1

PRINCIPLES OF FINITE METHODS IN ELECTROMAGNETIC SCATTERING

M. A. Morgan

1.1 **Introduction**
1.2 **Finite Methods**
  a. Finite Difference Approximation
  b. Simple Finite Difference Example
  c. A Brief Visit with the FD-TD Approach
  d. Weighted Residual Method
  e. Simple Finite Element Example
  f. Basics of Error Analysis
1.3 **Mesh Termination**
  a. Unimoment Method
  b. Boundary Integral Equation
  c. Field Feedback Formulation
  d. Radiation Boundary Conditions
1.4 **Conclusion**
References

1.1 Introduction

This book deals with the theory and applications of differential equation based numerical methods in the realm of electromagnetic scattering. To prepare the reader for the advanced discussions of this topic that will follow in subsequent chapters, it is appropriate to begin the book with a tutorial review of the underlying concepts and terminology. Such a review will form the heart of this chapter. In addition, some examples of applications that are not otherwise covered in this text will be considered. It will be assumed that the reader is familiar with the fundamental physical and mathematical aspects of electromagnetic theory at the graduate level. Some excellent references for this knowledge are those by Stratton [1], Harrington [2], Van Bladel [3] and Kong [4].

The International System (SI for the French equivalent) is used in this text. This is a form of meter-kilogram-second (mks) system, with electrical quantities being expressed in the units actually measured: volts, amperes, coulombs, ohms, watts, etc. Using our mathematical notation, vector quantities, including $N \times 1$ and $1 \times N$ arrays, will usually appear with an overbar, e.g. $\overline{V}$. Dyadics and $M \times N$ matrices are most often denoted by a double overbar, e.g. $\overline{\overline{G}}$ for a Green's dyadic. Other symbolism will sometimes appear, at the discretion of individual chapter authors. Examples are the use of an underbar to indicate a row vector of unknowns (e.g. $\underline{a}$), or brackets to represent arrays, as in $[A]$. Such chapter dependent notation will be defined locally, either formally or by way of an obvious context. Unit vectors are uniformly designated by a circumflex; for example, a unit normal vector on a surface is written as $\hat{n}$.

As is customary in electrical engineering, sinusoidally time-varying quantities (known as time-harmonic or frequency-domain) are represented by complex phasor quantities having assumed temporal variation of $e^{j\omega t}$. The conversion from a complex phasor, say $\overline{E}(\overline{r})$, to the corresponding time-harmonic function, $\overline{\mathcal{E}}(\overline{r}, t)$ is effected by the operation $\overline{\mathcal{E}} = Re\{\overline{E}e^{j\omega t}\}$, where Re denotes the real part. Time-harmonic fields will be considered in Chapters 2 through 6, with the time-domain case appearing in Chapters 7 and 8. In this introductory chapter, both frequency- and time-domain cases will be covered.

Numerical approximations will be considered in the next section for the solutions of differential equations using either finite differences or finite elements. These techniques will henceforth be referred to as the "finite methods". When continuous systems are discretized using finite methods, *local* interactions result between unknowns. Thus, unknowns at discrete spatial points (nodes) are explicitly related only to their nearest neighbors in space. This results in highly sparse matrices for the case of time-harmonic fields, while requiring causal updating due only to nearest spatial neighbors in time-domain applications. In contrast, integral equation formulations typically produce global interactions between discrete unknowns, thus generating full system matrices for the time-harmonic case and all-inclusive connectivity in the transient evolution of the time-domain equations. As a result of the numerical efficiency brought about from local, versus global, interactions, the finite methods have the potential to solve larger and more complex electromagnetic problems than can be handled by *volume* integral equations, given the same constraints on computer time and available memory.

Although the finite methods offer numerical efficiency they are, by necessity, formulated as boundary-value and initial-value problems. As such, the solution of scattering and radiation problems in unbounded (open) spatial regions requires a mechanism for coupling closed region numerical solutions to the exterior space. This procedure must also ensure that the proper radiation conditions on the scattered field are satisfied. With the exception of the "infinite element" approach [5], a closed region is used to bound the spatial mesh employed in the finite methods. Various procedures for properly terminating, at least in an approximate numerical sense, the outer boundary of the mesh are employed in the subsequent chapters of this book. We will consider these methods from a basic conceptual point of view, in section 1.3.

1.2 Finite Methods

The finite methods may be classified as the numerical techniques that provide local interaction discretization for solving continuous boundary-value and initial-value problems [6]. As such, the finite methods offer a means to approximate the solution of specified differential equations in one or more spatial dimensions, as well as in time.

We will first consider the general idea of polynomial interpolation, followed by its application to finite-differences, as supported by examples incorporating phasor and time-domain concepts. This will be followed by a discussion, with examples, of the finite element method. A particular point of view will be developed that unifies the finite methods by identifying finite differences as a special case of finite elements.

a. *Finite Difference Approximation*

The most common finite method has been that of finite differences, where discrete approximations to partial derivatives are obtained by differentiating a piecewise polynomial, or other approximating function, which has been *point-matched* to the actual unknown function at

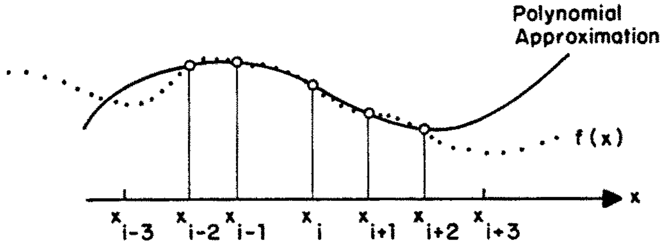

Figure 1 **Point matched polynomial approximation around $x=x_i$.**

the *nodes* of the problem. As a simple example, consider the finite difference (FD) approximation of the various derivatives, up to the n-th order, of an unknown function, $f(x)$, in one-dimension. The domain of x is partitioned into generally unequal segments, separated by ordered nodes, x_k, for $k = 1,2,3,\cdots,M$, as is illustrated in Fig. 1. In the region of x that contains the i-th node, at which point $x = x_i$, let $f(x)$ be approximated by an n-th order polynomial,

$$f(x) \approx a_n\, x^n + a_{n-1}\, x^{n-1} \cdots + a_1\, x + a_0 \tag{1}$$

The $n+1$ unknown coefficients can be found as linear functions of $n+1$ of the unknown nodal values of $f(x)$. Usually, these particular matching nodes are taken to bracket the i-th node except near the $i = 1$ and $i = M$ endpoints of the domain on x. The linear relationship between the coefficients and the nodal values of $f(x)$ is developed by point matching the polynomial in (1) at the $n+1$ nodes, resulting in the linear system defined by

$$\sum_{p=0}^{n} (x_k)^p\, a_p = f(x_k) \quad \text{for } n+1 \text{ values of } k \tag{2}$$

After inverting this system, the resultant linear functional form for each a_p can be substituted into (1). The FD formulas for each derivative up to the n-th order can then be obtained by analytically differentiating (1), followed by an evaluation at $x = x_i$. This procedure will yield a formula for each of the derivatives which is expressed as a linear function of the $n+1$ nodal values of $f(x)$.

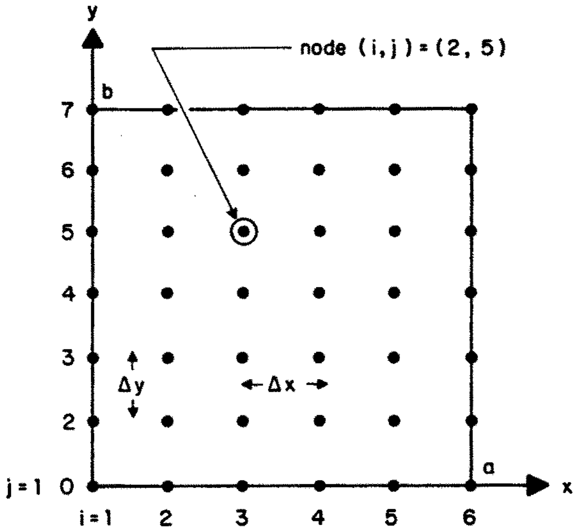

Figure 2 Finite difference grid in a rectangular region.

For example, if $n = 2$ with equal segments $x_i = i\,\Delta x$, then (1) results in a quadratic expression. By enforcing (1) at the three points $k = i-1$, $k = i$ and $k = i+1$, a 3×3 linear system,

$$\begin{bmatrix} 1 & x_{i-1} & x_{i-1}^2 \\ 1 & x_i & x_i^2 \\ 1 & x_{i+1} & x_{i+1}^2 \end{bmatrix} \cdot \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} f(x_{i-1}) \\ f(x_i) \\ f(x_{i+1}) \end{bmatrix} \tag{3}$$

is obtained. After inverting this system, we substitute the a_p's into (1) and differentiate once and then twice to obtain the well-known finite difference formulas for the first and second derivatives,

$$\frac{d\,f(x_i)}{dx} \approx \frac{f(x_{i+1}) - f(x_{i-1})}{2\Delta x} \tag{4a}$$

$$\frac{d^2 f(x_i)}{dx^2} \approx \frac{f(x_{i+1}) - 2\,f(x_i) + f(x_{i-1})}{(\Delta x)^2} \tag{4b}$$

Extensions of this idea can be made to higher order polynomial approximations and to multiple dimensions. As an example, consider overlaying a 2-D rectangular region, defined by $0 \leq x \leq a$ and $0 \leq y \leq b$, with a regular grid of nodes located at (x_i, y_j) for $i = 1, 2, \cdots, 6$ and $j = 1, 2, \cdots, 7$, as per Fig. 2. About a given point (x_i, y_j) in this mesh, we will approximate a continuous function, $f(x, y)$, by a 2nd-order polynomial form

$$f(x,y) \approx c_0 + c_1\, x + c_2\, y + c_3\, x^2 + c_4\, y^2 \tag{5}$$

The five coefficients are found by equating the expansion to the function at the five nodes: (x_{i-1}, y_j), (x_i, y_j), (x_{i+1}, y_j), (x_i, y_{j-1}), and (x_i, y_{j+1}), followed by an inversion of the resultant 5×5 linear system. Finite difference approximations to the partial derivatives are then found by differentiating (5). All of this work results in formulas which are identical in form to (4), e.g.

$$\frac{\partial^2 f(x_i, y_j)}{\partial x^2} \approx \frac{f(x_{i+1}, y_j) - 2\, f(x_i, y_j) + f(x_{i-1}, y_j)}{(\Delta x)^2} \tag{6}$$

This approach can be applied as well to irregular meshes, having non-rectangular and non-equispaced nodal arrangements. As will be shown in subsection 1.2d, the finite difference method can be thought of as a special case of a more generic technique, called the finite element method.

b. Simple Finite Difference Example

The finite difference solution of a boundary-value problem is set up by replacing the analytical derivatives contained in the differential equation by finite difference formulas at each nodal point where the solution function is to be found. Thus, there results a system of linear equations relating the unknown nodal values of the solution function to both the known excitations (drivers) of the differential equation and the known boundary-values of the solution function.

To illustrate this, let us set up the finite-difference frequency domain (FD-FD) solution to the Helmholtz equation in the same rectangular region, defined by $0 \leq x \leq a$ and $0 \leq y \leq b$, as depicted in Fig. 2. The partial differential equation (PDE) being considered is the inhomogeneous Helmholtz equation,

$$\{\nabla^2 + \kappa^2\}\, f(x,y) = g(x,y) \tag{7}$$

where the Laplacian operator is $\nabla^2 = \partial^2/\partial x^2 + \partial^2/\partial y^2$. The forcing function, $g(x,y)$, is to be specified and the solution, f, is assumed to satisfy homogeneous *Dirichlet* boundary conditions (BC's) on the perimeter: $f(x,0) = f(x,b) = f(0,y) = f(a,y) = 0$. This problem could be found, for instance, in solving for the TM eigenmodes of a metallic rectangular waveguide whose cross section encloses the solution region. In such a case, $f = E_z$, $g = 0$ and the eigenvalues of κ are assumed to have been found through a separate procedure.

The discrete unknowns are the nodal values of $f(x_i, y_j)$ for $i = 2, \cdots, 5$ and $j = 2, \cdots, 6$. We will denote the approximate numerical solution for these nodal values as $f_{i,j} \approx f(x_i, y_j)$. A five node "star" finite difference approximation to the ∇^2 operator can be constructed from the sum of (6) and the corresponding form for $\partial^2/\partial y^2$. Applying this to (7) at each interior node yields

$$\begin{aligned} &\frac{\{f_{i+1,j} + f_{i-1,j}\}}{(\Delta x)^2} + \frac{\{f_{i,j+1} + f_{i,j-1}\}}{(\Delta y)^2} \\ &+ \{\kappa^2 - \frac{2}{(\Delta x)^2} - \frac{2}{(\Delta y)^2}\}\, f_{i,j} = g(x_i, y_j) \end{aligned} \tag{8}$$

Enforcement of this equation at the interior nodes leads to a linear set of equations. To cast these equations into a matrix form, we need to define arrays for the unknowns, $f_{i,j}$ and the known nodal values of the driver, $g(x_i, y_j)$. By ordering these column vectors as

$$\overline{F} = [f_{2,2}, f_{3,2}, \cdots, f_{2,3}, \cdots, f_{5,6}]^T \tag{9}$$

and

$$\overline{G} = [g(x_2,y_2), g(x_3,y_2), \cdots, g(x_2,y_3), \cdots, g(x_5,y_6)]^T \tag{10}$$

where the T superscript indicates transpose, the linear system is given by

$$\overline{\overline{A}} \cdot \overline{F} = \overline{G} \tag{11}$$

The 30 × 30 matrix is symmetric, sparse and banded

$$\overline{\overline{A}} = \left[\begin{array}{cccccccccc} a_{11} & a_{12} & 0 & 0 & a_{15} & 0 & 0 & 0 & \cdots & \\ a_{21} & a_{22} & a_{23} & 0 & 0 & a_{26} & 0 & 0 & \cdots & \\ 0 & a_{32} & a_{33} & a_{34} & 0 & 0 & a_{37} & 0 & \cdots & \\ & & \ddots & \ddots & \ddots & & & \ddots & & \\ a_{51} & 0 & 0 & a_{54} & a_{55} & a_{56} & 0 & 0 & a_{59} & \\ & \ddots & & & \ddots & \ddots & \ddots & & & \ddots \end{array}\right] \tag{12a}$$

with

$$a_{m,n} = a_{n,m} \tag{12b}$$

$$a_{n,n} = \kappa^2 - \frac{2}{(\Delta x)^2} - \frac{2}{(\Delta y)^2} \tag{12c}$$

$$a_{n,n+1} = a_{k,k-1} = \frac{1}{(\Delta x)^2} \tag{12d}$$

$$a_{n,n+4} = a_{k,k-4} = \frac{1}{(\Delta y)^2} \tag{12e}$$

Note that the m-th row of $\overline{\overline{A}}$ represents the equation defined by enforcing (8) at the node (i,j) corresponding to $f_{i,j}$ of the m-th element of $\overline{F}$ in (9). Inversion of (11) can be accomplished by exploiting the sparsity and block-matrix nature of $\overline{\overline{A}}$. Such procedures will allow the practical solution of very large systems, as will be discussed further in subsection 1.2e and in Chapters 2 and 6.

c. *A Brief Visit with the FD-TD Approach*

In the example of the previous subsection we considered the finite difference solution of a simple boundary problem centered about the Helmholtz equation in 2-D. This approach can be extended to higher dimensions, including the time-variable. Such is the case in Chapter 8, where Taflove and Umashankar begin by constructing a finite-difference time-domain (FD-TD) discretization of Maxwell's equations in 4-D space-time. To set the stage for this development, we will consider a simple example of FD-TD as applied to the solution of an initial-value problem involving the 2-D (1-D space plus time) wave equation.

In particular, the wave equation for the scalar field $u(x,t)$ in a lossless, uniform and *non-dispersive* medium is given by the PDE,

$$\frac{\partial^2 u}{\partial x^2} - \frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} = 0 \tag{13}$$

where c represents the frequency-independent (hence non-dispersive) propagation velocity of the field u. In the electromagnetic case, this field can represent, for instance, the voltage or current on a uniform TEM x-directed transmission line. Another example is the transverse (to $\hat{x}$) electric or magnetic fields in an ideal plane wave propagating in the $\pm x$ direction. There are similar examples in acoustics and mechanical vibration.

To continue the example, let us first define the domain of the solution as $0 \leq x \leq a$ and $t \geq 0$. This domain is illustrated in Fig. 3. For the domain defined, (13) is to be solved subject to the following *mixed* initial-value and boundary-value conditions:

$$u(x,0) = q(x) \quad \text{for } 0 \leq x \leq a \tag{14a}$$

$$u(0,t) = 0 \qquad \text{for } t \geq 0 \tag{14b}$$

$$u(a,t) = 0 \qquad \text{for } t \geq 0 \tag{14c}$$

where $q(x)$ is specified so that $q(0) = q(a) = 0$. These conditions could define such phenomena as the deflection of a string which is pinned at $x = 0$ and $x = a$ or the $\mathcal{E}_y(x,t)$ electric field in a planar resonator model with conductor plates at $x = 0$ and $x = a$. The analytical solution can be obtained through a variety of elementary methods [7].

Since the wave equation is classified as a hyperbolic PDE, [8], the solution at any point will propagate along outbound "characteristics" which, for the example case, are lines in 2-D space-time defined by $\Delta x/\Delta t = \pm c$. Characteristic lines for the endpoints of the initial-values

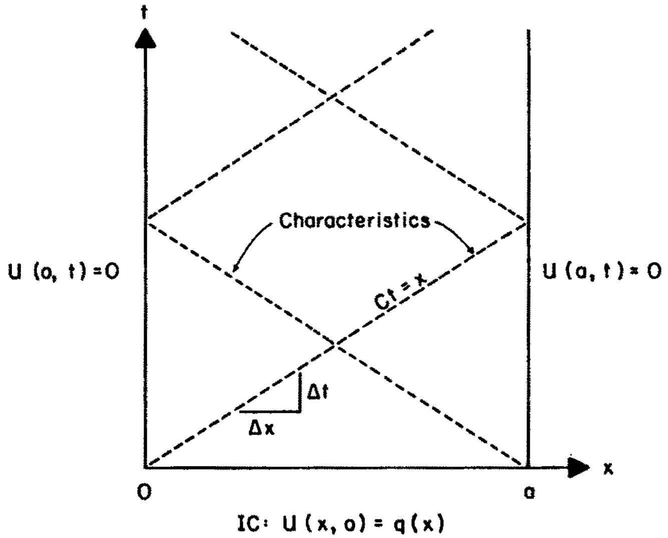

Figure 3 Space–time domain for wave equation solution.

are shown in Fig. 3. For problems in 3-D and 4-D the characteristics become, respectively, space-time cones and expanding spheres. The characteristics can be thought of as bounding the region of space-time which is causally influenced by an "event" at a given $(\bar{r}, t)$ point.

Characteristics play an important role in the numerical solution of processes defined by *hyperbolic* and *parabolic* (e.g. heat transport) PDE's. These processes are set up as equations of evolution and their numerical solution proceeds using a "marching in time" algorithm, with time step Δt. It was recognized as early as 1928, by Courant, Friedrichs and Lewy [9], that the ratio of spatial segmentation distance to time step size should be constrained so as *not* to violate the bounds defined by the characteristics. The result of ignoring the CFL criterion (also referred to as the "Courant limit") is the production of an ill-posed and unstable numerical solution, with rapidly diverging accuracy. This constraint applies not only to the finite method discretization of the original PDE problem but also to the time-stepping solution of the corresponding integral equation formulation for the same physical process. The Helmholtz equation is an example of an *elliptic* PDE. These do not have any characteristics in real space and, hence, we did not concern ourselves with this issue in setting up the previous example in subsection 1.1b. The Courant limit will be discussed further in the advanced time-domain applications considered in Chapters 7 and 8.

Let us return now to the problem at hand: discretizing (13) and (14) using finite differences. In analogy with the Helmholtz equation example, we first define a mesh to overlay the space-time region in Fig. 3. This is shown in Fig. 4. The node positions are located at $(x_i, t_j) = (i\Delta x, j\Delta t)$, where $\Delta x = a/N$. The numerical solution for the (x_i, t_j) nodal values are denoted as $u_{i,j}$. We next apply three-point central difference formulas for the partial derivatives in (13), at node (x_i, t_j). This is similar to what was done in (8) for the Helmholtz equation. The result, after rearranging terms, is the time-marching equation

$$u_{i,j+1} = \gamma^2 \left(u_{i+1,j} + u_{i-1,j}\right) + 2(1-\gamma^2)\, u_{i,j} - u_{i,j-1} \tag{15}$$

where the dimensionless constant is

$$\gamma = c\,\frac{\Delta t}{\Delta x} \tag{16}$$

The CFL condition for stability is that $\gamma \leq 1$. If c is not a spatial or temporal constant then the CFL condition on γ must be satisfied *locally* in space-time. To begin the evolution of (15) at $i = 1$ we note that the initial condition (IC) is given by $u_{i,0} = q(x_i) = q_i$, with $u_{i,-1}$ assumed to be zero. This yields

$$u_{i,1} = \gamma^2 \left(q_{i+1} + q_{i-1}\right) + 2(1-\gamma^2)\, q_i \tag{17}$$

where we note that $q_0 = q_N = 0$, as assumed in (14). After initiating the j=1 step using (17), the updating continues with (16).

An alternate approach, [10], is to use two first-order "state equations" to represent the wave equation in (13). This is done by defining two *state variable* functions: the original unknown $u(x,t)$, plus a new function, $w(x,t)$, where

$$\frac{\partial w}{\partial t} = c\,\frac{\partial u}{\partial x} \tag{18a}$$

$$\frac{\partial u}{\partial t} = c\,\frac{\partial w}{\partial x} \tag{18b}$$

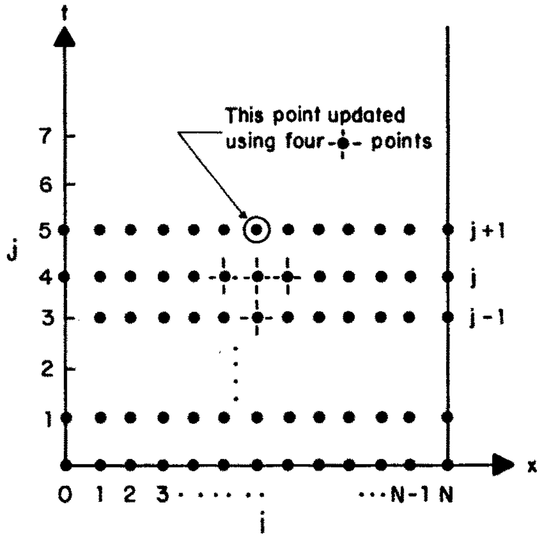

Figure 4 **Finite difference mesh for the wave equation showing forward stepping evolution point.**

This first-order coupled PDE system may be recognized as the well-known "telegrapher's equations" from transmission line analysis. The wave equation can be generated by differentiating (18a) with respect to x and (18b) with respect to t, followed by equating the resultant 2nd-order cross derivatives of w.

To discretize these equations, we will use a "leapfrog" mesh, [11: 489–496], as shown in Fig. 5. A key feature of this mesh is the space-time offset of the nodes for defining u and w. In particular, let us define

$$w_{i,j} = w(x_i, t_j) \quad \text{for } i = 0, 2, 4, \cdots, N\text{-}1; \;\; j = 1, 3, 5, \cdots \qquad (19a)$$

$$u_{i,j} = u(x_i, t_j) \quad \text{for } i = 1, 3, 5, \cdots, N \;\; ; \;\; j = 0, 2, 4, \cdots \qquad (19b)$$

Using central differences, we will enforce (18a) at (even,even) (i,j) nodes and (18b) at (odd,odd) (i,j) nodes. For $i = 2, 4, 6, \cdots, N$ and $j = 0, 2, 4, \cdots$, this procedure produces

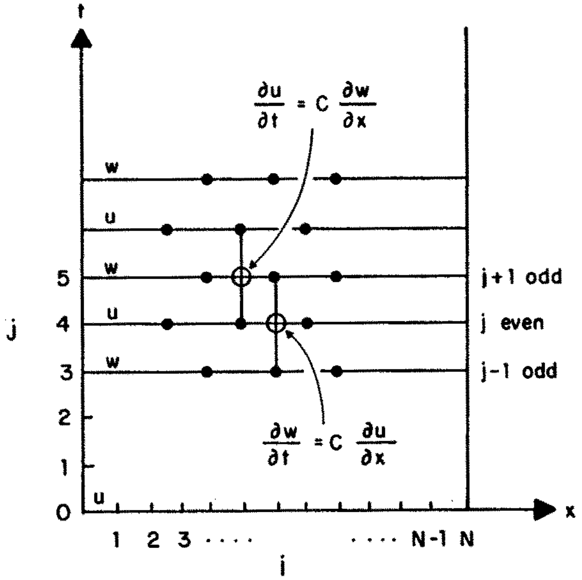

Figure 5 **Nodal topology for the leapfrog method showing central difference points to enforce coupled first-order PDE's.**

$$\frac{w_{i,j+1} - w_{i,j-1}}{2\,\Delta t} = c\,\frac{u_{i+1,j} - u_{i-1,j}}{2\,\Delta x} \tag{20a}$$

while for $i = 1, 3, 5, \cdots, N$, and $j = 1, 3, 5, \cdots$, it gives

$$\frac{u_{i,j+1} - u_{i,j-1}}{2\,\Delta t} = c\,\frac{w_{i+1,j} - w_{i-1,j}}{2\,\Delta x} \tag{20b}$$

where we have assumed N is odd for terminating the i-index. To begin the *leapfrog* evolution, we start by enforcing (20a) for j=0, with $w_{i,-1}$ assumed to be zero, resulting in

$$w_{i,1} = \frac{\Delta x}{\Delta t}\,(q_{i+1} - q_{i-1}) \tag{21}$$

The notation, $q_i = q(x_i) = u(x_i, 0)$ for the nodal values of the specified IC, was employed in (17). After the startup in (21), the relationship in (21b) is used to obtain $u(i,2)$, followed by alternating applications of (21a) and (21b). This leapfrog process has the same stability constraint as the original second-order system, namely the Courant limit:

$\Delta x/\Delta t \leq c$. To get around having double $2\,\Delta x$ and $2\,\Delta t$ increments in defining the central differences, a *half-index* scheme is often used. This will be the case in the time-marching algorithms presented in Chapters 7 and 8. There are numerous other methods for applying the finite-difference method to create time-stepping solutions to hyperbolic and parabolic PDE's. A comprehensive reference on this topic is given in [11].

An extension of this 2-D space-time leapfrog concept to Maxwell's curl equations in 4-D will result in the original *Yee algorithm* [12]. This early (1966) FD-TD method employs two interleaved 3-D rectangular spatial lattices, one for the vector $\overline{E}$-field and one for the vector $\overline{H}$-field. The details of the Yee algorithm are discussed in Chapter 8.

d. Weighted Residual Method

The finite element method can be approached for elliptic, hyperbolic or parabolic PDE's through the "weighted residual method", which we will henceforth call WRM [11]. Let us begin in a very general manner by considering the numerical solution of a specified PDE in an M-dimensional spatial region or space-time region, to be denoted in either case as V_0. Points in the M-D region are given by the ordered vector, $\overline{r} = (r_1, r_2, ..., r_M)$. For example, the Helmholtz equation example in part 1.2b had M=2, where $\overline{r}$ was (x, y), while for the wave equation example in the last subsection, M was also 2 and $\overline{r} = (x, t)$.

The generalized PDE can be written as

$$\overline{\overline{\mathcal{D}}}(\overline{r}) \cdot \overline{f}(\overline{r}) = \overline{g}(\overline{r}) \quad \text{for } r \in V_0 \tag{22}$$

where $\overline{\overline{\mathcal{D}}}(r)$ is the given dyadic differential operator, $\overline{f}$ is the unknown vector function and $\overline{g}$ is the known driving vector. Essential boundary data concerning $\overline{f}$ are known on a surface, S_0, which encloses V_0. An example is the scalar Helmholtz equation in 2-D, which we previously considered in (7).

To find the approximate numerical solution of (22), we use a *basis function* expansion to represent $\overline{f}$,

$$\overline{f}_a(\overline{r}) = \sum_{n=1}^{N} \overline{\overline{U}}_n(\overline{r}) \cdot \overline{C}_n \tag{23}$$

where the diagonal dyadic functions,

$$\overline{\overline{U}}_n(\overline{r}) = u_{n,1}\,\hat{r}_1\,\hat{r}_1 + u_{n,2}\,\hat{r}_2\,\hat{r}_2 + \cdots + u_{n,M}\,\hat{r}_M\,\hat{r}_M \tag{24}$$

are members of the basis set. Each basis function should, ideally, have the same order of differentiability as does the exact solution. As N is increased, the approximate expansion in (23) should converge in a *pointwise* sense to $\overline{f}(\overline{r})$. This last condition depends upon how completely the set of basis functions spans the subspace of functions occupied by the various solutions to (22). This quality is reflected in the linear independence of the function set.

Upon substituting (23) into (22), there results,

$$\overline{\overline{D}}\cdot\overline{f}_a(r) = \sum_{n=1}^{N} \{\overline{\overline{D}}(\overline{r})\cdot\overline{\overline{U}}_n(\overline{r})\}\cdot\overline{C}_n = \overline{g}(\overline{r}) \tag{25}$$

To solve for the N vector coefficients, $\overline{C}_n$, we enforce this equation with respect to a succession of N weighted vector integrations over V_0,

$$\begin{aligned}\langle\overline{\overline{W}}_k(\overline{r}),\,\overline{\overline{D}}\cdot\overline{f}(\overline{r})\rangle &= \sum_{n=1}^{N}\langle\overline{\overline{W}}_k(\overline{r}),\,\overline{\overline{D}}(\overline{r})\cdot\overline{\overline{U}}_n(\overline{r})\rangle\cdot\overline{C}_n \\ &= \langle\overline{\overline{W}}_k(\overline{r}),\,\overline{g}(\overline{r})\rangle \text{ for } k = 1, N\end{aligned} \tag{26}$$

with

$$\langle\overline{\overline{W}}(\overline{r}),\overline{\overline{G}}(\overline{r})\rangle = \int_{V_0}\overline{\overline{W}}(\overline{r})\cdot\overline{\overline{G}}(\overline{r})\,d^M\overline{r} \tag{27}$$

indicating an integration of the inner (dot) product of the two indicated functions over V_0.

The set of diagonal dyadics,

$$\overline{\overline{W}}_k(\overline{r}) = w_{k,1}\,\hat{r}_1\,\hat{r}_1 + w_{k,2}\,\hat{r}_2\,\hat{r}_2 + \cdots + w_{k,M}\,\hat{r}_M\,\hat{r}_M \tag{28}$$

is termed the "weighting functions" in WRM. This set is called the "testing functions" in the MoM. There are N vector equations in (28), each of which contains M scalar equations for the M components of each $\overline{C}_n$. Thus, the linear system has rank $N \cdot M$. By using a certain class of basis and weighting functions, this matrix system can be made very sparse. This will be demonstrated in the next subsection.

Before continuing with the development of the WRM, let us pause briefly to consider some of the terminology presented here. The procedure just considered could just as well have been applied to an integral operator equation having the form of (22). In fact, when applied to integral equations in electromagnetics, the WRM concept has historically been termed both the moment method and the method of moments (MoM), [2]. The reference to "moments" is due to the similarity of the inner product integral terms in (27) to statistical moments found in probability theory. Origins of the WRM terminology are in the area of finite elements, as applied to structural and fluid dynamics problems.

Let us now return to the details of the WRM. In setting up the linear system in (26) there are some additional considerations that need to be addressed. One of these concerns the support region of the basis functions: either *full-range* (over all of V_0) or *compact* (each being nonzero over only a portion of V_0). An example of full-range basis functions is the set of complex exponentials employed in Fourier series, where the set of $\overline{C}_n$'s are termed the "spectrum" of the expansion. Compact basis functions are common to finite element applications, as well as MoM solutions of integral equations. Usually, these basis functions are selected so that at each node of the discretized problem all basis functions, *except one*, are zero. At its associated node, where it is nonzero, the basis function will usually be set to unity. In such a case, the coefficients in (23), $\overline{C}_n$'s, represent the solution values of $\overline{f}(\overline{r})$ at the N nodes.

Another consideration involves the set of weighting functions. There is obviously an unlimited selection available. Three of the more common types are:

(1) *Point Collocation*, (also known as simply "collocation") uses a delta function, $w_{k,m} = \delta(\overline{r} - \overline{r}_k)$ for each diagonal component of $\overline{\overline{W}}_k$ in (28). The effect of this is to reduce the integration moments in (27) to simple point-matching at the respective nodes, resulting in nothing more than a special case of the finite difference method where the polynomial approximation in (1) and (2) is replaced by the basis function representation being employed to construct the $\overline{\overline{U}}_n$. Thus, point collocation can be viewed as equivalent to finite differences. An advantage of this approach is its relative simplicity in generating the matrix elements from (26), since integrations are reduced to enforcing the approximation at the node points. On the other hand, there is no control on the behavior of the solution

in-between the nodes, at least in the sense of solving the differential equation. This usually results in the least-accurate solution of the three methods being considered here, where we are assuming a similar node density and computer word-length in each case.

(2) *Subdomain Collocation*, (also known as the "subdomain method") which employs a set of mutually exclusive, unit-amplitude pulse functions, $w_{k,m}(\bar{r}) = 1$, in a defined region around the k-th node. These regions enclosing each node are non-overlapping and usually are directly adjacent to one another, without unfilled space. This approximation is usually more accurate than Point Collocation, but not as accurate as Galerkin's method, when self-adjoint operators are involved.

(3) *Galerkin Method*, which uses the same set of functions for both basis and testing, $\overline{\overline{W}}_k(\bar{r}) = \overline{\overline{U}}_k(\bar{r})$. For the case of a *self-adjoint* operator, $\overline{\overline{\mathcal{D}}}(\bar{r})$, it can be shown that the functional defined by

$$Q(\bar{f}) = \langle \bar{f}, \overline{\overline{\mathcal{D}}} \cdot \bar{f} \rangle - 2 \langle \bar{f}, \bar{g} \rangle$$

is stationary about the solution to the original operator equation, in (22), [13].

The variational principle considered above, when applied to the basis function expansion in (23), is termed the *Rayleigh-Ritz Method* and yields the Galerkin's result for the weighted residual approach. This procedure usually provides the most accurate solution and forms the foundation for most of the FEM work that has been done. Another means of achieving this same result (the Galerkin equations, with $\overline{\overline{W}}_k = \overline{\overline{U}}_k$) is by way of the classical Euler-Lagrange variational formulation. This has the advantage of reducing the order of differentiation on the basis functions, *vis-a-vis* a direct Galerkin approach, and will be demonstrated in the following section. Error bounds for the above three methods are considered in detail in [13]. We will also briefly consider this topic in subsection 1.2f.

A hybrid approach is often taken in applying the WRM to time-domain solutions, whereby a finite-element basis function expansion is used to represent the spatial variations of the solution while time derivatives are approximated using finite-differences. If interpreted in the WRM sense, this translates into using space-time weighting functions having delta function temporal variation, thus *sifting out* the

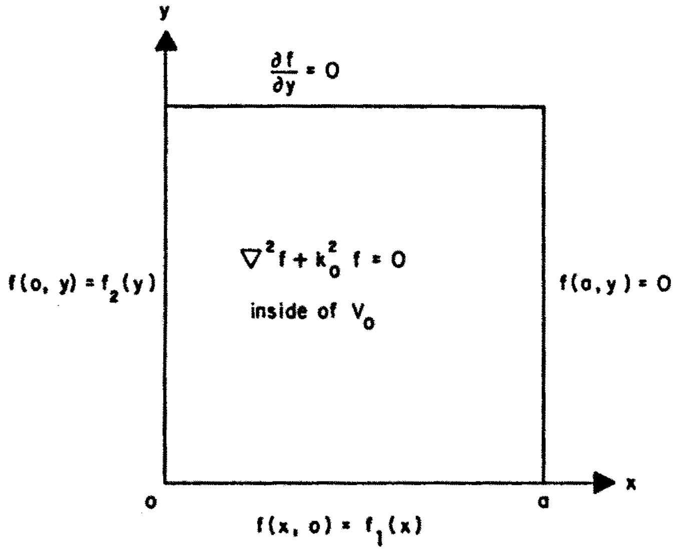

Figure 6 Helmholtz equation with mixed boundary conditions.

discrete time steps of the computational process. Such a procedure is used implicitly in the technique described in Chapter 7, resulting in time-domain leap-frogging of spatial basis function expansions for $\overline{E}$ and $\overline{H}$. A "boundary element" approach is also used in Chapter 7 for restricting irregularities in the space-time lattice (mesh) to the interfaces between dissimilar materials.

e. Simple Finite Element Example

Consider the simple problem of the undriven Helmholtz equation ($g = 0$) with free space wavenumber $k_0 = \omega/c$ (where ω is the radian frequency and $c = 3 \times 10^8$ m/s),

$$\{\nabla^2 + k_0^2\}\, f(x,y) = 0 \tag{29}$$

within the rectangular region of Fig. 6, with *mixed* (both Dirichlet and Neumann) BC's,

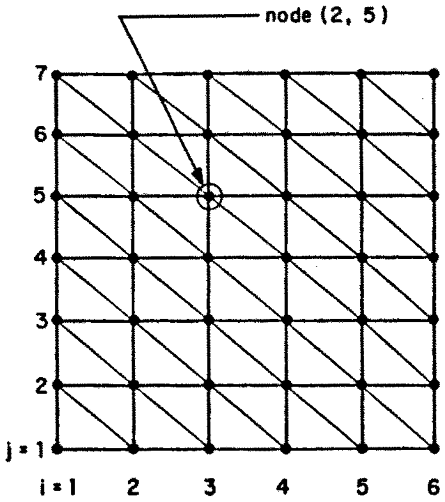

Figure 7 Rectangular region finite element mesh.

$$f(x,0) = f_1(x) \tag{30a}$$

$$f(0,y) = f_2(y) \tag{30b}$$

$$f(a,y) = 0 \tag{30c}$$

$$\frac{\partial f}{\partial y}(x,b) = 0 \tag{30d}$$

The rectangular region is discretized into the same grid of doubly-ordered nodes, as in Fig. 2. As before, the nodal values of the solution will be denoted as $f_{i,j} \approx f(x_i, y_j)$. The Dirichlet boundary condition nodes are specified for i=1, i=6 and j=1. Note that the nodal values along the top row (j=7) are now unknowns since the Neumann BC is specified there.

Proceeding with the FEM approach, triangular "finite elements" are used to divide up the space, as per Fig. 7. These elements are used

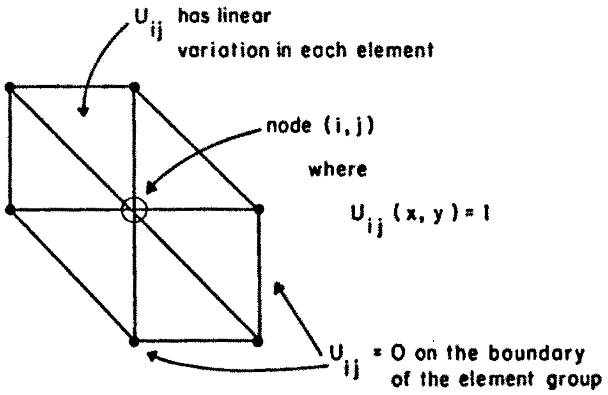

Figure 8 Node with surrounding element group.

to define support regions for the (compact support type) basis functions that will be used to approximate the solution around each node. We will employ piecewise linear *pyramid* basis functions, $U_{i,j}(x,y)$, to represent the solution. Following the general procedure indicated by (23), but for the simpler scalar function case, gives the basis function approximation to $f(x,y)$,

$$f_a(x,y) = \sum_{i=1}^{6} \sum_{j=1}^{7} f_{i,j} \; U_{i,j}(x,y) \tag{31}$$

The support region for a given $U_{i,j}$ is all elements which share the (i,j) node, as illustrated in Fig. 8 for an interior node. In the ℓ-th element, we will locally number the associated 3 nodes, (k=1,2,3) as shown in Fig. 9. Within this ℓ-th element, the linear basis function associated with the k-th node is given by the matrix product

$$U_k(x,y) = [x, \; y, \; 1] \cdot \overline{L}_k \tag{32}$$

where $\overline{L}_k$ is the k-th column of the element coordinate matrix,

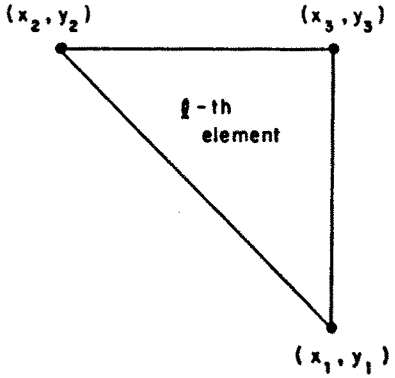

Figure 9 Local coordinates in an element.

$$\overline{\overline{L}} = \frac{1}{\Delta_\ell}\begin{bmatrix} (y_2 - y_3) & (y_3 - y_1) & (y_1 - y_2) \\ (x_3 - x_2) & (x_1 - x_3) & (x_2 - x_1) \\ (x_2 y_3 - x_3 y_2) & (x_3 y_1 - x_1 y_3) & (x_1 y_2 - x_2 y_1) \end{bmatrix} \tag{33}$$

The Δ_ℓ term is the determinant of the 3 × 3 matrix within the square brackets in (33). This determinant also equals twice the area of the ℓ-th triangle.

Employing the Euler-Lagrange formulation, we seek the nodal values in (31) about which the quadratic functional below is stationary,

$$Q(f) = \langle \nabla f, \nabla f \rangle - k_0^2 \langle f, f \rangle \tag{34}$$

Note that this functional has only first order derivatives inside of the $\langle *, * \rangle$ dot-product integrals over x, y. The proof that $Q(f)$ is stationary about the solution to (29), with BC's in (30), follows directly from the Euler equation, [14: pp.275–280],

$$\frac{\partial}{\partial x}\left[\frac{\partial Q}{\partial (D_x f)}\right] + \frac{\partial}{\partial y}\left[\frac{\partial Q}{\partial (D_y f)}\right] - \frac{\partial Q}{\partial f} = 0 \tag{35}$$

The stationary functional in (34) can also be obtained through applying Green's theorem (or multiple integration by parts) to the Galerkin equations considered in the previous subsection. The numerical approximation to the stationary solution is found by first substituting (31) into (34), to give

$$Q_a = \sum_{i=1}^{6}\sum_{j=1}^{7}\sum_{m=1}^{6}\sum_{n=1}^{7} f_{i,j}\, f_{m,n} \left\{\langle \nabla U_{i,j}, \nabla U_{m,n}\rangle - k_0^2 \left\langle U_{i,j}, U_{m,n}\right\rangle\right\} \tag{36}$$

We then differentiate this quadratic form with respect to each of the *unknown* nodal values of $f_{i,j}$ and $f_{m,n}$, setting the result to zero in each case. Each differentiation of the quadratic form will yield a linear equation representing the stationary requirement for the particular nodal value of f being considered. A linear system results, with one equation for each unknown nodal value as given by

$$\sum_{i=1}^{6}\sum_{j=1}^{7} f_{i,j} \left\{\langle \nabla U_{i,j},\ \nabla U_{m,n}\rangle - k_0^2 \langle U_{i,j},\ U_{m,n}\rangle\right\} = 0 \tag{37}$$

for m=2 to 5 and n=2 to 7. The terms in the summation involving *known* nodal values for $f_{i,j}$ can be transferred to the right hand side of (37) to put the equations in standard form. Note that for a given (i,j) node, only (m,n) nodes sharing at least one common element will provide a nonzero contribution to the moment integrations in (37). Thus, the resultant system matrix will be quite sparse, with most array elements being zero. As previously mentioned, this sparse matrix feature is produced by all finite methods when using basis and testing functions having compact support.

To put (37) into a block matrix form, let us denote the nodal unknowns across the j-th horizontal row of the mesh in Fig. 7 by the column vector (the T superscript indicates transpose)

$$\overline{F}_j = [f_{2,j}, f_{3,j}, f_{4,j}, f_5, j]^T \tag{38}$$

The matrix equation in (37) can then be written as a linear matrix-vector relationship between adjacent row vectors

$$\overline{\overline{A}}_j \cdot \overline{F}_{j-1} + \overline{\overline{B}}_j \cdot \overline{F}_j + \overline{\overline{C}}_j \cdot \overline{F}_{j+1} = \overline{P}_j \tag{39}$$

where the *block-matices* each have a banded structure

$$\overline{\overline{A}}_j \sim \begin{bmatrix} \times & \times & & \\ & \times & \times & 0 \\ 0 & & \times & \times \\ & & & \times \end{bmatrix} \tag{40a}$$

$$\overline{\overline{B}}_j \sim \begin{bmatrix} \times & \times & & \\ \times & \times & \times & 0 \\ 0 & \times & \times & \times \\ & & \times & \times \end{bmatrix} \tag{40b}$$

$$\overline{\overline{C}}_j \sim \begin{bmatrix} \times & & & \\ \times & \times & & 0 \\ 0 & \times & \times & \\ & & \times & \times \end{bmatrix} \tag{40c}$$

The nonzero "×" matrix elements as well as the components of the boundary condition vector, $\overline{P}_j$, are obtained in terms of the element integrals within the curly brackets in (37). Inside of the ℓ-th element, having $\overline{\overline{L}}$–matrix defined by (33), the integrands can be obtained directly from (32). Denoting the relationship between local (in element ℓ) and global node coordinates by $k = (m,n)$ and $q = (i,j)$, there results

$$U_k\, U_q = \overline{L}_k^T \cdot \begin{bmatrix} x^2 & xy & x \\ xy & y^2 & y \\ x & y & 1 \end{bmatrix} \cdot \overline{L}_q \tag{41a}$$

$$\nabla U_k \cdot \nabla U_q = \overline{L}_k^T \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \cdot \overline{L}_q \tag{41b}$$

The matrix entries are thus assembled from element integrations of the type

$$I_{r,s} = \iint\limits_{\Delta} x^r\, y^s\, dx\, dy \tag{42}$$

which are available in tabular form in a number of references on finite elements, e.g. [11].

Having loaded the block matrices relating adjacent row vectors of unknowns, the global matrix structure will have a tri-block form

$$\begin{bmatrix} \overline{\overline{B}}_2 & \overline{\overline{C}}_2 & & & & & \\ \overline{\overline{A}}_3 & \overline{\overline{B}}_3 & \overline{\overline{C}}_3 & & 0 & & \\ & \overline{\overline{A}}_4 & \overline{\overline{B}}_4 & \overline{\overline{C}}_4 & & & \\ & & \cdot & \cdot & \cdot & & \\ & 0 & & \cdot & \cdot & \cdot & \\ & & & & \cdot & \cdot & \cdot \\ & & & & & \overline{\overline{A}}_7 & \overline{\overline{B}}_7 \end{bmatrix} \cdot \begin{bmatrix} \overline{F}_2 \\ \overline{F}_3 \\ \overline{F}_4 \\ \cdot \\ \cdot \\ \cdot \\ \overline{F}_7 \end{bmatrix} = \begin{bmatrix} \overline{P}_2 \\ \overline{P}_3 \\ \overline{P}_4 \\ \cdot \\ \cdot \\ \cdot \\ \overline{P}_7 \end{bmatrix} \tag{43}$$

The finite element solution thus comes down to inverting a matrix. As was the case for the finite difference example, the global array is mostly filled with zeros. This is in contrast to integral equation methods, which produce full matrix structures. A sparse matrix allows highly economical inversion, for even very large matrix order, by any of a number of different algorithms [15]. In addition, by properly ordering the nodes, as was done here, the matrix can often be made to have a block structure. A block structured matrix can be efficiently inverted by way of the Riccati transform algorithm [16] (also see Chapters 2 and 6 of this book).

In developing the theory and examples for the finite methods, there was no discussion concerning expected accuracy. We will now briefly address this topic at an introductory level. Additional developments will be made in the following chapters for the specific cases to be considered.

f. Basics of Error Analysis

Let us now consider some of the concepts and terminology associated with computational error in numerical solutions. We will begin by using the simple example of the finite difference approximation of the Helmholtz equation as a vehicle to illustrate some of these ideas. The discussion could readily be expanded to include all finite and integral methods, as applied in both the frequency- and the time-domains.

In general, the "pointwise error" in a numerical solution at each node is simply the difference between the exact and computed values of the solution at the node. Referring back to Fig. 2, the pointwise error at node (x_i, y_j) becomes $\delta f_{i,j} = f(x_i, y_j) - f_{i,j}$. There are a number of ways to use the pointwise errors to attain some measure of the "global error" in the whole numerical solution. Some of the most popular and meaningful ways of defining global error are through "energy", "mean-square", and "root-mean-square" (RMS). The energy in the exact solution at the nodes would be defined, for the FD-FD example first considered, as

$$\Omega_f = \sum_{i=2}^{5} \sum_{j=2}^{6} |f(x_i, y_j)|^2 \tag{44}$$

where $|f(x,y)|^2 = f(x,y) \cdot f^*(x,y)$ indicates magnitude squared. Using the vector definition in (9) for the ordered nodal values of f, the energy becomes $\Omega_f = \overline{F} \cdot \overline{F}^*$.

Extending this energy definition to the pointwise error at each node gives

$$\Omega_{\delta f} = \sum_{i=2}^{5} \sum_{j=2}^{6} |\delta f_{i,j}|^2 \tag{45}$$

The *relative* mean-squared error is then defined as the ratio of error energy over solution energy. This leads to the definition for RMS error, which is often expressed as a percentage,

$$\varepsilon_{RMS} = \sqrt{\frac{\Omega_{\delta f}}{\Omega_f}} \times 100\% \tag{46}$$

Errors in the numerical solution result from two sources: discretization and roundoff. The net effect of these error sources depends upon several interrelated factors, including:

(1) Properties of the physical system being modeled,
(2) The basis and weighting functions being used,
(3) The level of segmentation being employed (number and spacing of nodes and/or time-steps) and

(4) The numerical resolution (wordlength) of the computer's numeric data processor and/or co-processor.

Physical system properties affect the *conditioning* (for frequency domain formulations) or the *stability* (for time-domain representations) of the numerical solution. Poorly conditioned or marginally stable physical systems tend to be intolerant of small errors in either the discretized model or the numerical roundoff of the solution process.

For frequency domain problems, this is manifested as error amplification in the linear system solution, as is reflected in the "condition number" of the system matrix [17]. Consider the FD-FD solution for the Helmholtz equation, which resulted in $\overline{\overline{A}} \cdot \overline{F} = \overline{G}$, as presented in (11). Relative errors in the solution for $\overline{F}$ will involve an amplification of the relative errors in computing both $\overline{\overline{A}}$ and $\overline{G}$. This amplification factor is the condition number of the A-matrix.

To be more specific we need to define an appropriate *norm*, denoted by $|| \quad ||$, for the arrays involved. The usual Euclidean vector norm just equals the RMS value of the vector (square-root of the energy). As an example, for the vector $\overline{F}$, as defined by the ordered nodal values in (9), we have $||F|| = \sqrt{\Omega_f}$. The norm of a square matrix, denoted by $||A||$, is generated in terms of the maximum norm of all vectors found from the product $\overline{\overline{A}} \cdot \overline{x}$, where $||x|| = 1$. The condition number of $\overline{\overline{A}}$ is then defined by $C(A) = ||A|| \cdot ||A^{-1}||$. Using this definition for the condition number, it can be shown [18] that the relative error in the solution is bounded by:

$$\frac{||\delta F||}{||F||} \leq \frac{C}{1 - C||\delta A||/||A||}\left[\frac{||\delta A||}{||A||} + \frac{||\delta G||}{||G||}\right] \tag{47}$$

where $\delta\overline{\overline{A}}$ and $\delta\overline{G}$ are the respective arrays formed from the cumulative numerical errors in the evaluation of $\overline{\overline{A}}$ and $\overline{G}$. The condition number is thus seen to amplify the errors in both matrix and driving vector approximations.

As mentioned earlier, the conditioning of a physical system is related to the sensitivity of the solution function to small changes in the input data. A closely related term is "ill-posed", which refers to a mathematical problem whose solution is not unique, at least not from the way it was specified. An example of a poorly conditioned physical problem would be in computing the field strength and modal configuration internal to a low-loss metal cavity with a small excitation port.

As a resonant frequency is approached, large relative changes in the internal field will be produced by small relative changes in the physical system, such as the cavity and excitation port geometry. This is a physical case which results in a poorly conditioned numerical solution, accompanied by a large condition number.

An additional factor which adversely influences the condition number is the rank of the numerical solution matrix. As the mesh density is increased for a fixed physical problem, an increasing matrix size results which is accompanied by a growing condition number. There is a two-fold reason for this. First, the increased number of calculations needed to invert the larger system will tend to enhance computational round-off errors. Secondly, after a certain point a further increase in the number of discrete equations begins to become somewhat redundant. This results in a lower degree of linear independence of the individual equations in the matrix system.

The numerical solution errors induced by the WRM discretizement process can be expressed in terms of an "order" which is inversely related to some power of the number of basis functions, N employed in (23). Assuming that the mesh resolution is enhanced by using more nodes one would expect the error in the numerical solution to decrease, at least to some lower bound. More specifically, consider a second-order PDE system in M-dimensions, such as either the Helmholtz equation or the wave equation in the examples that were considered. Note that coupled first-order state-equation systems (such as Maxwell's curl equations) are usually also second-order decoupled systems. For a second-order system, if p-th order polynomial basis functions are used for $\overline{\overline{U}}_n(\overline{r})$, while q-th order polynomial weighting functions are used for $\overline{\overline{W}}_k(\overline{r})$, then it follows from results derived in [13] that the discretization error is of order:

$$\varepsilon'_{RMS} = O(N^{-\alpha}) \tag{48a}$$

where

$$\alpha = \mathrm{Min}\{M(p+1), M(p+q)\} \tag{48b}$$

An obvious constraint for convergence to occur (namely, $\varepsilon'_{RMS} \to 0$ as $N \to \infty$) is that $\alpha \geq 1$. This requirement is satisfied for each of the three special cases of the WRM that we considered in subsection 1.2d. Recall, for instance, the 2-D example problems that we have previously encountered (with $M=2$). The FEM case used the *Galerkin* method with piecewise linear basis and weighting functions ($p=q=1$). This gives

$$\varepsilon'_{RMS} = O(N^{-4}) \qquad \text{Galerkin FEM Case} \tag{49}$$

which indicates a strong tendency towards quick convergence, as N is increased. On the other hand, the finite-difference solutions (both time and frequency domain examples) employed *point collocation* using quadratic basis functions ($p=2$) and Dirac weighting functions ($q=-1$), resulting in weaker convergence:

$$\varepsilon'_{RMS} = O(N^{-2}) \qquad \text{Finite Difference Cases} \tag{50}$$

A similar result to (50) is obtained for *subdomain collocation* using linear basis ($p=1$) and pulse weighting ($q=0$) functions.

It should be emphasized that the error in (48a) is *exclusive of* numeric round-off error and the associated condition number error amplification. This error is thus based on the assumption of an infinite wordlength computer. In reality, as N is increased the condition number also tends to increase, thus offsetting some of the reduction of error found by using more nodes. As N is increased further still, there comes a point where the growing round-off errors will overcome any further gains in accuracy, and the solution error will begin to creep upward with N. This behavior is illustrated in Fig. 10, where the effect of computer wordlength is depicted by different curves. In general, as the numeric wordlength is decreased the point of optimum accuracy occurs at lower N, beyond which the error increases. The effect of ill-conditioning in the physical system being discretized is to raise all of the $\varepsilon_{RMS}(N)$ curves.

In the time-domain case, poor stability of the system translates into more rapid divergence of the transient numerical solution than for a more stable system, all other parameters being equal. For the case of the wave equation, we have seen in subsection 1.2c that the Courant limit on space-time step size provides a well defined condition for stability – one which carries over into the temporal leapfrog solution of Maxwell's equations. There are similar, but not identical, stability conditions for *parabolic* PDE systems such as the diffusion equation. An excellent treatment of stability effects in time-domain integral equation solutions is given in [19] with much of this being valid for finite methods as well.

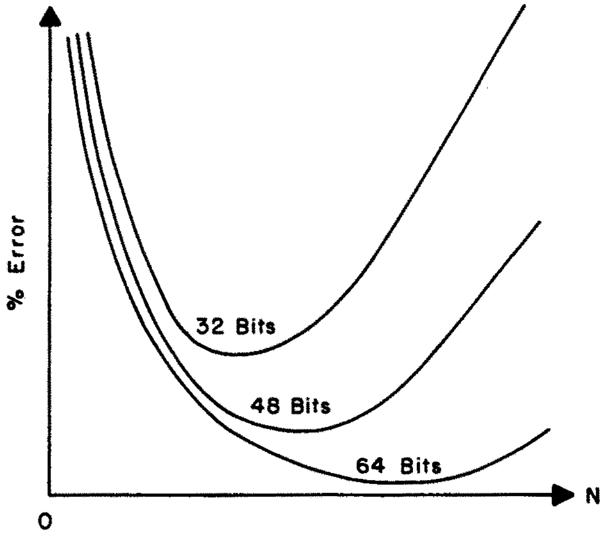

Figure 10 Solution convergence versus wordlength.

1.3 Mesh Termination

Having considered basic concepts related to discretizing frequency-domain (boundary-value) and time-domain (initial-value) electromagnetic problems in closed spatial regions, we will now turn our attention to some techniques for employing finite method solutions to field problems in open, unbounded regions.

a. *Unimoment Method*

The unimoment method, as developed by Mei [20], provides a self-consistent approach to coupling interior and exterior frequency-domain field problems across separable surface interfaces. This method was employed by Chang and Mei [21], Stovall and Mei [22] and Morgan and Mei [16] to open region scattering and antenna problems involving inhomogeneous 2-D and axisymmetric 3-D dielectrics. Further applications of the unimoment method, using finite elements, have been made to problems involving raindrop scattering [23], microwave energy deposition in the human head [24], scattering by multiple bodies [25] and even buried objects [26]. A recent unimoment application, considered by Fleming in Chapter 2 of this book, is scattering from composite axisymmetric objects having both metallic and penetrable parts.

As developed in the unimoment method, the exterior region fields are represented by a functional expansion in one of the separable coordinate systems for the vector Helmholtz equation [14: sect. 5.1]. The spatial interface for coupling the interior numerical solution to the unbounded exterior region is thus a constant coordinate surface of the separable system employed in the outside expansion. Spherical interfaces were utilized in [16] and [22–26] due to the relative ease of generation of exterior region spherical harmonic field expansions.

To understand the conceptual basis of the unimoment method, consider the solution of a scattering problem involving a 2-D cylindrical penetrable object of arbitrary cross section which is, perhaps, inhomogeneous. For either TE or TM (to $\hat{z}$) cases, having respectively $\overline{H} = f\,\hat{z}$ or $\overline{E} = f\,\hat{z}$, the undriven scalar Helmholtz equation in (29) is again applicable, but with a variable wavenumber, $\kappa = \sqrt{\epsilon(r,\theta)\,\mu(r,\theta)}$, within the scattering object

$$\{\nabla^2 + \kappa^2(r,\theta)\}\; f(r,\theta) = 0 \tag{51}$$

where we are using circular coordinates (r,θ).

The unimoment solution proceeds by enclosing the scatterer within a separable mathematical boundary, which we will choose to be circular in the 2-D (r,θ) cross section, as is illustrated in Fig. 11. Notice that there are two concentric circles, of radius r_1 and r_2, both of which enclose the scattering object. With specified Dirichlet BC's on the outer boundary, $f(r_1,\theta) = g(\theta)$, a finite method can be used to solve for nodal values of $f(r,\theta)$ for $r < r_1$. A necessary attribute of the interior mesh construction is that a set of the solution nodes lies on the inner circle, $r = r_2$. An example mesh, using linear triangular elements, is shown in Fig. 12.

Of course, in the scattering problem we do not know the total field, $f_1(\theta)$, on the r_1 boundary. With a knowledge of the incident field, we are to solve for the scattered field. To do this, we will first express the total field in the free space region (with wavenumber k_0) outside of the smallest circle centered at $r = 0$ which encloses the scatterer. The total field equals the sum of the known incident field and a truncated cylindrical harmonic expansion for the unknown scattered field

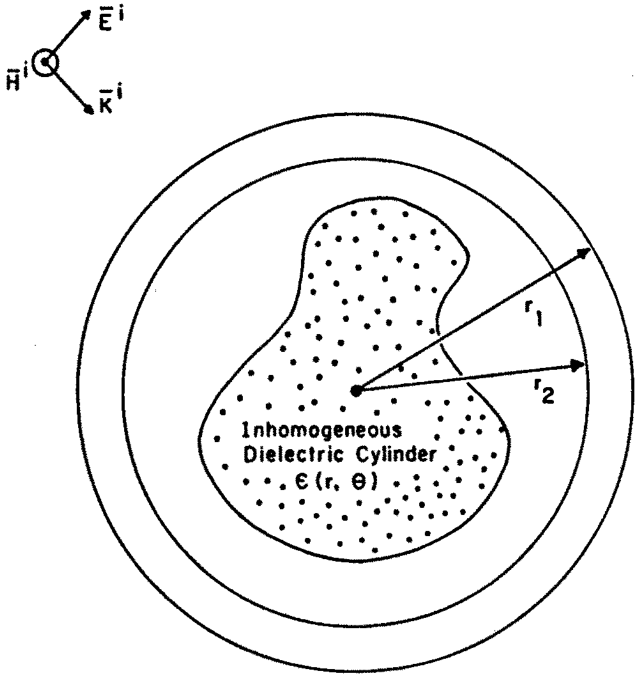

Figure 11 Unimoment matching contours for cylinder scattering problem.

$$f(r,\theta) = f^i(r,\theta) + a_0\, C_0(r) + \sum_{n=1}^{N} a_n\, C_n(r,\theta) + b_n\, S_n(r,\theta) \qquad (52a)$$

where the cylindrical harmonics for the scattered field are

$$C_n(r,\theta) = H_n^{(2)}(k_0 r)\ \cos(n\theta) \qquad (52b)$$

$$S_n(r,\theta) = H_n^{(2)}(k_0 r)\ \sin(n\theta) \qquad (52c)$$

with $H_n^{(2)}$ equal to the Hankel function of the second kind. It is important to note that each cylindrical harmonic term satisfies the 2-D radiation conditions in the far-field of the scatterer, as $r \to \infty$:

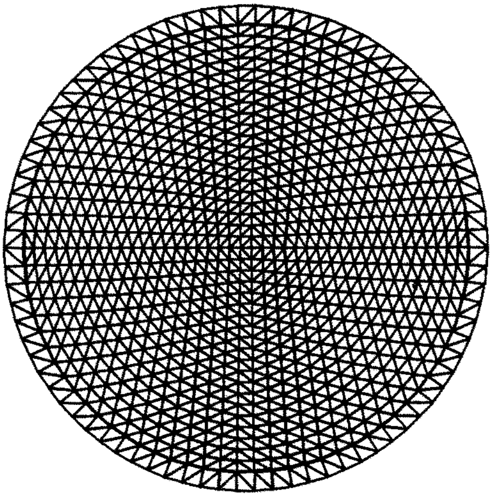

Figure 12 Cylindrical finite element mesh.

$$f(r,\theta) \to \chi(\theta)\,\frac{e^{-jk_0 r}}{\sqrt{k_0 r}} \tag{53a}$$

$$\frac{E_\theta}{H_z} \to \eta_0 \quad \text{(TE case)} \tag{53b}$$

$$-\frac{E_z}{H_\theta} \to \eta_0 \quad \text{(TM case)} \tag{53c}$$

where $\eta_0 = 120\pi\Omega$. Notice that the series representation in (52a) is only an approximation due to the truncation of the series at $N+1$ terms. In practice, this truncated series will converge rapidly to the exact scattered field for increasing N, when $N > k_0 r_2$.

To find the scattered field coefficients, a_n and b_n, we first solve the interior region problem for $2N+2$ separate BC's on $r = r_1$. These BC's are composed of the incident field, $f^i(r_1,\theta)$, and each of the scattered field modes: $C_n(r_1,\theta)$ for $n = 0,1,2,\cdots,N$ and $S_n(r_1,\theta)$ for $n = 1,2,\cdots,N$ (all evaluated on the outer boundary, r_1). The numerical solutions in the interior, $r < r_1$, which correspond to these applied BC's, are indicated by a tilde overbar. For example, the incident field BC, $f^i(r_1,\theta)$, produces an interior solution $\tilde{f}^i(r,\theta)$ for $r < r_1$ while a BC of $S_n(r_1,\theta)$ produces $\tilde{S}_n(r,\theta)$.

Using the principle of superposition, the numerical solution for the total field inside of the outer boundary will be given by

$$\tilde{f}(r,\theta) = \tilde{f}^i(r,\theta) + a_0\,\tilde{C}_0(r) + \sum_{n=1}^{N} a_n\,\tilde{C}_n(r,\theta) + b_n\,\tilde{S}_n(r,\theta) \tag{54}$$

To set up conditions obeyed by the the a_n and b_n coefficients, we simply equate the numerical solution in (54) to the analytical solution in (52) along the circular contour, $r = r_2$, resulting in

$$\sum_{n=1}^{N} a_n\left\{C_n(r_2,\theta) - \tilde{C}_n(r_2,\theta)\right\} + b_n\left\{S_n(r_2,\theta) - \tilde{S}_n(r_2,\theta)\right\}$$

$$+\,a_0\,C_0(r_2) = \tilde{f}^i(r_2,\theta) - f^i(r_2,\theta) \tag{55}$$

The unknown scattered field coefficients may be approximated by using the weighted residual method (WRM) to generate a system of linear equations. To employ the WRM, we integrate (55) with respect to each of $2N+1$ linearly independent weighting functions, $W_m(\theta)$ for $m = 0,1,2,\cdots,2N$. In this case our inner product integrations will be defined by

$$\langle W_m(\theta), Z(\theta)\rangle = \int_0^{2\pi} W_m(\theta)\,Z(\theta)\,d\theta \tag{56}$$

Using this definition, and representing the difference terms in (55) by a Δ notation, the WRM equations become

$$\sum_{n=1}^{N} a_n\,\langle W_m(\theta), \Delta C_n(\theta)\rangle + b_n\,\langle W_m(\theta), \Delta S_n(\theta)\rangle$$

$$+\,a_0\,\langle W_m(\theta), \Delta C_0(\theta)\rangle = \langle W_m(\theta), \Delta f^i(\theta)\rangle \tag{57}$$

enforced for $m = 0$ to $2N$.

By selecting Dirac function weights, $W_m(\theta) = \delta(\theta - \theta_m)$ with θ_m representing nodes along $r = r_2$, (57) will provide the point-matched form of the solution. A more accurate approach is to enforce (55) in the least-squares sense over the circle, including in-between the nodes. This is done by forming the least squares residual

$$\Omega_{\Delta f} = \int_0^{2\pi} |f(r_2,\theta) - \tilde{f}(r_2,\theta)|^2 \, d\theta \tag{58}$$

Upon substituting (57) and (58) into this equation, a quadratic form in the coefficients will result. The minimization of (58) is found by differentiating with respect to each of the unknown coefficients and nulling the result. In such a case, there results a system of linear equations for the coefficients having the form of (57), but with specified weighting functions which are proportional to the complex conjugates of the Δ-function differences

$$W_m(\theta) = \begin{cases} \Delta C_m^*(\theta) & \text{for } m = 0, 1, 2, \cdots, N \\ \\ \Delta S_m^*(\theta) & \text{for } m = N+1, \cdots, 2N+1 \end{cases} \tag{59}$$

The integrations to evaluate the matrix elements in (59) are performed either numerically, or semi-analytically by using the basis function expansions employed in the interior solution to represent both the difference functions and the weighting functions. In either case, the resultant $2N+1$ square matrix can be inverted to obtain the scattered field coefficients. The scattered field at any exterior point may then be obtained from its expansion in (52) and, if desired, the interior field can be found by using the weighted superposition of the stored interior field solutions. For $r \rightarrow \infty$, we have the well known asymptotic approximation [2–3]

$$H_n^{(2)}(k_0 r) \rightarrow \sqrt{\frac{2\jmath}{\pi k_0 r}} \, \jmath^n \, e^{-\jmath k_0 r} \tag{60}$$

which can be used to obtain far-field expressions for the scattering pattern, as defined in (53a),

$$\chi(\theta) = \sqrt{\frac{2\jmath}{\pi}} \left\{ a_0 + \sum_{n=1}^{N} \jmath^n \left[a_n \cos(n\theta) + b_n \sin(n\theta) \right] \right\} \tag{61}$$

In the original unimoment method [20], the expansion coefficients were found by equating the analytical radial derivative of (52a) at r_2 to

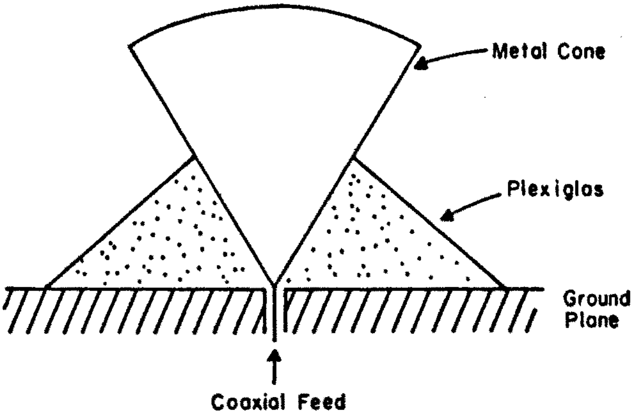

Figure 13 Experimental dielectrically loaded imaged conical antenna.

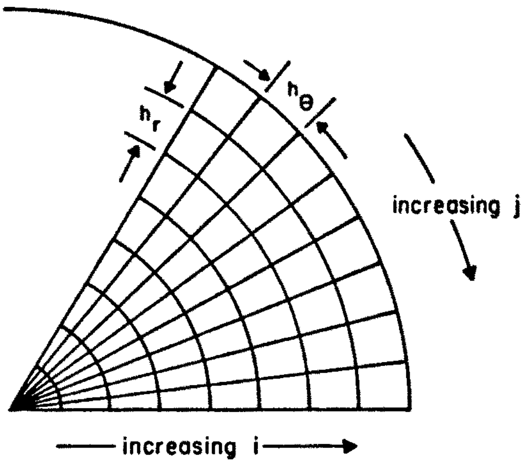

Figure 14 Finite difference mesh for the imaged conical antenna.

the finite difference approximation to the radial derivative found from use of the numerical solution in (54). The above procedure for matching fields, rather than derivatives, was adopted to both simplify the technique and to avoid extra error incurred by use of a finite difference derivative.

As mentioned earlier, the unimoment method has been employed in several computational efforts. One of the earliest of these involved

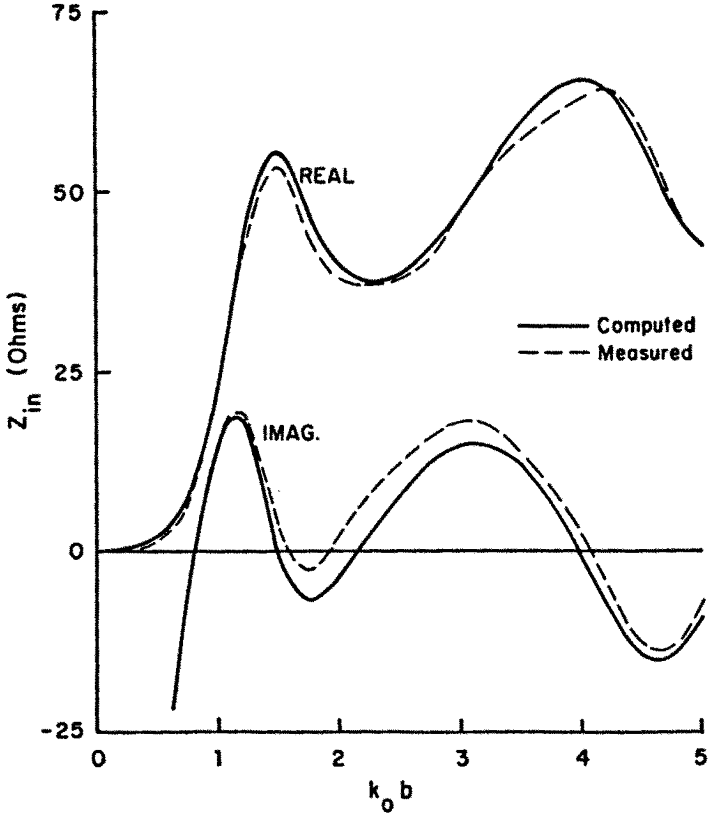

Figure 15 Comparison of measured and computed input impedance for loaded biconical antenna of half-height *b*.

the finite difference solution for radiation and input impedance of a finite length biconical antenna, loaded by various inhomogeneous dielectric configurations [22]. One such structure is depicted in Fig. 13. Since both the fields and material structure are axisymmetric (invariant to the ϕ-coordinate) the solution domain can be reduced to a single meridian plane, (r, θ) in spherical coordinates. A section of the the finite difference mesh is shown in Fig. 14. The interior region solution for this antenna problem was formulated using a special case of the "coupled azimuthal potential" (CAP) formulation, where the vector $\overline{E}$ and $\overline{H}$ fields are represented using two coupled scalar potential functions which are related to the azimuthal $(\widehat{\phi})$ field components. For the special case of axisymmetric fields in rotationally-symmetric material,

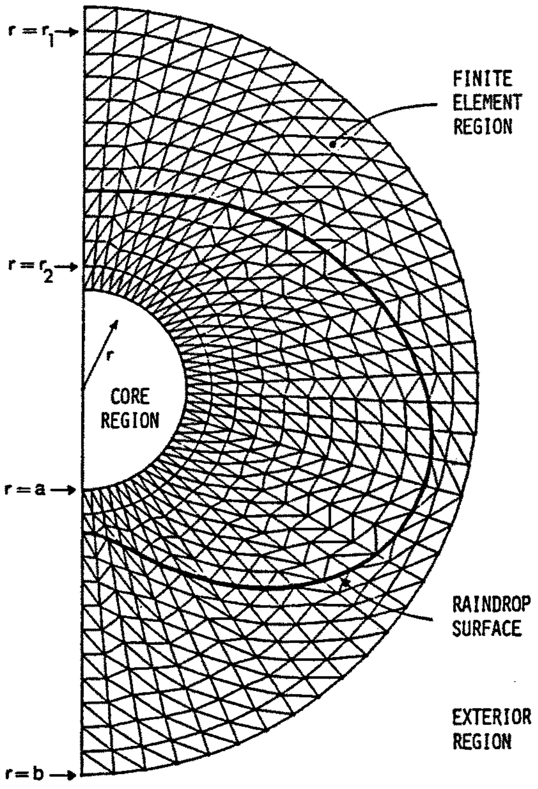

Figure 16 **Semi-annular conformal finite element mesh.**

as is found for the biconical antenna structure, only one azimuthal potential is needed. A sample result from [22] is shown in Fig. 15, which compares the computed and measured input impedance of a plexiglas loaded biconical antenna. The computation was performed at discrete frequencies over a 10:1 range, wherein the bicone height ranged from $.16\ \lambda_0$ to $1.6\ \lambda_0$.

Details of the original CAP formulation for axisymmetric material are discussed in Chapter 2 of this book, as well as in [27]. A recent extension of the CAP formulation to arbitrary 3-D inhomogeneous isotropic media is the subject of Chapter 6.

A second application of the unimoment method is that of scattering by inhomogeneous bodies of revolution [16]. This effort employed a tri-regional finite element mesh in the (r, θ) meridional plane, as shown

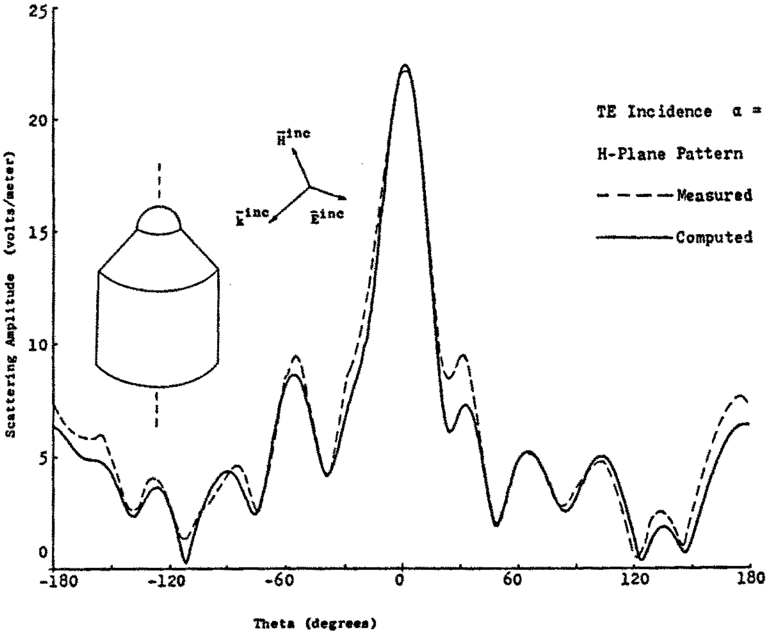

Figure 17 **Comparison of finite element computation for bistatic scattering to experimental results at 9.33 GHz.**

in Fig. 16. The CAP formulation was employed to represent the nonaxisymmetric fields using a Fourier series in the ϕ-coordinate. Spherical harmonic expansions were used to represent the scattered field outside of the mesh and the total fields within the spherical "core" region surrounded by the mesh. The sets of coefficients used in these field expansions were found by applying the various expansion modes for the potentials as BC's along the contours $r = a$ and $r = b$. In addition, the various incident fields being considered were applied along $r = b$. A finite element solution for each applied BC was then evaluated along the inner contours, $r = r_1$ and $r = r_2$. The total fields were assembled from these numerical solutions and equated in the least-squares sense to the original analytical expansions, resulting in a matrix equation for the coefficients. Such a procedure is a direct extension to that employed in our simple unimoment example which resulted in (59). Numerous comparisons to bistatic scattering measurements at 9.33 GHz were made for solid and hollow dielectric bodies of various shapes. A typical result is shown in Fig. 17, where the bistatic scattering is from a plexiglas body having cylindrical, conical and hemispherical portions.

Although the unimoment method offers a straightforward procedure for finite element or finite difference mesh termination in unbounded regions, the numerical solution within a 2-D circular or 3-D spherical interior region becomes inefficient for scatterer or radiator shapes that occupy only a small portion of the enclosed region. Examples of this are highly elongated or flattened structures such as thin cylinders and flat discs. Although it is possible to utilize a separable surface which is not circular or spherical to increase the numerical efficiency of the interior region solution, this will be offset by additional requirements in both generating the special functions that are needed in the exterior expansion and in computing the required moment integrations of these functions over the interface.

b. Boundary Integral Equation

Another approach to mesh termination is to replace the use of an exterior region modal expansion in the unimoment method with a surface integral equation system on the mesh boundary. After all, one purpose of the exterior field expansion was to ensure that the resultant numerical solution satisfied the radiation conditions in the far-field region of the scatterer or antenna. A properly formulated integral equation system will also ensure this, and will do so without any explicit restriction to use of a separable mesh boundary surface. Let us now look at this idea and the results of an early effort [28] in using it for scattering computations from 3-D axisymmetric objects.

A recent application to 2-D scattering by irregularly shaped and inhomogeneous objects is presented in detail in Chapter 3 of this book wherein the term "boundary element method" (BEM) is used. In the context of Chapter 3, BEM refers to the discretization procedure for solving surface (in 3-D) or contour (in 2-D) integral equations. The boundary element terminology is also used in Chapter 7, where it indicates special case finite elements being used at material interfaces to couple otherwise regular spatial finite difference lattices between homogeneous regions.

To consider the use of a boundary integral equation for mesh termination, let us return to our previous unimoment method example. This example considered 2-D scattering by an inhomogeneous dielectric object, as illustrated in Fig. 11. Rather than using a circular outer

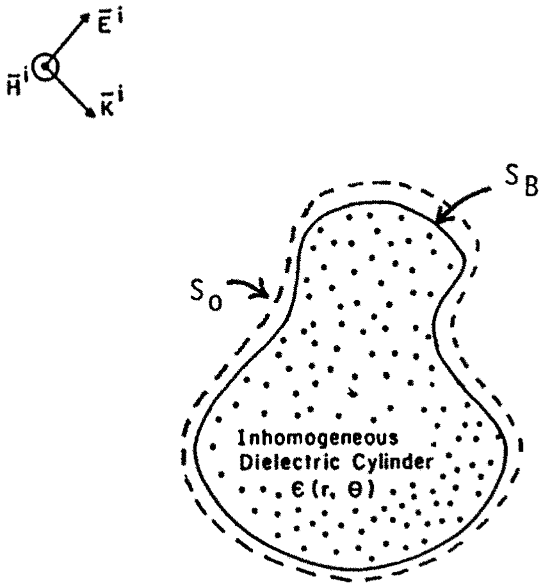

Figure 18 **Boundary integral equation contour enclosing inhomogeneous dielectric cylinder.**

mesh boundary, let us use a conformal boundary, S_0, as per Fig. 18. This boundary must enclose the scatterer and may, in fact, be congruent with the boundary surface of the object, to be denoted as S_B. Within the enclosed region, which contains inhomogeneous $\epsilon(\bar{r})$ and $\mu(\bar{r})$, we can use the finite element method to approximate the solution of the Helmholtz equation in (51) (but using x, y coordinates) for any specified Dirichlet BC on S_0. Let us define this BC using a basis function expansion (with unknown c_k coefficients) to represent the *total* field along the S_0 contour, having position variable s,

$$f(\bar{r})\Big|_{S_0} = \sum_{k=1}^{N} c_k \, g_k(s) \tag{62}$$

where g_k are the basis functions. These basis functions may have either full-range support (e.g. Fourier modes) or compact-support such as polynomial pulse functions. For the latter case, it may be expedient to use the unit amplitude finite element basis functions, u_k, associated with the S_0 nodes as the g_k's. With such a choice, the c_k's in (62) become the nodal values of the total field on S_0. The use of the finite element basis functions on the S_0 contour is adopted in Chapter 3. We will follow a more general route here in explaining the concept.

As in the unimoment example, let us use a tilde overbar to denote the finite element numerical solution which results from a given functional BC. For example, a BC of $g_k(s)$ on the boundary S_0 is said to produce a finite element solution within S_0 given by $\tilde{g}_k(\bar{r})$. Explicit enforcement of the radiation conditions in (52) will result by requiring the finite element solution to also satisfy a proper integral equation on the boundary, S_0. Such integral equations are readily found by using a principle-value limiting process, where an exterior field point is brought onto the S_0 surface [29]. For the special case of a smooth surface, one form of the integral equation is

$$\frac{1}{2}f(s) + \int_{S_{PV}} \left[G(s,s')\frac{\partial f(s')}{\partial n'} - f(s')\frac{\partial G(s,s')}{\partial n'} \right] ds' = f^i(s) \tag{63}$$

where s and s' are both points on S_0 and $\widehat{n}'$ is the outward unit normal at s'. The free-space Green's function is given by

$$G(s,s') = -\frac{j}{4} H_0^{(2)}(k_0|\bar{r} - \bar{r}'|) \tag{64}$$

with $|\bar{r} - \bar{r}'|$ equal to the distance between the s and s' contour points. The S_{PV} notation indicates Cauchy principle-value (PV) integration, where the point-wise singularity at $s = s'$ is removed. A similar integral equation is found for the TE case. Additional forms of 2-D integral equations are derived in Chapter 3.

To implement (64), using the BEM, the $\partial f/\partial n'$ integrand term needs to be approximated from the finite element solution. This approximation is obtained through numerically differentiating each $\tilde{g}_k(\bar{r})$ (due to the specified $g_k(s)$ BC). Using superposition, in terms of the still unknown c_k's, the integrand term becomes,

$$\left.\frac{\partial \tilde{f}}{\partial n'}\right|_{\bar{r}' \in S_0} = \sum_{k=1}^{N} c_k \frac{\partial \tilde{g}_k(s')}{\partial n'} \tag{65a}$$

where

$$\frac{\partial \tilde{g}_k(s')}{\partial n'} = \left. \widehat{n}' \cdot \nabla \tilde{g}_k(\bar{r}) \right|_{S_0} \tag{65b}$$

The BEM discretization is completed by substituting (62) and (65a) into (63), followed by rearrangement of terms to give

$$\sum_{k=1}^{N} c_k\, I_k(s) = f^i(s) \tag{66}$$

with

$$I_k(s) = \frac{1}{2}\tilde{g}_k(s) + \int_{S_{PV}} \left[G(s,s')\,\frac{\partial \tilde{g}_k(s')}{\partial n'} - \tilde{g}_k(s')\,\frac{\partial G(s,s')}{\partial n'} \right] ds' \tag{67}$$

A system of linear equations for the c_k's can be generated by weighted residual enforcement of (67)

$$\sum_{k=1}^{N} c_k \left\langle W_m(s), I_k(s) \right\rangle = \left\langle W_m(s), f^i(s) \right\rangle \quad \text{for } m = 1,2,\cdots,N \tag{68}$$

where the weighting functions $W_m(s)$ can correspond, for instance, to any of those discussed in subsection 1.2d. Once the total field solution is found on S_0, the far-zone scattered field can be obtained by using a simplified Green's function integration. This integration formula is developed in the next subsection.

Note that (68) represents a full matrix system. When using the finite element basis functions for the g_k's, the order, N will equal the number of nodes on the S_0 contour. If the material inside of S_0 had been *homogeneous*, we could have formulated the problem using just an integral equation on S_0; there would have been no need to take on the extra effort required for the finite element solution. On the other hand, for the case of a general inhomogeneous material inside of S_0, an integral equation approach would incorporate all unknowns spanning the internal region. A *very large* full matrix could result. The hybrid approach just considered incorporates a sparse matrix solution for the interior region while needing a full matrix only on the enclosing boundary. For problems involving multi-wavelength sized inhomogeneous scatterers, the computational savings incurred can be very significant.

An application of the boundary integral technique to scattering by axisymmetric 3-D objects was developed by Morgan, Chen, Hill and

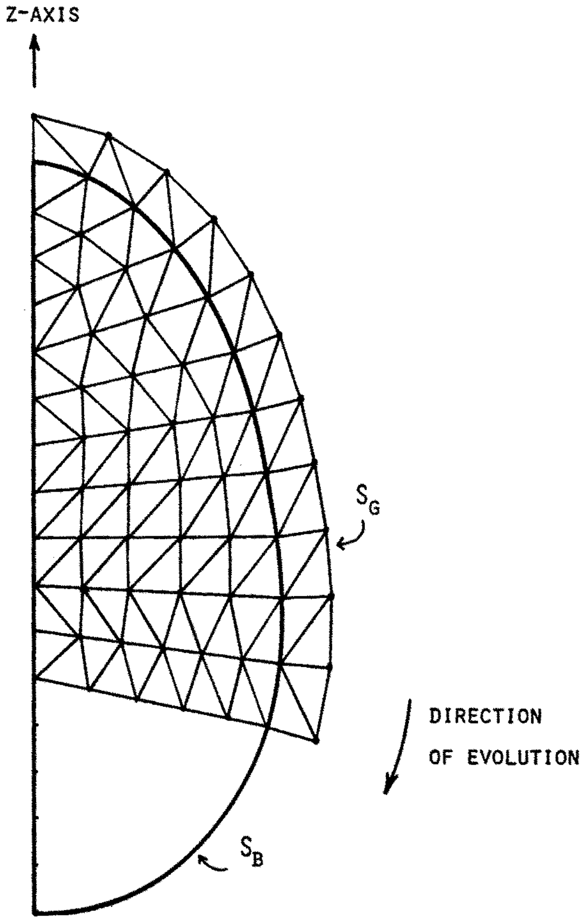

Figure 19 Surface conforming finite element mesh.

Barber [28]. This hybrid approach, named the "finite element boundary integral" (FEBI) method combines a finite element solution of the interior region with the surface integral equations found in Waterman's "extended boundary condition method" (EBCM) [30].

The FEBI procedure allows the use of a surface interface that conforms to the outer boundary of the scattering object, as is shown for the finite element mesh in Fig. 19. The finite element solution proceeds in a similar manner to that of the unimoment method, with incident field and scattered field spherical harmonic expansion modes being applied as BC's at the outer boundary, S_G. Numerical solutions are then found at the surface of the scattering body, S_B, for each of these applied BC's.

It should be noted, that for a specified origin, a truncated spherical harmonic scattered field expansion is uniformly convergent (with increasing truncation index) *outside* of the smallest geometrical sphere, centered about the origin, which encloses the scatterer. Since the boundaries at S_G and S_B do not usually conform to a spherical surface, the truncated spherical harmonic expansion for the scattered fields, as applied on S_B, may not be *complete* at all points on that surface. Such a phenomenon is related to the classical "Rayleigh hypothesis" [31]. As a result, we may not be able to obtain as a good match between the numerical solution at S_B and the original truncated analytical expansion, as we were able to do on the spherical surface for the unimoment method.

To evaluate the expansion coefficients for the boundary field in the FEBI, we can use a system of two combined field integral equations, as employed in the EBCM [30]. These integral equations are vector field 3-D versions of that in (63); they relate the tangential fields just inside of the boundary S_B to that just outside, without making use of a knowledge of the material structure inside of S_B.

The FEBI method has been shown to work well for scattering calculations involving moderately elongated lossy dielectric objects. An example computation, with comparison to that performed using the EBCM, is illustrated in Fig. 20. This approach tends to have convergence difficulties if the surface interface becomes extremely elongated or flattened (e.g. length to diameter ratios exceeding about 10). As just discussed, the culprit in this failure is the poorly convergent exterior region spherical harmonic expansion when used to represent the field over the surface of the scattering body.

The use of poorly convergent spherical harmonics as basis functions on radically non-spherical surfaces in the FEBI technique was a result of the attempt to combine the usual EBCM concept with finite elements. A better approach would be to have used the innate basis functions of the FEM on the surface. This has been done for 2-D cylinder scattering, as it appears in Chapter 3 and in a recent paper by Peterson [32].

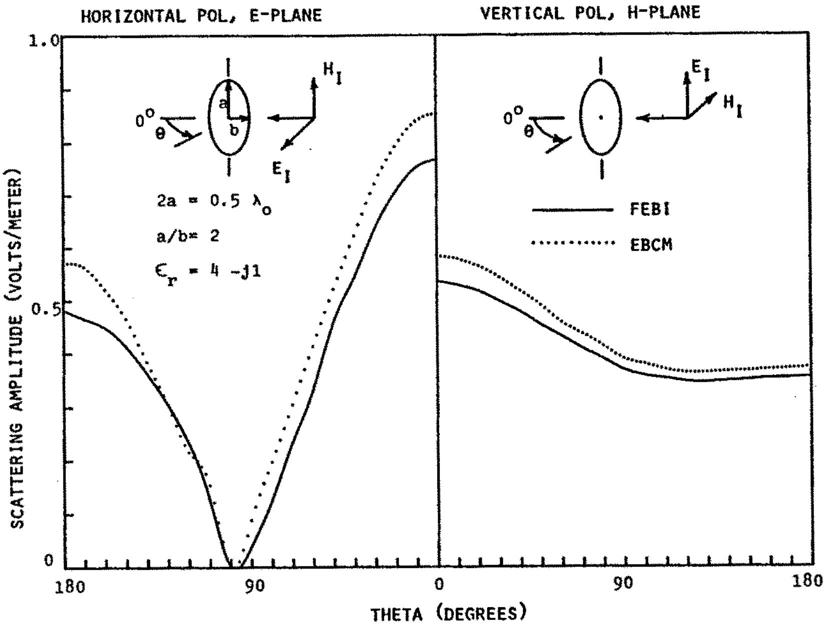

Figure 20 Scattering amplitude comparison for FEBI computation.

c. *Field Feedback Formulation*

The field feedback formulation (F^3) combines features from the unimoment method and the boundary integral equation technique to allow increased flexibility in coupling interior and exterior region time-harmonic field solutions [33]. As with these other two approaches, the interior boundary-value problem is initially decoupled from the outside region. The interior problem may then be formulated and computed using the most expedient approach that can accurately accommodate the level of material complexity that is present. Instead of using scattered field harmonic expansions (e.g. cylindrical in 2-D and spherical in 3-D), the exterior region scattered field is represented in terms of modes generated from Green's function surface integrations of equivalent currents which are obtained from the interior region solution. These modal scattered fields innately satisfy the radiation conditions but do not rely upon the use of separable coordinate surfaces for their completeness. A major advantage of the F^3 is its *modular* solution topology, where forward and feedback transfer matrices can be independently computed. These modules are connected for the complete scattering problem as a simple closed loop feedback network. This arrangement provides for either closed-form or iterative approaches as natural options in attaining the final scattering solution.

To explain the details of the F^3 approach via example, refer back to the problem of scattering by a dielectric cylinder, as depicted in Fig. 18. As shown there, the surface boundary contour of the object, S_B, is enclosed within a geometrical surface, S_0. Unlike the boundary integral equation method, where S_0 could have been equal to S_B, these two surfaces must now be distinct, with S_0 enclosing S_B. As before, the scattering problem is approached by first solving the Helmholtz equation inside of S_0 subject to the correct *total* field being applied as the boundary condition, as per (62). The $g_k(s)$ will now be denoted as $g_k(s_o)$ to emphasize that they are surface basis functions on S_0 and c_k are still to be determined.

Let us again assume that the finite element method is applied to the problem and that each of the $g_k(s_o)$ BC's on S_0 generates a numerical solution denoted by $\tilde{g}_k(\overline{r})$. We will further assume, although it is not necessary, that the g_k's are the unit-amplitude finite element basis functions evaluated on S_0, so that the c_k's become the corresponding nodal values of the total field on this surface. Next, we represent the numerical solution and its normal derivative ($D_n = \partial/\partial n$), evaluated on the object's surface S_B, as

$$\tilde{g}_k(s_b) = \sum_{m=1}^{M} a_{m,k}\, h_m(s_b) \tag{69a}$$

$$D_n\, \tilde{g}_k(s_b) = \sum_{m=1}^{M} a_{m,k}\, D_n\, h_m(s_b) \tag{69b}$$

where $h_m(s_b)$ is the m-th indexed unit-amplitude finite element basis function for nodes on S_B. The need for the normal derivative will be justified shortly. As an aside, the reader should note that we are assuming that there are N nodes on S_0 and M nodes on S_B.

Proceeding, we note that due to linearity a given c_k in (62) will produce a finite element solution and normal derivative along S_B given by $c_k\, \tilde{g}_k(s_b)$ and $c_k\, D_n\, \tilde{g}(s_b)$. Expanding on this idea, we can see that the total field on S_B is given in terms of its nodal values, d_m, as

$$\tilde{f}(s_b) = \sum_{m=1}^{M} d_m \, h_m(s_b) \tag{70a}$$

$$D_n \, \tilde{f}(s_b) = \sum_{m=1}^{M} d_m \, D_n \, h_m(s_b) \tag{70b}$$

with

$$d_m = \sum_{k=1}^{N} a_{m,k} \, c_k \tag{71}$$

As is apparent from (71), the finite element solution nodal values on S_B are given in terms of the nodal values of the BC on S_0 through the $M \times N$ array $\overline{\overline{A}}$, whose components are $a_{k,m}$. We will term this array as the "forward operator". By defining vectors of nodal values for the total fields on S_0 and S_B as

$$\overline{F}_0^t = [c_1, c_2, \cdots, c_N]^T \tag{72a}$$

$$\overline{F}_b^t = [d_1, d_2, \cdots, d_M]^T \tag{72b}$$

there results,

$$\overline{F}_b^t = \overline{\overline{A}} \cdot \overline{F}_0^t \tag{73}$$

Note that the c_k's in (72a) are the same total field coefficients as were defined in the basis function expansions of (62) and (65).

Our original assumption was that we had the total field available as a BC on S_0, via (62). Since only the incident field is known *a priori*, there remains the need to find the scattered field on S_0. On the other hand, if we know the total field and its normal derivative on S_B, we can evaluate the *scattered* field anywhere outside of S_B through a free-space Green's function convolution type integration, [14]. Applying this to find the scattered field on S_0 gives

$$f^s(s_o) = \int_{S_B} \left[G(s_o, s_b) \, \frac{\partial f(s_b)}{\partial n} - f(s_b) \, \frac{\partial G(s_o, s_b)}{\partial n} \right] ds_b \tag{74}$$

where $\hat{n}$ is the outward unit normal at s_b. The free-space Green's function is given by (64), with $\bar{r}$ positioned on S_0 and $\bar{r}'$ on S_B. It should be noted that by using the convolution operation in (74) we are implicitly enforcing the radiation conditions in (53) on the scattered field solution. This follows because the point source Green's function being used obeys these conditions. It was this property that guaranteed the proper far-field behavior of the scattered field produced by a surface integral equation solution.

As an aside, the total field and normal derivative within the Green's function integral in (74) can be replaced by the corresponding scattered field quantities on S_B. This is a well known result which follows from the equivalence theorem [2]. In essence, the Green's function integration of the incident field on S_B is zero for evaluation points outside of this contour, (e.g. the points s_o). Likewise, if s_o in (74) is changed to a point inside of S_B the result will be equal to the negative incident field, where the contribution of the scattered field component on S_B is zero.

Let us now return to the near field evaluation for points on S_0. Consider the scattered field on S_0 due to a single basis function, $h_m(s_b)$, existing on S_B,

$$\begin{aligned}\tilde{f}^s_m(s_o) &= \sum_{k=1}^{N} b_{k,m}\, g_k(s_o) \qquad (75)\\ &= \int_{S_B} \left[G(s_o, s_b)\, D_n\, h_m(s_b) - h_m(s_b)\, \frac{\partial G(s_o, s_b)}{\partial n} \right] ds_b\end{aligned}$$

Assuming for the moment that we have available the numerical solution on S_B, as indicated by (70), we can use superposition, along with (75), to obtain the corresponding scattered field on S_0,

$$\tilde{f}^s(s_o) = \sum_{k=1}^{N} e_k\, g_k(s_o) = \sum_{m=1}^{M} d_m\, \tilde{f}^s_m(s_o) \qquad (76)$$

where the nodal values of the scattered field are given by

$$e_k = \sum_{m=1}^{M} b_{k,m}\, d_m \qquad (77)$$

The role of the $N \times M$ array, $\overline{\overline{B}}$, whose elements are $b_{k,m}$, is to generate the *scattered* field on S_0 due to a specified *total* field on S_B. This array is termed the "feedback operator". Denoting the nodal total field vector on S_B using (72b) and the nodal scattered field vector on S_0 by

$$\overline{F}_0^s = [e_1, e_2, \cdots, e_N]^T \tag{78}$$

we obtain

$$\overline{F}_0^s = \overline{\overline{B}} \cdot \overline{F}_b^t \tag{79}$$

It may now be apparent what is going to happen next: we are going to "bootstrap" the solution by combining the scattered field in (76) with the known incident field, thus obtaining the total field on S_0. This, in turn, will allow us to find the total field on S_B, using (70), so we can then find the needed scattered field on S_0. Such a scheme describes a basic feedback system, as is illustrated in Fig. 21, where the *forward* and *feedback* operators represent the respective finite element solution and near-field Green's function integration. The $\overline{F}$-arrays represent the nodal values of the fields on the S_0 and S_B contours. On the outer geometric contour, S_0, the total field is the sum of the incident plus the scattered field: $\overline{F}_0^t = \overline{F}_0^i + \overline{F}_0^s$. Each of these nodal value arrays has N components. On the body contour, S_B, the total field nodal values are embodied in the M-element array, $\overline{F}_b^t$.

With the above ideas in mind, the matrix relationships in (73) and (79) provide the following geometric series representation for the total field on S_0

$$\overline{F}_0^t = [\overline{\overline{I}} + \overline{\overline{Q}} + \overline{\overline{Q}} \cdot \overline{\overline{Q}} + \overline{\overline{Q}} \cdot \overline{\overline{Q}} \cdot \overline{\overline{Q}} + \cdots] \cdot \overline{F}_0^i = \overline{\overline{T}} \cdot \overline{F}_0^i \tag{80}$$

where

$$\overline{\overline{Q}} = \overline{\overline{B}} \cdot \overline{\overline{A}} \tag{81}$$

is the "closed-loop gain" operator of the F^3 system.

In looking at Fig. 21, it should be apparent to the reader that the stability of the closed-loop system is an important issue. In particular, if the closed-loop gain operator has an array norm, $||Q|| \geq 1$, then the series in (80) will not converge, indicating an unstable system. For a

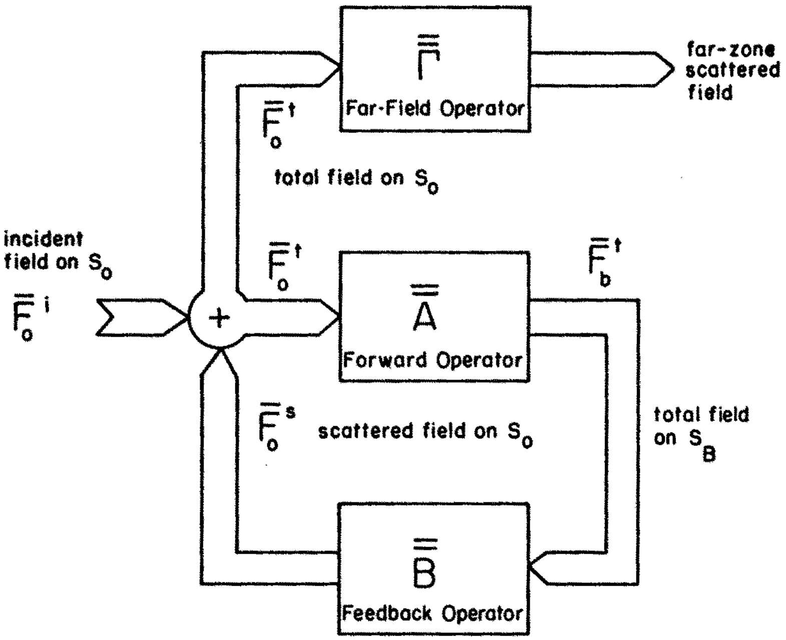

Figure 21 Field feedback system.

passive physical system, exhibiting either radiation damping and/or resistive dissipation, it can be argued that $||Q|| < 1$. On the other hand, real scattering structures exist, such as reentrant cavities, where, at selected resonant frequencies, large amounts of energy may be stored while relatively little power loss occurs due to radiation or dissipation per oscillation cycle. Such scatterers are said to be "high-Q" (borrowing the term *quality* factor from circuit theory). If the F^3 method is applied near the resonant frequencies of very high-Q structures, where $||Q|| \rightarrow 1$, then convergence of the iteration in (80) will be difficult to achieve numerically. This is not only a case of prohibitive computer time required to continue the iteration to large order in n; the major concern becomes the accumulation of roundoff error as this process proceeds. The difficulty in attaining accurate scattering solutions for ultra high-Q structures is not confined to the F^3 approach, but spans the whole gamut of numerical methods. An advantage of the F^3 is that the stability issue is obvious and well-defined.

Assuming that $||Q|| < 1$, where the concept of an *array norm* was defined in subsection 1.2f, the series in (80) can be summed to give the feedback system transfer matrix

$$\overline{\overline{T}} = [\overline{\overline{I}} - \overline{\overline{Q}}]^{-1} \tag{82}$$

This, of course, is analogous to the usual scalar transfer function formula for the total input to a feedback system. It may appear that we have overcome our convergence problem, due to roundoff, of numerically summing a truncated version of the series in (80). This is not the case; a marginally stable system having $||Q|| \rightarrow 1$ will produce a poorly conditioned matrix, $\overline{\overline{I}} - \overline{\overline{Q}}$, which needs to be inverted in (82). Referring back to the discussion of matrix conditioning which resulted in (47), we can use the same approach as employed in [18] to show that the condition number is bounded by

$$C(I - A) \leq \frac{1 + ||Q||}{1 - ||Q||} \tag{83}$$

It thus becomes obvious that as system stability is reduced, so is the conditioning of the direct matrix inversion in (82). It should be emphasized that in a recent study [34] no severe stability-conditioning problems were uncovered in computing the closed-loop gain, $||Q||$, for simple metallic and penetrable scatterers, including thin-wire structures. In the case of the highest Q scatterers, such as lossless cavities with small excitation ports at resonance, stability-conditioning problems should be expected in using the F^3, just as they should with other numerical methods.

An example 2-D finite element mesh for a semicircular dielectric cylindrical shell was shown in Fig. 12. Nodes are arranged to lie on both the body surface and the enclosing geometrical surface, as will usually be the case in implementing the F^3. Also, in this case, a single layer of triangular elements separates S_B from S_0 and $N = M$. The mesh could have been constructed with additional layers of elements between the two contours. The cost would have been additional unknowns to solve for in the finite element solution. If the extra element layer(s) had provided additional spatial separation between S_B and S_0 then the payoff would have been a less demanding accuracy in performing the near-field Green's function integration in (74) and (75).

After solving for $\overline{F}_0^t$ using either iteration, via (80), or closed form matrix inversion using (82), we may obtain the far-zone scattered fields by performing a simplified Green's function integration over S_0. Points on S_0 will be denoted by the vector $\overline{s}_o$. Starting with the integral form in (74), as applied to the S_0 contour rather than S_B, let us consider the case of the field point receding radially to infinity. We can then employ the asymptotic form for the Hankel function, as given in (60), to obtain a simplified integration for evaluating the complex scattering pattern, $\chi(\theta)$, using the definition in (53a). This results in

$$\chi(\theta) = \sqrt{\frac{\jmath}{8\pi}} \int_{S_0} \{\jmath D_n\, f^t(s_o) + k_0(\widehat{n} \cdot \widehat{r})\, f^t(s_o)\}\, e^{\jmath k_0 \overline{s}_o \cdot \widehat{r}}\, ds_o \qquad (84)$$

Using the basis function expansions for the total field in (62) and (65), we can evaluate the scattering pattern as

$$\begin{aligned} \chi(\theta) &= \sqrt{\frac{\jmath}{8\pi}} \sum_{k=1}^{N} c_k \int_{S_0} \{\jmath D_n\, g_k(s_o) + k_0(\widehat{n} \cdot \widehat{r})\, g_k(s_o)\}\, e^{\jmath k_0 \overline{s}_o \cdot \widehat{r}}\, ds_o \\ &= \sum_{k=1}^{N} c_k\, \gamma_k(\theta) = \overline{\Gamma}(\theta) \cdot \overline{F}_0^t \end{aligned} \qquad (85)$$

where

$$\overline{\Gamma}(\theta) = [\gamma_1(\theta), \gamma_2(\theta), \cdots, \gamma_N(\theta)] \qquad (86)$$

with

$$\gamma(\theta) = \sqrt{\frac{\jmath}{8\pi}} \int_{S_0} \{\jmath D_n\, g_k(s_o) + k_0(\widehat{n} \cdot \widehat{r})\, g_k(s_o)\}\, e^{\jmath k_0 \overline{s}_o \cdot \widehat{r}}\, ds_o \qquad (87)$$

The far-field Γ-operator is shown in Fig. 21 as a matrix. The $\jmath$-th row of this matrix is formed from the elements of $\overline{\Gamma}(\theta_j)$, where the set of θ_j is the discrete scattering angles desired for evaluating the pattern in (85).

The bistatic radar cross section (RCS) can be easily obtained from a knowledge of the scattering pattern. Assuming a unit magnitude incident field, $|f^i(\overline{r})| = 1$, the result is

$$\begin{aligned} \sigma(\theta) &= \lim_{r \to \infty} \frac{2\pi\, r\, |f^s(r,\theta)|^2}{|f^i|^2} \\ &= \lambda_0\, |\chi(\theta)|^2 \end{aligned} \qquad (88)$$

For the case of 2-D scattering, this RCS represents the effective cross sectional *width*. If the incident field power density is multiplied by $\sigma(\theta_0)$ the resultant power is equal to that which would be radiated by an isotropic scattering pattern, $\chi_0 = \chi(\theta_0)$, which is constant in θ.

Let's pause now to summarize the steps in the F^3 solution for scattering. These steps are:

(1) Enclose the object within a geometrical boundary, S_0,
(2) Formulate the numerical solution of the interior problem using an appropriate finite method,
(3) Find the interior solution for each of the applied basis functions on S_0, saving the nodal values of the function and its normal derivative on the boundary of the object, S_B (these solutions form the columns of the forward operator, A-matrix),
(4) Perform near-field Green's function integrations on S_B to find the nodal fields on S_0 due to each basis function on S_B (these integrations form the columns of the feedback operator, B-matrix),
(5) Form the closed-loop operator, $\overline{\overline{Q}} = \overline{\overline{B}} \cdot \overline{\overline{A}}$, and evaluate the T-matrix, which relates the total field on S_0 to the known incident field, using either the closed form matrix inversion approach of (82) or the series iteration of (80),
(6) Finally, perform far-zone integrations over S_0, per (84), thus computing the radiation field and the RCS in (88), if desired.

The initial demonstration of the F^3 was for scattering by a finite-length metallic thin-wire, with associated finite element mesh shown in Fig. 22. This is obviously a case where the unimoment method would be quite inefficient, requiring a meridian spherical mesh to enclose the wire unless a spheroidal mesh was adopted, with the associated difficulties of generating the spheroidal harmonic functions. Using the F^3, the mesh has only a single column of triangular elements, which produces an interior solution matrix (whose inverse is $\overline{\overline{A}}$) having a nonzero bandwidth of only 3 matrix elements in this case. The Riccati transform allows ultra-fast inversion of this matrix. Comparisons of the magnitude and phase of current on an $L = 1\lambda_0$ thin-wire, as computed from the F^3 and Hallen's integral equation, is shown in Fig. 23. One source of error in the F^3 computation was the use of linear basis functions in the finite element calculation. These were used to represent the ϕ-component of the magnetic field. In the immediate vicinity of the thin-wire, H_ϕ is characterized by a rapidly decaying

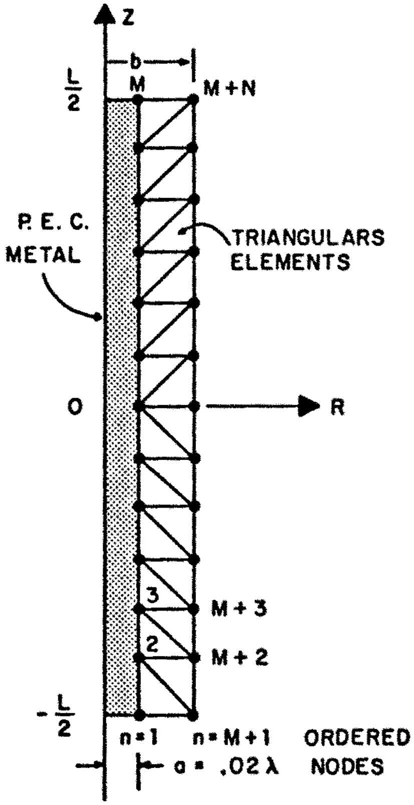

Figure 22 Finite element mesh for the thin-wire scatterer.

evanescence which is difficult to accurately interpolate using simple linear basis functions. More specialized basis functions would allow faster convergence of the solution for the thin-wire case. Much more rapid convergence was observed for thick cylinders.

A recent application of the F^3 has been made to scattering by 2-D dielectric cylinders, [35]. Uniformly excellent results in all validations were observed. An example is scattering by a dielectric semicircular shell of outer radius $0.3\lambda_0$, as originally computed by Richmond, using an E-field integral equation for the TM-case [36] while employing a cylindrical harmonic expansion approach for the TE-case [37]. The finite element mesh for the semicircular cylindrical shell is shown in Fig. 24, where the average element size is about $\lambda/20$ within the $\epsilon_r = 4$ dielectric material. A comparison of the computed normalized

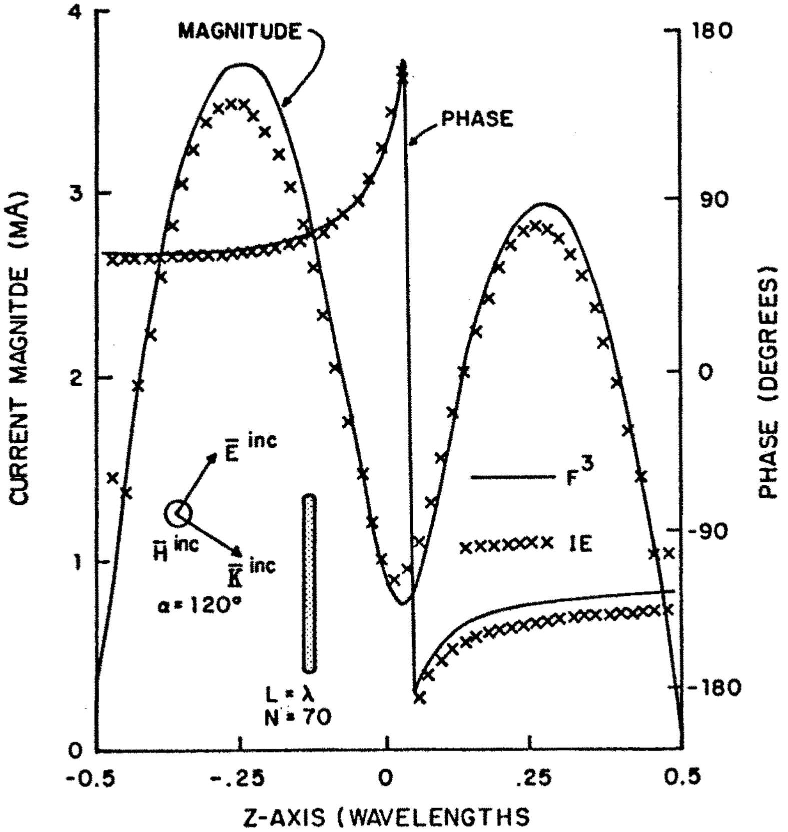

Figure 23 **Comparisons of computed magnitudes and phases of complex current on a thin-wire scatterer.**

bistatic RCS using the F^3 and Richmond's E-field integral equation is displayed for the TM-case in Fig. 25.

Efforts are underway to apply the F^3 to scattering by 3-D objects of general shape and composition using the generalized CAP formulation, as presented in Chapter 6, for computation of the forward A-operator. A major hurdle to overcome in applying the F^3 to scattering by *large* 3-D objects, having multiwavelength dimensions, is the tremendous CPU resources required to compute the near-field surface integrations in the feedback B-operator. The exact field at a point outside of the surface, S_B, depends upon the field (and normal derivative) over the entire S_B, as evidenced by the Green's function integral for-

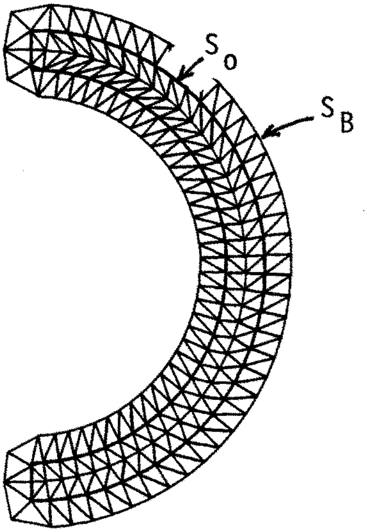

Figure 24 Conformal finite element mesh for a semicircular dielectric cylindrical shell.

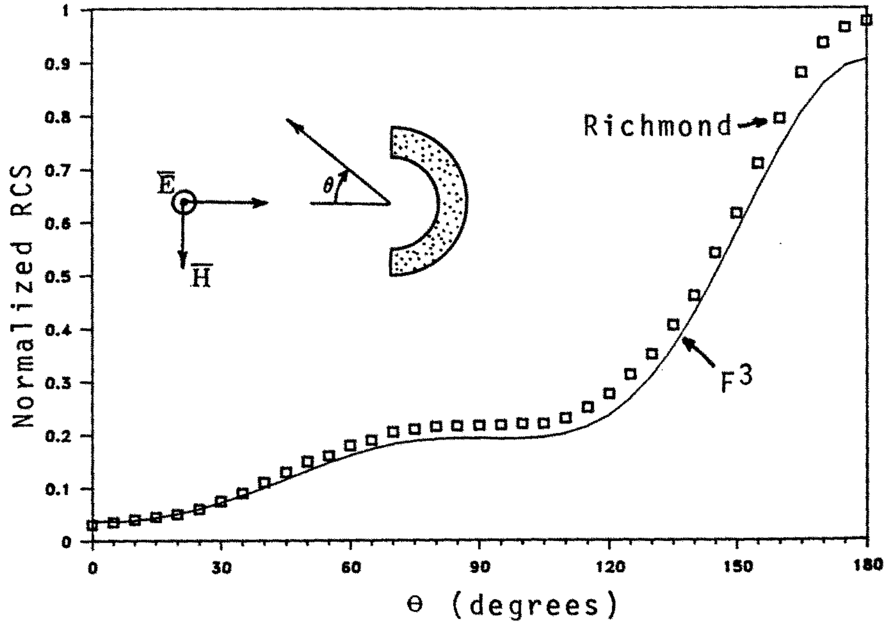

Figure 25 Comparisons of normalized RCS for a semicircular dielectric shell having $\epsilon_r=4$, inner radius$=0.25\lambda_0$ and outer radius$=0.30\lambda_0$.

mula in (74). In this integration, the effect on the field at $\bar{s}_o$ due to the S_B field at $\bar{s}_o$ dies away with increasing distance when $|\bar{s}_o - \bar{s}_b| > \lambda_0$. In 3-D, this decrease is even more rapid than in 2-D.

One approach to reducing the integration demands for large 3-D problems is to adaptively neglect the integration contributions outside of a local neighborhood of the field point. The cutoff would be dependent upon relative convergence of the integration. Along these same lines is the idea of local radiation boundary conditions for terminating the finite method mesh. Let us now briefly look at some of the basic concepts behind this topic. Our elementary level considerations will pave the way for the more advanced applications to be discussed in Chapters 4, 5, 7 and 8.

d. Radiation Boundary Conditions

Additional methods exist for employing finite method solutions in open region problems. One approach imposes a homogeneous Dirichlet BC on the tangential field components at a sufficiently large distance from the antenna or scatterer, [38]. For such an artificial assumption to result in any degree of accuracy, the mesh termination boundary must be far removed from the scatterer or antenna. This means that a large finite method mesh is needed, with an accompanying large matrix size in the frequency domain case. For time-domain calculations, the null BC condition also results in an extended mesh, requiring much computational effort in the causal updates of the many included nodes.

Another method employs "infinite elements", where unbounded angular sections are used to surround the scatterer [5]. Within these infinite elements, special basis functions are used which correspond to assumed far-zone behavior of the scattered field. This requires that the interface between the main mesh and the infinite elements lie in the *quasi* far-field region of the scatterer. The result is again a large mesh region, although not as big as that resulting from the zero BC assumption.

We have seen one form of the far-zone conditions in (53), as applied to 2-D scattering. For the 3-D case, the far-field behavior is embodied in the Sommerfeld radiation condition [39]. For a scalar scattered field, f^s, which satisfies the source-free Helmholtz equation in spherical coordinates, (r, θ, ϕ), this far-field condition becomes

$$\lim_{r \to \infty} r\, \mathcal{L}(r)\, f^s(r, \theta, \phi) = 0 \tag{89a}$$

where

$$\mathcal{L}(r) = \left[\frac{\partial}{\partial r} + \jmath k_0\right] \tag{89b}$$

By applying (89) to the $\hat{\theta}$ and $\hat{\phi}$ vector components of $\overline{E}^s$ and $\overline{H}^s$, followed by a substitution from Maxwell's curl equations, we will obtain the vector field form of Sommerfeld's conditions

$$\lim_{r\to\infty} r\left[\eta_0\,\hat{r}\times\overline{H}^s + \overline{E}^s\right] = 0 \tag{90a}$$

$$\lim_{r\to\infty} r\left[\hat{r}\times\overline{E}^s - \eta_0\,\overline{H}^s\right] = 0 \tag{90b}$$

Several efforts have used the Sommerfeld radiation condition, with applications in both the frequency and time-domains, as exemplified by [40–41].

To allow reduced mesh sizes, while retaining adequate accuracy of the enforced conditions on the outer mesh boundary, improved approaches have been developed in recent years involving radiation boundary operators. These operators come in two "flavors": extended annihilation [42–43], and one-way wave equations [44], both of which we will look at briefly. The equations that result from applying these operators at the mesh boundary are termed radiation boundary conditions (RBC's). A related RBC idea, as applied to time-harmonic surface integral equations, is the on-surface radiation condition (OSRC) [45]. The effect of the OSRC is to reduce the integral equation to either an ordinary differential equation on the surface, or a simple integration of *known* functions.

The annihilation operator approach can be considered as an extension to the far-field boundary conditions of Sommerfeld. As shown by Wilcox [46], outbound radiating fields which satisfy the Sommerfeld radiation condition in (89) can be represented by an inverse power series of the form

$$f^s(r,\theta,\phi) = e^{-\jmath k_0 r}\sum_{n=1}^{\infty}\frac{\chi_n(\theta,\phi)}{r^n} \tag{91}$$

As r becomes larger, the series can be truncated at a decreasing index for a fixed level of accuracy. Ultimately, as $r \to \infty$, only the $n{=}1$ term is needed, producing a far-field condition which is the 3-D version of the 2-D result in (53). The $\mathcal{L}(r)$ operator in (89), when applied to (91), is seen to reduce the leading order term from r^{-1} to r^{-2}. Thus, for increasing r, $\mathcal{L} \cdot f^s$ decays to zero much faster than does f^s.

To extend this idea, consider the possibility of discovering enhanced annihilation operators, $\mathcal{B}_m$, which can cancel the first m terms of the series in (91) at *finite* values of r. The approximate boundary conditions, $\mathcal{B}_m f^s \approx 0$, for increasing m, should then be applicable to surfaces of decreasing distance from the scattering object. The construction of such operators was the approach taken by Bayliss and Turkel in [42]. As an example, consider

$$\mathcal{B}_1(r) = \mathcal{L}(r) + \frac{1}{r} \tag{92}$$

which will annihilate the n=1 term in (91), while

$$\mathcal{B}_2(r) = \left\{\mathcal{L}(r) + \frac{3}{r}\right\} \mathcal{B}_1(r) \tag{93}$$

cancels both the n=1 and n=2 terms. This procedure can be continued through the recurrence formula

$$\mathcal{B}_{m+1}(r) = \left\{\mathcal{L}(r) + \frac{2m+1}{r}\right\} \mathcal{B}_m(r) \tag{94}$$

Since the RBC using (94) will have to be enforced using numerical differentiation in r, it becomes impractical to use large ordered $\mathcal{B}_m$'s; the m-th order operator contains $\partial/\partial r^m$ along with all lesser ordered derivatives. As it turns out, lower orders (e.g. m=2) are often sufficient to allow termination of the mesh close to the object's surface, at least for the class of 2-D convex homogeneous scatterers, as evidenced by results in Chapters 4, 5 and 8. For the 2-D case, slightly modified formulas are found for the $\mathcal{B}_m$'s since the Wilcox expansion in (91) must be changed to have the r^n term in the denominator replaced by $r^{n-1/2}$. Also, for the 2-D case, the 2nd-order radial derivatives in $\mathcal{B}_2$ can be replaced using a combination of a 2nd-order θ derivative with 1st-order radial derivative, using the Helmholtz equation. The annihilation operator RBC approach can be extended to the time-domain, as shown in [47] and Chapters 7 and 8.

Let us now investigate another method for generating the RBC. This is based upon the idea of a one-way wave equation, as developed by Engquist and Majda, [44]. As an example, let us reconsider the simple 2-D space-time wave equation that was covered in subsection 1.2c. In particular, assume we want to solve for the scattering due to an incident $+x$ directed pulse which impinges upon the Dirichlet BC, $u(a,t) = 0$, as illustrated in Fig. 26. A finite difference mesh as in Fig. 4 (or Fig. 5 for the leap-frog method) is overlaid on the *truncated* problem domain: $0 \leq x \leq a$ and $t \geq 0$. Since this is a scattering problem the actual spatial domain is unbounded, $-\infty < x \leq a$. The purpose of the RBC is to properly terminate the mesh at $x = 0$ so that the scattered wave is *absorbed* without reflection; hence, the term "absorbing boundary condition" is often used. Of course, this same terminology is equally applicable to the annihilation operator approach.

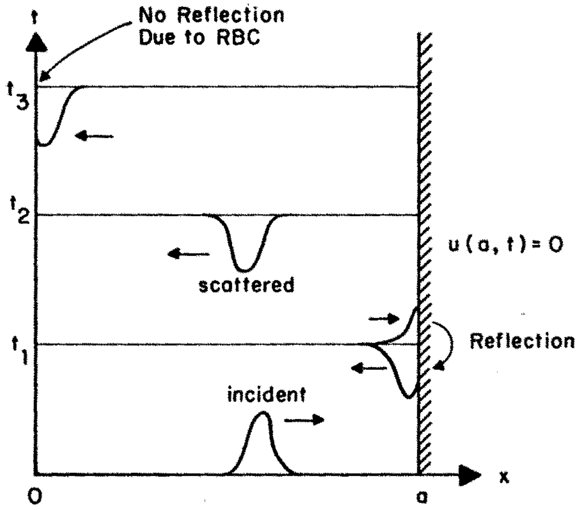

Figure 26 RBC applied to simple space-time scattering problem.

"Factoring" of the wave equation is a common technique for attaining one-way operators. Factoring of the (x,t) wave equation in (13) is particularly simple, resulting in the decomposition

$$\frac{\partial^2 u}{\partial x^2} - \frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} = \left[\frac{\partial}{\partial x} - \frac{1}{c}\frac{\partial}{\partial t}\right]\left[\frac{\partial}{\partial x} + \frac{1}{c}\frac{\partial}{\partial t}\right] u(x,t)$$

$$= L^+(x,t)\, L^-(x,t)\, u(x,t) = 0 \tag{95}$$

The L^+ operator cancels waves propagating in the $+x$ direction, such as the incident field in Fig. 26, while L^- cancels $-x$ traveling waves. Analytically applying L^- to the mesh termination boundary, at $x=0$, will provide an RBC to absorb the scattered wave without reflection

$$L^-(x,t)\, u(x,t)\Big|_{x=0} = 0 \tag{96}$$

For the grid in Fig. 4, with $u_{i,j} = u(x_i, t_j)$, the following finite difference approximation results by enforcing the RBC at $x = \Delta x/2$

$$u_{0,j+1} = u_{0,j-1} + 2\gamma\,(u_{1,j} - u_{0,j}) \tag{97}$$

where γ is given in (16). This relationship would be used in lieu of that in (15) for the i=0 boundary nodes.

Formulation of one-way space-time operators for the cases of two and three spatial dimensions follows from the factoring concept in (95), but with some additional approximations required. For example, in (x, y, t) coordinates, the one-way operators can be written as

$$L^{\pm}(x,y,t) = \frac{\partial}{\partial x} \pm \mathcal{R} \tag{98}$$

where $\mathcal{R}$ is a "pseudo-differential" operator which may be defined as

$$\mathcal{R} = \sqrt{\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial y^2}} \tag{99}$$

The implication of this symbolism is that a double application, $\mathcal{R} \cdot \mathcal{R}$, results in the operator within the radical in (99).

To be numerically usable, the radical in (99) may be interpreted as a convergent expansion of the derivatives contained within. An approximate RBC can then be generated by properly truncating the expansion and applying the resultant operator at the enclosing boundary. Such an approach was developed by Mur [48], who employed a two-term Taylor series. Accuracy of the resultant RBC becomes sensitive to the arrival angle of the scattered wave on the local boundary. Optimizations are possible using rational polynomial interpolates (Padé, Chebyshev, etc.) for the one-way operator. The objective is to improve the approximation of the operator and reduce the sensitivity to arrival angle. Further details on the development and application of one-way operators are given in Chapter 8. In addition, an excellent review article on RBC's is also available [49].

1.4 Conclusion

The objective in this chapter has been twofold: (1) to provide a tutorial overview of the concepts and terminology that are employed at more advanced levels in later chapters; and (2) to consider some scattering applications which are not subsequently covered in this book, such as those of the field feedback formulation (F^3). Our itinerary in developing the principles was to first consider finite difference and finite element methods (denoted as finite methods), as solved within enclosed spatial regions. In the second part of the chapter we turned our attention to various approaches for properly coupling closed region finite method solutions to the unbounded exterior domain of the scattering problem.

Having a confining boundary is, in fact, the natural setting for time-harmonic and time-domain PDE's, which form the basis for the finite methods. Examples were developed in the first half of this chapter for discretizing differential equation formulated problems within closed spatial regions. Finite difference grids were employed in the frequency and time-domains, and a basic finite element mesh was used for a time-harmonic example. Pertinent equations were deduced for solving each of the example problems and some specialized numerical procedures were introduced for effecting these solutions.

Along the way, an attempt was made to unify the finite method numerical procedures through introduction of the weighted residual method (also known as the moment method in the realm of integral operator equations). Within such a context, the finite difference technique can be seen as a special case of the finite element method – one which employs Dirac weighting functions. It was demonstrated that, when used with compact-support basis and weighting functions, the finite methods yield only *local* interactions between discretized unknowns. This is a key advantage of the finite methods over integral equation based techniques, where *global* type relationships usually exist among the unknowns. Local interactions, as produced by the finite methods, result in sparse matrix equations in the time-harmonic case and yield highly efficient causal updating algorithms in the time-domain. The topic of error analysis was also developed to give some basic insight into the numerical attributes of conditioning and stability, as associated respectively with frequency and time-domain solutions.

Since finite methods are innately posed as boundary- and initial-value problems, an *ad hoc* enforcement of the proper spatial "boundary conditions" at infinity (Sommerfeld radiation conditions) are required. Such an enforcement allows a finite-sized mesh to be used for the scattering problem in an unbounded domain. Different approaches are considered in section 1.3 for permitting mesh termination. Once again a sequence of examples was used as part of the tutorial to illustrate the basic ideas and methodology. Both time-harmonic and time-domain cases were considered, but with emphasis on the frequency domain.

One of the earliest of these techniques, the unimoment method, was presented here through a simple 2-D example. Some representative computational results were also given for 3-D axisymmetric scatterers and the unimoment method will be considered again in Chapter 2. For cases of highly elongated or flattened scattering structures, the unimoment method becomes somewhat inefficient. This is due to its fixed outer boundary shape which, for convenience, is usually circular in 2-D and spherical in 3-D. A large percentage of the enclosed mesh region becomes filled with a vacuum. In such cases, a more efficient approach is to use a mesh which is conformal to the scatterer.

The use of a boundary integral equation is one way to guarantee proper behavior of the scattered field. A simple 2-D example was presented to clarify the procedure. An early application to 3-D scattering was also considered using what was termed the finite-element boundary-integral method (FEBI). More will be said on the topic of boundary integral equations in Chapter 3, where a detailed application of this method is made to 2-D scattering from highly inhomogeneous structures.

The F^3 technique is similar in spirit to the boundary integral approach, but with increased flexibility. In fact, the boundary integral equation can be shown to be a special case of the F^3, when the two bounding surfaces become congruent and proper limiting conditions are applied. As with the boundary integral approach, the F^3 permits the use of a conformal mesh for field solutions within complicated scatterers composed of inhomogeneous and even anisotropic materials. As was shown, the F^3 casts the scattering problem into a vector-matrix equivalent of a simple feedback system, wherein the forward (interior problem) and feedback (radiated field) matrices may be computed *independently* and then combined to obtain a self-consistent scattering solution. Because of the F^3 system topology, either iterative or direct processing may be readily employed. The associated stability and condition number issues, and their relationship, were discussed. Results for scattering by 2-D penetrable objects and 3-D thin-wires were also presented. The future of the F^3 resides in extending the realm of practical electromagnetic calculations to electrically large, and quite complex, scattering configurations where volume integral equation techniques are numerically "bottlenecked" by their associated full matrices.

One of the most exciting (and controversial) recent developments in computational electromagnetics is in the area of radiation boundary conditions (RBC's). Such conditions are based upon numerical approximations to the behavior of outbound radiation fields. An RBC is applied at an enclosing boundary which does not need to be in the far-zone of the scatterer. In some cases, the boundary has been *very* close to the surface of the scattering object. RBC's have been developed through multiple approaches, two of which we briefly considered.

The computational power of RBC's stems from their potential for requiring only local interactions in the mesh termination equations. This is in contrast to the analytically exact boundary integral and F^3 approaches, where full global coupling at the mesh boundary is produced because of innate Green's function integrations. Within these boundary integrations, the scattered near-field at a given point is determined by the actual (or equivalent) source contributions over the entire integration surface. This dependence is stated as the classical Huygen's Principle [1]. However, the weighted effect of source contributions falls off with increasing distance from the field point. An important computational question concerns the possibility of ignoring the source contributions beyond some distance in exchange for a small increase in integration error. Although the RBC's do not explicitly perform the source integration, the implication of their success is that the scattered field at a near-field point is not strongly determined by other than the local behavior of equivalent sources in its immediate vicinity. This point is a major source of disagreement at the present time. Of particular concern are the cases of reentrant structures and highly resonant objects whose reactive fields are much stronger than their radiative fields over at least a portion of the equivalent source surface. In support of the RBC's is their record of success, as displayed by the results in Chapters 4, 5, 7 and 8. Very efficient computational algorithms result by combining the local mesh interactions in the finite methods with the local boundary interactions of an appropriate RBC. This global sparsity of interactions, if accurate, may ultimately drive virtually all large-scale electromagnetic computations (metallic as well as penetrable) into the use of finite methods.

References

[1] Stratton, J. A., *Electromagnetic Theory*, New York: McGraw-Hill, 1941.

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

[3] Van Bladel, J., *Electromagnetic Fields*, New York: McGraw-Hill, 1964; (Reprinted) New York: Hemisphere, 1985.

[4] Kong, J. A., *Electromagnetic Wave Theory*, New York: Wiley, 1986.

[5] Bettess, P., "Infinite elements," *Int. J. Num. Meth. Eng.*, **11**, 53–64, 1977.

[6] Mei, K. K., M. A. Morgan, and S. K. Chang, "Finite methods in electromagnetic scattering," Chap. 10 in *Electromagnetic Scattering*, P. L. E. Ushlenghi, Ed., New York: Academic Press, 1978.

[7] Greenberg, M. D., *Foundations of Applied Mathematics*, Engelwood Cliffs: Prentice-Hall, 1978, 540–547.

[8] Stakgold, I., *Green's Functions and Boundary Value Problems*, New York: Wiley, 1979, Chap. 8.

[9] Courant, R., K. Friedrichs, and H. Lewy, "Über die partiellen differenzengleichungen der mathematischen physik," *Mathematische Annalen*, **100**, 32–74, 1928.

[10] Fox, P., "The solution of hyperbolic partial differential equations by difference methods," Chap. 16 in *Mathematical Methods for Digital Computers*, A. Ralston and H. S. Wolf, Eds. New York: Wiley, 1964.

[11] Lapidus, L., and G. F. Pinder, *Numerical Solution of Partial Differential Equations in Science and Engineering*, New York: Wiley, 1982.

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

[13] Strang, G., and G. J. Fix, *An Analysis of the Finite Element Method*, Englewood Cliffs: Prentice-Hall, 1973.

[14] Morse, P. M., and H. Feshbach, *Methods of Theoretical Physics*, New York: McGraw-Hill, 1953.

[15] Tewarson, R. P., *Sparse Matrices*, New York: Academic, 1973.

[16] Morgan, M. A., and K. K. Mei, "Finite element computation of scattering by inhomogeneous penetrable bodies of revolution", *IEEE Trans. Antennas Propagat.*, **AP-27**, 202–214, 1979.

[17] Mittra, R., and C. A. Klein, "Stability and convergence of moment method solutions," Chap. 5 in *Numerical and Asymptotic Techniques in Electromagnetics*, R. Mittra, Ed., New York: Springer-Verlag, 1975.

[18] Isaacson E., and H. B. Keller, *Analysis of Numerical Methods*, New York: Wiley, 1966.

[19] Rynne, B. P., "Instabilities in time marching methods for scattering problems," *Electromagnetics*, **6**, 129–144, 1986.

[20] Mei, K. K., "Unimoment method of solving antenna and scattering problems," *IEEE Trans. Antennas Propagat.*, **AP-22**, 760–766, 1974.

[21] Chang, S. K., and K. K. Mei, "Application of the unimoment method to electromagnetic scattering of dielectric cylinders," *IEEE Trans. Antennas Propagat.*, **AP-24**, 35–42, 1976.

[22] Stovall, R. E., and K. K. Mei, "Application of a unimoment technique to a biconical antenna with inhomogeneous dielectric loading," *IEEE Trans. Antennas Propagat.*, **AP-23**, 335–341, 1975.

[23] Morgan, M. A., "Finite element computation of microwave scattering by raindrops," *Radio Science*, **15**, 1109–1119, 1980.

[24] Morgan, M. A., "Finite element calculation of microwave absorption by the cranial structure," *IEEE Trans. Biomed. Eng.*, **BME-28**, 687–695, 1981.

[25] Hunka, J. F., and K. K. Mei, "Electromagnetic scattering by two bodies of revolution, *Electromagnetics*, **1**, No. 3, 329–347, 1981.

[26] Chang, S. K., and K. K. Mei, "Multipole expansion technique for electromagnetic scattering by buried objects," *Electromagnetics*, **1**, No. 1, 73–89, 1981.

[27] Morgan, M. A., K. K. Mei, and S. K. Chang, "Coupled azimuthal potentials for electromagnetic field problems in inhomogeneous axially-symmetric media," *IEEE Trans. Antennas Propagat.*, **AP-25**, 413–417, 1977.

[28] Morgan, M. A., C. H. Chen, S. C. Hill, and P. W. Barber, "Finite element-boundary integral formulation for electromagnetic scattering," *J. Wave Motion,* **6**, 91–103, 1984.

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

[30] Waterman, P. C., "Scattering by dielectric obstacles," *Alta Frequenza*, **38**, (Speciale), 348–352, 1969.

[31] Miller, R. F., "Rayleigh hypothesis in scattering problems," *Electronics Letters*, **5**, No. 17, 416–418, 1969.

[32] Peterson, A. F., "A comparison of integral, differential and hybrid methods for TE-wave scattering from inhomogeneous dielectric cylinders," *J. Electromagnetic Waves Applicat.*, **3**, No. 2, 87–106, 1989.

[33] Morgan, M. A., and B. E. Welch, "The field feedback formulation for electromagnetic scattering problems," *IEEE Trans. Antennas Propagat.*, **AP-34**, 1377–1382, Dec. 1986.

[34] Morgan, M. A., "Stability considerations in the field feedback formulation," submitted to *IEEE Trans. Antennas Propagat.*, March 1989.

[35] Welch, T. B., *Electromagnetic Scattering from Two-Dimensional Objects Using the Field Feedback Formulation*, Engineer's Thesis, E.C.E. Dept., Naval Postgraduate School, Monterey, CA, March 1989.

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

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

[38] Mabaya, N., P. E. Lagasse and P. Vandenbulcke, "Finite element analysis of optical waveguides," *IEEE Trans. Microwave Theory Tech.*, **MTT-29**, 600–605, 1981.

[39] Sommerfeld, A., *Partial Differential Equations in Physics*, New York: Academic, 1949.

[40] Merewether, D. E., "Transient currents induced on a metallic body of revolution by an electromagnetic pulse," *IEEE Trans. Electromagn. Compat.*, **EMC-13**, 41–44, 1971.

[41] Kriegsmann, G. A., and C. S. Morawetz, "Numerical solutions of exterior problems with the reduced wave equation," *J. Comp. Phys.*, **28**, 181–197, 1978.

[42] Bayliss, A., and E. Turkel, "Radiation boundary conditions for wave-like equations," *Commun. Pure Appl. Math.*, **23**, 707–725, 1980.

[43] Kriegsmann, G. A., and C. S. Morawetz, "Solving the Helmholtz equation for exterior problems with variable index of refraction: I," *SIAM J. Sci. Stat. Comput.*, **1**, 371–385, 1980.

[44] Engquist, B., and A. Majda, "Absorbing boundary conditions for the numerical simulation of waves," *Math. Comp.*, **31**, 629–651, 1977.

[45] Kriegsmann, G. A., A. Taflove, and K. R. Umashankar, "A new formulation of electromagnetic wave scattering using an on-surface radiation boundary condition," *IEEE Trans. Antennas Propagat.*, **35**, 153–161, 1987.

[46] Wilcox, C. H., "An expansion theorem for electromagnetic fields," *Commun. Pure Appl. Math.*, **9**, 115–132, 1956.

[47] Taflove, A., and K. R. Umashankar, "The finite-difference time-domain (FD-TD) method for electromagnetic scattering and interaction problems," *J. Electromag. Waves Applicat.*, **1**, 243–267, 1987.

[48] Mur, G., "Absorbing boundary conditions for the finite-difference approximation of time-domain electromagnetic field equations," *IEEE Trans. Electromag. Compat.*, **EMC-23**, 377–382, 1981.

[49] Moore, T. G., J. G. Blaschak, A. Taflove, and G. A. Kriegsmann, "Theory and application of radiation boundary operators," *IEEE Trans. Antennas Propagat.*, **36**, 1797–1812, 1988.