Method of Moments as Applied to Electromagnetic Problems

MICHEL M. NEY, MEMBER, IEEE

(*Invited Paper*)

Abstract—**This paper reviews one of the most important general methods for solving electromagnetic-field problems, namely, the moment method. It begins with a brief mathematical foundation of the general method. Then, the various specializations are described, accompanied with relevant references to illustrate the pitfalls and shortcomings, as well as the advantages, as compared to other methods. Deterministic and eigenvalue problems are both discussed separately. Finally, two advanced techniques which have been found to be among the most efficient ones for solving matrix equations resulting from the moment method, namely, the conjugate gradient and the pseudo-inverse, are described. A version of their algorithm which is easily programmable on computer is also presented.**

I. Introduction

WITH THE EVER-INCREASING complexity of communication systems, there has been a need for engineers to predict the behavior of such systems by means of computer simulations. As a result, the equations involved in the mathematical description of electromagnetic quantities, which are of interest in most cases, have become more complex. Consequently, more sophisticated numerical methods have been developed to solve electromagnetic problems. These methods are gaining greater and greater success with the constant development of new powerful digital computers.

A theory is meant to extrapolate observations in order to make some prediction. As far as engineers are concerned, a theory is relevant if it can produce numbers in a finite number of steps performed in a reasonable period of time and with sufficient accuracy, taking into account the fact that computers have finite word length. Very elegant theories have been known for decades but had been useless for engineers owing to the lack of appropriate numerical algorithms to produce accurate numbers.

In recent years, most of those theories have received renewed interest with the developments of powerful high-speed computers which have made their numerical solutions within reach. Simultaneously, more sophisticated algorithms have been developed to obtain more accurate numbers while decreasing the number of operations.

Generally, before producing numbers, two steps are necessary. First, the original functional equations must be transformed into structures (such as systems of equations) that can be handled by computers. Then, appropriate techniques must be used to solve numerically the new form of equations hence produced. The degrees of difficulty that may be encountered in both steps are somehow interdependent. Indeed, if an efficient algorithm that transforms the functional equations is found, the size of the systems to be solved may be considerably decreased. However, efficient algorithms may require more computer time to numerically determine the various coefficients of the new form of equations.

This paper is a critical inspection of a general method used extensively in electromagnetic-field problems, namely, the method of moments. Shortcomings and pitfalls of the method will be discussed and some possible improvements proposed. New methods which are gaining interest to solve linear systems of equations that follow from the method of moments will be outlined, as solutions to such systems are an important step towards the production of numbers which are, probably, of most interest for engineers.

This paper is intended as an introduction for those totally unfamiliar with numerical methods, as well as for those with some experience in the field. It was felt by the author that it was more advisable to discuss the different specializations of the general method of moments, rather than to survey specific examples. Basic mathematical concepts are introduced to provide some help for further reading of the relevant literature. References found to be most useful for the specific examples are listed in the bibliography and the reader is urged to consult the papers relevant to his or her interest.

II. Method of Moments

Most of the solutions to linear functional equations can be interpreted in terms of projections onto subspaces of functional spaces. For computational reasons, these subspaces must be finite dimensional. For theoretical work, they may be infinite dimensional. The idea of transforming linear functional equations to linear matrix equations is rather old. Galerkin, a Russian engineer, developed the method, which carries his name, around 1920. It was a specialization of the more general method of moments which was presented later by R. F. Harrington in 1967 [1].

Most of the electromagnetic problems can be expressed under the form of a linear functional equation. One gener-

Manuscript received February 22, 1985; revised June 3, 1985.

The author is with the Department of Electrical Engineering, University of Ottawa, Ottawa, Ontario, K1N 6N5, Canada.

0018-9480/85/1000-0972$01.00 ©1985 IEEE

ally classifies electromagnetic problems in two categories, namely, *deterministic* and *eigenvalue* problems. In the first category, the linear functional equation enables one to determine the electromagnetic quantity directly. In the second category, parameters for which nontrivial solutions exist are found first. Then, the corresponding solutions called eigensolutions are determined. Both categories of problems can be handled by the method of moments.

A. Deterministic Problems

First, consider a deterministic problem for which the corresponding functional is given by

$$L \cdot f = y \tag{1}$$

where L is any linear operator, f is the unknown function to be determined, and y is the input also called the excitation. The space spanned by all functions resulting from the operation L is called the *range* of L. The set of all functions on which L can operate define the *domain* of L.

One needs an inner product $\langle\ \rangle$ associated with the problem, which must satisfy

$$\begin{aligned} \langle u, v\rangle &= \langle v, u\rangle^* \\ \langle \alpha u + \beta v, f\rangle &= \alpha\langle u, f\rangle + \beta\langle v, f\rangle \\ \langle f, f\rangle &\geq 0 \\ \text{if } \langle f, f\rangle &= 0 \text{ then } f = 0 \end{aligned} \tag{2}$$

where α and β are scalars, f, u, v any functions, and * denotes the complex conjugate. For instance, a suitable inner product for function spaces can be given by the functional

$$\langle u, v\rangle = \int_\Omega uv^* \, d\Omega \tag{3}$$

The above integral is performed over any N-dimensional subspace, depending on the application. Equation (3) is called an unweighted or standard inner product and it indicates a "projection" of u in the "direction" of v, from which the similarity between vector and function space becomes obvious. In some situations, one needs the *adjoint operator* of L and its domain defined by

$$\langle Lf, y\rangle = \langle f, L^a y\rangle \tag{4}$$

for all f in the domain of L. An operator is *self-adjoint* if $L^a = L$ and the domain of L^a is that of L. Self-adjointness depends strongly on the associated boundary conditions and also on the selection of an appropriate inner product. For instance, self-adjointness is largely determined by the boundary conditions in the case of differential operators [2]. However, for integral operators, self-adjointness is assured if the kernel of the integral possesses some symmetric properties [3].

The operator L is said to be positive/negative definite if

$$\langle Lf, f\rangle \gtrless 0 \tag{5}$$

for any $f \neq 0$ in the domain of L. The properties of the solution of (1) depend strongly on the properties of the operator L. For instance, if the operator is positive definite, the solution of (1) is unique. Indeed, suppose that u and v are two solutions of (1) such that $L \cdot u = y$ and $L \cdot v = y$. Then, by virtue of the linearity of L, $w = u - v$ is also a solution. Therefore, $Lw = 0$, and, since L is positive definite, w must be zero, yielding $u = v$.

Equations such as (1) can be analytically solved in a very few cases. Most of the time, they require methods that transform the original equation in the form of linear equation systems. The most well known are variational, finite difference, and moment methods. The first two methods are important, and, in certain cases, may have some advantage. However, for the sake of consistency, they will not be discussed here.

First of all, let us express the unknown function in terms of *basis* or *expansion* functions f_j in the domain of L

$$f = \sum_N \alpha_j f_j. \tag{6}$$

The set of basis functions can be finite or infinite. In the latter case, since in practical problems the summation must be truncated, the solution will be an approximation of the true solution. This is the case for orthogonal developments such as Fourier series. Using the property of linearity of L, (1) can now be written as

$$\sum_N \alpha_j L f_j = y. \tag{7}$$

If a set of *weighting* or *testing* functions w_i is chosen in *the range of the operator* L and the inner product of both sides of (7) is taken for each w_i, the original functional equation becomes a set of linear equations that can be written in the matrix form

$$[L]\vec{\alpha} = \vec{y}$$

where

$$\begin{aligned} L_{ij} &= \langle w_i, Lf_j\rangle \\ y_i &= \langle w_i, y\rangle \end{aligned} \tag{8}$$

and

$$\vec{\alpha} = [\alpha_1, \alpha_2, \cdots, \alpha_j \cdots]^T$$

in which T indicates transposition. If the matrix $[L]$ is regular, $[L]^{-1}$ exists, and the α_i's are given by

$$\vec{\alpha} = [L]^{-1}\vec{y} \tag{9}$$

and the solution is found using (6). The moment method can be interpreted as an error-minimization procedure with the concept of linear spaces. Let $R(L)$ be the range of the operator L. The right-hand side of (8) is the orthogonal projection of the subspace of $R(L)$ spanned by the operation of L on the exact solution f, i.e., the y, onto the subspace W spanned by the w_i's. The left-hand side of (8) is the projection of the subspace spanned by the operations Lf_j onto W. The moment method equates these two projections (see Fig. 1). Since the error (also called weighted residual) is orthogonal to the projection, it is of the second order and, consequently, the method is an error-minimization procedure.

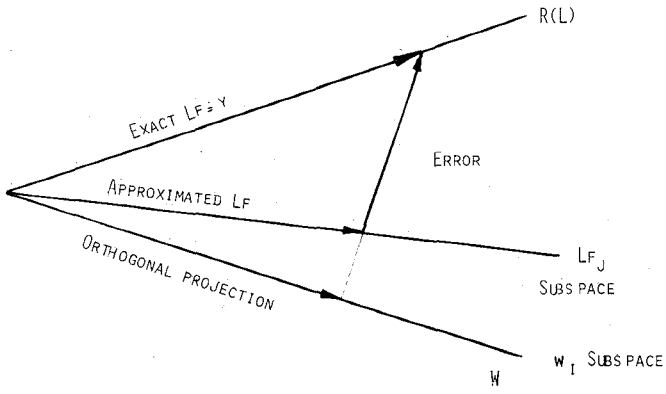


Fig. 1. Illustration of the method of moments in the function space.

There are infinitely many possible sets of basis and weighting functions. The most important task of the engineer for any particular problem is the selection of an appropriate set of f_j's and w_i's. Although the choice of these functions is specific to each problem, one can state rules that can be applied generally to optimize the chance of success by obtaining accurate results in a minimum time and computer memory storage. First of all, they should form a set of linearly independent functions. Second, using (6), the f_j's should approximate the (expected) function f reasonably well. Finally, the w_i's should be in $R(L)$ and so chosen that the inner products $\langle w_i, y \rangle$ depend on relatively independent properties of y. Some additional factors may influence the selection, such as

i) the desired accuracy of the solution,
ii) the size of the matrix $[L]$ to be inverted,
iii) the realization of a well-behaved matrix $[L]$,
iv) the ease of evaluating of the inner products.

The various selections of the w_i's lead to the different specializations of the moment method.

1) Galerkin's Method: In cases where the domain of L is identical to the domain of L^a, one can select $w_i = f_j$, which leads to the well-known Galerkin's method. For self-adjoint operators, the condition is automatically met and they are best suited for this method because, according to (4), the resulting matrix $[L]$ is symmetrical. This may have some numerical advantage for solving the corresponding linear system of equations. However, the elements of the matrix $[L]$ can be more difficult to evaluate than in other methods. This may outweigh the advantage of having a symmetric matrix $[L]$.

The Galerkin's method has been used extensively in electromagnetic problems. Numerous examples of application are given in [4]. This method has been found to yield accurate results with rapid convergence, as compared to others, in the case of low-order solutions, i.e., when few expansion functions are required.

Galerkin's method is also involved in a new method which is gaining interest in transmission-line problems, namely, the spectral-domain method [5]. In this approach, the coupled integral equations, relating field and current, which typically appear in the space domain, are expressed in the spectral domain via Fourier transform. As a result, the original equations are transformed into algebraic equations and convolutions into simple products. The application of the boundary conditions yields a system of linear algebraic equations relating the Fourier transform of both unknowns, respectively, the electric field at dielectric interface, and the current densities in the conducting strips. Finally, the application of the Galerkin's method in the spectral domain produces an homogeneous eigenvalue matrix equation. Equating the determinant of the matrix to zero leads to the solution of the propagation constants for the dominant and higher order modes [6].

2) Subsection and Point-Matching Method: For higher order solutions, i.e., when a large number of expansion functions is required to approximate the unknown function, there may be a certain advantage in using weighting functions that render the inner products, involved in the moment method, easy to determine. This is achieved by choosing the w_i's equal to Dirac's functions. Indeed, by virtue of the property of this function, the inner product which involves an integral in the function spaces, becomes trivial. This specialization of the moment method is called the *point matching* or *collocation* method. The elements of the matrix $[L]$ and the vector $\vec{y}$, hence, become

$$l_{ij} = Lf_j|_{\vec{r}=\vec{r}_i}$$
$$y_i = f_j(\vec{r}_i). \tag{10}$$

This is equivalent to enforce (7) at various points of interest, generally where boundary conditions must be met. The main advantage of this method resides in the ease with which the matrix elements are computed as compared to other specializations of the moment method. The major disadvantage is that for low-order solutions, the accuracy and the convergence of the solution generally depend on the location of the points at which (7) is matched (see, for example, [7]). For higher order solutions, Galerkin's method has been found to give better results and faster convergence in the majority of cases. However, equidistant points, in this case, give satisfactory results for the point-matching method.

Another important aspect is that the point-matching method has been proved inaccurate when the operations Lf_j yield symbolic functions such as Dirac's functions. This can be explained by the fact that the inner product of two distributions such as Dirac's functions is not defined [8]. On the other hand, integral operators do not produce such functions for practical problems. Consequently, they are better suited to be used with the point-matching method. A typical illustration is found in [9] in which the method is used for finding scattering fields produced by infinitely long dielectric cylinders with transverse magnetic irradiation. The operator is integral and the method yields excellent results. On the other hand, for transverse electric irradiation, the operator involved is integro-differential and, since pulse functions as basis are used, the results are less accurate [10].

For problems lacking symmetry, it is difficult to find basis functions that are defined over the entire domain of the solution, and they would imply rather involved calcula-

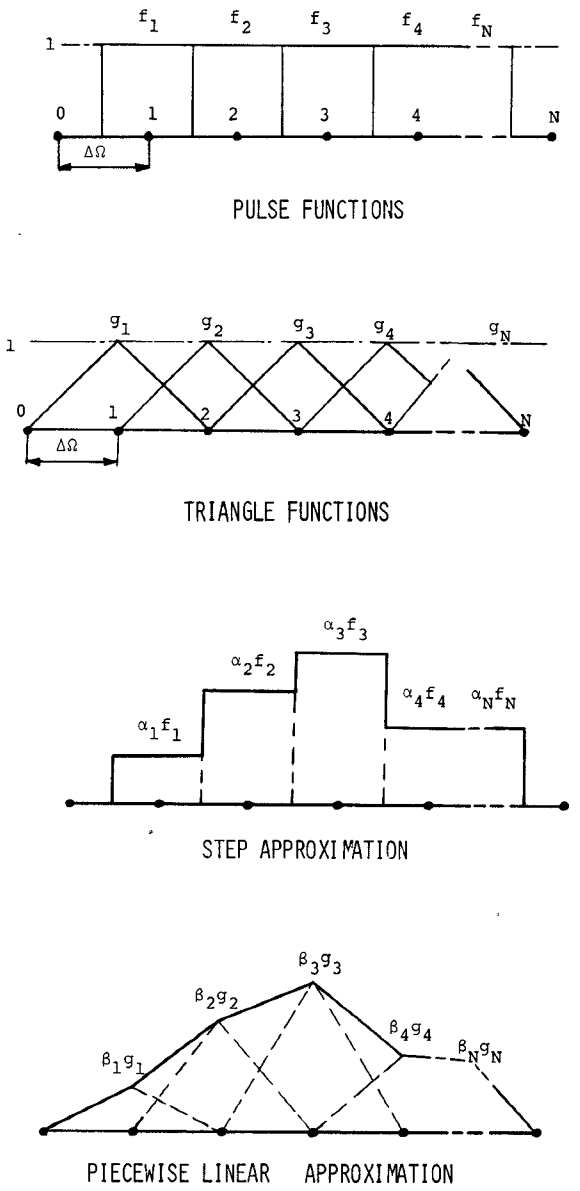


Fig. 2. Examples of subsectional basis and corresponding approximated function.

tions of the inner products. Consequently, it is more judicious to approximate the solution by basis functions which are defined only over subsections of the domain of f. The point-matching method is often used in conjunction with the subsectional basis. The main advantage is a certain facility to compute the elements of the matrix $[L]$ provided that the basis function be simple. For instance, the unknown function can be approximated using step or triangular functions (see Fig. 2). Care must be taken regarding the subsectional basis. The basis functions y should be in the domain of the operator L. In addition, the point-matching procedure should not be used if the operation of L on the f_j's yields symbolic functions, for reasons stated before. For instance, step functions should not be used at all when second-order differential operators are involved. Also, they should not be used for first-order differential operators if point-matching is intended.

The method would be very limited if such simple basis functions could not be used for this type of operators, which are often encountered in electromagnetic problems. Fortunately, the differential operators can always be approximated using a finite-difference procedure which is discussed in the next section. Another possibility is to extend the original domain of the operator such that it can operate on a wider class of functions without changing the operation in the original domain [4]. For problems involving integral operators, the method yields reasonably good results. However, the advantage inherent to the method described above can be outweighed by the size of the system. Indeed, in order to approximate the solution reasonably well, a rather large number of subsections is necessary. Consequently, the method is difficult to be applied in cases for which the wavelength is relatively short as compared to the dimension of the problem.

Many applications that use point-matching in conjunction with the subsection method can be found in the literature [4], [11]–[14]. They most frequently deal with scattering problems. For instance, an interesting comparison between Galerkin's and the point-matching methods for different choices of basis functions can be found in [11] for microstrip antenna problems. It is found that triangular expansions do not give significant improvement as compared to pulse functions. Three-dimensional problems can also be approached with the point-matching and subsection methods [12]. The electric field is expended by 3-D pulses (block model) inside lossy dielectric bodies and point-matching is applied within each subvolume. The method is limited to relatively low frequencies. Surface patch models of conducting objects are proposed in [13] for determining scattering fields of metallic objects with arbitrary shape. The use of triangular patches circumvents the existence of a fictitious line or point of charges at the edges of the subsections [14]. Indeed, by virtue of the continuity equation, line or point charges may appear at the limit of subsections if rectangular pulses are used for current expansion. Potential and field are not defined at those locations, and anomalies or inconsistencies usually appear in the solution. Point-matching used with appropriate basis functions can significantly decrease the order of the system to be solved. Hagman *et al.* [15] proposed a plane-wave correction for scattering problems. The order of the matrix is reduced by one order of magnitude as compared with simple pulse functions presented in [9]. However, the computations of the matrix elements are much more involved.

There are various techniques which have been proposed to improve the point-matching and subsection methods. For instance, in inverse scattering problems, it has been found that the solution is highly sensitive to the points of match [16]. In addition, the problem was found to be ill-posed. One possibility to render the solutions more accurate and less sensitive to the location of the matching points is to use least-square techniques [17]. Some constraints are enforced on the solution, and fairly good solutions are obtained even in the presence of realistic levels of noise in the input function of the integral equation. Another possibility is to match the equation at a larger number of points than required by the number of unknowns and to use least-squares optimization techniques

to solve the matrix equation [18]. A theoretical treatment of the so-called overdetermined collocation can be found in [19].

3) Approximate Operators: It is sometimes convenient to approximate the operator L. For instance, differential operators can be replaced by finite-difference operators [2], [4], [20]. This procedure is very useful for point-matching since it allows one to use discontinuous basis functions, such as pulses, for problems involving differential operators. The major inconvenience with the finite-difference approximation is that, for a given approximation of the operator, the solution does not converge when the distance between the mesh nodes decreases. In fact, the corresponding matrix becomes singular [21]. One possibility to obtain a faster convergence and better accuracy without increasing the number of meshes is to retain higher order terms in the finite-difference approximation. This automatically implies more computation for the matrix elements. In practical problems, it is convenient to make the nodes coincide with the boundaries on which the potential and its derivative are known. To achieve this, mesh refinement is required to a level at which the approximation of the boundaries does not affect the solution. Another way is to use irregularly spaced nodes [21], [22].

For integral operators, it is sometimes more convenient to approximate the Kernel of the integral operator [23], [24] so that analytical integration can take place. If this cannot be done, since the elements of $[L]$ generally involve the integral of Green's functions which behave poorly at the origin, numerical methods of integrations such as singularity extraction or statistical Riemann method of integration are necessary to evaluate the diagonal matrix elements. These methods are time consuming if a significant number of points are required, and eventually will not converge. They will not be discussed here.

4) Other Specializations: There are other selections of w_i's and f_j's which have been used. For instance, step functions for testing and triangular functions for expansion functions were used for microstrip antenna problems [11]. Results showed that no significant improvement can be observed as compared to Galerkin's method in which pulses were used for testing and expansion. Therefore, it is not always judicious to use more elaborate basis functions because the computations of the matrix elements may introduce more errors and require more computer time.

Among other possibilities, choosing $w_i = Lf_j$ yields the method of the minimum residual called more commonly the least-squares method. It can be readily seen from Fig. 1 that the error vector (residual) has a minimum norm when it is orthogonal to the space spanned by the Lf_j's.

B. Eigenvalue Problems

An eigenvalue equation is an homogeneous equation which can be written in the general linear case as

$$Lf = \lambda Mf \tag{11}$$

where L and M are linear operators. A solution of (11) exists only for particular values of λ called eigenvalues, associated with the corresponding solutions called eigenfunctions, eigenvectors, or eigensolutions. The method of moments transforms (11) to a matrix eigenvalue equation that can be solved by appropriate methods. Eigenvalue problems are important in electromagnetics. Indeed, the eigenvalues correspond to physical quantities which are of major importance for engineers, such as cutoff or resonance frequencies of a system. However, the numerical solution of the matrix eigenvalue equations is, in general, more complicated than that for deterministic matrix equations. Iterative schemes such as Jacobi method are most commonly used [25].

Using the same procedures as described for deterministic equations, (11) can be written as

$$\sum_N \alpha_i Lf_i = \lambda \sum_N \alpha_i Mf_i \tag{12}$$

where the f_j's are in the domain of the operators L and M. In a similar manner, a set of weighting functions w_i is chosen in the range of L and M and the inner product of (12) is taken for each w_i yielding the matrix system

$$[L]\vec{\alpha} = \lambda [M]\vec{\alpha} \tag{13}$$

where

$$m_{ij} = \langle w_i, Mf_j \rangle.$$

The above system can have a nontrivial solution only if

$$\det \|[L] - \lambda [M]\| = 0. \tag{14}$$

The determinant (14) is a polynomial in which roots $\lambda_1, \lambda_2, \cdots$, correspond to the eigenvalues of the matrix equation (13). They approximate the eigenvalues of the original functional equation (11). The corresponding vectors with coefficients $\alpha_{1n}, \alpha_{2n}, \cdots$, are the eigenvectors of the matrix equation (13) and

$$f_n = \sum_N \alpha_{in} f_i \tag{15}$$

approximates each eigenvector of the original functional equation (11).

If M possesses an inverse, (11) can be written in the canonical form

$$M^{-1}Lf = \lambda f. \tag{16}$$

Thus, the matrix $[L]$ must be multiplied by $[M]^{-1}$ before applying the method of solutions for eigenvalue matrix equations. It is worth noticing that if M and M^{-1} are the identity operators in (11) and (16), $[M]^{-1}$ is not the identity matrix in (13) and (14). Indeed, the elements of $[M]$ involve the scalar product of the weighting functions and the basis functions.

A judicious choice of w_i's is to select $w_i = f_j$ (Galerkin's procedure). It has been found that Galerkin's solutions give eigenvalues higher than the exact values for second-order differential operators, while they give smaller eigenvalues for first-order differential operators [4]. Like deterministic problems, it is sometimes convenient to extend the operator. However, when this is applied to eigenvalue problems, extraneous eigenvalues appear if the

basis functions associated with the extended operator violate the boundary conditions of the problem. Fortunately, there are several factors which make extraneous eigenvalues easily recognizable. First of all, they do not converge like the other eigenvalues. Then, even if the original operator is positive definite, they may have negative values because the extended operator is not necessarily positive definite and, finally, the corresponding eigenvectors tend to be irregular and do not generate eigenfunctions which respect the boundary conditions of the problem.

III. New Methods for Solving Matrix Equations

The generation of a linear matrix equation by using any specialization of the moment method is only a step towards the production of numbers which are of most interest for engineers. The last, but not necessarily the least, task is to solve numerically the matrix equation. For small-order well-conditioned systems, classical techniques such as Gauss, diagonal decomposition, and linear iterative techniques are efficient in the majority of cases. They are discussed in great detail in the literature [26], [27] and will not be surveyed here. For large systems, the classical schemes may not yield fast convergence or sufficient accuracy. In addition, if the matrix is ill-conditioned (as, for instance, in the point-matching method), more appropriate techniques must be applied. Finally, if the system is overdetermined because of the application of redundant data technique, the least-squares techniques must be used.

There are two methods for solving linear equation systems that result from the application of the moment method which are becoming increasingly popular among researchers in electromagnetics, namely, the conjugate gradient method and the pseudo-inverse technique. The reasons are the facility with which they can be implemented on a computer and their capability of handling ill-posed problems.

A. The Conjugate Gradient Method

Consider the following matrix equation which may result from the application of the moment method:

$$[L]\vec{\alpha} = \vec{y} \tag{17}$$

where the above quantities were defined before. It can be shown [28] that an iterative method, called conjugate gradient method, can produce the desired solution usually in a number of steps less than the order of the matrix $[L]$. The conjugate gradient method is similar to the steepest descent which involves the search for the minimum of a functional in a direction suggested by its negative gradient. A rigorous mathematical treatment of the method can be found in [29]. The conjugate gradient is a nonlinear iterative method, i.e., the new estimate is not a linear function of the past estimate.

The method starts with an initial guess that generates the first residual vector given by

$$\vec{r}_o = [L]\vec{\alpha}_0 - \vec{y} \tag{18}$$

and the direction vector

$$\vec{d}_1 = -[\tilde{L}]^*\vec{r}_0 \tag{19}$$

where $[\tilde{L}]^*$ is the transposed complex conjugate of $[L]$. Then, the successive iterative steps are given by

$$\vec{\alpha}_{n+1} = \vec{\alpha}_n + t_n \vec{d}_n \tag{20}$$

where

$$t_n = \frac{\|[\tilde{L}]^*\vec{r}_n\|^2}{\|[L]\vec{d}_n\|^2}$$

$$\vec{r}_{n+1} = \vec{r}_n + t_n \cdot [L]\vec{d}_n \tag{21}$$

$$\vec{d}_{n+1} = -[\tilde{L}]^*\vec{r}_{n+1} + q_n \vec{d}_n \tag{22}$$

where

$$q_n = \frac{\|[\tilde{L}]^*\vec{r}_{n+1}\|^2}{\|[\tilde{L}]^*\vec{r}_n\|^2}$$

in which $\|\cdot\|$ indicates the norm.

The conjugate gradient requires more memory storage as compared to linear iteration schemes [30]. However, it has the great advantage of having a rate of convergence practically insensitive to the initial guess. However, a good initial guess reduces considerably the number of iterations to obtain sufficient accuracy. In addition, as in iterative schemes, the round-off errors are confined in the final step of the solution, regardless of the condition number[1] of the matrix $[L]$. Consequently, the method is also well suited for ill-conditioned matrices. The round-off error can eventually be reduced if the ith residual is computed by $\vec{y} - [L]\vec{\alpha}$ rather than (21). However, more computer time is required in this case.

The application of the conjugate gradient method for electromagnetic problems was originated by Sarkar *et al.* for wire antenna scattering [31]. Solutions exhibit fast convergence. More recently, the problem of induced fields in lossy dielectric cylinders was investigated [32]. The matrix system resulting from a point-matching procedure was iteratively solved by the conjugate gradient method. Successive increases of the number of subsections were used to achieve a faster convergence. The order of the system that could be solved was considerably increased as compared with the direct method of solving matrix systems.

B. The Pseudoinverse

It was previously pointed out that if one wants to use the redundant data technique, the resulting system is over determined and, consequently, the least-squares techniques are necessary. These are generally classified into two categories, namely, unconstrained and constrained techniques. This section deals with a constrained least-squares method which has been proved to successfully handle ill-condi-

[1] The condition number of a matrix is defined as the ratio of its highest and its lowest eigenvalue.

tioned problems, in various disciplines such as optics [33], image processing [34], and inverse scattering [18], namely, the pseudoinverse.

Consider the matrix equation (17) in which $[L]$ is the matrix of a bounded operator. Rather than solving directly (17), an estimate $\vec{\alpha}_s$ is generated such that the norm of the residual produced by this estimate is minimum. Simultaneously, the norm of the estimate is constrained to be minimum. This can be written as follows:

$$\text{Among all } \vec{\alpha}_e \text{ which minimize } \|\vec{y}-[L]\vec{\alpha}\| \tag{23}$$

find the particular one which possesses the smallest norm.

The pseudoinverse generates a unique solution for (17) with the constraint (23) even if $[L]$ is singular [35].

A vector space approach is utilized to give a description of the pseudoinverse operation. Let L^a be the adjoint operator of L. Now, consider the nullspace of L generated by the solutions of the homogeneous equation $[L]\vec{\alpha}=\vec{0}$. If L is not onto nor one-to-one, $\vec{y}$ does not belong to the range of L. The orthogonal projection $\vec{y}_p$ of $\vec{y}$ onto the range of L yields a minimum norm of the residual error $\vec{y}-\vec{y}_p$. The set of vector $\vec{\alpha}_e$, which satisfies $\min(\|\vec{y}-\vec{y}_p\|)$, can be found by solving the system

$$[L]\vec{\alpha}_e=\vec{y}_p \tag{24}$$

in which $[L]$ is the matrix associated with the operator L. There are, generally, an infinite number of vectors which satisfy (24). The constraint of minimum norm determines the unique pseudoinverse solution. Using the adjoint operator of L, it can be shown that the minimum norm vector is found by the orthogonal projection of the $\vec{\alpha}_e$ onto the range of L^a [35]. This is a consequence of the orthogonality between the nullspace of L and the range of L^a.

Different techniques have been proposed for determining the pseudoinverse of a matrix, among them an iterative technique for sparse matrices [36]. For general applications, the mathematical description of a projection method is given in [37]. Here, a modified version of the projection method is proposed. It comprises a Gram–Schmidt orthogonalization procedure with pivoting, in order to minimize round-off error propagation. Indeed, when two nearly equal vectors are substracted, the error is likely to be significant in both magnitude and direction. Thus, the Gram–Schmidt procedure applied to nearly dependent vectors invariably results in substractions of nearly equal vectors. Error propagation can be avoided if the vectors with relatively small norm are not used until the end of the procedure.

Let $\{\vec{l}_1,\vec{l}_2\cdots\vec{l}_N\}$ be the set of vectors corresponding to the column of $[L]$ associated with the operator L. The following procedure is recommended for orthogonalizing.

i) Begin with the vector of the largest norm, say, $\vec{l}_k$ (pivot vector).
ii) Make all other vectors orthogonal to it using

$$\underline{l}_i=\vec{l}_i-\left(\langle\vec{l}_i,\vec{l}_k\rangle\|\vec{l}_k\|^2\right)\underline{l}_k. \tag{25}$$

iii) From the now modified vectors, choose the vector of the largest norm, say, $\underline{l}_{i+q}$ (the second pivot vector).
iv) Make all other vectors (excluding $\vec{l}_k$) orthogonal to it using (25).
v) Repeat iii) and iv) until the search for a new $\underline{l}_i$ finds no vectors whose norm is above a certain threshold.

In order to keep a record of the operations on the column of $[L]$, (25) is applied at each step on the column vectors of a identity matrix $[I]$.

At this stage, the matrices $[L]$ and $[I]$ have been transformed to new matrices $[B]$ and $[Q]$, respectively

$$\begin{matrix}[L]\\ \hline [I]\end{matrix}\rightarrow\begin{matrix}[\underline{l}_1,\underline{l}_2\cdots\underline{l}_m & \vec{0},\vec{0}\cdots\vec{0}]\\ \hline [\vec{q}_1,\vec{q}_2\cdots\vec{q}_m & \vec{q}_{m+1}\cdots\vec{q}_N]\end{matrix}=\begin{matrix}[B]\\ \hline [Q]\end{matrix}.$$

Since $[Q]$ records the operations performed on $[L]$, one has

$$[L][Q]=[B]. \tag{26}$$

Consequently, $\{\vec{q}_{m+1}\cdots\vec{q}_N\}$ is a basis for the nullspace of $[L]$. It is, then, orthogonalized with the same procedure described before. In addition, the set $\{\vec{l}_1,\vec{l}_2\cdots\underline{l}_m\}$ forms an orthogonal basis for the range of L. Consequently, the projection of $\vec{y}$ onto the range of L can be written as

$$\vec{y}_p=[B]\vec{a}=[L][Q]\vec{a} \tag{27}$$

with

$$\begin{aligned} a_i&=\langle\vec{y},\underline{l}_i\rangle/\|\underline{l}_i\|^2, && i\leqslant m\\ a_i&=0, && i>m. \end{aligned} \tag{28}$$

By virtue of (27), the solution of $[L]\vec{\alpha}_c=\vec{y}_p$ is clearly

$$\vec{\alpha}_e=[Q]\vec{a}. \tag{29}$$

Only the orthogonal projection of the solutions of (29) onto the range of L^a remains to be carried out. The set of vectors $\{\vec{q}_{m+1},\vec{q}_{m+2}\cdots\vec{q}_N\}$ constitutes a basis for the orthogonal complement of the range of L^a by virtue of the decomposition theorem [35]. Consequently, the pseudoinverse solution is found by the vector orthogonal to the nullspace of L, which is given by

$$\vec{\alpha}_s=\vec{\alpha}_e-\sum_{i=m+1}^{N}\left(\langle\vec{\alpha}_e,\vec{q}_i\rangle/\|\vec{q}_i\|^2\right)\vec{q}_i. \tag{30}$$

It can be noticed that the second term of the right-hand side of (30) represents the orthogonal projection of the vector given by (24) onto the nullspace of L.

The procedure described above is easily automated. The sets $\{\vec{l}_i\}$ and $\{\vec{q}_i\}$ need to be orthogonalized only once for a given problem. This is an advantage since most of the computer time involved in the whole procedure takes place during the Gram–Schmidt orthogonalization. Note that the adjoint operator need not be determined in this approach. If the pseudoinverse is to be calculated, the input vector is replaced by the standard basis vector for the N-dimensional space and the procedure is repeated for the N standard basis vectors yielding each time one column vector of the pseudoinverse of $[L]$. The Gram–Schmidt procedure has to be performed only once.

In conclusion, the method just described is, by its simplicity, very attractive. It has the disadvantage of requiring N more memories than the conjugate gradient method. If the matrix is $N \times N$ and has a low condition number, the pseudoinverse is identical to the standard inverse. The condition (23) prevails for ill-conditioned, over- or underdetermined systems. To prevent meaningless results, the orthogonalization procedure must be interrupted when the relative norm of the remaining column vectors of the matrix $[L]$ is too small. This is left to the judgment of the user. Decreasing or increasing the number of q_l's may severely affect the solution. Typical examples are illustrated in [18].

IV. Conclusion

The goal of this paper was to familiarize the reader with the principle of numerical analysis of electromagnetic-field problems. It was stressed that the success of obtaining accurate numbers with the method of moments lies mostly in the choice of basis functions. There is always a compromise to be made between the difficulty of computing the inner scalar products involved in the procedure and the size of the corresponding matrix equation. This is where the skill of the person who wishes to use the method can be a major factor yielding successful results.

It is felt that with the rapid development of computers, numerical techniques are becoming increasingly popular among engineers. The numerical methods can be applied for design applications and simulations for antenna, scattering, and transmission-line problems. The reader should be aware that there exist other numerical methods that are widely used in electromagnetic problems. They can be more appropriate in certain situations. One refers to finite-element and variational methods.

Two advanced numerical methods for solving equation systems, namely, the conjugate gradient and pseudoinverse method, were presented. They have the merit of being able to handle ill-posed problems which can easily occur in electromagnetic problems. Again, they are not the only methods that have this feature, but they have an advantage of being easily programmable. In addition, the pseudoinverse minimizes the norm of the solution which can be of practical interest in certain situations, such as for inverse scattering problems.

It is unrealistic to believe that computer technology will keep pace with problem-solving requirements. One has reached the machine capabilities and the only hope is the development of new algorithms.

References

[1] R. F. Harrington, "Matrix methods for fields problems," *Proc. IEEE*, vol. 55, pp. 136–149, Feb. 1967.

[2] A. Wexler, "Computation of electromagnetic fields," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-17, pp. 416–439, Aug. 1969.

[3] B. H. McDonald, M. Friedman, and A. Wexler, "Variational solution of integral equations," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-22, pp. 237–248, Mar. 1974.

[4] R. F. Harrington, *Field Computation by Moment Method*. New York: Macmillan, 1968.

[5] R. Mittra and T. Itoh, "A new technique for the analysis of the dispersion characteristics of microstrip lines," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-19, pp. 47–56, Jan. 1971.

[6] L.-P. Schmidt and T. Itoh, "Spectral domain analysis of the dominant and higher order modes in fin-lines," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-28, pp. 981–985, Sept. 1980.

[7] M. Gex-Fabry, J. R. Mosis, and F. E. Gardiol, "Reflection and radiation of an open-ended circular waveguide: Application to nondestructive measurement on materials," *Arch. Elek. Übertragung*, vol. 33, no. 12, pp. 473–478, 1979.

[8] M. Lighthill, *Introduction to Fourier Analysis and Generalized Functions*. Cambridge: University Press, 1962.

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

[10] 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, Apr. 1966.

[11] J. R. Mosis and R. E. Gardiol, "A dynamical radiation model for microstrip structures," *Advances in Electronics and Electron Physics*, vol. 59, pp. 139–235, 1982.

[12] D. E. Livesay and K. M. Chen, "Electromagnetic fields induced inside arbitrary shaped biological bodies," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-22, pp. 1273–1280, 1974.

[13] N. N. Wang, J. H. Richmond, and M. C. Gilreath, "Sinusodial reaction formulation for radiation and scattering from conducting surfaces," *IEEE Trans. Antennas Propagat.*, vol. AP-23, pp. 376–382, Mar. 1975.

[14] A. W. Glisson and D. R. Wilton, "Simple and efficient numerical methods for problems of electromagnetic radiation and scattering from surfaces," *IEEE Trans. Antennas Propagat.*, vol. AP-28, pp. 593–603, May 1980.

[15] M. J. Hagmann, O. P. Gandhi, and C. H. Durney, "Procedures for improving convergence of moment-method solutions in electromagnetics," *IEEE Trans. Antennas Propagat.*, vol. AP-26, pp. 743–748, May 1978.

[16] D. K. Ghodgaonkar, O. P. Gandhi, and M. J. Hagmann, "Estimation of complex permittivities of three-dimensional inhomogeneous biological bodies," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-31, pp. 442–446, June 1983.

[17] M. M. Ney, S. S. Stuchly, A. M. Smith, and M. Goldberg, "Electromagnetic imaging using moment methods," in *URSI Int. Symp. Dig. on Electromagnetic Theory* (Santiago de Compostela, Spain), 1983, pp. 104–107.

[18] M. M. Ney, A. M. Smith, and S. S. Stuchly, "A solution of electromagnetic imaging using pseudoinverse transformation," *IEEE Trans. Medical Imaging*, vol. MI-3, no. 4, pp. 155–162, Dec. 1984.

[19] E. D. Eason, "A review of least-squares methods for solving partial differential equations," *Int. Numerical Methods in Eng.*, vol. 10, pp. 1021–1046, 1976.

[20] G. E. Forsythe and W. R. Wasow, *Finite-Difference Methods for Partial Differential Equations*. New York: Wiley, 1960.

[21] Y. V. Vorbyev, *Method of Moments in Applied Mathematics*. New York: Gordon & Breach, 1965.

[22] M. G. Salvadori and M. L. Baron, *Numerical Methods in Engineering*. Englewood Cliffs, NJ: Prentice-Hall, 1964.

[23] E. H. Newman and P. Tulyathan, "Analysis of microstrip antennas using moment methods," *IEEE Trans. Antennas Propagat.*, vol. AP-29, pp. 47–53, 1981.

[24] J. R. Mosig and F. E. Gardiol, "Analytical and numerical techniques in the Green's function treatment of microwave antennas and scatterers," *Proc. Inst. Elec. Eng.*, pt. H (MOA), vol. 130, pp. 175–182.

[25] P. J. Eberlein, "A Jacobi-like method for the automatic computation of eigenvalues and eigenvectors of an arbitrary matrix," *J. SIAM*, vol. 10, no. 1, pp. 74–88, 1962.

[26] R. W. Hamming, *Numerical Methods for Scientists and Engineers*. New York: McGraw-Hill, 1962.

[27] J. R. Westlake, "A handbook of numerical matrix inversion and solution of linear equations. *Englewood-Cliffs, NJ: Prentice-Hall*, 1968.

[28] M. R. Hestenes and E. Stiefel, "Methods of conjugate gradients for solving linear systems," *J. Res. Nat. Bur. Stand.*, vol. 49, pp. 409–436, 1952.

[29] M. Hestenes, *Conjugate Direction Methods in Optimization*. New York: Springer-Verlag, 1980.

[30] T. K. Sarkar, K. R. Siarkiewicz, and R. F. Stratton, "Survey of numerical methods for solutions of large systems of linear equations

for electromagnetic field problems," *IEEE Trans. Antennas Propagat.* vol. AP-29, pp. 847–856, June 1981.

[31] T. K. Sarkar and S. M. Rao, "The application of the conjugate gradient method for the solution of electromagnetic scattering from arbitrary oriented wire antennas," *URSI Int. Symp. Dig. Electromagnetic Theory* (Santiago de Compostela, Spain), 1983, pp. 93–96.

[32] 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.

[33] R. Barakat and G. Newsam, "Numerically stable iterative method for the inversion of wave-front aberrations from measured point-spread-function data," *J. Opt. Soc. Amer.*, vol. 70, no. 10, pp. 1255–1263, 1980.

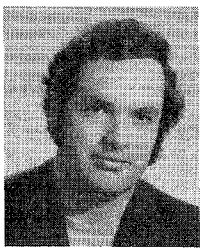
[34] W. K. Pratt and F. Davarian, "Fast computational techniques for pseudoinverse and Wiener image restoration," *IEEE Trans. Comput.*, vol. C-26, no. 6, pp. 571–580, 1977.

[35] C. N. Dorny, *A Vector Space Approach to Models and Optimization*. Hunington, New York: R. E. Krieger, 1980.

[36] A. Bjorck and T. Elfving, "Accelerated projection methods for computing pseudoinverse solutions of systems of linear equations," *J. BIT*, vol. 19, 10, pp. 145–163, 1979.

[37] K. Tanabe, "Projection method for solving a singular system of linear equations and its applications," *Numer. Math.*, vol. 17, pp. 203–214, 1971.

✠



Michel M. Ney (S'83–M'83) was born in Payerne, Switzerland, in 1950. He received the engineer degree from the Swiss Federal Institute of Technology of Lausanne, Switzerland, and the M.Sc. degree in electrical engineering from the University of Manitoba, Canada, in 1976 and 1978, respectively.

During 1978 and 1979, he was with the Laboratory of Electromagnetism and Acoustics at the Swiss Federal Institute of Lausanne as a Research Associate.

He joined the University of Ottawa, Canada, in 1979, where he completed his Ph.D. degree in 1983. Since that time, he has been with the Department of Electrical Engineering of the University of Ottawa as an Assistant Professor. His main research interests are electromagnetic engineering and numerical methods.