbounded modes appear guided.

This is an artifact, which is avoided in our method: For weakly or nonguided modes the interaction between consecutive supercells is no longer negligible and, therefore, they appear guided.

B. A slab with cylindrical corrugation

Our second test case is a similar slab but this time with a cylindrical corrugation. The slab dimensions, material, period length and discretization scheme are as above. However, the corrugation is formed of $y-$directional air cylinders (voids) with $r = 0.4P$. The structure is shown in Fig. 4.

Because the discretization is rectangular in both methods, we have averaged ε in the boundary cells in order to bring the average closer to its true value. The computed dispersion diagram is shown in Fig. 5.

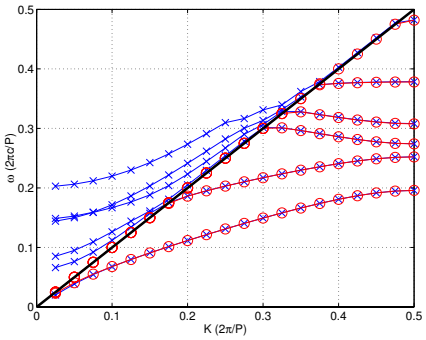


Fig. 5. Dispersion diagram for a slab with circular corrugation.

C. Field solution in a slab with Gaussian dielectric profile

Our third test case is a slab with thickness $h_z = 1$ and $x-$directional periodicity $L_1 = P = 1$. The dielectric function in the fundamental unit cell of the slab is

$$\varepsilon(x,z) = 1 + 9e^{-\left(\frac{0.5-x}{0.2}\right)^2} e^{\left(-\frac{0.5-z}{0.2}\right)^2}. \qquad (38)$$

Above and below the slab is free space. The dispersion diagram, which shows that this slab supports two TE-polarized modes, is presented in Fig. 6.

We solved the fields resulting from one y-directional dipole located at $(x,z) = (0.30, 0.27)$ for $(K,\omega) = (0.4, 0.2816)$. As can be seen from the dispersion diagram, this point corresponds to an eigenfrequency of the system and, therefore, the system matrix is (nearly)

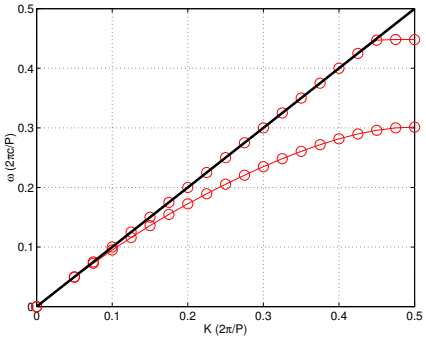


Fig. 6. Dispersion diagram for a slab with a Gaussian dielectric function as computed by our method.

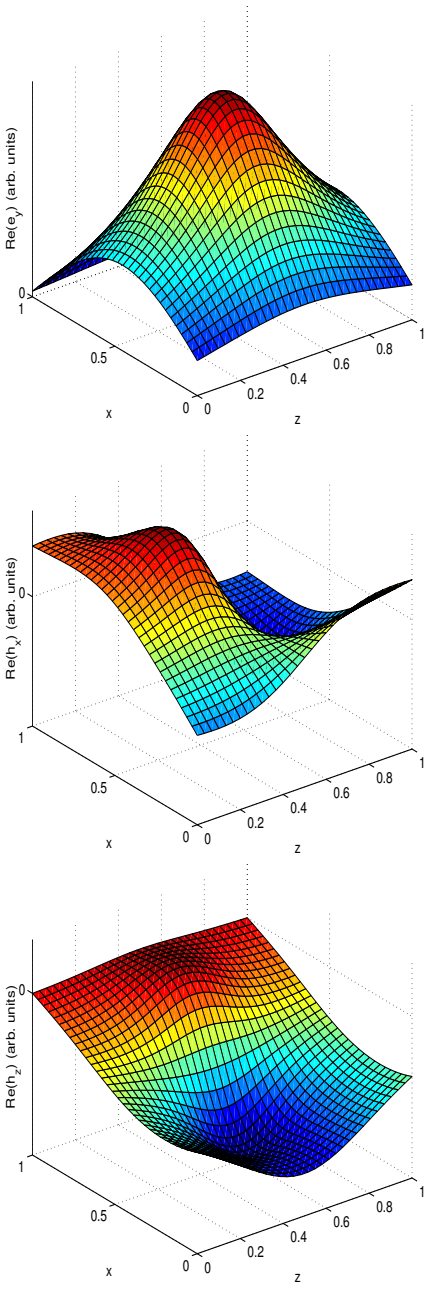


Fig. 7. Real part of eigenmode fi elds for a slab with Gaussian dielectric function, computed for $(K, \omega) = (0.4, 0.2816)$.

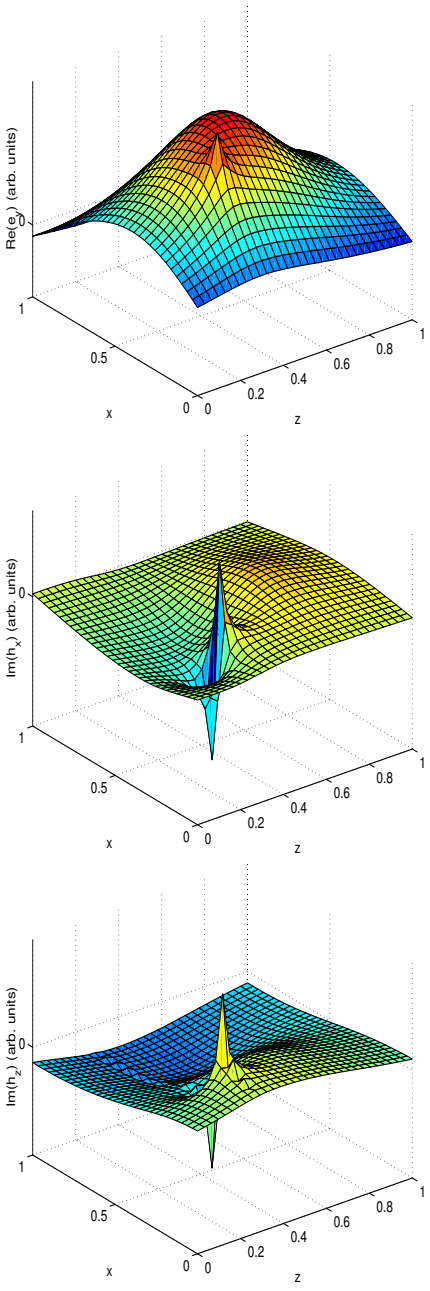


Fig. 8. Fields in a slab with Gaussian dielectric function, computed for $(K, \omega) = (0.5, 0.2816)$.

singular. Field distributions e_2 and h_1 presented in Fig. 7 are the transversal field components obtained directly by solving (34). Whereas the orthogonal field component h_3 is computed *a posteriori* using (5b). In the field patterns there is no trace of the dipole excitation. This can be understood as follows: A dipole excitation pumps energy into a resonating system and, therefore, the energy grows without bounds, until the excitation is negligible compared with the field strength. Mathematically speaking, the system matrix **A** has an eigenvalue 0 and, therefore, the solution of (34) is in the space which is spanned by eigenvectors corresponding to an eigenvalue 0. A more formal proof is given in [4].

Fields in response to an excitation in the same location but obtained for $(K,\omega) = (0.5, 0.2816)$ are shown in Fig. 8. Now the dipole excitation is clearly visible in the field patterns. Moreover, the ratio of field peak values at eigen frequency and this frequency, as returned by the iterative solver, is of order 10^8.

VII. Discussion of the Numerical Results and Conclusions

A. Searching for the minimum

The computational effort needed to solve the eigenmodes depends on how efficiently the system equation can be solved and how many times it has to be done. We have already addressed how accelerate the solver, here we discuss how minimize $F(\omega)$ efficiently for a given K. Currently we fit a second order polynomial to three previously computed points in order to estimate the location of the fourth one. When the curve approaches a minimum we decrease the step size trying to avoid overshooting the minimum. Practice has shown that it is good to aim at a point that changes the norm by 10% compared to the previous value. It is tempting to use bigger steps but then we risk jumping over a minimum without noticing its occurrence. This is especially crucial for closely spaced modes. Smaller steps on the other hand are more secure but then we end up solving the fields in unnecessarily many ω points. A typical search pattern is shown in Fig. 9.

When the minimum is bracketed, we switch to golden section search to iterate the minimum to the desired accuracy. The advantage of this procedure is that golden section search converges at a predetermided rate and it does not suffer from lock ups. The disadvantage is that the convergence rate is predetermined even though we could perform better since we have a good idea of the curve behavior. Most of the standard minimization techniques are not applicable because they assume parabolic behavior near the minimum. Instead, we could fit polynomials to both sides of the minimum and increase the degree as more points are solved. The crossing of the two polynomials would then be the next search point. However, this method is prone to lock ups

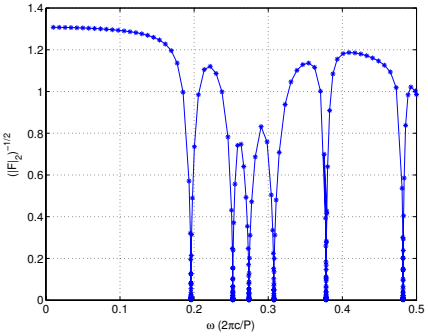


Fig. 9. Field energy function $F(\omega)$ for the test case with cylindrical corrugation. The phasing factor is chosen to be $K = 0.5$ (in Brillouin zone units). The behavior is very regular allowing effi cient optimization techniques in the quest for minima.

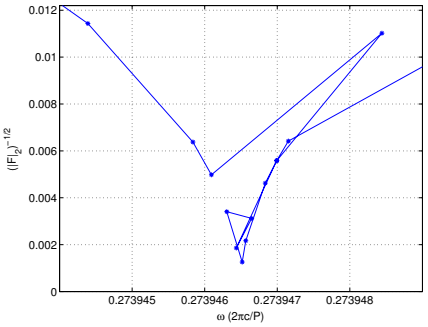


Fig. 10. Search pattern in the vicinity of the minimum for the third order mode with $K = 0.5$ (in Brillouin zone units) of the slab with cylindrical corrucation.

and has not been experimented so far. A typical search pattern in the vicinity of a minimum is shown in Fig. 10.

B. Convergence

1) Convergence of the iterative solver: It is difficult to make precise statements for the convergence behavior of the iterative solver because it strongly depends on a variety of parameters. With our preconditioner, the number of iterations needed usually varies between 2 and 20. Key factors are the dielectric function and the condition number of the matrix. In addition, the initial guess has an influence, even though a minor one.

Typically problems with $\partial\varepsilon(x,z)/\partial x$ being small converge fairly quickly. This is because large variations of ε in $x-$direction create large off-diagonal terms in the system equation. The three-diagonal preconditioner is capable of directly solving a matrix with only three diagonals but all off-diagonal terms are left for the it-

erative solver. Therefore, the bigger the off-diagonal terms are, the worse is the condition number of the preconditioned matrix and the more iterations are needed by the iterative solver. It is worth mentioning that if $\partial\varepsilon(x,z)/\partial x = 0$, there are no elements outside the three diagonals in the system matrix and only one application of the preconditioner solves the problem. No iterations are needed. It should also be pointed out that because of the finite difference approach in $z-$direction, $\partial\varepsilon(x,z)/\partial z$ has no effect on convergence. All variations with respect to z are eliminated by the preconditioner.

It appears that the condition number of the matrix is also a factor in the iterative solver convergence. This is unfortunate since the eigenmodes of the system are found exactly in the singular points of the system matrix. The effect is not dramatic; the number of the iterations required for convergence is maybe four fold as compared to a well behaved point, all other factors being held equal. Again, it is rather difficult to make precise statements because other factors often have a more significant effect. As an example, in the cylindrical-void slab problem with $n_x = n_z = 64$, convergence at $K = 0.5$, $\omega = 0.1$ requires three iterations (TFQMR requires two matrix products for each iteration) and at the lowest order eigenmode - $K = 0.5$, $\omega = 0.195719$ - 12 iterations are needed.

The number of discretizations, on the other hand, does not have a direct effect on convergence. Solving the above mentioned problem with $n_x = n_z = 384$, requires three and eleven iterations, respectively.

2) Convergence of eigenfrequencies: Our method gives precise results with a small number of discretization steps. As an example, we analyze the slab problem with cylindrical corrugation. We have solved the lowest order eigenfrequency for $K = 0.5$ with both, our method, and the planewave method, increasing discretization until the results converged. The convergence behavior is shown in Fig. 11.

In conclusion we summarize the distinct properties of our method: The difference between converged frequencies for $n = 16$ and $n = 384$ is only 0.024%. The relative difference between converged frequencies as solved with our method and the planewave method for $n = 384$ is 0.00204%.

VIII. Acknowledgements

This work was partially done while the second author (A. R. Baghai-Wadji) was visiting the Institute for Mathematical Sciences, National University of Singapore and Institute for High Performance Computing (IHPC) in 2003. This visit was supported by the Institute and IHPC.

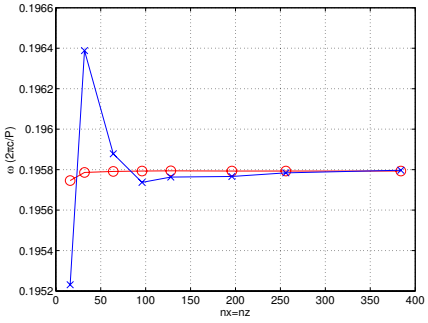


Fig. 11. Shows the lowest order eigenfrequency of $K = 0.5$ point for the slab with cylindrical corrugation as a function of the number of discretization points. Red curve with circular markings is the frequency as solved by our method, and blue with cross markings obtained by the planewave method. $x-$axis represents the number of discretization steps within one unit cell (same in both $x-$ and $z-$directions). For the planewave method, the number of planewaves in $z-$direction is $12n_z$ compensating the size of the supercell. Therefore both methods have equally many discretization points within the unit cell area.

References

[1] K. Varis, A.R. Baghai–Wadji, ''Hybrid Planewave/Finite-Difference Transfer Method for Solving Photonic Crystals in Finite Thickness Slabs,'' Proceedings EDMO, Electron Devices for Microwave and Optoelectronic Applications, pp. 161-166, Vienna University of Technology, Vienna, Austria, 15-16 November, 2001.

[2] K. Varis, A.R. Baghai–Wadji, ''A Novel 3D Pseudo-spectral Analysis of Photonic Crystal Slabs,'' ACES journal, **19**, no. 1 b, pp. 101–111, 2004.

[3] R. W. Freund, N. M. Nachtigal, ''A Transpose-Free Quasi-Minimal Residual Algorithm for Non-Hermitian Linear Systems,'' SIAM Journal on Scientifi c Computing, **14**, pp. 470-482, 1993.

[4] W. H. Press, S. A. Teukolsky, W. T. Wetterling, B. P. Flannery, ''Numerical recipes in C: the Art of Scientifi c Computing, 2nd Ed.,'' Cambridge University Press, Cambridge, pp. 493-495, 1992.

[5] S. G. Johnson, J. D. Joannopoulos, ''Block-Iterative Frequency-domain Methods for Maxwell's Equations in a Planewave Basis,'' Optics Express, **8**, no. 3, pp. 173-190, 2001.

Karri Varis was born in Espoo, Finland, in 1974. He received a M.Sc. degree in Electrical Engineering from Helsinki University of Technology in 1999. Since then, he has been a doctoral student in the Finnish Graduate School of Modern Optics and Photonics and working in the Optoelectronics Laboratory, Helsinki University of Technology. His current research interest includes the enhancement and development of computational techniques for the analysis of periodic and non-periodic systems, with an emphasis on photonic crystals and optical devices.

Ali R. Baghai-Wadji was born in Marand, Iran, on May 6, 1953. He has been with the Electrical Engineering and Information Technology Department at Vienna University of Technology (VUT), Vienna, Austria, since 1979. From 1979 to 1984 he was an associate researcher in Physical Electronics and Applied Electronics Groups, where he developed computer models for microelectronic and microacoustic devices. He earned his M.Sc. and Ph.D. in electrical engineering in 1984 and 1987, respectively, and obtained his *venia docendi* in physical electronics in 1994, all from VUT. Since 1997 he has been an associate professor at VUT. Currently he is heading the Accelerated Computational Technology (ACT) Group at the Institute for Fundamental and Theory of Electrical Engineering at VUT. Three times he was awarded the Kurt Gödel research scholarship from Austria, allowing him to spend a total of 10 months at UCI, University of California, Irvine, during the years 1990, 1991, and 1992. From 1994 to 1999 he was, on leave of absence from VUT, a principal engineer consultant in the United States serving more than four years for Motorola, Government System and Technology Group, Scottsdale, Arizona, and nearly one year for CTS-Wireless components in Albuquerque, New Mexico. In 1999 and 2000 he was (15 months) a visiting professor at Materials Science Laboratory, Helsinki University of Technology (HUT). That professorship was awarded by Nokia Research Foundation, and TEKES, a national science foundation in Finland. In addition, in the Fall 2000 he was awarded a Nokia Visiting Fellowship. In 2003 he was awarded a *venia docendi* for modeling and simulation of classical and quantum electronic devices and materials at HUT for an initial period of 6 six years. In 2003 he was four months an invited senior member of the Institute for Mathematical Sciences, and a visiting professor at the Institute for High Performance Computing, in Singapore. Since 1995 he has also been affiliated with Arizona State University as an adjunct professor at the Department for Mathematics and Statistics. In 2002 he was elected an honorary member of the Electromagnetics Academy, Massachusetts, USA. He has supervised five PhD dissertations, and eleven Masters' theses. He has lectured 12 short courses at various IEEE conferences internationally. He has authored more than 120 publications in reviewed journals and conference proceedings, and has one patent. His current research interest includes the development of accelerated computational modeling techniques, quantum mechanics, photonic crystals, and molecular electronics. Since 1995 he has been a senior member of the IEEE, an IEEE-UFFC associate editor, and an IEEE-UFFC technical program committee member. He was the guest editor for a special issue on *Modeling, Optimization, and Design of Surface and Bulk Acoustic Wave Devices* in IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control (Sept. 2001, Vol. 48, Num. 5). He has been appointed the general chairman of the PIERS'5 conference.

The behavior of smart obstacles in electromagnetic scattering: mathematical models as optimal control problems

Lorella Fatone[1], Maria Cristina Recchioni[2], Adina Scoccia[2], Francesco Zirilli[3]

[1] Dipartimento di Matematica Pura e Applicata, University of Modena e Reggio Emilia, Modena, Italy

[2] Dipartimento di Scienze Sociali "D. Serrani", University Politecnica delle Marche, Ancona, Italy

[3] Dipartimento di Matematica "G. Castelnuovo", University of Roma "La Sapienza", Roma, Italy

Abstract

We consider a bounded obstacle characterized by a boundary electromagnetic impedance contained in the three dimensional real Euclidean space filled with a homogeneous isotropic medium. When an incoming electromagnetic field illuminates the obstacle a scattered field is generated. A smart obstacle is an obstacle that in the scattering process, circulating a surface electric current density on its boundary, tries to achieve a given goal. We consider four possible goals: making the obstacle undetectable (i.e.: furtivity problem), making the obstacle to appear with a shape and impedance different from its actual ones (i.e.: masking problem), making the obstacle to appear in a location different from its actual one eventually with a shape and impedance different from its actual ones (i.e.: ghost obstacle problem) and finally one of the previous goals limited to a given subset of the frequency space (i.e.: definite band problems). We consider the problem of determining the optimal electric current density to achieve the given goal. The relevance in many application fields (i.e. stealth technology, electromagnetic noise control, etc.) of these problems is well known. The previous problems are modelled as optimal control problems for the Maxwell equations. Some numerical results on test problems obtained solving the optimal control problems proposed are shown.

1. Introduction

In recent years the development of new technologies has made possible to build a vast class of "smart" objects. This wave of innovation has moved from cutting edge military applications to everyday life objects such as, for example, washing machines. In this paper we consider the problem of formulating adequate mathematical models of smart obstacles in the context of electromagnetic scattering. The general mathematical model that we have in mind to describe the behavior of a "smart" object is an optimal control problem. The problems considered in electromagnetic obstacle scattering are described by partial differential equations so that we deal with optimal control problems for partial differential equations. Optimal control problems are widely used in engineering as mathematical models. However their

use is mainly limited to the control of systems governed by ordinary differential equations and their use in electromagnetic scattering is rather uncommon. The development of computer technology and numerical methods occurred in the last decades makes possible now to consider optimal control problems for systems of partial differential equations such as the Maxwell equations, that is makes possible the use of optimal control to model electromagnetic scattering problems. We consider four examples: furtivity problem (i.e.: the obstacle wants to be undetectable), masking problem (i.e.: the obstacle wants to appear with a shape different from its actual shape eventually with a boundary impedance different from its actual one), ghost obstacle problem (i.e.: the obstacle wants to appear in a location different from its actual location eventually with a shape and impedance different from its actual ones) and finally definite band problems (i.e.: the obstacle pursues one of the previous goals on a given subset of the frequency space). Recently similar problems in the context of time dependent acoustic and electromagnetic obstacle scattering have been studied from the point of view of formulating adequate mathematical models and of developing highly parallelizable numerical methods to solve them (see [1], [2], [3], [4], [5] and the websites: *http://www.econ.univpm.it/recchioni/w6*, *http://www.econ.univpm.it/recchioni/w8*, *http://www.econ.univpm.it/recchioni/w9*, *http://www.econ.univpm.it/recchioni/w10*, *http://www.econ.univpm.it/recchioni/w11*). Note that in these papers "smart" and "active" obstacles are synonyms. More in detail in [1] the furtivity problem in acoustic time dependent obstacle scattering has been modelled as an optimal control problem and the first order optimality conditions to solve it have been obtained as a system of coupled partial differential equations, finally a highly parallelizable numerical solver for this system of partial differential equations has been developed. Later in [2], in [4] and in [5] the masking problem in acoustics and the furtivity and the masking problems in electromagnetics have been studied and finally in [3] the definite band ghost obstacle problem in acoustics has been solved.

The practical interest of the mathematical models of smart obstacles proposed consists in the fact that these models can be used to design smart obstacles of practical value. Hence, for example, in the realization of radar absorbers the approach proposed can be a way of approaching the design of phase-switched screens (see for example [6], [7], [8]). In fact the phase-switched screen is an object that does not absorb the incident energy but shifts it in frequency using phase modulation so that the reflected energy falls outside the receiver bandwidth. That is, a phase-switched screen in our language can be seen as a smart obstacle that pursues the goal of being furtive in a given subset of the frequency space.

In Section 2 we formulate the mathematical models of the electromagnetic smart obstacles considered. In Section 3 we show some numerical results obtained solving the model proposed in Section 2 concerning the definite band furtivity problem.

2. Mathematical models of electromagnetic smart obstacles

Let us begin introducing some notations. Let $\mathbf{R}$ be the set of real numbers, $\mathbf{R}^3$ be the three dimensional real Euclidean space and $\mathbf{x} = (x_1, x_2, x_3)^T \in \mathbf{R}^3$ be a generic vector, where the superscript T means transposed. We denote with $(\cdot,\cdot)$ the Euclidean scalar product in $\mathbf{R}^3$, with $\|\cdot\|$ the corresponding Euclidean vector norm and with $[\cdot,\cdot]$ the usual vector product. Let $\mathbf{R}^3$ be filled with a homogeneous isotropic medium of constant electric permittivity $\epsilon > 0$, constant magnetic permeability $\upsilon > 0$ and zero electric conductivity. Moreover we assume that there are no free electric charges or currents. Let us suppose that $\mathbf{R}^3$ contains an obstacle Ω given by a bounded set without holes and internal cavities, more technically, a bounded simply connected open set Ω, with locally Lipschitz boundary $\partial\Omega$. Let $\overline{\Omega}$ denote the set

$\Omega \cup \partial\Omega$ and $\mathbf{n}(\mathbf{x}) = (n_1(\mathbf{x}), n_2(\mathbf{x}), n_3(\mathbf{x}))^T \in \mathbf{R}^3$, $\mathbf{x} \in \partial\Omega$ be the outward unit normal vector to $\partial\Omega$ in $\mathbf{x} \in \partial\Omega$. In the following Ω will be the scatterer, that is the obstacle responsible for the scattering of the incoming electromagnetic field. We assume that Ω has a known constant real boundary electromagnetic impedance $\chi \geq 0$. The limit case of perfectly insulating obstacles (i.e. $\chi = +\infty$) can be treated with straightforward modifications of the material presented here.

We begin modelling the standard direct obstacle scattering problem that is, the scattering problem relative to an obstacle that does not pursue any goal. We refer to this obstacle as a "passive" obstacle.

We consider an incoming electromagnetic field $(\boldsymbol{E}^i(\mathbf{x},t), \boldsymbol{B}^i(\mathbf{x},t))$, $(\mathbf{x},t) \in \mathbf{R}^3 \times \mathbf{R}$. The electric vector field $\boldsymbol{E}^i(\mathbf{x},t) \in \mathbf{R}^3$, $(\mathbf{x},t) \in \mathbf{R}^3 \times \mathbf{R}$ and the magnetic induction vector field $\boldsymbol{B}^i(\mathbf{x},t) \in \mathbf{R}^3$, $(\mathbf{x},t) \in \mathbf{R}^3 \times \mathbf{R}$ associated to the incoming electromagnetic field satisfy the Maxwell equations, that is equations (1), (2), for $(\mathbf{x},t) \in \mathbf{R}^3 \times \mathbf{R}$. We use the M.K.S. unit system to write equations (1), (2) (see [9], p. 16). When the incoming electromagnetic field $(\boldsymbol{E}^i(\mathbf{x},t), \boldsymbol{B}^i(\mathbf{x},t))$, $(\mathbf{x},t) \in \mathbf{R}^3 \times \mathbf{R}$, hits the scatterer Ω generates a scattered electromagnetic field $(\boldsymbol{E}^s(\mathbf{x},t), \boldsymbol{B}^s(\mathbf{x},t))$, $(\mathbf{x},t) \in (\mathbf{R}^3 \setminus \overline{\Omega}) \times \mathbf{R}$, solution of an exterior problem for the Maxwell equations. That is the scattered electric vector field $\boldsymbol{E}^s(\mathbf{x},t) \in \mathbf{R}^3$, $(\mathbf{x},t) \in (\mathbf{R}^3 \setminus \overline{\Omega}) \times \mathbf{R}$ and the scattered magnetic induction vector field $\boldsymbol{B}^s(\mathbf{x},t) \in \mathbf{R}^3$, $(\mathbf{x},t) \in (\mathbf{R}^3 \setminus \overline{\Omega}) \times \mathbf{R}$ satisfy the following equations,

$$\left(\operatorname{curl} \boldsymbol{E}^s + \frac{\partial \boldsymbol{B}^s}{\partial t}\right)(\mathbf{x},t) = \mathbf{0},$$
$$\left(\operatorname{curl} \boldsymbol{B}^s - \frac{1}{c^2}\frac{\partial \boldsymbol{E}^s}{\partial t}\right)(\mathbf{x},t) = \mathbf{0},$$
$$(\mathbf{x},t) \in (\mathbf{R}^3 \setminus \overline{\Omega}) \times \mathbf{R}, \quad (1)$$

$$\operatorname{div} \boldsymbol{B}^s(\mathbf{x},t) = 0, \operatorname{div} \boldsymbol{E}^s(\mathbf{x},t) = 0,$$
$$(\mathbf{x},t) \in (\mathbf{R}^3 \setminus \overline{\Omega}) \times \mathbf{R}, \quad (2)$$

with the boundary condition,

$$[\mathbf{n}(\mathbf{x}), \boldsymbol{E}^s(\mathbf{x},t)] - c\chi\,[\mathbf{n}(\mathbf{x}), [\mathbf{n}(\mathbf{x}), \boldsymbol{B}^s(\mathbf{x},t)]] =$$
$$[\mathbf{n}(\mathbf{x}), \boldsymbol{b}(\mathbf{x},t)]\,, (\mathbf{x},t) \in \partial\Omega \times \mathbf{R}, \quad (3)$$

where,

$$\boldsymbol{b}(\mathbf{x},t) = -\boldsymbol{E}^i(\mathbf{x},t) + c\chi\left[\mathbf{n}(\mathbf{x}), \boldsymbol{B}^i(\mathbf{x},t)\right],$$
$$(\mathbf{x},t) \in \partial\Omega \times \mathbf{R}, \quad (4)$$

the condition at infinity and the radiation condition given respectively by,

$$[\boldsymbol{B}^s(\mathbf{x},t), \hat{\mathbf{x}}] - \frac{1}{c}\boldsymbol{E}^s(\mathbf{x},t) = o\left(\frac{1}{r}\right),$$
$$\boldsymbol{E}^s(\mathbf{x},t) = O\left(\frac{1}{r}\right),\ r \to +\infty,\ t \in \mathbf{R}, \quad (5)$$

where $\mathbf{0} = (0,0,0)^T$, $c = 1/\sqrt{\epsilon\upsilon}$, $\hat{\mathbf{x}} = \mathbf{x}/\|\mathbf{x}\|$, $\mathbf{x} \neq \mathbf{0}$, $r = \|\mathbf{x}\|$, curl $\cdot$ and div$\cdot$ denote respectively the curl and the divergence of $\cdot$ with respect to the $\mathbf{x}$ variable, $\partial \cdot /\partial t$ denotes the time derivative of $\cdot$, and $o(\cdot)$, $O(\cdot)$ are the Landau symbols. When we consider the case $\chi = +\infty$ the boundary condition (3) must be "rewritten". The two conditions contained in (5) imply the vanishing of the magnetic induction vector field at infinity, that is $\boldsymbol{B}^s(\mathbf{x},t) = O(1/r)$, $r \to +\infty$, $t \in \mathbf{R}$. Moreover we assume that the incoming electromagnetic field vanishes when $t \to -\infty$, that is $\boldsymbol{E}^i(\mathbf{x},t)$, $\boldsymbol{B}^i(\mathbf{x},t) \to \mathbf{0}$, $\mathbf{x} \in \mathbf{R}^3$, $t \to -\infty$, that implies that the scattered electromagnetic field vanishes when $t \to -\infty$ as well, that is $\boldsymbol{E}^s(\mathbf{x},t)$, $\boldsymbol{B}^s(\mathbf{x},t) \to \mathbf{0}$, $\mathbf{x} \in (\mathbf{R}^3 \setminus \overline{\Omega})$, $t \to -\infty$.

The scattering problem for a "passive" obstacle Ω can be stated as follows:

Scattering Problem (passive obstacle). Given the incident electromagnetic field $(\boldsymbol{E}^i, \boldsymbol{B}^i)$, the obstacle Ω and its boundary electromagnetic impedance χ, solve the time dependent Maxwell equations (1)-(3), (5) in the unknowns $(\boldsymbol{E}^s, \boldsymbol{B}^s)$.

Let us study the possibility of transforming the "passive" obstacle into a "smart" obstacle.

Problem 1. Furtivity Problem. Given the incident electromagnetic field $(\boldsymbol{E}^i, \boldsymbol{B}^i)$, the obstacle Ω and its boundary electromagnetic impedance χ choose a control vector field (i.e. a surface electric current density) defined for $(\mathbf{x},t) \in \partial\Omega \times \mathbf{R}$ in a suitable class of admissible controls, in order to

minimize a cost functional that roughly speaking measures the "magnitude" of the electromagnetic field $(\boldsymbol{E}^s, \boldsymbol{B}^s)$ scattered by Ω, χ (when the control vector field is active) when hit by the incoming field $(\boldsymbol{E}^i, \boldsymbol{B}^i)$ and the "magnitude" of the control vector field employed.

To obtain a satisfactory formulation of the furtivity problem we modify the boundary condition (3) as follows,

$$[\mathbf{n}(\mathbf{x}), \boldsymbol{E}^s(\mathbf{x},t)] - c\chi\,[\mathbf{n}(\mathbf{x}), [\mathbf{n}(\mathbf{x}), \boldsymbol{B}^s(\mathbf{x},t)]] = [\mathbf{n}(\mathbf{x}), \boldsymbol{b}(\mathbf{x},t)] + (1+\chi)[\mathbf{n}(\mathbf{x}), \boldsymbol{\Phi}(\mathbf{x},t)], \quad (\mathbf{x},t) \in \partial\Omega \times \mathbf{R}. \tag{6}$$

The quantity $\boldsymbol{\Phi}(\mathbf{x},t)$, $(\mathbf{x},t) \in \partial\Omega \times \mathbf{R}$ has the dimension of an electric field and is related to the control variable that transforms the obstacle Ω from being passive to being smart. We assume that $\lim_{t\to\pm\infty} \boldsymbol{\Phi}(\mathbf{x},t) = \mathbf{0}$, $\mathbf{x} \in \partial\Omega$.

Let us define $\boldsymbol{\psi}(\mathbf{x},t) = \frac{\partial \boldsymbol{\Phi}}{\partial t}(\mathbf{x},t)$, $(\mathbf{x},t) \in \partial\Omega \times \mathbf{R}$ and let V be the space of the admissible controls, that we leave undetermined in this paper (see [5]) for a definition of V). Note that $\boldsymbol{\psi} = \frac{\partial \boldsymbol{\Phi}}{\partial t}$ has the dimensions of an electric (surface) current density. The furtivity problem can be formulated as the following optimal control problem,

$$\min_{\boldsymbol{\psi} \in V} \mathcal{F}_{\lambda,\mu}(\boldsymbol{\psi}), \tag{7}$$

subject to the constraints (1), (2), (5), (6) and $\mathcal{F}_{\lambda,\mu}$ is the following functional,

$$\mathcal{F}_{\lambda,\mu}(\boldsymbol{\psi}) = (1+\chi)\left\{\lambda |||\,[\mathbf{n}, \boldsymbol{E}^s]\,|||^2 + \lambda c^2 |||\,[\mathbf{n}, \boldsymbol{B}^s]\,|||^2 + \mu\varsigma |||\,[\mathbf{n}, \boldsymbol{\psi}]\,|||^2\right\}. \tag{8}$$

The quantity ς is a positive dimensional constant and $\lambda \geq 0$, $\mu \geq 0$ are adimensional constants such that $\lambda + \mu = 1$. Moreover the norms $|||\cdot|||$ appearing in (8) are norms on a suitable space of functions defined on $\partial\Omega \times \mathbf{R}$ (see [5]). For example the square root of the integral over $\partial\Omega \times \mathbf{R}$ of the square of the vector norm of $\cdot$ is such a norm. Note that the solution of problem (7), (1), (2), (5), (6) when $\lambda = 0$, $\mu = 1$ is $[\mathbf{n}(\mathbf{x}), \boldsymbol{\psi}(\mathbf{x},t)] = \mathbf{0}$, $(\mathbf{x},t) \in \partial\Omega \times \mathbf{R}$, that is in this case $(\boldsymbol{E}^s, \boldsymbol{B}^s)$ is the electromagnetic field scattered by the passive obstacle. On the other hand when $\lambda = 1$, $\mu = 0$ the solution of the same problem gives an obstacle completely undetectable since the minimization of (8) in this case gives $[\mathbf{n}(\mathbf{x}), \boldsymbol{E}^s(\mathbf{x},t)] = \mathbf{0}$, $[\mathbf{n}(\mathbf{x}), \boldsymbol{B}^s(\mathbf{x},t)] = \mathbf{0}$, $(\mathbf{x},t) \in \partial\Omega \times \mathbf{R}$ that implies $\boldsymbol{E}^s(\mathbf{x},t) = \mathbf{0}$, and $\boldsymbol{B}^s(\mathbf{x},t) = \mathbf{0}$, $(\mathbf{x},t) \in (\mathbf{R} \setminus \overline{\Omega}) \times \mathbf{R}$. However when $\lambda = 1$, $\mu = 0$ the cost functional (8) does not contain a term that depends on the control employed $\boldsymbol{\psi}$. Note that $\boldsymbol{E}^s$, $\boldsymbol{B}^s$ depends implicitly on $\boldsymbol{\psi}$ through the boundary condition (6). The remaining cases, that is $0 < \lambda < 1$, correspond to nontrivial formulations of the furtivity problem.

Problem 2. Masking Problem. In the circumstances of Problem 1 given an obstacle $D \subseteq \Omega$, and its electromagnetic boundary impedance χ_D, choose a control vector field $\boldsymbol{\psi}(\mathbf{x},t)$, $(\mathbf{x},t) \in \partial\Omega \times \mathbf{R}$ in a suitable class of admissible controls, in order to minimize a cost functional that roughly speaking measures the "magnitude of the difference" between the electromagnetic field $(\boldsymbol{E}^s, \boldsymbol{B}^s)$ scattered by Ω, χ (when the control vector field is active) and the electromagnetic field $(\boldsymbol{E}^s_D, \boldsymbol{B}^s_D)$ scattered by D, χ_D when hit by the incoming field $(\boldsymbol{E}^i, \boldsymbol{B}^i)$ and the "magnitude" of the control vector field employed. The couple D, χ_D will be called the "mask". For simplicity we assume the mask to be a passive obstacle.

The Masking Problem can be modelled as the optimal control problem (7), (1), (2), (5), (6) if the functional $\mathcal{F}_{\lambda,\mu}$ that appears in (7) is defined as follows,

$$\mathcal{F}_{\lambda,\mu}(\boldsymbol{\psi}) = (1+\chi)\left\{\lambda |||\,[\mathbf{n}, \boldsymbol{E}^s - \boldsymbol{E}^s_D]\,|||^2 + \lambda c^2 |||\,[\mathbf{n}, \boldsymbol{B}^s - \boldsymbol{B}^s_D]\,|||^2 + \mu\varsigma |||\,[\mathbf{n}, \boldsymbol{\psi}]\,|||^2\right\} \tag{9}$$

Problem 3. Ghost Obstacle Problem. In the circumstances of Problem 1 given an obstacle G such that $G \neq \emptyset$, $\overline{G} \cap \overline{\Omega} = \emptyset$, its electromagnetic boundary impedance χ_G, and a bounded set without holes and internal cavities Ω_1 such that $\overline{\Omega}$, $\overline{G}$ are contained in Ω_1 and $\partial\Omega_1$ is a sufficiently regular surface, choose a control vector field $\boldsymbol{\psi}(\mathbf{x},t)$, $(\mathbf{x},t) \in \partial\Omega \times \mathbf{R}$ in a suitable class of admissible

controls, in order to minimize a cost functional that roughly speaking measures in $(\mathbf{R}^3 \setminus \overline{\Omega}_1) \times \mathbf{R}$ the "magnitude of the difference" between the electromagnetic field $(\boldsymbol{E}^s, \boldsymbol{B}^s)$ scattered by Ω, χ (when the control vector field is active) and the electromagnetic field $(\boldsymbol{E}^s_G, \boldsymbol{B}^s_G)$ scattered by G, χ_G when hit by the incoming field $(\boldsymbol{E}^i, \boldsymbol{B}^i)$ and the "magnitude" of the control vector field employed. The couple G, χ_G will be called "ghost obstacle". For simplicity we assume the "ghost obstacle" to be a passive obstacle. The Ghost Obstacle Problem can be modelled as the optimal control problem (7), (1), (2), (5), (6) if the functional $\mathcal{F}_{\lambda,\mu}$ that appears in (7) is defined as follows,

$$\mathcal{F}_{\lambda,\mu}(\boldsymbol{\psi}) = (1+\chi)\left\{\lambda \| [\mathbf{n}, \boldsymbol{E}^s - \boldsymbol{E}^s_G] \|_1^2 + \lambda c^2 \| [\mathbf{n}, \boldsymbol{B}^s - \boldsymbol{B}^s_G] \|_1^2 + \mu_\varsigma \| [\mathbf{n}, \boldsymbol{\psi}] \|^2 \right\}, \quad (10)$$

where $\| \cdot \|_1$ is a norm on a suitable space of functions defined on $\partial\Omega_1 \times \mathbf{R}$.

Finally we formulate the so called Definite Band Problems.

Let $K \subseteq \mathbf{R}$ be an assigned set of the frequency space that we assume to be an open interval symmetric with respect to the origin, let $\check{I}_K(t)$, $t \in \mathbf{R}$ be the inverse Fourier transform of the characteristic function of the set K and let us denote with $f * g$ the convolution product with respect to the time variable of the functions f and g. The set K is the definite band in the frequency space where the smart obstacle pursues its goal.

Problem 4. Definite Band Furtivity Problem. In the circumstances of Problem 1 given K choose a control vector field $\boldsymbol{\psi}(\mathbf{x}, t)$, $(\mathbf{x}, t) \in \partial\Omega \times \mathbf{R}$ in a suitable class of admissible controls, in order to minimize a cost functional that roughly speaking measures the "magnitude" in the frequency band K ($K \subset \mathbf{R}$) of the electromagnetic field $(\boldsymbol{E}^s, \boldsymbol{B}^s)$ scattered by Ω, χ (when the control vector field is active) when hit by the incoming field $(\boldsymbol{E}^i, \boldsymbol{B}^i)$ and the "magnitude" of the control vector field employed.

The Definite Band Furtivity Problem can be modelled as the optimal control problem (7), (1), (2), (5), (6) if the functional $\mathcal{F}_{\lambda,\mu}$ that appears in (7) is defined as follows,

$$\mathcal{F}_{\lambda,\mu}(\boldsymbol{\psi}) = (1+\chi)\left\{\lambda \| \check{I}_K * [\mathbf{n}, \boldsymbol{E}^s] \|^2 + \lambda c^2 \| \check{I}_K * [\mathbf{n}, \boldsymbol{B}^s] \|^2 + \mu_\varsigma \| [\mathbf{n}, \boldsymbol{\psi}] \|^2 \right\}. \quad (11)$$

Similarly we can consider the remaining goals on a definite band:

Problem 5. Definite Band Masking Problem. In the circumstances of Problem 1 given K, an obstacle $D \subseteq \Omega$, and its electromagnetic boundary impedance χ_D, choose a control vector field $\boldsymbol{\psi}(\mathbf{x}, t)$, $(\mathbf{x}, t) \in \partial\Omega \times \mathbf{R}$ in a suitable class of admissible controls, in order to minimize a cost functional that roughly speaking measures the "magnitude of the difference" in the frequency band K between the electromagnetic field $(\boldsymbol{E}^s, \boldsymbol{B}^s)$ scattered by Ω, χ (when the control vector field is active) and the electromagnetic field $(\boldsymbol{E}^s_D, \boldsymbol{B}^s_D)$ scattered by D, χ_D when hit by the incoming field $(\boldsymbol{E}^i, \boldsymbol{B}^i)$ and the "magnitude" of the control vector field employed.

The Definite Band Masking Problem can be modelled as the optimal control problem (7), (1), (2), (5), (6) if the functional $\mathcal{F}_{\lambda,\mu}$ that appears in (7) is defined as follows,

$$\mathcal{F}_{\lambda,\mu}(\boldsymbol{\psi}) = (1+\chi)\left\{\lambda \| \check{I}_K * [\mathbf{n}, \boldsymbol{E}^s - \boldsymbol{E}^s_D] \|^2 + \lambda c^2 \| \check{I}_K * [\mathbf{n}, \boldsymbol{B}^s - \boldsymbol{B}^s_D] \|^2 + \mu_\varsigma \| [\mathbf{n}, \boldsymbol{\psi}] \|^2 \right\}. \quad (12)$$

Problem 6. Definite Band Ghost Obstacle Problem. In the circumstances of Problem 1 given K, an obstacle G such that $G \neq \emptyset$, $\overline{G} \cap \overline{\Omega} = \emptyset$, its electromagnetic boundary impedance χ_G, and a bounded set without holes and internal cavities Ω_1 such that $\overline{\Omega}$, $\overline{G} \subset \Omega_1$, and $\partial\Omega_1$ is sufficiently regular choose a control vector field $\boldsymbol{\psi}(\mathbf{x}, t)$, $(\mathbf{x}, t) \in \partial\Omega \times \mathbf{R}$ in a suitable class of admissible controls, in order to minimize a cost functional that roughly speaking measures the "magnitude of the difference" in the frequency band K between the electromagnetic field $(\boldsymbol{E}^s, \boldsymbol{B}^s)$ scattered by Ω, χ (when the control vector field is active) and the electromagnetic field $(\boldsymbol{E}^s_G, \boldsymbol{B}^s_G)$ scattered by G, χ_G when hit by the incoming field

$(\boldsymbol{E}^i, \boldsymbol{B}^i)$ and the "magnitude" of the control vector field employed.

The Definite Band Ghost Obstacle Problem can be modelled as the optimal control problem (7), (1), (2), (5), (6) if the functional $\mathcal{F}_{\lambda,\mu}$ that appears in (7) is defined as follows,

$$\mathcal{F}_{\lambda,\mu}(\boldsymbol{\psi}) = (1+\chi)\big\{\lambda \| \check{I}_K * [\mathbf{n}, \boldsymbol{E}^s - \boldsymbol{E}^s_G] \|_1^2 + \lambda c^2 \| \check{I}_K * [\mathbf{n}, \boldsymbol{B}^s - \boldsymbol{B}^s_G] \|_1^2 + \mu\varsigma \| [\mathbf{n}, \boldsymbol{\psi}] \|^2 \big\}. \quad (13)$$

Note that the Definite Band Problems formulated, that is Problems 4, 5, 6, are generalizations of Problems 1, 2, 3. In fact when we choose $K = \mathbf{R}$ the Definite Band Furtivity, Masking and Ghost Obstacle Problems reduce respectively to the Furtivity, Masking and Ghost Obstacle Problems. The advantage of solving the Definite Band Problems rather than the corresponding problems on the entire frequency space is that the "price" to be paid in term of the control variable employed is smaller when the Definite Band Problems are considered. In fact as shown in [10] in the acoustic case in the Definite Band Furtivity and Ghost Obstacle Problems the "quantity" of the control variable, measured by the norm used in the cost functional, required to get a given furtivity effect (or to get a given "ghost" effect) in the frequency band K is smaller than the "quantity" of the control variable needed to get the same effect on the entire frequency space (i.e. when $K = \mathbf{R}$).

A straightforward mode to solve the six control problems formulated here is the use of an optimization routine and a numerical solver for the Maxwell equations. This approach is computationally very expansive since it implies the solution of the Maxwell equations (several times due to the necessity of estimating gradient and eventually "Hessian" of the cost functionals involved in the control problems) at each iteration of the optimization procedure. A computationally cheaper approach can be obtained using the Pontryagin maximum principle. In fact under some hypotheses using the Pontryagin maximum principle it is possible to write the first order optimality conditions corresponding to these control problems as a system of partial differential equations with the necessary boundary, initial and final conditions. Highly parallelizable numerical methods can be developed to solve these systems of partial differential equation. For brevity we refer the interested reader to [4], and [5], [3], [10]. The numerical results obtained in Section 3 have been obtained using the Pontryagin maximum principle. In fact, we have derived the first order optimality conditions, i.e. a system of partial of differential equations for Problem 1 and a similar system for Problem 6 and then we have solved these systems developing suitable solvers based on the operator expansion method presented in [5] and [11].

3. Some numerical results

We present some numerical results relative to two experiments involving smart obstacles. In both experiments we choose $c = 1$, $\varsigma = 1$ and the following electromagnetic incoming field,

$$\boldsymbol{E}^i(\mathbf{x}, t) = (1,0,0)^T e^{-[x_3 - t]^2},$$
$$\boldsymbol{B}^i(\mathbf{x}, t) = (0,1,0)^T e^{-[x_3 - t]^2},$$
$$(\mathbf{x}, t) \in \mathbf{R}^3 \times \mathbf{R}. \quad (14)$$

The smart obstacle of the first experiment is a sphere of center the origin and radius 2 (Figure 1a)) with boundary impedance $\chi = 2$ that pursues the goal of being undetectable (i.e. Furtivity Problem, Problem 1).

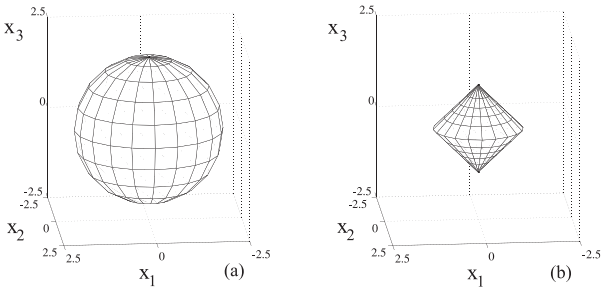


Figure 1. Obstacles.

The smart obstacle of the second experiment is a perfectly conducting (i.e. $\chi = 0$) double cone (see Figure 1b)) that pursues the goal of being

undetectable in the subset $K = (-1, 1)$ of the frequency space (Definite Band Furtivity Problem, Problem 4). The double cone consists of two cones of the same height 1.2 and base (a circle having center the origin and radius 1.2) one upon the other through their bases.

The numerical results relative to the first experiment are shown in Table I. Let us describe these results. Let B_{R_i}, $i = 1, 2, 3$ be spheres having center the origin and radii $R_i = 2.0 + (i-1) * 0.5$, $i = 1, 2, 3$ and let $t_\nu = -2 + \nu$, $\nu = 1, 2, 3$ be three time values such that the incident field is beginning to hit the scatterer ($t = t_1 = -1$), is going through the body of the scatterer ($t = t_2 = 0$) and is leaving the scatterer ($t = t_3 = 1$) respectively. Note that the spheres B_{R_i}, $i = 1, 2, 3$ contain or coincide with the smart obstacle, $\Omega \subseteq B_{R_i}$, $i = 1, 2, 3$. Let $\boldsymbol{E}_a^s$ and $\boldsymbol{E}_p^s$ denote the electric field generated by the smart obstacle when the optimal surface electric current density is used, and by the same obstacle considered as a passive obstacle respectively. For $i = 1, 2, 3$, $\nu = 1, 2, 3$ we define the following quantities,

$$\epsilon^{a,\lambda}_{E,R_i,\nu} = \left[\int_{\partial B_{R_i}} \|\boldsymbol{E}_a^s(\mathbf{x}, t_\nu)\|^2 ds_{\partial B_{R_i}}(\mathbf{x}) \right]^{1/2}, \tag{15}$$

$$\epsilon^{p}_{E,R_i,\nu} = \left[\int_{\partial B_{R_i}} \|\boldsymbol{E}_p^s(\mathbf{x}, t_\nu)\|^2 ds_{\partial B_{R_i}}(\mathbf{x}) \right]^{1/2}, \tag{16}$$

and

$$\epsilon^{\lambda}_{E,R_i} = \min_{\nu=1,2,3} \frac{|\epsilon^{p}_{E,R_i,\nu} - \epsilon^{a,\lambda}_{E,R_i,\nu}|}{|\epsilon^{p}_{E,R_i,\nu}|}, \ i = 1, 2, 3, \tag{17}$$

where $ds_{\partial B_{R_i}}$ is the surface measure on ∂B_{R_i}, $i = 1, 2, 3$. Note that the quantity $\epsilon^{a,\lambda}_{E,R_i,\nu}$, $\epsilon^{p}_{E,R_i,\nu}$ $i = 1, 2, 3$, $\nu = 1, 2, 3$ are a sample of the "magnitude" of the electric fields generated by the smart obstacle and by the passive obstacle respectively. The quantity $\epsilon^{\lambda}_{E,R_i}$, $i = 1, 2, 3$ is a measure of how the electric field generated by the smart obstacles is small when compared with the electric field generated by a passive obstacle that is, is a measure of the furtivity effect achieved. The results obtained are satisfactory when $\epsilon^{\lambda}_{E,R_i}$ is close to one, in fact when $\epsilon^{a,\lambda}_{E,R_i,\nu} = 0$ we have $\epsilon^{\lambda}_{E,R_i} = 1$.

Table I . Furtivity Effect

$\lambda = 0.1, \quad \mu = 0.9$				
R_i	ν	$\epsilon^{a,\lambda}_{E,R_i}$	ϵ^{p}_{E,R_i}	$\epsilon^{\lambda}_{E,R_i}$
2.0	1	1.651	2.323	0.287
2.5	1	1.082	1.421	0.238
3.0	1	0.753	0.911	0.173
$\lambda = 0.5, \quad \mu = 0.5$				
R_i	ν	$\epsilon^{a,\lambda}_{E,R_i}$	ϵ^{p}_{E,R_i}	$\epsilon^{\lambda}_{E,R_i}$
2.0	1	0.808	2.323	0.652
2.5	1	0.525	1.421	0.630
3.0	1	0.367	0.911	0.597
$\lambda = 0.9, \quad \mu = 0.1$				
R_i	ν	$\epsilon^{a,\lambda}_{E,R_i}$	ϵ^{p}_{E,R_i}	$\epsilon^{\lambda}_{E,R_i}$
2.0	1	0.191	2.323	0.918
2.5	1	0.123	1.421	0.913
3.0	1	0.085	0.911	0.906

Note that the column denoted with ν in Table I contains the minimizer of formula (17). Results similar to those shown in Table I have been obtained for the magnetic induction vector field (see [5] for further details).

Note that the furtivity effect increases when λ increases and that it ranges from 17% when $\lambda = 0.1$ to 90% when $\lambda = 0.9$ (see Table I).

Finally Figures 2, 3 show the numerical results relative to the second experiment. In this experiment we choose $\lambda = 0.9$, $\mu = 0.1$, $K = (-1, 1)$. As above, let $\boldsymbol{E}_a^s$, $\boldsymbol{B}_a^s$ and $\boldsymbol{E}_p^s$, $\boldsymbol{B}_p^s$ be the electric vector field and the magnetic induction vector field scattered by the smart double cone when the optimal surface electric current density is employed and by the passive double cone respectively. Figure 2 shows from left to right in the colour scale shown the Euclidean norms of the convolution products $\check{I}_K * \boldsymbol{E}_a^s$, $\check{I}_K * \boldsymbol{E}_p^s$, $\check{I}_K * \boldsymbol{B}_a^s$, $\check{I}_K * \boldsymbol{B}_p^s$ on the sphere B_{R_2} as a function of the polar angles (θ, ϕ), for three different values of the time variables that is, $t = 0$, $t = 2$, $t = 3$. Note that $\Omega \subset B_{R_2}$ and that the norms of the vector fields $\check{I}_K * \boldsymbol{E}_a^s$, $\check{I}_K * \boldsymbol{B}_a^s$ are negligible compared to the

corresponding norms $\check{I}_K * \boldsymbol{E}_p^s$, $\check{I}_K * \boldsymbol{B}_p^s$.

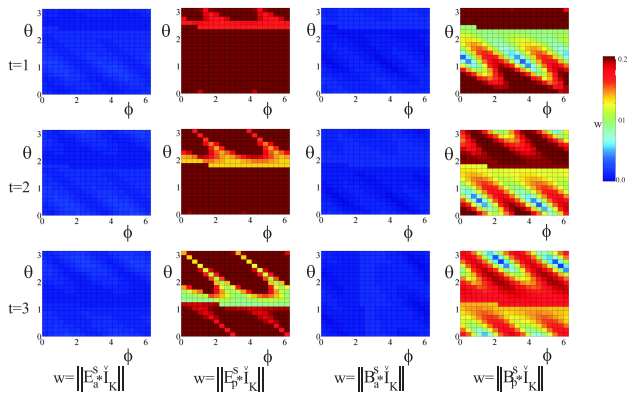


Figure 2. Furtivity effect in the frequency band K.

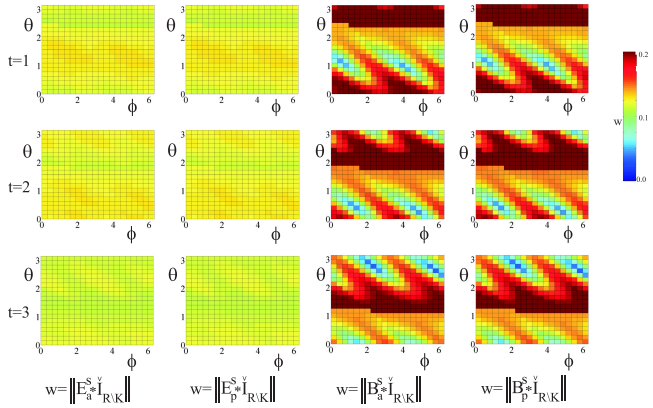


Figure 3. Furtivity effect outside the frequency band K.

Similarly Figure 3 shows from left to right in the colour scale shown the norms of the convolution products $\check{I}_{\mathbf{R}\setminus K} * \boldsymbol{E}_a^s$, $\check{I}_{\mathbf{R}\setminus K} * \boldsymbol{E}_p^s$, $\check{I}_{\mathbf{R}\setminus K} * \boldsymbol{B}_a^s$, $\check{I}_{\mathbf{R}\setminus K} * \boldsymbol{B}_p^s$ on the sphere B_{R_2} as a function of the polar angles (θ, ϕ) for $t = 0$, $t = 2$, $t = 3$. Note that the norms of the vector fields $\check{I}_{\mathbf{R}\setminus K} * \boldsymbol{E}_a^s$, $\check{I}_{\mathbf{R}\setminus K} * \boldsymbol{B}_a^s$ are similar to the corresponding norms of $\check{I}_{\mathbf{R}\setminus K} * \boldsymbol{E}_p^s$, $\check{I}_{\mathbf{R}\setminus K} * \boldsymbol{B}_p^s$. That is, outside of the frequency band K no furtivity effect is present. Note that the colour scales used in Figures 2 and 3 to represent the data are the same.

4. Conclusions

In this paper we have shown how mathematical models, such as optimal control problems, can be used profitably to design smart objects able to pursue non trivial goals. The main advantage of the mathematical formulation of the electromagnetic scattering problem involving smart obstacles proposed in this paper is that it allows to reduce the solution of the scattering problem to the solution of an optimal control problem whose optimal solution can be determined as the solution of a suitable system of coupled partial differential equations. This fact guarantees a great computational efficiency. In fact the most standard approaches solve the optimal control problem iteratively. That is at each step of the iterative procedure a system of partial differential equations must to be solved.

References

[1] F. Mariani, M. C. Recchioni, and F. Zirilli, "The Use of the Pontryagin Maximum Principle in a Furtivity Problem in Time-Dependent Acoustic Obstacle Scattering", *Waves in Random Media,* Vol. 11, 549-575, 2001.

[2] L. Fatone, M. C. Recchioni, and F. Zirilli, "A Masking Problem in Time Dependent Acoustic Obstacle Scattering", *ARLO - Acoustics Research Letters Online,* Vol. 5, Issue 2, 25-30, 2004.

[3] L. Fatone, M. C. Recchioni, and F. Zirilli, "Mathematical models of "active" obstacles in acoustic scattering", in *Control and Boundary Analysis*, J. Cagnol and J. P. Zolesio Editors, Marcel Dekker/CRC Press, Boca Raton, Fl. USA, Lecture Notes in Pure and Applied Mathematics, Vol. 240, 119-129, 2005.

[4] L. Fatone, M. C. Recchioni, and F. Zirilli, "Some Control Problems for the Maxwell Equations Related to Furtivity and Masking Problems in Electromagnetic Obstacle Scattering", in G.C. Cohen, E. Heikkola, P. Joly and P. Neittaanmaki editors, *Mathematical and Numerical Aspects of Wave Propagation. Waves 2003*, Springer Verlag, Berlin, 189-194, 2003.

[5] L. Fatone, M. C. Recchioni, and F. Zirilli, "Furtivity and Masking Problems in Time Dependent Electromagnetic Obstacle Scattering", *Journal of Optimization Theory and Applications,* Vol. 121, 223-257, 2004.

[6] B. Chambers, "A Smart Radar Absorber", *Smart Materials and Structures,* Vol. 8, 64-72, 1999.

[7] K. L. Ford, and B. Chambers, "A Smart Microwave Absorber", *IEE Electronics Letters,* Vol. 36, 50-52, 2000.

[8] B. Chambers, and A. Tennant, "Influence of Switching Waveform Characteristics on the Performance of a Single Layer Phase-Switched", *IEEE Transactions on Electromagnetic Compatibility,* Vol. 44, 434-441, 2002.

[9] J. A. Stratton, "Electromagnetic Theory", McGraw-Hill, New York, 1941.

[10] L. Fatone, G. Pacelli, M. C. Recchioni, and F. Zirilli, "The Use of Optimal Control Methods to Study Two New Classes of Smart Obstacles in Time Dependent Acoustic Scattering", submitted to *Journal of Engineering Mathematics,* 2005.

[11] M. C. Recchioni, and F. Zirilli, "A New Formalism for Time-Dependent Electromagnetic Scattering from a Bounded Obstacle", *Journal of Engineering Mathematics,* Vol. 47, 17-43, 2003.

Performing 3-D FDTD Simulations in less than 3 Seconds on a Personal Computer and its Application to Genetic Algorithm Antenna Optimization

L.A. Griffiths, and C.M. Furse
ECE Department, University of Utah
50 S. Campus Drive, MEB 3280
Salt Lake City, UT 84112
USA

***Abstract* —** FDTD simulations generally require significant computational resources and time. This paper systematically reduces the number of time steps and the grid size to determine the shortest simulation time that returns results with tolerable error for microstrip antenna simulations and their optimization of insertion loss with the genetic algorithm. Although the error would generally be unacceptable for traditional antenna simulations, it is sufficiently small to optimize their design. Simulations in less than 3 seconds on a P4 2.8 GHz processor were shown to be usable, with error approximately equal to manufacturing tolerances. A dual band 'waffle' antenna is designed that has better performance than the traditional dual band "F" antenna.

I. Introduction

The finite difference time domain (FDTD) method [1], has become a mainstay of electromagnetic computation. It has been applied to a seemingly endless array of applications, with little limitation of geometry, frequency band, materials, etc. FDTD has an advantage over the method of moments which is also commonly used to model antennas [2], because it can simulate a fully heterogeneous antenna or antenna environment, however this does come at a significant cost. Computer simulation time and memory requirements are inherently large in heterogeneous simulations of this type, and FDTD is no exception. This computational cost has a serious impact on the numerical optimization of antennas, such as is often done using the genetic algorithm (GA) [3]. The GA will create a "population" of many (in our case 16) antennas, run their simulations (an individual FDTD simulation for each antenna), reject the poor performers, mutate/cross-over the good performers, and repeat this for many (50-100) generations resulting in 800-1600 FDTD simulations. A simple pioneering FDTD simulation with a grid size of 20x20x40 took over 38 minutes to run 600 time steps [4]. Since that time, much larger simulations have been run requiring days and even weeks [5].

This paper focuses on ways to reduce the overall time required by reducing the time for each FDTD simulation. Many efforts have been made to reduce the time required to run FDTD simulations. Methods to make FDTD run faster include: subdividing the problem and running on multiple computers or processors in parallel [5-10], using an initially smaller grid that expands with time [11], efficient processing of fields to extract useful information [12], using variable cell sizes [13], and exploiting symmetry to reduce the model size [14, 15 and others] to name a few. Dedicated hardware has also been developed specifically for FDTD to circumvent the limitation of general purpose processors [16]. A single FDTD simulation can now readily be done for most if not all field analysis applications of interest. When multiple simulations are required, however, the computational requirements can become prohibitive.

Researchers have used the GA to optimize antennas, but have generally relied on methods other than FDTD [2, 17]. Some researchers have used GA-FDTD schemes, but found computational constraints to be a limiting factor [18, 19]. They have limited the GA search to a small set of parameters, thereby limiting its usefulness as an optimization method. This paper shows that with the proper selection parameters, FDTD simulations can be run quickly enough on a personal computer to be used to design a dual band antenna using the GA in a very short period of time. Unlike typical numerical solutions where we need excellent accuracy and precision, it was found that relatively "sloppy" FDTD simulations, while not perfectly accurate, can yield results that are sufficient to determine the relative performance of similar antennas, and hence the design of optimal antennas.

Section II describes a traditional dual band monopole antenna that is used for comparison and simulation purposes. Section II also introduces QFDTD, an FDTD program well suited to use with genetic algorithms. Section III analyzes the how reducing the run time of the FDTD algorithm affects accuracy. The number of time steps is reduced, and the FDTD grid is reduced in a systematic manner while error is measured. Section IV applies the results from Section III to the design of a dual band antenna using the genetic algorithm. Section V gives instructions on how to apply these methods generally, and Section VI draws conclusions from the results.

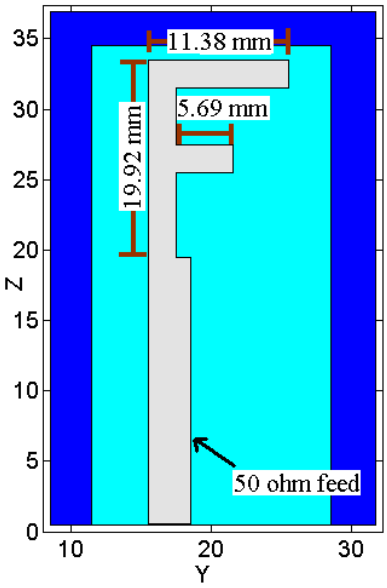


Fig. 1. Model of dual band 2.4/5.2 GHz "F" antenna as created by the authors. ΔX = 0.762 mm, ΔY=ΔZ= 1.423 mm. A ground plane on the back of the substrate extends from Z=1 to the end of the 50 ohm feed at Z=19.

II. Antenna Model and FDTD

In [20], a dual band monopole "F" antenna is designed on a microwave substrate at 2.4/5.2 GHz. In this paper, we optimize the design of a similar antenna structure shown in Fig. 1. The antenna is designed to have low insertion loss at 2.4 and 5.2 GHz as shown in Fig. 2. The FDTD model cell size is 0.762 x 1.423 x 1.423 mm^3. The microwave substrate is 0.060 inches (1.524 mm) thick and has a relative dielectric constant of 2.6. The cell dimension X is chosen to be half the substrate thickness, and the Y-Z dimension is chosen to simulate a 50 Ω microstrip feed. This antenna model is used in section IV to determine key FDTD simulation parameters.

FDTD models Maxwell's equations in a spatial grid consisting of electric and magnetic field components, which are alternately computed as the algorithm steps through time. Each cell in the grid contains six orthogonal electric and magnetic field components. Because the cells in the grid must be a minimum of $\lambda/10$ for numeric accuracy and are typically $\lambda/20$ to $\lambda/60$, the grid can be very large for many structures, including antennas. As an example, a 100x100x100 grid that can model roughly a λ^3 antenna has six million field components that need to be updated with each of about 2000 time steps.

The major sources of error in FDTD calculations are due to numerical dispersion, reflections due to imperfect absorbing boundary conditions

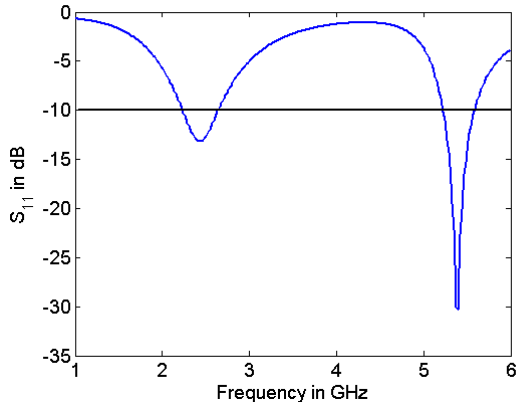


Fig. 2. Insertion loss of "F" antenna with 50 ohm feed. Note that the "F" antenna is frequency shifted at 5.2 GHz due to the course grid requirements of FDTD.

(ABCs), and poor modeling because of a discrete rectangular grid. Numerical dispersion can be minimized by reducing the cell size. When the grid sampling density is 10 points per free-space wavelength, the numerical dispersion is approximately 1%, which is considered the minimum special sampling rate for accurate simulations [21]. By increasing the grid sampling density to 20 points per wavelength, the dispersion is reduced to 0.2% (and the grid size increases by a factor of 8). This paper uses a spatial sampling frequency of about 35 points per wavelength to accurately model the feed, reducing dispersion errors to negligible levels.

An in depth review of analytical boundary operators are covered by [22]. This review explains that the approximations used to create ABCs cause them to be imperfect. Waves traveling normal to the 2nd order Mur boundary are absorbed well. As the angle of incidence increases, the reflection coefficient increases. By increasing the grid size (and the computation time), the maximum angle of incidence is reduced, reducing reflections. Also, fewer waves will reflect back onto the antenna, because the antenna is located further from the boundary.

For fast FDTD simulations, care needs to be taken to ensure accurate simulations, while keeping the grid to a minimal size. In addition, the boundary conditions need to be computationally efficient. The commercial software package QFDTD uses simple update equations that assume a non-dispersive media. It also uses the computationally efficient Mur 2nd order boundary condition. It is written in FORTRAN90, allowing the user to modify it and port it to any desired platform. Additionally, it uses text files for all input-output operations, allowing the GA to easily create new models and access output data [23]. QFDTD runs two simulations to analyze a structure. The first simulation

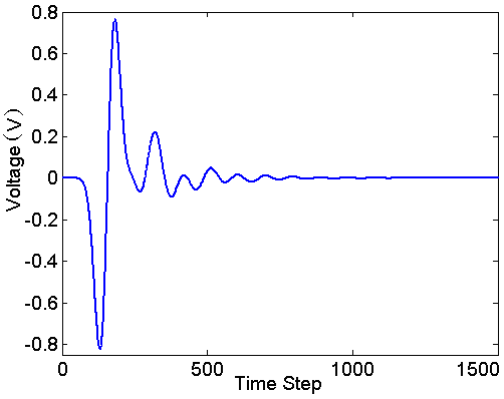


Fig. 3. Voltage on microstrip feed versus time.

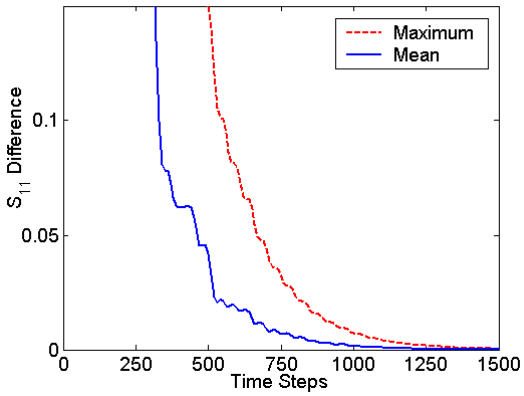


Fig. 4. Difference in insertion loss power over 1-6 GHz frequency band when number of time steps is reduced, compared to a simulation running 5000 time steps.

measures the incident signal at the microstrip feed. The second simulation measures the total signal. By subtracting the incident signal from the total signal, the scattering parameters are calculated over a wide frequency range. This two-simulation setup can be exploited with GA optimization. If the feed doesn't change, the first simulation results can be used for all subsequent simulations, and its contribution to overall computation time is negligible.

III. Simulations Speed and Accuracy

Well known ways to reduce FDTD simulation time include reducing the number of operations by reducing the model size and/or running fewer time steps. However, there are fundamental limits on how much reduction in size and time can be done before inaccuracies are introduced. This section of the paper assesses the impact of each time reduction method on the accuracy and speed of the program.

A. Run for fewer time steps

Perhaps the most obvious method to speed up the FDTD simulation is to reduce the total number of time steps in the simulation. The "F" antenna model was simulated for a long time, to determine the minimum number of time steps needed for transients to die down to a sufficient level as shown in Fig. 3. Then S_{11} is calculated from 1-6 GHz in 0.1 GHz increments. The simulation is then repeated many times stopping at 300 to 1500 time steps. S_{11} is calculated for each simulation, and the change in reflected power is computed at each frequency increment. The maximum and average change in reflected power is given in Fig. 4. Simulations show that S_{11} changes very little after 1500 time steps. Fig. 4 shows that if the number of time steps is reduced to 900, the error is relatively small, but the computation time is almost cut in half. Reducing the number of time steps below 600 creates significant errors, especially in the higher frequencies.

The number of time steps required is also dependent on the size of the discrete time step. Generally, Δt should be the maximum value that meets the Courant stability criterion unless lossy or active components are embedded into the FDTD grid [23]. Using the highest stable value of Δt will avoid unnecessary time steps. QFDTD and other commercial FDTD programs used by the authors automatically calculate the correct value for Δt.

B. Reduce grid size

Reducing the grid size can have an even greater effect on simulation time than the number of time steps. If the dimensions of a 3-D grid are halved in each direction, the number of cells is reduced by a factor of eight, but the grid size reduction effectively brings the outer boundary closer and causes reflections at the boundary to increase. To determine how much grid size affects S_{11}, the "F" antenna model was simulated on a grid with 37x48 cells in the YxZ direction and a variable number of cells in the X direction. It is first simulated on a 100x37x48 grid for comparison. After the comparison simulation is run, the X grid is reduced to 10 and expanded in the X direction with each simulation, while the antenna is held at the center of the grid. Fig. 5a shows that as the grid expands in the X dimension, the change in reflected S_{11} power is reduced.

The next test enlarged the grid in all three dimensions and compared the results to a 100x100x100 cell grid. As can be seen from Fig. 5b the difference in S_{11} continued to decrease as the grid is enlarged, but the results have not fully converged, even when the grid is a 100x100x100 cell. Expanding the grid beyond 100x100x100 cells to reduce reflections proved unreasonable. Rather a smaller grid would be implemented with a more effective, but computationally costlier, boundary condition. Thus a

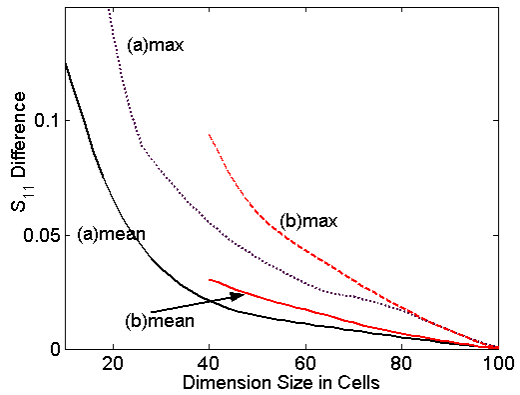


Fig. 5. Change in S_{11} power when grid size is changed. (a) YZ dimensions are 37 and 48. X dimension vary from 10-100 cells. (b) XYZ dimensions vary from 40-100 cells.

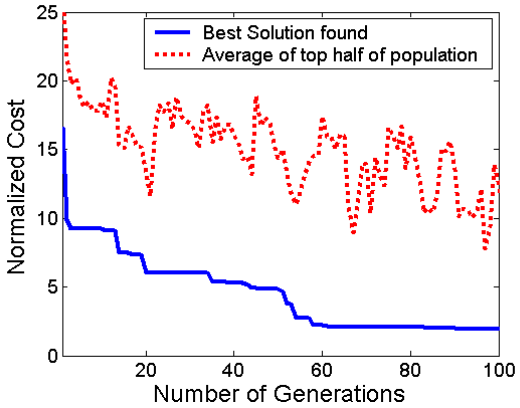


Fig. 6. Cost (or fitness) of antennas at each generation of the GA. The cost is the percentage of reflection loss over the frequency range of 2.2 – 2.6 GHz, and 5.0-5.5 GHz.

100x100x100 grid for these simulations could be considered large enough when using the Mur 2nd order boundary condition. Getting the fastest execution time precludes the use of high cost boundary conditions even for small simulations.

It is significant to note that reducing the grid size caused more error in the lower frequency part of the 1 to 6 GHz range. Reducing the grid size moves the outer boundary closer to the antenna model. Because distance to the outer boundary is relative to the wavelength, lower frequencies will be closer to the boundary than higher frequencies. It was also observed that reducing the number of time steps produced more error in the higher frequency part of the 1 to 6 GHz simulations (although we can't explain why and it may be model-dependent). From these observations we can conclude that a relatively narrow band design can be run faster. A low frequency design can run for fewer time steps, and a high frequency design can simulate on a smaller grid, while maintaining simulation accuracy.

C. Other execution speed factors

The FDTD executable needs to be optimized for speed. The authors found that different compilers and compiler settings can affect execution speed by more than 300%. Also, storing more information than necessary increase the simulation time and memory usage.

Each speed increase factor is multiplied by the next factor. Implementing all speedup factors results in a dramatic decrease in FDTD run time. This speedup makes FDTD a viable simulator for running hundreds or even thousands of simulations needed by the GA optimizer.

IV. Application Example-Design of GA Antenna

To show that high speed FDTD simulations can produce useful results, the authors remove the branches from the "F" antenna model to produce a simple monopole. This simple monopole is then placed in a 25x30x42 cell grid to be simulated for 900 time steps. A rectangle with either non-metal or metal covered cells is placed over the top of the monopole model. The cells with and without metal correspond directly to the "1s" and "0s" in a binary chromosome that is controlled by the GA.

The GA creates a population of 16 antennas, and simulates them for 100 generations. Each antenna is given a cost which is based on how much insertion loss it generates over the frequency bands 2.2-2.6 GHz and 5.0-5.5 GHz. Single point crossover is used, and population decimation is used as the selection criteria. The mutation method randomly picks bits in each chromosome and replaces the previous value with a random bit. This mutation method is different than standard mutation schemes, but is necessary because it may be desirable to have a higher percentage of "1" bits. Fig. 6 shows that the GA is able to evolve an antenna with low return loss after 60 generations. Fig. 7 shows the optimized antenna model. After the GA completes optimization, the final model is simulated for 1500 time steps on a larger 100x100x100 grid for comparison.

Fig. 8 shows that the smaller model didn't produce the most accurate results, but the optimized antenna performance is still very good. In addition, the change in S_{11} from the small to the large model is comparable to manufacturing tolerances found by previous GA antenna designers [16, 17]. The larger comparison simulation shows that the GA antenna has a

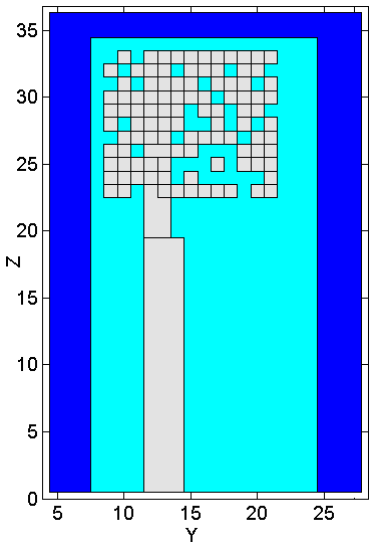


Fig. 7. Antenna design created by GA using structure similar to "F" antenna design in Fig. 1. The GA created a population of 16 structures and evolved them for 100 generations to create this final design found in the 91st generation. The substrate dimensions are 17x34 cells or 24.19mm x 48.38mm. The substrate is 1.524 mm thick.

-10 dB insertion loss bandwidth from 2 to 2.6 GHz and 5.1 to 5.7 GHz. This is better than the results obtained by the authors and [20] for the "F" antenna. It is true that the GA design shown isn't much better than the "standard" design. The point of this paper is that it can quickly be designed using the GA if a small FDTD model is used. We purposely used a simple dual band design because we are emphasizing FDTD. Once the simulation time is reduced, very wide or multiband designs can be quickly produced that would be nearly impossible to produce using conventional design techniques. We have obtained better results using smaller cells, but we wanted to step to the limits in this paper.

Our first optimization simulations took over 17 minutes. To run 1600 simulations would take 19 days, and we often had to make changes in our model starting the process over again. Using the techniques described in this paper (along with faster processors) has reduced simulation time to less than 3 seconds for the models presented here. With faster simulation times, several designs can be produced in a single day.

The entire optimization presented took only 82 minutes and ran 1600 FDTD simulations. That corresponds to 3.06 seconds for the GA to create each model, run the FDTD simulation, and evaluate its cost. The simulation was run on a Pentium 4 running at 2.8 GHz. The FDTD executable was created using the Intel FORTRAN Compiler Version 8.0 for Linux.

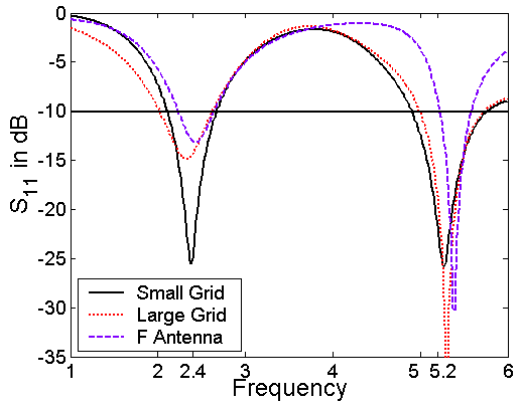


Fig. 8. Optimized S_{11} results as calculated on small grid (25x30x42 cells) for GA optimization, and large grid (100x100x100 cells) for comparison with "F" antenna. Even though the small grid introduces simulation error, the -10 dB bandwidth is very close on both simulations, justifying the use of the small grid for fast optimization.

V. General Application of Speed Increase Methods

Carefully choosing the cell and grid size is essential to fast and accurate FDTD simulations. Minimizing the grid size and number of time steps can reduce simulation time 1-2 orders of magnitude compared to poorly chosen parameters.

As explained in the introduction, the cell size should be chosen to accurately model the structures dimensions, and meet the minimum of 10 cells per wavelength at the lowest frequency. A general guideline for choosing the grid size is to have at least 1/4 wavelength between the structure and the outer boundary at the lowest frequency of interest. For the example given, the wavelength at 2.2 GHz is 136 mm and 1/4 wavelength is 34 mm. The cell size in the X direction is 0.762 mm or about 45 cells. These are similar to the dimensions used for the 100x100x100 cell grid (about 45 cells above and below the model). The problem with this approach is that each simulation takes approximately 94 seconds or almost 42 hours for a 1600 simulation GA optimization. By reducing the grid size, additional error will be introduced into the simulation. The plots in Fig. 5 give the reader a general guide to how much error is introduced, and the grid size should be chosen based on how much error can be tolerated. Again, more error will be introduced at the lower frequencies as shown in Fig. 8.

Correctly choosing the number of time steps is also critical to a fast, accurate simulation. In the example shown, a modulated Gaussian pulse is used.

After 900 time steps the reflected signal at the feed is reduced to 0.5% of the pulses maximum amplitude. For fast simulations, the reflected signal should decay to between 0.5% and 1.0%. After 1500 time steps, negligible error was introduced, and the reflected signal was 0.1% of the maximum incident pulse. Note that highly resonant structures may require much longer simulation times.

Advanced absorbing boundaries such as PML have lower reflections, but a much higher computational cost. It should therefore be possible to reduce the grid size without inducing as much error as when reducing the grid size using the Mur boundary. Optimizing using the GA on a larger grid after it has been optimized on a small grid is an excellent way to apply the results of the small grid optimization. When only a few variables are present, a hybrid GA that consists of a GA and local optimizer works extremely well and will outperform the GA alone [24]. However, a local optimizer wouldn't be appropriate in this situation because each cell is considered a variable, and local optimizers are not efficient for a large number of variables.

Several antenna prototypes have been successfully built using photo-etching techniques, and measured data has matched well with simulations [25]. Using these techniques, broad and multiband designs have been created that minimize size requirements while achieving extremely low return loss.

VI. Conclusions

FDTD is a viable solution for the GA simulator on a PC if FDTD parameters are chosen for quick simulation. Even though individual simulation results may not be extremely accurate, their relative values are sufficient for the GA to find a good solution. When the FDTD simulation is optimized for speed by running for a minimum number of time steps, using a minimal FDTD grid, and storing only necessary information, it can be fast enough to compete with other simulation techniques such as method of moments. Simulation time is no longer the limiting factor with the GA-FDTD combination, allowing for more complex designs to be generated than previously possible, including the 2.4 GHz and 5.2 GHz dual band antenna described in this paper.

References

[1] K. S. Yee, "Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media," *IEEE Trans. Antennas Propagat*., vol. AP-14, pp. 302-307, May 1966.

[2] J. M. Johnson, Y. Rahmat-Samii, "Genetic Algorithms and Method of Moments (GA/MOM) for the Design of Integrated Antennas," *IEEE Trans. Antennas Propag*., vol. 47, no. 10, pp. 1606-1614, Oct. 1999.

[3] R. L. Haupt, "An Introduction to Genetic Algorithms for Electromagnetics," *IEEE Antennas Propagat. Mag.*, vol. 37, No. 2, April 1995.

[4] A. Taflove, and M.E. Brodwin, "Computation of the Electromagnetic Fields and Induced Temperatures Within a Model of the Microwave-Irradiated Human Eye," *IEEE Trans. Microwave Theory Tech.*, pp. 888-896, Nov. 1975.

[5] A. D. Tinniswood, C. M. Furse, and O. P. Gandhi, "Computations of SAR Distributions for Two Anatomically Based Models of the Human Head Using CAD Files of Commercial Telephones and the Parallelized FDTD Code," *IEEE Trans. Antennas Propag*., vol. 46, no. 6, pp. 829-833, June 1998.

[6] K. D. Tatalias and J. M. Bornholdt, "Mapping Electromagnetic Field Computations to Parallel Processors," *IEEE Trans. Magn*., vol. 25, no. 4, pp. 2901-2906, 1989.

[7] J. V. Mullan, C. J. Gillan, and V. F. Fusco, "Optimizing the Parallel Implementation of a Finite Difference Time Domain Code on a Multi-User Network of Workstations," *Applied Computational Electromagnetics Society J.*, vol. 13, no. 2, pp. 168-178, 1998.

[8] K. C. Chew and V. F. Fusco, "A Parallel Implementation of the Finite Difference Time-Domain Algorithm," *Int. J. Numerical Modeling: Electronic Networks, Devices and Fields*, vol. 8, no. 3, pp. 293-299, 1995.

[9] L. Catarinucci, P. Palazzari, and L. Tarricone, "Human Exposure to the Near Field of Radiobase Antennas-A Full-Wave Solution Using Parallel FDTD," *IEEE Trans. Microwave Theory Tech.*, vol. 51, no. 3, pp. 935-940, March 2003.

[10] M. Sypniewski, M. Celuch-Marcysiak, J. Rudnicki, W. Gwarek, and A. Wieckowski, "Faster Analysis of Microwave Engineering Problems with Multithread FDTD Multiprocessor PCs," *Int. Conf. on Microwaves, Radar and Wireless Communications, MIKON-2000*, Wroclaw, Poland, vol. 1, pp. 275-278, May 2000.

[11] X. Wu, D. Conn, J. Song, and K. Nickerson, "Faster Finite-Difference Time-Domain Method Using Moving Spatial Boundaries," *Electronics Lett.*, vol. 28, no. 24, pp. 2229-2231, 1992.

[12] C. M. Furse, "Faster than Fourier: Ultra-Efficient Time-to-Frequency-Domain Conversions for FDTD Simulations," *IEEE Antennas Propagat. Mag.*, vol. 42, no. 6, pp. 24-34, Dec. 2000.

[13] S. D. Gedney, and F. Lansing, "Chapter 11. Nonorthogonal and Unstructured Grids," in A. Taflove, "Computational Electrodynamics: The Finite Difference Time Domain Method," 2nd Edition, Artech House Inc., 2000.

[14] D. Arakaki, W. Yu, and R. Mittra, "On the Solution of a Class of Large Body Problems with Partial Circular Symmetry (Multiple Asymmetries) by Using a Hybrid-Dimensional Finite-Difference Time-Domain (FDTD) Method," *IEEE Trans. Antennas Propag.*, vol. 49, no. 3, pp. 354-360, March 2001.

[15] J. G. Maloney, G. S. Smith, W. R. Scott, Jr., "Accurate computation of the radiation from simple antennas using the finite-difference time-domain method," *IEEE Trans. Antennas Propag.*, vol. 38, pp. 1059–1068, July 1990.

[16] Available on the World Wide Web at: www.emphotonics.com.

[17] H. Choo, H. Ling, "Design of Multiband Microstrip Antennas Using a Genetic Algorithm," *IEEE Microwave Wireless Component Lett.*, vol. 12, no. 9, pp. 345-347, Sep. 2002.

[18] P. Pinho, J. F. Rocha Pereira, "Design of a PIFA Antenna using FDTD and Genetic Algorithms," *IEEE Antennas and Propagation Society International Symposium*, 2001, vol.4, pp. 700-703.

[19] A. F. Muscat, C.G. Parini, "Novel compact handset antenna," *Antennas and Propagation, Eleventh International Conference on (IEE Conf. Publ. No. 480)*, vol.1, pp. 336-339, 2001.

[20] S.H. Yeh and K. L. Wong, "Integrated F-shaped monopole antenna for 2.4/5.2 GHz dual band operations," *Microwave Opt. Tech. Lett.*, vol. 34, pp. 24-26, July 5, 2002.

[21] Allen Taflove and Susan C. Hagness, "Computational Electrodynamics, the finite-difference time-domain method", 2nd ed., Boston:Artech House, 2000, pp. 118.

[22] Thomas G. Moore, Jeffrey G. Blaschak, Allen Taflove and Gregory A. Kriegsmann, "Theory and Application of Radiation Boundary Operators", *IEEE Trans. Antennas Propag.*, vol. 36, no. 12, Dec. 1988.

[23] Y. Qian, and T. Itoh., "FDTD Analysis and Design of Microwave Circuits and Antennas: Software and Applications," Tokyo, Japan, Realize Inc., 1999.

[24] Randy Haupt and You Chung Chung, "Optimizing Backscattering from Arrays of Perfectly Conducting Strips", *IEEE Antennas and Propag., Mag.*, vol. 45, no. 5, Oct. 2003.

[25] Lance Griffiths, Cynthia Furse and You Chung Chung, "Wide and Tri-Band Microstrip LAN Antenna Design and GUI Tool Using a GA and FDTD," *IEEE International Antennas and Propagation Society Symposium and USNC/URSI National Radio Science Meeting*, June 22-27, 2004, Montery, CA.

Lance Allen Griffiths was born in Pocatello, Idaho on November 2, 1974. He received a minor in computer science and B.S. and M.S. degrees in electrical engineering from Utah State University in 1999 and 2002 respectively. He is currently working toward a Ph.D. in electrical engineering at the University of Utah.

From Sept. 1999 to Nov. 2000 he worked for Trans-Lux Corporation in Logan Utah as the Sustaining Engineering Manager. From May 2003 to the present he has worked for RF and Sensor Innovation as a Research Associate. His research interests include wide and multiband antenna design, computational electromagnetics, and optimization using the genetic algorithm.

Cynthia Furse received her B.S. in electrical engineering with a mathematics minor at the University of Utah in 1985, an M.S. degree in electrical engineering in 1988 at the University of Utah, and her Ph.D. in electrical engineering also at University of Utah in 1994. From 1997-2002 she was a professor at Utah State University. She is currently an associate professor at the University of Utah where her

research interests include high frequency applications in bioelectromagnetics, geophysics, and advanced sensing systems.

She has taught electromagnetics, wireless communication, computational electromagnetics, microwave engineering, and antenna design. Dr. Furse was Professor of the Year in the College of Engineering in 2000 and Utah State University Outstanding Faculty Employee in 2001. She is Director of the Center of Excellence for Smart Sensors and heads an active, funded research program in electromagnetics for remote sensing and bioelectromagnetics, and is the lead engineer in the development of low-cost miniaturized frequency domain reflectometers for testing of aging aircraft wiring in the Smart Wire System design. She has been an NSF CISE Graduate Fellow, IEEE Microwave Theory and Techniques Graduate Fellow, and President's Scholar at the University of Utah.

Cubic-Spline Expansion with GA for Half-Space Inverse Problems

Wei Chien and Chien-Ching Chiu

Electrical Engineering Department, Tamkang University
Tamsui, Taiwan, R.O.C.
Email:chiu@ee.tku.edu.tw

Abstract — In this paper we address an inverse scattering problem whose aim is to determine the geometrical as well as the physical properties of a perfectly conducting cylindrical body buried in a half-space. We use cubic-spline method instead of trigonometric series to describe our shape and reformulated into an optimization problem and solved by the genetic algorithm. The genetic algorithm is employed to find out the global extreme solution of the object function. As a result, the shape of the scatterer, which is described by using cubic-spline, can be reconstructed. In such a case, fourier series expansion will fail. Even when the initial guess is far away from the exact one, the cubic-spline description and genetic algorithm can avoid the local extreme and converge to a global extreme solution. Numerical results are given to show that the shape description using cubic-spline method is much better than the Fourier series.

Index Terms — Inverse Problem, Cubic-spline, Fourier series.

I. INTRODUCTION

Due to large domain of applications such as non-destructive problem, geophysical prospecting and determination of underground tunnels and pipelines, etc., the inverse scattering problems related to the buried bodies has a particular importance in the scattering theory. In the past 20 years, many rigorous methods have been developed to solve the exact equations [1]-[9]. However, inverse problems of this type are difficult to solve because they are ill-posed and nonlinear [10]. As a result, many inverse problems are reformulated into optimization ones and then numerically solved by different iterative methods such as the Newton-Kantorovitch method [1]-[5], the Levenberg-Marquardt algorithm [6]-[8], and the successive-overrelaxation method [9]. Most of these approaches employ the gradient-based searching scheme to find the extreme of the cost function, which are highly dependent on the initial guess and usually get trapped in the local extreme. The genetic algorithm (GA) [11] is an evolutionary algorithm that uses the stochastic mechanism to search through the parameter space. As compared to the gradient-based searching techniques, the genetic algorithm is less prone to converge to a local extreme. This renders it an ideal candidate for global optimization. Recently, researchers have applied GA together with electromagnetic solver to attack the inverse scattering problem mainly in two ways. One is surface reconstruction approach; the other is volume reconstruction approach. Chiu [12] first applied the GA for the inversion of a perfectly conducing cylinder with the geometry described by a Fourier series (surface reconstruction approach), while Takenaka [13], Meng [14] and Zhou [15] used the concept of local shape function to describe the conducting objects (volume reconstruction approach). Alternatively, Chien [16], Zhou [17] and Qing [18] used b-splines to describe the geometry of a perfect conducting cylinder. The 2-d perfectly conducting cylinders are denoted by local shape functions $\rho = F(\theta)$ with respect to their local origins, which can be continuous or discrete. However, to the best of our knowledge, there are still no numerical results, which compared the cubic-spline and Fourier series, shape description with the genetic algorithm for the buried conducting scatterers. In this paper, we present a computational method based on the genetic algorithm to recover the shape of a buried cylinder. In Section II, a theoretical formulation for the inverse scattering is presented. The general principles of genetic algorithms and the way we applied them to the inverse problem are described. Numerical results for reconstructing objects of different shapes are given in Section III. Finally, some conclusions are drawn in Section IV.

II. THEORETICAL FORMULATION

Let us consider a perfectly conducting cylinder buried in a lossy homogeneous half-space, as shown in Fig 1. The media in regions 1 and 2 are characterized by the permittivity and conductivity $(\varepsilon_1, \sigma_1)$ and $(\varepsilon_2, \sigma_2)$, respectively, while the permeability μ_0 is used for each region, i.e., only non-magnetic media are concerned here. The cross-section of the cylinder is described by

polar coordinates in the xy plane through the shape function $\rho = F(\theta)$. The cylinder is illuminated by a plane wave with time dependence $\exp(j\omega t)$, of which the electric field is assumed parallel to the z-axis (i.e., transverse magnetic or TM polarization). Let E^{inc} denote the incident E-field from region 1 to region 2 with incident angle ϕ_1. Owing to the interface between region 1 and region 2, the incident plane wave would generate two waves in the absence of the conducting object: the reflected wave (for $y \le -a$) and the transmitted wave (for $y > -a$). Thus the unperturbed field is given by

$$\vec{E}_i(\vec{r}) = E_i(x,y)\hat{z} \tag{1}$$

with

$$E_i(x,y) = \begin{cases} E_1(x,y) = e^{-jk_1[x\sin\phi_1 + (y+a)\cos\phi_1]} \\ \qquad + R_1 e^{-jk_1[x\sin\phi_1 - (y+a)\cos\phi_1]}, & y \le -a \\ E_2(x,y) = Te^{-jk_2[x\sin\phi_2 + (y+a)\cos\phi_2]}, & y > -a \end{cases}$$

where

$$R_1 = \frac{1-n}{1+n},\ T = \frac{2}{1+n},\ n = \frac{\cos\phi_2}{\cos\phi_1}\sqrt{\frac{\varepsilon_2 - j\sigma_2/\omega}{\varepsilon_1 - j\sigma_1/\omega}}$$

$$k_1 \sin\phi_1 = k_2 \sin\phi_2$$

$$k_i^2 = \omega^2\varepsilon_i\mu_0 - j\omega\mu_0\sigma_i,\ i = 1,2 \quad \mathrm{Im}(k_i) \le 0 .$$

Note that each point can be expressed by (x, y) in Cartesian coordinates or (r, θ) in polar coordinates.

As the buried object is present, the scattered field can be expressed by

$$E_s(x,y) = -\int_0^{2\pi} G(x,y;F(\theta'),\theta')J(\theta')d\theta' \tag{2}$$

with

$$\vec{r} = (x,y),\ J(\theta) = -j\omega\mu_0\sqrt{F^2(\theta) + F'^2(\theta)}J_s(\theta)$$

$$G(x,y,x',y') = \begin{cases} G_1(x,y,x',y') & , y \le -a \\ G_2(x,y,x',y') = G_f(x,y,x',y') + G_s(x,y,x',y') & , y > -a \end{cases} \tag{3}$$

where

$$G_1(x,y,x',y') = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{j}{\gamma_1 + \gamma_2} e^{j\gamma_1(y+a)} e^{-j\gamma_2(y'+a)} e^{-j\alpha(x-x')} d\alpha \tag{3a}$$

$$G_f(x,y,x',y') = \frac{j}{4} H_0^{(2)}[k_2\sqrt{(x-x')^2 + (y-y')^2}] \tag{3b}$$

$$G_s(x,y,x',y') = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{j}{2\gamma_2}\left(\frac{\gamma_2 - \gamma_1}{\gamma_2 + \gamma_1}\right) e^{-j\gamma_2(y+2a+y')} e^{-j\alpha(x-x')} d\alpha \tag{3c}$$

$$\gamma_i^2 = k_i^2 - \alpha^2,\ i = 1,2,\ \mathrm{Im}(\gamma_i) \le 0,\ y' > a .$$

Here, $J_s(\theta)$ is the induced surface current density, which is proportional to the normal derivative of the electric field on the conductor surface. $G(x, y; x', y')$ is the Green's function which can be obtained by Fourier transform [3]. Note that we might face some difficulties in calculating the Green's function, since the Green's function given by (3) takes the form of an improper integral that must be evaluated numerically. However, the integral converges very slowly when (x, y) and (x', y') approach the interface, for which the acceleration of converging speed is possible by rewriting the Green's function as a closed-form term plus a rapidly converging integral (see Appendix). In (3b), $H_0^{(2)}$ is the Hankel function of the second kind of order zero. The boundary condition for a perfectly conducting object is

$$\hat{n} \times \vec{E} = 0 \tag{4}$$

where $\hat{n}$ is the outward unit vector normal to the surface of the scatterer. The boundary condition at the surface of the scatterer given by (4) leads to an integral equation for $J(\theta)$:

$$E_2(F(\theta),\theta) = \int_0^{2\pi} G_2(F(\theta),\theta;F(\theta'),\theta')J(\theta')d\theta' . \tag{5}$$

The total field E^{out} in region 1 is given by

$$E^{out}(r) = E_1(r) - \int_0^{2\pi} G_1(r;F(\theta'),\theta')J(\theta')d\theta' \quad (y \le -a) . \tag{6}$$

The direct problem is to compute the total field in region 1 when the shape function $F(\theta)$ is given. This can be achieved by first solving for J from equation (5) and then calculating E^{out} by equation (6).

For numerical calculation of the direct problem, the contour is first divided into sufficient small segments so that the induced surface current can be considered constant over each segment. Then the moment method is used to solve equations (5) and (6) with a pulse basis function for current expansion and the Dirac delta function for testing [19].

Let us consider the following inverse problem, given the scattered electric field E_s measured outside the scatterer, and determine the shape $F(\theta)$ of the object.

(A) Using Fourier-series to describe the shape:

Assume the approximate center of the scatterer, which in fact can be any point inside the scatterer, is known. Then the shape function $F(\theta)$ can be expanded as:

$$F(\theta) \cong \sum_{n=0}^{\frac{N}{2}} B_n \cos(n\theta) + \sum_{n=1}^{\frac{N}{2}} C_n \sin(n\theta) \tag{7}$$

where B_n , and C_n , are real coefficients to be determined, and $N+1$ is the number of expanded terms.

(B) Using Cubic-spline to describe the shape:

The geometry of the cubic-spline is shown in Fig. 2. First, we separate the boundary of the shape with N pieces and we have $N+1$ separated points. We denote the separated points by polarized-coordinate

expression (ρ_0,θ_0) , (ρ_1,θ_1) ,..., (ρ_N,θ_N) , where $(0° \le \theta_i \le 360°)$, i=0…N, $\theta_0 = 0°$, $\theta_N = 360°$, and $\theta_0 < \theta_1 < ... < \theta_N$ ρ_i is the distance from point (ρ_i,θ_i) to the center point (x_0, y_0).

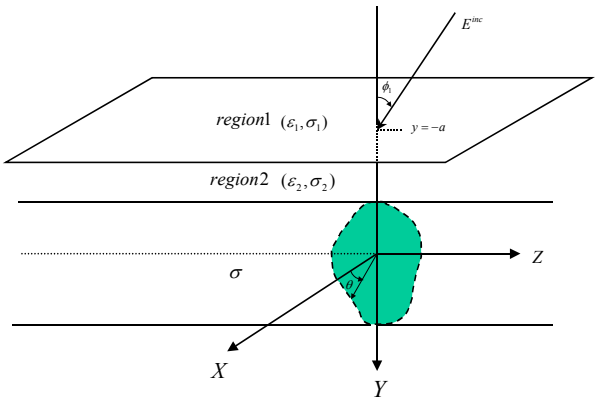


Fig. 1. Geometry of the problem in (x,y) plane.

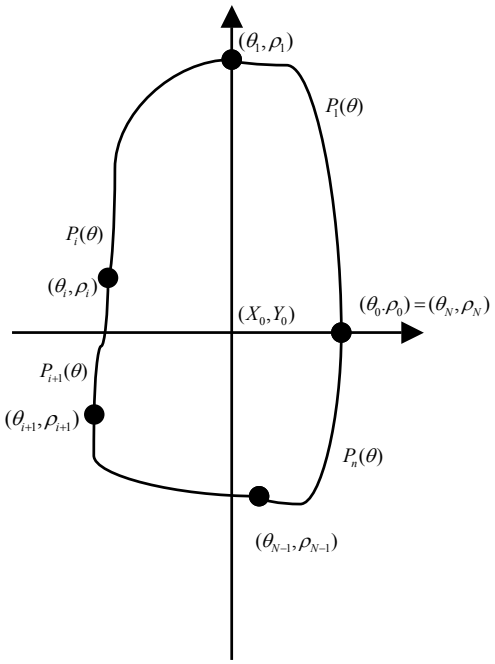


Fig. 2. Geometry of the cubic-spline.

Note that the discretization number of $J(\theta)$ for the inverse problem must be different from that for the direct problem. Since it is crucial that the synthetic data generated by a direct solver are not like those obtained by the inverse solver, the discretization number for the direct problem is twice of that for the inverse problem in this study. For the inversion procedure, the genetic algorithm is employed to maximize the following objective function:

$$SF = \left\{ \sum_{x=1}^{X_T} \left| E_s^{\exp}(\vec{r}_x) - E_s^{cal}(\vec{r}_x) \right|^2 \Big/ \sum_{x=1}^{X_T} \left| E_s^{\exp}(\vec{r}_x) \right|^2 + \beta \left| F'(\theta) \right|^2 \right\}^{-1/2} \tag{8}$$

where X_T is the total number of measured points. $E_s^{\exp}(\vec{r})$ and $E_s^{cal}(\vec{r})$ are the measured scattered field and the calculated scattered field, respectively. The minimization of $\beta|F'(\theta)|^2$ can, to a certain extent, be interpreted as the smoothness requirement for the boundary of $F(\theta)$. Therefore, the maximization of SF can be interpreted as the minimization of the least-square error between the measured and the calculated fields with the constraint of smooth boundary. Typical values of β range from 0.00001 to 10. The optimal value of β depends mostly on the dimensions of the geometry. One can always choose a large enough value to ensure the convergence, although overestimation would result in a very smooth reconstructed image. Technically, we can let the value of β decrease gradually during the course of convergence [4].

Genetic algorithms are the global optimization methods based on the genetic recombination and evolution in nature [11]. They use the iterative optimization procedures that start with a randomly generated population of potential solutions, and then gradually evolve toward a better solution through the application of the genetic operators. Genetic algorithms typically operate on a discretized and coded representation of the parameters rather than on the parameters themselves. These representations are considered as the "chromosomes", while the elements that constitute the chromosome are called "genes". Simple but often very effective chromosome representations for optimization problem involving several continuous parameters can be obtained through the juxtaposition of discretized binary representations of the individual parameters. In our problem, parameters B_n , C_n, and ρ_i are given by the following equation. As an example B_n is shown

$$B_n = p_{\min} + \frac{p_{\max} - p_{\min}}{2^L - 1} \sum_{i=0}^{L-1} b_i^{B_n} 2^i \tag{9}$$

where $b_0^{B_n}$, $b_1^{B_n}$, …and $b_{L-1}^{B_n}$ (gene) are the L-bit string of the binary representation of B_n , and P_{min} and P_{max} are the minimum and the maximum values admissible for B_n. Similar expressions exist for the parameters C_n and ρ_i and are omitted here for brevity. Here, P_{min} and P_{max} can be determined by prior knowledge of the object. Also, the finite resolution with B_n (C_n or ρ_i) can be tuned in practice by changing the number of bits

assigned to it. The basic GA for which a flowchart is shown in Fig. 3 starts with a large population containing a total of X candidates. Each candidate is described by a chromosome. Then, the initial population can simply be created by taking X random chromosomes. Then, the GA iteratively generates a new population that offspring from the previous population through the application of the reproduction, crossover, and mutation operators.

The new population contains increasingly better chromosomes and will eventually converge to a population that consists of the optimal chromosomes. As soon as the cost function (CF) changes by <1% in two successive generations or exceeds 1000 generations, the genetic algorithm will be terminated and the final solution is then obtained.

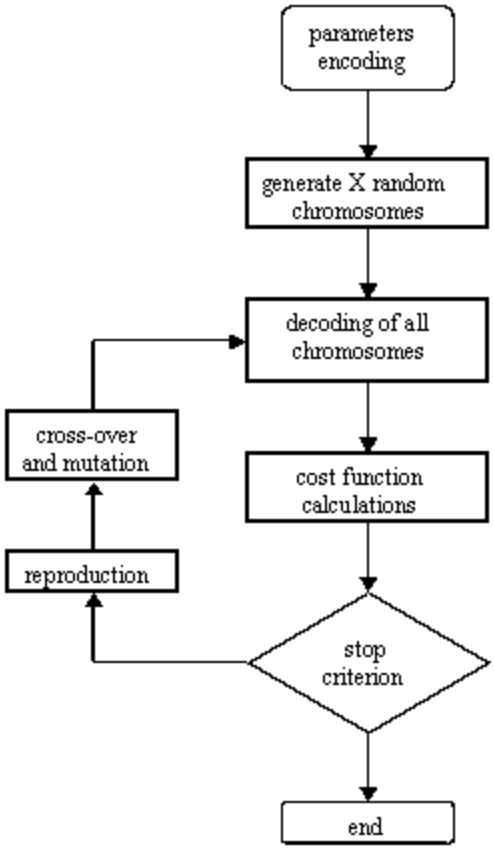


Fig. 3. The flowchart of GA.

III. NUMERICAL RESULT

Let us consider a perfectly conducting cylinder buried in a lossless half-space ($\sigma_1 = \sigma_2 = 0$). The permittivities in region 1 and region 2 are characterized by $\varepsilon_1 = \varepsilon_0$ and $\varepsilon_2 = 2.56\varepsilon_0$, respectively. A TM polarized plane wave of unit amplitude is incident from region 1 upon the object in region 2 as shown in Fig. 1. The frequency of the incident wave is chosen to be 3GHz, of which the wavelength λ_0 in free space is 0.1m. The object is buried at a depth $a \cong \lambda_0$ and the scattered fields are measured on a probing line along the interface between region 1 and region 2. Our purpose is using the Fourier-series and cubic-spline shape expressions to reconstruct the shape and comparing which is better in the inverse problem. The object is illuminated by three incident waves from different directions, while 20 measurement points at equal spacing are used along the interface $y = -a$ for each incident angle. There are 60 measurement points in each simulation. The measurement is taken from x=0 to 0.2m for incident angle $\phi_1 = -60°$, from x=-0.1 to 0.1m for incident angle $\phi_1 = 0°$, and from x=-0.2 to 0m for incident angle $\phi_1 = 60°$. To save computing time, the number of unknowns is set to be 7, and the population size is chosen as 300 (i.e. X =300). The binary string length of the unknown coefficient, B_n (C_n and ρ_l), is set to 20 bit (i.e., L=20). The search range for the unknown coefficient of the shape function is chosen to be from 0 to 0.1. The extreme values of the coefficients of the shape function can be determined by some priori knowledge of the objects. Here, the prior knowledge means that we can get the approximate position and the size of the buried cylinder by first using tomography technique, and then get the exact solution by the genetic algorithm. The crossover probability p_c and mutation probability p_m are set to be 0.8 and 0.1, respectively. The value of β in equation (6) is chosen to be 0.001.

In the first example, the shape function is given by $F(\theta) = (0.03 + 0.015\cos 3\theta)$ m and we use Fourier-series and cubic-spline expressions to recover it. The reconstructed shape function for the best population member (chromosome) is plotted in Fig. 4(a), respectively, with the error shown in Fig. 4(b). Here, DR, which is called shape function discrepancies, is defined as

$$DR = \{\frac{1}{N'}\sum_{i=1}^{N'}[F^{cal}(\theta_i) - F(\theta_i)]^2 / F^2(\theta_i)\}^{1/2} \qquad (10)$$

where N' is set to 60. The quantities DR provide measures of how well $F^{cal}(\theta)$ approximates $F(\theta)$. From Fig. 4, it is clear that reconstruction of the shape function is quite good for both Fourier-series and cubic-spline expressions. To investigate the sensitivity of the imaging algorithm against random noise, two independent Gaussian noises with zero mean have been added to the real and imaginary parts of the simulated scattered fields. Normalized standard deviations of 10^{-5} , 10^{-4} , 10^{-3} , 10^{-2} , and 10^{-1} are used in the

simulations. The normalized standard deviation mentioned earlier is defined as the standard deviation of the Gaussian noise divided by the rms value of the scattered fields. Here, the signal-to-noise ratio (SNR) is inversely proportional to the normalized standard deviation. The numerical result for this example is plotted in Fig. 4(b).

In the second example, we selected cubic-spline to describe the shape $\rho_1 = 0.02$ m, $\rho_2 = 0.02\sqrt{3}$ m, $\rho_3 = 0.02\sqrt{3}$ m, $\rho_4 = 0.02$ m, $\rho_5 = 0.02\sqrt{3}$ m, $\rho_6 = 0.02\sqrt{3}$ m. We can see that the 7-terms Fourier-series expression cannot recover the shape. The purpose of this example is to show that cubic-spline method is able to reconstruct a scatter while the Fourier-series fails. Both the shape results are shown in Fig. 5(a) and the relative error of Cubic-spline expand is shown in Fig. 5(b).

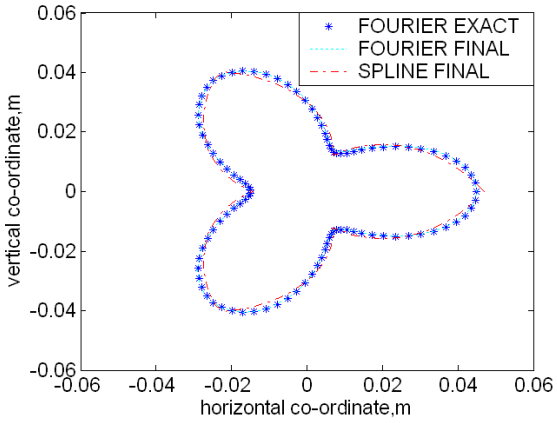


Fig. 4(a). Shape function for example 1. The star curve represents the exact shape by the Fourier-series, while the curve of short imaginary line is calculated shape by the Fourier-series and the curve of long imaginary line represents calculated shape by the cubic-spline in final result.

In the third example, the shape and conductivity function are selected to be $F(\theta) = (0.03 + 0.005\cos\theta + 0.005\cos 2\theta + 0.005\cos 3\theta)$ m. Note that the shape function is not symmetrical about either x-axis or y-axis. Both Fourier-series and cubic-spline expressions can recover it. Refer to Fig. 6(a) and Fig. 6(b) for details.

In the fourth example, we selected cubic-spline to describe the shape ρ_0 = 0.03 m, ρ_1 = 0.02 m, ρ_2 = 0.01 m, ρ_3 = 0.01 m, ρ_4 = 0.01 m, ρ_5 = 0.03 m, and slope is 5. Again 7-terms Fourier series expression cannot recover the shape. Both the shape results are shown in Fig. 7(a) and the relative error of Cubic-spline expand is shown in Fig. 7(b).

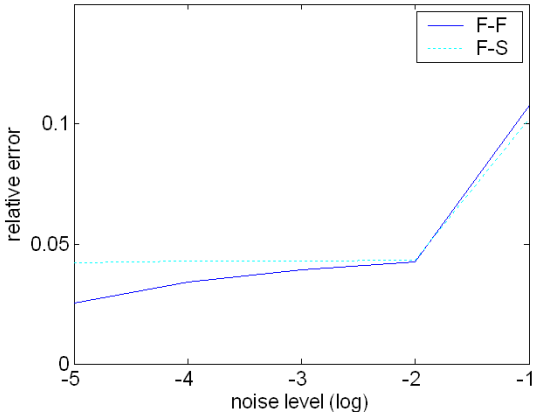


Fig. 4(b). Shape function error in each represented method. The F-F means that the shape functions both in direct and inverse problems are described by the Fourier-series. The F-S means that the shape function in the direct problem is described by the Fourier-series and in the inverse problem is described by the cubic-spline.

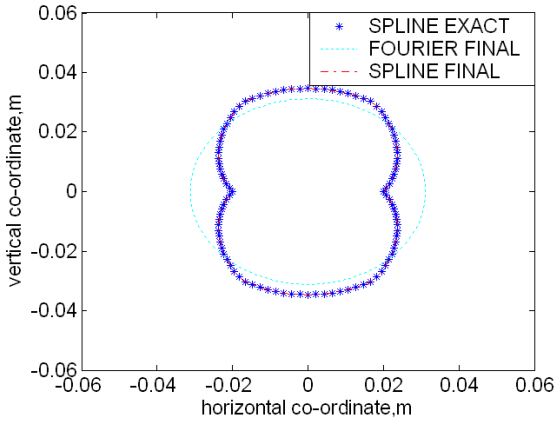


Fig. 5(a). Shape function for example 2. The star curve represents the exact shape by the Fourier-series, while the curve of short imaginary line is calculated shape by the Fourier-series and the curve of long imaginary line represents calculated shape by the cubic-spline in final result.

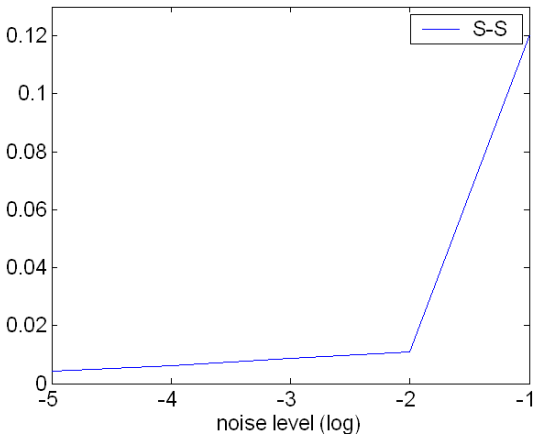


Fig. 5(b). Shape function error in each represented method. The S-S means that the shapefunctions both in the direct and inverse problem are described by the cubic-spline.

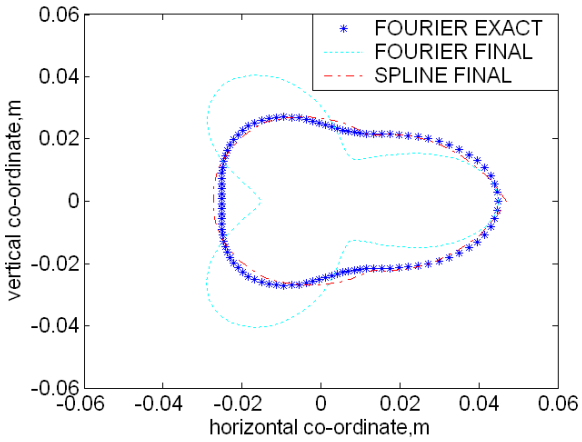


Fig. 6(a). Shape function for example 3. The star curve represents the exact shape by the cubic-spline, while the curve of short imaginary line is calculated shape by the Fourier-series and the curve of long imaginary line represents calculated shape by the cubic-spline in final result.

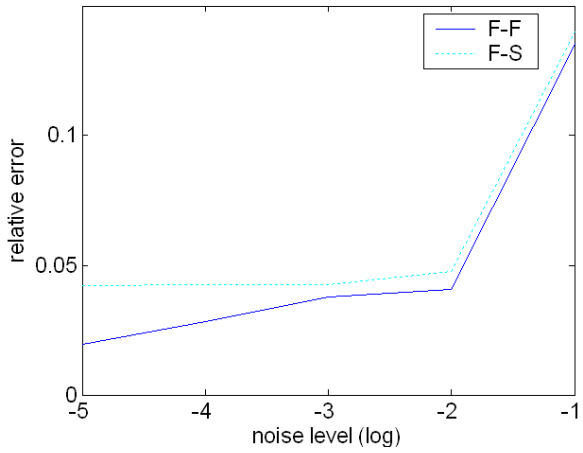


Fig. 6(b). Shape function error in each represented method. The F-F means that the shape functions both in the direct and inverse problems are described by the Fourier-series. The F-S means that the shape function in the direct problem is described by the Fourier-series and in the inverse problem is described by the cubic-spline.

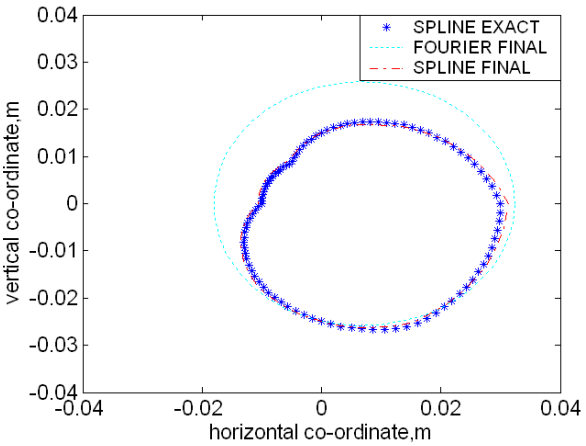


Fig. 7(a). Shape function for example 4. The star curve represents the Fourier-series exact shape, while the curve of short imaginary line is the Fourier-series calculated shape and the curve of long imaginary line is represents the cubic-spline calculated shape in final result.

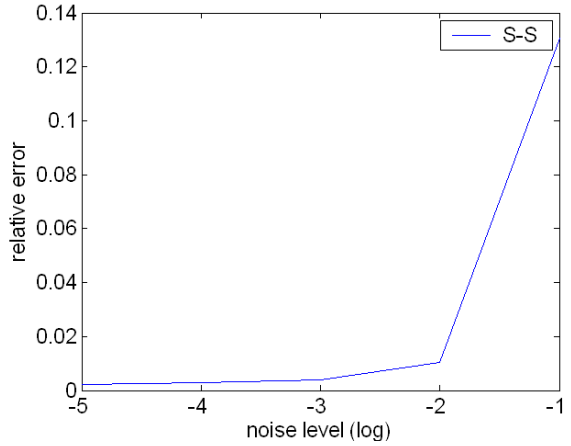


Fig. 7(b). Shape function error in each represented method. The S-S means that the shape functions both in the direct and inverse problems are described by the cubic-spline.

IV. CONCLUSION

We have presented a study of applying the genetic algorithm to reconstruct the shape of a buried metallic object through the measurements of scattered E fields. Based on the boundary condition and the measured scattered fields, we have derived a set of nonlinear integral equations and reformulated the imaging problem into an optimization one. The contours of the cylinders are denoted by cubic-spline local shape functions in local polar coordinate instead of trigonometric series local functions to guarantee the nonnegative definiteness. Experiment results show that the variable searching ability of GA has its limitation, and Fourier-series expression cannot recover the arbitrary shape in finite terms. In our numerical results, it is shown that using cubic-spline expand to describe the shape in the half-space inverse problem is more suitable than Fourier-series expression.

ACKNOWLEDGEMENT

This work was supported by National Science Council, Republic of China, under Grant NSC-91-2219-E-032-004.

Autobiography

Wei Chien was born in Taipei, Taiwan, Republic of China, on April 21, 1974. He received the B.S.E.E. degree from Tatung Institute of Technology, Taipei, Taiwan, in 1997 and M.S.E.E. degree from Tamkang University, Taipei, Taiwan, in 1999. From 1999 to 2001, he served in the ROC Army Force as a second lieutenant. He is working toward Ph.D degrees in the Department of Electrical Engineering, Tamkang University, now. His current research interests include numerical techniques in electromagnetics and wireless communications.

Chien-Ching Chiu was born in Taoyuan, Taiwan, Republic of China, on January 23, 1963. He received the B.S.C.E. degree from National Chiao Tung University, Hsinchu, Taiwan, in 1985 and M.S.E.E. and ph.D degrees from National Taiwan University, Taipei, Taiwan, in 1987 and 1991 respectively. From 1987 to1989, he served in the ROC Army Force as a communication officer. In 1992 he joined the faculty of the Department of Electrical Engineering, Tamkang University, where he is now a Professor. He was a visiting scholar at the MIT and University of Illinois, Urbana from 1998 to 1999. His current research interests include microwave imaging, numerical techniques in electromagnetics and indoors wireless communications.

APPENDIX

To calculate the Green's function, we can use the following formula,

$$\int_u^\infty x^{r-1}e^{-\beta x}\cos\delta x dx = \frac{1}{2}(\beta+j\alpha)^{-r}\Gamma[r,(\beta+j\delta)u]+\frac{1}{2}(\beta-j\alpha)^{-r}\Gamma[r,(\beta-j\delta)u]$$

for $\mathrm{Re}\,\beta > |\mathrm{Im}\,\delta|$ (A1)

where $\Gamma(\alpha,Z)=\int_z^\infty e^{-t}t^{\alpha-1}dt$

Γ is the incomplete Gamma function that satisfies

$$\Gamma(-n,z)=\frac{(-1)^n}{n!}\left[\Gamma(0,Z)-e^{-z}\sum_{m=0}^{n-1}(-1)^m\frac{m!}{z^{m+1}}\right]$$

$$\Gamma(0,z)=-\gamma-\ln z-\sum_{n=1}^{\infty}(-1)^n\frac{z^n}{(n+1)!}\qquad [|\arg(z)|<\pi]$$

(A2)

γ is Euler's constant, i.e., γ =0.5772156649. Let us consider the following integral

$$G_1=\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{j}{\gamma_1+\gamma_2}e^{j\gamma_1(y+a)}e^{-j\gamma_2(y'+a)}e^{-j\alpha(x-x')}d\alpha$$

$$=\frac{1}{\pi}\int_0^{\infty}\frac{j}{\gamma_1+\gamma_2}e^{j\gamma_1(y+a)}e^{-j\gamma_2(y'+a)}\cos\alpha\,(x-x')\,d\alpha$$

where $\gamma_i^2=k_i^2-\alpha^2, i=1,2,3$ $Im(\gamma_i)\le 0$, $y'>-a$.

The integral G_1 may be rewritten as follows

$$G_1=\frac{1}{\pi}\int_0^{\infty}\frac{j}{\gamma_1+\gamma_2}e^{j\gamma_1(y+a)}e^{-j\gamma_2(y'+a)}\cos\alpha(x-x')d\alpha$$

$$+\frac{1}{2\pi}\int_{\alpha_0}^{\infty}\frac{1}{\alpha}e^{-\alpha(y'-y)}\cos\alpha(x-x')d\alpha$$

$$-\frac{1}{2\pi}\int_{\alpha_0}^{\infty}\frac{1}{\alpha}e^{-\alpha(y'-y)}\cos\alpha(x-x')d\alpha$$

In general, we choose $\alpha_0 >> |k_i|$, $i = 1, 2$. From Eq. (A1), we get

$$-\frac{1}{2\pi}\int_{\alpha_0}^{\infty}\frac{1}{\alpha}e^{-\alpha(y'-y)}\cos\alpha(x-x')d\alpha$$
$$=-\frac{1}{4\pi}\left\{\Gamma\left[0,\left[(y-y')+j(x-x')\right]\alpha_0\right]+\Gamma\left[0,\left[(y-y')-j(x-x')\right]\alpha_0\right]\right\}.$$

Using the above relation, we obtain

$$G_1=\frac{1}{\pi}\int_0^{\infty}\frac{j}{\gamma_1+\gamma_2}e^{j\gamma_1(y+a)}e^{-j\gamma_2(y'+a)}\cos\alpha(x-x')d\alpha$$
$$+\frac{1}{2\pi}\int_{\alpha_0}^{\infty}\frac{1}{\alpha}e^{-\alpha(y'-y)}\cos\alpha(x-x')d\alpha$$
$$-\frac{1}{4\pi}\left\{\Gamma\left[0,\left[(y-y')+j(x-x')\right]\alpha_0\right]+\Gamma\left[0,\left[(y-y')-j(x-x')\right]\alpha_0\right]\right\}. \quad \text{(A3)}$$

Now, the integral G_1 is written as a rapidly converging integral plus a dominant integral, which can be easily calculated by means of Simpson's rule. Similarly, we have

$$G_s=\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{J}{2\gamma_2}\left(\frac{\gamma_1-\gamma_2}{\gamma_1+\gamma_2}\right)e^{-j\gamma_2(y+2a+y')}e^{-j\alpha(x-x')}d\alpha$$
$$=\begin{bmatrix}\frac{1}{\pi}\int_0^{\infty}\frac{j}{2\gamma_2}\left(\frac{\gamma_2-\gamma_1}{\gamma_2+\gamma_1}\right)e^{-j\gamma_2(y+2a+y')}\cos\alpha(x-x')d\alpha\\ -\frac{k_2^2-k_1^2}{8\pi}\int_{\alpha_0}^{\infty}\frac{1}{\alpha^3}e^{-\alpha(y+2a+y')}\cos\alpha(x-x')d\alpha\end{bmatrix}$$
$$+\frac{k_2^2-k_1^2}{16\pi}\begin{Bmatrix}\left[(y+2a+y')+j(x-x')\right]^2\Gamma\times\\ \left[-2,\left[(y+2a+y')+j(x-x')\right]\alpha_0\right]+\\ \left[(y+2a+y')-j(x-x')\right]^2\Gamma\times\\ \left[-2,\left[(y+2a+y')-j(x-x')\right]\alpha_0\right]\end{Bmatrix}.$$

References

[1] A. Roger, "Newton-Kantorovitch algorithm applied to an electromagnetic inverse problem," *IEEE Trans. Antennas Propagat*., vol. AP-29, pp. 232-238, Mar. 1981.

[2] C. C. Chiu and y. W. Kiang, "inverse scattering of a buried conducting cylinder,"*Inverse Problems*, vol. 7, pp. 187-202, April 1991.

[3] C. C. Chiu and y. W. Kiang, "microwave imaging of multiple conducting cylinders," *IEEE Trans. Antennas propagat*., vol. 40, pp. 933-941, Aug. 1992.

[4] G. P. Otto and W. C. Chew, "Microwave inverse scattering-local shape function imaging for improved resolution of strong scatterers." *IEEE Trans. Microwave Theory Tech*., vol. 42, pp. 137-142, Jan. 1994.

[5] Kress R. "A Newton method in inverse obstacle scattering Inverse Problem in Engineering Mechanics," ed H D Bui et al (Rotterdam: Balkema), pp. 425-432, 1994.

[6] D. Colton and P. Monk, "A novel method for solving the inverse scattering problem for time-harmonic acoustic waves in the resonance region II," *SIAM J. Appl. Math*., vol. 46, pp. 506-523, June 1986.

[7] A. Kirsch, R. Kress, P. Monk, and A. Zinn, "Two methods for solving the inverse acoustic scattering problem, " *Inverse Problems*, vol. 4, pp. 749-770, Aug. 1998.

[8] F. Hettlich, "Two methods for solving an inverse conductive scattering problem," *Inverse Problems*, vol. 10, pp. 375-385, 1994.

[9] R. E. Kleinman and P. M. Van Den Berg, "Two-dimensional location and shape reconstruction," *Radio Sci.*, vol. 29, pp. 1157-1169, July-Aug. 1994.

[10] T. Hohage, "Iterative methods in inverse obstacle scattering: regularization theory of linear and nonlinear exponentially ill-posed problems," Dissertation Linz, 1999.

[11] D. E. Goldgreg, Genetic Algorithm in Search, Optimization and Machine Learning, Addison-Wesley, 1989.

[12] C. C. Chiu and P. T. Liu, "Image reconstruction of a perfectly conducting cylinder by the genetic algorithm," *IEE Proc.-Micro. Antennas Propagat.*, vol. 143, pp. 249-253, June 1996.

[13] T. Takenaka and Z. Q. Meng, T. Tanaka, W. C. Chew "Local shape function combined with genetic algorithm applied to inverse scattering for strips", *Microwave and Optical Technology Letters,* vol. 16, pp. 337-341, Dec. 1997.

[14] Z. Q. Meng, T. Takenaka and T. Tanaka, "Image reconstruction of two-dimensional impenetrable objects using genetic algorithm", *Journal of Electromagnetic Waves and Applications,* vol. 13, pp. 95-118, 1999.

[15] Y. Zhou and H. Ling "Electromagnetic inversion of Ipswich objects with the use of the genetic algorithm", *Microwave and Optical Technology Letters,* vol. 33, pp. 457-459, June 2002.

[16] W. Chien, "Using the Genetic Algorithm to reconstruct the two-dimensional conductor" Master Thesis, National Tamkang University, Department of Electrical Engineering, June 1999.

[17] Y. Zhou, J. Li and H. Ling; "Shape inversion of metallic cavities using hybrid genetic algorithm combined with tabu list", *Electronics Letters,* vol. 39, pp. 280 -281, Feb. 2003.

[18] A. Qing "An experimental study on electromagnetic inverse scattering of a perfectly conducting cylinder by using the real-coded genetic algorithm", *Microwave and Optical Technology Letters,* vol. 30, pp. 315-320, Sept. 200.1

[19] F. M. Tesche, "On the inclusion of loss in time domain solutions of electromagnetic interaction problems," *IEEE Trans. Electromagn. Compat.*, vol. 32, pp. 1-4, 1990.

[20] E. C. Jordan and K. G. Balmain, Electromagnetic Waves and Radiating systems. Englewood Cliffs, NJ: Prentice-Hall, 1968.

[21] S. Nakamura, "Applied Numerical Method in C," *Prentice-Hall int*. 1993.

[22] "A pratical Guide to Splines," *New York: Spring-Verlag*, 1987.

On the Physical Interpretation of the Sobolev Norm in Error Estimation

Clayton P. Davis*and Karl F. Warnick†
*Electromagnetics Laboratory
The University of Illinois at Urbana-Champaign
1406 W. Green Street, Urbana, IL 61801
Email: cpdavis2@uiuc.edu
† Department of Electrical and Computer Engineering
Brigham Young University, Provo, UT 84602
Email: warnick@ee.byu.edu

***Abstract*— Error estimates for the moment method have been obtained in terms of Sobolev norms of the current solution. Motivated by the historical origins of Sobolev spaces as energy spaces, we show that the Sobolev norm used in these estimates is related to the forward scattering amplitude, for the case of 2D scattering from a PEC circular cylinder and for 3D scattering from a PEC sphere. These results provide a physical meaning for solution error estimates in terms of the power radiated by the error in the current solution. We further show that bounds on the Sobolev norm of the current error imply a bound on the error in the computed backscattering amplitude.**

***Index Terms*— Sobolev space, error analysis, method of moments**

I. Introduction

Since the introduction of the method of moments for solving electromagnetic radiation and scattering problems, error analysis of numerical methods has received much attention in the mathematics literature. This effort has led to fundamental results on the convergence of the method of moments. Typical of this work are proofs that under various assumptions about the algorithm and scattering problem, as the mesh is refined or the number of degrees of freedom of the approximate solution increases, numerical solutions converge to exact solutions. Theorems of this kind have been obtained for 2D smooth closed curves and screens [1], dielectric polygons [2], and have been verified by numerical studies [3]. For smooth screens in 3D, similar results are available for scalar fields [2], [4], [5]. These results are of great importance because they place the algorithms of computational electromagnetics on solid theoretical ground.

The approach taken in this work by the numerical analysis community is to place the integral operators of radiation and scattering in a Sobolev space setting. This leads to asymptotic solution error estimates of the form

$$\|\Delta u\|_{\mathcal{H}^s} \leq Ch^r \tag{1}$$

where the norm is defined on the Sobolev space $\mathcal{H}^s$, with $s = -1/2$ for the TM polarization and $s = 1/2$ for TE [6]. Δu is the difference between the exact current solution and a numerical solution, and h is the mesh element width or discretization length. The convergence rate r is typically $1/2$ for low order basis functions. All dependence on the physical problem and implementation details of the numerical method, including the incident field, frequency, scatterer geometry, and choice of basis functions, is lumped into the unknown constant C.

While the estimate (1) shows in an abstract sense that a numerical solution converges as the discretization length becomes small, it cannot be used to determine the error in a specific numerical solution because the Sobolev norm $\|\cdot\|_{\mathcal{H}^s}$ can be difficult to compute [7] and the constant C is unknown. Furthermore, it is not obvious how the Sobolev norm may relate to a directly measurable, physical quantity.

Motivated by the historical origin of Sobolev spaces as energy spaces, we show in this paper that the Sobolev norm in Eq. (1) is related to a readily computable, physical quantity: the power supplied by a surface current to its surroundings. Heuristically, a Sobolev space for fields in a volumetric region consists of those functions which have finite energy, where the energy measure is induced by a particular partial differential equation. Sobolev spaces of surface currents are defined slightly differently, as they consist of functions on the surface that radiate finite energy [7]. This definition is motivated by Poynting's theorem,

$$\int_{\mathcal{S}} \boldsymbol{E}^* \cdot \boldsymbol{J}_s \, d\mathcal{S} = \frac{i\omega}{2} \int_V \epsilon |\boldsymbol{E}|^2 - \mu |\boldsymbol{H}|^2 \, dV - \oint_{\partial V} \boldsymbol{S}^* \cdot \hat{n} \, dV$$

where the terms are defined as is usual in electromagnetic theory. The Sobolev space of fields $\boldsymbol{E}$ and $\boldsymbol{H}$ is essentially defined by requiring that the volume integral on the right-hand side be finite. In order to obtain consistent function spaces for fields and surface currents, at least nonrigorously, the Sobolev space of surface currents should include all functions on S for which the left-hand side is finite. If the surface current is produced by an incident field illuminating a PEC scatterer, then the left-hand side of Poynting's theorem with a suitable normalization becomes the forward scattering amplitude of the scatterer. This suggests a connection between the Sobolev norm in (1) and the forward scattering amplitude.

Based on this connection, we derive a direct relationship between the forward scattering amplitude and the Sobolev norm $\|\cdot\|_{\mathcal{H}^s}$. Proofs of the result are given for the specific

cases of the circular cylinder and sphere, and we conjecture that similar relationships hold for more general geometries. This relationship between the Sobolev norm and the forward scattering amplitude is used to provide a physical interpretation for error estimates of the form of (1). We further show that the bound (1) implies a bound on the error in the computed backscattering amplitude solution. These results provide a link between abstract results of numerical analysis and the physical quantities that are the desired results of practical CEM simulations. Some of the results in this paper for the 2D case were presented in [8].

II. PEC INFINITE CIRCULAR CYLINDER

For a plane wave incident in the $-x$ direction on a PEC circular cylinder, the induced current u may be written as a Fourier series, $u = (2\pi)^{-1/2}\sum_q U_q e^{iq\phi}$, where ϕ is the azimuthal angle and the Fourier coefficients are given by

$$U_q = \sqrt{\frac{2}{\pi}}\frac{2}{\eta ka}\frac{i^{-q}}{H_q^{(1)}(ka)} \tag{2}$$

for the TM case, and the TE case is identical by replacing the Hankel function $H_q^{(1)}(ka)$ with its derivative. Here, η is the characteristic impedance of free space, k is the wavenumber, and a is the cylinder radius.

In general, the currents induced on 2D PEC scatterers lie in fractional order Sobolev spaces, where the order s is $-1/2$ for the TM polarization and $1/2$ for the TE polarization. For closed surfaces, the Sobolev norm is computable and is given by [9], [10] as

$$\|u\|_{\mathcal{H}^s}^2 = \sum_q |U_q|^2\left(1+q^2\right)^s. \tag{3}$$

When s is an integer, the Sobolev norm (3) reduces to the more usual definition in terms of the L^2 norm of the current and its derivatives up to order s. For example, for $s = 0$, $\|u\|^2 = \sum_q |U_q|^2$, which by Parseval's relation is the L^2 norm of the current. The relationship between Eq. (3) for fractional s and a physical quantity is not immediately apparent.

Extrapolating the foundational relationship between Sobolev spaces and physical energy, we will show that the Sobolev norm (3) is equivalent to the forward scattering amplitude, which is given by

$$P(u) = -\frac{k\eta}{4}\int_{\mathcal{C}} E^{s*}u\,dl = \frac{k\eta}{4}\int_{\mathcal{C}}(\mathcal{L}u)^* u\,dl, \tag{4}$$

where E^s is the tangential component of the scattered field and $\mathcal{L}$ is the EFIE operator so that $\mathcal{L}u = E^i$. Note that P is the left-hand side of Poynting's theorem (2) scaled by $-k\eta/4$. If E^i is a plane wave, P is the power scattered in the direction of the plane wave. Otherwise, it may be viewed as a generalized forward scattering amplitude. It will be convenient to express the forward scattering amplitude in series form. This may be done for an arbitrary current u by decomposing the Green's function in $\mathcal{L}$ as a sum over Bessel functions [11]:

$$\mathcal{L}u = \frac{2}{\pi ka\eta}\sum_q {\alpha_q}^*\int_0^{2\pi} u(\phi')e^{iq(\phi-\phi')}d\phi' \tag{5}$$

where the coefficients are given by

$$\alpha_q = \frac{\pi(ka\eta)^2}{8}\times\begin{cases} J_q(ka)H_q^{(2)}(ka) & \text{TM} \\ J_q'(ka)H_q^{(2)\prime}(ka) & \text{TE}\end{cases}, \tag{6}$$

and $J_q(ka)$ is the usual Bessel function. Substituting Eq. (5) into Eq. (4) yields

$$P(u) = \sum_q \alpha_q |U_q|^2. \tag{7}$$

To establish a rigorous relationship between the Sobolev norm and the forward scattering amplitude, we will use the notion of equivalent norms. If two norms are equivalent, then if $x \to 0$ in either norm, then it will vanish in both norms. Formally, two norms $\|\cdot\|$ and $\|\cdot\|'$, defined for the same space X, are said to be equivalent if there exists constants $c_1, c_2 > 0$ such that

$$c_1\|x\| \le \|x\|' \le c_2\|x\| \tag{8}$$

for every $x \in X$. We will also need the definition of a quasinorm. A quasinorm is a functional with the following properties:

1) $\|x\| \ge 0$ with equality iff x is everywhere 0.
2) $\|\alpha x\| = |\alpha|\|x\|$ for all $\alpha \in \mathbb{C}$.
3) $\|x_1 + x_2\| \le K(\|x_1\| + \|x_2\|)$ for all x_1, $x_2 \in X$ and for some $K \ge 1$.

A quasinorm differs from a norm in that for a norm, we have $K = 1$. We will prove that the quantity $\|\cdot\|_P \equiv \sqrt{|P(\cdot)|}$ is a quasinorm and is equivalent to the Sobolev norm (extending the notion of equivalency to include quasinorms).

To prove that $\|\cdot\|_P$ is a quasinorm, all three properties above must be shown. Property (1) is satisfied by further stipulating that there are no internal resonant modes, i.e., $\alpha_q \neq 0$ for all q. This is equivalent to saying that the interior Dirichlet (TM) or Neumann (TE) problem does not have a non-trivial solution. Satisfaction of property (2) is seen by substituting αU_q in for U_q in the scattering amplitude expression (7). Since $\|u\|_{\mathcal{H}^s}$ is a norm, it satisfies property (3) above with $K = 1$. Using this with the equivalency statement (8) yields the inequality

$$\|u_1 + u_2\|_P \le \frac{c_2}{c_1}\left(\|u_1\|_P + \|u_2\|_P\right). \tag{9}$$

This proves property (3) and $\|\cdot\|_P$ is a quasinorm. If $\mathcal{L}$ were self-adjoint, then $\|\cdot\|_P$ would be a norm. This is a minor point, since this paper relies on the equivalency relationship (8) to relate the Sobolev norm to a physical quantity, particularly in the sense as the Sobolev norm of an error current vanishes. The properties of norms and quasinorms are not used, except to couch the problem in a more familiar framework.

It remains to find constants c_1 and c_2 that satisfy Eq. (8) with $\|\cdot\| = \|\cdot\|_P$ and $\|\cdot\|' = \|\cdot\|_{\mathcal{H}^s}$. The constant c_1 is found by directly comparing the $\|\cdot\|_P$ norm with the Sobolev norm (3) term by term. This yields

$$c_1 = \left[\max_q |\beta_q|\right]^{-1/2} \approx \frac{2.5}{\eta}(ka)^{(4s-5)/6}, \tag{10}$$

where $\beta_q \equiv \alpha_q(1+q^2)^{-s}$. The approximation was made analytically using results of [12] for the TM case ($s = 1/2$)

and extended numerically to the TE case ($s = -1/2$), where estimating $\max_q |\beta_q|$ is more difficult.

A constant c_2 satisfying Eq. (8) can be derived by first classifying as low order modes those terms in (7) for which $|q| < q_0$, where q_0 is a positive integer to be specified. Splitting the sum in the scattering amplitude (7) into high and low order modes and also into real and imaginary parts, we define

$$\begin{aligned} R_l &\equiv \sum_{|q|<q_0} \mathrm{Re}(\alpha_q)|U_q|^2 \quad I_l \equiv \sum_{|q|<q_0} \mathrm{Im}(\alpha_q)|U_q|^2 \\ R_h &\equiv \sum_{|q|\geq q_0} \mathrm{Re}(\alpha_q)|U_q|^2 \quad I_h \equiv \sum_{|q|\geq q_0} \mathrm{Im}(\alpha_q)|U_q|^2 \end{aligned} \tag{11}$$

We then rewrite the forward scattering amplitude (7) as

$$P(u) = R_l + R_h + i(I_l + I_h). \tag{12}$$

To compare the forward scattering amplitude to the Sobolev norm term by term, we will compare the low order terms of the Sobolev norm (3) with R_l since it allows us to guarantee that c_2 is finite except at resonance frequencies. We obtain the following relationship:

$$|R_l| \geq r_l \sum_{q:|q|<q_0} |U_q|^2(1+q^2)^s, \tag{13}$$

where $r_l = \min_{q:|q|<q_0} \mathrm{Re}(\beta_q)$. We can compare the high order terms of the Sobolev norm to I_h, since they both fall off at the same rate in q. Asymptotic expansions of $\mathrm{Im}(\alpha_q)$ show that the integer q_0 can be chosen sufficiently large such that the α_q have the same sign for all $|q| > q_0$. This allows us to bring the absolute value operator inside the summation in the definition of I_h and derive the relationship

$$|I_h| = \sum_{|q|\geq q_0} |\mathrm{Im}(\alpha_q)||U_q|^2 \geq i_h \sum_{|q|\geq q_0} |U_q|^2(1+q^2)^s, \tag{14}$$

where $i_h = \min_{q:|q|\geq q_0} |\mathrm{Im}(\beta_q)|$. Since R_h and I_l are extraneous, we will discard them. By the definition of α_q, R_l and R_h have the same sign and R_h can be immediately eliminated from $P(u)$ to give the lower bound $|P(u)| \geq |R_l + i(I_l + I_h)|$. We can remove I_l by noting that

$$|I_l| \leq \max_{q:|q|<q_0} \left|\frac{\mathrm{Im}(\beta_q)}{\mathrm{Re}(\beta_q)}\right| |R_l| \leq M|R_l|, \tag{15}$$

where M is a constant given by

$$M \equiv \max\left(1, \max_{q:|q|<q_0} \left|\frac{\mathrm{Im}(\beta_q)}{\mathrm{Re}(\beta_q)}\right|\right). \tag{16}$$

We have defined M to guarantee that it satisfies $M \geq 1$. This allows us to apply inequality (43) from the appendix, yielding

$$|P(u)| \geq \frac{1}{3M}\left(|R_l| + |I_h|\right). \tag{17}$$

Substituting in the term by term comparisons (13) and (14) yields

$$3M|P(u)| \quad \geq \quad \min(r_l, i_h) \sum_q |U_q|^2(1+q^2)^s. \tag{18}$$

Simplifying and taking the square root of each side, gives $\|u\|_{\mathcal{H}^s} \leq c_2\|u\|_P$, where

$$c_2 = \sqrt{\frac{3M}{\min(r_l, i_h)}}. \tag{19}$$

Note that c_2 depends only on ka and not on the current u, as required. It can be proved from the definition of the α_q that c_2 is finite, except at resonance frequencies. This is consistent with Eq. (8) where, as ka approaches a resonance, $\|u\|_P$ vanishes if u is a resonant mode, but $\|u\|_{\mathcal{H}^s}$ does not. Thus, if $\|u\|_{\mathcal{H}^s} \leq c_2\|u\|_P$ is to be maintained, we must have $c_2 \to \infty$ at these frequencies. We have thus obtained

$$c_1\|u\|_P \leq \|u\|_{\mathcal{H}^s} \leq c_2\|u\|_P \tag{20}$$

which relates the Sobolev norm of a current to the forward scattering amplitude, a physically meaningful quantity.

III. PEC SPHERE

We now derive a similar relationship between a 3D Sobolev norm and the forward scattering amplitude for scattering from a PEC sphere. In general, any function tangential to a surface may be expressed in terms of its surface Helmholtz decomposition

$$\boldsymbol{J} = \boldsymbol{J}^{cf} + \boldsymbol{J}^{df} \tag{21}$$

where $\boldsymbol{J}^{cf}$ is curl-free (irrotational) and $\boldsymbol{J}^{df}$ is divergence-free (solenoidal). In [7], it is shown that $\boldsymbol{J}^{cf}$ and $\boldsymbol{J}^{df}$ on a sphere can be expanded as

$$\boldsymbol{J}^{cf} = \nabla^t \sum_{n=1}^{\infty}\sum_{m=-n}^{n} d_{mn}^{1/2} u_{nm}^{cf} P_n^{|m|}(\cos\theta) e^{im\phi} \tag{22}$$

and

$$\boldsymbol{J}^{df} = \hat{n} \times \nabla^t \sum_{n=1}^{\infty}\sum_{m=-n}^{n} d_{mn}^{1/2} u_{nm}^{df} P_n^{|m|}(\cos\theta) e^{im\phi}, \tag{23}$$

where the normalizing factor is

$$d_{mn} = \frac{(n-|m|)!\,(2n+1)}{(n+|m|)!\,4\pi n(n+1)}. \tag{24}$$

Here, the $P_n^{|m|}(\cdot)$ is the associated Legendre function of the first kind ∇^t is the surface gradient. A Sobolev space for currents on 3D bodies is denoted by $H_{div}^{-1/2}$ and for a sphere the norm is given by

$$\|\boldsymbol{J}\|^2_{H_{div}^{-1/2}} \equiv \sum_{n=1}^{\infty} \left[U_n^{cf}(1+n^2)^{1/2} + U_n^{df}(1+n^2)^{-1/2}\right], \tag{25}$$

where

$$U_n^{cf} = \sum_{m=-n}^{n} |u_{nm}^{cf}|^2 \,,\; U_n^{df} = \sum_{m=-n}^{n} |u_{nm}^{df}|^2. \tag{26}$$

Using the orthogonality relationships in [7, Sec. VIII], it can be shown that the forward scattering amplitude decomposes as

$$P(\boldsymbol{J}) = P^{cf}(\boldsymbol{J}^{cf}) + P^{df}(\boldsymbol{J}^{df}), \tag{27}$$

where

$$P^{cf} = \sum_{n=1}^{\infty} \alpha_n^{cf} U_n^{cf} \ , \ P^{df} = \sum_{n=1}^{\infty} \alpha_n^{df} U_n^{df}. \qquad (28)$$

Here, P^{cf} is the forward scattering amplitude due to the curl-free component of the current and P^{df} is similarly defined for the divergence-free component. Note the similarity of Eq. (28) to the 2D expression (7). In these expressions,

$$\alpha_n^{cf} = -\frac{i(k\eta)^2}{4\pi}[ka\ j_n(ka)]'[ka\ h_n^{(2)}(ka)]' \qquad (29)$$
$$\alpha_n^{df} = -\frac{i(k\eta)^2}{4\pi}[ka\ j_n(ka)][ka\ h_n^{(2)}(ka)],$$

where $j_n(ka)$ and $h_n^{(2)}(ka)$ are the usual spherical Bessel and spherical Hankel functions, respectively. Performing the same term by term comparison as was done for the circular cylinder, we obtain

$$c_1^{3D}\sqrt{|P(\boldsymbol{J})|} \le \|\boldsymbol{J}\|_{H_{div}^{1/2}}, \qquad (30)$$

where $c_1^{3D} = \left[\max\left(c_1^c, c_1^d\right)\right]^{-1/2}$ and

$$c_1^c = \max_{n:n\ge 1} |\alpha_n^c (1+n^2)^{-1/2}|. \qquad (31)$$

The constant c_1^d is defined similarly, replacing α_n^c with α_n^d and $-1/2$ with $1/2$ in the exponent. Numerically, $c_1^{3D} \approx 4.5/(k\eta)\,(ka)^{-2/3}$. As in the two-dimensional problem, Eq. (30) implies that if the current $\boldsymbol{J}$ vanishes in the Sobolev norm, then the forward scattering amplitude must also vanish. For scattering from a circular cylinder, we proved a stronger equivalency relationship. Because of possible cancellation between radiation from curl-free and divergence-free modes, a constant analogous to c_2 cannot be obtained for the sphere. Fortunately, this stronger equivalency is not essential to provide a physical interpretation of the Sobolev norm in error estimates, as will be seen.

IV. NUMERICAL EXAMPLES

To illustrate the relationship between current measures in 2D, we consider two example TM currents. The current $u^{(1)}$ is induced by an incident plane wave and $u^{(2)}$ is a single mode $e^{iq'\phi}$ that is nearest to resonance ($|\alpha_{q'}| \le |\alpha_q|,\ |q| \le q_0$) for a given value of ka. The corresponding Fourier coefficients are given by Eq. (2) for $u^{(1)}$ and by $U_q^{(2)} = \sqrt{2\pi}\delta_{qq'}$, where $\delta_{qq'}$ is the Kronecker delta. Figure 1 shows the ratio $\|\cdot\|_{\mathcal{H}^s}/\|\cdot\|_P$ for $u^{(1)}$ and $u^{(2)}$ as a function of electrical size ka. We plot on the same axes c_1 and c_2. The ratio of norms is always bounded below by c_1 and above by c_2, as proved. Near resonances, the bound c_2 becomes large, but away from resonances it is on the order of 0.01.

Similarly, define a current $\boldsymbol{J}$ on a sphere that is induced by a plane wave of unit amplitude, $\hat{\mathbf{x}}$ polarized and traveling in the negative z direction. In this circumstance, the current coefficients can be obtained using results of [13] and are given by

$$u_{nm}^{cf} = \frac{\sqrt{\pi}}{k\eta} i^{-n-1} \frac{\sqrt{2n+1}}{ka\, h_n^{(1)}(ka)} (\delta_{m,-1} - \delta_{m,1}) \qquad (32)$$

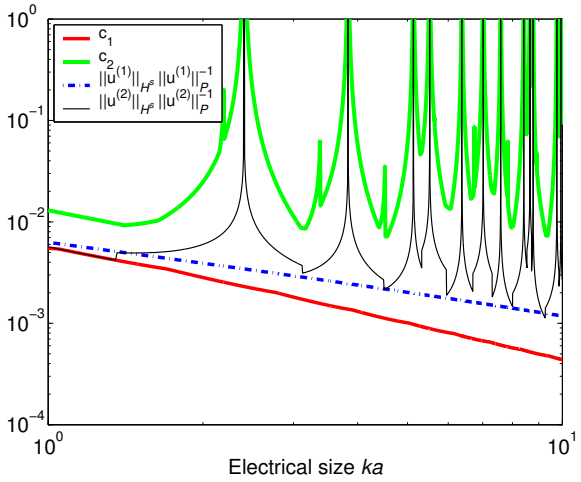


Fig. 1. Illustration of norm equivalence for two different test currents (TM polarization). The constants c_1 and c_2 always bound the ratio $\|\cdot\|_{\mathcal{H}^s}/\|\cdot\|_P$, which ratio is shown for the current induced by a plane wave ($u^{(1)}$) and a nearest-to-resonance, single-mode current $u^{(2)}$. Similar results are obtained for the TE polarization.

and

$$u_{nm}^{df} = \frac{\sqrt{\pi}}{k\eta} i^{-n-1} \frac{\sqrt{2n+1}}{[ka\, h_n^{(1)}(ka)]'} (\delta_{m,-1} + \delta_{m,1}). \qquad (33)$$

Figure 2 verifies the bound (30) for this surface current. Note that the ratio $\|\boldsymbol{J}\|_{H_{div}^{1/2}} |P(\boldsymbol{J})|^{-1/2}$ is always greater than c_1^{3D}, as predicted by Eq. (30). We also see from Figs. (1) and (2) that away from resonances

$$\|u\|_P \approx \frac{1}{c_1}\|u\|_{\mathcal{H}^s} \ , \ \sqrt{|P(\boldsymbol{J})|} \approx \frac{1}{c_1^{3D}}\|\boldsymbol{J}\|_{H_{div}^{1/2}}. \qquad (34)$$

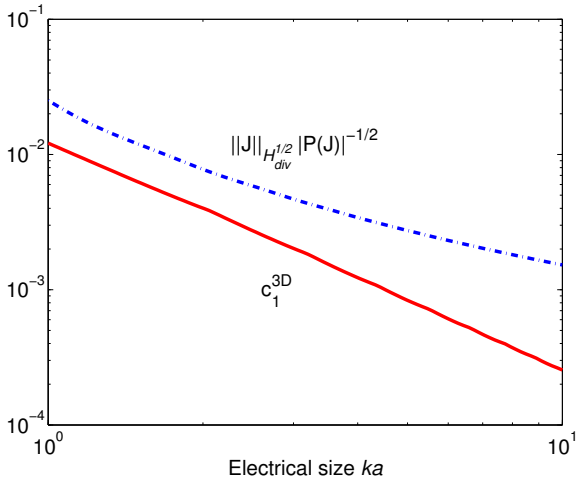


Fig. 2. Ratio of the Sobolev norm to the square root of the magnitude of the forward scattering amplitude for a plane wave induced current on a PEC sphere. The ratio is always bounded below by c_1^{3D}, as proved.

V. APPLICATION TO ERROR ANALYSIS

The equivalency statement (8) provides a physical interpretation for the Sobolev norm of the current solution error in the estimate (1). Suppose a moment method solution $\hat{u}$ to

$\mathcal{L}u = E^i$ on a PEC circular cylinder is generated, with current error defined by $\Delta u = u - \hat{u}$. The Sobolev norm of the error is approximately proportional within a range specified by the constants c_1 and c_2 to the total power radiated by the error current Δu if it were impressed on the scatterer contour $\mathcal{C}$. This transforms (1) into an error estimate for a physically meaningful quantity:

$$|P(\Delta u)| \leq c_1^{-2} \|\Delta u\|_{\mathcal{H}^s}^2 \leq (C/c_1)^2 \, h^{2r}, \tag{35}$$

which implies that the forward scattering amplitude or total supplied power associated with the current error Δu must decay at least as quickly as h^{2r}.

The quantity $\|\Delta u\|_P^2 = |P(\Delta u)|$ in Eq. (35) is the forward scattering amplitude of the error current, which is not the error in the forward scattering amplitude as computed from the numerical current solution $\hat{u}$. In terms of the EFIE operator $\mathcal{L}$, we have for the forward scattering amplitude of the error current

$$\|\Delta u\|_P^2 = |P(\Delta u)| = \frac{k\eta}{4} \left| \int_S (\mathcal{L}\Delta u)^* \Delta u \, ds \right|, \tag{36}$$

whereas the error in the computed scattering amplitude is

$$|\Delta S(\phi^i, \phi^s)| = |S - \hat{S}| = \frac{k\eta}{4} \left| \int_S E^{s*} \Delta u \, ds \right|. \tag{37}$$

To relate the Sobolev norm to a direct error quantity requires the scattering error to be put in a form containing two Δu terms. This may be done by defining an adjoint equation, $\mathcal{L}^a u^a = E^s$, where $\mathcal{L}^a$ is the adjoint of $\mathcal{L}$. Assuming that the adjoint equation is solved using the same procedure as the EFIE, but exchanging the roles of testing and basis functions, the following result is obtained [14]–[18]

$$|\Delta S(\phi^i, \phi^s)| = \frac{k\eta}{4} \left| \int_S (\mathcal{L}\Delta u) \, (\Delta u^a)^* \, ds \right| \tag{38}$$

where $\Delta u^a = u^a - \hat{u}^a$. Because $\mathcal{L}$ is not self-adjoint, Δu is not simply related to Δu^a for an arbitrary scattered field and Eq. (38) does not behave like an induced norm for Δu. However, in the backscatter direction $E^s = E^{i*}$, $u^a = u^*$, and assuming that $\hat{u}^a$ and $\hat{u}$ are expanded in the same basis (Galerkin testing), we also have $\hat{u}^a = \hat{u}^*$. This yields

$$|\Delta S(\phi^i, \phi^i)| = \frac{k\eta}{4} \left| \int_S (\mathcal{L}\Delta u) \, \Delta u \, ds \right|. \tag{39}$$

Note that $|\Delta S(\phi^i, \phi^i)|$ (39) differs from $\|\Delta u\|_P^2$ (36) only by a conjugate on the $\mathcal{L}\Delta u$ term. For a circular cylinder, ΔS in the backscattering direction is therefore similar to Eq. (7) and is given by

$$\Delta S(\phi^i, \phi^i) = -\sum_q \alpha_q^* (\Delta U_q)^2 \tag{40}$$

where we have used the fact that $\Delta U_{-q} = \Delta U_q$. The derivation of the lower constant c_1 applies equally well to the series (40) as it does to the series (7), therefore the inequality (35) is valid replacing $|P(\Delta u)|$ with $|\Delta S(\phi^i, \phi^i)|$, giving finally

$$|\Delta S(\phi^i, \phi^i)| \leq c_1^{-2} \|\Delta u\|_{\mathcal{H}^s}^2 \leq (C/c_1)^2 \, h^{2r}. \tag{41}$$

This is a new bound on the backscattering error, subject to the Galerkin testing condition. It shows that the error in the backscattering amplitude must decay at least as quickly as h^{2r}.

The curves in Fig. 3 were computed by generating a moment method solution for the EFIE and computing the coefficients ΔU_q numerically. A triangle (piecewise linear) expansion was chosen to avoid the Gibbs phenomenon associated with the Fourier coefficients of discontinuous functions. We see that the inequalities (35) and (41) are evident in the figure because both $|\Delta S(\phi^i, \phi^i)|$ and $\|u\|_P^2$ are both less than $1/c_1^2 \, \|\Delta u\|_{\mathcal{H}^s}^2$. Further, we see that the error measures $\|\Delta u\|_{\mathcal{H}^s}^2$, $\|u\|_P^2$, and $|\Delta S|$ converge asymptotically at the same rate as the mesh is refined. While this is required of the first two error measures by the equivalency statement (8), $|\Delta S(\phi^i, \phi^i)|$ may actually converge faster than the Sobolev measure $\|\Delta u\|_{\mathcal{H}^s}^2$ without violating any inequality derived in this paper. We also note that all three error measures converge as h^5, which rate is proved for $|\Delta S|$ analytically in [12]. This is much faster than the $2r = 1$ rate predicted by the Sobolev bound (1), implying that these bounds are not tight. To achieve this convergence rate required a quadrature rule that combined lin-log Gaussian quadrature [19] with a Gauss-Legendre rule.

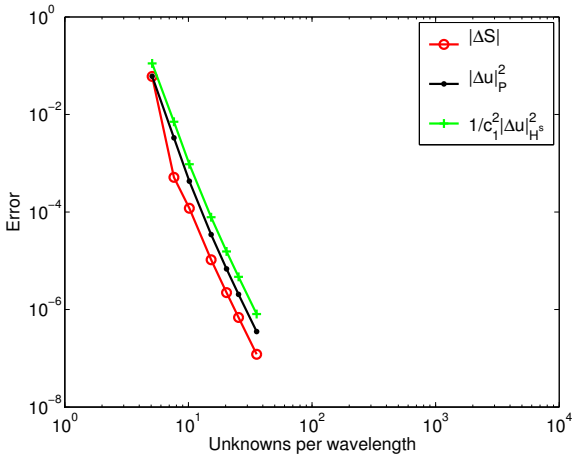


Fig. 3. Three different error measures for a moment method, TM polarized current solution, $ka = \pi/4$. The horizontal axis is λ/h. The backscattering amplitude error $|\Delta S|$ (circles) is computed from an MoM solution for $\mathcal{L}u = E^i$ with triangle expansion functions and Galerkin testing. The forward scattering amplitude of the current error, $\|\Delta u\|_P$ (dots) is equivalent in the rigorous sense to the Sobolev measure $\|\Delta u\|_{\mathcal{H}^s}$ (pluses). These error measures are related by Eqs. (35) and (41). For this particular value of ka, we have $c_1 \approx 0.006$.

We have given numerical examples of moment method error measures for scattering from a circular cylinder. Since the variational expression (38) applies also in three dimensions, we have

$$|\Delta S| \leq (c_1^{\mathrm{3D}})^{-2} \|\Delta \boldsymbol{J}\|_{H_{div}^{1/2}}^2. \tag{42}$$

Here, ΔS is the error in the backscattering amplitude for moment method solutions to scattering from a sphere. (To the authors' knowledge, there are no bounds analogous to Eq. (1) for 3D PEC scattering problems, although [2], [4] give Sobolev-type bounds for 3D scalar problems.) To give a numerical example similar to Fig. 3, computing the Sobolev measure $\|\Delta \boldsymbol{J}\|_{H_{div}^{1/2}}^2$ would require computing the inner products of $\Delta \boldsymbol{J}$

with Legendre polynomials and complex exponentials. Since this is tedious, we omit it. The impracticality of computing the Sobolev norm of error currents is one reason a physically meaningful alternative to the Sobolev norm is desirable.

VI. CONCLUSIONS

We have related the abstract Sobolev norm of an arbitrary current on an infinite, PEC circular cylinder to the forward scattering amplitude associated with that current. A slightly weaker result was derived for 3D currents on a PEC sphere. This equivalency was used to show that a small error current measured in the Sobolev norm implies that the error current radiates little energy. Therefore, Sobolev error estimates prove that moment method solutions converge in the sense that the energy radiated by the current error vanishes as the mesh is refined.

Further, a direct relationship was derived between the Sobolev norm of the current error to the error in the computed backscattering amplitude solution. This provides a link between error estimates in the Sobolev literature to physical quantities in practical CEM simulations. We conjecture that these observations hold for more general scatterers.

APPENDIX
INEQUALITY FOR COMPUTING c_2

Let a, b, c be real numbers with $|b| \leq M|a|$ and $M \geq 1$. Then we have the inequality

$$|a + i(b + c)| \geq \frac{|a| + |c|}{3M}. \tag{43}$$

Proof: Assume that $|b| > |c|$. Then we have

$$|a + i(b + c)| \geq |a| = \frac{|a|}{2} + \frac{|a|}{2} \tag{44a}$$

$$\geq \frac{|a|}{2} + \frac{|b|}{2M} \tag{44b}$$

$$\geq \frac{|a|}{2} + \frac{|c|}{2M} \tag{44c}$$

$$\geq \frac{|a| + |c|}{3M}. \tag{44d}$$

The second step (44b) follows from the given $|b| \leq M|a|$ and (44c) from the case statement $|b| > |c|$. The fourth line (44d) follows from the given $M \geq 1$. Now assume instead that $|b| \leq |c|$. It can be shown that

$$\sqrt{2}|a + i(b + c)| \geq |a| + |b + c| \geq |a| + |c| - |b|, \tag{45}$$

where we have used the triangle inequality and $|b| \leq |c|$. Claiming that

$$|a| + |c| - |b| \geq \frac{|a| + |c|}{2M}, \tag{46}$$

inequality (43) immediately follows. We can prove claim (46) by contradiction. Suppose that

$$|a| + |c| - |b| < \frac{|a| + |c|}{2M} \tag{47}$$

is true. Then the following inequalities are implied

$$(2M - 1)(|a| + |c|) < 2M|b| \tag{48a}$$

$$(2M - 1)(|b|/M + |c|) < 2M|b| \tag{48b}$$

$$|c| < \left(\frac{2M}{2M - 1} - \frac{1}{M}\right)|b|. \tag{48c}$$

For $M \geq 1$, the expression $2M/(2M - 1) - 1/M$ is less than one. This implies that $|c| < |b|$, a contradiction to the case statement $|b| \leq |c|$. Thus, the assumption (47) must be false and Eq. (46) must hold, completing the proof.

REFERENCES

[1] M. Feistauer, G. C. Hsiao, and R. E. Kleinman, "Asymptotic and a posteriori error estimates for boundary element solutions of hypersingular integral equations," *SIAM J. Numer. Anal.*, vol. 33, pp. 666–685, Apr. 1996.

[2] M. Costabel and E. P. Stephan, "A direct boundary integral equation method for transmission problems," *J. Math. Anal. Appl.*, vol. 106, pp. 367–413, 1985.

[3] M. Maischak, P. Mund, and E. P. Stephan, "Adaptive multilevel BEM for acoustic scattering," *Comp. Meth. Appl. Mech. Engrg.*, vol. 150, pp. 351–367, 1997.

[4] E. P. Stephan, "Boundary integral-equations for screen problems in IR-3," *Integral Equations and Operator Theory*, vol. 10, pp. 236–257, 1987.

[5] H. Holm, M. Maischak, and E. P. Stephan, "The *hp*-version of the boundary element method for helmholtz screen problems," *Computing*, vol. 57, pp. 105–134, 1996.

[6] E. F. Kuester, "Computable error bounds for variational functionals of solutions of a convolution integral equations of the first kind," *Wave Motion*, vol. 22, pp. 171–185, 1995.

[7] G. C. Hsiao and R. E. Kleinman, "Mathematical foundations for error estimation in numerical solutions of integral equations in electromagnetics," *IEEE Trans. Ant. Propag.*, vol. 45, pp. 316–328, March 1997.

[8] C. P. Davis and K. F. Warnick, "The physical meaning of the Sobolev norm in error estimation," in *IEEE Antennas and Propagation Society Digest*, vol. 3, (Monterey, CA), pp. 3377–3380, IEEE Antennas and Propagation Society Symposium, June 2004.

[9] J. Marti, *Introduction to Sobolev Spaces and Finite Element Solution of Elliptic Boundary Value Problems*. London: Academic Press, 1986.

[10] S. Amini and S. Kirkup, "Solution of Helmholtz equation in the exterior domain by elementary boundary integral methods," *Journal of Computational Physics*, vol. 118, pp. 208–221, 1995.

[11] M. Abramowitz, I. A. Stegun, and et. al., *Handbook of Mathematical Functions*. Dover Publications, 1965.

[12] K. F. Warnick and W. C. Chew, "Accuracy of the method of moments for scattering by a cylinder," *IEEE Trans. Micr. Th. Tech.*, vol. 48, pp. 1652–1660, Oct. 2000.

[13] J. J. Bowman, T. B. A. Senior, and P. L. E. Uslenghi, *Electromagnetic and Acoustic Scattering by Simple Shapes*. New York: Hemisphere, 1987.

[14] J. H. Richmond, "On the variational aspects of the moment method," *IEEE Trans. Ant. Propag.*, vol. 39, pp. 473–479, April 1991.

[15] J. Mautz, "Variational aspects of the reaction in the method of moments," *IEEE Trans. Ant. Propag.*, vol. 42, pp. 1631–1638, Dec. 1994.

[16] M. I. Aksun and R. Mittra, "Choices of expansion and testing functions for the method of moments applied to a class of electromagnetic problems," *IEEE Trans. Microw. Theory Tech.*, vol. 41, pp. 503–509, Mar. 1993.

[17] A. F. Peterson, D. R. Wilton, and R. E. Jorgenson, "Variational nature of Galerkin and non-Galerkin moment method solutions," *IEEE Trans. Ant. Propag.*, vol. 44, pp. 500–503, April 1996.

[18] D. G. Dudley, "Comments on 'Variational nature of Galerkin and non-Galerkin moment method solutions'," *IEEE Trans. Ant. Propag.*, vol. 45, p. 1062, June 1997.

[19] J. Ma, V. Rokhlin, and S. Wandzura, "Generalized gaussian quadrature rules for systems of arbitrary weight functions," *SIAM J. Numer. Anal.*, vol. 33, pp. 971–996, June 1996.

Clayton P. Davis received the BS degree (*cum laude*) and the MS degree in 2004, both from Brigham Young University (BYU), Provo, UT. He is currently working toward the PhD degree at the University of Illinois at Urbana-Champaign. His research interests include electromagnetic theory, numerical methods, and signal processing.

Karl F. Warnick received the BS degree with University Honors in 1994 and the PhD degree in 1997, both from Brigham Young University (BYU), Provo, UT. He was a recipient of the National Science Foundation Graduate Research Fellowship. From 1998 to 2000, he was a postdoctoral research associate and Visiting Assistant Professor in the Center for Computational Electromagnetics at the University of Illinois at Urbana-Champaign. Since 2000, he has been an Assistant Professor in the Department of Electrical and Computer Engineering at Brigham Young University. Research interests include computational electromagnetics, antenna theory, rough surface scattering, remote sensing, inverse scattering, and numerical analysis. He has coauthored a book chapter and over 50 conference presentations and scientific journal papers.

Induced Currents on a Moving and Vibrating Perfect Plane Under the Illumination of Electromagnetic Pulse: One-Dimensional Simulation using Characteristic-Based Algorithm

Mingtsu Ho
Department of Electronic Engineering, WuFeng Institute of Technology
117 Jian Kuo Road Section 2, Min Shong, Chia Yi, Taiwan
homt@mail.wfc.edu.tw

Abstract: This paper provides one-dimensional simulation results of the induced currents on constantly moving and vibrating perfect conductors under the normal illumination of plane Gaussian electromagnetic pulses. The characteristic-based algorithm is employed for the solutions of time-dependent Maxwell curl equations. In the numerical model, the size of the computational cell adjacent to the moving boundary, and its corresponding numerical time step become time-dependent since the boundary is not stationary. By comparing the computational results with the theoretical Doppler shift values, we show that the present method successfully predicts the induced currents on the perfect conductor surface. The computed electric and magnetic field intensities and induced currents are demonstrated as well.

***Introduction*:**

The effects on electromagnetic waves caused by uniformly traveling or oscillating targets are usually neglected if the velocity of movement or the resultant instantaneous speed of vibration is relatively small. The study of these topics becomes important wherever researchers have to deal with them. Several analytic studies can be found and the following remarks can be drawn: perfect conductors undergoing translational motion result in the well-known Doppler shift in the reflected fields; an oscillating target changes not only the phase but also the magnitude of the scattered fields [1–3].

A variety of computational techniques are developed for the solutions of the electromagnetic scattering problems for the past half century. The two most commonly approaches for solving electromagnetic problems are the method of moments (MoM) and the finite-difference time-domain (FDTD) technique. A recently proposed method applied to the solution of various electromagnetic problems is the characteristic-based algorithm that numerically approximates the time-dependent Maxwell curl equations. Whitfield and Janus applied this characteristic-based algorithm to the solutions of the Navier-Stokes equations for the fluid dynamic problems in the early 80s [4]. A decade later, Shang employed this method to solve the time-domain Maxwell's equations [5] through the application of explicit central-difference scheme. The implicit formulation was developed for the same purpose and its results were found to agree with data produced by FDTD [6]. Unlike MoM and FDTD, all field quantities are placed in the center of grid cell in the characteristic-based approach. It directly solves Maxwell's equations by balancing the net flux across all cell faces within each computational cell. The present numerical method is then considered a better approach over MoM and FDTD for problems involved with time varying cells, such as cases where object is moving or vibrating.

Governing Equations:

The governing equations for electromagnetic problems in source-free region are the time-dependent Maxwell curl equations:

$$\frac{\partial \mathrm{B}}{\partial \mathrm{t}} + \nabla \times \mathrm{E} = 0 \tag{1}$$

$$\frac{\partial \mathrm{D}}{\partial \mathrm{t}} - \nabla \times \mathrm{H} = 0 . \tag{2}$$

Since one-dimensional models are used, we can only consider a two-dimensional numerical formulation. To begin with, the characteristic-based algorithm requires the transformation of the governing equations from the

Cartesian coordinate system (t, x, y) into the body-fitted coordinate system (τ, ξ, η). We rewrite (1) and (2) as

$$\frac{\partial Q}{\partial \tau} + \frac{\partial F}{\partial \xi} + \frac{\partial G}{\partial \eta} = 0 \tag{3}$$

where

$$Q = J\, q \tag{4}$$

$$F = J\,(\xi_x f + \xi_y g) \tag{5}$$

$$G = J\,(\eta_x f + \eta_y g) \tag{6}$$

and

$$J = \left| \frac{\partial x}{\partial \xi}\frac{\partial y}{\partial \eta} - \frac{\partial x}{\partial \eta}\frac{\partial y}{\partial \xi} \right|. \tag{7}$$

The symbol J in above equations stands for the Jacobian of the inverse transformation, and the three variable vectors are respectively given by

$$q = [\, B_x\,,\, B_y\,,\, D_z\,]^T \tag{8}$$

$$f = [\, 0\,,\, -E_z\,,\, -H_y\,]^T \tag{9}$$

$$g = [\, E_z\,,\, 0\,,\, H_x\,]^T. \tag{10}$$

Shown as in (3) is called the Maxwell's equations in form of the Euler equation. The numerical procedure is formulated by applying the central difference operator

$$\delta_k(\varphi) = (\varphi)_{k+½} - (\varphi)_{k-½} \tag{11}$$

to (3). Then it becomes

$$\frac{Q^{n+1} - Q^n}{\Delta \tau} + \frac{\delta_i F}{\Delta \xi} + \frac{\delta_j G}{\Delta \eta} = 0\,. \tag{12}$$

In (11) the half-integer index represents the interface between two adjacent computational cells where flux is evaluated. The superscripts "n" and "n+1" on variable vector Q in (12) are two consecutive time levels. The numerical method approximates Maxwell's curl equation in curvilinear coordinate system by solving for the flux change for each grid cell within each numerical time step. The flux vector splitting technique and the Newton iterative method are also applied followed by the lower-upper approximate factorization scheme for the solution of the system of linear equations.

Boundary Conditions:

The boundary conditions (BC's) used in the present method are derived from the concepts of characteristic variable (CV) boundary conditions and the relativistic boundary conditions. According to the definition, every CV associates with one particular eigenvalue and is defined as the product of the instantaneous variable vector and the eigenvector corresponding to that particular eigenvalues [4]. Since every eigenvalue indicates the direction and velocity of the information propagating across the cell face, the use of CV for the evaluation of boundary variables should increase the accuracy of scheme. In order to incorporate the relativistic effects on the perfectly conducting surface, we combine the characteristic variable boundary conditions and the relativistic relation to evaluate the boundary values of variables. The relativistic boundary conditions are given by

$$\hat{n} \times \vec{E}^b = (\hat{n} \cdot \vec{v})\, \vec{B}^b \tag{13}$$

where $\vec{v}$ and $\hat{n}$ are the velocity and unit vector normal of the perfectly conducting surface, respectively. The superscript "b" stands for the boundary values of the electric and magnetic field variables evaluated right on the perfectly conducting surface. By definition, the CV arriving on the boundary is given by

$$CV^b = \hat{n} \times \vec{B} + \eta_o\, \vec{D} \tag{14}$$

with η_o being the impedance of free space and $\vec{B}$ and $\vec{D}$ are variables of the cell next to the boundary. Note that, this CV^b is tangential to the perfectly conducting surface and contains information propagating from the adjacent cell as indicated by the corresponding eigenvalue.

***The Problem*:**

The incident electromagnetic pulse used in the simulation is specified as follows. It is a Gaussian-windowed plane electromagnetic pulse with a cutoff level of 100 dB from the peak value, initially propagates in the positive x-direction in source-free region, and normally illuminates upon a perfect plane that is either at rest or in motion. This Gaussian electromagnetic pulse with its electric intensity being normalized to unity has a width of

about 1.902 ns measured from the center to $e^{-0.5}$.

For a motionless boundary, the grid system is stationary where the cell number and cell dimension are constant and uniform as shown in Figure 1(a). In the present simulation, due to the motion of boundary, both cell number and cell size are changing as time advances. As shown in Figures 1(b) and 1(c), provided that the cell immediate next to the moving boundary is the cell N, portion of the N^{th} cell may be truncated by the boundary at certain instance of time; and a moment later an extra fractional cell, the $(N+1)^{th}$, may be introduced into the grid system. The number of cells eliminated from or added into the grid system may be multiple and subjects to the oscillation amplitude and the grid density. These variations must be taken into account by updating the effective cell area and the resultant numerical time step to maintain decent accuracy of scheme.

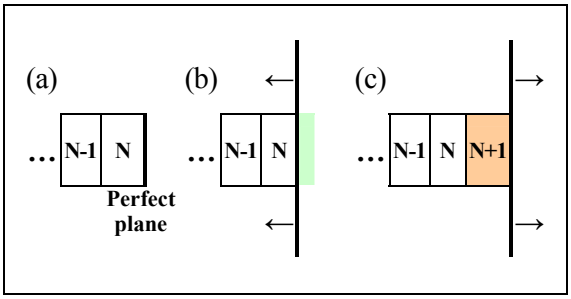


Figure 1. Computational cell indexing: (a) stationary grid system, (b) the N^{th} cell is truncated, (c) the $(N+1)^{th}$ cell is introduced.

In order to easily observe the effects of the moving object on the induced currents, we make the following arrangements. The perfectly conducting surfaces are set to constantly move at a velocity of 10 percent of the speed of light ($C = 3 \times 10^8$ m/s), and/or vibrate with a constant frequency and a constant amplitude so that the extreme instantaneous velocity equal to ± 0.1 C. The vibration frequency and amplitude are set to be 1 GHz (an impractical high value) and 4.775 mm to result in an extreme speed of 0.1 C near the equilibrium position. The resulting velocities of conductor are illustrated in Figure 2 where they are superposed if conductor moves and vibrates simultaneously. The two ratios of the translational velocity and oscillatory instantaneous velocity to the light speed are β_τ and β_v, respectively. Since the latter ranges from –0.1 C to +0.1 C, symbol $|\beta_v|$ is used for the magnitude. The value of β_τ and β_v is positive if conductor and the incident pulse move in the same direction and negative if they approach each other.

The numerical setups are as follows: the numerical electromagnetic pulse is plane and only has the components E_z and B_y. The excitation pulse is three meters in spatial span from the peak to the cutoff point; the number of grid cell for a six meters span is 800 points; the numerical time step is set so that the numerical electromagnetic pulse takes forty steps for one grid cell. Note that, for an oscillation amplitude as previously stated the moving boundary covers 1.273 cells peak-to-peak.

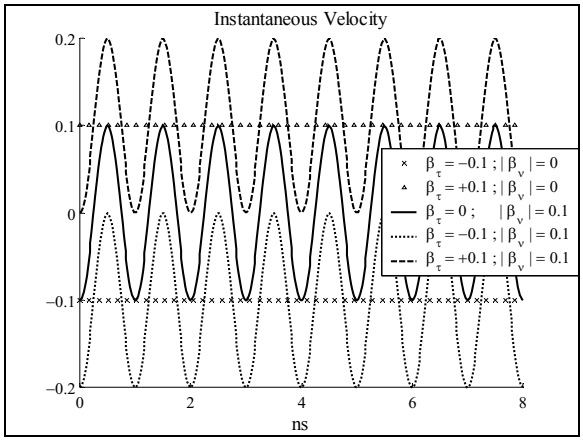


Figure 2. Instantaneous velocities of moving and/or vibrating conductors.

The induced currents are computed by taking the cross product of the unit vector normal and the magnetic field intensity where the latter is the resultant boundary values. Another field quantity sampled at the same location is the electric field intensity. If conductor is stationary the electric field is always zero in magnitude. Yet, the electric field is no longer vanished if conductor moves so that the relativistic effects cannot be dismissed. Under such circumstances, if the resultant velocity of conductor is $\vec{v}$, the boundary values of the field

components can be solved directly from equations (13) and (14) and are respectively given

$$\vec{B}^b = \frac{1}{\vec{v}-1} CV^b \tag{15}$$

$$\vec{E}^b = \frac{\vec{v}}{\vec{v}-1} CV^b . \tag{16}$$

To obtain above expressions, we take the convenience that both unit vector normal of the surface and velocity are along the x-axis. Note also that $\vec{v}$ used in (15) and (16) is the combined velocity of conductor and that the sign of this velocity is dependent upon the relativistic motion between the electromagnetic pulse and conductor as previously mentioned. The induced current flows in the positive z-direction can be computed as

$$\vec{J}_z = \hat{n} \times \vec{H}^b \tag{17}$$

where $\vec{H}^b$ is the magnetic field intensity evaluated on the boundary.

***Results*:**

To illustrate the interaction between the electromagnetic pulse and the moving/vibrating perfect conductor, two time sequences of the electric field intensity are given in Figure 3. It is observed that on the perfectly conducting surface the electric fields are not always zero in strength due to the application of the relativistic boundary conditions. Plotted in Figures 4 and 5 are the boundary values of the electric and magnetic fields computed through (15) and (16). Note that all field quantities are normalized to unity henceforward and that the induced currents are similar to those of Figure 4 since it can be obtained by (17). If we take the difference between (15) and (16), we expect by the mathematical expression that the oscillatory behaviors of the fields would be cancelled out. The computational results are calculated and shown in Figure 6.

The Doppler effects on the induced current can be investigated on both magnitude and pulse width. The magnitude of the induced current is predictable by the relation $\frac{2}{1+(\beta_\tau+\beta_v)}$ and the resulted pulse width by $\sqrt{\frac{1-\beta_\tau}{1+\beta_\tau}}$ where the oscillatory behavior is ignored for easy estimation. Listed in Tables 1 and 2 are the calculated maximum shifts in magnitude along with the theoretical values. For instance, if the boundary moves and vibrates at the same time, when the maximum instantaneous velocity is –0.2 C, the corresponding magnitude is equal to 2.5; the pulse width is 1.2247 times that of the incident pulse, which is 1.7208. It is noted that the computational results are in good agreement with the analytical calculations.

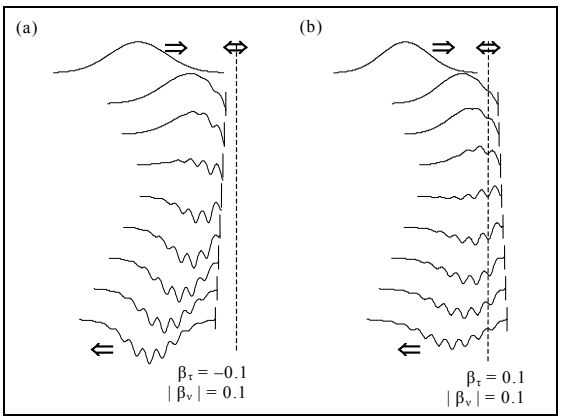


Figure 3. Interaction of electromagnetic pulse with moving perfect planes: (a) Vibrating and approaching, (b) Vibrating and receding (only electric fields are shown).

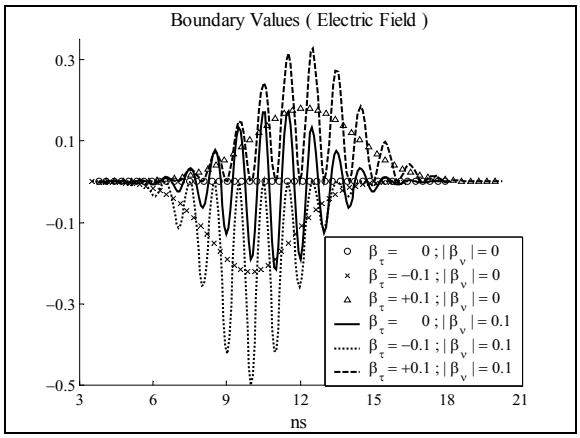


Figure 4. Calculated electric field using equation (3).

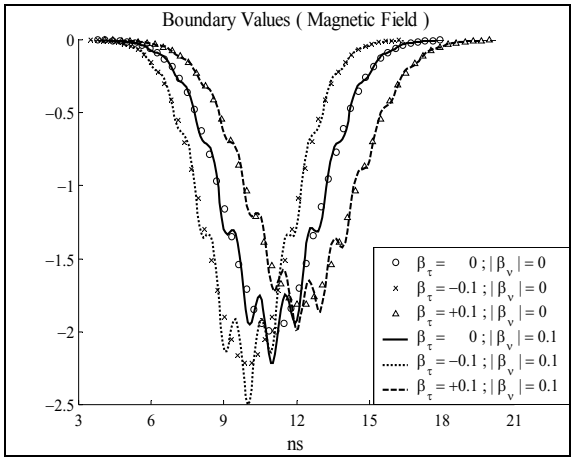


Figure 5. Calculated magnetic fields using equation (4).

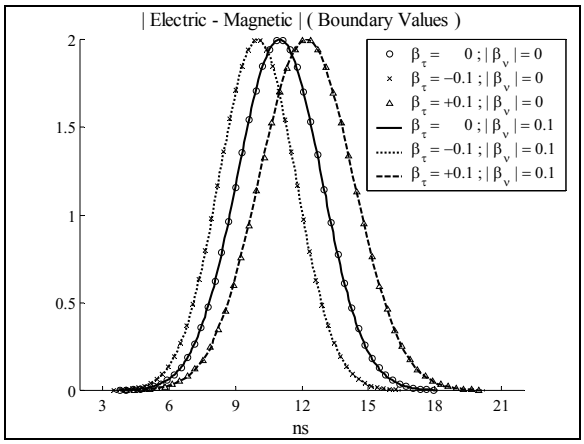


Figure 6. Magnitude of (E_z^b - H_y^b).

Table 1: Doppler shifts in pulse width.

Velocities		From center to $e^{-0.5}$	
β_τ	$\lvert \beta_v \rvert$	Calculated	Theoretical
0	0.0	1.9028	1.9024
− 0.1	0.0	1.7294	1.7208
+ 0.1	0.0	2.1153	2.1032

Table 2: Doppler shifts in the induced current.

Velocities		Maximum $\lvert J_z \rvert$	
β_τ	$\lvert \beta_v \rvert$	Calculated	Theoretical
0	0.0	1.9999	2.0000
− 0.1	0.0	2.2223	2.2222
+ 0.1	0.0	1.8176	1.8182
0	0.1	2.2220	2.2222
− 0.1	0.1	2.5001	2.5000
+ 0.1	0.1	1.9908	2.0000

***Conclusion*:**

This paper has shown that the characteristic-based algorithm successfully simulates the induced currents on the surface of moving and/or vibrating perfect conductors in one dimension. The computational results of the induced current magnitudes and pulse widths as consequences of the moving conductors are compared with the theoretical values. They are in good agreement. It is our future work to develop the existing code to problems with objects of finite dimension and problems involved with moving medium.

References

[1] R. E. Kleinman, R. B. Mack, "Scattering by linearly vibrating objects," IEEE Trans. Antennas Propagation, vol. AP-27, no. 3, pp. 344-352, May 1979.

[2] J. Cooper, "Scattering of electromagnetic fields by a moving boundary: the one-dimensional case," IEEE Trans. Antennas Propagation, vol. AP-28, no. 6, pp. 791-795, November 1980.

[3] F. Harfoush, A. Taflove, and G. Kriegsmann, "A numerical technique for analyzing electromagnetic wave scattering from moving surfaces in one and two dimensions," IEEE Trans. Antennas Propagation, vol. 37, pp. 55-63, January 1989.

[4] D. L. Whitfield and J. M. Janus, "Three-dimensional unsteady Euler equations solution using flux splitting," AIAA Paper, no. 84-1552, June 1984.

[5] J. S. Shang, "A characteristic-based algorithm for solving 3-d time-domain Maxwell equations," Electromagnetics, vol. 10, pp. 127, 1990.

[6] J. P. Donohoe, J. H. Beggs, and M. Ho, "Comparison of finite-difference time-domain results for scattered electromagnetic fields: Yee algorithm vs. a characteristic based algorithm," 27th IEEE Southeastern Symposium on System Theory, pp. 325-328, March 1995.

Mingtsu (Mark) Ho was born in Chia Yi City, Taiwan, R.O.C. in 1958. He received the B.S. degree in Physics in 1983 from the National Cheng Kung University, Tainan, Taiwan, R.O.C., the M.S. degrees in Physics and Electrical Engineering from the Mississippi State University, U.S.A., and the Ph.D. degree in Electrical Engineering from the Mississippi State University, U.S.A., in 1997.

He was a senior engineer in the R&D department, TECO Electric and Machinery Co., Ltd. from March 1998 till August 1999. Since then he has been an assistant professor in the Department of Electronic Engineering, Wufeng Institute of Technology, Chia Yi, Taiwan, R.O.C. His research interests include the use of the characteristic-based method for numerical simulations of the following problems: scattering of EM waves from moving/vibrating objects, EM interferences among strips, and the propagation of EM waves/pulses inside media having complex higher-order susceptibilities.

2005 INSTITUTIONAL MEMBERS

AUSTRALIAN DEFENCE LIBRARY
Northcott Drive
Campbell, A.C.T. 2600 AUSTRALIA

BEIJING BOOK COMPANY, INC
701 E Lindon Ave.
Linden, NJ 07036-2495

DARMSTADT U. OF TECHNOLOGY
Schlossgartenstrasse 8
Darmstadt, Hessen
GERMANY D-64289

DARTMOUTH COLL-FELDBERG LIB
6193 Murdough Center
Hanover, NH 03755-3560

DEFENCE RESEARCH ESTAB. LIB.
3701 Carling Avenue
Ottowa, ON, K1A 0Z4 CANADA

DPS/LIBRARY (EABV)
Alion Science & Technology
185 Admiral Cochrane Drive
Annapolis, MD 214017307

DSTO-DSTORL EDINBURGH
Jets AU/33851-99, PO Box 562
Milsons Point, NSW
AUSTRALIA 1565

DTIC-OCP/LIBRARY
8725 John J. Kingman Rd. Ste 0944
Ft. Belvoir, VA 22060-6218

ELSEVIER
Bibliographic Databases
PO Box 2227
Amsterdam, Netherlands 1000 CE

ENGINEERING INFORMATION, INC
PO Box 543
Amsterdam, Netherlands 1000 Am

ETSE TELECOMUNICACION
Biblioteca, Campus Lagoas
Vigo, 36200 SPAIN

FLORIDA INTERNATIONAL UNIV.
ECE Dept./EAS-3983
10555 W. Flagler St
Miami, FL 33174

GEORGIA TECH LIBRARY
225 North Avenue, NW
Atlanta, GA 30332-0001

HRL LABS, RESEARCH LIBRARY
3011 Malibu Canyon
Malibu, CA 90265

IEE INSPEC/Acquisitions Section
Michael Faraday House
6 Hills Way
Stevenage, Herts UK SG1 2AY

IND CANTABRIA
PO Box 830470
Birmingham, AL 35283

INSTITUTE FOR SCIENTIFIC INFO.
Publication Processing Dept.
3501 Market St.
Philadelphia, PA 19104-3302

LIBRARY of CONGRESS
Reg. Of Copyrights
Attn: 40T Deposits
Washington DC, 20559

LINDA HALL LIBRARY
5109 Cherry Street
Kansas City, MO 64110-2498

MISSISSIPPI STATE UNIV LIBRARY
PO Box 9570
Mississippi State, MS 39762

MIT LINCOLN LABORATORY
Periodicals Library
244 Wood Street
Lexington, MA 02420

NA KANSAI KINOKUNNA CO.
Attn: M. MIYOSHI
PO Box 36 (NDLA KANSAI)
Hongo, Tokyo, JAPAN 113-8688

NAVAL POSTGRADUATE SCHOOL
Attn:J. Rozdal/411 Dyer Rd./ Rm 111
Monterey, CA 93943-5101

NAVAL RESEARCH LABORATORY
C. Office, 4555 Overlook Avenue, SW
Washington, DC 20375

OHIO STATE UNIVERSITY
1320 Kinnear Road
Columbus, OH 43212

PAIKNAM ACAD. INFO CTR LIB.
Hanyang U/17 Haengdang-Dong
Seongdong-ki, Seoul, S Korea 133-791

PENN STATE UNIVERSITY
126 Paterno Library
University Park, PA 16802-1808

RENTON TECH LIBRARY/BOEING
PO BOX 3707
SEATTLE, WA 98124-2207

SOUTHWEST RESEARCH INST.
6220 Culebra Road
San Antonio, TX 78238

SWETS INFORMATION SERVICES
160 Ninth Avenue, Suite A
Runnemede, NJ 08078

SYRACUSE UNIVERSITY
EECS, 121 Link Hall
Syracuse, NY 13244

TECHNISCHE UNIV. DELFT
Mekelweg 4, Delft, Holland, 2628 CD
NETHERLANDS

TIB & UNIV. BIB. HANNOVER
DE/5100/G1/0001
Welfengarten 1B
Hannover, GERMANY 30167

TOKYO KOKA UNIVERSITY
1404-1 Katakura-Cho
Hachioji, Tokyo, JAPAN 192-0914

UNIV OF CENTRAL FLORIDA LIB.
4000 Central Florida Boulevard
Orlando, FL 32816-8005

UNIV OF COLORADO LIBRARY
Campus Box 184
Boulder, CO 80309-0184

UNIVERSITY OF MISSISSIPPI
John Davis Williams Library
PO Box 1848
University, MS 38677-1848

UNIV OF MISSOURI-ROLLA LIB.
1870 Miner Circle
Rolla, MO 65409-0001

USAE ENG. RES. & DEV. CENTER
Attn: Library/Journals
72 Lyme Road
Hanover, NH 03755-1290

ACES COPYRIGHT FORM

This form is intended for original, previously unpublished manuscripts submitted to ACES periodicals and conference publications. The signed form, appropriately completed, MUST ACCOMPANY any paper in order to be published by ACES. PLEASE READ REVERSE SIDE OF THIS FORM FOR FURTHER DETAILS.

TITLE OF PAPER:

AUTHORS(S)

PUBLICATION TITLE/DATE:

RETURN FORM TO:

Dr. Atef Z. Elsherbeni
University of Mississippi
Dept. of Electrical Engineering
Anderson Hall Box 13
University, MS 38677 USA

PART A - COPYRIGHT TRANSFER FORM

(NOTE: Company or other forms may not be substituted for this form. U.S. Government employees whose work is not subject to copyright may so certify by signing Part B below. Authors whose work is subject to Crown Copyright may sign Part C overleaf).

The undersigned, desiring to publish the above paper in a publication of ACES, hereby transfer their copyrights in the above paper to The Applied Computational Electromagnetics Society (ACES). The undersigned hereby represents and warrants that the paper is original and that he/she is the author of the paper or otherwise has the power and authority to make and execute this assignment.

Returned Rights: In return for these rights, ACES hereby grants to the above authors, and the employers for whom the work was performed, royalty-free permission to:

1. Retain all proprietary rights other than copyright, such as patent rights.
2. Reuse all or portions of the above paper in other works.
3. Reproduce, or have reproduced, the above paper for the author's personal use or for internal company use provided that (a) the source and ACES copyright are indicated, (b) the copies are not used in a way that implies ACES endorsement of a product or service of an employer, and (c) the copies per se are not offered for sale.
4. Make limited distribution of all or portions of the above paper prior to publication.
5. In the case of work performed under U.S. Government contract, ACES grants the U.S. Government royalty-free permission to reproduce all or portions of the above paper, and to authorize others to do so, for U.S. Government purposes only.

ACES Obligations: In exercising its rights under copyright, ACES will make all reasonable efforts to act in the interests of the authors and employers as well as in its own interest. In particular, ACES REQUIRES that:

1. The consent of the first-named author be sought as a condition in granting re-publication permission to others.
2. The consent of the undersigned employer be obtained as a condition in granting permission to others to reuse all or portions of the paper for promotion or marketing purposes.

In the event the above paper is not accepted and published by ACES or is withdrawn by the author(s) before acceptance by ACES, this agreement becomes null and void.

AUTHORIZED SIGNATURE — TITLE (IF NOT AUTHOR)

EMPLOYER FOR WHOM WORK WAS PERFORMED — DATE FORM SIGNED

Part B - U.S. GOVERNMENT EMPLOYEE CERTIFICATION

(NOTE: if your work was performed under Government contract but you are not a Government employee, sign transfer form above and see item 5 under Returned Rights).

This certifies that all authors of the above paper are employees of the U.S. Government and performed this work as part of their employment and that the paper is therefor not subject to U.S. copyright protection.

AUTHORIZED SIGNATURE — TITLE (IF NOT AUTHOR)

NAME OF GOVERNMENT ORGANIZATION — DATE FORM SIGNED

PART C - CROWN COPYRIGHT

(NOTE: ACES recognizes and will honor Crown Copyright as it does U.S. Copyright. It is understood that, in asserting Crown Copyright, ACES in no way diminishes its rights as publisher. Sign only if *ALL* authors are subject to Crown Copyright).

This certifies that all authors of the above Paper are subject to Crown Copyright. (Appropriate documentation and instructions regarding form of Crown Copyright notice may be attached).

AUTHORIZED SIGNATURE TITLE OF SIGNEE

NAME OF GOVERNMENT BRANCH DATE FORM SIGNED

Information to Authors

ACES POLICY

ACES distributes its technical publications throughout the world, and it may be necessary to translate and abstract its publications, and articles contained therein, for inclusion in various compendiums and similar publications, etc. When an article is submitted for publication by ACES, acceptance of the article implies that ACES has the rights to do all of the things it normally does with such an article.

In connection with its publishing activities, it is the policy of ACES to own the copyrights in its technical publications, and to the contributions contained therein, in order to protect the interests of ACES, its authors and their employers, and at the same time to facilitate the appropriate re-use of this material by others.

The new United States copyright law requires that the transfer of copyrights in each contribution from the author to ACES be confirmed in writing. It is therefore necessary that you execute either Part A-Copyright Transfer Form or Part B-U.S. Government Employee Certification or Part C-Crown Copyright on this sheet and return it to the Managing Editor (or person who supplied this sheet) as promptly as possible.

CLEARANCE OF PAPERS

ACES must of necessity assume that materials presented at its meetings or submitted to its publications is properly available for general dissemination to the audiences these activities are organized to serve. It is the responsibility of the authors, not ACES, to determine whether disclosure of their material requires the prior consent of other parties and if so, to obtain it. Furthermore, ACES must assume that, if an author uses within his/her article previously published and/or copyrighted material that permission has been obtained for such use and that any required credit lines, copyright notices, etc. are duly noted.

AUTHOR/COMPANY RIGHTS

If you are employed and you prepared your paper as a part of your job, the rights to your paper initially rest with your employer. In that case, when you sign the copyright form, we assume you are authorized to do so by your employer and that your employer has consented to all of the terms and conditions of this form. If not, it should be signed by someone so authorized.

NOTE RE RETURNED RIGHTS: Just as ACES now requires a signed copyright transfer form in order to do "business as usual", it is the intent of this form to return rights to the author and employer so that they too may do "business as usual". If further clarification is required, please contact: The Managing Editor, R. W. Adler, Naval Postgraduate School, Code EC/AB, Monterey, CA, 93943, USA (408)656-2352.

Please note that, although authors are permitted to re-use all or portions of their ACES copyrighted material in other works, this does not include granting third party requests for reprinting, republishing, or other types of re-use.

JOINT AUTHORSHIP

For jointly authored papers, only one signature is required, but we assume all authors have been advised and have consented to the terms of this form.

U.S. GOVERNMENT EMPLOYEES

Authors who are U.S. Government employees are not required to sign the Copyright Transfer Form (Part A), but any co-authors outside the Government are.

Part B of the form is to be used instead of Part A only if all authors are U.S. Government employees and prepared the paper as part of their job.

NOTE RE GOVERNMENT CONTRACT WORK: Authors whose work was performed under a U.S. Government contract but who are not Government employees are required so sign Part A-Copyright Transfer Form. However, item 5 of the form returns reproduction rights to the U. S. Government when required, even though ACES copyright policy is in effect with respect to the reuse of material by the general public.

APPLIED COMPUTATIONAL ELECTROMAGNETICS SOCIETY JOURNAL
http://aces.ee.olemiss.edu

INFORMATION FOR AUTHORS

PUBLICATION CRITERIA

Each paper is required to manifest some relation to applied computational electromagnetics. **Papers may address general issues in applied computational electromagnetics, or they may focus on specific applications, techniques, codes, or computational issues.** While the following list is not exhaustive, each paper will generally relate to at least one of these areas:

1. **Code validation.** This is done using internal checks or experimental, analytical or other computational data. Measured data of potential utility to code validation efforts will also be considered for publication.

2. **Code performance analysis.** This usually involves identification of numerical accuracy or other limitations, solution convergence, numerical and physical modeling error, and parameter tradeoffs. However, it is also permissible to address issues such as ease-of-use, set-up time, run time, special outputs, or other special features.

3. **Computational studies of basic physics.** This involves using a code, algorithm, or computational technique to simulate reality in such a way that better, or new physical insight or understanding, is achieved.

4. **New computational techniques**, or new applications for existing computational techniques or codes.

5. **"Tricks of the trade"** in selecting and applying codes and techniques.

6. **New codes, algorithms, code enhancement, and code fixes.** This category is self-explanatory, but includes significant changes to existing codes, such as applicability extensions, algorithm optimization, problem correction, limitation removal, or other performance improvement. **Note: Code (or algorithm) capability descriptions are not acceptable, unless they contain sufficient technical material to justify consideration.**

7. **Code input/output issues.** This normally involves innovations in input (such as input geometry standardization, automatic mesh generation, or computer-aided design) or in output (whether it be tabular, graphical, statistical, Fourier-transformed, or otherwise signal-processed). Material dealing with input/output database management, output interpretation, or other input/output issues will also be considered for publication.

8. **Computer hardware issues.** This is the category for analysis of hardware capabilities and limitations of various types of electromagnetics computational requirements. Vector and parallel computational techniques and implementation are of particular interest.

Applications of interest include, but are not limited to, antennas (and their electromagnetic environments), networks, static fields, radar cross section, shielding, radiation hazards, biological effects, electromagnetic pulse (EMP), electromagnetic interference (EMI), electromagnetic compatibility (EMC), power transmission, charge transport, dielectric, magnetic and nonlinear materials, microwave components, MEMS technology, MMIC technology, remote sensing and geometrical and physical optics, radar and communications systems, fiber optics, plasmas, particle accelerators, generators and motors, electromagnetic wave propagation, non-destructive evaluation, eddy currents, and inverse scattering.

Techniques of interest include frequency-domain and time-domain techniques, integral equation and differential equation techniques, diffraction theories, physical optics, moment methods, finite differences and finite element techniques, modal expansions, perturbation methods, and hybrid methods. This list is not exhaustive.

A unique feature of the Journal is the publication of unsuccessful efforts in applied computational electromagnetics. Publication of such material provides a means to discuss problem areas in electromagnetic modeling. Material representing an unsuccessful application or negative results in computational electromgnetics will be considered for publication only if a reasonable expectation of success (and a reasonable effort) are reflected. Moreover, such material must represent a problem area of potential interest to the ACES membership.

Where possible and appropriate, authors are required to provide statements of quantitative accuracy for measured and/or computed data. This issue is discussed in "Accuracy & Publication: Requiring, quantitative accuracy statements to accompany data," by E. K. Miller, *ACES Newsletter*, Vol. 9, No. 3, pp. 23-29, 1994, ISBN 1056-9170.

EDITORIAL REVIEW

In order to ensure an appropriate level of quality control, papers are peer reviewed. They are reviewed both for technical correctness and for adherence to the listed guidelines regarding information content.

JOURNAL CAMERA-READY SUBMISSION DATES

March issue	deadline 8 January
July issue	deadline 20 May
November issue	deadline 20 September

Uploading an acceptable camera-ready article after the deadlines will result in a delay in publishing this article.

STYLE FOR CAMERA-READY COPY

The ACES Journal is flexible, within reason, in regard to style. However, certain requirements are in effect:

1. The paper title should NOT be placed on a separate page. The title, author(s), abstract, and (space permitting) beginning of the paper itself should all be on the first page. The title, author(s), and author affiliations should be centered (center-justified) on the first page.

2. An abstract is REQUIRED. The abstract should be a brief summary of the work described in the paper. It should state the computer codes, computational techniques, and applications discussed in the paper (as applicable) and should otherwise be usable by technical abstracting and indexing services.

3. Either British English or American English spellings may be used, provided that each word is spelled consistently throughout the paper.

4. Any commonly-accepted format for referencing is permitted, provided that internal consistency of format is maintained. As a guideline for authors who have no other preference, we recommend that references be given by author(s) name and year in the body of the paper (with alphabetical listing of all references at the end of the paper). Titles of Journals, monographs, and similar publications should be in italic font or should be underlined. Titles of papers or articles should be in quotation marks.

5. Internal consistency shall also be maintained for other elements of style, such as equation numbering. As a guideline for authors who have no other preference, we suggest that equation numbers be placed in parentheses at the right column margin.

6. The intent and meaning of all text must be clear. For authors who are NOT masters of the English language, the ACES Editorial Staff will provide assistance with grammar (subject to clarity of intent and meaning).

7. Unused space should be minimized. Sections and subsections should not normally begin on a new page.

PAPER FORMAT

The preferred format for initial submission and camera-ready manuscripts is 12 point Times Roman font, single line spacing and double column format, similar to that used here, with top, bottom, left, and right 1 inch margins. Manuscripts should be prepared on standard 8.5x11 inch paper.

Only camera-ready electronic files are accepted for publication. The term **"camera-ready" means that the material is neat, legible, and reproducible.** Full details can be found on ACES site, Journal section.

ACES reserves the right to edit any uploaded material, however, this is not generally done. It is the author(s) responsibility to provide acceptable camera-ready pdf files. Incompatible or incomplete pdf files will not be processed, and authors will be requested to re-upload a revised acceptable version.

SUBMITTAL PROCEDURE

All submissions should be uploaded to ACES server through ACES web site (http://aces.ee.olemiss.edu) by using the upload button, journal section. Only pdf files are accepted for submission. The file size should not be larger than 5MB, otherwise permission from the Editor-in-Chief should be obtained first. The Editor-in-Chief will acknowledge the electronic submission after the upload process is successfully completed.

COPYRIGHTS AND RELEASES

Each primary author must sign a copyright form and obtain a release from his/her organization vesting the copyright with ACES. Copyright forms are available at ACES, web site (http://aces.ee.olemiss.edu). To shorten the review process time, the executed copyright form should be forwarded to the Editor-in-Chief immediately after the completion of the upload (electronic submission) process. Both the author and his/her organization are allowed to use the copyrighted material freely for their own private purposes.

Permission is granted to quote short passages and reproduce figures and tables from and ACES Journal issue provided the source is cited. Copies of ACES Journal articles may be made in accordance with usage permitted by Sections 107 or 108 of the U.S. Copyright Law. This consent does not extend to other kinds of copying, such as for general distribution, for advertising or promotional purposes, for creating new collective works, or for resale. The reproduction of multiple copies and the use of articles or extracts for commercial purposes require the consent of the author and specific permission from ACES. Institutional members are allowed to copy any ACES Journal issue for their internal distribution only.

PUBLICATION CHARGES

ACES members are allowed 12 printed pages per paper without charge; non-members are allowed 8 printed pages per paper without charge. Mandatory page charges of $75 a page apply to all pages in excess of 12 for members or 8 for non-members. Voluntary page charges are requested for the free (12 or 8) pages, but are NOT mandatory or required for publication. A priority courtesy guideline, which favors members, applies to paper backlogs. Authors are entitled to 15 free reprints of their articles and must request these from the Managing Editor. Additional reprints are available to authors, and reprints available to non-authors, for a nominal fee.

ACES Journal is abstracted in INSPEC, in Engineering Index, and in DTIC.