ELECTROMAGNETIC COMMUNICATION LABORATORY
TECHNICAL REPORT NO. 85-9

December 1985

ON THE IMPLEMENTATION AND PERFORMANCE OF ITERATIVE METHODS FOR COMPUTATIONAL ELECTROMAGNETICS

A. F. Peterson
R. Mittra

Supported by

Office of Naval Research
Contract No. N00014-81-K-0245
ONR Fellowship
Texas Instruments

ELECTROMAGNETIC COMMUNICATION LABORATORY
DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN
URBANA, ILLINOIS 61801

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) ON THE IMPLEMENTATION AND PERFORMANCE OF ITERATIVE METHODS FOR COMPUTATIONAL ELECTROMAGNETICS		5. TYPE OF REPORT & PERIOD COVERED Technical Report 6. PERFORMING ORG. REPORT NUMBER EC 85-9; UILU-ENG-85-2571
7. AUTHOR(s) A. F. Peterson and R. Mittra		8. CONTRACT OR GRANT NUMBER(s) N00014-81-K-0245 ONR Fellowship Texas Instruments
9. PERFORMING ORGANIZATION NAME AND ADDRESS Electromagnetic Communication Laboratory Department of Elec. and Computer Engg. University of Illinois, Urbana, Illinois 61801		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Project No. 410
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Department of the Navy Arlington, Virginia 22217		12. REPORT DATE December 1985 13. NUMBER OF PAGES 197
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Distribution Unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Moment Method; Electromagnetic Scattering; Spectral Iterative Technique; K-space Method; Conjugate Gradient Method

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The numerical solution of electromagnetic scattering problems involves approximating an exact equation by a finite-dimensional matrix equation. The use of an iterative algorithm to solve the matrix equation sometimes results in a considerable savings in computer memory requirements. For a fixed amount of computer memory, this approach permits the analysis of scatterers that are an order of magnitude larger electrically.

(over)

Iterative algorithms of the conjugate gradient class are examined and applied to a variety of typical electromagnetic scattering problems, in order to evaluate their performance in practice. ~~In contrast with the simple iterativ algorithms used in the past, which often diverged when applied to electromagnetics problems,~~ these algorithms never diverge and usually converge at a quick rate.

Depending on the geometry of the scatterer under consideration, it may be possible to build symmetries into the matrix representation and effect the necessary storage reduction. Two distinct approaches for creating these symmetries are examined. An alternate procedure, which requires some of the matrix elements to be regenerated as needed by the iterative algorithm in use, does not rely on symmetries and is applicable to a larger set of geometries. Both procedures are applied to several scattering problems. Execution time comparisons show that the approaches based on symmetries are the most efficient, and that both procedures can be superior to noniterative techniques for large scatterers.

UILU-ENG-85-2571

Electromagnetic Communication Laboratory Report No. 85-9

ON THE IMPLEMENTATION AND PERFORMANCE OF ITERATIVE METHODS FOR COMPUTATIONAL ELECTROMAGNETICS

by

A. F. Peterson and R. Mittra

Technical Report

December 1985

Supported by

Office of Naval Research
Contract No. N00014-81-K-0245
ONR Fellowship
Texas Instruments

Electromagnetic Communication Laboratory
Department of Electrical and Computer Engineering
University of Illinois at Urbana-Champaign
Urbana, Illinois 61801

ABSTRACT

The numerical solution of electromagnetic scattering problems involves approximating an exact equation by a finite-dimensional matrix equation. The use of an iterative algorithm to solve the matrix equation sometimes results in a considerable savings in computer memory requirements. For a fixed amount of computer memory, this approach permits the analysis of scatterers that are an order of magnitude larger electrically.

Iterative algorithms of the conjugate gradient class are examined and applied to a variety of typical electromagnetic scattering problems, in order to evaluate their performance in practice. In contrast with the simple iterative algorithms used in the past, which often diverged when applied to electromagnetics problems, these algorithms never diverge and usually converge at a quick rate.

Depending on the geometry of the scatterer under consideration, it may be possible to build symmetries into the matrix representation and effect the necessary storage reduction. Two distinct approaches for creating these symmetries are examined. An alternate procedure, which requires some of the matrix elements to be regenerated as needed by the iterative algorithm in use, does not rely on symmetries and is applicable to a larger set of geometries. Both procedures are applied to several scattering problems. Execution time comparisons show that the approaches based on symmetries are the most efficient, and that both procedures can be superior to noniterative techniques for large scatterers.

ACKNOWLEDGEMENTS

The work in this report was supported in part by the Office of Naval Research under Grant N00014-81-K-0245 to the University of Illinois and through a Predoctoral fellowship support to A. Peterson. It was also supported in part by the Texas Instruments Company of Dallas, Texas. Helpful discussion with various members of the Electromagnetics and Electromagnetic Communication Laboratories and with Dr. Oren Kesler of the Antenna Department of TI are gratefully acknowledged.

Finally, thanks are due to the Publication Office of Electrical and Computer Engineering Department for compiling the manuscript.

TABLE OF CONTENTS

Page

LIST OF TABLES

TABLE Page

LIST OF FIGURES

Figure Page

Figure Page

Figure Page

Figure Page

1. INTRODUCTION

Electromagnetic scattering problems comprise a variety of physical phenomena, from the interaction of radio waves with aircraft to the biological effects of microwaves on the human body. Three distinct approaches have evolved for the analysis of scattering problems. The first, exact analytical techniques, include approaches such as separation of variables [1], [2] and the Wiener-Hopf method [3]. Although exact techniques have been developed and applied to electromagnetics since the appearance of Maxwell's work [4], they remain limited to problems which can be described by relatively simple geometries. The second, approximate analytical methods, such as the variational methods [5] and the asymptotic approaches including geometric optics [6] and geometric theory of diffraction [7], can sucessfully treat many types of problems not amenable to exact solution. The asymptotic methods are usually limited to the analysis of electrically large scatterers whose geometry can be described in terms of the few canonical shapes for which diffraction coefficients are available. The success of these methods is usually highly dependent on the intuition and skill of the user, and systematic estimates of the accuracy are usually impossible to obtain. The final category, numerical solutions, encompasses the many computer-aided approaches developed over the past several decades [8], [9]. Numerical solutions are not fundamentally restricted to scatterers with certain canonical shapes or materials, and in principle they can be carried out to obtain any level of accuracy. They are, however, limited in practice by the available computer resources. Because of this limitation, conventional numerical methods are best suited for the analysis of electrically small scatterers, i.e., scatterers whose maximum dimensions are several wavelengths or less. Recently, specialized numerical techniques have been introduced for the efficient treatment of larger

scatterers [10] - [13]. The purpose of this thesis is to review these recent developments and extend these techniques to other applications.

Every numerical solution process can be separated into three parts. First, the physical problem must be cast into the form of a mathematical equation. Since few equations arising in practice are convenient to solve analytically, the second part of the process involves replacing the original equation by a finite dimensional matrix equation, which is then amenable to solution. This is known as discretization. The final part of the process is to solve the matrix equation, either by direct methods such as Gaussian elimination or by iterative methods.

Initially, the problem may be posed as a differential equation with boundary conditions, an integral equation, or some variational principle to be satisfied. Furthermore, the problem may be posed in the time or frequency domain. The most popular approach for the numerical treatment of electromagnetic scattering problems has been the integral equation formulation [14] - [16], although progress has been reported using other formulations [17]. Often, these integral equations are convolutional in form, and this property is of central importance for many of the specialized methods to be considered below.

The discretization process may involve the direct sampling of pertinent quantities over a grid of points, often used in connection with finite-difference approximations to derivatives [18]. If the problem is posed in terms of a variational principle, the discretization may take the form of the Ritz procedure where the unknown quantity is approximated by several trial functions [19], [20]. The approach widely used to discretize the integral equations of electromagnetics is a generalization of the Ritz procedure known as the method of moments (MoM) [21] or the weighted residual method [22]. The MoM received widespread attention during the 1960's, and excellent reviews of its

early use in electromagnetics are provided by Richmond [23], Harrington [24] - [26], and Tanner and Andreasen [27]. The MoM procedure requires the unknown function to be approximated by N expansion functions, and reduces the task to that of finding N unknown coefficients. The resulting equation is made orthogonal to N testing functions, to yield an NxN matrix equation. This procedure is illustrated in Chapter 5, and additional information regarding the MoM may be found in several recent textbooks [8], [25], [28], [29].

Algorithms for the solution of matrix equations are also available in many texts [30] - [34]. Conventional numerical analysis usually incorporates a direct method, such as Gaussian elimination, to solve the matrix equation. The specialized techniques developed to treat large scatterers usually use an iterative method to solve the system because general purpose elimination algorithms require the NxN matrix to be stored in computer memory. Computer memory consists of "fast access" in-core memory and "slow access" out-of-core memory, the latter including peripheral devices such as disk and magnetic tape facilities. Because of physical limitations, only relatively small order (i.e., orders of several hundred) matrices can be stored in most computers without some use of out-of-core memory. Unfortunately, the order of the matrix increases with the electrical size of the scatterer, and frequently exceeds the fast-access limits of a given machine. Although advances in computer architecture continuously improve the efficiency of transferring data to and from peripheral storage facilities, such transfers remain extremely slow compared to in-core access times. Therefore, the specialized techniques of interest attempt to reduce the necessary storage to a level which can be handled by in-core computer memory alone, or at least minimize the transfers to out-of-core devices. Iterative algorithms do not require the NxN matrix to be stored in computer memory, and thus, are highly compatible with the specialized techniques

because they can take into account any sparseness or redundancy of the matrix to reduce the storage burden.

Originally, specialized techniques for large scatterers incorporated "simple" iterative algorithms of the Jacobi, Gauss-Seidel, or Successive Over-Relaxation (SOR) variety, which are described in detail in the literature [30] - [34]. These algorithms can be problematic, because they sometimes diverge. More "sophisticated" iterative algorithms, such as the gradient methods [30], [33], are more complicated to implement but never diverge. This report will only consider three algorithms based on gradient methods, all of which are developed in Chapter 2.

There is one major drawback to the use of iteration for computational electromagnetics. The efficiency of iteration is poor compared to direct methods of solution whenever multiple systems involving the same NxN matrix must be treated. Because direct methods in effect generate an inverse matrix, they can efficiently solve a matrix equation involving any number of right-hand sides. To date, no general, systematic iterative procedure is available with similar efficiency. The relative inefficiency of iteration holds to some degree in spite of the fact that out-of-core storage procedures may be necessary for the implementation of a direct method. For electromagnetic scattering problems, there is a trade-off between these two approaches and this trade-off apparently has not been studied in detail to date.

All of the well-known frequency-domain based techniques for the efficient analysis of electrically large scatterers have incorporated three features [10] - [13]. First, they involve a convolutional integral equation formulation. Second, they are restricted to geometries which can be discretized in such a manner as to preserve the convolution form in the matrix equation. (Such a system consists of one or more discrete convolution operations, meaning that it

possesses a significant degree of redundancy.) Finally, they incorporate iterative algorithms to solve the matrix, and attempt to implement these algorithms so as to minimize the computer memory requirements.

Although this report will not discuss time-domain based techniques in any detail, they deserve some mention because they have also been applied to the efficient analysis of electrically large scatterers [35] - [38]. Time-domain approaches can require considerably less computer memory and execution time than conventional frequency-domain based methods, and appear to be comparable in efficiency to the specialized frequency-domain methods considered here. General reviews of numerical time-domain techniques are provided by Mittra [39], Bennett and Ross [40], and Miller and Landt [41].

The first specialized frequency-domain based technique for the analysis of electrically large scatterers was published in a comprehensive report by Bojarski in 1971 [10]. Although Bojarski's "K-space" formulation made extensive mention of the Fourier transform and the fast-Fourier transform (FFT) algorithm, the distinguishing features of the approach as a numerical technique included its use of a convolutional integral equation, an evenly-spaced sampling grid which preserved the convolution form in the matrix equation, and a "simple" iterative algorithm to solve the system. Bojarski made extensive use of the FFT, which allows a large order discrete convolution to be performed more efficiently than possible by direct matrix multiplication [42], [43]. However, the FFT can only be directly applied to an infinite-periodic sequence of numbers; the treatment of non-periodic sequences requires the incorporation of zero-padding [42]. Because of this, any use of the FFT algorithm for non-periodic scattering problems increases the array sizes and the computer memory requirements considerably.

Ko and Mittra [44], and later Kastner and Mittra [11], [45], [46], developed a technique similar to Bojarski's that they applied to a variety of applications. Entitled the "Spectral-Iterative" technique (SIT), it differed from the "K-space" method in two respects. First, the "K-space" method required a three-dimensional FFT and three-dimensional zero-padding to treat a three-dimensional non-periodic scatterer. The SIT required only a two-dimensional FFT to treat a three-dimensional geometry, and thus eliminated some of the cumbersome zero-padding. Since the SIT required the discrete convolution to be summed explicitly in the remaining dimension, it would require somewhat more computation per iteration step than the "K-space" method for large scatterers. It is important to realize that an explicit summation can be substituted for the FFT whenever desired, and may be necessary if the additional storage constraints imposed by the FFT algorithm exceed the fast-access limits of a given machine. The second difference between the SIT and the "K-space" method lies in the discretization employed. The discretization used within the SIT required the direct sampling of the analytical Fourier transform of the kernel (Green's function) appearing in the integral equation, as opposed to a sampling of the kernel itself as was done in the "K-space" method. The difference between these two approaches is manifested in the difference between the Fourier transform and the FFT, although other factors are also involved.

Since the FFT treats a sequence as periodic in both the original and the transform domains, the validity of mixing the analytical transform and the FFT for non-periodic functions comes under question. Because of this, Bojarski specifically recommended against the type of discretization that was later used with the SIT [10]. In a variety of applications, some of which are illustrated in Chapter 5, it is desirable to sample the Fourier transform directly because the Green's function is not well suited for numerical computation. In order to

ascertain the validity of the type of discretization in which this is done, and to develop guidelines for its use, Chapter 5 presents a detailed comparison of the two discretization procedures.

Although both the "K-space" method and the SIT were applied to a variety of problems, both methods were inconvenient to use in practice because the "simple" iterative algorithm used to solve the matrix equation often diverged. Recently, Sultan and Mittra incorporated a "sophisticated" iterative algorithm, the conjugate gradient method (CGM), into the SIT procedure [13]. The CGM has only recently received widespread attention in electromagnetics, and thus its performance in practice is largely undocumented. In an attempt to alleviate this situation, Chapter 2 describes the CGM and two related algorithms in detail, and Chapter 3 discusses the performance of the algorithms for a variety of scattering problems.

Other iterative-based methods have been developed to efficiently treat scattering problems. Nyo and Harrington [12], [47] and Borup and Gandhi [48] applied the MoM discretization to geometries which preserved the discrete convolutional form of the original equation. The resulting matrix equations were solved with a simple iterative algorithm, although Borup and Gandhi later incorporated the CGM into their approach [49]. This discretization procedure can be systematically applied to many problems, and is named the discrete-convolutional method of moments (DCMoM). The DCMoM is actually a generalization of the approach originally used by Bojarski [10], and is briefly described in Chapter 5. Ray and Mittra [50] and Hurst and Mittra [51] used a DCMoM discretization with the CGM to find the current density induced on large plates and finite cylinders, and showed that the procedure could be efficiently applied to treat matrix equations with orders in excess of 4000.

Van den Berg used the CGM and several related algorithms to analyze scattering from conducting strips and dielectric cylinders, and presented a convergence rate comparison which shows that these algorithms are superior to a "simple" iterative algorithm [52]. Bokhari and Balakrishnan used a simple iterative algorithm for the anaylsis of a cylindrical antenna, and reported a way of always ensuring the convergence of the algorithm by the use of an appropriate amount of zero-padding [53].

All of the above approaches rely on the presence of discrete-convolutional symmetries in the matrix equation to permit a reduction in the necessary computer memory. These are a slightly more general form of the Toeplitz and block-Toeplitz symmetries for which specialized direct methods of solution have been developed [54], [55]. These direct algorithms permit the same storage reduction as iterative techniques, and in addition may be more efficient than iteration. However, slightly perturbed Toeplitz systems and Toeplitz systems embedded in larger matrix equations often arise in practice, and are not easily treated by direct algorithms. For instance, the pertinent matrix equations often have discrete-convolutional symmetries everywhere except along the main diagonal, which upsets the Toeplitz form of the system without changing the significant redundancy features. Examples of almost-Toeplitz systems are discussed in Chapters 6 and 7. Historically, it appears that the primary use of iteration in electromagnetics has been to solve equations having slightly perturbed Toeplitz symmetries.

Because of the reliance on discrete-convolutional symmetries, the specialized techniques discussed above are only optimum for a limited set of problems. These include geometries which can be described by surfaces of constant curvature (such as flat plates or right-circular cylinders) and geometries which lend themselves to volume discretization (such as penetrable

dielectric bodies). Although it may always be possible to treat an arbitrarily shaped structure with a volume discretization, this approach often requires a large number of fictitious unknowns and can not be considered optimum for problems other than penetrable bodies. In other words, in a given situation it may not be possible to cast the matrix equation into discrete-convolutional form. An alternate possibility for the efficient analysis of large scatterers that does not require the matrix to possess discrete-convolutional symmetries is the matrix element regeneration (MER) scheme recently proposed by Peterson, Sultan, and Mittra [56] - [59]. Chapter 4 discusses the implementation of the MER. Since the MER does not rely on symmetries in the system, the approach is not generally as efficient as the specialized techniques discussed above.

The above discussion has emphasized recent trends in the numerical analysis of electrically large scattering problems. The purpose of this report is to interpret these trends, extend the state of knowledge concerning certain aspects of these approaches which have heretofore gone unreported, and effectively implement similar techniques for other applications. Specifically, Chapter 2 presents several iterative algorithms related to the conjugate gradient method and shows how these algorithms can be obtained from two different viewpoints. This material may be beneficial to anyone attempting to modify existing algorithms to improve their rate of convergence, as has been recently proposed [60]. Chapter 3 illustrates the convergence of the iterative algorithms of Chapter 2 when applied to integral equations of typical electromagnetic scattering problems. Chapter 4 begins the topic of efficient implementation of iterative algorithms, with an investigation into the matrix element regeneration (MER) approach. Chapter 5 provides an in-depth comparison of two discretizations often used in conjunction with iterative methods, the discrete-convolutional method of moments (DCMoM) and the spectral-domain fast-Fourier

transform (SDFFT) procedures, the latter denoting the discretization used with the original SIT [11]. Chapters 6 and 7 concern the efficient implementation of iterative approaches for the problems of scattering from a dielectric body and scattering by a finite hollow conducting cylinder. Conclusions and suggestions for future work are given in Chapter 8.

The work reported within is intended to build upon the previous efforts of others, especially R. Kastner, M. Sultan, and S. Ray. Based upon the foundation set by these and other researchers, the application of iterative methods to numerical electromagnetics has today reached a stage where the systematic implementation of a variety of problems is well-understood. Appropriate choices may be made from available discretizations and available iterative algorithms, in a relatively independent manner. As improvements are made in either of these two areas, they can be incorporated into the existing approaches. This study examines several aspects of iterative computational electromagnetics, and should be considered a small step in the continuing attempts to improve the efficiency of numerical techniques for solving electromagnetic scattering problems. Preliminary results from this investigation have been previously published [56], [57], [59], [61], [62].

2. THE CONJUGATE GRADIENT METHOD AND RELATED ALGORITHMS

2.1. Introduction

The computer-aided analysis of electrically large scatterers is often based upon an iterative solution algorithm [10] - [13]. In this chapter, several iterative "gradient" methods are developed and presented in a form suitable for this type of analysis. Of central importance is the conjugate gradient method (CGM), which is used extensively in later chapters. In the literature, the CGM is usually presented in a specialized form, which is useful only for symmetric, positive definite systems [30], [31], [33], [63], [64]. Here, the general algorithm is developed from two different perspectives. In addition, several approaches that are related to the CGM are discussed.

Initially, the CGM is developed from a generalized variational procedure, i.e., the minimization of an error functional. The CGM can also be developed from an orthogonal expansion process, without reference to a functional. However, it will be apparent that these two constructs are actually linked, that is, each functional has associated with it a certain type of orthogonal expansion. The algorithms reduce to the process of generating the orthogonal functions and finding the proper coefficients to represent the desired solution.

It suffices to consider the nonsingular matrix equation

$$Lf = h \tag{2.1}$$

where L denotes a linear operator (N × N matrix), f is an unknown function to be determined (N × 1 matrix), and h is a given right-hand side (also an N × 1 matrix). An inner product and its associated norm are defined

$$\langle f,g \rangle = \overline{g}^T f \tag{2.2}$$

$$||f|| = \sqrt{\langle f, f \rangle} \tag{2.3}$$

where the 'T' denotes matrix transpose and the bar complex conjugation. It is desired to iteratively generate the solution f, so that at some stage of the process we seek an estimate of the solution in the form

$$f_n = f_{n-1} + \alpha_n p_n \tag{2.4}$$

where f_{n-1} is a previous estimate of the solution, p_n is a function which is as yet undetermined, and α_n is a scalar coefficient.

2.2. Functional Minimization

Assuming a function p_n is available, the coefficient α_n in Equation (2.4) can be chosen so that f_n is a "better" estimate of the solution than f_{n-1}, in the sense that f_n minimizes some measure of the error. For instance, the quadratic functionals

$$F_1(f_n) = ||Lf_n - h||^2 \tag{2.5}$$

$$F_2(f_n) = ||f - f_n||^2 \tag{2.6}$$

provide two possible definitions of the error in f_n. The functional

$$F_3(f_n) = \langle Lf_n - h, f - f_n \rangle \tag{2.7}$$

can also be used to define the error in f_n, although a useful iterative procedure can be based on F_3 only if L is self-adjoint and positive definite in the sense of Griffel [65]. Each of the above functionals gives rise to a different iterative procedure for the solution of Equation (2.1); other functionals could also be used and would lead to other algorithms.

Functional F_1, as defined in Equation (2.5), is associated with the standard form of the CGM for arbitrary linear operators. For a given p_n, the coefficient α_n of Equation (2.4) that minimizes F_1 is found by setting the first variation of F_1 with respect to α_n to zero, which is equivalent to

$$2\mathrm{Re}\{\delta(\overline{\alpha_n})(\langle Lf_{n-1} - h, Lp_n\rangle + \alpha_n ||Lp_n||^2)\} = 0 \tag{2.8}$$

for arbitrary variation $\delta(\overline{\alpha_n})$. This is satisfied for

$$\alpha_n = \frac{-\langle R_{n-1}, Lp_n\rangle}{||Lp_n||^2} \tag{2.9}$$

where for convenience we define

$$R_n = Lf_n - h \tag{2.10}$$

If it is desired to simultaneously find two coefficients which minimize $F_1(\tilde{f}_n)$ where

$$\tilde{f}_n = f_{n-1} + \tilde{\alpha}_n p_n + \tilde{\gamma}_n q_n \tag{2.11}$$

the above process can be extended to produce

$$\tilde{\alpha}_n = \frac{-\langle R_{n-1}, Lp_n\rangle ||Lq_n||^2 + \langle R_{n-1}, Lq_n\rangle\langle Lq_n, Lp_n\rangle}{||Lp_n||^2\ ||Lq_n||^2 - |\langle Lp_n, Lq_n\rangle|^2} \tag{2.12}$$

$$\tilde{\gamma}_n = \frac{-\langle R_{n-1}, Lq_n\rangle ||Lp_n||^2 + \langle R_{n-1}, Lp_n\rangle\langle Lp_n, Lq_n\rangle}{||Lp_n||^2\ ||Lq_n||^2 - |\langle Lp_n, Lq_n\rangle|^2} \tag{2.13}$$

From Equations (2.12) and (2.13) it is apparent that if the functions p_n and q_n satisfy

$$\langle Lp_n, Lq_n \rangle = 0 \tag{2.14}$$

the expressions for the coefficients decouple and reduce to the form presented in Equation (2.9). This indicates that if a sequence of functions satisfying Equation (2.14) is used to expand the solution f of Equation (2.1), the coefficient of each function may be found independently. Under these conditions, the process cannot be improved upon by simultaneously computing several coefficients.

It is also of interest to consider a solution estimate of the form

$$\tilde{f}_n = f_{n-1} + \alpha_n(p_n + \beta_n q_n) \tag{2.15}$$

In the manner outlined above, the β_n that minimizes $F_1(\tilde{f}_n)$ assuming optimum α_n is given by

$$\beta_n = \frac{\langle R_{n-1}, Lq_n \rangle ||Lp_n||^2 - \langle R_{n-1}, Lp_n \rangle \langle Lp_n, Lq_n \rangle}{\langle R_{n-1}, Lp_n \rangle ||Lq_n||^2 - \langle R_{n-1}, Lq_n \rangle \langle Lq_n, Lp_n \rangle} \tag{2.16}$$

Suppose that p_n and q_n are arbitrary functions which do not satisfy Equation (2.14). Suppose also that q_n has already been used in the expansion process, so that

$$f_{n-1} = f_{n-2} + \alpha_{n-1} q_n \tag{2.17}$$

where α_{n-1} is found according to Equation (2.9). It immediately follows that

$$\langle R_{n-1}, Lq_n \rangle = \langle R_{n-2} + \alpha_{n-1} Lq_n, Lq_n \rangle = 0 \tag{2.18}$$

$$\beta_n = \frac{-\langle Lp_n, Lq_n \rangle}{||Lq_n||^2} \tag{2.19}$$

and

$$\langle Lq_n, L(p_n + \beta_n q_n)\rangle = 0 \tag{2.20}$$

Thus, the process of selecting coefficients that minimize the functional F_1 is optimized when functions satisfying Equation (2.14) are used. If arbitrary functions are used, and their coefficients adjusted simultaneously to produce a minimum F_1, the process will adjust these coefficients in order to produce functions satisfying Equation (2.14).

2.3. The Conjugate Direction Method

The previous section presented a procedure for finding coefficients of trial functions that minimize an error functional, for the solution of Equation (2.1). It was seen that a sequence of functions $\{p_n\}$ satisfying

$$\langle Lp_i, Lp_j\rangle = 0 \qquad i \neq j \tag{2.21}$$

arose naturally in connection with the error functional F_1 defined in Equation (2.5). If we assume the existence of such a sequence, and note that in a finite dimensional vector space N of these p-functions will span the space, the solution can be written

$$f = f_0 + \alpha_1 p_1 + \ldots + \alpha_n p_n + \ldots + \alpha_N p_N \tag{2.22}$$

where f_0 is an arbitrary initial estimate of the solution. Because of the orthogonality expressed in Equation (2.21), each coefficient can be found independently as

$$\alpha_n = \frac{-\langle R_0, Lp_n\rangle}{||Lp_n||^2} \tag{2.23}$$

where R_n is defined in Equation (2.10). Note that Equation (2.23) and Equation (2.9) appear to differ; it will be shown below that they are, in fact, equal.

From the above relationships, it is apparent that

$$R_n = R_0 + \alpha_1 Lp_1 + \ldots + \alpha_n Lp_n \tag{2.24}$$

and recursive relationships are given as

$$R_n = R_{n-1} + \alpha_n Lp_n \tag{2.25}$$

$$f_n = f_{n-1} + \alpha_n p_n \tag{2.26}$$

The "residual norm" defined by

$$N_n = \frac{||R_n||}{||h||} = \frac{||Lf_n - h||}{||h||} \tag{2.27}$$

is one indication of the accuracy of the solution at the n-th step in the expansion process. Because of the orthogonality of the p-functions, Equation (2.24) can be used to show that

$$||R_n||^2 = ||R_{n-1}||^2 - |\alpha_n|^2 \, ||Lp_n||^2 \tag{2.28}$$

and, thus, the residual norm decreases monotonically as n increases. Of course, the residual norm is directly proportional to the error functional F_1 discussed above, and must decrease at each step if a minimization is performed. From Equations (2.21), (2.23), and (2.24) we can readily deduce that

$$\langle R_n, Lp_m \rangle = \begin{cases} \langle R_0, Lp_m \rangle & n < m \\ 0 & n \geq m \end{cases} \tag{2.29}$$

This relationship shows the equivalence of Equations (2.9) and (2.23) and will prove useful below.

Functions that satisfy Equation (2.21) are said to be mutually conjugate with respect to the operator $L^A L$, where L^A is the adjoint to L defined by

$$\langle L^A R, P \rangle = \langle R, LP \rangle \tag{2.30}$$

The above process of expanding a solution in conjugate functions is known as the "conjugate direction method," after Hestenes and Stiefel [63]. The conjugate direction method does not specify a procedure for the generation of mutually conjugate functions. We now consider such a procedure.

2.4. The Conjugate Gradient Method (CGM)

An approach that is based on the conjugate direction method and includes an algorithm for the generation of mutually conjugate p-functions was developed by Hestenes and Stiefel and named the "conjugate gradient" method [63]. The process begins with the choice

$$p_1 = -L^A R_0 \tag{2.31}$$

Subsequent functions are found from

$$p_{n+1} = -L^A R_n + \beta_n p_n \tag{2.32}$$

where the scalar coefficient β_n is chosen to ensure

$$\langle L^A L p_n, p_{n+1} \rangle = 0 \tag{2.33}$$

It will be shown that due to the special nature of the R-functions, this process yields a mutually conjugate set. First, some preliminary relationships are presented.

Based upon Equation (2.32), we write

$$\langle p_{n+1}, L^A R_m \rangle = -\langle L^A R_n, L^A R_m \rangle + \beta_n \langle p_n, L^A R_m \rangle \tag{2.34}$$

From Equation (2.29), the first and last inner products in Equation (2.34) vanish for $m > n$, leaving

$$\langle L^A R_n, L^A R_m \rangle = 0 \qquad m \neq n \tag{2.35}$$

Equations (2.32) and (2.29) can be combined to yield

$$\langle L^A R_n, p_{n+1} \rangle = -\langle L^A R_n, L^A R_n \rangle \tag{2.36}$$

It follows from Equation (2.23) that

$$\alpha_n = \frac{||L^A R_{n-1}||^2}{||L p_n||^2} \tag{2.37}$$

From Equation (2.25), we have

$$L^A R_n = L^A R_{n-1} + \alpha_n L^A L p_n \tag{2.38}$$

An inner product between $L^A R_m$ and Equation (2.38) leads to the result

$$\langle L^A L p_n, L^A R_m \rangle = \begin{cases} \dfrac{||L^A R_m||^2}{\alpha_n} & m = n \\ \dfrac{-||L^A R_m||^2}{\alpha_n} & m = n - 1 \\ 0 & \text{otherwise} \end{cases} \tag{2.39}$$

Using Equation (2.39) with $m = n$, we find the value of β_n from Equations (2.32) and (2.33) to be

$$\beta_n = \frac{||L^A R_n||^2}{||L^A R_{n-1}||^2} \tag{2.40}$$

To see that β_n guarantees the proper orthogonality between functions when Equation (2.21) is not enforced explicitly, consider the above process as an iterative method. Given some "initial guess" f_o, we generate R_0, p_1, R_1, p_2, ... in the prescribed manner. Since Equation (2.33) is explicitly enforced we know that

$$\langle L^A R_0, L^A R_1 \rangle = 0 \tag{2.41}$$

$$\langle L^A R_0, L^A R_2 \rangle = 0 \tag{2.42}$$

$$\langle L^A R_1, L^A R_2 \rangle = 0 \tag{2.43}$$

What remains is the validity of

$$\langle L^A L p_1, p_3 \rangle = 0 \tag{2.44}$$

By direct computation p_3 can be written as

$$p_3 = -||L^A R_2||^2 \sum_{i=0}^{2} \frac{L^A R_i}{||L^A R_i||^2} \tag{2.45}$$

and it follows that

$$\langle L^A L p_1, p_3 \rangle = -||L^A R_2||^2 \sum_{i=0}^{2} \frac{\langle L^A L p_1, L^A R_i \rangle}{||L^A R_i||^2} \tag{2.46}$$

But, by the relationship established in Equation (2.39), the above reduces to

$$\langle L^A L p_1, p_3 \rangle = -||L^A R_2||^2 \left\{ \frac{-1}{\alpha_1} + \frac{1}{\alpha_1} \right\} = 0 \tag{2.47}$$

Thus, in an inductive fashion we see that the functions generated by the above procedure do in fact satisfy the assumed orthogonality properties of the conjugate direction method.

The conjugate gradient method is given in its entirety as follows:

Initial Steps

Guess f_o

$R_0 = Lf_o - h$

$p_1 = -L^A R_0$

Iterate n = 1,2,...

$$\alpha_n = -\frac{\langle R_{n-1}, Lp_n \rangle}{||Lp_n||^2} = \frac{||L^A R_{n-1}||^2}{||Lp_n||^2}$$

$$f_n = f_{n-1} + \alpha_n p_n$$

$$R_n = Lf_n - h = R_{n-1} + \alpha_n Lp_n$$

$$\beta_n = \frac{||L^A R_n||^2}{||L^A R_{n-1}||^2}$$

$$p_{n+1} = -L^A R_n + \beta_n p_n$$

A discussion of the theoretical convergence rate of the CGM and the convergence in practice is presented in Chapter 3. The remainder of this chapter considers alternate approaches that are related to the algorithms discussed above.

2.5. Generation of Mutually Conjugate Functions

The CGM delineates one approach for the generation of mutually conjugate functions. Given a set of linearly independent functions, a Gram-Schmidt procedure can be used to produce a conjugate set. A form of this process is discussed by Beckman [64] and summarized below.

If $\{\phi_i\}$ denotes a set of linearly independent functions, mutually $L^A L$ - orthogonal functions $\{p_i\}$ can be generated from the formulas

$$p_1 = \phi_1 \tag{2.48}$$

$$p_n = \phi_n + \sum_{k=1}^{n-1} \gamma_{nk} p_k \tag{2.49}$$

where Equation (2.21) is enforced to get

$$\gamma_{nk} = \frac{-\langle L\phi_n, Lp_k \rangle}{||Lp_k||^2} \qquad k < n \tag{2.50}$$

The CGM utilizes a special case of the general Gram-Schmidt procedure, with

$$\phi_n = -L^A R_{n-1} \tag{2.51}$$

Because of the specific orthogonality properties arising within the CGM, the process diagonalizes so that

$$\gamma_{nk} = 0 \qquad k < n - 1 \tag{2.52}$$

Because of this, the CGM generates $L^A L$ - orthogonal functions recursively, without the need to permanently store the p-functions and R-functions in computer memory. This process is summarized as follows, based on a given function R_0:

$$p_1 = -L^A R_0 \tag{2.53}$$

$$R_n = R_{n-1} + ||L^A R_{n-1}||^2 \frac{Lp_n}{||Lp_n||^2} \tag{2.54}$$

$$P_{n+1} = -L^A R_n + \frac{||L^A R_n||^2}{||L^A R_{n-1}||^2} P_n \tag{2.55}$$

As an illustration of a scheme that produces the same R-functions but normalizes the p-functions differently, consider

$$P_1 = \frac{-L^A R_0}{||L^A R_0||^2} \tag{2.56}$$

$$R_n = R_{n-1} + \frac{Lp_n}{||Lp_n||^2} \tag{2.57}$$

$$P_{n+1} = P_n - \frac{L^A R_n}{||L^A R_n||^2} \tag{2.58}$$

Other normalizations could also be used, and would be equivalent to the CGM procedure.

2.6. Minimization in the Domain of the Operator (OGM)

The CGM as presented above was shown to minimize the functional F_1 defined in Equation (2.5) at each step of the procedure. Functional F_1 measures the error in a given estimate of the solution as seen in the range of the operator L. CGM-like algorithms can be developed for many other functionals [66], [67]. For instance, a similar algorithm can be based upon functional F_2 defined in Equation (2.6), and repeated for convenience below:

$$F_2(f_n) = ||f - f_n||^2 \tag{2.59}$$

F_2 provides an error measure made in the domain of the operator L. If a solution is sought in the form

$$f_n = f_{n-1} + \alpha_n q_n \tag{2.60}$$

the value of the coefficient α_n which minimizes F_2 is given by

$$\alpha_n = \frac{\langle f - f_{n-1}, q_n \rangle}{||q_n||^2} \tag{2.61}$$

Since we seek a solution to Equation (2.1), we do not know f at this point. However, we can use the identity

$$(L^{-1})^A = (L^A)^{-1} \tag{2.62}$$

and the previous definition

$$R_n = L(f_n - f) \tag{2.63}$$

to rewrite Equation (2.61) as

$$\alpha_n = \frac{-\langle R_{n-1}, (L^A)^{-1} q_n \rangle}{||q_n||^2} \tag{2.64}$$

which is shown below to be a useful form.

At this point, consider a solution to Equation (2.1) in the form

$$f = f_0 + \alpha_1 q_1 + \ldots + \alpha_N q_N \tag{2.65}$$

where the functions $\{q_n\}$ span the N-dimensional space and satisfy the orthogonality requirement

$$\langle q_i, q_j \rangle = 0 \qquad i \neq j \tag{2.66}$$

It is apparent that the coefficients are given by

$$\alpha_n = \frac{-\langle R_0, (L^A)^{-1} q_n \rangle}{||q_n||^2} \tag{2.67}$$

without regard to minimizing any specific functional. By analogy with the CGM development described previously, we conclude that the functional F_2 is associated with an orthogonal expansion of the type appearing in Equation (2.66). Note that, at this point in the development, $(L^A)^{-1}$ is not known. However, by making the choice

$$q_1 = -L^A R_0 \tag{2.68}$$

and

$$q_{n+1} = -L^A R_n + \beta_n q_n \tag{2.69}$$

we can construct an algorithm that does not require the explicit use of either $(L^A)^{-1}$ or L^{-1}. Furthermore, it can be shown that

$$R_n = R_{n-1} + \alpha_n L q_n \tag{2.70}$$

and

$$\langle R_n, R_m \rangle = 0 \qquad n \neq m \tag{2.71}$$

$$\alpha_n = \frac{||R_{n-1}||^2}{||q_n||^2} \tag{2.72}$$

$$\beta_n = \frac{||R_n||^2}{||R_{n-1}||^2} \tag{2.73}$$

by carrying out a development which parallels that given for the CGM case in Section 2.4. We note that, strictly speaking, the minimization-in-the-domain algorithm is not a conjugate direction method, because the orthogonality between expansion functions is simple orthogonality, as expressed in Equation (2.66). Perhaps it is clearer to name this method an "orthogonal gradient" method (OGM).

The complete OGM algorithm is summarized below:

Initial Steps

Guess f_o

$$R_0 = Lf_o - h$$

$$q_1 = -L^A R_0$$

Iterate n = 1,2,...

$$\alpha_n = \frac{||R_{n-1}||^2}{||q_n||^2}$$

$$f_n = f_{n-1} + \alpha_n q_n$$

$$R_n = Lf_n - h = R_{n-1} + \alpha_n L q_n$$

$$\beta_n = \frac{||R_n||^2}{||R_{n-1}||^2}$$

$$q_{n+1} = -L^A R_n + \beta_n q_n$$

2.7. A Gradient Algorithm Using Approximate Inverse Operators (AIGM)

From the above discussion of the CGM and the OGM, it is clear that the algorithms generate a sequence of expansion functions and then find their coefficients to minimize some measure of the error. Furthermore, the CGM and OGM both generate expansion functions which are mutually orthogonal, and thus, theoretically terminate in a finite number of iterations. These algorithms are in some sense equivalent to generating an inverse operator to a prescribed degree of accuracy. In many applications, an approximate inverse operator is readily available, but not accurate enough to produce a meaningful solution

directly. In this section, we consider a gradient algorithm that incorporates an approximate inverse operator in order to systematically improve an estimate of the solution. The rate of convergence of this type of algorithm depends on the accuracy of the approximate inverse operator; unfortunately, there is no theoretical guarantee that convergence will occur in a finite number of iterations. This procedure is a generalization of an approach used by van den Berg [52].

If the standard form of the CGM is modified so that

$$p_1 = \widetilde{-L^{-1}} R_0 \tag{2.74}$$

$$p_{n+1} = \widetilde{-L^{-1}} R_n + \beta_n p_n \tag{2.75}$$

where $\widetilde{L^{-1}}$ is the approximate inverse operator, and if β_n is chosen so that

$$\langle Lp_{n+1}, Lp_n \rangle = 0 \tag{2.76}$$

the resulting algorithm provides a systematic way to generate the solution to Equation (2.1). The motivation for this choice stems from the fact that

$$f = f_{n-1} - L^{-1} R_{n-1}$$

which follows from Equation (2.10). Thus, if $\widetilde{L^{-1}}$ is the exact inverse, the process terminates after the generation of p_1 according to Equation (2.74). Note that the above choice of β_n does not ensure that the sequence $\{p_n\}$ is mutually $L^A L$ - orthogonal, as it did with the CGM. The orthogonality relationships that were described for the CGM do not apply to this algorithm, and a finite number of the p-functions will not, in general, span the solution space. Even if they did, their coefficients are not adjusted accordingly, but rather with the formula of Equation (2.9). We emphasize that the usefulness of this approach,

which we name the "approximate inverse gradient" method (AIGM), is directly proportional to the accuracy of $\widetilde{L^{-1}}$.

The complete AIGM algorithm is given as follows:

Initial Steps

Guess f_o

$$R_0 = Lf_o - h$$

$$P_1 = -\widetilde{L^{-1}}R_0$$

Iterate n = 1,2,...

$$\alpha_n = \frac{-\langle R_{n-1}, Lp_n \rangle}{||Lp_n||^2}$$

$$f_n = f_{n-1} + \alpha_n p_n$$

$$R_n = Lf_n - h = R_{n-1} + \alpha_n Lp_n$$

$$\beta_n = \frac{\langle L\widetilde{L^{-1}}R_n, Lp_n \rangle}{||Lp_n||^2}$$

$$p_{n+1} = -\widetilde{L^{-1}}R_n + \beta_n p_n$$

This AIGM is based on minimizing functional F_1 defined in Equation (2.5). Similar algorithms could be based on other functionals.

2.8. Summary

This chapter has presented several procedures that can be used for the iterative solution of matrix equations, with the intent of emphasizing the important aspects of their development. The CGM algorithm is the common form

of the conjugate gradient method for arbitrary matrix equations, and is sometimes known as the minimization-in-the-range method. The OGM is a different algorithm which is sometimes known as the minimization-in-the-domain version of the CGM. The AIGM is a procedure that is based on a minimization-in-the-range procedure, and allows the systematic use of an approximate inverse operator. An important feature of these three algorithms is that they will not diverge, as simple iterative methods such as the Jacobi and Gauss-Seidel algorithms [30] - [34] occasionally do when applied to problems of practical interest. A discussion of the performance of the CGM, OGM, and AIGM is presented in Chapter 3, along with some theoretical results concerning the convergence of the CGM.

3. CONVERGENCE OF ITERATIVE ALGORITHMS WHEN USED FOR ELECTROMAGNETIC SCATTERING APPLICATIONS

3.1. Introduction

The previous chapter presented several iterative algorithms but omitted a discussion of their convergence. This chapter provides an illustration of the convergence for several types of problems arising in electromagnetic scattering. A variety of examples, intending to represent typical performance, are presented. While the conclusions from this study are not necessarily applicable to every problem that may arise, they will serve as useful guidelines in the characterization of the iterative algorithms for many practical electromagnetic scattering problems. Furthermore, it appears that a familiarity with the performance of the iterative algorithms is useful in a different way: it is helpful in identifying situations where the numerical modeling is inadequate to produce a meaningful solution.

Chapter 2 presented three algorithms, the conjugate gradient method (CGM), the orthogonal gradient method (OGM) and the approximate-inverse gradient method (AIGM). Here, implementations of all three of the algorithms are considered, the majority being examples of the CGM. The rate of convergence of the OGM is very similar to that of the CGM, and in general, the performance of the AIGM depends on the particular approximate-inverse operator in use. In all cases discussed here, the approximate inverse is obtained by inverting the main diagonal of the matrix. This specific implementation of the AIGM appears to converge slower than the CGM for most of the examples presented herein, although it is a more general implementation of the AIGM than that used by van den Berg [52]. Van den Berg presented an example for which a more accurate approximation to the inverse was available, enabling the AIGM to consistently outperform the CGM [52].

Primarily, we are concerned with the solution of the matrix equation arising from a moment-method discretization of an integral (or integro-differential) equation, as discussed by Harrington [25] and Wilton and Butler [68]. The examples used for illustration are all equations representing electromagnetic scattering problems. The formulation used for each example is not discussed in detail here, but is available in later chapters or elsewhere in the literature. Furthermore, at this point we are not concerned with the implementation of these algorithms for the efficient treatment of large-order systems of equations. This topic is reserved for later chapters. Thus, most of the examples presented here are based upon matrix equations of relatively small order, i.e., below order 150.

The convergence behavior of the iterative algorithms will be illustrated via the residual norm

$$N_k = \frac{||Lf_k - h||}{||h||} \tag{3.1}$$

Experimentation with this quantity has shown it to be reliable provided that the matrix L is not badly conditioned. Normally, a criterion of

$$N_k < 10^{-4} \tag{3.2}$$

is adequate to ensure several digits of accuracy in the solution. The condition of the matrix L has a significant effect on the performance of the iterative algorithms, as will be discussed in Section 3.2.

All of the iterative algorithms require the user to provide an initial estimate or "guess" of the solution. In all cases examined here, a zero estimate is used. All of the algorithms considered here will converge in theory for arbitrary initial estimates, a feature not shared by many of the iterative

algorithms used in the past. In addition, the rate of convergence is not a strong function of the initial guess, as is indicated by the theoretical rates of convergence presented in Section 3.2.

A desirable feature of any algorithm for the solution of electromagnetic scattering problems is the ability to solve Equation (2.1) for many different right-hand sides. Specifically, it may be necessary to generate solutions corresponding to many different orientations of the scatterer and excitation in space, all of which can be described by the same matrix operator L. Some computational savings could arise by simultaneously generating several solutions with the same set of expansion functions (the p-functions discussed in Section 2.3). This approach is examined in Section 3.7, but it does not appear to yield the expected savings due to properties of the iterative algorithms in use.

3.2. Aspects of the Theoretical Convergence of the CGM

From the discussion of the CGM in Chapter 2, it is apparent that in the absence of any machine errors in the various calculations, the CGM will terminate in N steps or less for an order N system. In fact, the number of iterations required for the solution of Equation (2.1) is normally equal to the number of independent eigenvalues in the matrix $L^A L$ [34], [69]. Furthermore, if f_n is the estimate of the solution f produced by the CGM after n iterations, f_n must satisfy

$$||f_n - f|| \leq ||f_m - f|| , \quad n > m \tag{3.3}$$

Thus, the estimates produced by the CGM converge monotonically to the solution.

Equation (3.3) may be obtained by considering the general form of Equation (2.45)

$$p_n = -||L^A R_{n-1}||^2 \sum_{i=0}^{n-1} \frac{L^A R_i}{||L^A R_i||^2} \tag{3.4}$$

From Equation (3.4) and the orthogonality of the set $\{L^A R_i\}$ as expressed in Equation (2.35), we have

$$\langle p_i, p_j \rangle \geq 0 \tag{3.5}$$

From the definition of f_n, it follows that

$$f_n - f_m = \sum_{i=m+1}^{n} \alpha_i p_i \qquad n > m \tag{3.6}$$

Note that the $\{\alpha_i\}$ are nonnegative by Equation (2.37). Equations (3.5) and (3.6) can be combined to yield the inequality

$$\langle f_n - f_m, f_N - f_n \rangle \geq 0 \qquad N > n > m \tag{3.7}$$

Finally, Equation (3.7) can be combined with the identity

$$||f_m - f||^2 = ||f_m - f_n||^2 + ||f_n - f||^2 + 2\mathrm{Re}\{\langle f_n - f_m, f - f_n \rangle\} \tag{3.8}$$

to prove Equation (3.3). It is interesting to note that both the CGM and the OGM algorithms produce solution estimates which satisfy Equation (3.3); the AIGM algorithm does not. Of course, the presence of round-off errors may invalidate Equation (3.3) to some extent.

Because round-off errors degrade the finite-step termination property of the CGM, in practice it is often considered a purely iterative algorithm. In this context, an upper bound on its rate of convergence is given as [67], [69]

$$\frac{||f_n - f||}{||f_0 - f||} \leq 2\left[\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1}\right]^n \tag{3.9}$$

where κ is the condition number of $L^A L$ (the ratio of the maximum to minimum eigenvalues of the matrix $L^A L$). In addition, it has been observed that approximately $n\sqrt{\kappa}$ iterations are normally required to ensure accuracy to n decimal places [69]. Data to follow support the conclusion that the dependence of convergence rate on the number of decimal places is approximately linear.

These approximate rates of convergence are independent of the right-hand side of the system and the initial estimate of the unknown, traits which were usually observed in the numerical examples examined throughout this study. In an effort to reduce the amount of iterative computation as much as possible, appreciable effort is often expended to produce a "good" initial estimate of the solution. For instance, asymptotic solutions to electromagnetic scattering problems are often used as initial estimates in iterative solutions [44]. Based upon the above observations, an initial estimate will probably not affect the rate of convergence, only the distance from the starting point to the actual solution. Since the rate of convergence is roughly linear with the required number of decimal places of accuracy, an initial estimate can be considered "good" only if it reduces to the initial value of the residual norm by a significant amount below unity. (Unity is the value of N_0 produced by an initial guess of zero.) As an example, suppose that it is desired to produce a solution with residual norm $N = 10^{-4}$. An initial estimate of the solution which produces $N_0 = 10^{-2}$ should reduce the amount of iterative computation to half of what would have been required if an initial estimate of zero had been used. Similarly, an initial estimate with $N_0 = 10^{-1}$ should reduce the required iterative computation by one quarter, and so on.

Whether considered a finite-step method or a purely iterative method, the rate of convergence of the CGM depends on the eigenvalues of the matrix. The NxN matrix produced by a moment-method discretization of an integral equation can be considered an approximation to the integral operator. Properties of the integral operator, such as its eigenvalue spectrum, are projected onto the matrix. In general, both the accuracy of the solution to the system and the accuracy of the eigenvalue representation will be linked to the order of the discretization, and should improve as the discretization is refined.

Suppose that a given integral equation is discretized via the moment method to yield an N-th order matrix equation, which is then solved by the CGM. Let us also suppose for the moment that the discrete system we obtain will "converge" in some sense to the integral equation as the order of the discretization is increased. If a solution to the N-th order system is found in n steps, where n is much smaller than the order of the system, the theoretical property of the CGM suggests that there are n dominant eigenvalues in the spectrum of the associated integral operator. A further check on this result can be obtained by rediscretizing the integral equation to yield a larger-order matrix equation, for which the CGM should again converge in n steps. This behavior is typical for problems where the matrix is able to represent the eigenvalue structure of the integral operator, as illustrated in the examples to follow.

Suppose, on the other hand, convergence is slow in that n≈N. If the CGM is then applied to a larger system arising from a refinement of the discretization, and many additional iterations are necessary for convergence, it is a clear indication that the original matrix is not able to represent the eigenvalue structure of the integral operator. Under such conditions the particular matrix equation is probably not an accurate model of the integral equation, in that the level of discretization is inadequate to represent the problem. Alternatively,

it is possible that the formulation on which the matrix equation is based may be invalid, which could cause a significant change in the typical behavior of the iterative algorithm. An example of the latter situation arises when an integral equation exhibits non-unique solutions, such as occur in connection with the "internal resonance" problems discussed briefly in Section 3.3.

Although the preceding remarks are somewhat speculative, numerical results support the notion that the CGM provides feedback which can be useful for identifying situations where the accuracy of the numerical modeling process appears questionable. This behavior is illustrated throughout the remainder of this chapter. However, not every example will conform to the same type of behavior. User-familiarity with the typical behavior of the algorithm for the particular type of problem involved is essential before unusual behavior can be identified; even then the above ideas are not absolutes and the behavior may be misleading. If the operator is ill-conditioned, severe round-off errors will disrupt the finite step termination property of the CGM and mask the convergence behavior. Normally, under such conditions the slow convergence of the algorithm is a clear indication that the system is poorly conditioned, which itself may be useful information.

Round-off errors occur to some degree whenever machine computations are performed, and thus mandate the choice of algorithms which minimize their effects whenever possible. Round-off errors affect the CGM, which generates expansion functions recursively to satisfy the orthogonality of Equation (2.21), by degrading the actual orthogonality of these functions. If the round-off errors are severe, the functions may not even approximately span the N-dimensional solution space. Due to this and the additional fact that the coefficients of the expansion functions are also incorrect due to round-off errors,

the algorithm may not converge in practice. Simon illustrates the loss of orthogonality in recursively generated functions for a similar algorithm [70]. It has been noted that the stability of the CGM to round-off errors can be very low [64].

Although it is possible to examine the theoretical stability of numerical algorithms to round-off errors, it is much easier and perhaps more meaningful to experiment with different procedures and observe their performance in practice. Round-off errors did not appear to present a problem for the examples taken from electromagnetic scattering problems illustrated in the remainder of this chapter. These examples were generated with the aid of a CDC Cyber-175 computer, which uses a single-precision word length of 60 bits. However, when poorly conditioned systems were constructed for the sole purpose of testing the CGM, the algorithm was not able to solve them to any reasonable degree of accuracy. This behavior had prompted past research into CGM-like algorithms with better numerical stability, and several of these are discussed by Stoer [69]. A different avenue of approach has been the use of preconditioning techniques, which essentially attempt to convert the system into a better conditioned matrix equation with the same solution [71] - [73]. Still another technique, the reorthogonalization of the p-functions [70], generally requires the storage of the set $\{p_i\}$ and thus is not immediately compatible with the goals of this study.

3.3. TM Scattering from Perfectly Conducting Cylinders

One of the simplest moment method formulations is the scattering of TM plane waves from an infinite, perfectly conducting cylinder. The formulation is general and can treat parallel cylinders of arbitrary shape, including thin strips. The matrix equation arising from a moment method discretization of the electric-field integral equation (EFIE) has been discussed by Harrington [74]. An implementation of the CGM for the solution of the matrix equation and a

discussion of the accuracy of the numerical solution are described in a report by Peterson and Mittra [57], and briefly summarized in Chapter 4.

The following figures illustrate the rate of convergence for several examples of the above moment method formulation. Since in theory the CGM and OGM terminate in N steps or less for an order-N system, it is convenient to normalize the number of iterations to N. This approach facilitates the observation of trends which might otherwise be obscured by examples of different order. Thus, the figures show plots of the residual norm N_n as defined in Equation (3.1) versus the normalized iteration step n/N.

In all cases, we arbitrarily consider the system "solved" once a solution is produced which satisfies $N_n < 0.0001$. Note that in practice, if the matrix involved is ill-conditioned, the residual norm is not always a valid indication of an accurate solution. Under these conditions, it may be advisable to terminate the algorithm at a lower value of N_n.

Various parameters describing the data presented in these figures have been controlled in order to investigate any general effects they may have on relative rates of convergence. These parameters include the cell density per wavelength used within the moment method discretization, the presence or absence of simple symmetries in the equation and solution, and the degree to which different sized cells (or in later sections, mixtures of different permittivities representing inhomogeneous dielectric materials) are incorporated into the model. Furthermore, we consider examples of the CGM, OGM, and AIGM. The specific form of the AIGM in use has been discussed in Section 3.1.

Figure 3.1 shows the convergence of the CGM for three examples, all with the cell density fixed at 10.0 cells/λ_o. Two of these systems possessed a degree of symmetry, resulting from symmetry planes in both the scatterer geometry and in the excitation being preserved by the discretization. Often, the CGM is

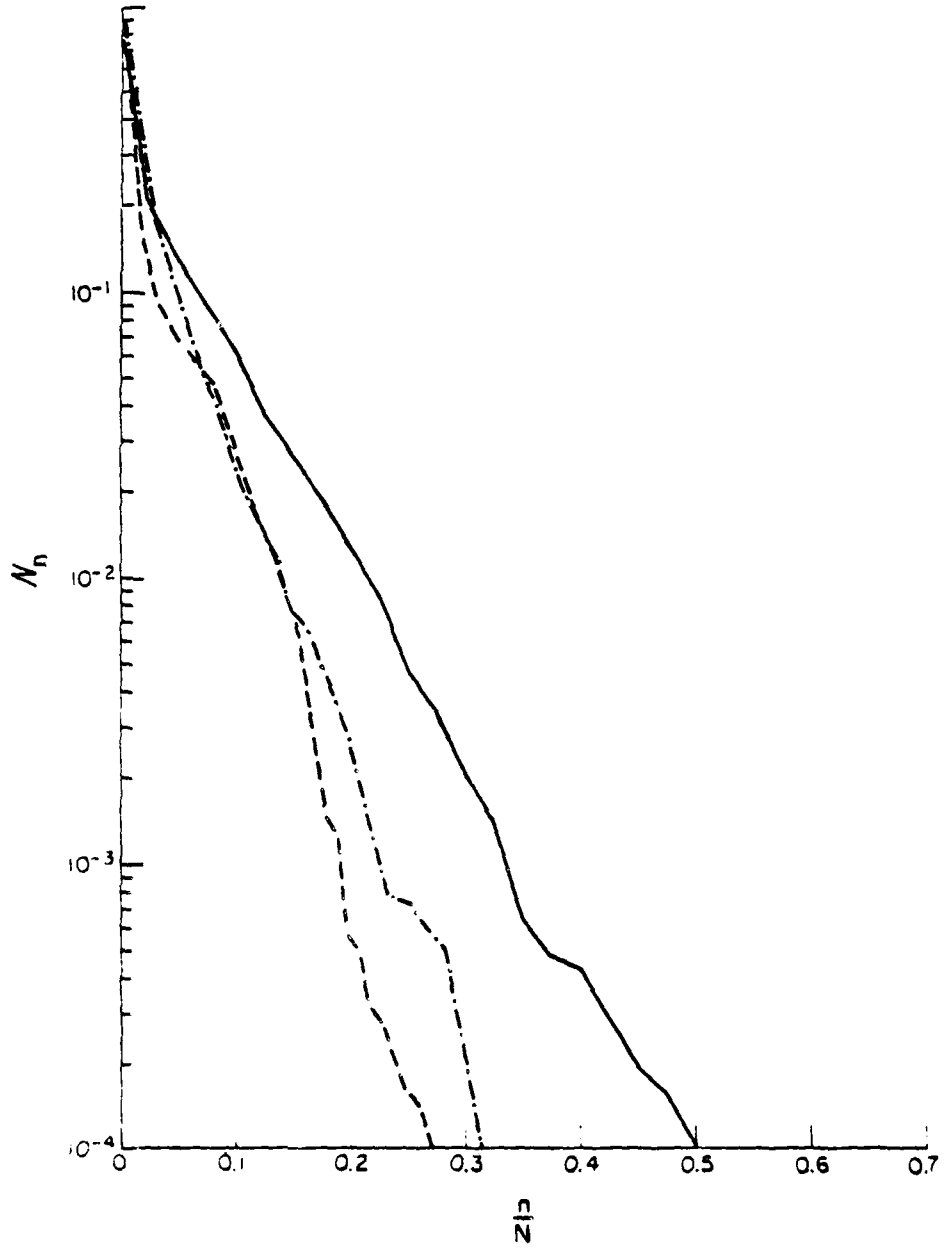

Figure 3.1 Convergence of the CGM for examples of PEC cylinders, TM polarization.
---- N = 100, 10.0 cells/λ_0, symmetry, pie-shaped cylinder
-·-· N = 60, 10.0 cells/λ_0, symmetry, 2 coplanar strips
—— N = 40, 10.0 cells/λ_0, no symmetry, 2 coplanar strips

observed to converge at a faster rate for systems with this type of symmetry as compared to similar equations with no symmetry. An examination of the expansion functions produced by the CGM for a symmetric example showed that the p-functions contained the same type of symmetry as the solution. This may explain the quicker convergence, since only half of the normal set of expansion functions is needed to span the symmetric part of the solution space.

Figure 3.2 shows the convergence of the CGM for three examples of moment method systems created with different cell densities. These equations all posessed a degree of symmetry. The data illustrate a general trend which has been observed in a wide variety of examples, namely, that the rate of convergence is dependent upon the sampling density, and is usually faster for problems with a higher density of cells per free-space wavelength.

The interdependence of convergence rate and sampling density is related to the discussion in Section 3.2 concerning the eigenvalues of the integral operator and their projections into the matrix operator. The concept is better illustrated by Table 3.1, which shows values of the residual norm produced by the CGM for a system representing a circular cylinder illuminated by a plane wave. Four different levels of discretization are shown. Note that while only three iterations are required to solve the 4x4 system, 5 iterations are required to solve the 8x8, 16x16, and 32x32 systems. This process was carried out for 64x64 and 128x128 matrix representations as well (not shown) for which the CGM also required 5 iterations to reduce the residual norm below 0.0001. The implications here are that the corresponding integral operator has five independent eigenvalues and that the matrix is "converging" toward the integral as the order of the system increases. (This particular geometry yields an exact solution, and the numerical solution appears to converge to the exact result as the order of the system increased. The comparison is presented in Section 4.3.) Note

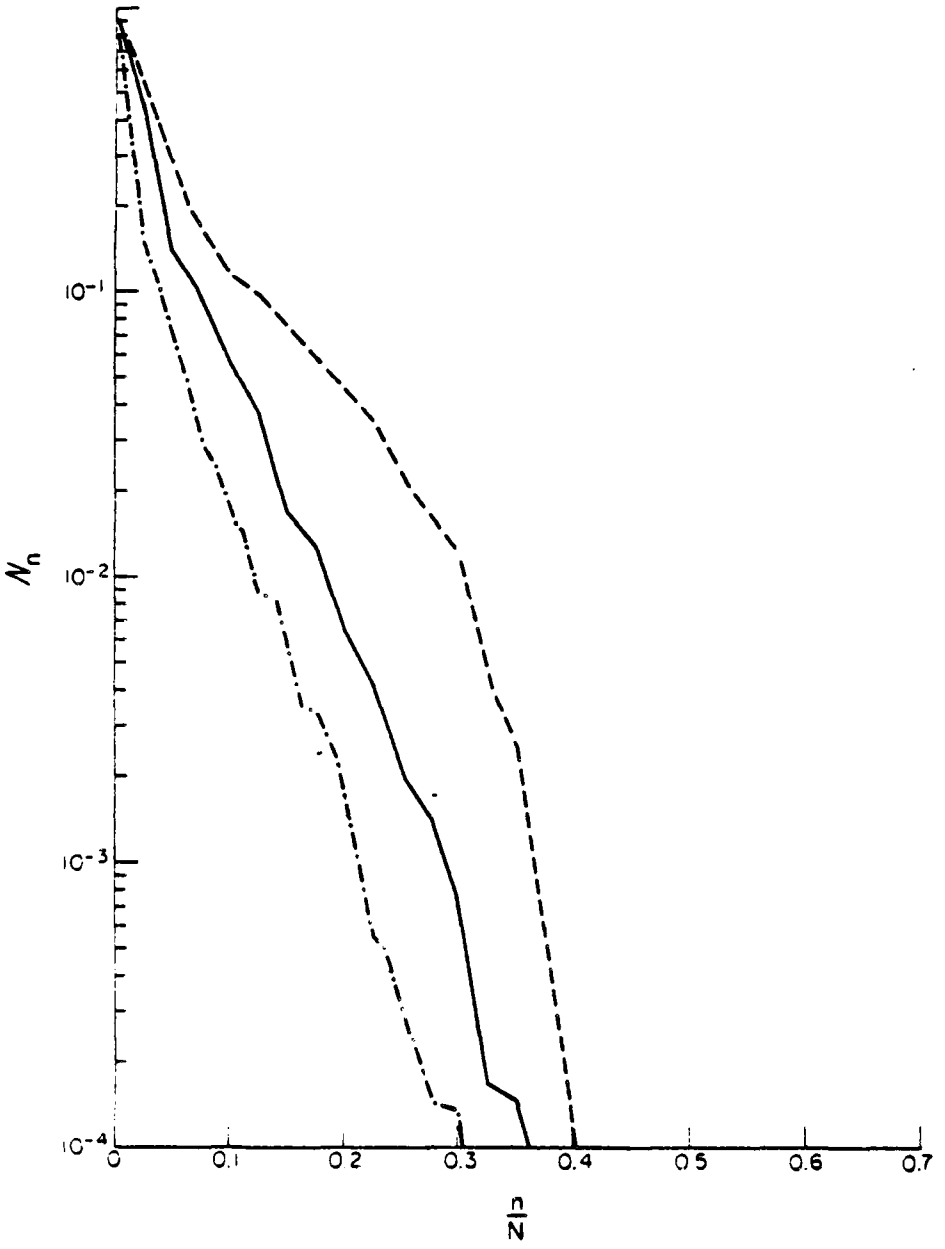

Figure 3.2. Convergence of the CGM for examples of PEC cylinders, TM polarization.
---- N = 40, 3.2 cells/λ_0, symmetry, circular cylinder
—— N = 40, 10.0 cells/λ_0, symmetry, 2 coplanar strips
-·-· N = 80, 20.0 cells/λ_0, symmetry, 2 coplanar strips

TABLE 3.1

VALUES OF THE RESIDUAL NORM PRODUCED BY THE CGM VERSUS ITERATION STEP FOR FOUR DIFFERENT DISCRETIZATIONS OF THE SAME INTEGRAL EQUATION

order of system	N = 4	8	16	32
cell density	$4.0 \frac{\text{cells}}{\lambda_0}$	8.0	16.0	32.0
n = 0	$N_n = 1.0$	1.0	1.0	1.0
1	0.359	0.366	0.361	0.358
2	0.100	0.114	0.115	0.115
3	8.9×10^{-10}	0.0142	0.0161	0.0161
4	-	0.000706	0.00128	0.00132
5	-	2.2×10^{-7}	6.9×10^{-5}	8.0×10^{-5}

also that the value of the residual norm at each step appears to converge as the cell density increases, possibly to the value of the residual norm which would be produced by an application of the CGM to the integral equation itself (if such an application could be carried out analytically, which has not been done to date). In practice, a cell density of 10.0 cells/λ_o is usually considered necessary to produce reasonably accurate numerical solutions to these integral equations. According to Table 3.1, the numerical values of the residual norm stabilize at about this cell density. Additional examples of the type shown in Table 3.1 are presented in later sections.

Figure 3.3 shows the convergence rates of 2 examples constructed with mixed cell densities in the "accurate" range. The particular models in use involved large and small cells placed immediately adjacent to each other, and contained no simple symmetry planes. Note that the rate of convergence in these examples is slower than in the previous examples. Jennings [30] suggests that the convergence of the CGM can be improved by scaling so that the main diagonal of the matrix has identical elements. No attempt was made here to incorporate scaling, which may explain the relatively slower convergence for the mixed cell cases.

Figure 3.4 shows the convergence behavior for 2 other symmetric systems, one corresponding to a circular cylinder and the other a pie-shaped cylinder. Both of these scatterers are also internally resonant cavities, such that the integral equation describing the external scattering problem fails because of the nontrivial homogeneous solution due to the internal cavity problem. In other words, there is no unique solution to the EFIE for the specific geometries under consideration. Although the CGM was able to accurately solve the matrix equations in these cases, (as was verified by independent computations and not merely the value of the residual norm), the numerical solution was not the desired result for the external scattering problem. The "internal resonance"

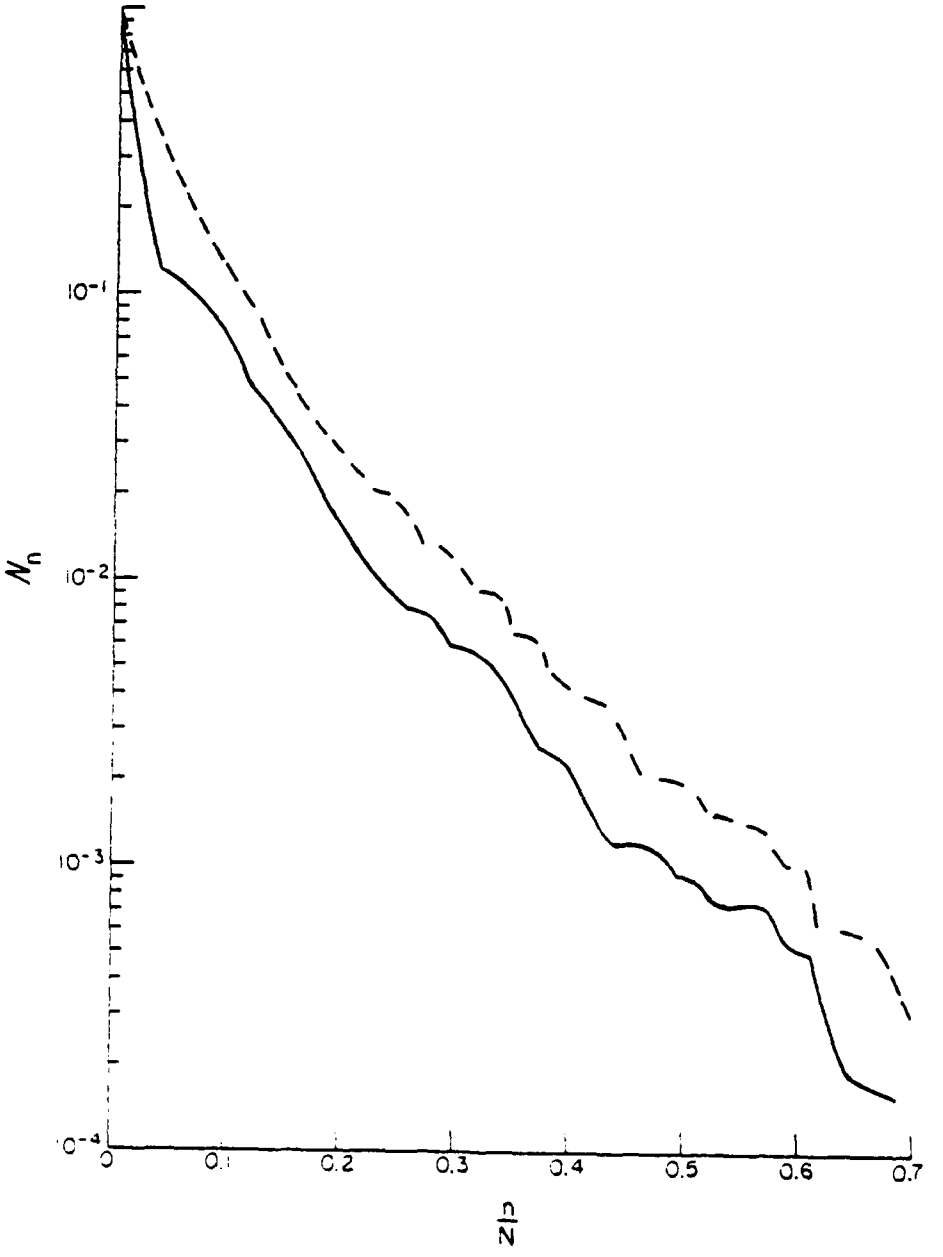

Figure 3.3. Convergence of the CGM for examples of PEC cylinders, TM polarization.
---- N = 63, 10.0-50.0 cells/λ_0, no symmetry, 1 bent strip, 1 flat strip, mixed cell sizes on each strip
—— N = 51, 10.0-50.0 cells/λ_0, no symmetry, 2 coplanar strips

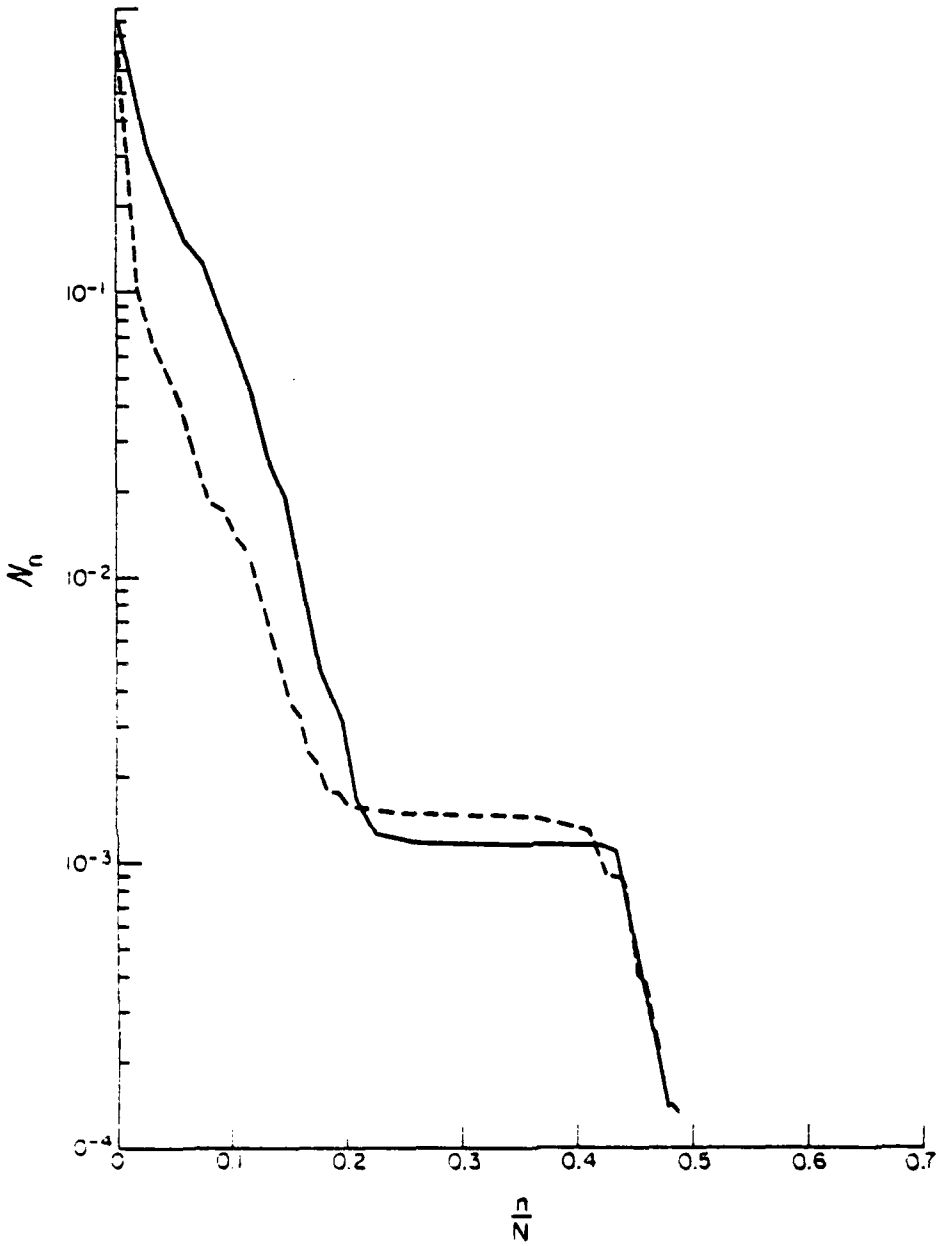

Figure 3.4. Convergence of the CGM for examples of PEC cylinders, TM polarization, where geometries are "internally resonant."
—— N = 67, 3.4 cells/λ_0, symmetry, circular cylinder
---- N = 120, 10.0 cells/λ_0, symmetry, pie-shaped cylinder

situation and various remedies have been discussed elsewhere [75] - [78]. In practice, the characteristic leveling of the curves shown in Figure 3.4 indicates an ill-conditioning problem, and this behavior may be useful in some instances as a flag to identify problem situations. In spite of the ill-conditioning, round-off errors do not appear problematic in these examples, which did converge at relatively quick rates. The symmetry in these systems may be responsible for the quick convergence.

Figures 3.5 and 3.6 show the performance of the OGM for seve-al examples. As expected, the residual norm does not always decrease monotonically with the OGM, and rates of convergence are similar to the CGM. In fact, the CGM and OGM usually require the same number of iterations to solve a given system, which is to be expected because of the "finite-step convergence in a number of steps equal to the number of eigenvalues" property shared by both of these algorithms. Figure 3.7 shows the residual norm produced by the CGM and OGM for a given example involving highly mixed cell sizes and no symmetry planes. Both algorithms converge at similar rates.

Figure 3.8 shows convergence data obtained with the AIGM. As implemented here, the AIGM is observed to converge at a slower rate than the CGM and OGM algorithms. Figure 3.9 shows a comparison of the CGM and AIGM for the same system. Figure 3.10 compares all three algorithms for an ill-conditioned system corresponding to an "internally resonant" circular cylinder. The convergence of the AIGM in this case is very slow. The behavior of the OGM illustrated in Figure 3.10 is interesting because the residual norm decreases, then increases to a value exceeding its initial value. Although this behavior is observed for a special type of example, in particular one where the condition number of the matrix $L^{A}L$ is in excess of 300000, it suggests that it may be misleading to terminate the OGM based upon the value of an indicator which does not decrease

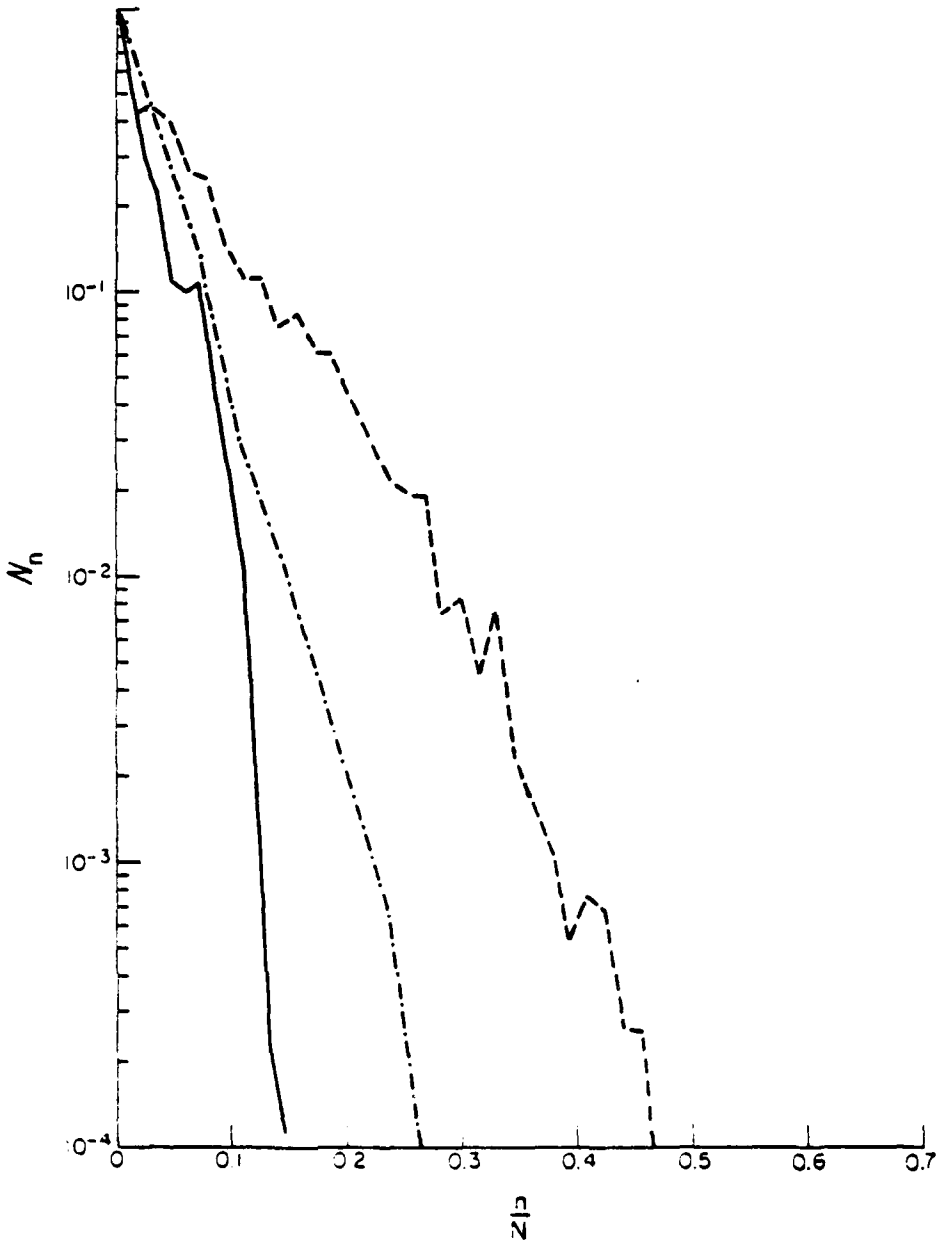

Figure 3.5. Convergence of the OGM for examples of PEC cylinders, TM polarization.

—— N = 85, 10.4 cells/λ_0, symmetry, circular cylinder

-·-· N = 30, 9.5 cells/λ_0, symmetry, circular cylinder

---- N = 64, 10.0 cells/λ_0, no symmetry, semi-circular cylinder with 2 coplanar strips

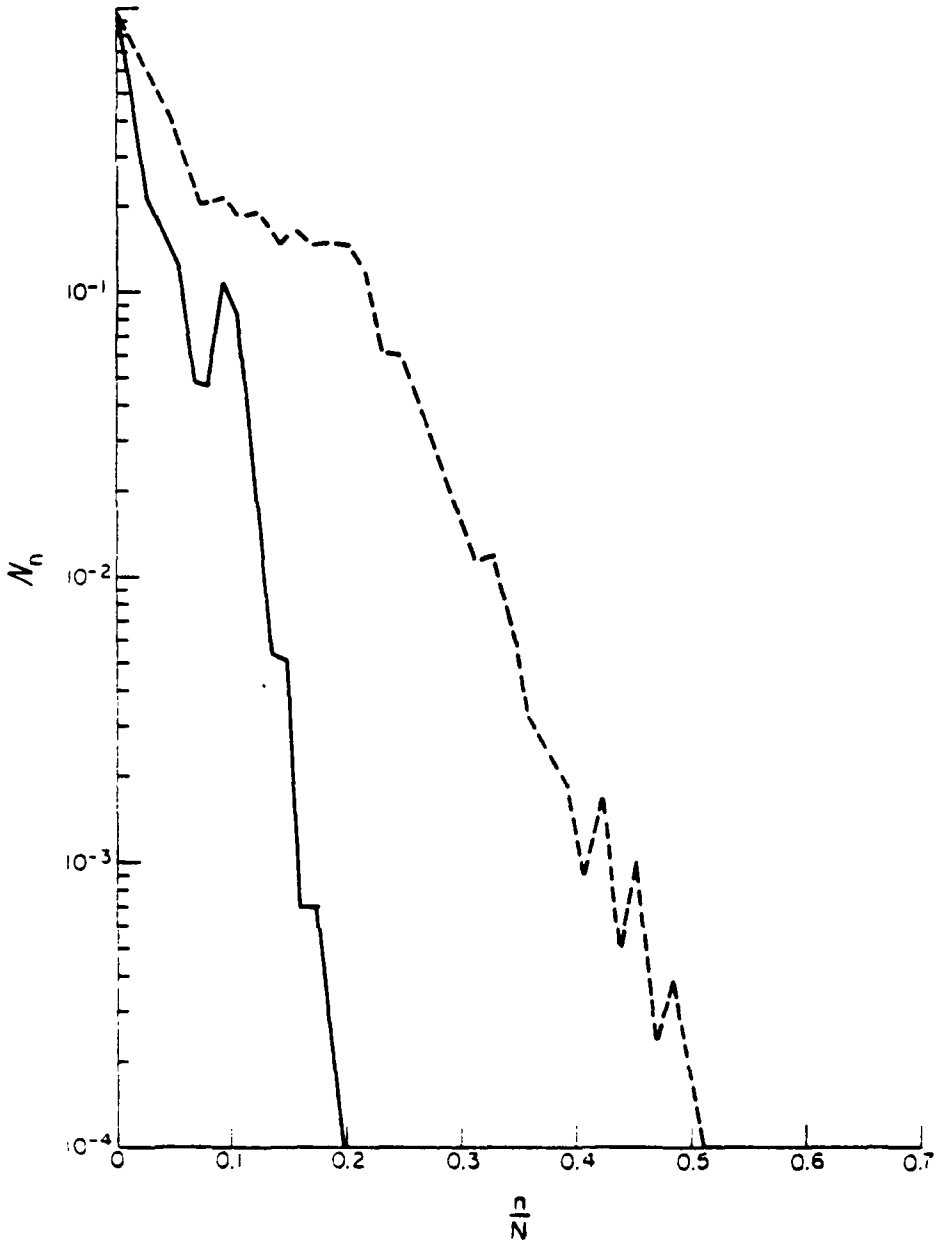

Figure 3.6. Convergence of the OGM for examples of PEC cylinders, TM polarization.
——— N = 75, 10.0 cells/λ_0, symmetry, circular cylinder
---- N = 64, 10.0 cells/λ_0, no symmetry, semi-circular cylinder with 2 coplanar strips

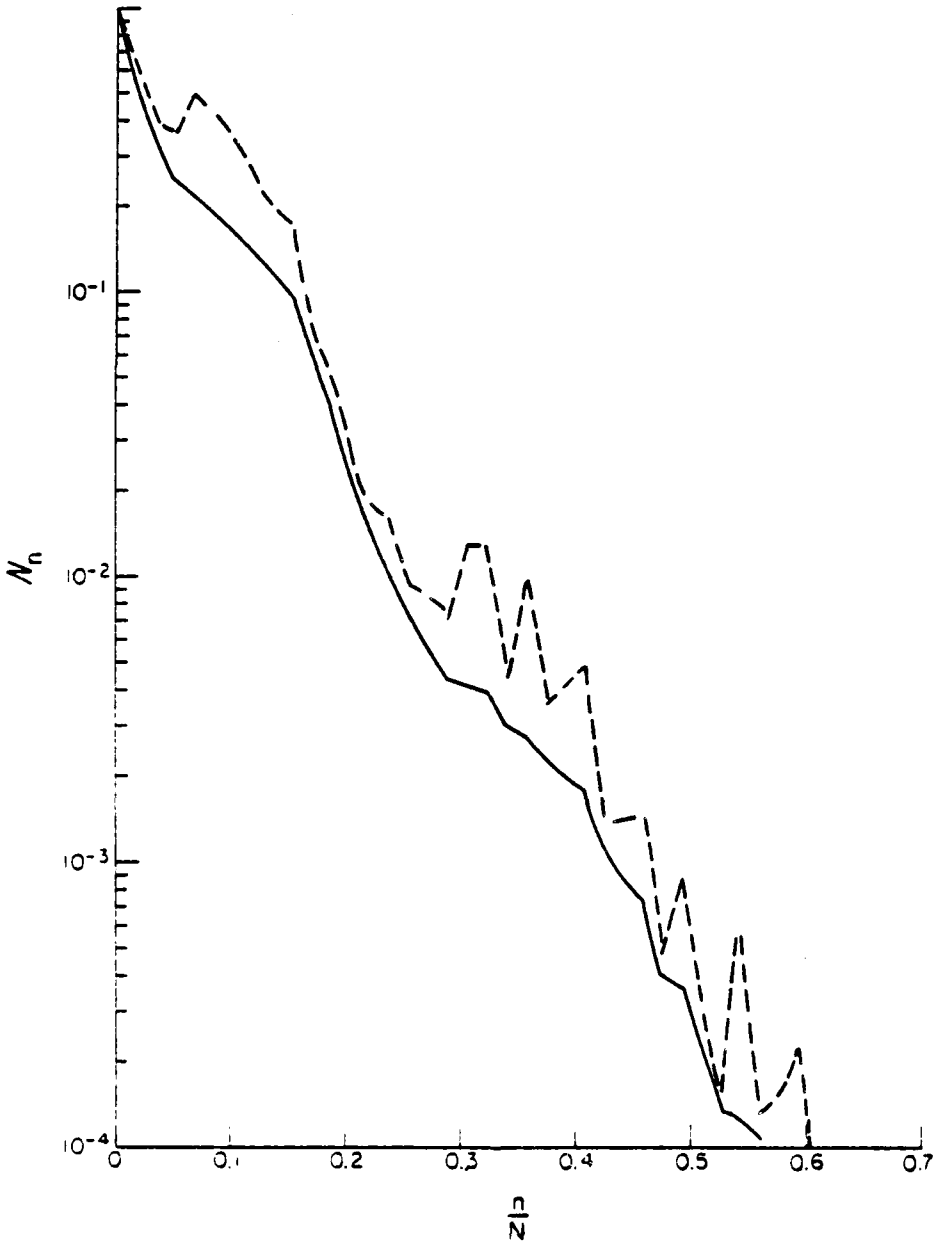

Figure 3.7. Convergence of CGM and OGM for a configuration of curved and flat PEC strips.
N = 59, 10.0-20.0 cells/λ_0, no symmetry, highly mixed cells in model
—— CGM
---- OGM

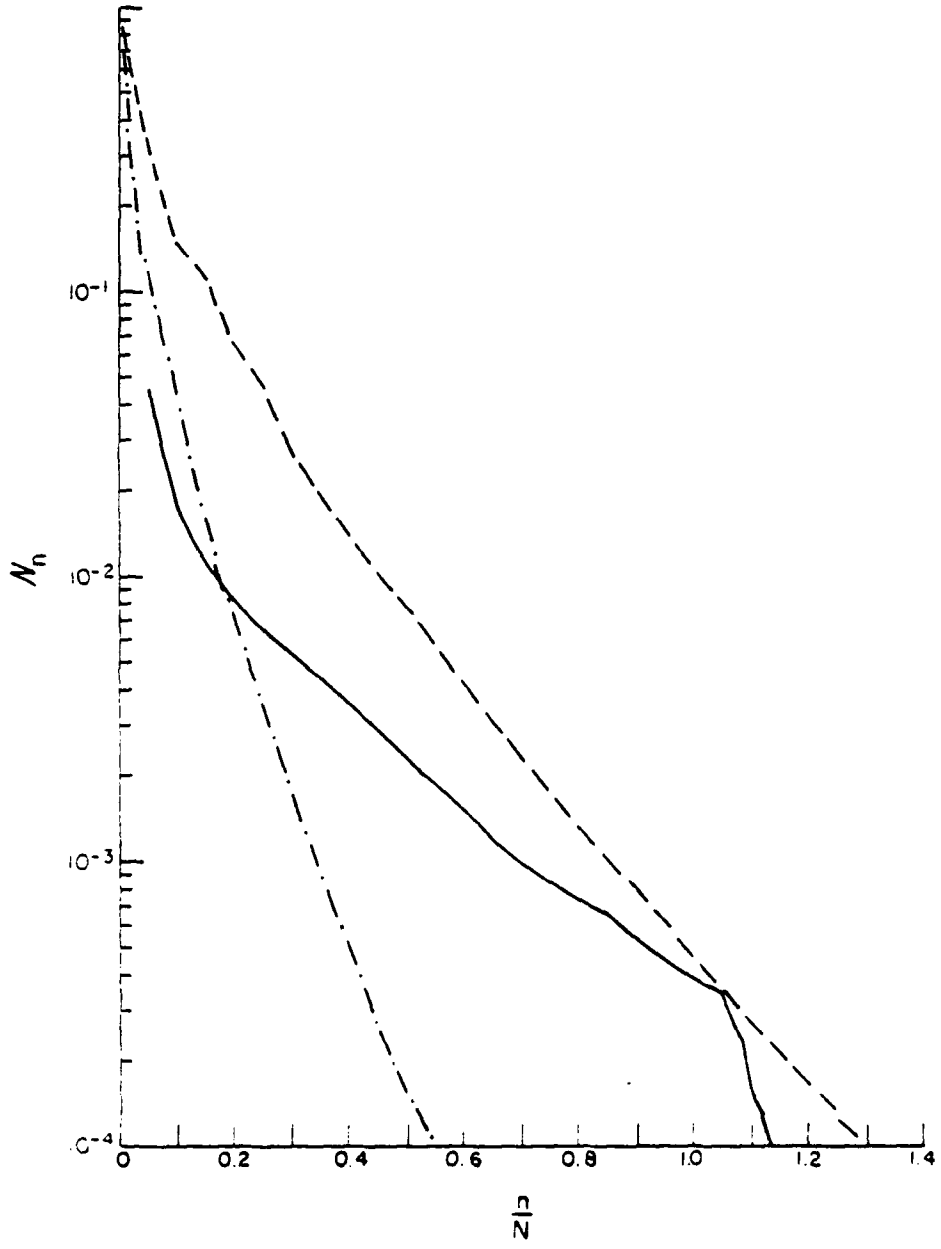

Figure 3.8. Convergence of the AIGM for examples of PEC cylinders, TM polarization.
-·-· N = 60, 10.0 cells/λ_0, no symmetry, 2 coplanar strips
—— N = 60, 100.0 cells/λ_0, symmetry, 2 coplanar strips
---- N = 20, 10.0 cells/λ_0, no symmetry, 2 coplanar strips

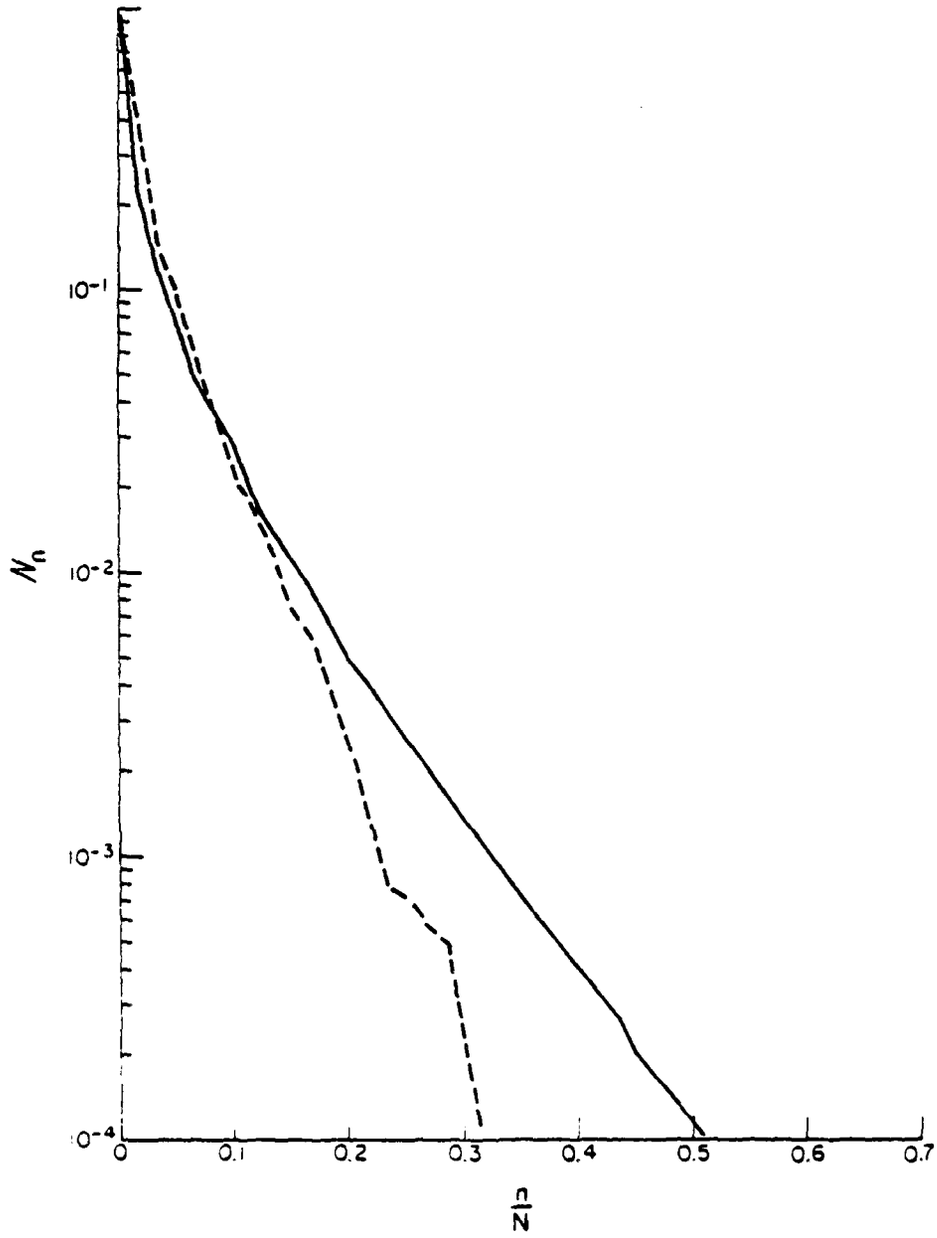

Figure 3.9. Comparison of convergence of the CGM and AIGM for a system with N = 60, 10.0 cells/λ_0, symmetry, 2 coplanar strips.
---- CGM
—— AIGM

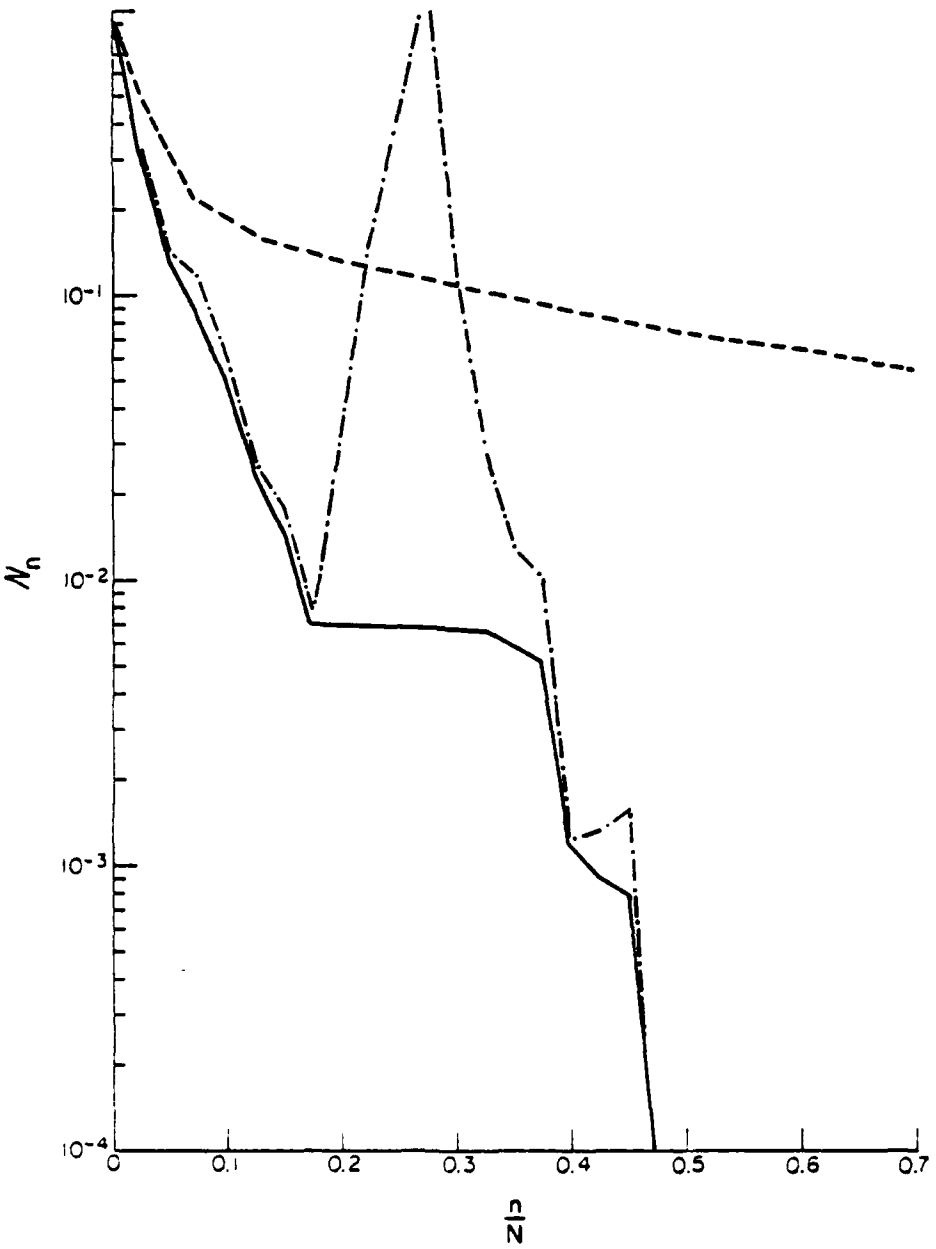

Figure 3.10. Comparison of the CGM, OGM, and AIGM for an example of a PEC cylinder, TM polarization, where geometry is "internally resonant" circular cylinder, N = 40, 7.8 cells/λ_0, symmetry.

——— CGM

-•-•- OGM

---- AIGM

monotonically. It also raises some questions about the validity of the residual norm itself as an indicator.

3.4. TM Scattering from Dielectric Cylinders

A scheme for the analysis of arbitrary, lossy or lossless dielectric cylinders illuminated with plane TM waves was developed by Richmond [79]. Here, we use the Richmond formulation but solve the resulting moment method matrix equation with the iterative algorithms of Chapter 2.

Figure 3.11 shows the rates of convergence of the CGM for three systems representing dielectric cylinders discretized with different cell densities. As was observed for previous examples, a general trend indicating faster convergence when smaller cell sizes are used is indicated. For cylinder examples (infinite cylinders described by their cross-sections) involving dielectric material, cell densities refer to the number of cells per unit cross-sectional area, where area is measured in terms of the wavelength in the dielectric material. If λ_o is the free-space wavelength, λ_d is the wavelength in the dielectric material defined by

$$\lambda_o = \sqrt{|\varepsilon_r|}\ \lambda_d \tag{3.10}$$

where ε_r is the complex relative permittivity. Note that the cell density of 9.6 cells/λ_d^2 used in one of the examples is normally considered far below what is required for the moment-method discretization to accurately model the scattering problem.

Figure 3.12 shows data for 2 examples where mixed permittivities were used in the model. This did not appear to slow the rate of convergence over cases where a constant permittivity was used throughout the model.

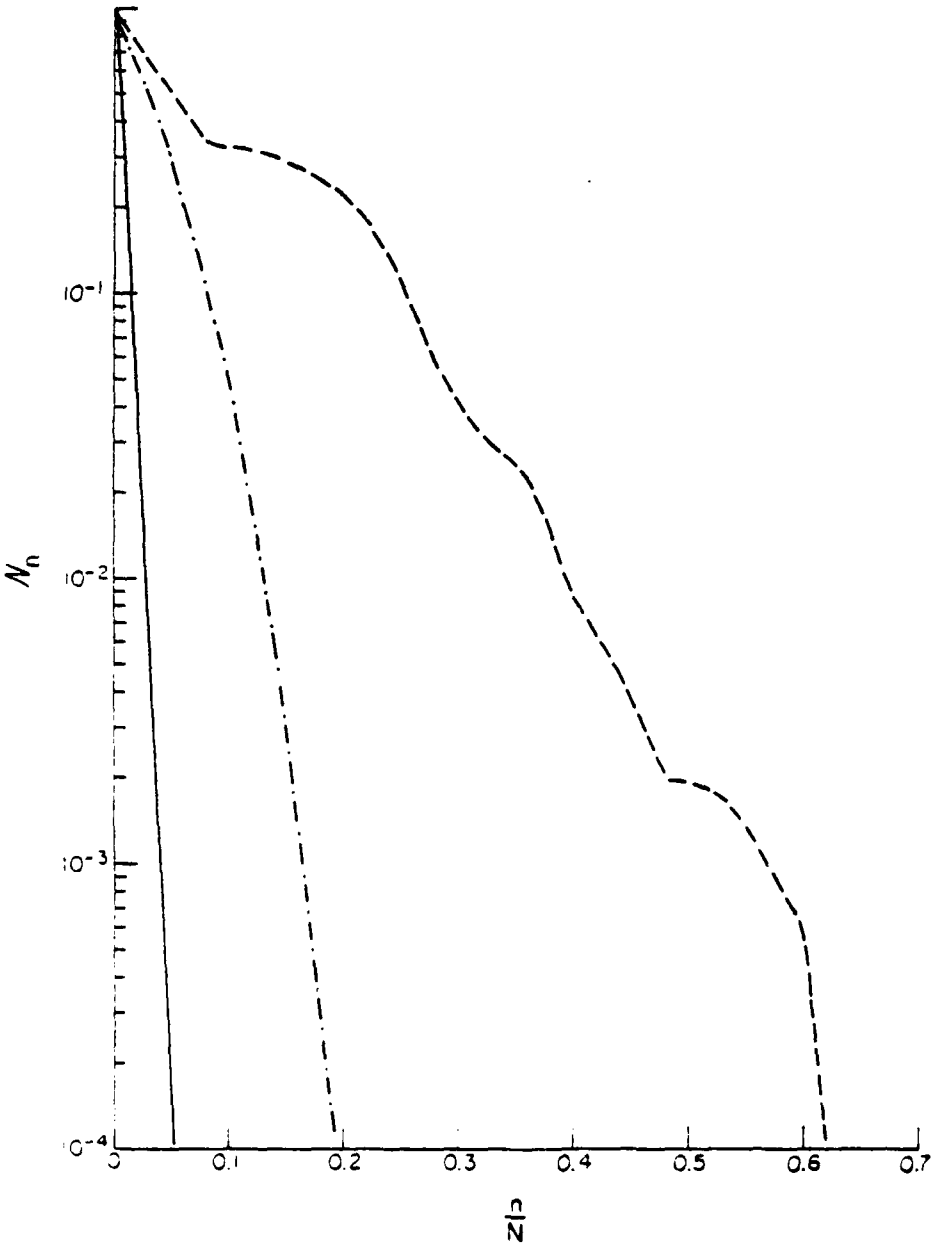

Figure 3.11. Convergence of the CGM for examples of dielectric cylinders, TM polarization.
—— N = 63, 835 cells/λ_d^2, homogeneous, rectangular cylinder, ε_r = 3
-·-· N^r = 21, 104 cells/λ_d^2, homogeneous, circular cylinder, ε_r = 2.56
---- N = 25, 9.6 cells/λ_d^2, homogeneous, square cylinder, ε_r = 2.56

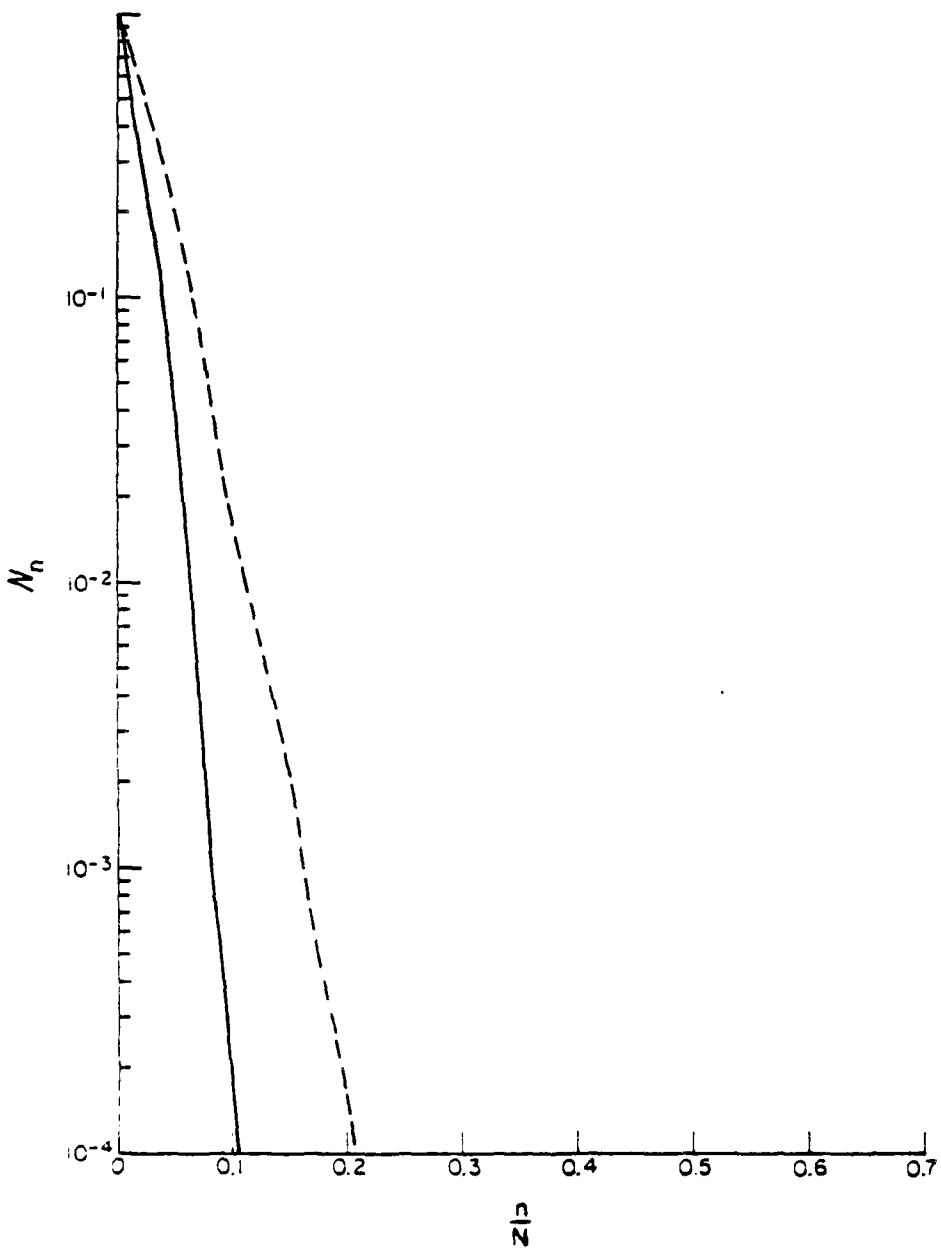

Figure 3.12. Convergence of the CGM for examples of dielectric cylinders, TM polarization.
—— N = 74, 100-502 cells/λ_d^2, inhomogeneous rectangular slab with holes, ε_r = 4-j3 and 3-j25
---- N = 41, 100-795 cells/λ_d^2, rectangular cylinder with holes, ε_r = 4 - j0

Figures 3.13 and 3.14 show the performance of the OGM, CGM, and AIGM applied to geometries modeled with mixed cell sizes and permittivities. Again, previous remarks concerning the algorithms apply.

Table 3.2 shows the residual norm versus the unnormalized iteration step for three different models of a homogeneous, circular dielectric cylinder, illuminated by a plane TM wave. The three models consist of 21, 61, and 101 square cells configured to approximate the circular cross-section of the cylinder. As the number of cells in the model increases, the order of the system increases as well, but in each case only 5 iterations are required for CGM solution. As discussed in Section 3.3, the matrix representation appears to be "converging" to something as the discretization is refined. This example is studied in Chapter 4, where it is shown that the solution to the discrete system appears to converge to the analytical solution for the circular cylinder.

3.5 TE-wave Scattering from Dielectric Cylinders

A computational method for TE-wave scattering from arbitrary dielectric cylinders was formulated by Richmond [80]. This approach uses the EFIE with pulse basis functions and point matching, and an iterative implementation of this scheme is developed in Chapter 6. In this section, the CGM and AIGM are applied to solve the moment method system, and their rates of convergence are presented via graphs showing the residual norm versus the normalized iteration step. In general, these rates of convergence depend upon the cell density of the particular model as did the previous examples of this chapter. The cell density is defined in Section 3.4. Note that the convergence rates of the AIGM are slightly faster than the CGM for these equations. As throughout the rest of the chapter, the AIGM is implemented using an approximate inverse obtained by inverting the main diagonal of the matrix.

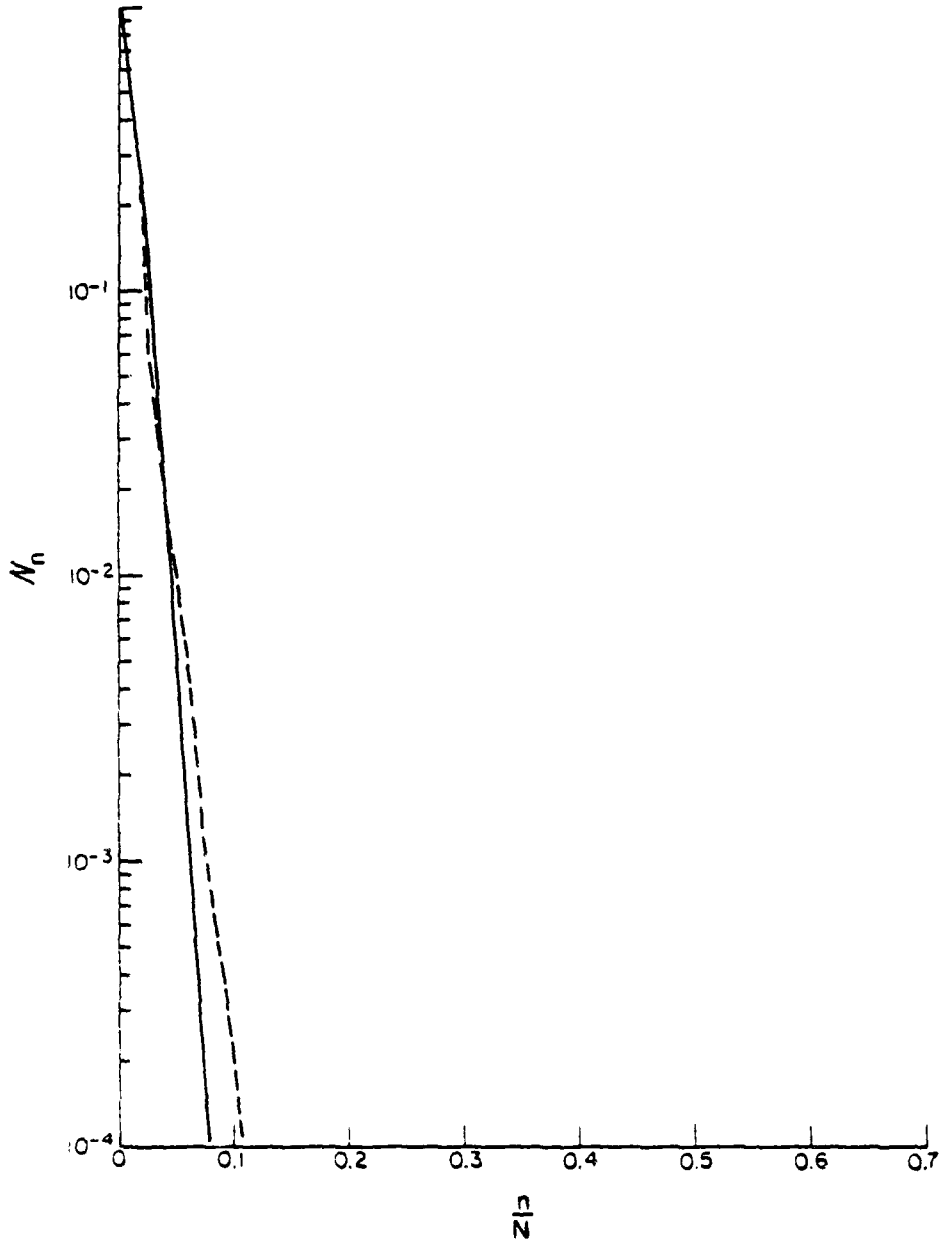

Figure 3.13. Convergence of the CGM, OGM, and AIGM for examples of dielectric cylinders, TM polarization.
N = 78, 225-645 cells/λ_d^2, inhomogeneous cylinder with ε_r ranging from 2-j1 to 4.5-j2
—— CGM and OGM
---- AIGM

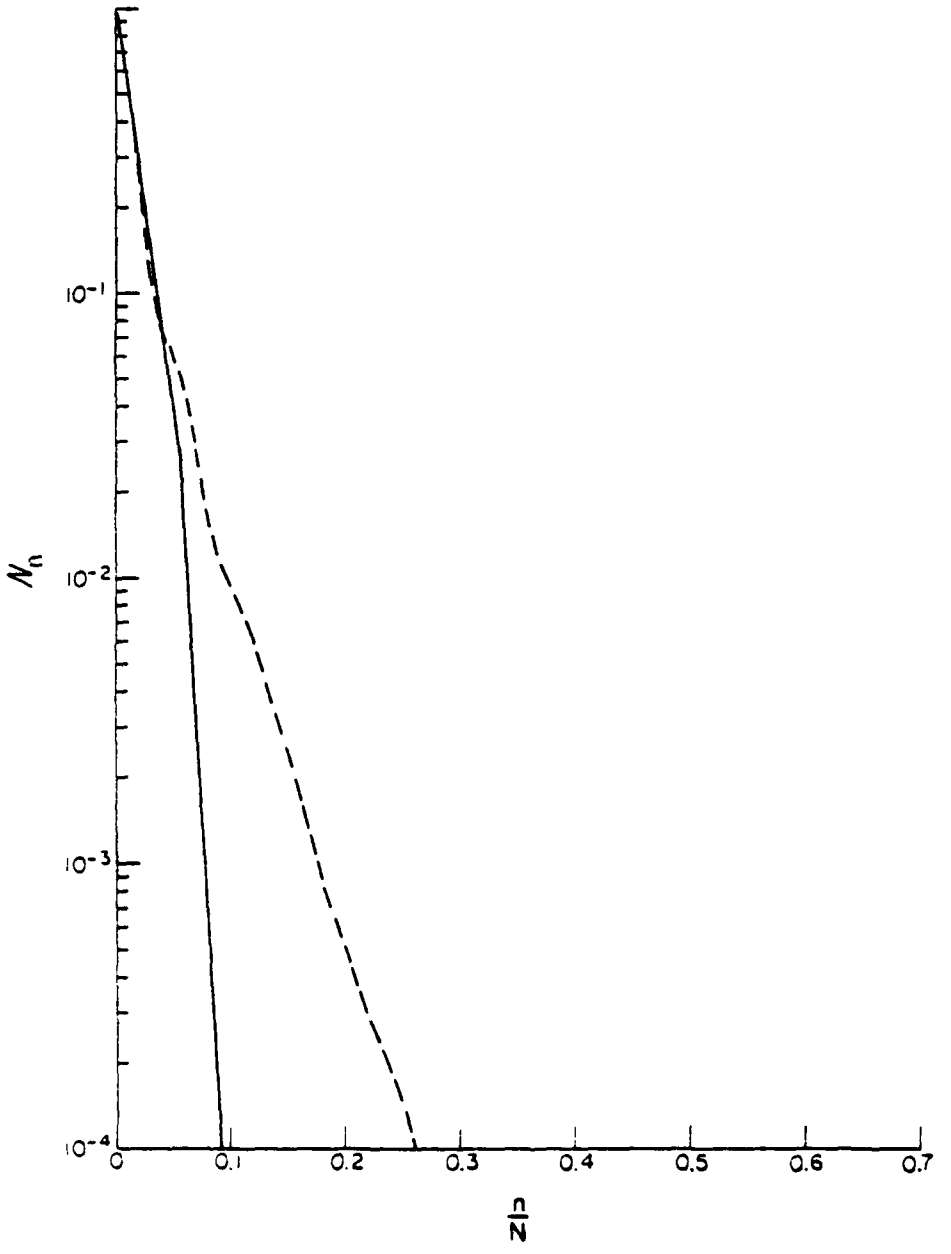

Figure 3.14. Convergence of CGM, OGM and AIGM for a dielectric cylinder, TM polarization.
N = 54, 60 cells/λ_d^2, homogeneous, rectangular slab with ε_r = 3-j10
—— CGM and OGM
---- AIGM

TABLE 3.2.

CONVERGENCE OF THE CGM FOR THREE MODELS OF CIRCULAR DIELECTRIC CYLINDERS WITH RADIUS = 0.08176 λ_0, ε_r = 10, TM POLARIZATION

order of system	N = 21	61	101
cell density	100 $\frac{\text{cells}}{\lambda^2}$	291	481
n = 0	N = 1	1	1
1	0.242	0.249	0.249
2	0.0941	0.101	0.102
3	0.0128	0.0189	0.0197
4	0.000499	0.00102	0.00117
5	0.0000208	0.0000437	0.0000545

Figure 3.15 shows the convergence rates of the CGM when applied to several examples, each modeled with similar cell densities. These examples serve to emphasize the fact that while we observe general trends, individual examples can deviate somewhat from these.

Figure 3.16 shows convergence rates of the CGM for three examples, each with different cell densities. These examples again illustrate the general trend that convergence is usually faster for models having higher cell densities. Note that the cell density of 18.5 cells/λ_d^2 is considered inadequate to produce accurate numerical solutions to the scattering problem.

Table 3.3 shows values of the residual norm versus the unnormalized iteration step for three models of the same scattering problem, in this case a homogeneous circular cylinder. As the cell density increases, the number of iterations required stabilizes, and the values of the residual norm $_n$ appear to stabilize also. As discussed in Section 3.3, we interpret this as an indication that the discrete system is "converging" as the order of the discretization is refined. The question of whether the discrete system is converging to the desired integral equation is reserved for discussion in Chapter 6. Observe that the process seems to stabilize once the cell density exceeds about 100 cells/λ_d^2, an indication that the eigenvalue structure of the matrix has stabilized.

Figure 3.17 shows the convergence of the CGM when applied to models consisting of mixed cell sizes and permittivities, and posessing no simple types of symmetry.

Figure 3.18 shows the convergence rates of the AIGM when applied to several models of dielectric cylinders. Figures 3.19 and 3.20 compare the convergence rates of the CGM and AIGM for two matrix equations. These figures illustrate a trend which has been observed for many examples of the TE dielectric cylinder

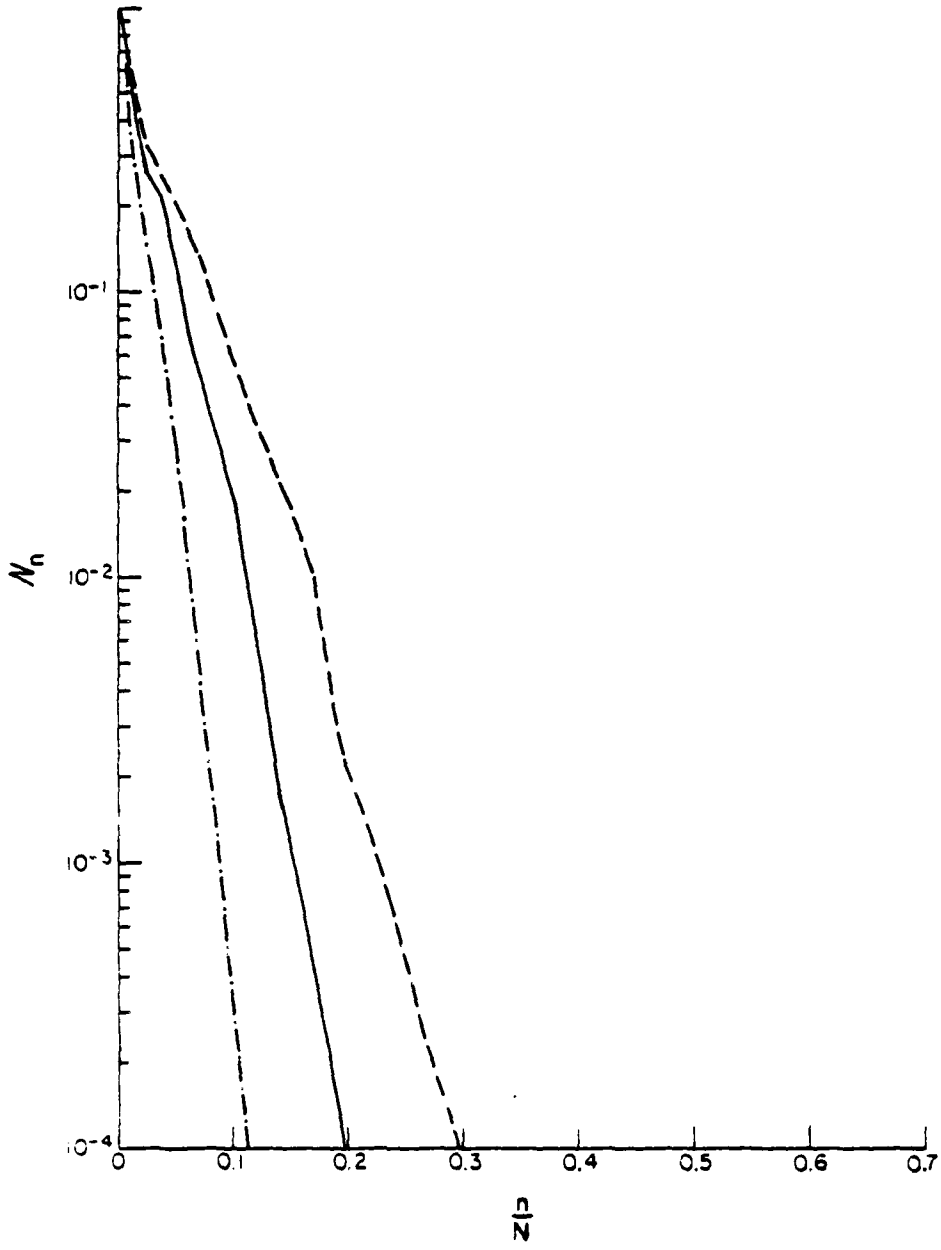

Figure 3.15. Convergence of the CGM for examples of dielectric cylinders, TE polarization.
-·-·- N = 168, 104 cells/λ_d^2, homogeneous circular cylinder with ε_r = 2.56
—— N = 80, 132 cells/λ_d^2, homogeneous rectangular cylinder with ε_r = 3-j0.2
---- N = 42, 104 cells/λ_d^2, homogeneous circular cylinder with ε_r = 2.56

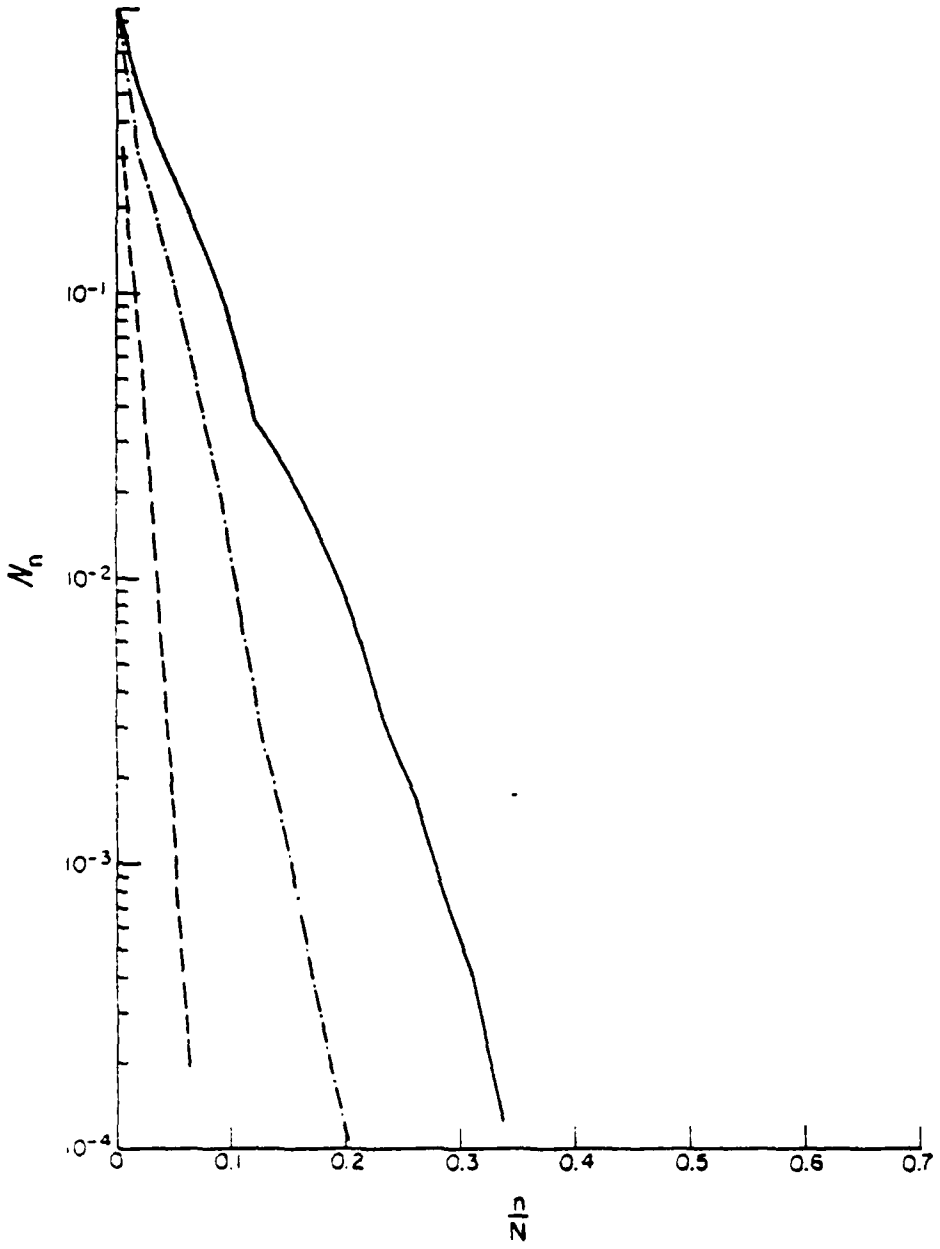

Figure 3.16. Convergence of the CGM for examples of dielectric cylinders, TE polarization.
---- n = 202, 497 cells/λ_d^2, homogeneous circular cylinder, ε_r = 2.56
-·-· N = 80, 132 cells/λ_d^2, homogeneous rectangular cylinder, ε_r = 3-j0.2
—— N = 42, 18.5 cells/λ_d^2, homogeneous circular cylinder, ε_r = 2.56-j2.56

TABLE 3.3.

CONVERGENCE OF THE CGM FOR THREE MODELS OF A CIRCULAR DIELECTRIC CYLINDER WITH RADIUS = 0.3183 λ_0, ε_r = 2.56, TE POLARIZATION

order of system	N = 42	168	672
cell density	$26 \frac{\text{cells}}{\lambda_d^2}$	104	412
n = 0	N = 1	1	1
1	0.587	0.585	0.584
2	0.320	0.340	0.347
3	0.249	0.258	0.263
4	0.156	0.178	0.181
5	0.0881	0.118	0.124
6	0.0564	0.0776	0.0831
7	0.0372	0.0513	0.0570
8	0.0250	0.0366	0.0412
9	0.0127	0.0235	0.0272
10	0.00723	0.0146	0.0166
11	0.00359	0.00880	0.0107
12	0.00214	0.00497	0.00652
13	0.00137	0.00270	0.00361
14	0.000649	0.00161	0.00191
15	0.000238	0.000941	0.00108
16	0.0000876	0.000638	0.000663
17	-	0.000378	0.000457
18	-	0.000171	0.000258

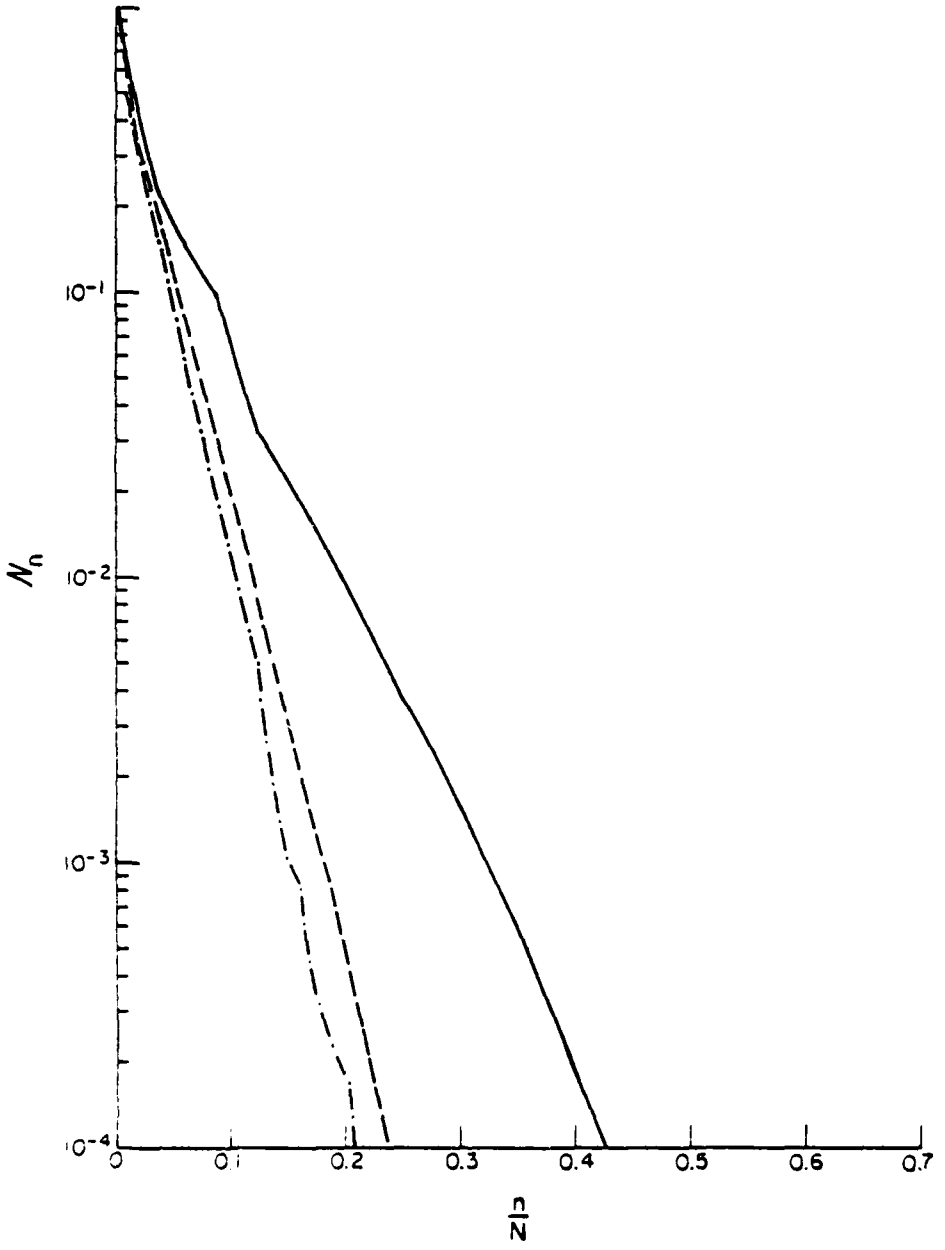

Figure 3.17. Convergence of the CGM for examples of dielectric cylinders, TE polarization.
-·-· N = 90, 125-1063 cells/λ_d^2, 5 square dielectric cylinders in close proximity, ε_r = 2 - j0 to ε_r = 5 - j0.8.
---- N = 80, 100-196 cells/λ_d^2, inhomogeneous skewed-rectangular cylinder with ε_r = 2-j0.4, 3-j0.2, and 4-j0.1
—— N = 80, 992-1962 cells/λ_d^2, inhomogeneous skewed-rectangular cylinder with ε_r = 5-j1 and 10-j1

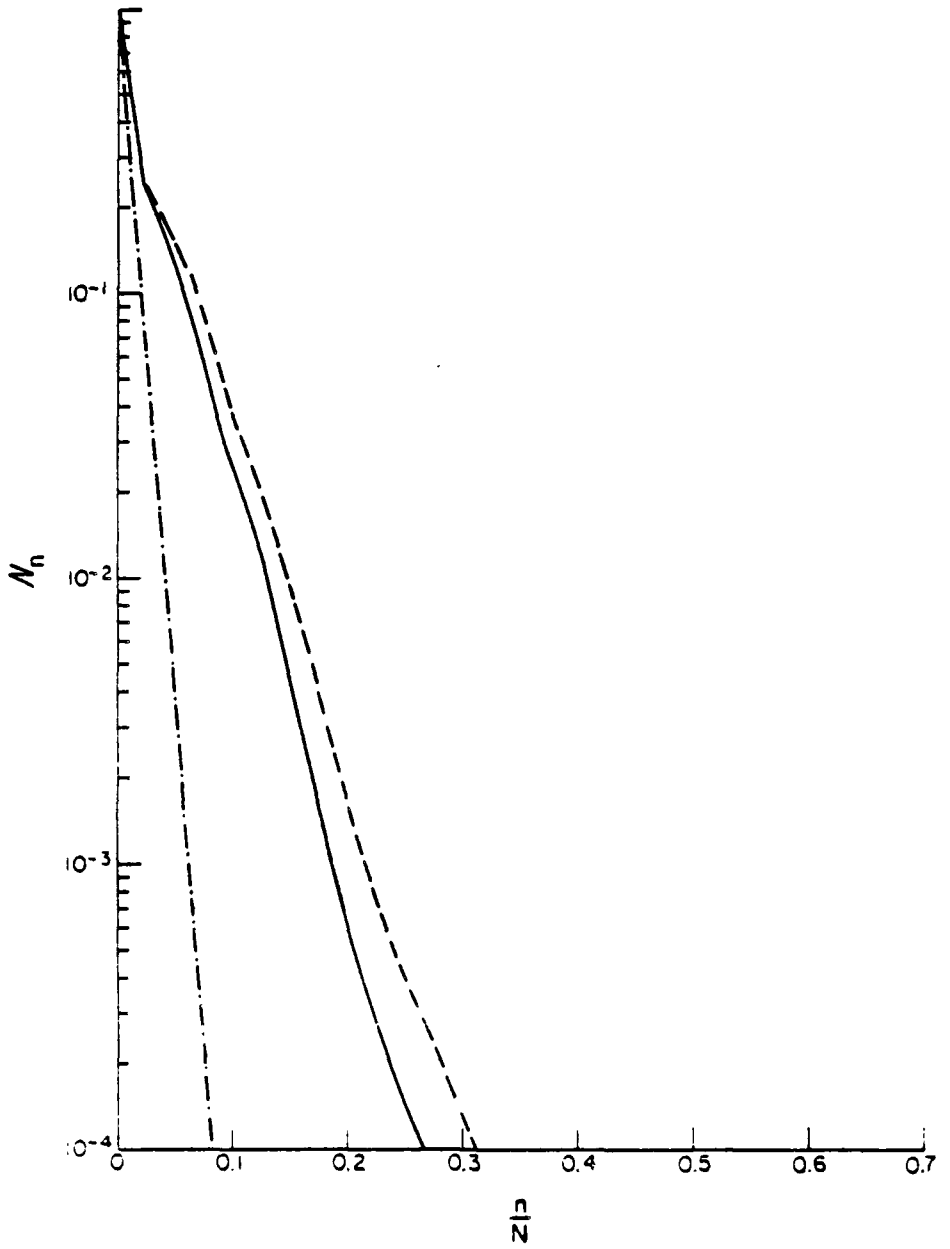

Figure 3.18. Convergence of the AIGM for examples of dielectric cylinders, TE polarization.
-·-· N = 90, 365 cells/λ_d^2, homogeneous rectangular cylinder with ε_r = 3-j0.5
—— N = 42, 100 cells/λ_d^2, homogeneous circular cylinder with ε_r = 5-j2
---- N = 42, 98 cells/λ_d^2, homogeneous circular cylinder with ε_r = 5-j5

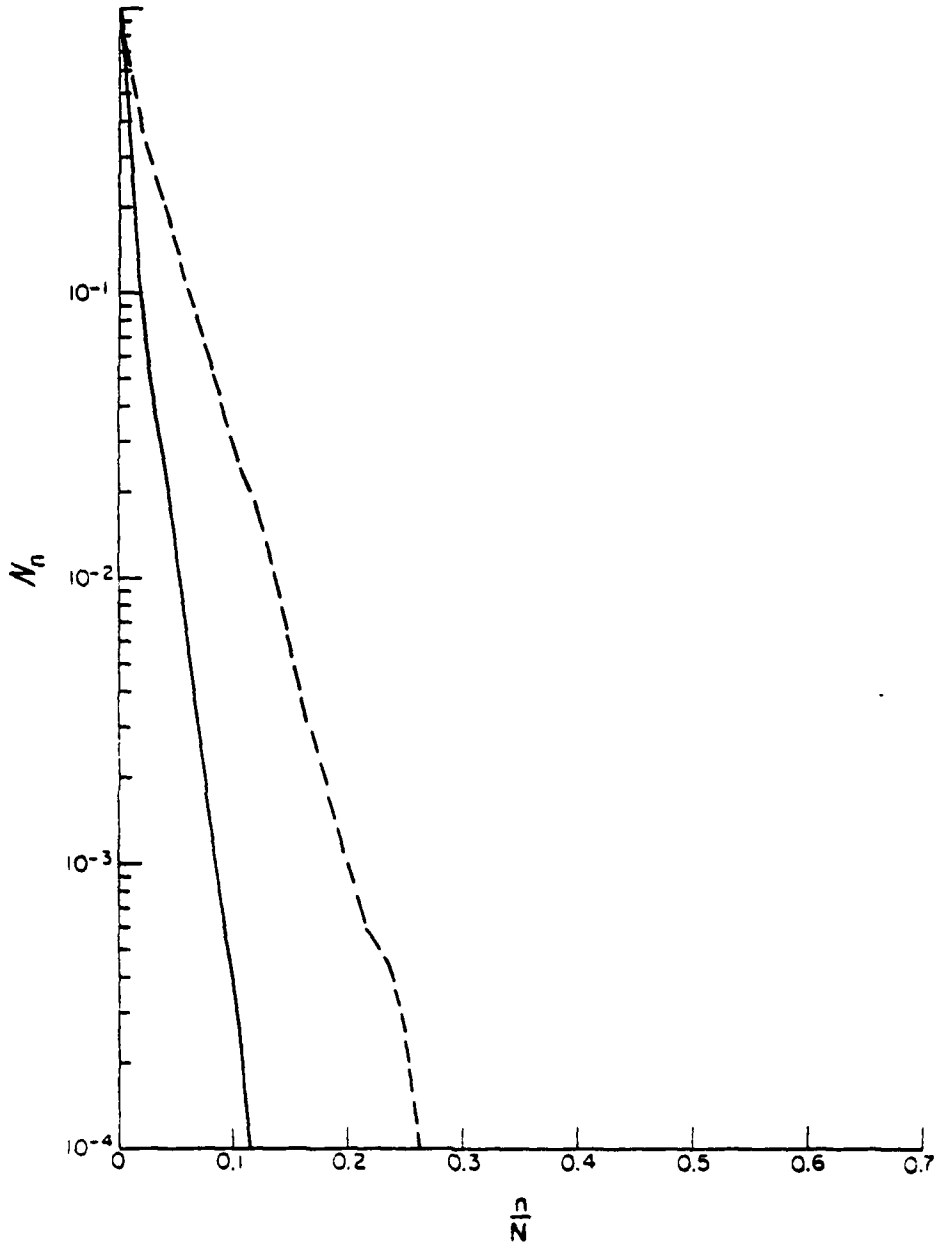

Figure 3.19. Comparison of the convergence of the CGM and AIGM for an example of a dielectric cylinder, TE polarization.
N = 74, 177-222 cells/λ_d^2, 3 square dielectric cylinders in close proximity with ε_r = 2-j0, 5-j0.2, 3.5-j0.4
---- CGM
—— AIGM

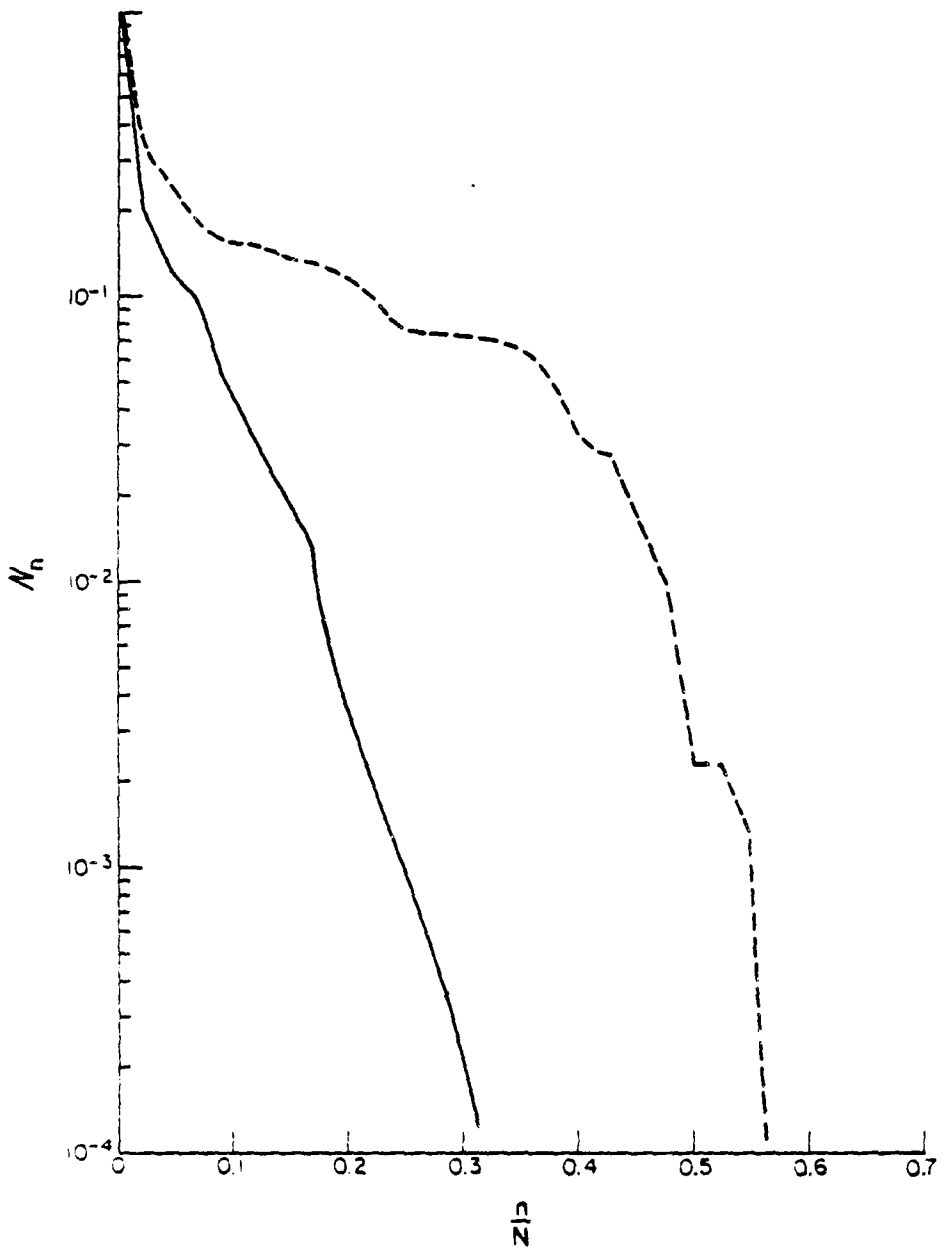

Figure 3.20. Comparison of the convergence of the CGM and AIGM for an example of a dielectric cylinder, TE polarization.
N = 42, 102 cells/λ_0^2, homogeneous circular cylinder with
ε_r = 15-j5
—— AIGM
---- CGM

problem, namely, that the AIGM appears to converge at a faster rate than the CGM. The convergence of the CGM as illustrated in Figure 3.20 is much slower than would normally be expected, considering the previous performance of the CGM for similar problems. In fact, the plot of Figure 3.20 bears some resemblance to the plots of the "internal resonance" problems of Figures 3.4 and 3.10, suggesting that there might be a problem with the accuracy of the numerical result in this case.

To facilitate a study of this type of behavior, Figure 3.21 shows the performances of the CGM and AIGM for an example where the cell density is clearly inadequate to represent the scattering problem in question. As should be expected, the convergence of both algorithms is very slow. The slow convergence of the CGM is indicative of inadequate sampling of the eigenvalue structure of the integral equation; yet the process converges in this case because round-off errors are apparently not severe enough to prevent the finite-step termination of the algorithm. The behavior of the CGM is very similar in the examples of Figures 3.20 and 3.21, yet the AIGM behavior is very different. This is probably related to the accuracy of the approximate inverse in the case where $|\varepsilon_r|$ becomes large. In both cases, the performance of the CGM suggests that there might be a problem with the matrix equation in use. Further study has shown that the accuracy of the particular moment method formulation is poor for large values of $|\varepsilon_r|$, and this topic is examined in detail in Chapter 6. Although we have no reason a priori to expect the formulation to fail in this case, as it turns out there is a problem which might have gone unnoticed were we not familiar with the normal convergence rate of the CGM.

The convergence rate of the CGM for an additional example of a problem which was expected to adequately represent a scattering problem is depicted in Figure 3.22. The numerical solution to the matrix in question bore no

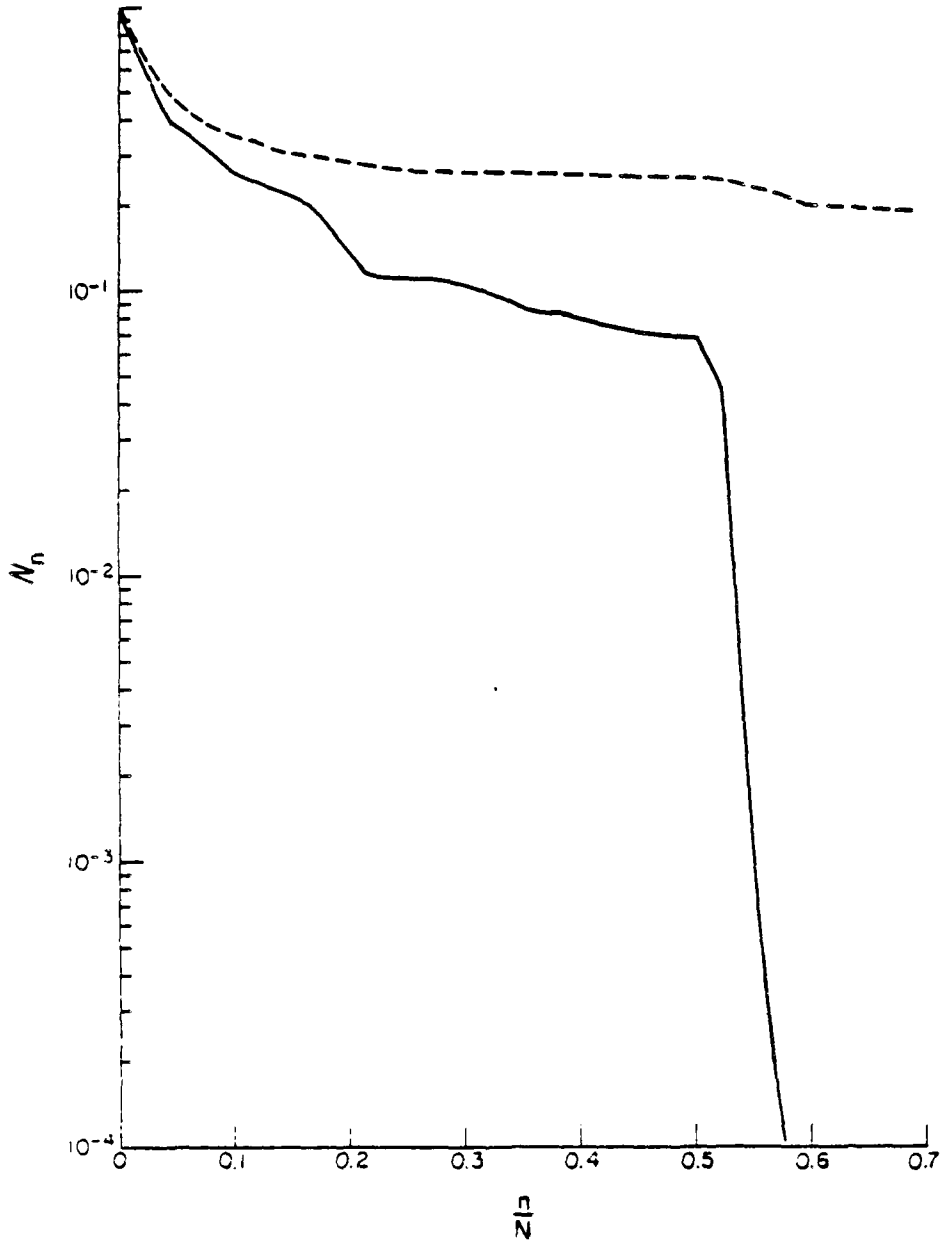

Figure 3.21. Comparison of the convergence of the CGM and AIGM for an example of a dielectric cylinder, TE polarization, with N = 42, 0.0018 cells/λ_d^2, homogeneous circular cylinder, ε_r = 1-j60000.
—— CGM
---- AIGM

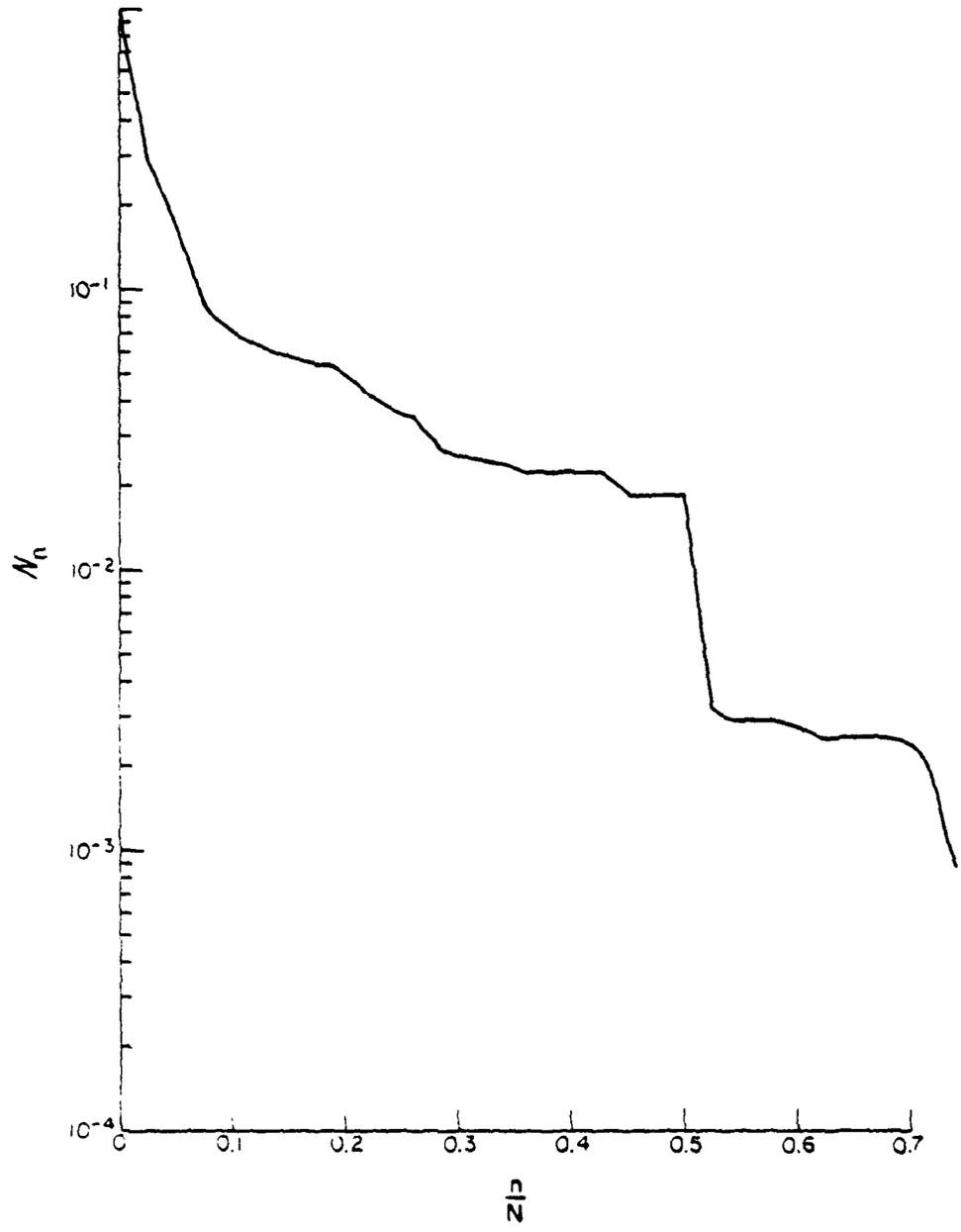

Figure 3.22. Convergence of the CGM for an example of a dielectric cylinder, TE polarization, with N = 42, 231 cells/λ_d^2, homogeneous circular cylinder with ε_r = 76-j278.

resemblence to the analytical solution, which was available for this geometry (a homogeneous, circular cylinder illuminated by a plane TE wave). The permittivity involved here was $\varepsilon_r = 76\text{-}j278$, in the range often used to model certain biological media. More information on the behavior of this problem is presented in Chapter 6.

3.6. Scattering from Cylinders Containing Both Perfectly Conducting and Dielectric Materials for Both TM and TE Polarizations

Chapter 4 considers the iterative implementation of a moment method analysis for cylinders modeled by both perfectly conducting (PEC) and dielectric materials. Specifically, the TM polarization is treated using the EFIE with pulse basis functions and point matching. The TE polarization is treated by combining the EFIE for the dielectric material with the magnetic field integral equation (MFIE) for the PEC material, and again using pulse basis functions and point matching. Because the resulting matrix equations contain terms representing the interaction between the different types of material present, they are somewhat more complicated than the systems presented thus far in this chapter. For instance, the matrix equations considered in Sections 3.3 and 3.4 were of the form

$$\begin{bmatrix} G_{zz} \end{bmatrix} \begin{bmatrix} J_z \end{bmatrix} = \begin{bmatrix} E_z^i \end{bmatrix} \tag{3.11}$$

while the more complicated example of Section 3.5 involved a system of the form

$$\left[\begin{array}{c|c} G_{xx} & G_{xy} \\ \hline G_{yx} & G_{yy} \end{array}\right] \left[\begin{array}{c} J_x \\ \hline J_y \end{array}\right] = \left[\begin{array}{c} E_x^i \\ \hline E_y^i \end{array}\right] \tag{3.12}$$

Equation (3.11) represents a scalar integral equation. Equation (3.12) represents a vector equation which has been separated into $\hat{x}$ and $\hat{y}$ components. The blocks labelled G_{xy} and G_{yx} represent interactions between different vector components. In the combined dielectric and PEC case the matrix equation takes the form

$$\left[\begin{array}{c|c} G_{zz}^{dd} & G_{zz}^{dp} \\ \hline G_{zz}^{pd} & G_{zz}^{pp} \end{array}\right] \left[\begin{array}{c} J_z^d \\ \hline J_z^p \end{array}\right] = \left[\begin{array}{c} E_z^i \\ \hline E_z^i \end{array}\right] \tag{3.13}$$

for the TM polarization and

$$\left[\begin{array}{cc|c} G_{xx}^{dd} & G_{xy}^{dd} & G_x^{dp} \\ G_{yx}^{dd} & G_{yy}^{dd} & G_y^{dp} \\ \hline G_x^{pd} & G_y^{pd} & G^{pp} \end{array}\right] \left[\begin{array}{c} J_x^d \\ J_y^d \\ \hline J^p \end{array}\right] = \left[\begin{array}{c} E_x^i \\ E_y^i \\ \hline H_z^i \end{array}\right] \tag{3.14}$$

for the TE polarization. The sub-matrices appearing in Equations (3.13) and (3.14) represent interactions between different vector components or different materials, and are generally different orders of magnitude than the blocks located along the main diagonal. This affects the condition of the equation, and prompts us to consider scaling the different parts of the matrix to improve the convergence of the iterative algorithms when used to solve these systems. This section shows the performance of the CGM and AIGM when applied to Equations (3.13) and (3.14). Additional information on the discretizations in use may be found in Chapter 4.

Consider the TM polarization described by Equation (3.13). In order to alleviate any problems which may arise due to the presence of different orders-

of-magnitude in the blocks which make up the matrix, a scale factor is introduced to normalize the different sub-matrices to the same order of magnitude. Equation (3.13) takes the form

$$\left[\begin{array}{c|c} G^{dd} & S^{TM}G^{dp} \\ \hline G^{pd} & S^{TM}G^{pp} \end{array}\right] \left[\begin{array}{c} J_z^d \\ \hline \tilde{J}_z^p \end{array}\right] = \left[\begin{array}{c} E_z^i \\ \hline E_z^i \end{array}\right] \tag{3.15}$$

where

$$J_z^p = S^{TM}\,\tilde{J}_z^p \tag{3.16}$$

and the scale factor S^{TM} is chosen to give the two types of unknowns J_z^d and $\tilde{J}_z^p$, the same order of magnitude. For the following examples, the value $S^{TM} = 0.025$ was used with the specific form of the equation described in Chapter 4.

Figure 3.23 shows the convergence of the CGM for three different circular PEC cylinders coated in each case with a single, uniform layer of lossless dielectric material. Overall, the convergence behavior is similar to that seen in previous sections of this chapter. Note that in keeping with the previous results there is a general trend of relatively quick convergence (normalized to the order of the system) as higher cell densities are used in the models.

Figure 3.24 shows the convergence of the CGM for three examples involving a single, circular PEC cylinder in the presence of one or more rectangular dielectric cylinders.

Figure 3.25 compares the convergence of the CGM and AIGM for a matrix equation representing a circular PEC cylinder with a uniform dielectric coating.

Although we introduced a scale factor, experimentation showed that the rate of convergence of the iterative algorithms was not a strong function of the specific scale factor in use, for the TM polarization. The system representing the

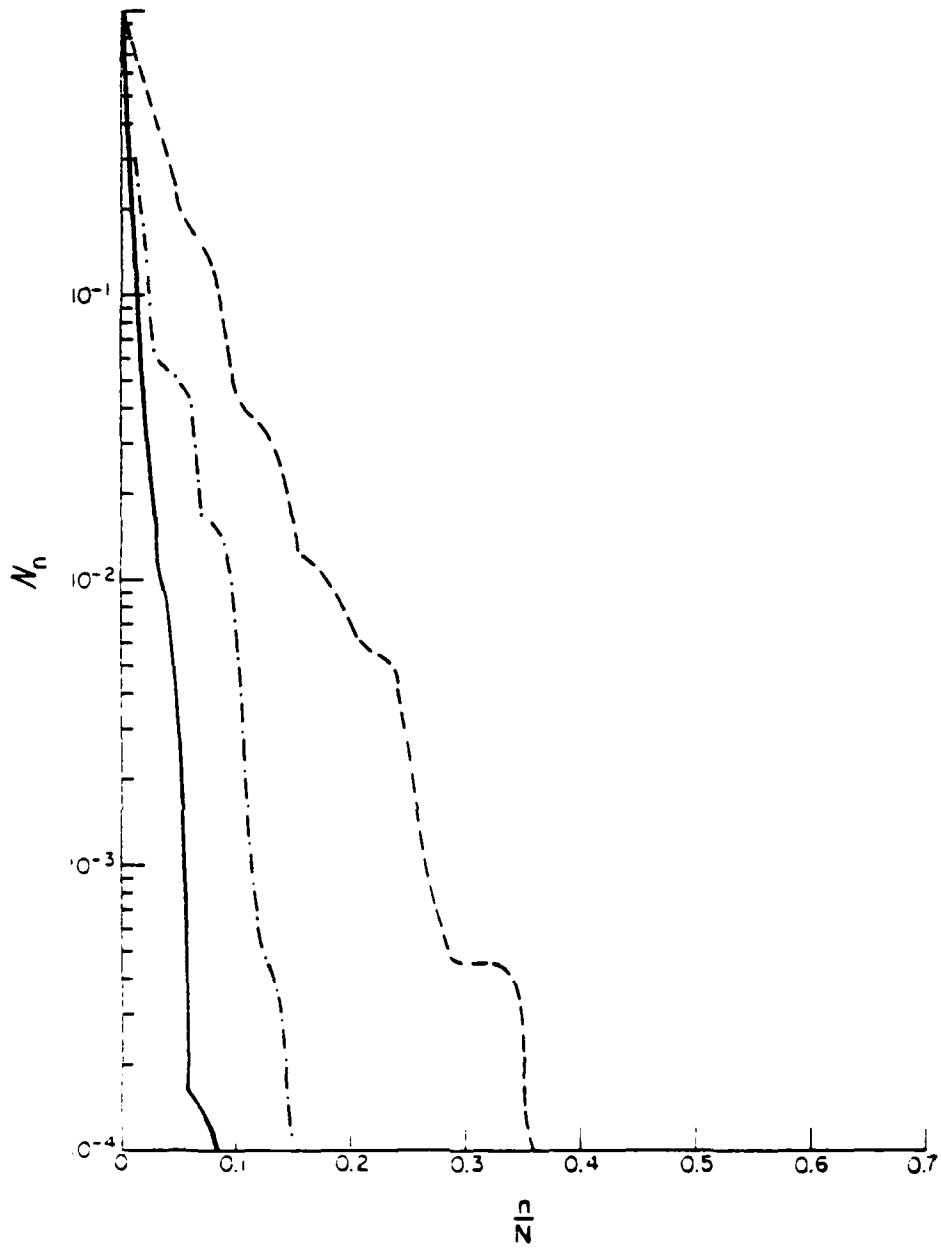

Figure 3.23. Convergence of the CGM for examples of combination dielectric/PEC cylinders, TM polarization, circular PEC cylinders with homogeneous dielectric coating, scale factor $S^{TM} = 0.025$.
—— N = 99, 66.0 cells/λ_0 PEC, 6180 cells/λ_d^2 dielectric
-·-· N = 99, 16.5 cells/λ_0 PEC, 384 cells/λ_d^2 dielectric
---- N = 38, 10.0 cells/λ_0 PEC, 151 cells/λ_d^2 dielectric

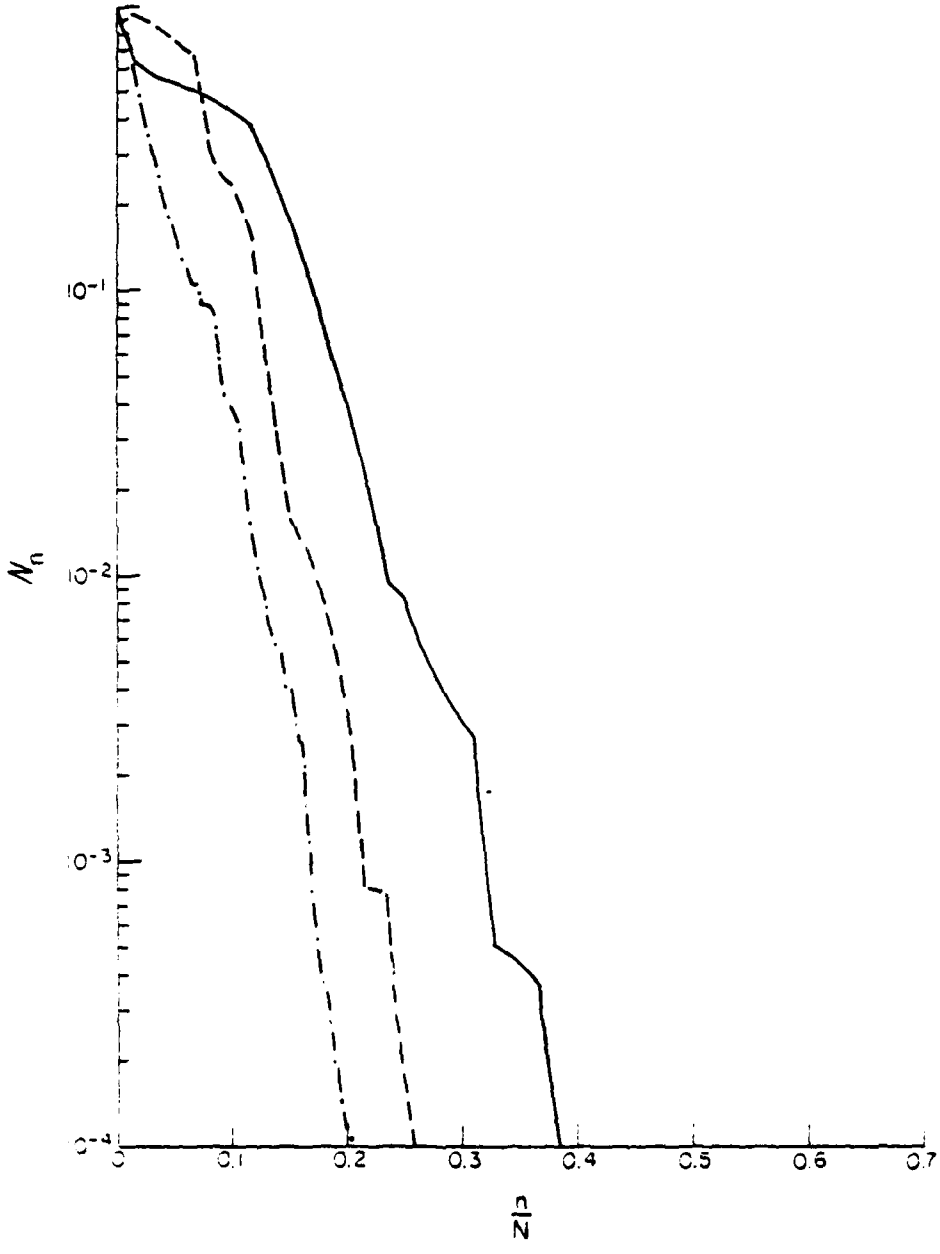

Figure 3.24. Convergence of the CGM for examples of combination dielectric/PEC cylinders, TM polarization, with scale factor $S^{TM} = 0.025$.
- -·- N = 272, 10.0 cells/λ_0 PEC, 110 cells/λ_d^2 dielectric, circular PEC cylinder behind square cylinder, with ε_r = 2-j0.3
- ---- N = 60, 10.0 cells/λ_0 PEC, 132-279 cells/λ_d^2 dielectric, circular PEC cylinder near 3 square dielectric cylinders with ε_r = 1.5-j0.1, 3.0-j0.3, 4-j0
- —— N = 52, 10.0 cells/λ_0 PEC, 269-424 cells/λ_d^2 dielectric, circular PEC cylinder near 2 rectangular dielectric cylinders with ε_r = 3-j0.5, 4-j0.8

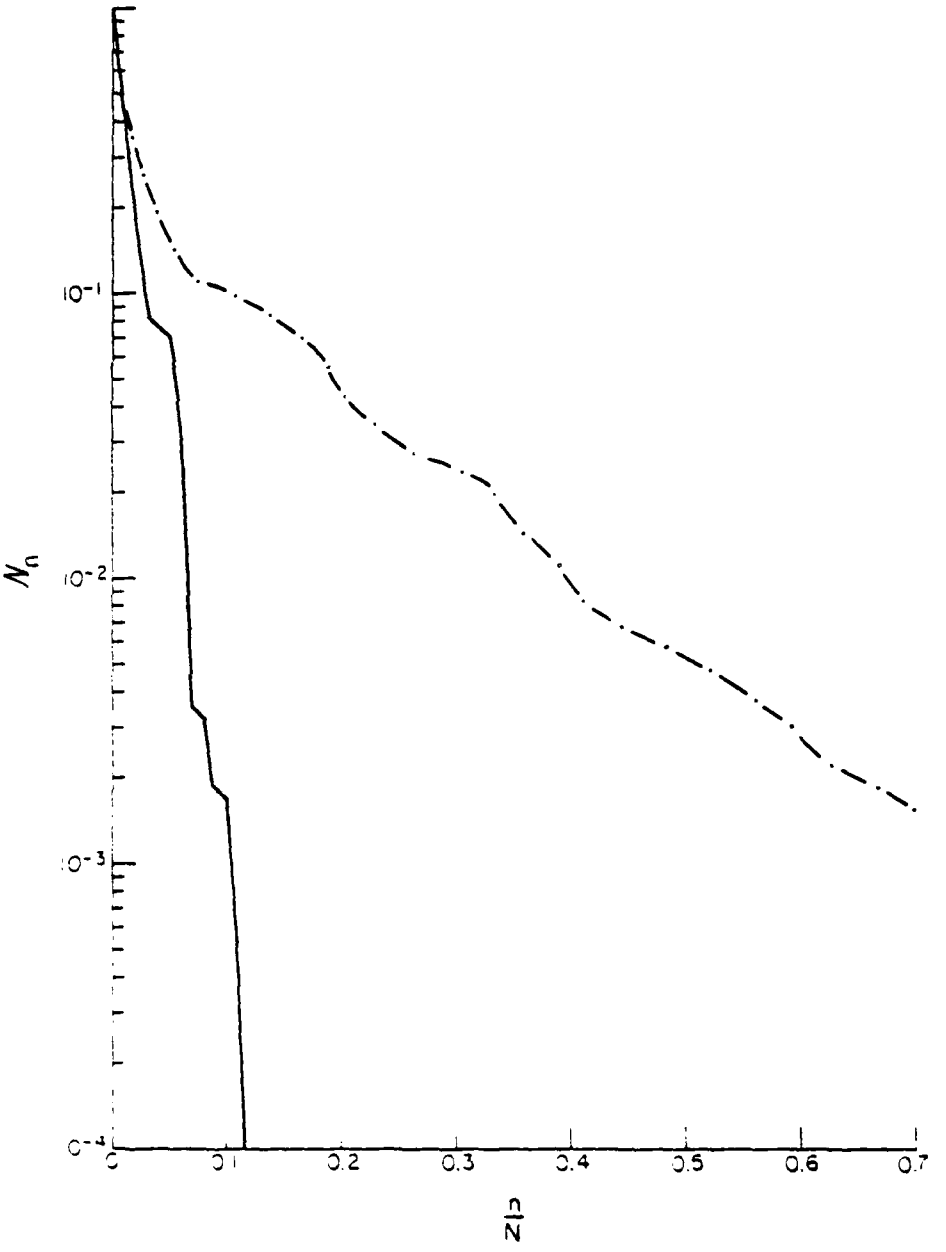

Figure 3.25. Comparison of CGM and AIGM for example of circular PEC cylinder coated with homogeneous dielectric layer, TM polarization, $S^{TM} = 0.025$, N = 102, 25.5 cells/λ_0 PEC, 161 cells/λ_d^2 dielectric, $\varepsilon_r = 3.5\text{-}j0.5$.
—— CGM
---- AIGM

TE polarization appears to be much more sensitive to proper scaling, as illustrated below.

Equation (3.14), representing the TE polarization, can be scaled to take the form

$$\left[\begin{array}{c|c|c} G_{xx}^{dd} & G_{xy}^{dd} & S_1^{TE} G_x^{dp} \\ \hline G_{yx}^{dd} & G_{yy}^{dd} & S_1^{TE} G_y^{dp} \\ \hline S_2^{TE} G_x^{pd} & S_2^{TE} G_y^{pd} & S_1^{TE} S_2^{TE} G^{pp} \end{array}\right] \left[\begin{array}{c} J_x^d \\ J_y^d \\ \tilde{J}^p \end{array}\right] = \left[\begin{array}{c} E_x^i \\ E_y^i \\ S_2^{TE} H_z^i \end{array}\right] \tag{3.17}$$

where

$$J^p = S_1^{TE} \tilde{J}^p \tag{3.18}$$

The scale factors $S_1^{TE} = 0.01$ and $S_2^{TE} = 377$ appear to work well with the CGM to ensure normal convergence behavior. As an example, Figure 3.26 shows the convergence of the CGM for a circular PEC cylinder with a uniform dielectric coating. Note that without scaling, the CGM does not converge to a solution.

Figure 3.27 shows the convergence of the CGM for geometries involving a single circular PEC cylinder in the presence of one or more rectangular dielectric cylinders. Typically, convergence was fairly rapid for this type of model.

Figure 3.28 shows the convergence of the CGM and AIGM for an example involving a circular PEC cylinder and two square dielectric cylinders. Although the AIGM is unable to produce a solution with $N_n < 0.0001$ in this case, the solution found after 36 iterations was reasonably accurate. The slow convergence of the CGM for this example suggests the possibility of a problem; in fact, the

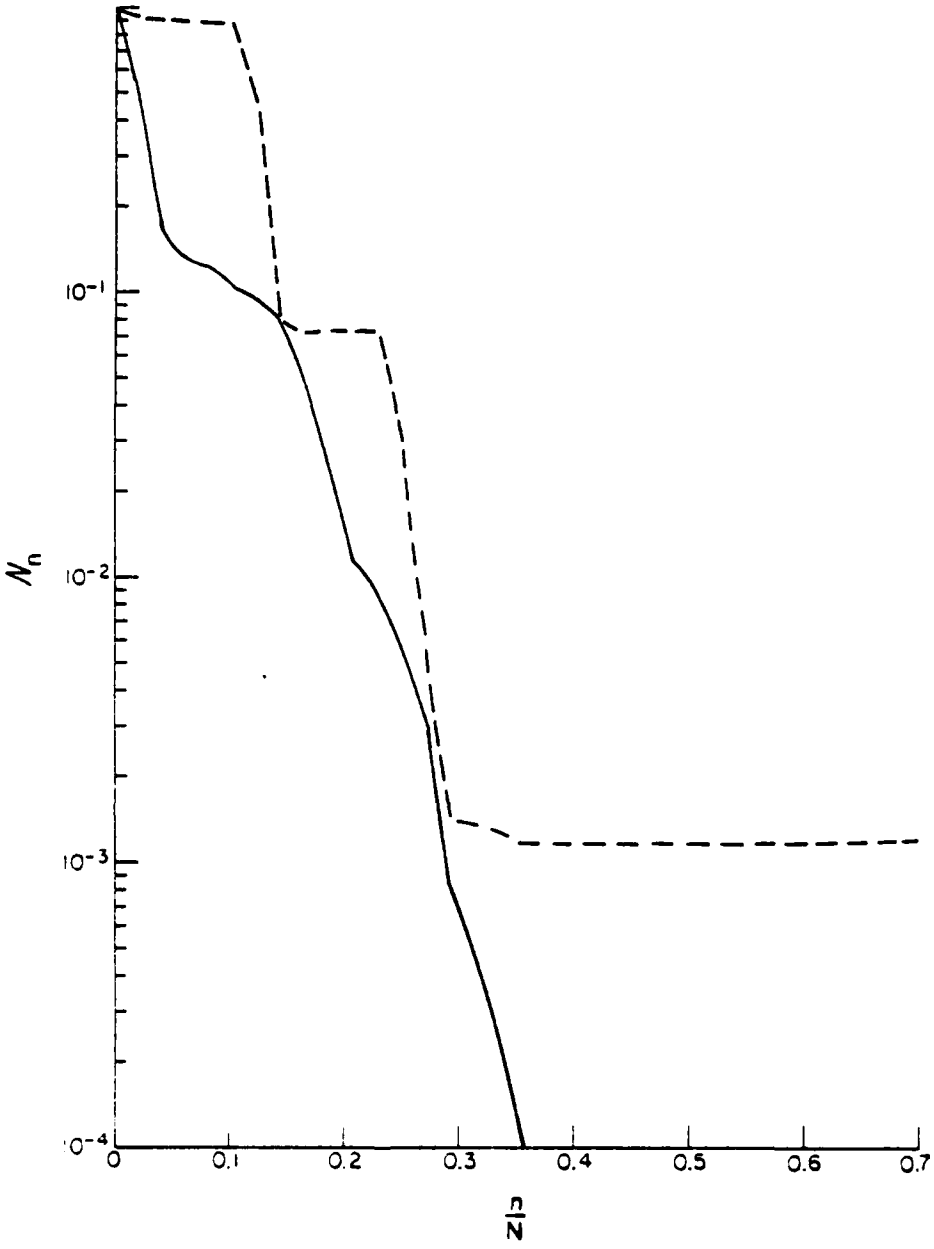

Figure 3.26. Convergence of CGM for an example of a circular PEC cylinder coated with a homogeneous dielectric layer, TE polarization.
N = 48, 25.5 cells/λ_0 PEC, 137 cells/λ_d^2 dielectric, ε_r = 3-j0
——— S_1^{TE} = 0.01 S_2^{TE} = 377
---- S_1^{TE} = 1 S_2^{TE} = 1

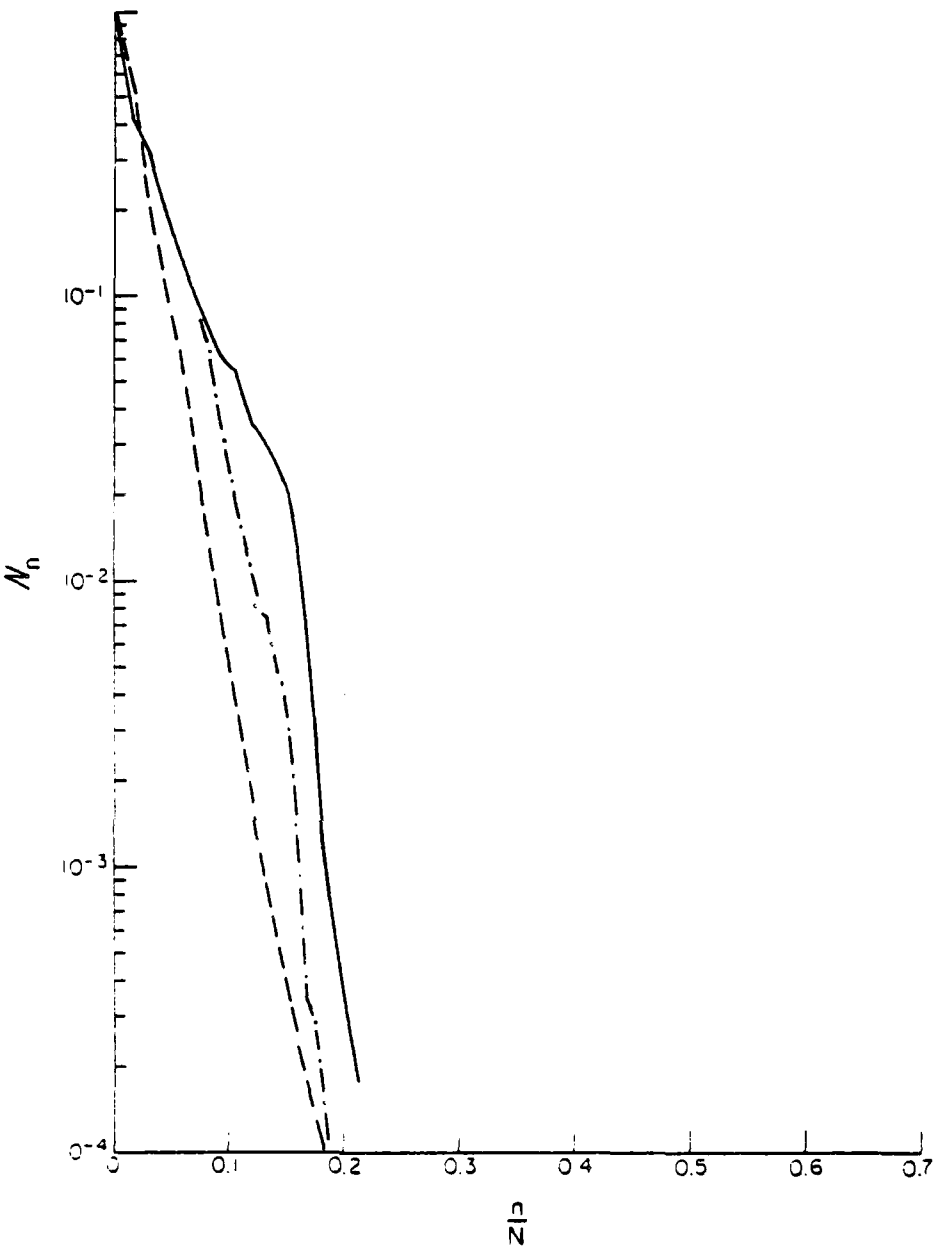

Figure 3.27. Convergence of the CGM for examples of combined dielectric/PEC cylinders, TE polarization $S_1^{TE} = 0.01$, $S_2^{TE} = 377$.
---- N = 110, 10.0 cells/λ_0 PEC, 132-279 cells/λ_d^2 dielectric, circular PEC cylinder with 3 square dielectric cylinders, ε_r = 1.5-j0.1, 3-j0.3, 4-j0
•-•- N = 137, 10.0 cells/λ_0 PEC, 110 cells/λ_d^2 dielectric, circular PEC cylinder and square dielectric cylinder, ε_r = 2-j0.3
—— N = 67, 10.0 cells/λ_0 PEC, 193 cells/λ_d^2 dielectric, circular PEC cylinder and square dielectric cylinder with ε_r = 2-j0.5

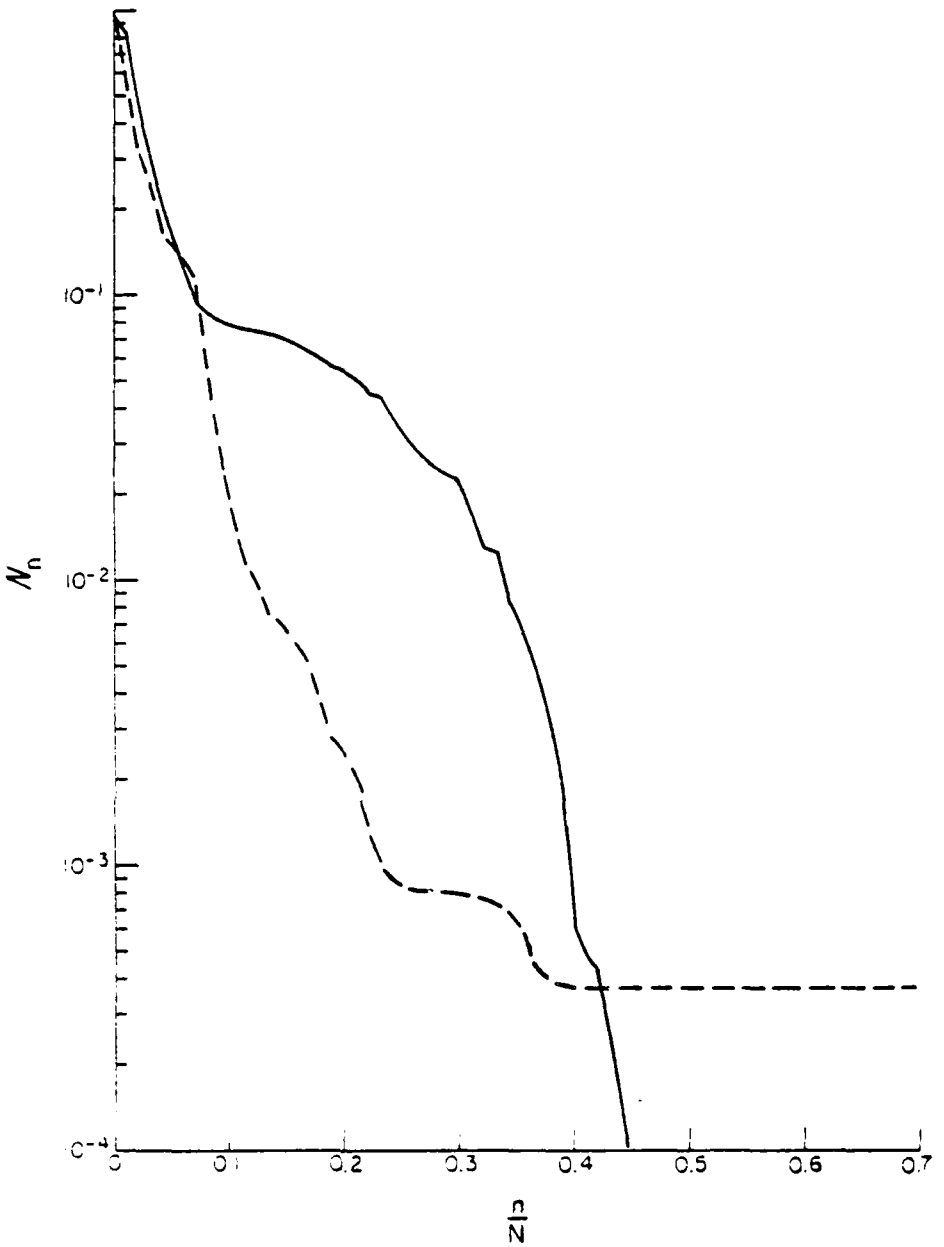

Figure 3.28. Comparison of the convergence of the CGM and AIGM for an example of a cylinder containing both dielectric and PEC material. N = 90, 10.0 cells/λ_0 PEC, 204-369 cells/λ_d^2 dielectric, ε_r = 3-j0 and 3-j0.5, S_1^{TE} = 0.01, S_2^{TE} = 377, circular PEC cylinder and 2 square dielectric cylinders.

—— CGM

---- AIGM

performance of this numerical approach is very sensitive to some of the approximations made in the evaluation of the matrix elements. A detailed discussion of these approximations and their effects on the accuracy of the final solution is provided in Chapter 4. The particular example illustrated in Figure 3.28 is borderline in that both the rate of convergence and the accuracy of the numerical solution could be improved somewhat by better evaluation of the matrix elements.

3.7. Use of the CGM to Treat Multiple Excitations

As mentioned in Section 3.1, the amount of calculation per iteration performed by the CGM algorithm is roughly divided between the tasks of generating the expansion functions (the p-functions of Section 2.3) and computing their coefficients. Thus, one way of efficiently treating multiple right-hand sides is to simultaneously expand several solutions in terms of a single set of expansion functions. In theory, this approach will permit us to save about half of the cost per iteration of treating each additional solution (corresponding to each additional right-hand side) without adding significantly to the storage requirements. In order to test this idea in practice, the following data were generated.

Figure 3.29 shows the rate of convergence of the CGM as illustrated by the residual norm for five different systems arising in the analysis of a single, circular PEC cylinder by the moment method approach of Section 3.3. Each solution corresponds to a different excitation, in this case, plane waves impinging on the scatterer from five different angles. The p-functions from the 0 degree incidence wave were used to expand all five solutions. The process, however, evidently does not yield the desired improvement in efficiency, as only the 0 degree solution is obtained to the necessary accuracy. Some of the other

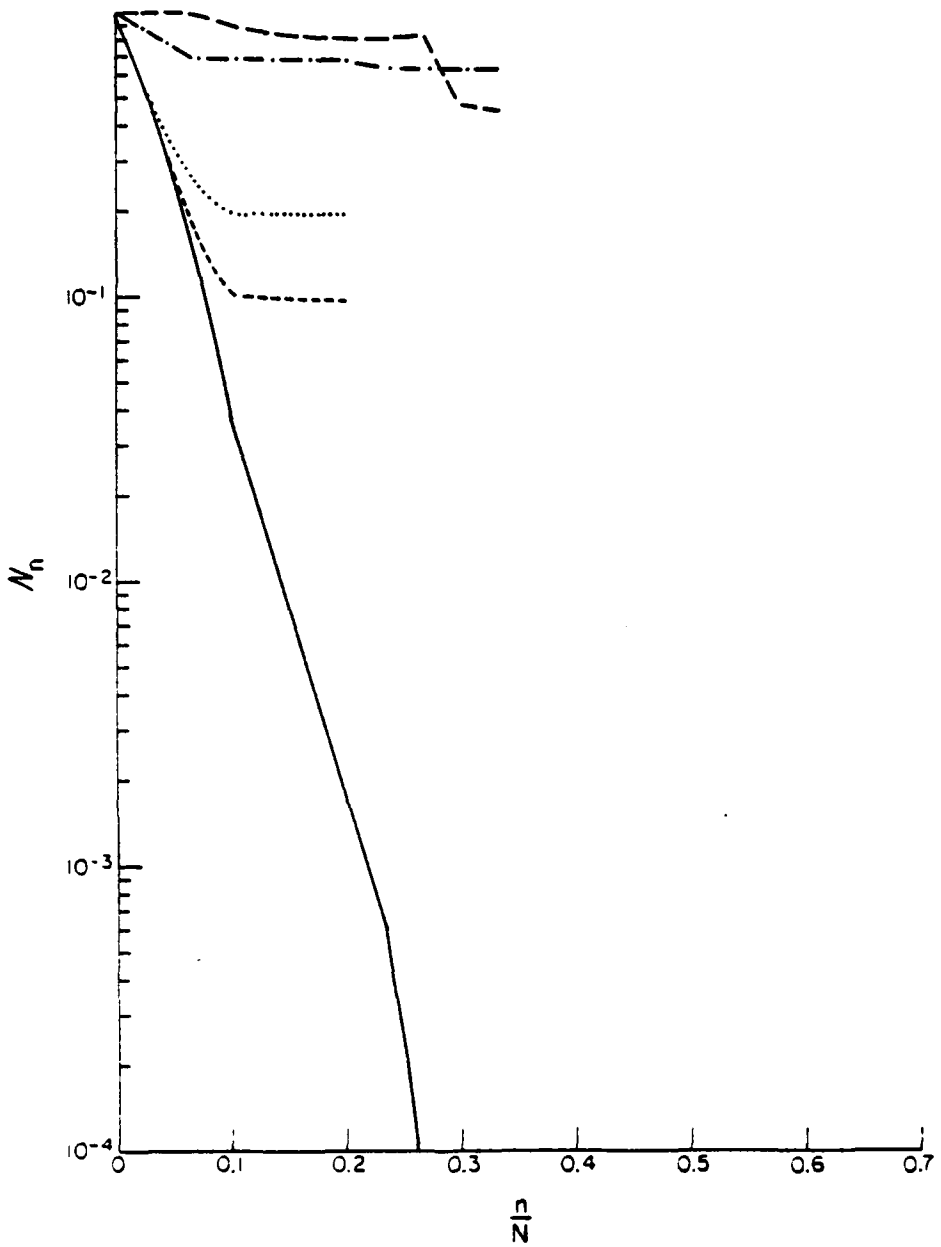

Figure 3.29. Convergence of the CGM modified to treat simultaneous excitations for an example of a circular PEC cylinder, TM polarization. N = 30, 9.6 cells/λ_0.

——— $\theta^i = 0°$
------- $\theta^i = 2.5°$
..... $\theta^i = 5.0°$
·-·-·-· $\theta^i = 90°$
— — — — $\theta^i = 45°$

solutions are clearly no better than their initial estimate, even after the solution to 0 degree incidence is obtained.

Figure 3.30, representing to a different scatterer, illustrates similar behavior when the CGM is applied to the simultaneous treatment of three different excitations. Clearly, in this case it would be more efficient to treat each system separately.

The reasons for the failure of the above technique are necessarily related to the strong point of the CGM -- namely, its quick convergence in most cases. Because the expansion functions are geared to represent the solutions corresponding to the specific right-hand side in use, the convergence of the algorithm is relatively fast. These functions, however, are not geared to represent any function, even though in theory a finite number of them span the space. Because of round-off errors affecting the orthogonality, they apparently do not span the space (in spite of the fact that they can represent the single solution they were generated for). In addition, it has been mentioned that the p-functions may take on the symmetries, if any, of the corresponding solutions. Thus, they clearly would not be useful for the representation of solutions which did not possess the same symmetries. For all of these reasons, the simultaneous expansion of several solutions does not appear effective for the efficient treatment of multiple systems.

An alternate approach, that of using a previous solution to generate an initial estimate of a different solution, may prove feasible for the treatment of multiple right-hand sides. However, as evidenced by the theoretical convergence rates given in Section 3.2 and by observations made in practice, the rate of convergence is relatively independent of the initial guess. Thus, a significant improvement in efficiency is only possible by finding a fairly accurate initial estimate of the solution.

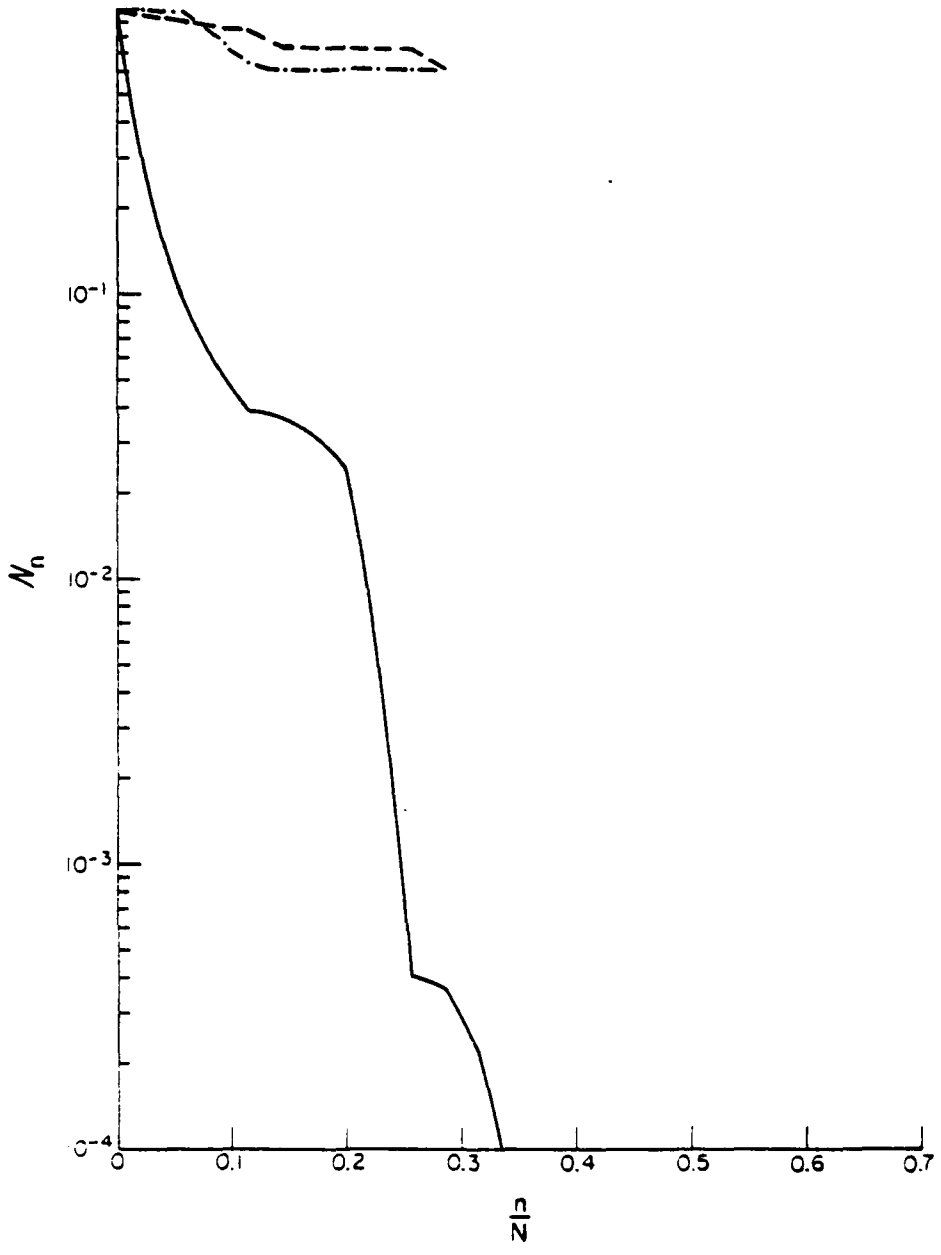

Figure 3.30. Convergence of the CGM modified to treat simultaneous excitations for an example of a circular PEC cylinder, TM polarization. N = 35, 9.3 cells/λ_0.
—— $\theta = 0°$
•–• $\theta = 90°$
– – – $\theta = 45°$

Generally speaking, the goals of fast convergence and ability to treat multiple right-hand sides are somewhat diametrically opposed, as complete flexibility in the latter mandates that the full set of expansion functions be available (which negates the former). A better approach for dealing with multiple right-hand sides would require an algorithm which does not converge fast for a single problem, but does generate useful expansion functions for the treatment of other systems. A non-iterative method such as Gaussian elimination may be the best currently available example of such an algorithm, in spite of the additional storage required by all general direct solution procedures.

3.8. Summary

This chapter has presented a discussion of the theoretical convergence of the CGM, and has illustrated the convergence of all three of the iterative algorithms of Chapter 2 for a variety of electromagnetic scattering problems. The examples presented were selected from a wide range of test cases, and are believed to represent the extremes that typically arise in practice.

Based upon these data, the CGM solution of a moment method system usually requires N/4 to N/2 iterations, assuming that the cell densities in use are sufficient to ensure adequate sampling of the original integral equation, and where N is the order of the matrix equation. However, it is not uncommon for the algorithm to require as many as 3N/4 iterations, especially if the moment method formulation involves mixed cell sizes as was observed in Section 3.3. As higher cell densities are used, faster rates of convergence (relative to N) are obtained. If convergence is much slower than this, or if the residual norm remains virtually constant for many iterations, it is possible that the matrix is not an accurate model of the original equation and should be modified. On the other hand, very fast convergence, i.e., N/10 iterations or less, has never been observed for the examples considered here unless the cell densities in use are

extremely high. The convergence behavior of the CGM is linked to the eigenvalue structure of the matrix, which itself is an approximation to the eigenvalue spectrum of the integral operator. Thus, the convergence behavior of the CGM is feedback which appears to be useful for indicating the adequacy of the discretization used to form the matrix. These conclusions are based upon observations made with several types of integral equations representing electromagnetic scattering problems and may not be applicable to other types of equations.

The AIGM algorithm converged slower than the CGM and OGM for all of the systems except the TE dielectric cylinder matrix equation. However, this example shows that even with the simple idea of generating an approximate inverse operator by inverting the main diagonal of the matrix, the technique is feasible and may be more efficient than the CGM. It may be advantageous to expend additional effort to generate a better inverse, such as inverting a banded approximation to the matrix. For a specific application, more information can be brought to bear on the problem of finding a good approximate inverse. Our purpose here was to examine general algorithms that were not tied to any particular geometry or symmetry, and thus we did not attempt to find a better approximate inverse for any of the examples considered.

4. THE MATRIX ELEMENT REGENERATION (MER) APPROACH

4.1. Introduction

Iterative techniques have been employed in the solution of electromagnetic scattering problems primarily because they permit any sparseness or redundancy of the associated matrix equation to be fully exploited. It has long been appreciated that the computer storage required for the direct solution of a large system places a practical limit on the electrical size of the scatterer to be analyzed. Thus, specialized approaches have been developed which build sparseness or redundancy into the matrix equation, permitting more efficient use of fast-access memory. Examples of these approaches are discussed in Chapters 5 - 7.

The discretizations used with the above-mentioned "special" approaches limit the scope of these methods in most cases to problems involving surfaces of constant curvature or volumes which can be represented by evenly-spaced subdomains. The present chapter investigates a different implementation of iterative algorithms, one that is suitable for the treatment of arbitrary geometries. In general, these methods are not as efficient as the specialized techniques of Chapters 5 - 7, because they require more computation per iteration step. However, they may provide an effective alternative for geometries which are not easily treated by the specialized methods.

4.2. The Matrix Element Regeneration (MER) Approach

The approach to be investigated is based upon the simple idea of "recomputing" each matrix element as it is needed in the iterative solution of the matrix equation. We denote this the matrix element regeneration (MER) approach. The goal of this investigation is to determine if the MER can be implemented efficiently for different types of electromagnetic scattering

problems. Earlier work on this topic dealt primarily with problems represented by a simple scalar integral equation, and proved encouraging [54] - [57]. Here, the MER is applied to a more complicated integral equation representing a scatterer modeled by a combination of perfectly conducting and lossy dielectric materials.

The idea behind the MER approach is that it may be more efficient to generate each matrix element as it is needed than to repeatedly transfer the necessary numerical values from an out-of-core storage device such as a disk or magnetic tape unit. In order to test this idea, the relative execution times of both approaches are compared.

Since simple discretization schemes, such as the moment method using pulse basis functions and point matching, are likely to produce matrix equations with relatively simple matrix elements, they appear to be the best candidates for efficient MER implementation. The methods to be considered here are all based upon these simple basis and testing functions. However, the use of simple discretizations has sometimes led to inaccuracy in the final numerical solution [81]. In order to determine the overall value of the methods, the accuracy of each numerical approach should be investigated thoroughly for a variety of problems. While we do not attempt an exhaustive study here, several examples are given in order to compare numerical results with exact analytical solutions.

The objective of the MER process is to reduce the amount of storage required for an NxN system to some small multiple of N. Information describing the geometry and materials can be stored directly, as can the matrix elements comprising the main diagonal of the NxN system. Savings are obtained by reducing the required storage for the off-diagonal matrix elements.

4.3. TM-wave Scattering by Conducting Cylinders

As an example, consider the problem of TM-wave scattering by perfectly conducting infinite cylinders. A detailed description of the formulation and

discretization of the electric-field integral equation are provided by Harrington [74]. Specifically, the approach requires a conducting surface to be modeled by subsectional strips, over which the unknown current density is treated as a constant (i.e., pulse basis functions). The equation is enforced by point matching at the center of each strip. The off-diagonal matrix elements can be found approximately by replacing the pulse basis functions by Dirac delta functions located at the center of each strip, and take the form

$$Z_{mn} = \alpha_n H_0^{(2)}(\rho_{mn}) \qquad m \neq n \tag{4.1}$$

In Equation (4.1), ρ_{mn} represents the distance from the center of strip n to the center of strip m, and $H_0^2(*)$ is the zeroth-order Hankel function of the second kind. The Hankel function can be computed efficiently using linear interpolation from a look-up table, a task requiring only a few arithmetic operations. Note that although the size of the look-up table varies in proportion with the maximum linear dimension of the scatterer geometry under consideration, it is typically much smaller than the NxN matrix it replaces.

Execution time data for small-order systems are provided in a report by Peterson and Mittra [57], and are reproduced in Figure 4.1. The original data show the execution time required to iteratively solve the system using the CGM and to compute the bistatic radar cross section, for each of two cases. In the first case the entire NxN matrix is stored in fast-access memory; in the second case all of the off-diagonal matrix elements are recomputed whenever needed in accordance with the MER procedure. The original data have been augmented with data showing the corresponding execution time when each of the off-diagonal matrix elements is transferred from a disk storage device as needed. The CDC CYBER 175 computer used for the study is a main-frame time-sharing machine, and

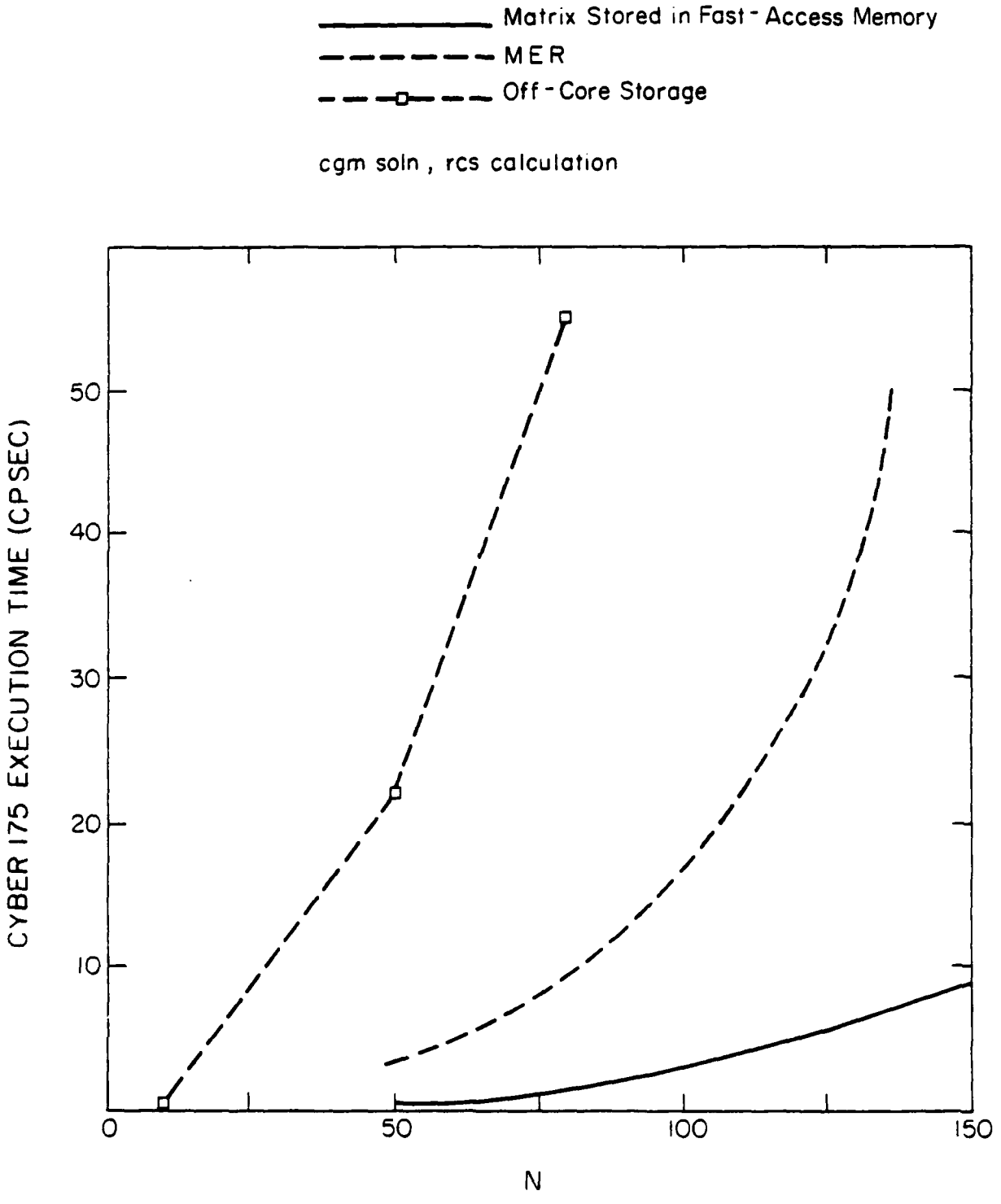

Figure 4.1. Relative execution times for PEC cylinders, TM polarization.

the data may be machine dependent to some extent. In addition, no attempt was made to transfer blocks of data rather than single elements, which might improve the efficiency. Under these conditions, the MER process is more efficient than out-of-core storage. Note that the above procedure was done as a test; in practice, it is beneficial to use as much of fast-access memory as is available, and examples to follow are based upon a more efficient approach.

In order to judge the accuracy of the particular discretization employed for this problem, data from different levels of discretization are compared to the exact solution for a circular cylinder illuminated by a plane wave. Table 4.1 shows the convergence of the monostatic radar cross section as the order of discretization is refined. (A definition of radar cross section for this problem is available in [57].) Tables 4.2 and 4.3 show the current density in the center of the shadow region and at the specular point on the cylinder. The numerical solutions appear to be converging to the exact values. (In this and other examples to follow, in order to show numerical convergence, it is necessary to replace the look-up tables used for the Bessel functions for MER implementation with more accurate values. The look-up tables for the Hankel function of order zero exhibit a maximum error of 1 percent.) The numerical convergence observed in Tables 4.1 - 4.3 suggests that the various approximations used within the moment method formulation are acceptable.

4.4. TM-wave Scattering by Dielectric Cylinders

The problem of TM-wave scattering by lossy, inhomogeneous dielectric cylinders was formulated by Richmond [79] using pulse basis functions and point-matching. The MER implementation of this problem is discussed by Sultan and Mittra [13], [58], although no execution time comparisons are provided. Since the off-diagonal matrix elements are identical in form to Equation (4.1), an MER implementation yields relative efficiencies identical to those of the perfectly

TABLE 4.1.

MONOSTATIC RCS FOR CIRCULAR PEC CYLINDER WITH ONE WAVELENGTH CIRCUMFERENCE, TM POLARIZATION

N	RCS (dB λ_0^2)
4	-2.672
8	-2.030
16	-2.082
32	-2.096
64	-2.105
128	-2.109
Exact	-2.113

TABLE 4.2.

CURRENT DENSITY AT CENTER OF SHADOW REGION INDUCED BY PLANE TM WAVE ON CIRCULAR PEC CYLINDER WITH ONE WAVELENGTH CIRCUMFERENCE

N	$\lvert J_z \rvert / \lvert E_z \rvert$	$\angle J_z$
4	0.001109	179.22
8	0.000845	154.16
16	0.000784	152.36
32	0.000773	152.68
64	0.000766	153.00
128	0.000763	153.17
Exact	0.000760	153.35

TABLE 4.3.

CURRENT DENSITY AT THE SPECULAR POINT INDUCED BY PLANE TM WAVE ON CIRCULAR PEC CYLINDER WITH ONE WAVELENGTH CIRCUMFERENCE

N	$\lvert J_z \rvert / \lvert E_z \rvert$	$\angle J_z$
4	0.006273	34.557
8	0.006391	41.261
16	0.006302	41.134
32	0.006271	40.792
64	0.006254	40.567
128	0.006245	40.451
Exact	0.006237	40.335

conducting cylinder case discussed in Section 4.3. Thus, Figure 4.1 suffices to describe the dielectric cylinder problem also.

To evaluate the accuracy of the moment method formulation, numerical results based on several different models are compared to the exact solution for a homogeneous, circular cylinder. The models of the cylinder approximate the circular cross-section by a superposition of square cells, with the areas of the model cross-section normalized to the area of the desired cylinder. Table 4.4 shows the monostatic radar cross section obtained from this process, and Table 4.5 shows values of the electric field at the center of the cylinder. These results suggest the validity of the pulse basis and point matching formulation, although they do not necessarily indicate the level of accuracy that might be obtained for a more complicated geometry, such as one comprised of inhomogeneous material.

4.5. TM-wave Scattering by Cylinders Modeled by a Combination of Dielectric and Conducting Materials

The problem of TM-wave scattering from an infinite cylinder containing both dielectric and perfectly conducting (PEC) materials can be attacked by combining the techniques discussed in the previous two sections. The electric field integral equation may be written as

$$E_z^{inc}(x,y) = E_z^{tot}(x,y) + \frac{k\eta}{4} \int_{p.e.c.} J_z(\ell') \, H_0^{(2)} \left(k\sqrt{(x-x')^2 + (y-y')^2}\right) d\ell'$$

$$+ j\frac{k^2}{4} \iint_{dielectric} (\varepsilon_r - 1) \, E_z^{tot}(x',y') \, H_0^{(2)}\left(k\sqrt{(x-x')^2 + (y-y')^2}\right) dx'dy' \qquad (4.2)$$

where the first integral is to be taken around the outer surface of the perfect conductor and the second integral throughout the cross-sectional volume of the

TABLE 4.4.

MONOSTATIC RCS OBTAINED FOR A HOMOGENEOUS, CIRCULAR DIELECTRIC CYLINDER WITH $\varepsilon_r = 10$, AND CIRCUMFERENCE OF 0.5137 λ_0

N	RCS (dB λ_0^2)
21	-1.8484
61	-1.8469
101	-1.8442
exact	-1.8426

TABLE 4.5.

ELECTRIC FIELD AT THE CENTER OF A HOMOGENEOUS, CIRCULAR CYLINDER WITH $\varepsilon_r = 10$, CIRCUMFERENCE = $0.5137\ \lambda_0$

N	$\lvert E_z \rvert / \lvert E_z^{inc} \rvert$	$\angle E_z$
21	0.770	-94.75
61	0.779	-94.66
101	0.779	-94.75
Exact	0.780	-94.82

dielectric material. Note that E_z vanishes for points (x,y) on the surface of the perfect conductor. Figure 4.2 illustrates the type of geometry under consideration. In Equation (4.2) and throughout this report, k is the free-space wavenumber, η is the intrinsic impedance of free space, and ε_r is the complex relative permittivity of a dielectric material.

A given cylinder can be modeled by a superposition of N_p perfectly conducting strips and N_d homogeneous dielectric cylinder cells. If the current density on each strip and fields throughout each cell are assumed constant, and Equation (4.2) is enforced by point matching in the center of each strip and each cell, the resulting system can be written

$$\begin{bmatrix} \underline{E}^i \end{bmatrix} = \left[\begin{array}{c|c} \underline{\underline{G}}^{dd} & \underline{\underline{G}}^{dp} \\ \hline \underline{\underline{G}}^{pd} & \underline{\underline{G}}^{pp} \end{array}\right] \begin{bmatrix} \underline{U}^d \\ \underline{U}^p \end{bmatrix} \tag{4.3}$$

In Equation (4.3), $\underline{\underline{G}}^{dd}$ denotes an $N_d x N_d$ matrix, $\underline{\underline{G}}^{pp}$ denotes an $N_p x N_p$ matrix, $\underline{\underline{G}}^{dp}$ denotes an $N_d x N_p$ matrix, and $\underline{\underline{G}}^{pd}$ denotes an $N_p x N_d$ matrix. $\underline{U}^d$ and $\underline{U}^p$ denote arrays containing the coefficients of the pulse basis functions used to represent the field and current densities, which are the unknowns to be determined. $\underline{E}^i$ denotes the sampled values of the incident electric field.

If we number the dielectric cells from 1 to N_d and the conducting strips from N_d+1 to N_d+N_p, the elements of the NxN matrix, where $N=N_d+N_p$, are given by

$$G_{mn}^{dd} = \delta_n^m + \frac{jk^2(\varepsilon_{r_n} - 1)}{4} \iint_{\substack{cell \\ n}} H_0^{(2)}\left(k\sqrt{(x_m-x')^2 + (y_m-y')^2}\right) dx'dy \tag{4.4}$$

$$G_{mn}^{pp} = \frac{k\eta}{4} \int_{\substack{strip \\ n}} H_0^{(2)}\left(k\sqrt{(x_m-x')^2 + (y_m-y')^2}\right) d\ell' \tag{4.5}$$

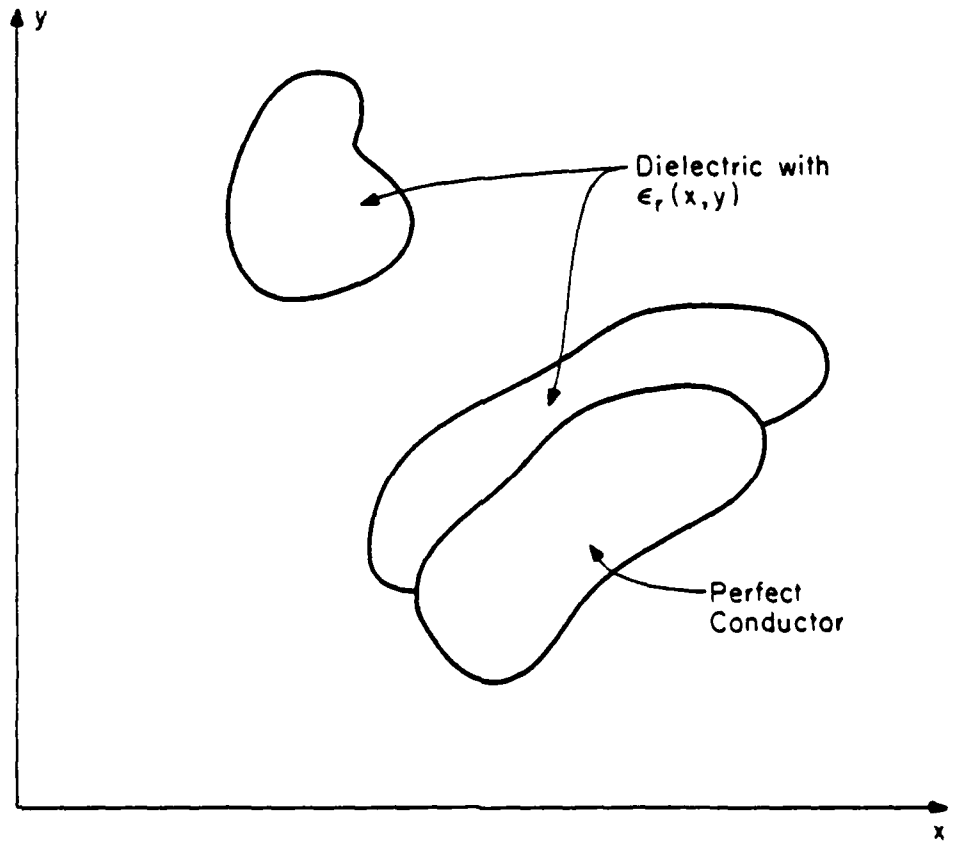

Figure 4.2. Cross-section of the type of cylinder under consideration.

In Equation (4.4), indices m and n range from 1 to N_d. In Equation (4.5), m and n range from N_d+1 to N_d+N_p. G_{mn}^{dp} may be found from Equation (4.5) by treating the index n as ranging from N_d+1 to N_d+N_p, and index m as ranging from 1 to N_d. Similarly, G_{mn}^{pd} may be found from Equation (4.4) with n between 1 and N_d and m between N_d+1 and N_d+N_p. Using the approximations of Richmond [79], the integral in Equation (4.4) can be reduced so that

$$G_{mn}^{dd} = \begin{cases} 1 + (\varepsilon_{r_n} - 1)\left[\frac{j\pi ka_n}{2} H_1^{(2)}(ka_n) + 1\right] & m = n \\ \frac{j\pi ka_n}{2} J_1(ka_n)\, H_0^{(2)}(k\rho_{mn})(\varepsilon_{r_n} - 1) & m \neq n \end{cases} \tag{4.6}$$

where a_n is the radius of a circular cylinder with cross-sectional area equal to that of cell n, and ρ_{mn} is the distance from the center of cell n to the center of cell m. Using a simple approximation suggested by Harrington [74], the integral of Equation (4.5) can be simplified to yield

$$G_{mn}^{pp} = \begin{cases} \frac{\eta k w_n}{4}\left[1 - j\frac{2}{\pi}\ln\left(\frac{kw_n}{6.105}\right)\right] & m = n \\ \frac{\eta k w_n}{4} H_0^{(2)}(k\rho_{mn}) & m \neq n \end{cases} \tag{4.7}$$

where w_n is the width of the n-th strip.

It is not surprising that the off-diagonal matrix elements are of the form of Equation (4.1), since the method is a combination of two approaches which individually satisfy this condition. The accuracy of these two procedures has been discussed in Sections 4.3 and 4.4, and additional convergence tests are not provided. Figures 4.3 and 4.4 illustrate the accuracy of the combined methods for the surface current density and internal fields found for a circular PEC cylinder coated with a uniform dielectric layer.

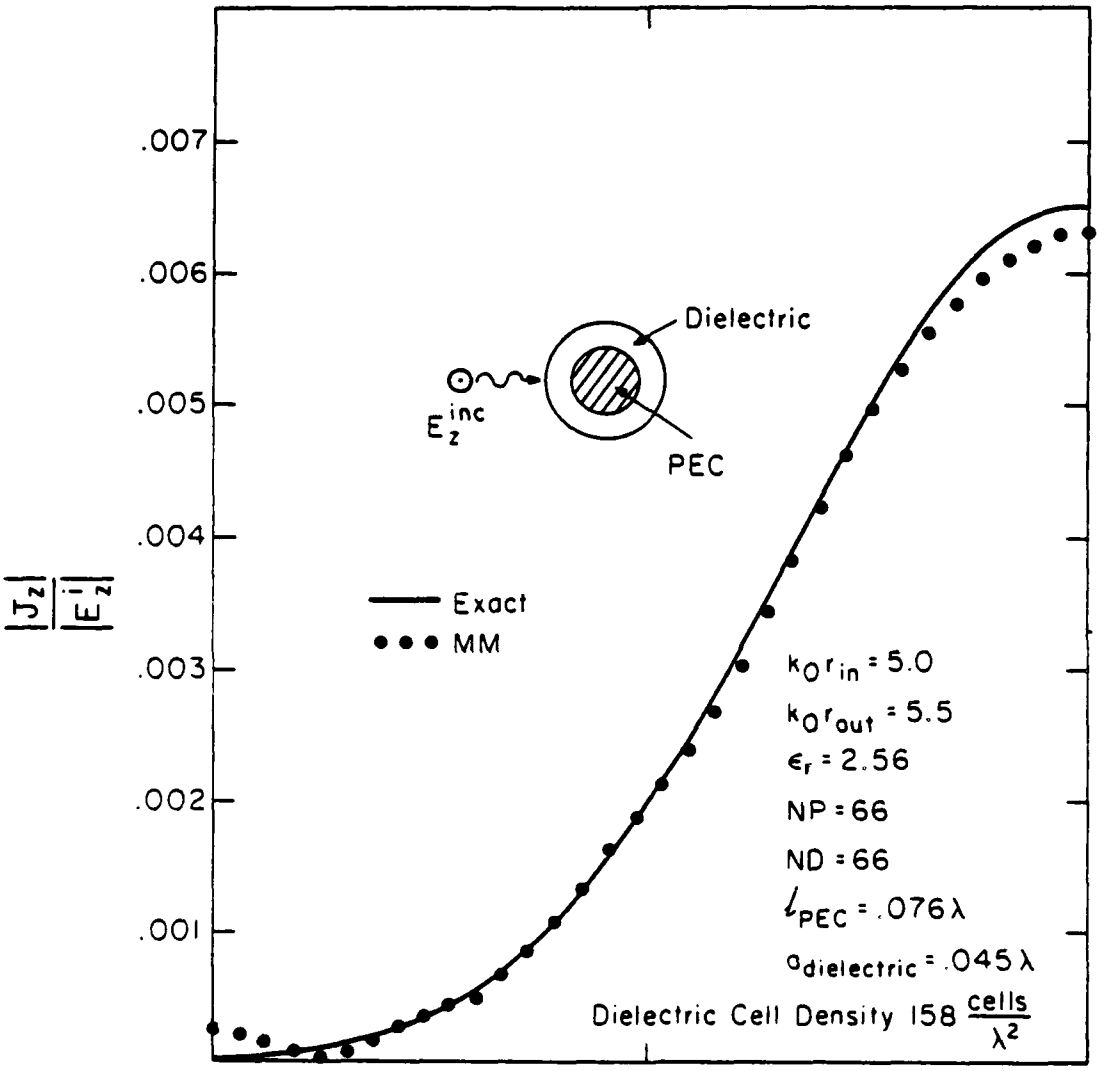

Figure 4.3. Comparison of exact and numerical results for the current density induced on a PEC cylinder with a dielectric coating.

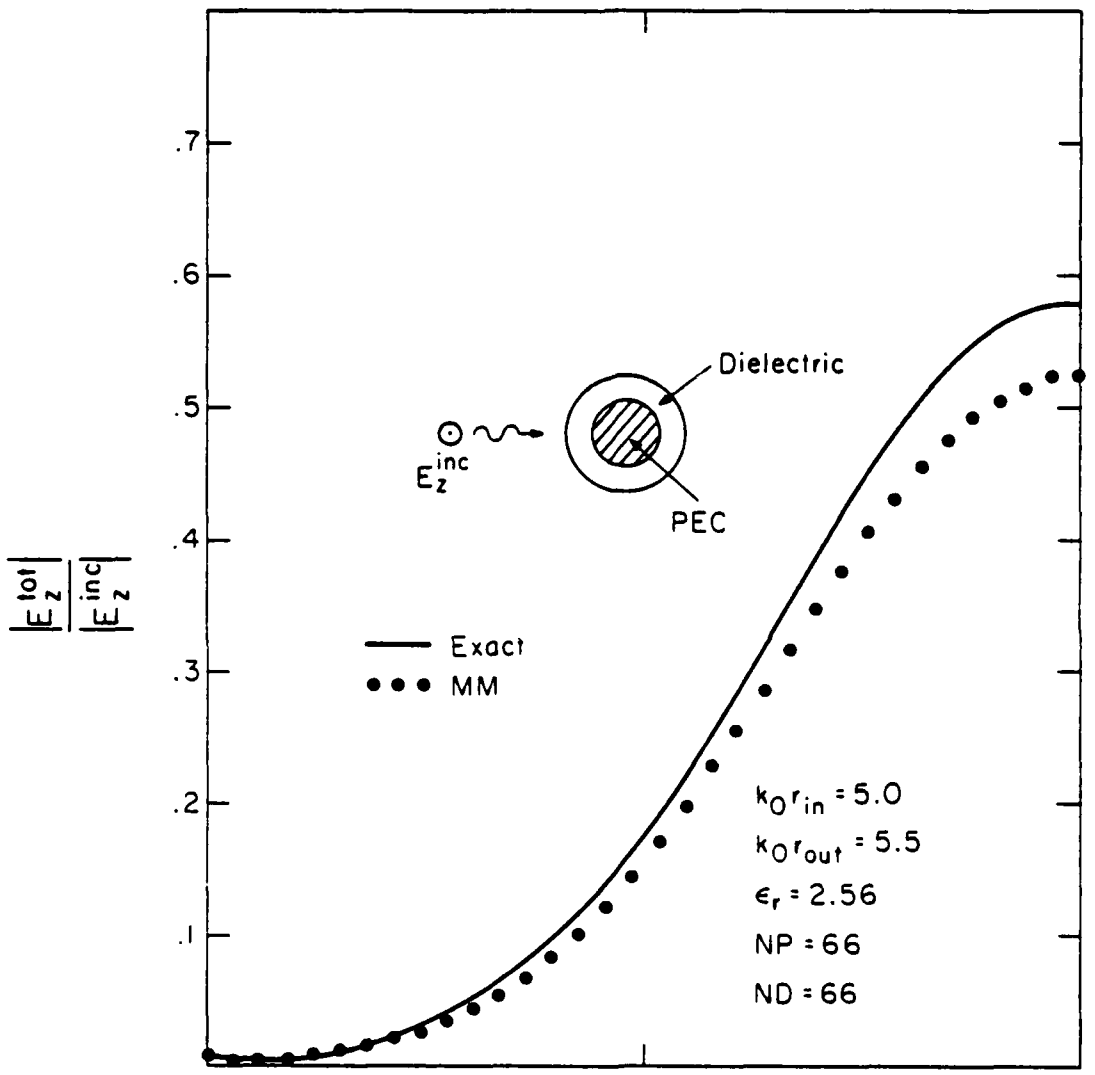

Figure 4.4. Comparison of exact and numerical results for the electric field in the center of a dielectric coating around a PEC cylinder.

The relative efficiency of the MER implementation of the above technique is illustrated by Figure 4.5, which shows execution times for both the MER implementation and one where matrix elements are transferred from a disk storage device. In both cases a 125x125 portion of the matrix was stored in fast-access memory on the CDC CYBER 175, and the remaining matrix elements either generated or read from disk as needed. The system was solved by the iterative CGM algorithm of Chapter 2. Note that the data shown represent the total execution time, which is proportional to the number of iterations required for the iterative algorithm to converge. Identical examples were used for both the MER and the off-core storage data, and should be a valid indication of the relative efficiency. The total execution time for other examples may differ appreciably from these data if a different number of iterations are necessary to produce a solution. From the data, it is clear that the MER approach is the more efficient procedure.

4.6. TE-wave Scattering by Cylinders Modeled by a Combination of Dielectric and Conducting Materials

A numerical solution to the TE-polarization counterpart of the composite dielectric - conducting cylinder problem of Section 4.5 can be based on a variety of mathematical formulations. Anticipating an MER implementation, the matrix elements are to be kept as simple as possible. One of the approaches involves the magnetic field integral equation (MFIE) for the perfect conducting material combined with the electric field integral equation (EFIE) for the dielectric material. An EFIE could be used to treat the conducting material, but would require more complicated matrix elements. Similarly, the MFIE has been applied to the dielectric material for this polarization, but simple discretizations appear to yield more accurate results when used with the EFIE under similar conditions [82].

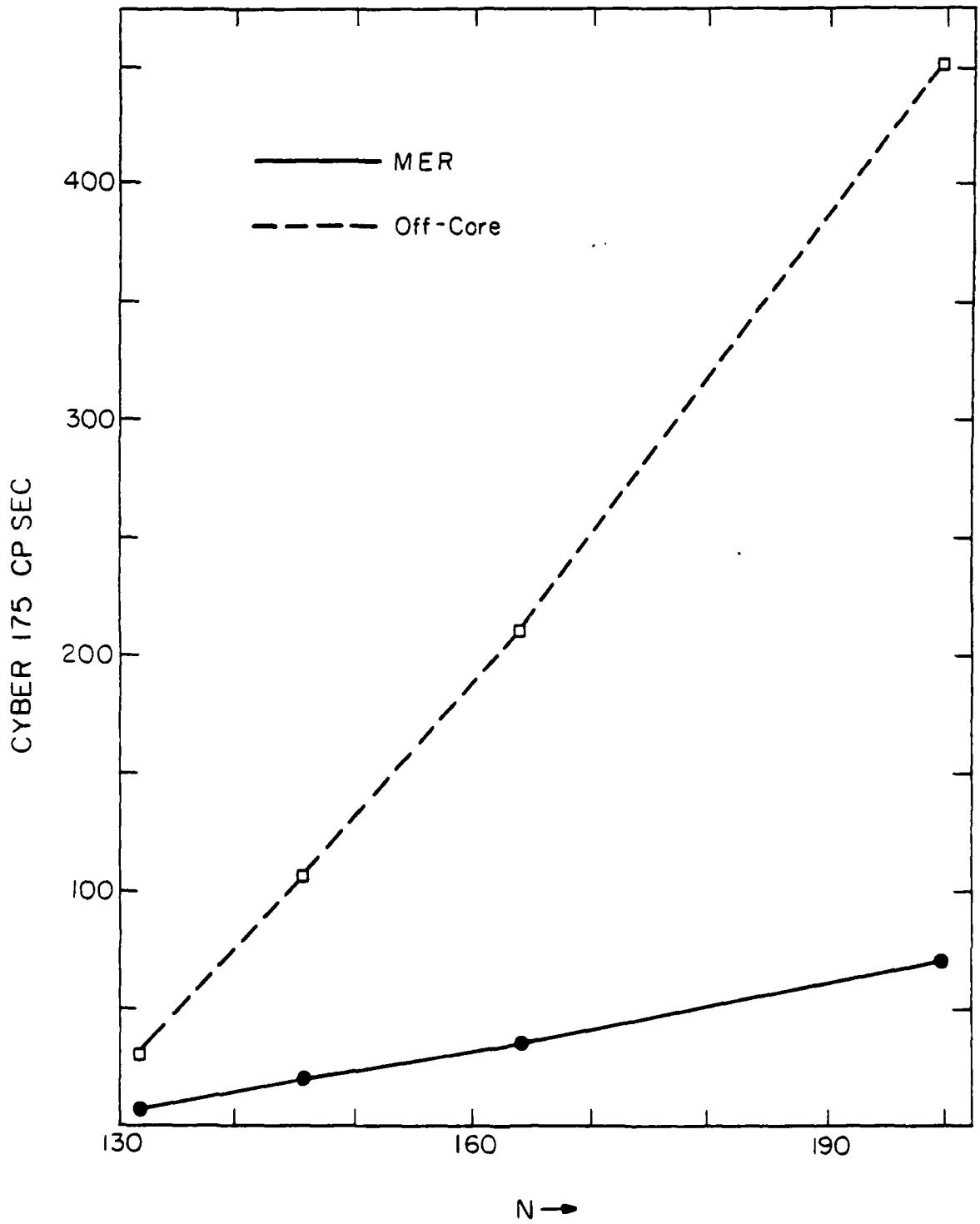

Figure 4.5. Relative execution times for PEC - dielectric cylinders, TM polarization.

The coupled system can be written

$$\left.\begin{aligned} \overline{E}^{inc} &= \overline{E}^{tot} - \frac{\text{curl curl } \overline{A}_d - j\omega\varepsilon_0(\varepsilon_r - 1)\,\overline{E}^{tot} + \text{curl curl } \overline{A}_p}{j\omega\varepsilon_0} \\ H_z^{inc} &= -J_\phi - \hat{z}\cdot\text{curl } \overline{A}_p - \hat{z}\cdot\text{curl } \overline{A}_d \end{aligned}\right\} \tag{4.8}$$

where $\overline{A}$ is the magnetic vector potential defined for the conducting material as

$$\overline{A}_p(x,y) = \int \overline{J}(\ell')\,\frac{1}{4j}\,H_0^{(2)}\,(kR)\,d\ell' \tag{4.9}$$

and for the dielectric material as

$$\overline{A}_d(x,y) = \frac{k}{4\eta}\iint (\varepsilon_r - 1)\,\overline{E}^{tot}\,H_0^{(2)}(kR)\,dx'dy' \tag{4.10}$$

where

$$R = \sqrt{(x - x')^2 + (y - y')^2} \tag{4.11}$$

Note that the EFIE of Equation (4.8) is to be enforced throughout the dielectric material, while the MFIE is to be enforced over the surface of the conducting material.

If the geometry to be analyzed is modeled exactly as in Section 4.5, and expansion functions identical to those of Section 4.5 are used for the unknown quantities, Equation (4.8) can be enforced at the center of each strip and cell to yield a system of the form

$$\begin{bmatrix} \underline{\underline{G}}_{xx}^{dd} & \underline{\underline{G}}_{xy}^{dd} & \underline{\underline{G}}_{x}^{dp} \\ \underline{\underline{G}}_{yx}^{dd} & \underline{\underline{G}}_{yy}^{dd} & \underline{\underline{G}}_{y}^{dp} \\ \underline{\underline{G}}_{x}^{pd} & \underline{\underline{G}}_{y}^{pd} & \underline{\underline{G}}^{pp} \end{bmatrix} \begin{bmatrix} \underline{E}_{x}^{d} \\ \underline{E}_{y}^{d} \\ \underline{J}^{p} \end{bmatrix} = \begin{bmatrix} \underline{E}_{x}^{i} \\ \underline{E}_{y}^{i} \\ \underline{H}_{z}^{i} \end{bmatrix} \tag{4.12}$$

If the model contains N_d dielectric cells and N_p conducting strips, there will be a total of $(2N_d+N_p)$ unknowns. We number the x component of the fields in the dielectric cells from 1 to N_d, the y component of the fields in the dielectric cells from N_d+1 to $2N_d$, and the current density on the conducting strips from $2N_d+1$ to $2N_d+N_p$. The matrix elements G_{dd} have been evaluated approximately by Richmond [80], and are given as

$$G_{xx\ mn}^{dd} = \gamma_n\left[H_0(R)\,\frac{Y^2}{R^2} + H_1(R)\,\frac{X^2 - Y^2}{R^3}\right] \tag{4.13}$$

$$G_{yy\ mn}^{dd} = \gamma_n\left[H_0(R)\,\frac{X^2}{R^2} + H_1(R)\,\frac{Y^2 - X^2}{R^3}\right] \tag{4.14}$$

$$G_{xy\ mn}^{dd} = G_{yx\ mn}^{dd} = \gamma_n[2H_1(R) - RH_0(R)]\,\frac{XY}{R^3} \tag{4.15}$$

for the off-diagonal elements, where

$$\gamma_n = (\varepsilon_{r_n} - 1)\,\frac{j\pi k a_n\, J_1(k a_n)}{2} \tag{4.16}$$

$$X = k(x_m - x_n) \tag{4.17}$$

$$Y = k(y_m - y_n) \tag{4.18}$$

$$R = \sqrt{X^2 + Y^2} \tag{4.19}$$

$$H_0(R) = H_0^{(2)}(R) \tag{4.20}$$

$$H_1(R) = H_1^{(2)}(R) \tag{4.21}$$

In each of the above expressions the indices refer to the proper range of parameters for that sub-matrix. The diagonal elements are given by

$$G^{dd}_{xx\ mm} = G^{dd}_{yy\ mm} = 1 + (\varepsilon_{r_m} - 1)\left[\frac{j\pi ka_m H_1^{(2)}(ka_m)}{4} + 1\right] \tag{4.22}$$

and the diagonals of the sub-matrices $\underline{\underline{G}}^{dd}_{xy}$ and $\underline{\underline{G}}^{dd}_{yx}$ vanish. The elements in the $\underline{\underline{G}}_x{}^{pd}$ and $\underline{\underline{G}}_y{}^{pd}$ sub-matrices can be found by a simple extension of Richmond's formulas [80] to give

$$G^{pd}_{x\ mn} = -(\varepsilon_{r_n} - 1)\frac{\pi ka_n}{2\eta} J_1(ka_n)\, H_1(R)\, \frac{Y}{R} \tag{4.23}$$

$$G^{pd}_{y\ mn} = (\varepsilon_{r_n} - 1)\frac{\pi ka_n}{2\eta} J_1(ka_n)\, H_1(R)\, \frac{X}{R} \tag{4.24}$$

Note that the m-index in Equation (4.23) is to range from $(2N_d+1)$ to $(2N_d+N_p)$ and the n-index from 1 to N_d. The n-index in Equation (4.24) ranges from N_d+1 to $2N_d$.

In order to evaluate the matrix elements for $\underline{\underline{G}}^{pp}$, $\underline{\underline{G}}_x{}^{dp}$, and $\underline{\underline{G}}_y{}^{dp}$, we consider an approach used by Harrington [83]. A single strip is depicted in Figure 4.6.

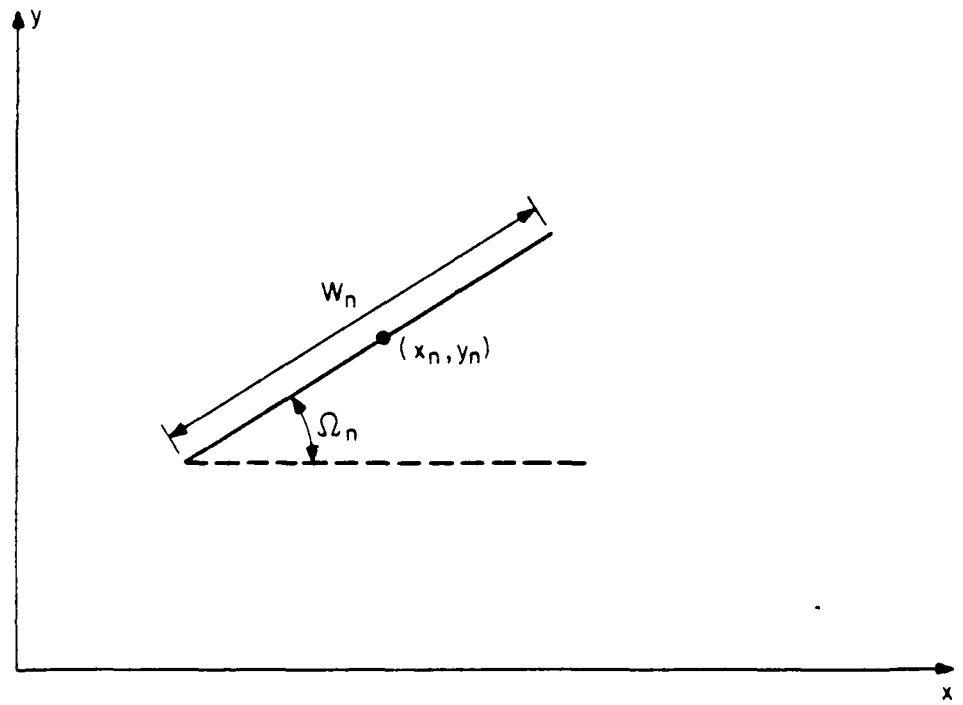

Figure 4.6. Parameters describing the model of a PEC strip for the TE polarization.

Note that we need to introduce an orientation parameter Ω, which was not needed for the treatment of the TM polarization in Section 4.5.

If the strips in the model are flat, the diagonal elements of $\underline{\underline{G}}^{PP}$ are

$$G_{mm}^{PP} = -\frac{1}{2} \tag{4.25}$$

The off-diagonal elements are given by Harrington [83] as

$$G_{mn}^{PP} \approx j\frac{kw_n}{4} H_1(R)\left[\frac{Y}{R}\cos\Omega_n - \frac{X}{R}\sin\Omega_n\right] \tag{4.26}$$

Equation (4.26) is actually an approximation of the expression

$$G_{mn}^{PP} = \frac{j}{4}\int_{-\frac{kw_n}{2}}^{\frac{kw_n}{2}} H_1(\tilde{R})\left\{\frac{\tilde{Y}}{\tilde{R}}\cos\Omega_n - \frac{\tilde{X}}{\tilde{R}}\sin\Omega_n\right\} dL \tag{4.27}$$

where

$$\tilde{Y} = ky_m - ky_n - L\sin\Omega_n \tag{4.28}$$

$$\tilde{X} = kx_m - kx_n - L\cos\Omega_n \tag{4.29}$$

and

$$\tilde{R} = \sqrt{\tilde{X}^2 + \tilde{Y}^2} \tag{4.30}$$

Similarly, we find

$$G^{dp}_{x\,mn} = -\frac{\eta}{4}\int_{-\frac{kw_n}{2}}^{\frac{kw_n}{2}} \left\{\sin \Omega_n \tilde{X}\,\tilde{Y}\,\frac{H_0(\tilde{R})\,\tilde{R} - 2H_1(\tilde{R})}{\tilde{R}^3}\right.$$

$$\left. - \cos \Omega_n\left[H_0(\tilde{R})\left(\frac{\tilde{Y}}{\tilde{R}}\right)^2 + H_1(\tilde{R})\,\frac{\tilde{X}^2 - \tilde{Y}^2}{\tilde{R}^3}\right]\right\} dL \qquad (4.31)$$

$$G^{dp}_{y\,mn} = -\frac{\eta}{4}\int_{-\frac{kw_n}{2}}^{\frac{kw_n}{2}} \left\{\cos \Omega_n \frac{\tilde{X}\tilde{Y}}{\tilde{R}^3}\left[H_0(\tilde{R})\,\tilde{R} - 2H_1(\tilde{R})\right]\right.$$

$$\left. -\sin \Omega_n\left[H_0(\tilde{R})\left(\frac{\tilde{X}}{\tilde{R}}\right)^2 + H_1(\tilde{R})\,\frac{\tilde{Y}^2 - \tilde{X}^2}{\tilde{R}^3}\right]\right\} dL \qquad (4.32)$$

$$G^{dp}_{x\,mn} \simeq -\frac{\eta k w_n}{4}\left\{\sin \Omega_n XY\,\frac{H_0(R)\,R - 2H_1(R)}{R^3}\right.$$

$$\left. - \cos \Omega_n\left[H_0(R)\left(\frac{Y}{R}\right)^2 + H_1(R)\,\frac{X^2 - Y^2}{R^3}\right]\right\} \qquad (4.33)$$

$$G^{dp}_{y\,mn} \simeq -\frac{\eta k w_n}{4}\left\{\cos \Omega_n XY\,\frac{H_0(R)\,R - 2H_1(R)}{R^3}\right.$$

$$\left. - \sin \Omega_n\left[H_0(R)\left(\frac{X}{R}\right)^2 + H_1(R)\,\frac{Y^2 - X^2}{R^3}\right]\right\} \qquad (4.34)$$

The right-hand side of Equation (4.12) represents the sampled incident field at the centers of the cells and strips. The unknown column array contains coefficients of the basis functions (pulses in this case) used to represent the fields in the dielectric and the current density on the conductor.

Due to the relative simplicity of the formulas given in Equations (4.26), (4.33) and (4.34), as compared to the exact expressions which would normally require some form of numerical integration for evaluation, they are the initial choices for MER implementation. If based upon these approximate expressions, the MER approach is more efficient than transferring needed data from disk. As an example, a circular perfectly conducting cylinder coated by a homogeneous dielectric layer was described by a 150x150 matrix equation. If a 125x125 portion of this equation is stored in fast-access memory on the CDC CYBER 175 computer, and the CGM algorithm of Chapter 2 is used to solve the system, the MER approach is approximately three times faster per iteration than reading the necessary matrix elements from disk.

The time necessary to transfer needed data from disk appears to be relatively repeatable, and for the preceding example about 4.3 seconds per iteration were necessary to perform the transfers and other computations. The MER implementation required approximately 1.3 seconds per iteration. For the TM polarization as discussed in Section 4.5, an example involving a 150x150 matrix equation, with a 125x125 part stored in fast access memory, requires about 0.75 second per iteration for MER implementation. Clearly, the efficiency of the MER approach is affected by the additional complexity of the matrix elements for the TE polarization.

Furthermore, experimentation indicates that the approximate formulas are not very accurate when the model under consideration involves close spacings between conducting strips and dielectric cells. Table 4.6 illustrates the accuracy in

TABLE 4.6.

MAGNITUDE OF SURFACE CURRENT DENSITY INDUCED ON PEC CIRCULAR CYLINDER BY PLANE TE-POL WAVE, WHEN THE CYLINDER IS COATED WITH A HOMOGENEOUS LAYER OF DIELECTRIC.
$r_{in} = 0.05\ \lambda_0$, $r_{out} = 0.0875\ \lambda_0$, $\varepsilon_r = 3\text{-}j0$, $N_p = 12$, $N_d = 12$

PEC CELL DENSITY: 38 strips/λ_0
DIELECTRIC CELL DENSITY: 247 cells/λ_d^2

Numerical values produced using the approximate Equations (4.26), (4.33), and (4.34) compared to the values obtained by a numerical integration of Equations (4.27), (4.31), and (4.32). The exact eigenfunction solution is shown for comparison.

ϕ	$\lvert J^{\text{app elem}}\rvert / \lvert H_z^{\text{inc}}\rvert$	$\lvert J^{\text{num int}}\rvert / \lvert H_z^{\text{inc}}\rvert$	$\lvert J^{\text{exact}}\rvert / \lvert H_z^{\text{inc}}\rvert$
0°	1.776	1.114	1.106
30	1.719	1.075	1.064
60	1.629	1.018	1.000
90	1.670	1.059	1.040
120	1.874	1.209	1.198
150	2.094	1.363	1.358
180	2.184	1.425	1.422

the surface current density (magnitude only) for a circular PEC cylinder coated with a uniform, homogeneous layer of dielectric. In this case, large errors occur in the solution due to the use of the approximate formulas of Equations (4.26), (4.33), and (4.34). The accuracy of the approximations can be investigated based on direct numerical calculations, and for illustration Table 4.7 shows the values of Equations (4.33) and (4.31) for a range of observation coordinates. It appears that the approximate expression is reasonably accurate except when the observation point is located within a circle of radius 0.2 λ_0, centered at the source strip. For close spacings, other approximations or some form of numerical integration should be used to evaluate the expressions given in Equations (4.31) and (4.32), to ensure accurate solutions. If implemented carefully, the more accurate expressions should not significantly interfere with the efficiency of the MER approach.

4.7. Summary

This chapter has explored the iterative MER approach for the solution of electromagnetic scattering problems. The MER approach requires a portion of the system matrix to be recomputed whenever needed by the iterative algorithm in use, and has the advantage that the matrix does not need specific symmetries for effective iterative implementation. Because the procedure is not based upon symmetries in the matrix equation, it is not as efficient as the other iterative approaches in use. Experimentation indicates that the MER approach can be more efficient than transferring needed matrix elements from disk whenever needed in the solution process, and thus may be a favorable alternative for the solution of large matrix equations. However, the efficiency of the MER process depends on the simplicity of the matrix elements, and will vary with the problem under consideration.

TABLE 4.7.

COMPARISON OF NUMERICAL VALUES OF EQUATION (4.31) AND EQUATION (4.33) FOR A STRIP WITH ORIENTATION $\Omega = 0$, $w_n = 0.1\ \lambda_0, x_n = 0$, $y_n = 0$.

x_m	y_m	Eq. (4.31)	Eq. (4.33)	% error in Eq. (4.33)
0.1	0	28.03 + j147.8	28.13 + j114.5	22%
0.07	0.07	26.67 - j15.92	26.78 + j8.36	78%
0	0.1	25.21 - j78.46	25.32 - j98.23	24%
0.2	0	24.03 + j29.16	24.10 + j27.31	5%
0.14	0.14	19.12 - j9.252	19.19 - j7.510	8%
0	0.2	13.83 - j41.13	13.91 - j42.77	4%
0.3	0	18.21 + j5.920	18.25 + j5.441	2.5%
0.21	0.21	8.884 - j14.87	8.923 - j14.51	2.1%
0	0.3	-1.087 - j34.22	-1.052 - j34.67	1.3%
0.5	0	5.415 - j6.614	5.358 - j6.751	2%
0.35	0.35	-8.707 - j10.10	-8.725 - j10.05	0.5%
0	0.5	-23.28 - j12.53	-23.33 - j12.67	0.6%
1	0	-1.987 + j2.201	-1.999 + j2.250	1.7%
0.7	0.7	6.035 + j7.199	6.100 + j7.210	0.7%
0	1	15.04 + j11.25	15.02 + j11.30	0.3%

Specific examples of cylinders containing both perfectly conducting and dielectric materials were studied for the MER approach. The TM polarization involves relatively simple off-diagonal matrix elements, and appears to be a good candidate for the MER approach. The TE polarization requires more complicated matrix elements, and the numerical efficiency is not as good as that of the TM polarization. Furthermore, for the specific formulation used with the TE polarization, care must be taken when evaluating the matrix elements since simple approximations are poor if the model requires close spacings between dielectric and conducting cells. Although more accurate evaluation of the matrix elements (when required by close spacings) reduces the efficiency of the MER procedure, the approach is still a viable alternative to storing the needed matrix elements on disk.

Since the efficiency of the MER approach depends on the complexity of the matrix elements, candidates for MER implementation must be evaluated on an individual basis. Clearly, the complexity of some of the state-of-the-art techniques for the numerical solution of scattering problems might preclude their use in the MER procedure. However, simple approaches which offer reasonable accuracy should be good candidates for MER solution. The MER approach should be considered an alternative for the treatment of large scatterers that do not possess the necessary symmetries to be treated efficiently by other techniques.

5. A COMPARISON OF TWO PROCEDURES FOR THE DISCRETIZATION OF CONVOLUTIONAL INTEGRAL EQUATIONS

5.1. Introduction

Electromagnetic scattering problems can often be described by convolutional integral equations. These are equations having the form

$$E(x) = \int_b^a J(x')\, K(x - x')\, dx' \qquad a < x < b \tag{5.1}$$

where $J(x)$ is the unknown function to be determined. An approximate solution for $J(x)$ can be obtained by replacing Equation (5.1) by a finite-dimensional discrete system. If the convolutional form is preserved by the discretization process, the discrete system can be expressed

$$e_m = \sum_{n=1}^{N} j_n\, g_{m-n} \qquad m = 1,2,\ldots,N \tag{5.2}$$

We adopt the notation of using lower-case letters to denote sequences and upper-case to denote functions. The discrete system of Equation (5.2) may be written as an N-th order matrix equation

$$\underline{e} = \underline{\underline{g}}\ \underline{j} \tag{5.3}$$

where

$$\underline{\underline{g}} = \begin{bmatrix} g_0 & g_{-1} & g_{-2} & & \\ g_1 & g_0 & g_{-1} & \cdot\ \cdot\ \cdot & \\ g_2 & g_1 & g_0 & & \\ & \cdot & & & \\ & \cdot & & & \\ & \cdot & & & \end{bmatrix} \tag{5.4}$$

Note that the NxN matrix is Toeplitz, so that each element is a repetition of the first row or column according to Equation (5.4). Thus, there is a considerable amount of redundancy present in the system. Because of this redundancy, a Toeplitz equation is one example of the type of system for which iterative solutions are well suited. It is not a particularly good example, however, because direct methods for solving Toeplitz systems are also available and may be more efficient than iterative algorithms [54], [55], [84] - [86]. In practice, convolutional integral equations representing electromagnetic scattering problems can often be converted to discrete systems containing Toeplitz or almost-Toeplitz symmetries. An example of an almost-Toeplitz system is a matrix with Toeplitz symmetries everywhere except along the main diagonal, as might arise in connection with integral equations describing an inhomogeneous geometry. These are not usually well-suited for direct solution, but are easily treated efficiently using iterative algorithms. Additional examples of the iterative solution of almost-Toeplitz systems are presented in Chapters 6 and 7.

Two types of discretization procedure have been used in the recent past to convert convolutional integral equations to matrix equations with discrete-convolutional symmetries. Because of the utility of these procedures in connection with iterative solution algorithms, a firm understanding of their implementation is of central importance. The first technique to be considered is the discrete-convolutional method of moments (DCMoM), which is a special form of the general moment method procedure [12]. The second is the spectral-domain fast-Fourier transform (SDFFT) method, which is the name we choose to denote the discretization used in connection with the original "spectral-iterative" technique (SIT) [11], [44] - [46]. In contrast with the DCMoM, which is well understood after years of research on the method of moments, the SDFFT discretization has not received much exposition to date in the literature. This chapter

presents the SDFFT procedure in a general form, so that an equivalence can be made between it and the DCMoM approach. It will be shown that the two procedures can be identical in principle, although in practice the SDFFT can only be considered an approximation to the DCMoM approach. Guidelines for using the SDFFT discretization are given.

5.2. The Discrete-convolutional Method of Moments (DCMoM) Procedure

Consider a convolutional integral equation of the form of Equation (5.1). E and K are known over the interval of interest and J is an unknown function to be determined. Equation (5.1) can be used to describe scattering from a strip or wire of constant curvature, and is representative of a variety of other electromagnetic scattering problems. A discretization of Equation (5.1) according to the moment-method procedure requires that J be replaced by a finite expansion of the form

$$J(x) \approx \sum_{n=1}^{N} j_n B_n(x) \tag{5.5}$$

where the $\{B_n(x)\}$ are known basis functions and the j_n unknown coefficients. If the expansion is substituted into Equation (5.1) and the resulting equation is made orthogonal to N independent testing functions $\{T_m(x)\}$, the result is a matrix equation of the form

$$e_m = \sum_{n=1}^{N} j_n \, g_{m,n} \qquad m = 1,2,\ldots,N \tag{5.6}$$

where

$$e_m = \int_a^b T_m(x) \, E(x) \, dx \tag{5.7}$$

$$g_{m,n} = \int_a^b T_m(x) \int_a^b B_n(x') \, K(x - x') \, dx'dx \tag{5.8}$$

In the general case, g_{mn} represents a fully-populated matrix whose NxN entries satisfy no symmetry or redundancy condition.

If the choice of basis and testing functions is restricted to the form

$$B_n(x) = B(x - x_n) \tag{5.9}$$

$$T_m(x) = T(x - x_m) \tag{5.10}$$

where

$$x_n = x_0 + n\Delta x \qquad n = 1,2,\ldots,N \tag{5.11}$$

and if the basis and testing functions do not overlap the endpoints of the interval (a,b), the discrete system described in Equation (5.6) can be written as

$$e_m = \sum_{n=1}^{N} j_n g_{m-n} \qquad n = 1,2,\ldots,N \tag{5.12}$$

The "g" appearing in Equation (5.12) is completely described by only (2N-1) entries of the NxN system, and often symmetry considerations reduce the number of independent entries to N. This system is exactly the Toeplitz form described in Equations (5.2) - (5.4). The moment method application embodied in Equations (5.9) - (5.12) is denoted the discrete-convolutional method of moments (DCMoM), because the summation appearing in Equation (5.12) is a discrete convolution. Examples of the DCMOM are given in Chapters 6 and 7, and elsewhere in the literature [12], [47] - [51].

5.3. The Spectral-domain Fast-Fourier Transform (SDFFT) Method

The DCMoM procedure is one type of discretization which is readily compatible with iterative solution algorithms. A different approach, applicable to convolutional integral equations, requires the Green's function K appearing in Equation (5.1) to be sampled in the Fourier transform or "spectral" domain. In a multidimensional case, the Fourier transform is to be taken with respect to one or more of the spatial variables. The motivation for this alternate approach stems from the fact that often the Fourier transform of K is much simpler and easier to compute numerically than K itself. For instance, problems involving planar, stratified media give rise to spatial Green's functions in terms of infinite integrals [87] or infinite summations [88]. If written in the transform domain, these Green's functions usually become algebraic expressions [89]. Unfortunately, in practice many of the integral equations of interest are only valid over a region of finite support, and cannot be transformed entirely into the spectral domain. Instead, the fast-Fourier transform algorithm is used to connect the two domains. Although in the past the procedure was developed from a different perspective [44], it can be considered as nothing more than an alternate way to produce the matrix "g" employed in Equation (5.12), by applying the inverse FFT to the sequence obtained by sampling the analytical Fourier transform of K(x). We denote this approach the spectral-domain fast-Fourier transform (SDFFT) discretization.

Because the SDFFT approach requires a mixture of analytical Fourier transform techniques and numerical applications of the FFT algorithm, it is important to note the differences beteeen these two tools. A detailed discussion may be found in Brigham [42]. Specifically, the FFT is equivalent to the Fourier transform only when the latter is applied to functions which are discrete and periodic in both the original and transform domains. As an aside, for this

reason the SDFFT procedure appears well-suited for the discretization of equations which are already periodic in one or more spatial variables. Some examples of periodic equations in electromagnetics arise in the analysis of antenna arrays [90] and frequency selective surfaces [91]. A discussion of the SDFFT applied to periodic problems is reserved for a later section of this chapter. The topic of interest here is the application of the SDFFT to non-periodic problems, for it is in these cases that additional care must be taken to ensure the accuracy of the process.

In order to study the SDFFT approach in detail, we will construct the type of function for which the FFT and Fourier transforms are equivalent, and describe the method using the analytical Fourier transform. The Fourier transform is defined

$$F\{H(x)\} = \tilde{H}(f) = \int_{-\infty}^{\infty} H(x)\, e^{-j2\pi fx}\, dx \tag{5.13}$$

and the inverse transform as

$$F^{-1}\{\tilde{H}(f)\} = H(x) = \int_{-\infty}^{\infty} \tilde{H}(f)\, e^{j2\pi fx}\, df \tag{5.14}$$

We adopt the conventional practice of denoting convolution as

$$A(x) * B(x) = \int_{-\infty}^{\infty} A(x')\, B(x - x')\, dx' \tag{5.15}$$

Since the FFT is equivalent to the Fourier transform of a discrete, periodic function, it is necessary to convert continuous aperiodic quantities to discrete periodic quantities. Consider the functions

$$S(x) = \sum_{m=-\infty}^{\infty} \delta(x - m\Delta x) \tag{5.16}$$

$$P(x) = \sum_{q=-\infty}^{\infty} \delta(x - q\Delta X) \tag{5.17}$$

Multiplication with S(x) is equivalent to sampling at intervals of Δx; convolution with P(x) produces a periodic function with period ΔX. The Fourier transforms of the above are

$$\tilde{S}(f) = \Delta F \sum_{m=-\infty}^{\infty} \delta(f - m\Delta F) \tag{5.18}$$

$$\tilde{P}(f) = \Delta f \sum_{q=-\infty}^{\infty} \delta(f - q\Delta f) \tag{5.19}$$

where

$$\Delta F = \frac{1}{\Delta x} \tag{5.20}$$

$$\Delta f = \frac{1}{\Delta X} \tag{5.21}$$

In the transform domain, convolution with $\tilde{S}(f)$ produces a periodic function with period ΔF; multiplication with $\tilde{P}(f)$ produces a discrete function sampled at intervals Δf. In practice, the periods and sampling intervals are related by an integer M, so that

$$\Delta X = M\ \Delta x \tag{5.22}$$

$$\Delta F = M\ \Delta f \tag{5.23}$$

Equations (5.15) - (5.19) will be used to convert functions to discrete, periodic sequences in order to model the FFT algorithm.

According to the discussion of the SIT in the literature [11], the SDFFT procedure involves sampling the transform $\tilde{K}(f)$ in such a manner as to create a

discrete, periodic "spectral Green's function" of the form

$$\tilde{G}_1(f) = \tilde{S}(f) * [\tilde{P}(f)\ \tilde{W}(f)\ \tilde{K}(f)] \tag{5.24}$$

where $\tilde{K}(f)$ represents the analytical Fourier transform of $K(x)$, $\tilde{S}(f)$ and $\tilde{P}(f)$ are defined above, and $\tilde{W}(f)$ is a windowing function introduced to truncate the support of $\tilde{K}(f)$, typically to one period. This construction allows us to immediately transform the discrete spectral Green's function to the spatial domain, and make a comparison with the analogous quantity arising from the DCMoM procedure. The discrete spatial Green's function obtained from Equation (5.24) is given by

$$G_1(x) = P(x) * [S(x)\{W(x) * K(x)\}] \tag{5.25}$$

An alternate way of denoting this is

$$g^{(1)}_{\ell-m} = \sum_{q=-\infty}^{\infty} W(x) * K(x)\Big|_{x\ =\ (\ell-m)\Delta x - q\Delta X} \tag{5.26}$$

If the DCMoM process is generalized to produce an infinite-periodic sequence that coincides with the previous DCMoM sequence "g" over the interval of interest in the spatial domain, a similar procedure can be used to produce the discrete Green's function

$$G_2(x) = P(x) * [S(x)\ U(x)\{T(-x) * B(x) * K(x)\}] \tag{5.27}$$

or, equivalently,

$$g^{(2)}_{\ell-m} = T(-x) * B(x) * K(x)\Big|_{x = (\ell-m)\Delta x} \tag{5.28}$$

B(x) is the basis function introduced in Equation (5.9), and T(-x) is a space-reversal of the testing function T(x) appearing in Equation (5.10). The notable difference between the functional form of Equations (5.25) and (5.27) is the appearance of U(x) in the DCMoM function. U(x) is necessary to truncate the spatial kernel K(x) to the period in order to avoid aliasing errors. For instance, U(x) may be of the form

$$U(x) = \begin{cases} 1 & x \in (a,b) \\ 0 & \text{otherwise} \end{cases} \tag{5.29}$$

Of course, the period may be larger than the interval (a,b), and U(x) may vary accordingly. The aliasing errors due to the absence of U(x) are clearly illustrated by the infinite summation in Equation (5.26).

In view of the above comparison, it appears that the SDFFT process should be generalized to incorporate a function corresponding to the U(x) used with the DCMoM technique. However, $\tilde{U}(f)$ appears within a convolution in the spectral domain, and the $\tilde{U}(f)$ corresponding to Equation (5.29) is a so-called sinc function, with support over the entire x-axis. Because of this, in general it is difficult to include the convolution with $\tilde{U}(f)$ in a numerical implementation.

There appear to be two ways in which the effects of U(x) could be included approximately in the SDFFT procedure. The first is simply to extend the period to some large interval, and approximate the transform $\tilde{U}(f)$ by a Dirac delta function (which it approaches as the period becomes sufficiently large). This is the technique used in the literature [11], [13], [44] - [46]. (Note that a given function is not altered after convolution with a delta function, and thus

Equation (5.25) suffices to describe the process.) An alternate approach is to replace the rectangular truncation function of Equation (5.29) with a smoother function, in order to obtain a $\tilde{U}(f)$ with finite support (at least approximately). The smoother U(x) must be sufficiently flat over the spatial interval of interest, in order to avoid distorting the desired spatial Green's function, yet yield a transform which can be conveniently included in the convolution operation of the discrete spectral Green's function

$$\tilde{G}_1(f) = \tilde{S}(f) * [\tilde{U}(f) * \tilde{P}(f)\ \tilde{W}(f)\ \tilde{K}(f)] \tag{5.30}$$

The first approximate method for including U(x) has the advantage that the convolution with $\tilde{U}(f)$ disappears from Equation (5.30), simplifying the calculation of $\tilde{G}_1(f)$. However, the period must be significantly larger than would otherwise be the case, in order to ensure that the summation of Equation (5.26) is an adequate approximation to the desired result (the q=0 term alone). Examples to follow will attempt to determine the size period necessary for reasonable accuracy. It is important to note that the large period is only necessary for the initial construction of $\tilde{G}(f)$; once a suitable discrete spectral Green's function is obtained it can be transformed (via the FFT) to the spatial domain and truncated to the interval of interest. The second approach has the drawback that additional computation will be necessary to include the effects of the approximate U(x), assuming such a function can be found. An advantage of the second approach is that a singularity often present in $\tilde{K}(f)$ is explicitly smoothed by convolution with $\tilde{U}(f)$. Ray and Mittra have attributed irregular numerical results to improper treatment of the singularity in $\tilde{K}(f)$ [92].

It is clear from a comparison of Equations (5.26) and (5.28) that the function obtained by convolving the basis function B(x) and the space-reversed testing function T(-x) in the DCMoM process plays an analogous role to the inverse transform of the windowing function $\tilde{W}(f)$ used in the SDFFT procedure. This suggests that the SDFFT is equivalent to the DCMoM, if the latter uses basis functions W(x) and Dirac delta testing functions. This fact was recently noted by Nyo and Harrington [47]. Previous presentations of the SDFFT approach in the context of SIT have omitted the point that the spatial-domain basis function is implicitly chosen as the inverse transform of the windowing function. Of course, this equivalency only holds if the aliasing effects of the summation in Equation (5.26) are negligible.

Because the windowing function $\tilde{W}(f)$ appears as a multiplication with $\tilde{K}(f)$, the SDFFT approach can easily incorporate a variety of windowing functions. Based upon the comparison with the DCMoM, it appears that a primary consideration for the choice of $\tilde{W}(f)$ should be the corresponding spatial domain basis function selected implicitly in the process. For instance, the choice of a rectangular window for $\tilde{W}(f)$ corresponds to the implicit choice of a sinc function for the basis function. Since sinc functions have unbounded support, they do not appear to be appropriate approximations to subsectional basis functions, and will apparently have considerable support outside the original domain of interest (i.e., outside the original scatterer). Typical subsectional basis functions such as a piecewise constant or a triangle function thus correspond to windowing functions $\tilde{W}(f)$ with unbounded support, which seems to suggest that the "proper" windowing function to use with the SDFFT is one which allows considerable aliasing in the spectral domain. Thus, the incorporation of a windowing function that corresponds to a subsectional basis function may be

complicated by the need to deliberately overlap many periods of the function $\tilde{K}(f)$ when constructing $\tilde{G}_I(f)$.

In practice, the selection of $\tilde{W}(f)$ may be thought of as a scheme to produce a reasonable approximation to a subsectional basis function and a reasonable approximation to a windowing function with finite support. For instance, Figures 5.1 and 5.2 show rectangular and exponential windowing functions, and their inverse transforms. The exponential window might be an improvement over the rectangular window because the basis function associated with the exponential window has its support reasonably confined, yet is similar to that produced by the rectangular window over the desired interval. Furthermore, the exponential window only requires approximately two periods to overlap in the calculation of $\tilde{G}_I(f)$.

By analogy with the DCMoM procedure, it is obvious that a testing function could be incorporated into the SDFFT process, as may be necessary if the excitation in a given problem is highly localized. In principle, the choice of $\tilde{W}(f)$ can be made to correspond to both a basis and a testing function, and the excitation sequence "e" can be computed according to Equation (5.7).

5.4. Iterative Implementation of the DCMoM and SDFFT Systems

From the above discussion, it is apparent that the DCMoM system and the SDFFT system can both be represented by Equation (5.12), with the upper limit of the summation limited to N. In this case, N represents the number of points lying in the original domain (a,b). Note that there are 2N-1 pertinent values of "g" appearing in Equation (5.12). Thus, both the SDFFT and DCMoM systems can be implemented in exactly the same manner. While this implementation can be accomplished in many ways, for instance by solving an N-th order matrix equation, our interest here is centered on the iterative algorithms of Chapter 2.

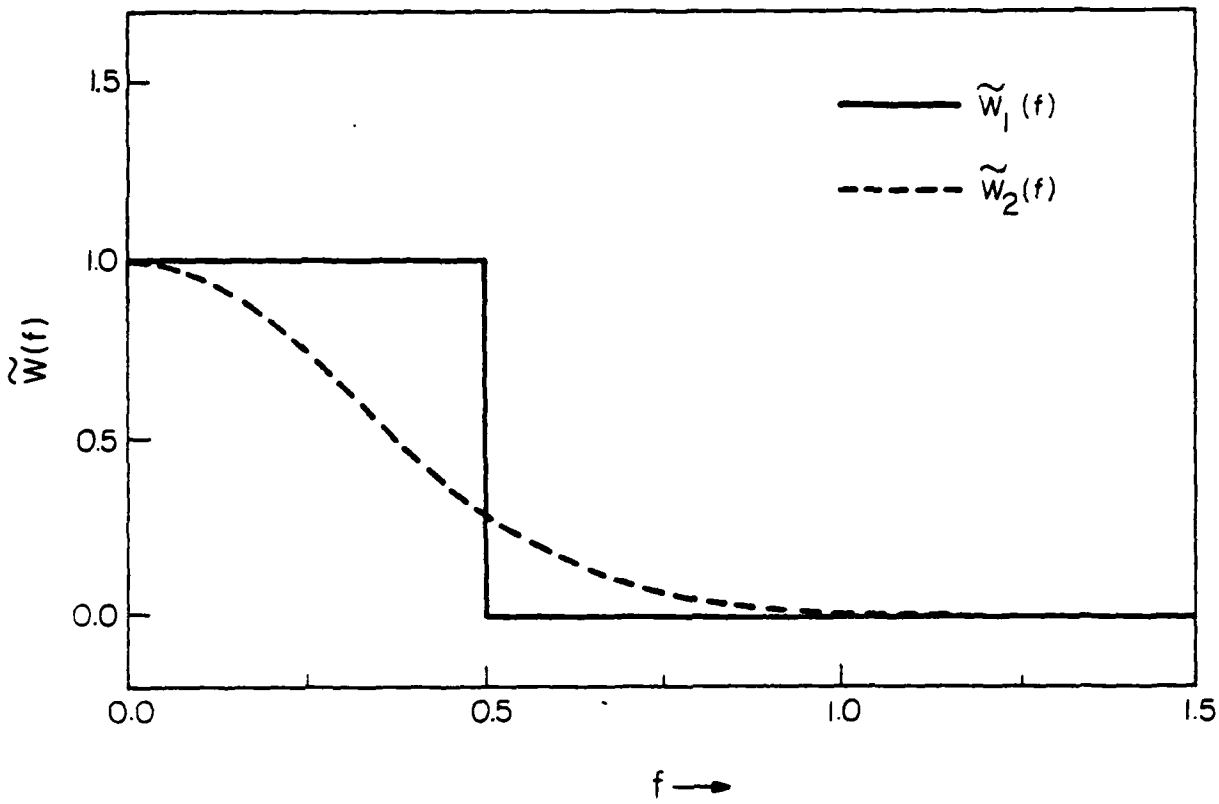

Figure 5.1. Comparison of the windowing functions

$$\tilde{W}_1(f) = \begin{cases} 1 & |f| < 0.5 \\ 0 & \text{otherwise} \end{cases}$$

and

$$\tilde{W}_2(f) = \exp(-\{0.7\ \pi f\}^2).$$

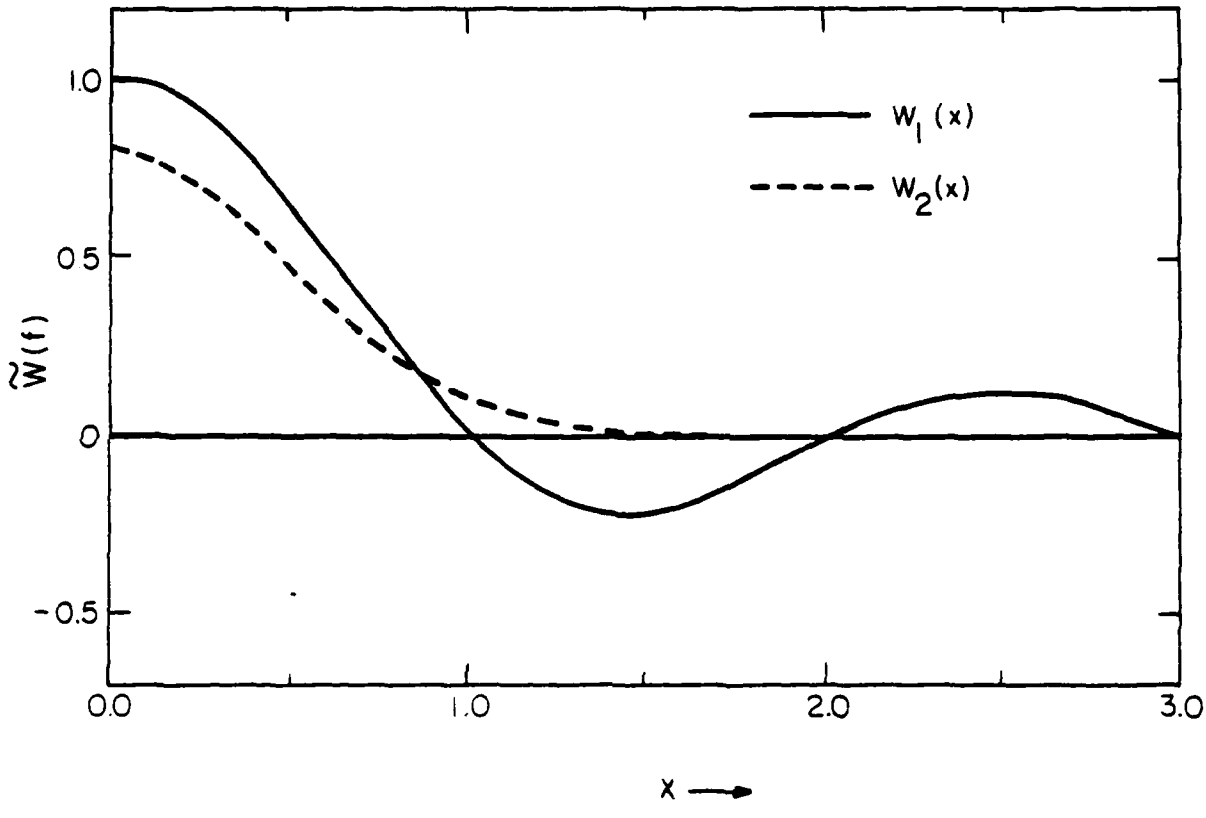

Figure 5.2. Comparison of the implicit basis functions

$$W_1(x) = \sin(\pi x)/(\pi x)$$

and

$$W_2(x) = \frac{1}{0.7\sqrt{\pi}} \exp\left(-\left(\frac{x}{0.7}\right)^2\right) .$$

These algorithms require successive applications of the operator (NxN matrix) described in Equation (5.12).

The operator can be applied in several ways. One approach is to perform the matrix multiplication explicitly whenever needed, which does not imply that the full NxN matrix must be stored in computer memory. A popular alternative is to make use of the discrete convolution theorem [42], and use the FFT algorithm to perform the convolution of Equation (5.12), according to

$$\sum_{n=1}^{N} j_n g_{m-n} = \sum_{n=1}^{2N-1} j_n g_{m-n} \tag{5.31}$$

$$= FFT^{-1}\{\tilde{j}\ \tilde{g}\} \tag{5.32}$$

where "j" must be zero-padded according to

$$j_n = 0 \qquad n = N+1, N+2, \ldots, 2N-1 \tag{5.33}$$

This approach typically requires 2Mlog(M) operations, where M = 2N-1 and the logarithm is with respect to base 2, as opposed to N^2 if the operator was implemented as a matrix multiplication. Thus, there are fewer computations required using the FFT assisted approach for a one-dimensional system as long as N is greater than 20. Two-dimensional systems of the form

$$\sum_{n=1}^{N} \sum_{m=1}^{M} j_{nm}\, g_{k-n,\ell-m} \tag{5.34}$$

can be treated in a similar manner, as can three-dimensional equations. Note that for two-dimensional systems, the total number of unknowns must be approximately 125 before the FFT-assisted approach exceeds the efficiency of direct matrix multiplication. Because of the zero-padding requirements, the FFT

approach requires larger array sizes, which also affects the trade-off between the two approaches.

5.5. Example: TM-wave Scattering by Strips

As an example of the implementation of the DCMoM and SDFFT procedures, and a means to compare the two to develop guidelines for the use of the latter approach, consider the problem of TM-wave scattering by a perfectly conducting flat strip. The integral equation for a one wavelength strip has the form

$$E(x) = \int_{-0.05}^{0.95} J(x')\, K(x - x')\, dx' \qquad -0.05 < x < 0.95 \tag{5.35}$$

The kernel in this case is given by

$$K(x) = \frac{1}{j4} H_0^{(2)} (2\pi|x|) \tag{5.36}$$

and its Fourier transform is

$$\tilde{K}(f) = \begin{cases} \dfrac{1}{j4\pi\sqrt{1 - f^2}} & |f| < 1 \\ \dfrac{1}{4\pi\sqrt{f^2 - 1}} & |f| > 1 \end{cases} \tag{5.37}$$

If ten basis functions are used with the DCMoM procedure, specifically piecewise constant or "pulse" functions defined by

$$B(x) = \begin{cases} 1 & x \in (-0.05,\ 0.05) \\ 0 & \text{otherwise} \end{cases} \tag{5.38}$$

and if the testing functions are Dirac delta functions

$$T(x) = \delta(x) \tag{5.39}$$

where

$$x_n = -0.1 + n(0.1) \qquad n = 1,2,\ldots,10 \tag{5.40}$$

then the numerical values of the spatial domain sequence representing "g" are given in Table 5.1. This sequence was obtained from Equation (5.8), and is equivalent to the values of the discrete spatial Green's function $G_2(x)$ described in Equations (5.27) or (5.28).

Consider an SDFFT discretization of Equation (5.35), assuming that the function $\tilde{U}(f)$ is approximated by a Dirac delta function as discussed in Section 5.3. Initially, a rectangular window is used for $\tilde{W}(f)$. The objective of this example is to determine the value of M from Equation (5.22) that produces a reasonably accurate discrete spatial Green's function over the interval (-0.05,0.95), if it is possible to do so and assuming that the values produced by the DCMoM are interpreted as the "correct" result. Note that the data from Table 5.1 are not the numerical values which would be produced by the SDFFT procedure even if an infinite amount of zero-padding was incorporated into the process, because $B(x)$ and $\tilde{W}(f)$ are not a transform pair.

Tables 5.2, 5.3, and 5.4 show values of the sequence "g" produced by the SDFFT procedure for values of M of 99, 255, and 1023. These tables also show the relative difference between the data of Table 5.1 and the SDFFT data. In spite of the fact that we do not expect perfect agreement, it appears that the numerical values are similar provided that M is large. From a study using a variety of strip sizes, it appears that the equivalent spatial period necessary for agreement between the DCMoM sequence and the SDFFT sequence must exceed 25 wavelengths in order to obtain agreement within five percent in the first few

TABLE 5.1

DISCRETE SPATIAL DOMAIN SEQUENCE PRODUCED BY DCMOM WITH PULSE BASIS FUNCTIONS AND DIRAC DELTA TESTING FUNCTIONS FOR A STRIP OF LENGTH 1 λ_0 WITH 10 CELLS.

n	$\lvert g_n \rvert$	$\angle g_n$ (degrees)
0	0.0440	-34.64
1	0.0237	-71.33
2	0.0171	-111.38
3	0.0141	-149.03
4	0.0123	174.06
5	0.0110	137.49
6	0.0101	101.10
7	0.0093	64.81
8	0.0087	28.59
9	0.0082	-7.58

TABLE 5.2

DISCRETE SPATIAL DOMAIN SEQUENCE PRODUCED BY SDFFT USING A RECTANGULAR WINDOW W(f) FOR M=99.

n	$\lvert g_n \rvert$	$\angle g_n$ (degrees)	% diff. as compared to Table 5.1
0	0.0469	-26.65	16 %
1	0.0226	-56.75	25 %
2	0.0159	-110.65	7 %
3	0.0170	-146.70	21 %
4	0.0172	-168.59	54 %
5	0.0130	171.04	65 %
6	0.0078	130.44	50 %
7	0.0087	64.03	7 %
8	0.0125	35.37	45 %
9	0.0130	20.48	84 %

TABLE 5.3

DISCRETE SPATIAL DOMAIN SEQUENCE PRODUCED BY THE SDFFT USING A RECTANGULAR WINDOW W(f) FOR M=255.

n	$\lvert g_n \rvert$	$\angle g_n$ (degrees)	% diff. as compared to Table 5.1
0	0.0425	-34.35	3.5 %
1	0.0227	-72.14	4.2 %
2	0.0174	-114.94	6.4 %
3	0.0140	-146.48	4.5 %
4	0.0113	178.67	11 %
5	0.0092	136.67	16 %
6	0.0091	95.37	13 %
7	0.0094	61.78	5.4 %
8	0.0086	32.21	8.1 %
9	0.0070	-0.69	18 %

TABLE 5.4

DISCRETE SPATIAL SEQUENCE PRODUCED BY THE SDFFT USING A RECTANGULAR WINDOW W(f) WITH M=1023.

n	$\lvert g_n \rvert$	$\angle g_n$ (degrees)	% diff. as compared to Table 5.1
0	0.0433	-35.37	1.9 %
1	0.0237	-72.47	2.0 %
2	0.0177	-114.32	6.0 %
3	0.0139	-148.42	2.1 %
4	0.0117	173.18	5.3 %
5	0.0104	132.60	10 %
6	0.0101	96.09	8.8 %
7	0.0097	62.70	5.1 %
8	0.0085	29.61	3.0 %
9	0.0075	-8.95	9.5 %

values of the sequence "g." The period must exceed 100 wavelengths if the first few values of the sequence are to agree within one percent.

In an effort to make a more systematic comparison, consider the DCMoM and SDFFT procedures applied to Equation (5.35) when both use the same basis functions. The basis function

$$B(x) = \frac{1}{0.7\sqrt{\pi}} \exp\left\{-\left\{\frac{x}{0.7\Delta x}\right\}^2\right\} \tag{5.41}$$

is easily used within the DCMoM procedure since it can be approximated by a function of finite support for numerical calculations. Its Fourier transform

$$\tilde{B}(f) = \Delta x \exp\left\{-\left\{\frac{0.7\pi f}{\Delta F}\right\}^2\right\} \tag{5.42}$$

is also a good approximation to a function of finite support, and can be used as a windowing function within the SDFFT process. (These functions are depicted in Figures 5.1 and 5.2.) Numerical values of the sequence "g" obtained by the SDFFT process with M equal to 255 and 1023 are presented in Tables 5.5 and 5.6, and are compared to the sequence produced by the DCMoM process using the basis functions of Equation (5.41). The relative accuracy obtained with a certain level of zero-padding agrees well with the previous results of Tables 5.3 and 5.4, and supports the above conclusions.

The difference between the SDFFT sequence and the DCMoM sequence as illustrated above appears to be entirely due to the fictitious periodic nature of the SDFFT representation (i.e., the finite M) and not to direct numerical difficulties associated with the singularity in $\tilde{K}(f)$. Of course, if M is taken to be an integer multiple of ΔF, a sample point will coincide with the singularity at $|f|=1$ and the numerical procedure will fail. As described in Section 5.3, it should be possible to incorporate a truncation function into the SDFFT

TABLE 5.5

SPATIAL DOMAIN SEQUENCE PRODUCED BY THE SDFFT USING AN EXPONENTIAL WINDOWING FUNCTION W(f) WITH M=255. THE RELATIVE DIFFERENCE BETWEEN THE SDFFT AND DCMOM SEQUENCE FOR THE SAME BASIS FUNCTIONS IS SHOWN (COMPARISONS NOT AVAILABLE FOR n=0 AND n=1)

n	$\lvert g_n \rvert$	$\angle g_n$ (degrees)	difference with DCMoM
0	0.0369	-39.45	-
1	0.0228	-68.00	-
2	0.0166	-111.26	2.8 %
3	0.0136	-145.30	4.8 %
4	0.0108	179.86	13 %
5	0.0089	138.05	17 %
6	0.0087	95.93	14 %
7	0.0090	62.61	4.6 %
8	0.0082	33.70	8.5 %
9	0.0067	-0.13	20 %

TABLE 5.6

SPATIAL DOMAIN SEQUENCE PRODUCED BY THE SDFFT USING AN EXPONENTIAL WINDOWING FUNCTION W(f) WITH M=1023. THE RELATIVE DIFFERENCE BETWEEN THE SDFFT AND DCMOM SEQUENCE FOR THE SAME BASIS FUNCTIONS IS SHOWN. (COMPARISONS NOT AVAILABLE FOR n=0 AND n=1).

n	$\lvert g_n \rvert$	$\angle g_n$ (degrees)	difference with DCMoM
0	0.0378	-40.46	-
1	0.0238	-68.53	-
2	0.0169	-110.67	1.9 %
3	0.0135	-147.21	2.4 %
4	0.0111	174.35	7.2 %
5	0.0099	133.86	9.9 %
6	0.0097	96.59	8.7 %
7	0.0092	63.52	3.4 %
8	0.0081	30.11	4.1 %
9	0.0071	-8.42	11 %

process, which will reduce the necessary zero-padding and explicitly smooth the singularity in $\tilde{K}(f)$. This may be necessary in a multidimensional problem in order to keep the necessary array sizes within the storage constraints of the computer in use.

It is interesting that the SDFFT sequence based upon the implicit basis function

$$W(x) = \frac{\sin(\pi x/\Delta x)}{(\pi x/\Delta x)} \tag{5.43}$$

agrees well with the sequence produced by the DCMoM with pulse basis functions (assuming the aliasing effects are suppressed). Since it appears that most of the previous results obtained with the SDFFT used implicit basis functions of the form of Equation (5.43), this may explain the reported success of the procedure [11], [13], [44] - [46], [50].

5.6. Application of the SDFFT Procedure to Periodic Equations

Integral equations representing infinite-periodic geometries such as idealized antenna arrays [90] or frequency selective surfaces [91] involve a Green's function K(x) that is periodic. If Equation (5.1) represents a periodic problem, it can be discretized using the SDFFT procedure without the detrimental effects introduced by the periodic nature of the FFT algorithm. For instance, since the Fourier transform of a periodic function is discrete, Equation (5.30) simplifies to

$$\tilde{G}_1(f) = \tilde{S}(f) \star [\tilde{W}(f)\, \tilde{K}(f)] \tag{5.44}$$

and Equation (5.26) is given by

$$g^{(1)}_{\ell-m} = W(x) \star K(x)\Big|_{x=(\ell-m)\Delta x} \tag{5.45}$$

Thus, the SDFFT process and the DCMoM process are equivalent for periodic equations if W(x) is chosen to be the desired basis function (or basis-testing function pair). Since the spatial Green's function for a periodic problem is usually an infinite summation, it may be easier to work with the SDFFT than with the DCMoM. Examples of the SDFFT applied to periodic equations are available in the literature [91], [93].

5.7. Summary

This chapter presents a comparison of two procedures for the discretization of convolutional integral equations, the discrete-convolutional method of moments (DCMoM) and the spectral-domain fast-Fourier transform method (SDFFT). Forms of the DCMoM have been used by Bojarski [10], Nyo and Harrington [12], [47], Borup and Gandhi [48], [49], Ray and Mittra [50] and Hurst and Mittra [51]. The SDFFT discretization has been used in connection with the spectral-iterative technique (SIT) developed by Ko and Mittra [44], Kastner and Mittra [11], [45], [46], Tsao and Mittra [91], and Sultan and Mittra [13]. For multi-dimensional problems, the SIT actually used a DCMoM discretization in one variable and an SDFFT discretization in the others [11].

For problems involving planar stratified media or problems that are periodic in one or more spatial variable, the SDFFT process might be easier to implement and computationally more efficient than the DCMoM because of the ease of working directly with the Fourier transform of the Green's function. However, for non-periodic equations, a large amount of zero-padding may be required to initially construct the discrete spectral Green's function from the analytical transform. Because of this, it may be difficult to ensure an arbitrary degree of accuracy in the discrete spectral Green's function. Since the DCMoM procedure works directly in the spatial domain, it appears to be preferable to the SDFFT

for problems where the Green's function is readily available in the spatial domain.

Based upon a comparison between the DCMoM and SDFFT procedures, it is apparent that both produce a matrix equation with Toeplitz or almost-Toeplitz symmetries that can be interpreted as an approximation to the integral equation of interest. Iterative algorithms for solving the SDFFT equation can be applied in an identical manner to solve the DCMoM equation, and vice versa. Both methods should be considered alternative possibilities, with relative advantages and disadvantages dependent on the specific integral equation under consideration.

6. EFFICIENT ITERATIVE IMPLEMENTATION FOR TE-SCATTERING BY DIELECTRIC CYLINDERS

6.1. Introduction

Preceding chapters have presented various aspects of iterative approaches for the numerical solution of electromagnetic scattering problems. It was mentioned that the most efficient approaches all rely on discrete-convolutional symmetries in the matrix equation. Some indication of how these symmetries can arise was given in Chapter 5, which discussed discretization procedures such as the DCMoM. This chapter presents specific details concerning the analysis of TE-wave scattering by lossy, inhomogeneous dielectric cylinders. It is shown that a DCMoM discretization of the electric field integral equation can yield a 2x2 block system, where each of the 4 blocks is itself a block-Toeplitz or almost-block-Toeplitz matrix. The low-storage iterative implementation of this problem requires the organization of the matrix equation to best exploit the symmetries. This chapter presents both the implementation of the approach and an evaluation of the accuracy of the overall numerical process.

A detailed description of the problem of TE-wave scattering by dielectric cylinders is provided by Richmond [80], who used pulse basis functions and point matching to discretize an electric field integral equation. Below, we briefly review the Richmond formulation applied to a special type of model. In particular, if the model under consideration is restricted to that whose cross-section is a lattice of evenly-spaced square cells, such as depicted in Figure 6.1, the resulting matrix equation contains the desired discrete-convolutional symmetries. In general, each cell of such a model may represent a region of different permittivity without affecting the significant symmetry features. Lossy regions may be modeled by the use of complex-valued permittivities.

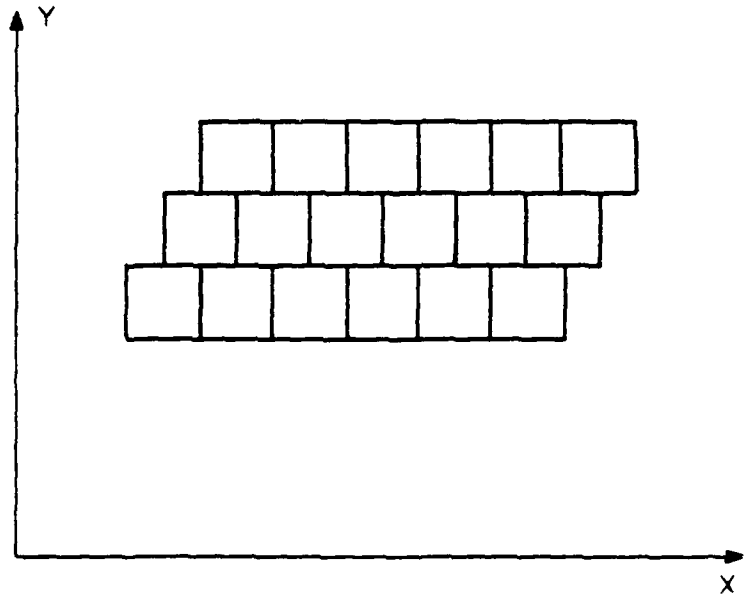

Figure 6.1. Cross-section of the type of model under consideration.

6.2. Formulation of the Matrix Equation

Suppose the cylindrical scatterer depicted in Figure 6.1, which is characterized by a complex-valued relative permittivity $\varepsilon_r(x,y)$, exists in the presence of a source $\overline{J}_{source}$. (Note that $\varepsilon_r=1$ outside the cylinder.) If the scatterer were absent, the source $\overline{J}_{source}$ would radiate a TE field $\overline{E}^{inc}$. Because of the presence of the cylinder, the actual field differs from $\overline{E}^{inc}$. The task is to find this field, which we denote $\overline{E}^{tot}$. These fields are related at all points in space by [80]

$$\overline{E}^{inc} = \overline{E}^{tot} - \frac{\text{curl curl } \overline{A} - \overline{J}}{j\omega\varepsilon_0} \tag{6.1}$$

where

$$\overline{A}(x,y) = \frac{-j}{4} \iint_{\Sigma} \overline{J}\, H_0^{(2)}(kR)\, dx'dy' \tag{6.2}$$

$$R = \sqrt{(x-x')^2 + (y-y')^2} \tag{6.3}$$

$$\overline{J} = j\omega\varepsilon_0(\varepsilon_r - 1)\, \overline{E}^{tot} \tag{6.4}$$

The integration in Equation (6.2) is over the support of the polarization current density $\overline{J}$, which vanishes outside the scatterer.

If restricted to the support of the scatterer, Equation (6.1) is an integro-differential equation for the unknown $\overline{J}(x,y)$. Once $\overline{J}$ is determined, secondary quantities of interest can be found from the magnetic vector potential $\overline{A}$ given in Equation (6.2), using standard formulas such as

$$\overline{H}^{tot} = \overline{H}^{inc} + \text{curl } \overline{A} \tag{6.5}$$

$$\overline{E}^{tot} = \overline{E}^{inc} + \frac{\text{curl curl } \overline{A} - \overline{J}}{j\omega\epsilon_0} \tag{6.6}$$

Since Equation (6.1) is usually impossible to solve exactly, the method of moments is employed to generate a finite dimensional matrix approximation to Equation (6.1). The process begins with the expansion of $\overline{J}$ in terms of basis functions, in this case, pulse functions that are defined as

$$\rho_n(x,y) = \begin{cases} 1 & (x,y) \in \text{cell } n \\ 0 & (x,y) \notin \text{cell } n \end{cases} \tag{6.7}$$

We make the assumption that the permittivity of each cell of the model is constant, and number the cells from 1 to N. Because there are two polarizations to consider, there are 2N unknowns to be determined. These are related to $\overline{J}$ by

$$\frac{\overline{J}(x,y)}{j\omega\epsilon_0} \approx \sum_{n=1}^{N} \{\hat{x}\, X_n\, \rho_n(x,y) + \hat{y}\, Y_n\, \rho_n(x,y)\} \tag{6.8}$$

where X_n and Y_n are unknown coefficients. From Equation (6.8), we obtain

$$\frac{\text{curl curl } \overline{A}}{j\omega\epsilon_0} = \sum \{\hat{x}\left[Y_n \frac{\partial^2 f_n}{\partial x \partial y} - X_n \frac{\partial^2 f_n}{\partial y^2}\right] + \hat{y}\left[X_n \frac{\partial^2 f_n}{\partial x \partial y} - Y_n \frac{\partial^2 f_n}{\partial x^2}\right]\} \tag{6.9}$$

where

$$f_n(x,y) = \iint_{\text{cell } n} \frac{1}{4j} H_0^{(2)}(kR)\, dx'dy' \tag{6.10}$$

Equation (6.1) may be separated into $\hat{x}$ and $\hat{y}$ vector components and enforced at the center of each of the cells to yield the 2Nx2N system

$$E^{inc}_{x\,m} = \frac{\epsilon_{rm}}{\epsilon_{rm} - 1} X_m - \sum_{n=1}^{N} \left\{ Y_n \frac{\partial^2 f_n}{\partial x \partial y} \right\} + \sum_{n=1}^{N} X_n \frac{\partial^2 f_n}{\partial y^2} \qquad m = 1,2,\ldots,N$$

$$E^{inc}_{y\,m} = \frac{\epsilon_{rm}}{\epsilon_{rm} - 1} Y_m - \sum_{n=1}^{N} X_n \frac{\partial^2 f_n}{\partial x \partial y} + \sum_{n=1}^{N} Y_n \frac{\partial^2 f_n}{\partial x^2} \qquad m = 1,2,\ldots,N \tag{6.11}$$

An essential feature of the above formulation is that $f_n(x,y)$ can be approximated analytically and then differentiated to yield

$$\frac{\partial^2 f_n}{\partial x^2}(x_m,y_m) = \begin{cases} \dfrac{j\pi ka}{4} H_1(ka) & (x_m,y_m) \in \text{cell } n \\ \dfrac{j\pi a J_1(ka)}{2} \left[H_0(k\rho) \dfrac{kx^2}{\rho^2} + H_1(k\rho) \left(\dfrac{y^2 - x^2}{\rho^3} \right) \right] & (x_m,y_m) \notin \text{cell } m \end{cases} \tag{6.12}$$

$$\frac{\partial^2 f_n}{\partial x \partial y}(x_m,y_m) = \begin{cases} 0 & (x_m,y_m) \in \text{cell } n \\ \dfrac{j\pi a J_1(ka)}{2} \left[H_0(k\rho) \dfrac{kxy}{\rho^2} + H_1(k\rho) \left(\dfrac{-2xy}{\rho^3} \right) \right] & (x_m,y_m) \notin \text{cell } n \end{cases} \tag{6.13}$$

$\frac{\partial^2 f_n}{\partial y^2}$ may be found from $\frac{\partial^2 f_n}{\partial x^2}$ by exchanging x and y in the above expression. If the suppressed time dependence is $e^{j\omega t}$, H_0 and H_1 represent the zero and first-order Hankel functions of the second kind. An assumption involved in the above process is that the cells are able to be approximated by a circular cross section of radius 'a.' The variables (x_m,y_m) refer to the center of cell m, $x=x_m-x_n$, $y=y_m-y_n$, and $\rho = \sqrt{x^2 + y^2}$.

6.3. The Symmetry Structure Imposed by a Lattice Geometry

The above expressions can be used with any configuration of cells in a model. For geometries of the type depicted in Figure 6.1, the cells are arranged in a

two-dimensional lattice. Upon investigation, Equations (6.12) and (6.13) are seen to preserve the translational structure of the lattice, so that if the numbering system were changed to the type depicted in Figure 6.2,

$$\frac{\partial^2 f_{11}}{\partial x^2}(\tilde{x}_2, y_2) = \frac{\partial^2 f_{22}}{\partial x^2}(\tilde{x}_3, y_3) \tag{6.14}$$

Thus, the efficient numbering system of Figure 6.2 enables us to exploit the two-dimensional structure. If NX and NY are used to denote the lattice dimensions, Equation (6.11) can be expressed in the form

$$E^{inc}_{x\,\alpha\beta} = \left(\frac{\varepsilon_{r\,\alpha\beta}}{\varepsilon_{r\,\alpha\beta}^{-1}}\right) X_{\alpha\beta} + \sum_{m=0}^{NY-1}\sum_{n=0}^{NX-1} Y_{nm}\, GYX_{\alpha-n,\beta-m} + \sum_{m=0}^{NY-1}\sum_{n=0}^{NX-1} X_{nm}\, GXX_{\alpha-n,\beta-n}$$

$$E^{inc}_{y\,\alpha\beta} = \left(\frac{\varepsilon_{r\,\alpha\beta}}{\varepsilon_{r\,\alpha\beta}^{-1}}\right) Y_{\alpha\beta} + \sum_{m=0}^{NY-1}\sum_{n=0}^{NX-1} X_{nm}\, GYX_{\alpha-n,\beta-m} + \sum_{m=0}^{NY-1}\sum_{n=0}^{NX-1} Y_{nm}\, GYY_{\alpha-n,\beta-m}$$

$$\alpha = 0,1,\ldots,NX-1 \tag{6.15}$$

$$\beta = 0,1,\ldots,NY-1$$

which clearly illustrates the symmetries present in the discrete system. Equation (6.15) is equivalent to the 2Nx2N matrix equation (where N=NXNY)

$$\begin{bmatrix} \underline{E}^{inc}_{x} \\ \hline \underline{E}^{inc}_{y} \end{bmatrix} = \left[\begin{array}{c|c} \underline{\underline{A}} & \underline{\underline{B}} \\ \hline \underline{\underline{B}} & \underline{\underline{D}} \end{array}\right] \begin{bmatrix} \underline{X} \\ \hline \underline{Y} \end{bmatrix} \tag{6.16}$$

Each of the blocks $\underline{\underline{A}}$, $\underline{\underline{B}}$, and $\underline{\underline{D}}$ is itself an NX-th order block-Toeplitz or almost block-Toeplitz matrix, whose individual entries are Toeplitz (or almost

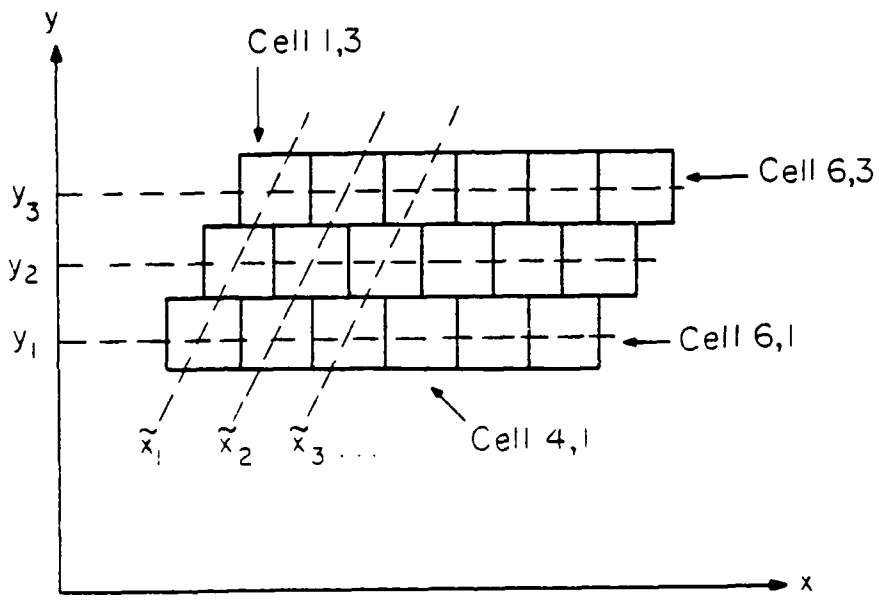

Figure 6.2. Numbering system used to build translational symmetries into the problem.

Toeplitz) matrices of order NY. (The system could be configured so that the block-Toeplitz systems are order NY and their entries order NX Toeplitz systems.) For instance, $\underline{\underline{A}}$ may have the form

$$A = \begin{bmatrix} \underline{\underline{G}}_0^{(1)} & \underline{\underline{G}}_1 & \underline{\underline{G}}_2 & \cdots & \underline{\underline{G}}_{NX} \\ \underline{\underline{G}}_1 & \underline{\underline{G}}_0^{(2)} & \underline{\underline{G}}_1 & & \\ \underline{\underline{G}}_2 & \underline{\underline{G}}_1 & \underline{\underline{G}}_0^{(3)} & & \\ & & & & \\ \vdots & & & & \vdots \\ & & & & \\ \underline{\underline{G}}_{NX} & & & & \underline{\underline{G}}_0^{(N_X)} \end{bmatrix} \tag{6.17}$$

where each of the off-diagonal entries are Toeplitz matrices

$$\underline{\underline{G}} = \begin{bmatrix} g_0 & g_1 & g_2 & \cdots & g_{NY} \\ g_1 & g_0 & g_1 & & \\ g_2 & g_1 & g_0 & & \\ & & & & \\ \vdots & & & & \\ & & & & \\ g_{NY} & & & & g_0 \end{bmatrix} \tag{6.18}$$

If the original dielectric cylinder is inhomogeneous, i.e., if ε_r varies from cell to cell, the diagonal entries of $\underline{\underline{A}}$ are not exactly Toeplitz, and are referred to as almost-Toeplitz.

As compared to the system as written in Equation (6.11), the significant redundancy features of the latter matrix equation permit the fully populated 2Nx2N matrix to be generated from only 5[illegible] entries. Thus, the equation is an excellent candidate for iterative solution.

6.4. Fast-Fourier Transform Implementation of the Matrix Operator

An iterative solution of the above matrix equation can be obtained using one of the algorithms of Chapter 2. These algorithms do not require the matrix operator explicitly, only that it (and its adjoint or approximate inverse) be available to operate on given vectors throughout the iterative process. Thus, the manner in which the operator is implemented is of no consequence to the algorithm, and typically is accomplished in as storage-efficient manner as is possible. One approach is simply to perform the matrix multiplication directly, which requires N^2 complex operations and a storage of 2.5N complex variables, where $N=2(NX \times NY)$.

Since the matrix operator is primarily a superposition of discrete convolution operations, it can be implemented with the aid of the fast-Fourier transform (FFT) algorithm, as described in many texts [42], [43] and in Chapter 5. For the two dimensional non-periodic geometry involved in this example, the task requires each of the four discrete convolution operations to be replaced by three successive FFT operations, with an additional vector multiply operation required in each case. (Actually, only two FFT operations are necessary per convolution during each iteration, as one could be performed initially and the result stored. The discussion below assumes that this is the case). The advantage of using the FFT is the computational efficiency for large systems, since approximately Mlog(M) operations are necessary to compute a one-dimensional FFT or inverse FFT of length M, where the logarithm is of base 2. Thus, each of the two-dimensional discrete convolution operations contained in the matrix operator

will require approximately 2(2NX-1)(2NY-1)log(2NX-1)(2NY-1) + (2NX-1)(2NY-1) operations via the FFT, as compared to (NXNYNXNY) if the convolution is performed by direct multiplication. (These estimates assume that the two dimensional FFT's are obtained from one dimensional FFT's, as is typical practice. Furthermore, the computational effects of zero-padding are included in the estimates.)

As illustrated in the above discussion, the effects of zero-padding impact the computational efficiency as well as the storage efficiency, and specific calculations are considered in order to judge the overall efficiency of using the FFT algorithm to implement this particular operator. For instance, if the model in question is a 10x10 lattice, direct matrix multiplication requires 40,000 operations and a storage of 500 elements. The FFT assisted implementation of the matrix operator requires 25,980 operations, and a storage of more than 2166 elements. Thus, for large problems, the FFT approach is computationally efficient, but always requires more storage than the direct approach. In addition, the overhead imposed by the FFT subroutine adds to the overall program length and the available storage space, which have not been incorporated into the above estimates.

In summary, the matrix operator can be implemented either as a direct matrix multiplication or as an FFT-assisted computation. Depending on the specific geometry involved, one approach may be more efficient than the other. In a given situation, specific parameters describing the FFT algorithms available for use will determine the actual relative efficiency. It may be worth noting that the commercial "array processors" frequently supply special FFT algorithms, and their availability may shift the relative efficiency toward the FFT-assisted approach.

6.5. Performance of the Numerical Approach

The performance of the above approach can be studied by comparing numerical solutions and exact analytical results for homogeneous, circular cylinders. For instance, three models of circular cross sections are illustrated in Figure 6.3. Figures 6.4 to 6.12 show values of the total electric field produced by the moment method approach when applied to these models and compared to exact solutions based on an eigenfunction expansion [94,. In this case, the cylinder under consideration is homogeneous and lossy, with ε_r=2.56-j2.56, and has a circumference of two free-space wavelengths. As the models are refined with smaller cell sizes, the numerical solutions appear to approach the exact.

The computer program that generated these data used a mixed radix FFT algorithm to implement the matrix operator, and required 3.4, 24.2, and 72.4 CPU seconds of execution time on the CDC Cyber 175 computer to solve the N=42, N=202, and N=512 systems using the CGM algorithm of Chapter 2. For large N, this approach clearly exceeds the efficiency of the MER approach discussed in Chapter 4. However, execution time data can be misleading if not interpreted correctly. The N=512 equation only required 18 iterations to converge, and used approximately 3.4 seconds of CPU time per iteration. The fast convergence is related to the large cell density in use, as explained in Chapter 3.

For values of $|\varepsilon_r|$ exceeding 10, the accuracy of the numerical solutions is poor, regardless of cell density. Furthermore, the convergence of the CGM algorithm becomes less and less rapid as $|\varepsilon_r|$ increases, suggesting that the system becomes ill-conditioned for large $|\varepsilon_r|$. This seems to be true whether ε_r is complex-valued (representing a lossy media) or not. For ε_r=76-j278, a value often used to model biological media [81], the approach fails to produce numerical solutions that exhibit any kind of convergence behavior as larger cell den-

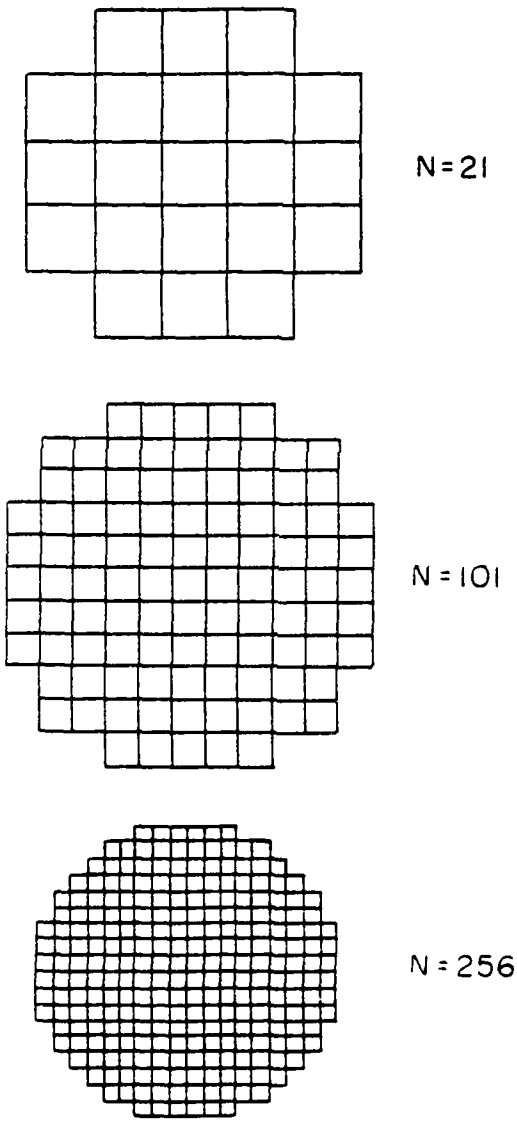

Figure 6.3. Three models of dielectric cylinders with circular cross-sections.

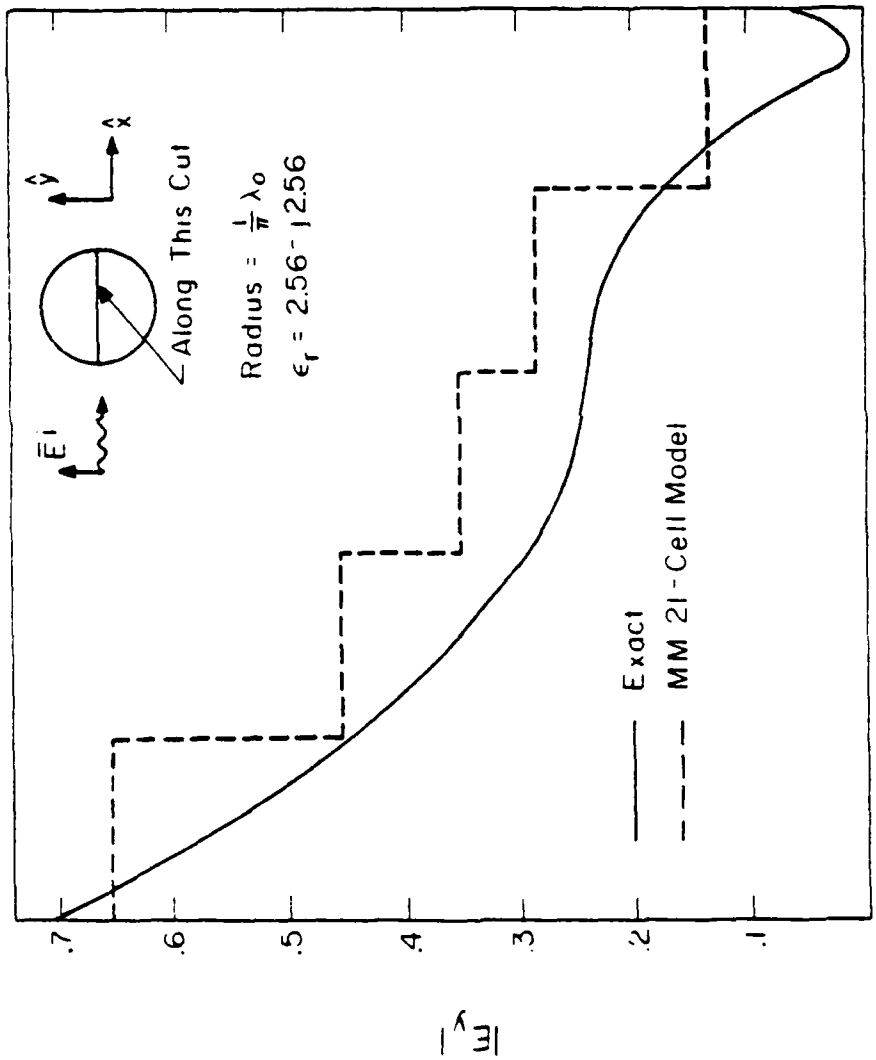

Figure 6.4. Comparison of the exact solution and the numerical solution obtained with the 21 cell model. The cell density is 18 cells/λ_d.

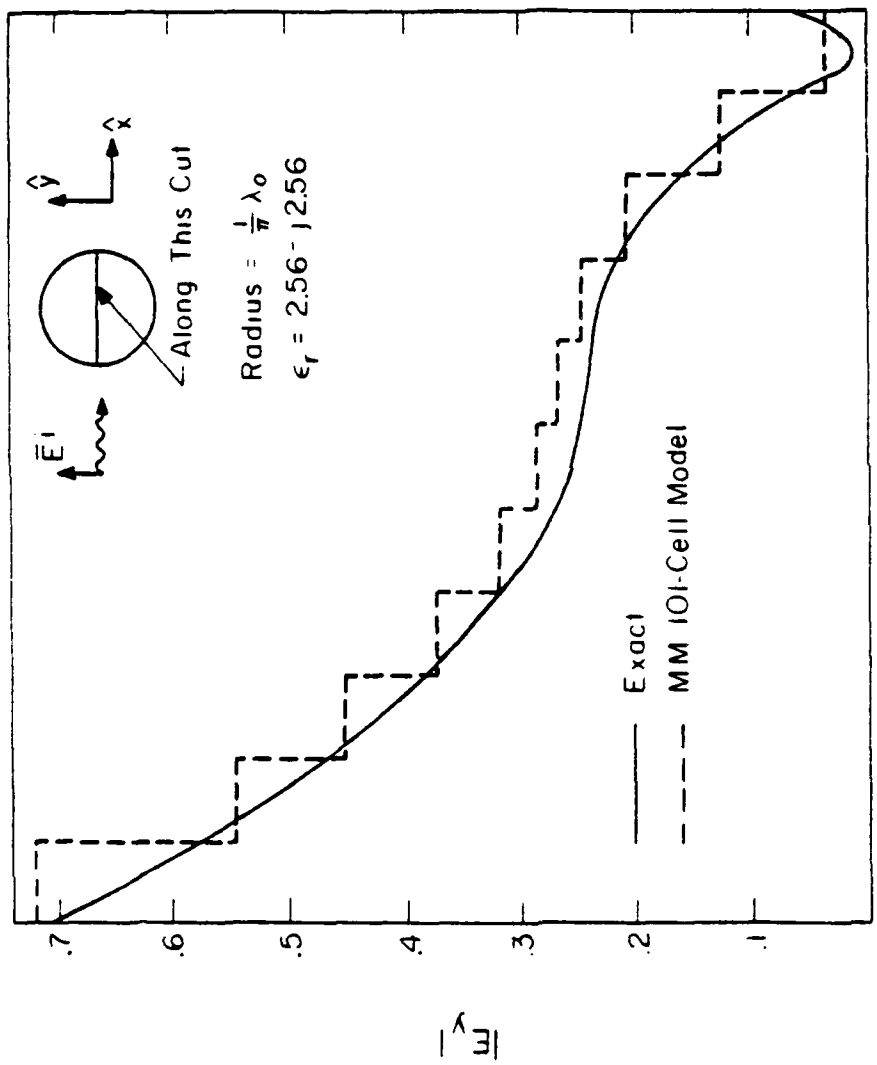

Figure 6.5. Comparison of the exact solution and the numerical solution obtained with the 101 cell model. The cell density is 88 cells/λ_d^2.

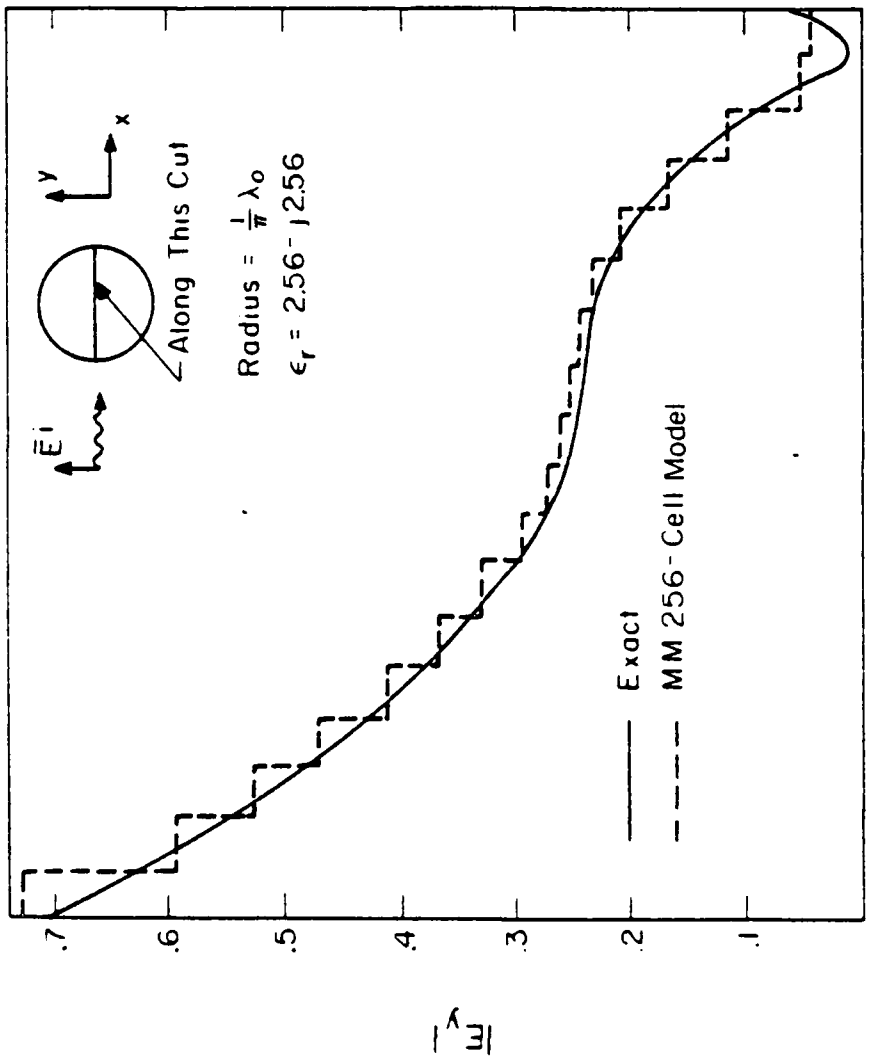

Figure 6.6. Comparison of the exact solution and the numerical solution obtained with the 256 cell model. The cell density is 222 cells/λ_o^2.

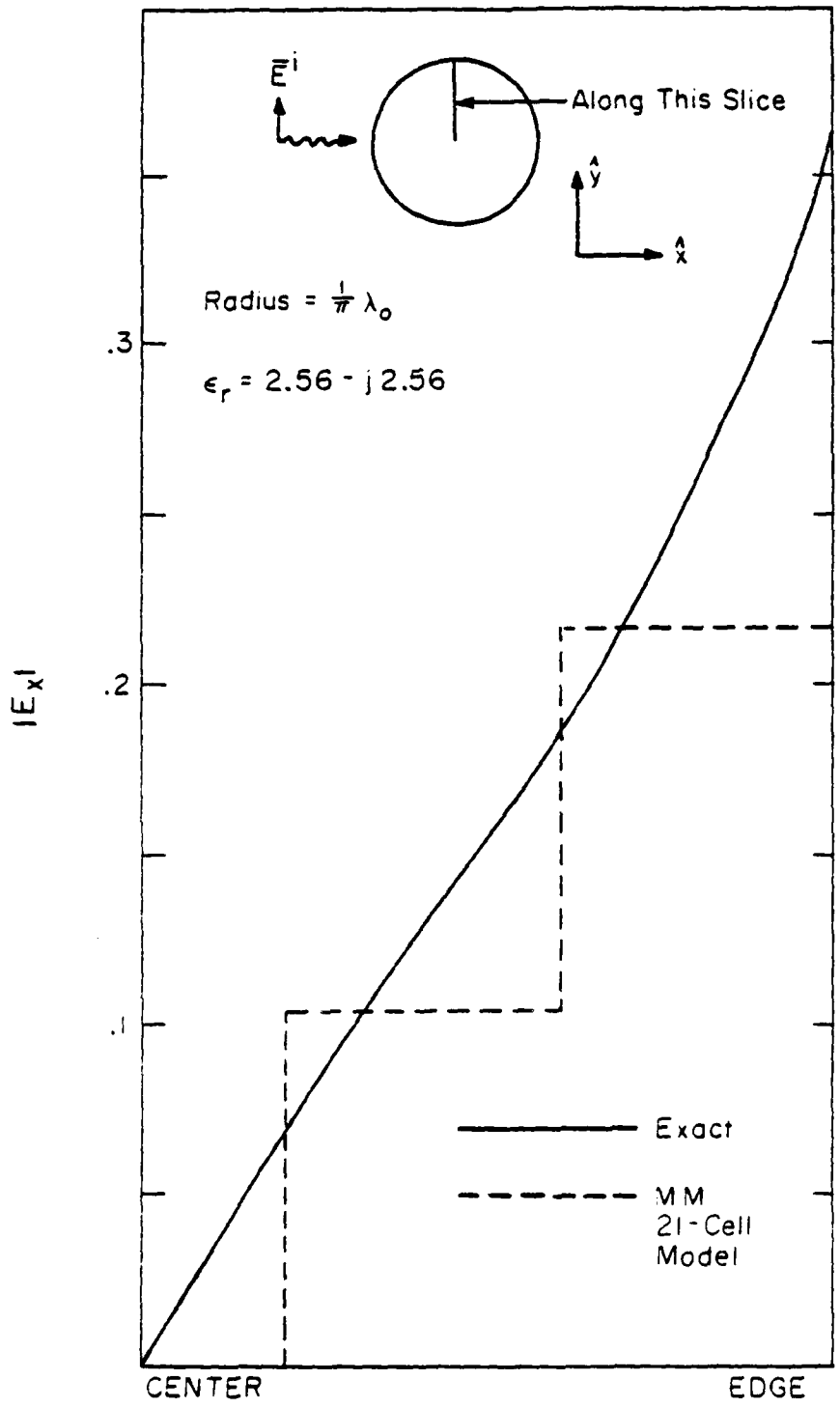

Figure 6.7. Comparison of the exact solution and the numerical solution obtained with the 21 cell model.

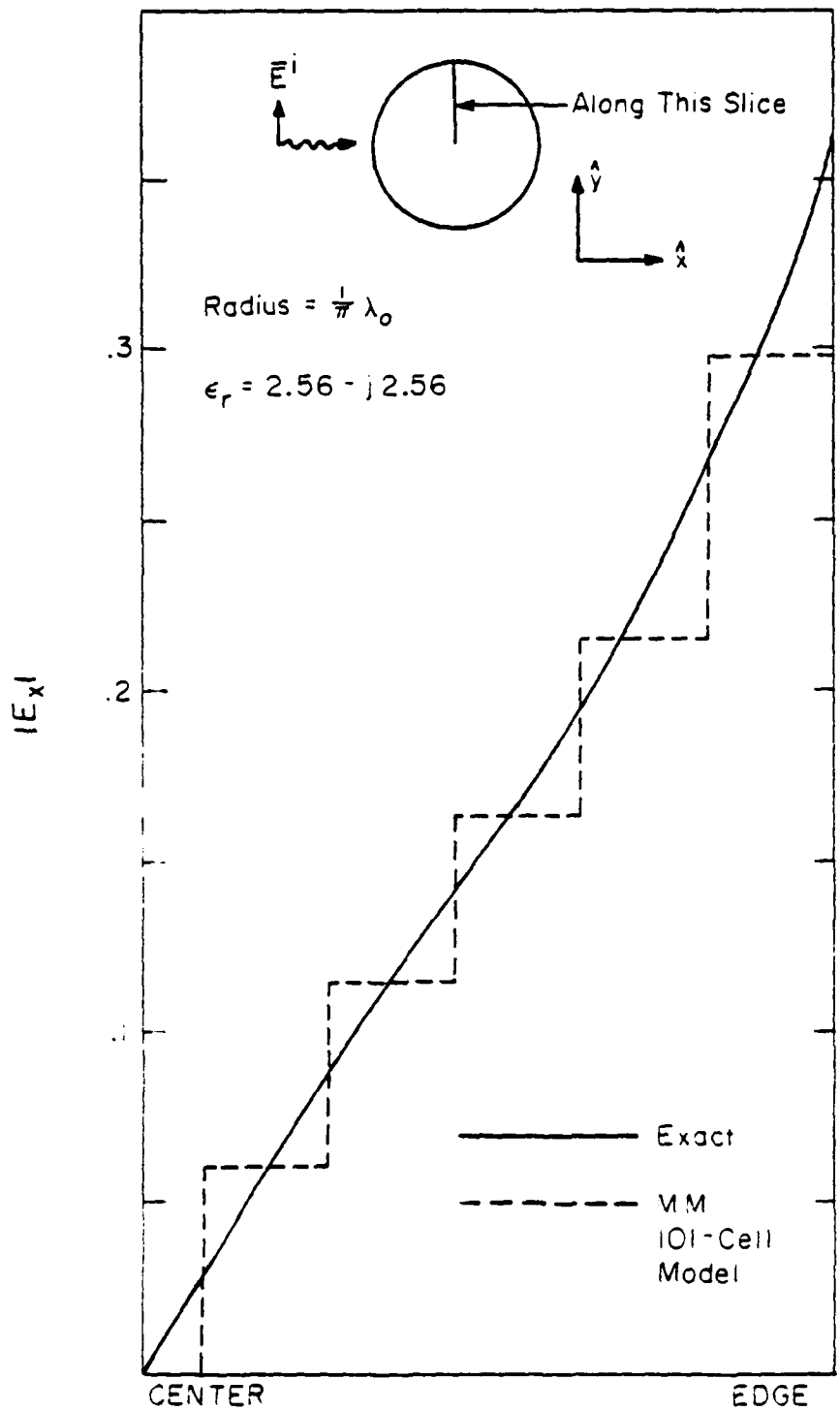

Figure 6.8. Comparison of the exact solution and the numerical solution obtained with the 101 cell model.

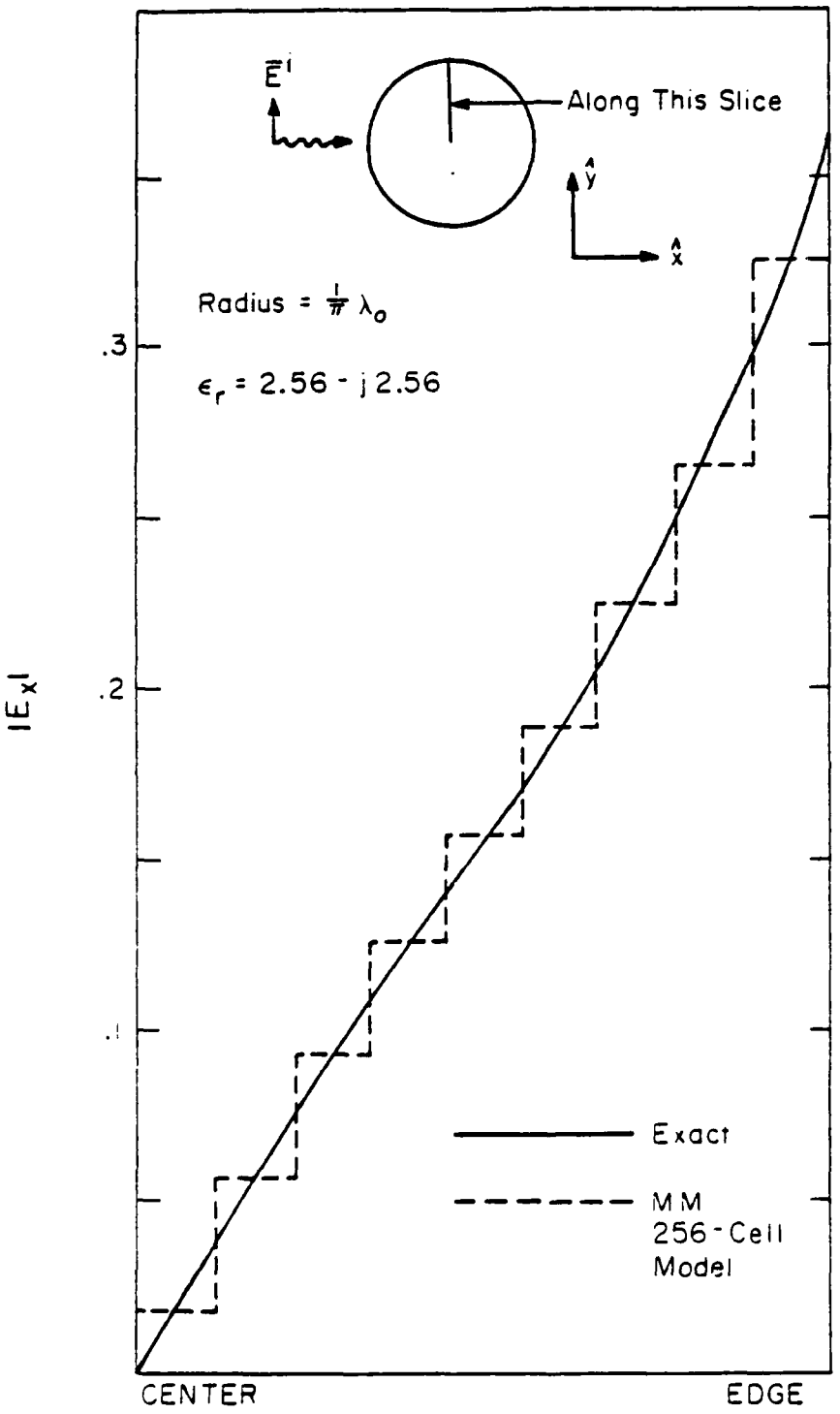

Figure 6.9. Comparison of the exact solution and the numerical solution obtained with the 256 cell model.

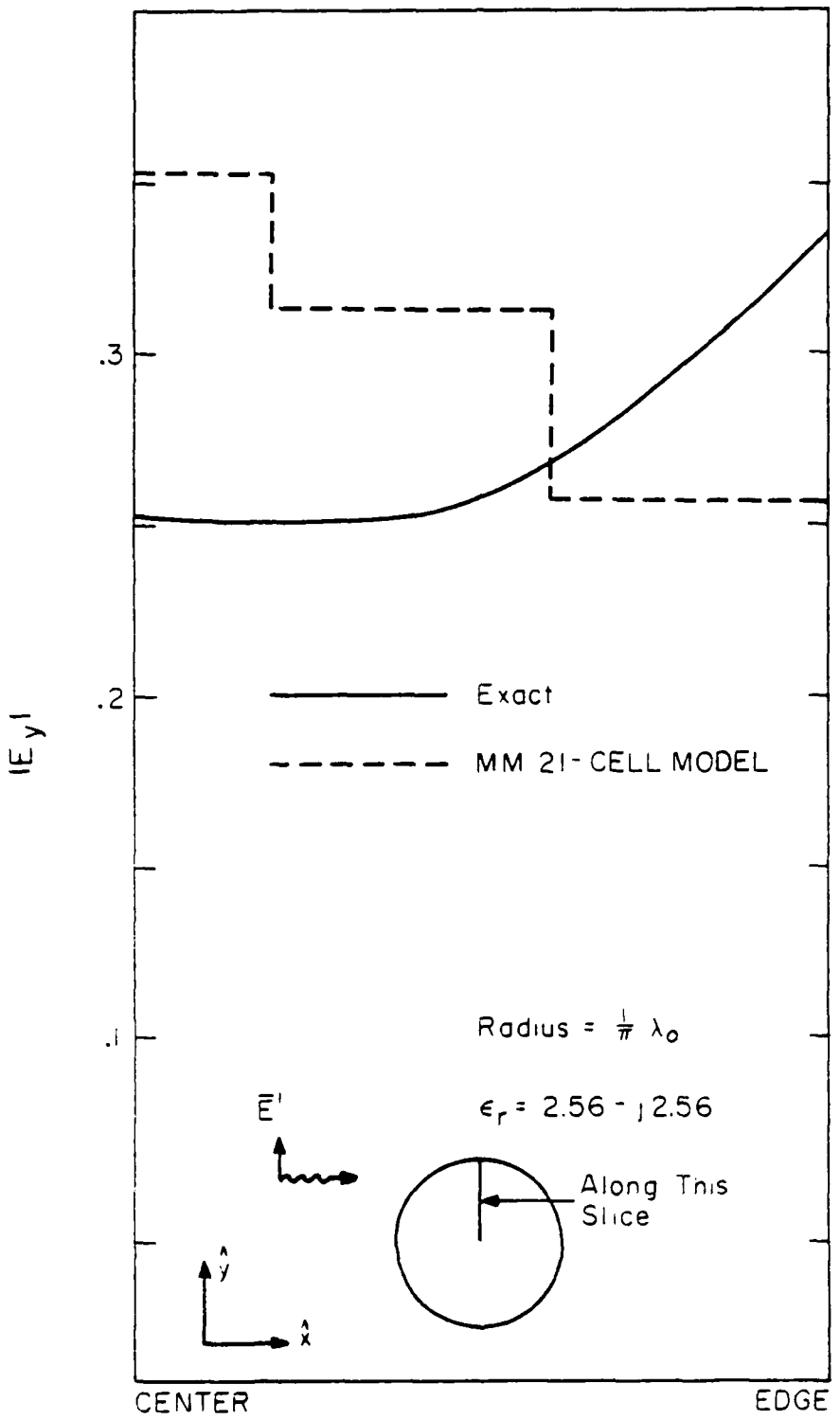

Figure 6.10. Comparison of the exact solution and the numerical solution obtained with the 21 cell model.

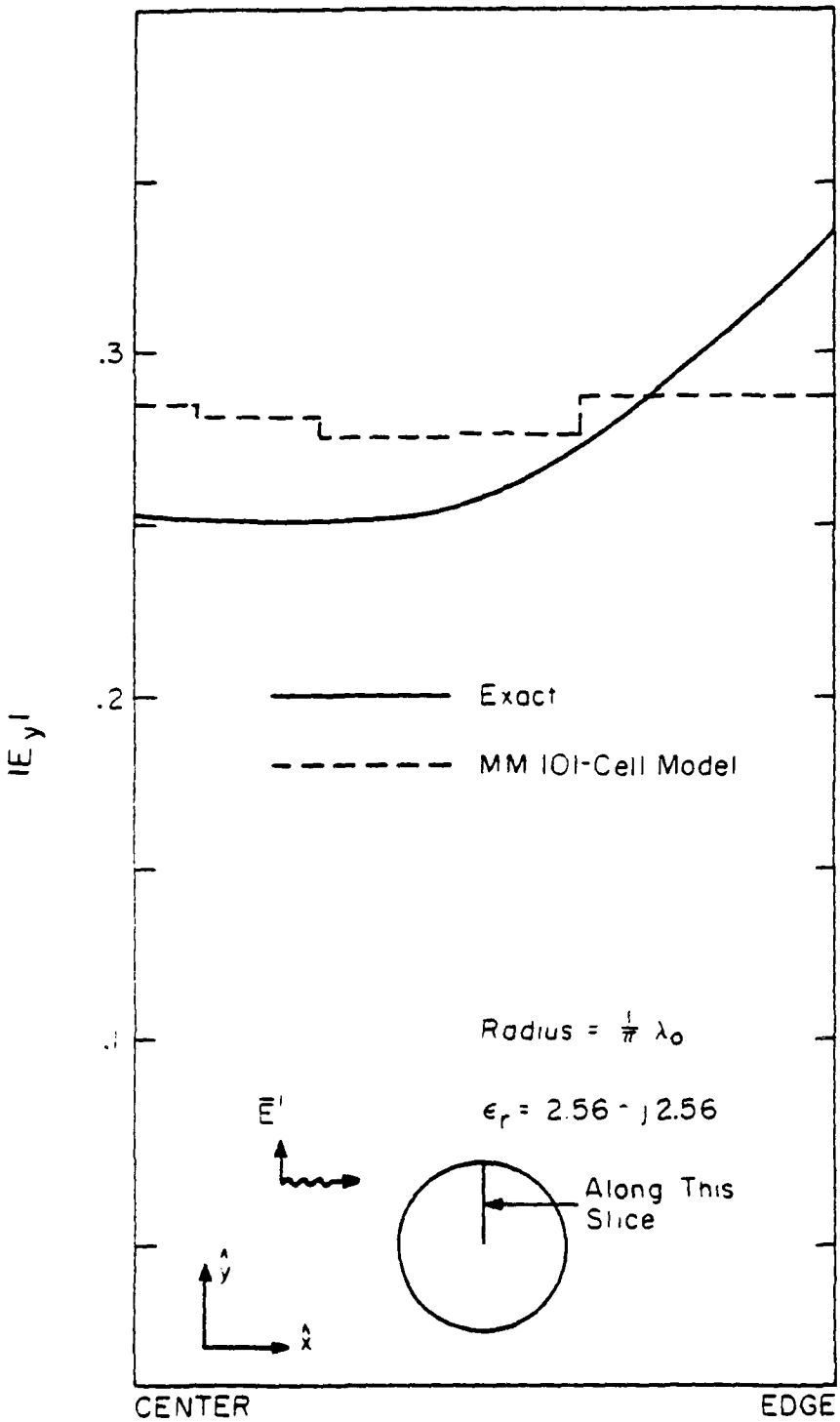

Figure 6.11. Comparison of the exact solution and the numerical solution obtained with the 101 cell model.

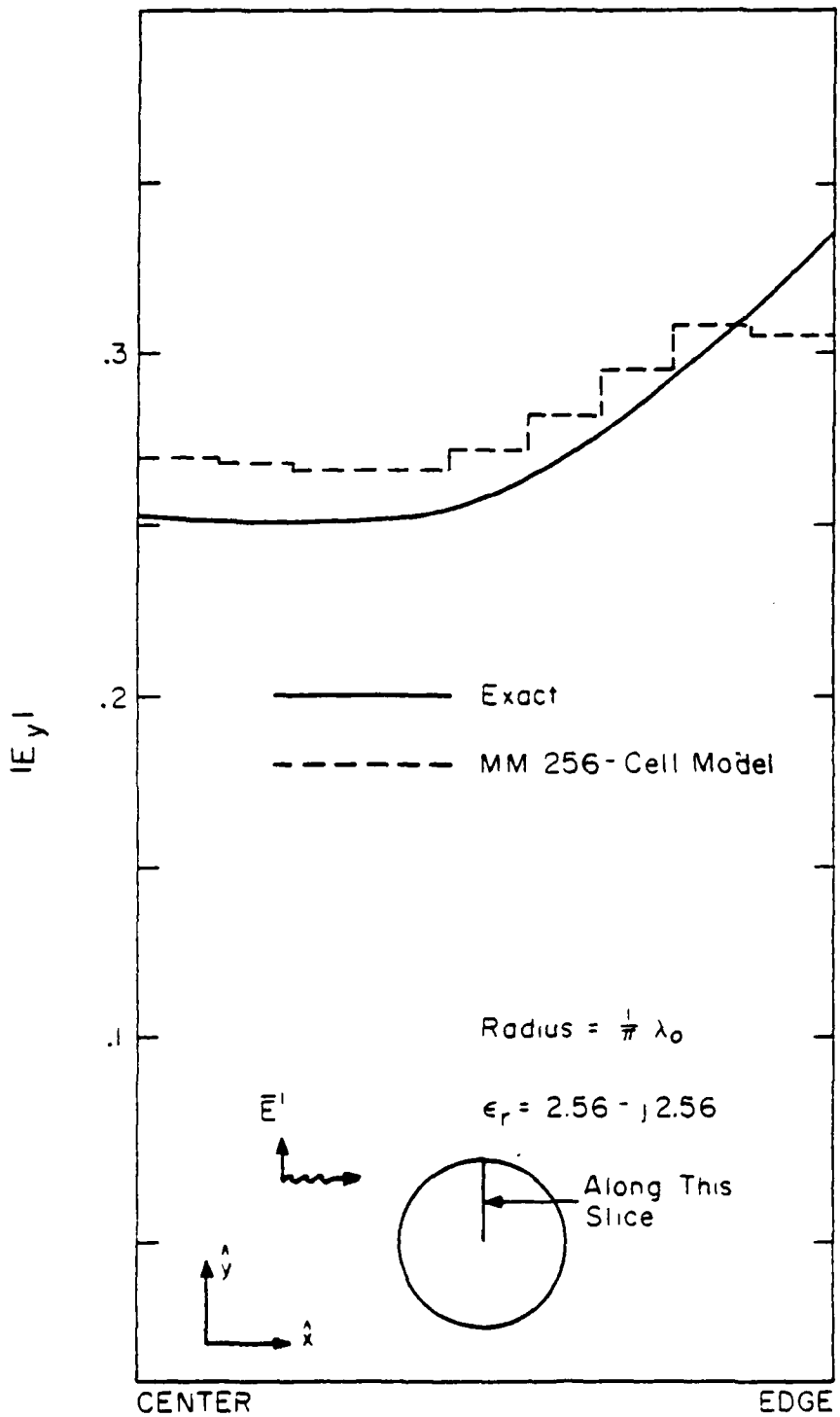

Figure 6.12. Comparison of the exact solution and the numerical solution obtained with the 256 cell model.

sities are used. The reason for the failure of the approach for large $|\varepsilon_r|$ is not as yet understood. However, it is not a direct consequence of the general DCMoM procedure or the iterative solution. It may be a result of the integral equation or the specific basis and testing functions used within the discretization.

6.6. Summary

The DCMoM procedure can be applied to discretize the electric field integral equation representing scattering by inhomogeneous, lossy dielectric cylinders. The resulting matrix equation is convenient to solve iteratively because of the discrete-convolutional redundancy built into the system. In this manner, efficient use can be made of available fast-access memory, enabling the treatment of electrically larger scatterers than possible by the conventional approach [80] with the same storage constraints.

From a comparison of numerical solutions and exact solutions for homogeneous, circular cylinders, the process appears to produce accurate solutions for $|\varepsilon_r|$ less than 10. The numerical solutions for problems with large $|\varepsilon_r|$ fail to show convergence behavior as higher cell densities are used in the discretization procedure, and do not appear to be reliable approximations to the desired solution. The reasons for the failure of the method are not as yet understood.

7. ITERATIVE ANALYSIS OF HOLLOW, FINITE-LENGTH CIRCULAR CONDUCTING CYLINDERS

7.1. Introduction

Most of the examples presented in previous chapters involved a moment-method discretization with simple pulse basis and Dirac delta testing functions. In this chapter, a DCMoM discretization using higher-order basis and testing functions is applied to the analysis of wave scattering by a finite-length, hollow, circular conducting cylinder.

Because circular cylinders are periodic in the azimuthal variable, all quantities can be represented by a Fourier series. As a result, Fourier harmonics decouple and can be found independently, then superimposed to produce the total solution. This procedure is known as a "body of revolution" formulation, and is desirable because the matrix equation that must be solved for each of the harmonics is much smaller than the single equation arising from conventional techniques [95], [96]. The geometry under consideration is restricted to circular cylinders; thus, the integral equations for each harmonic can be discretized using the DCMoM. Since the resulting system has discrete-convolutional symmetries, an iterative solution could require significantly less storage than a conventional direct solution and add to the savings already achieved by the use of the "body of revolution" formulation. Therefore, the combination of the "body of revolution" formulation and the DCMoM discretization permits the treatment of significantly larger scatterers than otherwise possible with the same storage constraints.

A different approach for the iterative analysis of finite circular cylinders was presented by Hurst and Mittra [51]. Their approach was based upon a straightforward DCMoM discretization without the "body of revolution" formulation and required the solution of one large matrix equation as opposed to a

solution of a number of smaller order systems. An example involving a cylinder with length 3.0 λ_0 was treated via the iterative solution of one matrix equation of order 4032 [51]. Using the "body of revolution" formulation, the same problem could be treated (to the same degree of accuracy) by the solution of a number of systems of order 63. The trade-off in computational efficiency between these two approaches depends on the number of Fourier harmonics excited by the incident field and on the radius of the cylinder. In general, the "body of revolution" approach will be much more efficient than the other if the incident waves are propagating in a direction nearly parallel to the axis of the cylinder, for few harmonics are excited in this case.

7.2. Formulation of the Matrix Equation

The geometry under consideration is shown in Figure 7.1. If the vector components of the current density are expanded in Fourier harmonics

$$J_z(z,\phi) = \sum_{m=-\infty}^{\infty} J_{zm}(z)\, e^{jm\phi} \tag{7.1}$$

$$J_\phi(z,\phi) = \sum_{m=-\infty}^{\infty} J_{\phi m}(z)\, e^{jm\phi} \tag{7.2}$$

coupled integral equations for the m-th harmonic are

$$-E^{inc}_{\phi m}(z) = \frac{2\pi a}{j\omega\epsilon_0} \int_{z'=z_0}^{z_n} \left\{k^2 J_{\phi m}(z') \frac{G_{m-1}(z-z') + G_{m+1}(z-z')}{2} + \left[-\frac{m^2}{a^2} J_{\phi m}(z') + j\frac{m}{a}\frac{\partial J_{zm}}{\partial z'}\right] G_m(z-z')\right\} dz' \tag{7.3}$$

$$-E^{inc}_{zm}(z) = \frac{2\pi a}{j\omega\epsilon_0} \int_{z'=z_0}^{z_n} k^2 J_{zm}(z')\, G_m(z-z')\, dz' + \frac{2\pi a}{j\omega\epsilon_0} \frac{\partial}{\partial z} \int_{z'=z_0}^{z_n} \left\{j\frac{m}{a} J_{\phi m}(z') + \frac{\partial J_{zm}}{\partial z'}\right\} G_m(z-z')\, dz' \tag{7.4}$$

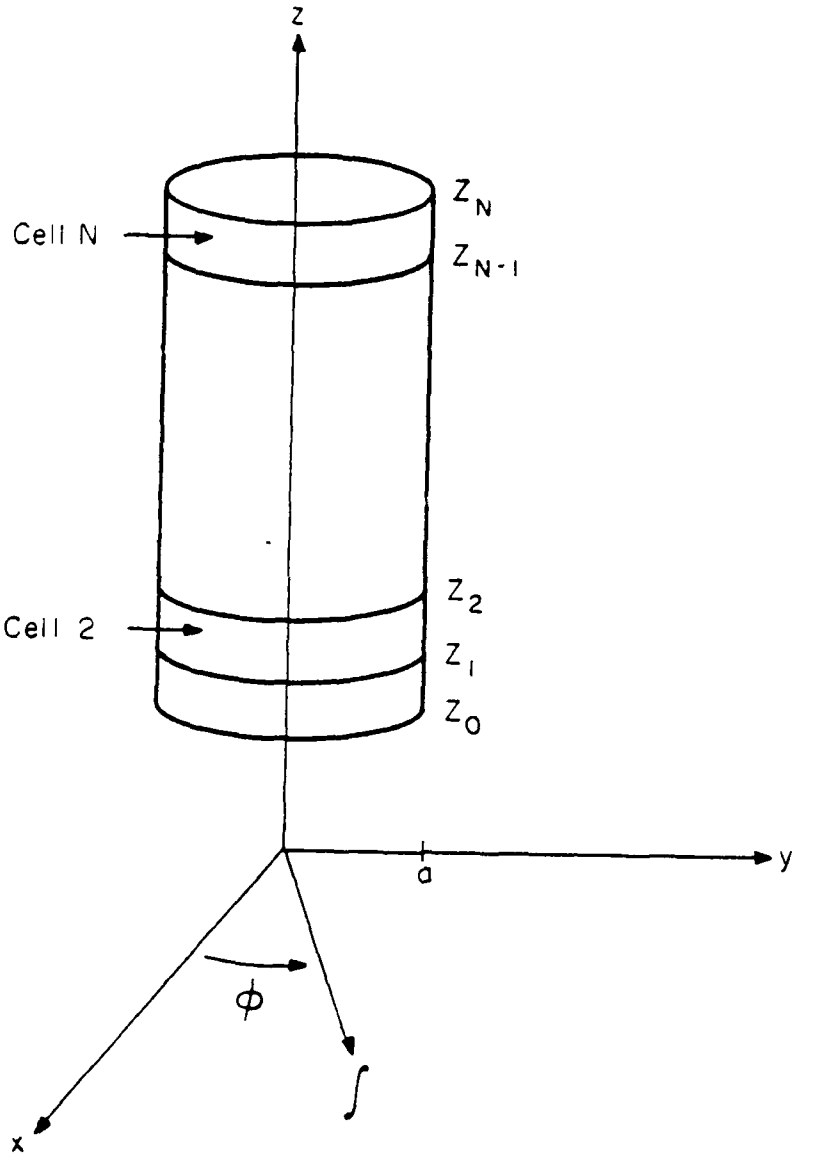

Figure 7.1. Geometry of the hollow cylinder under consideration.

where

$$E^{inc}_{\phi m}(z) = \frac{1}{2\pi} \int_{-\pi}^{\pi} E^{inc}_{\phi}(a,z,\phi)\, e^{-jm\phi}\, d\phi \tag{7.5}$$

$$E^{inc}_{zm}(z) = \frac{1}{2\pi} \int_{-\pi}^{\pi} E^{inc}_{z}(a,z,\phi)\, e^{-jm\phi}\, d\phi \tag{7.6}$$

$$G_m(z-z') = \frac{1}{\pi} \int_{0}^{\pi} \frac{e^{-jkR}}{4\pi R} \cos(m\alpha)\, d\alpha \tag{7.7}$$

$$R = \sqrt{(z - z')^2 + 4a^2 \sin^2\left(\frac{\alpha}{2}\right)} \tag{7.8}$$

These equations have been derived elsewhere in a more general form [95], [96], [97].

Consider a discretization of Equations (7.3) and (7.4) where the current densities are replaced by

$$J_{\phi m}(z) \cong \sum_{n=1}^{N} j_{\phi n}\, p(z;z_{n-1},z_n) \tag{7.9}$$

$$J_{zm}(z) \cong \sum_{n=1}^{N-1} j_{zn}\, t(z;z_{n-1},z_n z_{n+1}) \tag{7.10}$$

The expansion functions are defined in Figure 7.2. The integral equations can be enforced approximately by integrating them with testing functions. Specifically, Equation (7.3) is to be tested with

$$T_{\phi p}(z) = \delta\left(z - \frac{z_{p-1} + z_p}{2}\right) \qquad p = 1,2,\ldots,N \tag{7.11}$$

and Equation (7.4) is to be tested with

$$T_{zp}(z) = p\left(z;\ \frac{z_{p-1} + z_p}{2},\ \frac{z_p + z_{p+1}}{2}\right) \qquad p = 1,2,\ldots,N-1 \tag{7.12}$$

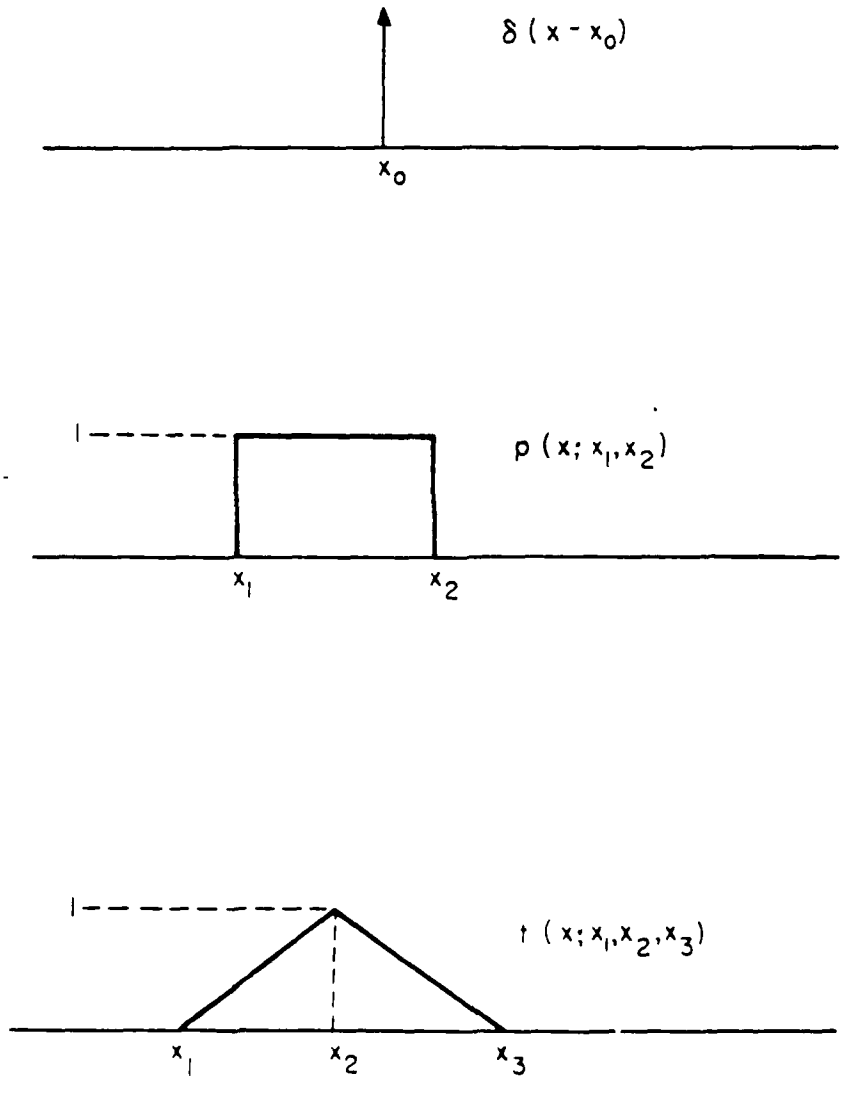

Figure 7.2. Definition of basis and testing functions.

If the cells in the model as depicted in Figure 7.1 are identical in size, the discretization process reduces to a form of the DCMoM and produces a matrix equation for the m-th harmonic of the form

$$\begin{bmatrix} \underline{\underline{A}} & \underline{\underline{B}} \\ \underline{\underline{C}} & \underline{\underline{D}} \end{bmatrix} \begin{bmatrix} \underline{j}_\phi \\ \underline{j}_z \end{bmatrix} = \begin{bmatrix} \underline{e}_\phi \\ \underline{e}_z \end{bmatrix} \tag{7.13}$$

where

$$A_{pn} = -ja\eta\Delta z\left(k^2 K^m_{p-n} - \frac{m^2}{a^2} I^m_{p-n}\right) \tag{7.14}$$

$$B_{pn} = m\eta\left(I^m_{p-n} - I^m_{p-n-1}\right) \tag{7.15}$$

$$C_{pn} = m\eta\left(I^m_{p-n+1} - I^m_{p-n}\right) \tag{7.16}$$

$$D_{pn} \cong \frac{-ja\eta}{\Delta z}\left\{(k^2\Delta z^2 - 2)\, I^m_{p-n} + I^m_{p-n-1} + I^m_{p-n+1}\right\} \tag{7.17}$$

$$I^m_q = \int_{-\Delta z/2}^{\Delta z/2} G_m(q\Delta z - z')\, dz' \tag{7.18}$$

$$K^m_q = \frac{I^{m-1}_q + I^{m+1}_q}{2} \tag{7.19}$$

The sub-matrices A and D are Toeplitz, i.e.,

$$\underline{\underline{A}} = \begin{bmatrix} a_0 & a_1 & a_2 & \\ a_1 & a_0 & a_1 & \cdots \\ a_2 & a_1 & a_0 & \\ & \vdots & & \end{bmatrix} \tag{7.20}$$

$$\underline{\underline{D}} = \begin{bmatrix} d_0 & d_1 & d_2 & \\ d_1 & d_0 & d_1 & \cdots \\ d_2 & d_1 & d_0 & \\ & \vdots & & \end{bmatrix} \tag{7.21}$$

where

$$a_n = -ja\eta\Delta z\left(k^2 K_n^m - \frac{m^2}{a^2} I_n^m\right) \tag{7.22}$$

$$d_n = \frac{-ja\eta}{\Delta z}\left\{(k^2\Delta z^2 - 2)\, I_n^m + I_{n-1}^m + I_{n+1}^m\right\} \tag{7.23}$$

Sub-matrices B and C are not square matrices, but they are almost Toeplitz and take the form

$$\underline{\underline{B}} = \begin{bmatrix} -b_1 & -b_2 & -b_3 & & -b_{N-1} \\ b_1 & -b_1 & -b_2 & \cdots & \\ b_2 & b_1 & -b_1 & & \\ & \vdots & & & \\ & & & b_1 & -b_1 \\ b_{N-1} & & & b_2 & b_1 \end{bmatrix} \tag{7.24}$$

$$\underline{\underline{C}} = \begin{bmatrix} b_1 & -b_1 & -b_2 & & -b_{N-1} \\ b_2 & b_1 & -b_1 & \cdots & \\ b_3 & b_2 & b_1 & & \\ & \vdots & & & \\ & & & -b_1 & -b_2 \\ b_{N-1} & & & b_1 & -b_1 \end{bmatrix} \tag{7.25}$$

where

$$b_n = mn\left(I_n^m - I_{n-1}^m\right) \tag{7.26}$$

Note that

$$\underline{\underline{B}} = -\underline{\underline{C}}^T \tag{7.27}$$

where the superscript "T" denotes matrix transpose.

The computation of I_q^m requires two-dimensional numerical integration, and is complicated by the presence of a singularity in the integrand when q=0. Techniques for computing I_0^m are available in the literature [68], [97]. For the examples to follow, numerical integration was implemented using standard adaptive library routines.

The formulation of the above matrix representation is identical to a general approach presented by Glisson and Wilton [97], if their approach is specialized to the geometry of Figure 7.1. Detailed information about the matrix elements, symmetries in the equations, and the decomposition of a plane wave field into Fourier harmonics may be found in the literature [96], [97].

7.3. Incorporation of an Impedance Boundary Condition

The above integral equation is based upon the boundary condition

$$\overline{E}_{tan} = 0 \tag{7.28}$$

which is to be enforced on the surface of a perfect electric conductor (PEC). A generalization of this condition is

$$\overline{E}_{tan} = Z_s \overline{J} \tag{7.29}$$

which happens to coincide with the classical definition of surface impedance for an imperfect but thick conductor [98]. A boundary condition of this same form has also been used under certain conditions to describe electrically thin materials, and may apply to the hollow cylinder geometry under consideration [99], [100].

If the surface impedance Z_s is invariant to azimuthal variation, boundary conditions of the form of Equation (7.29) can be incorporated into the above integral equations without affecting the significant symmetry features of the DCMoM matrix equation. The effects of the additional terms in Equations (7.3) and (7.4) are to modify the diagonal entries of A and D, and may destroy the Toeplitz nature of these sub-matrices. However, the off-diagonal entries retain the discrete-convolutional symmetries which permit efficient iterative treatment of the system.

7.4. Performance of the Method

To test the efficiency of the above approach, two computer programs were created using the iterative CGM algorithm to solve Equation (7.13). The first program used explicit matrix multiplication to implement the operator and adjoint operator required by the CGM. The other program used a mixed-radix FFT algorithm to perform the discrete convolutions, as described in Chapters 5 and 6. As expected, the FFT-assisted approach is more efficient if the order of the matrix equation is large. However, for the specific computer codes in use, the

trade-off between the two approaches was not as favorable as expected. For instance, when used to treat a 99 × 99 system of equations on the CDC Cyber 175 computer, the FFT-assisted approach required approximately 0.35 second of computation time per iteration, as opposed to 0.24 second per iteration for direct matrix multiplication. For a 399 × 399 system, the FFT-assisted approach requires about 1.8 seconds per iteration, as opposed to 4.0 seconds per iteration for direct matrix multiplication. (These execution times should be compared to those presented in Chapters 4 and 6.) Although the efficiency of the FFT-assisted approach could be improved by using a radix-2 algorithm [50], the approach is obviously not as efficient as expected for matrix equations of relatively small order. Furthermore, use of the FFT algorithm increases storage requirements somewhat. Thus, the FFT-assisted approach may not be beneficial unless the matrix equations of interest are of large order and the additional storage constraints do not exceed the limits of a given computer.

Figure 7.3 shows the maximum magnitudes of the surface current density induced on a large perfectly conducting cylinder by an axially incident plane wave, according to the above integral equations. This result required the solution of a single 399 × 399 system, and the CGM algorithm produced a satisfactory solution after 79 iterations. The current density shows interference effects caused by superimposing interior and exterior currents, as is necessary when using the electric field integral equation to model thin structures. There is no known exact analytical solution for the finite cylinder geometry, and thus the accuracy of the numerical result is not determined.

7.5. Summary

An electric field integral equation representing scattering from hollow, finite-length circular cylinders can be discretized using the DCMoM procedure

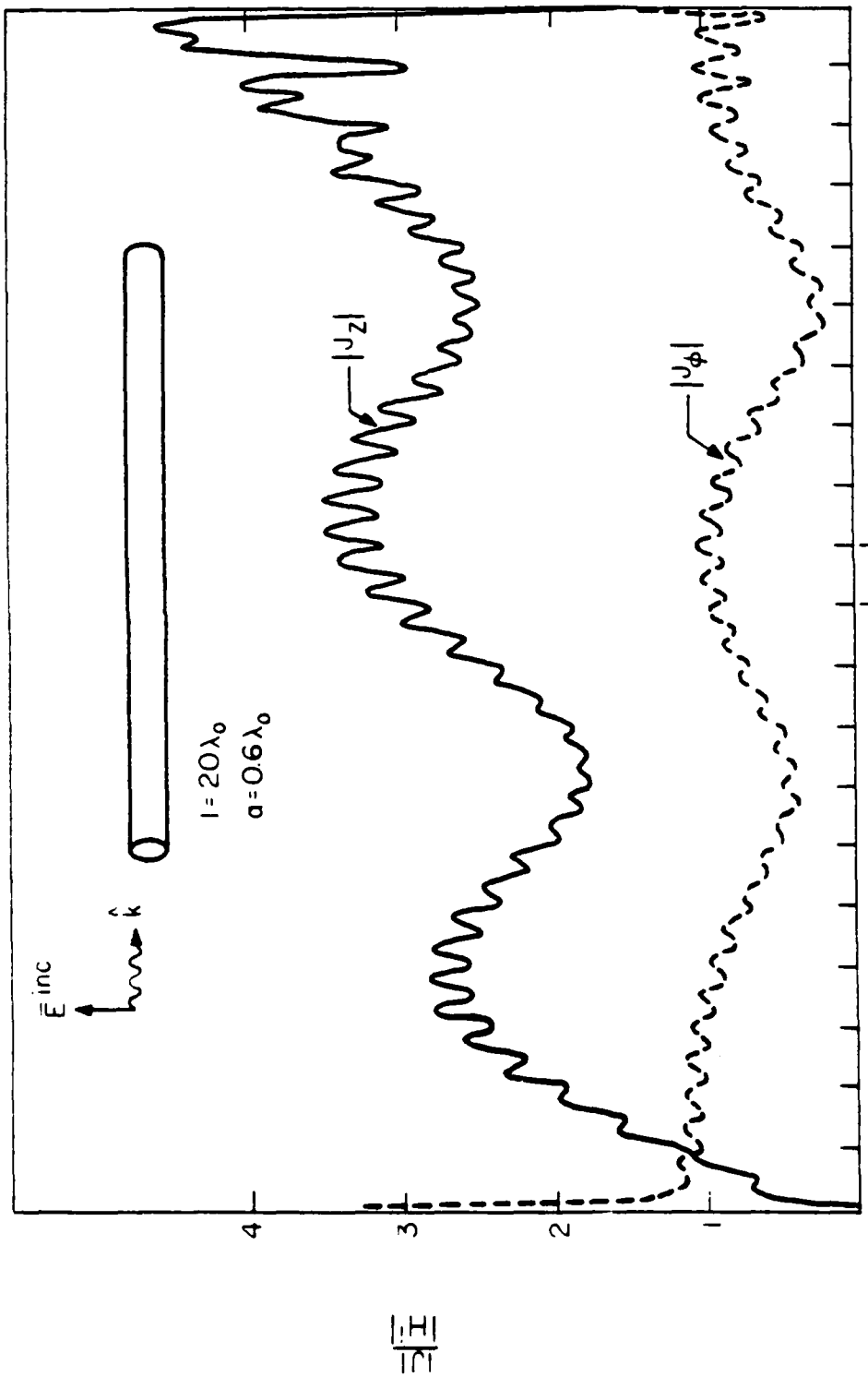

Figure 7.3. Numerical solution for maximum magnitudes of current density induced on a perfectly conducting cylinder.

combined with a "body of revolution" formulation. The system for each Fourier harmonic is readily solved iteratively, so that electrically large cylinders can be treated with minimal storage requirements.

8. SUMMARY AND RECOMMENDATIONS FOR FUTURE WORK

This report describes the application of iterative algorithms to the numerical solution of electromagnetic scattering problems. Iterative algorithms can sometimes be incorporated into conventional techniques in such a manner as to save considerable computer storage and permit the treatment of larger discrete systems within the same storage constraints. Issues addressed in connection with iterative methods include the iterative algorithms available, their performance in theory and in practice, and the various ways in which they can be incorporated into numerical analysis. Three iterative algorithms related to the conjugate gradient method are presented in Chapter 2, and the performance of these algorithms when applied to typical electromagnetic scattering problems is described in Chapter 3. One way of implementing iterative algorithms is the matrix-element regeneration (MER) approach discussed in Chapter 4. The MER requires no special symmetries in the matrix equation, but is not as efficient as approaches based on symmetries. Chapter 5 discusses two discretization procedures for building symmetries into the matrix equations, and detailed examples of this type of approach are presented in Chapters 6 and 7.

Since the goal of iterative methods is numerical efficiency, execution time data have been emphasized where pertinent. Whenever possible, numerical solutions have been compared to exact analytical results for verification, and their accuracy was studied for different models of the same scatterer.

Because of the emphasis on discretizations that build symmetry into the matrix equations, it is apparent that the problems best suited for iterative solution are usually those involving simple geometric shapes. Many additional cases arise in which parts of a given scatterer will conform to the type of geometry easily treated iteratively, and other parts will not. For these

problems, iteration may be efficient if a large enough part of the system matrix can be made to possess the necessary symmetry features. Sometimes, it may be necessary to resort to rather crude modeling in order to achieve this type of symmetry. Depending on the problem under consideration, the MER approach may be a viable alternative to crude modeling techniques.

Although the use of iterative algorithms based on gradient methods has eliminated the problems with diverging numerical solutions that plagued earlier researchers, faster convergence is always a desired goal. In practice, the algorithms described in Chapter 2 may be expensive if applied to large systems. It is expected that future efforts in the fields of engineering and computer science will improve the available algorithms, and improvements could and should be incorporated into existing methods as quickly as they become available. Existing algorithms based on preconditioning the matrix equation require knowledge concerning the eigenvalues of the typical systems that arise. Based upon the performance of the algorithms as illustrated in Chapter 3, there appear to be similarities in the eigenvalue behavior of different scattering problems that could be identified and used in preconditioning algorithms. This may motivate an investigation into the eigenvalue structure of the typical integral equations of electromagnetics.

The topic of multiple right-hand sides remains a vexing problem for electromagnetic scattering problems, and deserves additional attention. However, as discussed in Chapter 3, the goals of fast convergence and ability to treat multiple right-hand sides are at odds, and algorithms well-suited to the first will probably be poor at the second. It appears that attempts to treat multiple right-hand sides should be based upon orthogonal expansions with more flexibility than those produced by the CGM, which are geared to one solution of a given system.

The state-of-the-art techniques for the solution of discrete systems need to reflect current progress in hardware as well as in software. The availability of large blocks of random access memory masquerading as dummy disks and the enhancement of computer architecture to allow larger chunks of directly addressable memory both have an immediate impact on the size problem which can be treated, and the efficiency of direct and iterative algorithms. These issues must be kept in mind when considering a method for a specific problem. Of course, it is expected that the size problem that one would like to be able to solve will always exceed the capacity of existing machines.

It is evident from this work that there are many alternatives and trade-offs to be addressed when considering an iterative computational method for the solution of electromagnetic scattering problems. The character of a specific problem, the computational facilities available, and the desired accuracy in modeling and in the numerical solution will determine the best approach for a given problem.

REFERENCES

[1] P. Moon and D. E. Spencer, Field Theory for Engineers. New York: D. Van Nostrand, 1961.

[2] P. Moon and D. E. Spencer, Field Theory Handbook. New York: Springer-Verlag, 1971.

[3] R. Mittra and S. W. Lee, Analytical Techniques in the Theory of Guided Waves. New York: Macmillan, 1971.

[4] J. C. Maxwell, "A dynamical theory of the electromagnetic field," Phil. Trans. Royal Soc. (London), vol. 155, p. 459, 1864.

[5] L. Cairo and T. Kahan, Variational Techniques in Electromagnetism. London: Blackie & Son, Ltd., 1965.

[6] M. Kline and I. W. Kay, Electromagnetic Theory and Geometrical Optics. New York: Wiley, 1965.

[7] G. L. James, Geometrical Theory of Diffraction for Electromagnetic Waves. New York: Peter Peregrinus Ltd., 1980.

[8] R. Mittra (Ed.), Computer Techniques for Electromagnetics. New York: Pergamon, 1973.

[9] F. L. Ng, "Tabulation of methods for the numerical solution of the hollow waveguide problem," IEEE Trans. Microwave Theory Tech., vol. MTT-22, pp. 322-329, March 1974.

[10] N. N. Bojarski, "K-space formulation of the electromagnetic scattering problem," Tech. Rep. AFAL-TR-71-75, Air Force Avionics Laboratory, Wright-Patterson Air Force Base, Ohio, March 1971.

[11] R. Kastner and R. Mittra, "Spectral-domain iterative techniques for analyzing electromagnetic scattering from arbitrary bodies," Electromagnetics Lab. Tech. Rep. 82-1, Department of Electrical Engineering, University of Illinois, Urbana, IL, 1982.

[12] H. L. Nyo and R. F. Harrington, "The discrete convolution method for solving some large moment matrix equations," Tech. Rep. 21, Department of Electrical and Computer Engineering, Syracuse University, Syracuse, New York, 1983.

[13] M. F. Sultan and R. Mittra, "Iterative methods for analyzing the electromagnetic scattering from dielectric bodies," Electromagnetics Lab. Tech. Rep. 84-4, Department of Electrical and Computer Engineering, University of Illinois, Urbana, IL, 1984.

[14] A. J. Poggio and E. K. Miller, "Integral equation solutions of three-dimensional scattering problems," in Computer Techniques for Electromagnetics, R. Mittra, Ed. New York: Pergamon, 1973.

[15] D. Colton and R. Kress, Integral Equation Methods in Scattering Theory. New York: Wiley, 1983.

[16] M. A. Jaswon and G. T. Symm, Integral Equation Methods in Potential Theory and Elastostatics. New York: Academic, 1977.

[17] K. K. Mei, "Unimoment method of solving antenna and scattering problems," IEEE Trans. Antennas Propagat., vol. AP-22, pp. 760-766, Nov. 1974.

[18] G. D. Smith, Numerical Solution of Partial Differential Equations: Finite Difference Methods. Oxford: Clarendon Press, 1978.

[19] L. E. Elsgolc, Calculus of Variations. London: Pergamon, 1962.

[20] R. Weinstock, Calculus of Variations. New York: Dover, 1974.

[21] L. V. Kantorovich and G. P. Akilov, Functional Analysis. New York: Pergamon, 1982.

[22] M. Becker, The Principles and Applications of Variational Methods. Cambridge: The M.I.T. Press, 1964.

[23] J.H. Richmond, "Digital computer solutions of the rigorous equations for scattering problems," Proc. IEEE, vol. 53, pp. 796-804, Aug. 1965.

[24] R. F. Harrington, "Matrix methods for solving field problems," Proc. IEEE, vol. 55, pp. 136-149, Feb. 1967.

[25] R. F. Harrington, Field Computation by Moment Methods. Malabar, Florida: Krieger, 1982.

[26] R. F. Harrington, "Origin and development of the method of moments for field computation," in Applications of the Method of Moments to Electromagnetic Fields, B. Strait, Ed. St. Cloud, Florida: SCEEE Press, 1981.

[27] R. L. Tanner and M. G. Andreasen, "Numerical solution of electromagnetic problems," IEEE Spectrum, vol. 4, pp. 53-61, Sept. 1967.

[28] R. Mittra (Ed.), Numerical and Asymptotic Techniques in Electromagnetics. New York: Springer-Verlag, 1975.

[29] J. Moore and R. Pizer (Eds.), Moment Methods in Electromagnetics. New York: Wiley, 1984.

[30] A. Jennings, Matrix Computation for Engineers and Scientists. New York: Wiley, 1977.

[31] G. H. Golub and C. F. Van Loan, Matrix Computations. Baltimore: The Johns Hopkins University Press, 1983.

[32] R. S. Varga, Matrix Iterative Analysis. Englewood Cliffs: Prentice-Hall, 1962.

[33] H. R. Schwartz, Numerical Analysis of Symmetric Matrices. Englewood Cliffs: Prentice-Hall, 1973.

[34] T. K. Sarkar, K. R. Siarkiewicz, and R. F. Stratton, "Survey of numerical methods for solution of large systems of linear equations for electromagnetic field problems," IEEE Trans. Antennas Propagat., vol. AP-29, pp. 847-856, Nov. 1981.

[35] A. Taflove and M. E. Brodwin, "Numerical solution of steady-state electromagnetic scattering problems using the time-dependent Maxwell's equations," IEEE Trans. Microwave Theory Tech., vol. MTT-23, pp. 623-630, Aug. 1975.

[36] 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., vol. MTT-23, pp. 888-896, Nov. 1975.

[37] A. Taflove and K. Umashankar, "A hybrid moment method/finite difference time domain approach to electromagnetic coupling and aperture penetration into complex geometries," IEEE Trans. Antennas Propagat., vol. AP-30, pp. 617-627, July 1982.

[38] N. N. Bojarski, "The k-space formulation of the scattering problem in the time domain," J. Acoust. Soc. Am., vol. 72, part II, pp. 570-584, Aug. 1982.

[39] R. Mittra, "Integral equation methods for transient scattering" in Transient Electromagnetic Fields, L. B. Felsen, Ed. New York: Springer-Verlag, 1976.

[40] C. L. Bennett and G. F. Ross, "Time-domain electromagnetics and its applications," Proc. IEEE, vol. 66, pp. 299-318, Mar. 1978.

[41] E. K. Miller and J. A. Landt, "Direct time-domain techniques for transient radiation and scattering from wires," Proc. IEEE, vol. 68, pp. 1396-1423, Nov. 1980.

[42] E. O. Brigham, The Fast Fourier Transform. Englewood Cliffs: Prentice-Hall, 1974.

[43] A. V. Oppenheim and R. W. Schafer, Digital Signal Processing. Englewood Cliffs: Prentice-Hall, 1975.

[44] W. L. Ko and R. Mittra, "A new approach based on a combination of integral equation and asymptotic techniques for solving electromagnetic scattering problems," IEEE Trans. Antennas Propagat., vol. AP-25, pp. 187-197, March 1977.

[45] R. Kastner and R. Mittra, "A spectral-iteration technique for analyzing scattering from arbitrary bodies, Part I: Cylindrical scatterers with E-wave incidence," IEEE Trans. Antennas Propagat., vol. AP-31, pp. 499-506, May 1983. "Part II: Conducting cylinders with H-wave incidence," IEEE Trans. Antennas Propagat., vol. AP-31, pp. 535-537, May 1983.

[46] R. Kastner and R. Mittra, "A new stacked two-dimensional spectral iterative technique (SIT) for analyzing microwave power deposition in biological media," IEEE Trans. Microwave Theory Tech., vol. MTT-31, pp. 898-904, Nov. 1983.

[47] H. L. Nyo and R. F. Harrington, "Solution of some additional electromagnetic problems by the discrete convolution method," Tech. Rep. No. 23, Department of Electrical and Computer Engineering, Syracuse University, Syracuse, NY, Feb. 1984.

[48] D. T. Borup and O. P. Gandhi, "Fast-Fourier transform method for calculation of SAR distributions in finely discretized inhomogeneous models of biological bodies," IEEE Trans. Microwave Theory Tech., vol. MTT-32, pp. 355-360, Apr. 1984.

[49] D. T. Borup and O. P. Gandhi, "Calculation of high-resolution SAR distributions in biological bodies using the FFT algorithm and conjugate gradient method," IEEE Trans. Microwave Theory Tech., vol. MTT-33, pp. 417-419, May 1985.

[50] S. Ray and R. Mittra, "Spectral-iterative analysis of electromagnetic radiation and scattering problems," Electromagnetic Communication Laboratory Tech. Rep. 84-11, Department of Electrical and Computer Engineering, University of Illinois, Urbana, IL, July 1984.

[51] M. Hurst and R. Mittra, "Scattering center analysis for radar cross section modification," Electromagnetic Communication Laboratory Tech. Rep. 84-12, Department of Electrical and Computer Engineering, University of Illinois, Urbana, IL., July 1984.

[52] P. M. van den Berg, "Iterative computational techniques in scattering based upon the integrated square error criterion," IEEE Trans. Antennas Propagat., vol. AP-32, pp. 1063-1071, Oct. 1984.

[53] S. A. Bokhari and N. Balakrishnan, "Analysis of cylindrical antennas - a spectral iteration technique," IEEE Trans. Antennas Propagat., vol. AP-33, pp. 251-258, Mar. 1985.

[54] D. C. Farden, "Solution of a Toeplitz set of linear equations," IEEE Trans. Antennas Propagat., vol. AP-24, pp. 906-907, Nov. 1976.

[55] W. L. Stutzman and G. A. Thiele, Antenna Theory and Design. New York: Wiley, 1981, pp. 577-591.

[56] A. F. Peterson, M. F. Sultan, and R. Mittra, "Solution of large-body scattering problems using the method of conjugate gradients," in Digest of 1984 IEEE Antennas and Propagation Society International Symposium. New York: Institute of Electrical and Electronics Engineers, pp. 351-354, 1984.

[57] A. F. Peterson and R. Mittra, "On the method of conjugate gradients for scattering by PEC cylinders," Electromagnetics Laboratory Tech. Rep. 84-3, Department of Electrical Engineering, University of Illinois, Urbana, IL, Jan. 1984.

[58] M. F. Sultan and R. Mittra, "An iterative moment method for analyzing the electromagnetic field distribution inside inhomogeneous lossy dielectric objects," IEEE Trans. Microwave Theory Tech., vol. MTT-33, pp. 163-168, Feb. 1985.

[59] A. F. Peterson and R. Mittra, "Method of conjugate gradients for the numerical solution of large-body electromagnetic scattering problems," J. Opt. Soc. Amer. A, vol. 2, pp. 971-977, June 1985.

[60] R. Mittra and C. H. Chan, "Iterative solution of the electromagnetic boundary value problems in the spectral domain," Digest of the 1985 North American Radio Science Meeting, International Union of Radio Science, p. 102, June 1985.

[61] A. F. Peterson and R. Mittra, "The method of conjugate gradients for iterative solution of matrix and operator equations," Electromagnetic Communication Laboratory Tech. Rep. 84-6, Department of Electrical and Computer Engineering, University of Illinois, Urbana, IL, Mar. 1984.

[62] A. F. Peterson and R. Mittra, "The method of conjugate gradients for the numerical solution of large-body electromagnetic scattering problems," Electromagnetic Communication Laboratory Tech. Rep. 85-1, Department of Electrical and Computer Engineering, University of Illinois, Urbana, IL, Jan. 1985.

[63] M. R. Hestenes and E. Stiefel, "Methods of conjugate gradients for solving linear systems," J. Res. Nat. Bur. Stand., vol. 49, pp. 409-435, 1952.

[64] F. S. Beckman, "The solution of linear equations by the conjugate gradient method," in Mathematical Methods for Digital Computers, A. Ralston and H. S. Wilf, Eds. New York: Wiley, 1960.

[65] D. H. Griffel, Applied Functional Analysis. New York: Wiley, 1981.

[66] R. Fletcher and C. M. Reeves, "Functional minimization by conjugate gradients," Computer J., vol. 7, pp. 149-154, 1964.

[67] J. W. Daniel, The Approximate Minimization of Functionals. Englewood Clifs: Prentice-Hall, 1971.

[68] D. R. Wilton and C. M. Butler, "Effective methods for solving integral and integro-differential equations," Electromagnetics, vol. 1, pp. 289-308, 1981.

[69] J. Stoer, "Solution of large linear systems of equations by conjugate gradient type methods," in Mathematical Programming: The State of the Art, A. Bachem, M. Grötschel, B. Korte, Eds. New York: Springer-Verlag, 1983.

[70] H. D. Simon, "The Lanczos algorithm with partial reorthogonalization," Mathematics of Computation, vol. 42, pp. 115-142, Jan. 1984.

[71] C. C. Paige and M. A. Saunders, "A bidiagonalization algorithm for sparse linear equations and least-square problems," Systems Optimization Laboratory Tech. Rep. 78-19, Stanford University, Stanford, CA, Oct. 1978.

[72] A. Bjorck and T. Elfving, "Accelerated projection methods for computing pseudoinverse solutions of systems of linear equations," B.I.T., vol. 19, pp. 145-163, 1979.

[73] M. A. Saunders, H. D. Simon, E. L. Yip, "Two conjugate-gradient type methods for sparse unsymmetric linear equations," Boeing Computer Services Tech. Rep. ETA-TR-18, Boeing Computer Services, Seattle, WA, June 1984.

[74] R. F. Harrington, Field Computation by Moment Methods. Malabar, Florida: Krieger, 1982, pp. 42-46.

[75] J. C. Bolomey and W. Tabbara, "Numerical aspects on coupling between complementary boundary value problems," IEEE Trans. Antennas Propagat., vol. AP-21, pp. 356-363, May 1973.

[76] R. Mittra and C. A. Klein, "Stability and convergence of moment method solutions," in Numerical and Asymptotic Techniques in Electromagnetics, R. Mittra, Ed. New York: Springer-Verlag, 1975.

[77] J. R. Mautz and R. F. Harrington, "H-field, E-field, and combined field solution for conducting bodies of revolution," A.E.U., vol. 32, pp. 157-163, 1978.

[78] A. D. Yaghjian, "Augmented electric and magnetic field equations," Radio Sci., vol. 16, pp. 987-1001, 1981.

[79] J. H. Richmond, "Scattering by a dielectric cylinder of arbitrary cross-section shape," IEEE Trans. Antennas Propagat., vol. AP-13, pp. 334-341, May 1965.

[80] J. H. Richmond, "TE-wave scattering by a dielectric cylinder of arbitrary cross-section shape," IEEE Trans. Antennas Propagat., vol. AP-14, pp. 460-464, July 1966.

[81] H. Massoudi, C. H. Durney, and M. F. Iskander, "Limitations of the cubical block model of man in calculating SAR distributions," IEEE Trans. Microwave Theory Tech., vol. MTT-32, pp. 746-751, Aug. 1984.

[82] A. F. Peterson, "A study of transverse electric wave scattering by dielectric cylinders," M.S. Thesis, Department of Electrical an Computer Engineering, University of Illinois, Urbana, IL, April 1983.

[83] R. F. Harrington, Field Computation by Moment Methods. Malabar, Florida: Krieger, 1982, pp. 42-46.

[84] D. H. Pries, "The Toeplitz matrix: Its occurrence in antenna problems and a rapid inversion algorithm," IEEE Trans. Antennas Propagat., vol. AP-20, pp. 204-206, March 1972.

[85] H. Akaike, "Block-Toeplitz matrix inversion," SIAM J. Appl. Math., vol. 24, pp. 234-241, March 1973.

[86] R. Kumar, "A fast algorithm for solving a Toeplitz system of equations," IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-33, pp. 254-267, Feb. 1985.

[87] T. Kitazawa and R. Mittra, "Analysis of asymmetric coupled striplines," IEEE Trans. Microwave Theory Tech., vol. MTT-33, pp. 643-646, July 1985.

[88] P. Silvester, "TEM wave properties of microstrip transmission lines," Proc. IEE, vol. 115, pp. 43-48, Jan. 1968.

[89] T. Itoh, "Spectral domain immitance approach for dispersion characteristics of generalized printed tansmission lines," IEEE Trans. Microwave Theory Tech., vol. MTT-28, pp. 733-736, July 1980.

[90] S. M. Wright, "Efficient analysis of infinite microstrip arrays on electrically thick substrates," Ph.D. dissertation, University of Illinois, Urbana, IL, 1984.

[91] C. H. Tsao and R. Mittra, "A spectral iteration approach for analyzing scattering from frequency selective surfaces," IEEE Trans. Antennas Propagat., vol. AP-30, pp. 303-308, March 1982.

[92] S. Ray and R. Mittra, "Spectral-iterative analysis of radiation and scattering problems," Electromagnetic Communication Laboratory Tech. Rep. 84-11, Department of Electrical and Computer Engineering, University of Illinois, Urbana, IL, July 1984, p. 16.

[93] T. Cwik and R. Mittra, "Scattering from frequency selective screens," Electromagnetics, vol. 5, Feb. 1986 (to appear).

[94] R. F. Harrington, Time-harmonic Electromagnetic Fields. New York: McGraw Hill, 1961, p. 261.

[95] M. G. Andreasen, "Scattering from bodies of revolution," IEEE Trans. Antennas Propagat., vol. AP-13, pp. 303-310, March 1965.

[96] J. R. Mautz and R. F. Harrington, "Radiation and scattering from bodies of revolution," Appl. Sci. Res., vol. 20, pp. 405-435, June 1969.

[97] A. W. Glisson and D. R. Wilton, "Simple and effective numerical techniques for treating bodies of revolution," University of Mississippi, RADC-TR-79-22, Rome Air Development Center, Griffiss AFB, NY, March 1979.

[98] E. C. Jordan and K. G. Balmain, Electromagnetic Waves and Radiating Systems. Englewood Cliffs: Prentice-Hall, Inc., 1968, p. 153.

[99] K. M. Mitzner, "Effective boundary conditions for reflection and transmission by an absorbing shell of arbitrary shape," IEEE Trans. Antennas Propagat., vol. AP-16, pp. 706-712, Nov. 1968.

[100] T. B. A. Senior, "Scattering by resistive strips," Radio Sci., vol. 14, pp. 911-924, Sept. 1979.