Electromagnetic Modeling of Composite Metallic and Dielectric Structures

For a complete listing of the *Artech House Antennas and Propagation Library,* turn to the back of this book.

Electromagnetic Modeling of Composite Metallic and Dielectric Structures

Branko M. Kolundzija
Antonije R. Djordjević

Artech House
Boston • London
www.artechhouse.com

Library of Congress Cataloging-in-Publication Data
Kolundzija, Branko M.
Electromagnetic modeling of composite metallic and dielectric structures / Branko Kolundzija.
p. cm. — (Artech House antennas and propagation library)
Includes bibliographical references and index.
ISBN 0-89006-360-5 (alk. paper)
1. Dielectric devices—Mathematical models. 2. Metallic composites—Electric properties—Mathematical models. 3. Electromagnetic waves—Mathematical models. 4. Moments method (Statistics). I. Djordjević, Antonije. II. Title. III. Series.
TK7872.D53 K65 2002
621.382'24—dc21 2002027962

British Library Cataloguing in Publication Data
Kolundzija, B. M.
Electromagnetic modeling of composite metallic and dielectric structures. — (Artech House antennas and propagation library)
1. Metallic composites—Computer simulation 2. Dielectrics—Computer simulation
3. Electromagnetic fields—Mathematical models I. Title II. Djordjević, A.
620.1'6

ISBN 0-89006-360-5

Cover design by Igor Valdman

685 Canton Street
Norwood, MA 02062

International Standard Book Number: 0-89006-360-5
Library of Congress Catalog Card Number: 2002027962

10 9 8 7 6 5 4 3 2 1

Contents

1

Introduction

In the theory and practice of antennas, microwave and radar techniques, electromagnetic compatibility, and so forth, various problems of the analysis of electromagnetic fields are often encountered. From them, only a limited number of problems can be solved exactly using classical mathematical analysis [1, 2]. In most cases where classical mathematical analysis cannot be applied, scientists and engineers have to perform tedious experiments and measurements to find acceptable solutions.

With the advent of digital computers in the mid-1960s, new possibilities for solving electromagnetic-field problems became available. To exploit the power of the computers fully, a variety of new techniques were introduced. During the last three decades, a tremendous effort has been made to develop methods for the numerical analysis of various classes of antennas, scatterers, passive microwave circuits, and other similar devices. The aim of these methods is to determine the distribution of the electromagnetic field (including currents and charges) within the analyzed structures, both in their immediate vicinity and at a distance. Such an analysis is often referred to as *electromagnetic modeling*.

Within electromagnetic modeling, various structures can be treated following similar guidelines. Antennas, scatterers, passive microwave circuits, and so forth, represent structures consisting of metallic and dielectric parts collectively referred to as *composite metallic and dielectric structures*.[1] If all the parts are made of linear materials, analysis can be facilitated, running in the domain of the *complex frequency* rather than the *time domain*; however, if one or more of

1. In this book, the term *composite metallic and dielectric structures* will also be used in a broader sense to include magnetic materials.

these parts is made of a nonlinear material, the time-domain approach must be applied. Although the time-domain approach is steadily increasing in importance, in most cases electromagnetic problems are considered in the frequency domain.

In this book, we present an efficient and unified approach to the electromagnetic modeling of practically arbitrary composite metallic and dielectric structures in the frequency domain. This approach is based on the *method of moments* (MoM), a general technique for solving numerical problems of electromagnetic fields. Specific topics related to the application of the MoM are considered. Techniques developed by various researchers in the past are compared or combined with our original techniques, some of which are published in this book for the first time. Based on these comparisons, optimal combinations of techniques are suggested, always keeping flexibility, ease of implementation, accuracy, and efficiency in mind. The power of the proposed approach is demonstrated on real engineering problems, such as a TV-UHF panel antenna with a radome, a paraboloidal reflector antenna with a waveguide feed and feed struts, shielded solid dielectric horn antenna, and so forth.

1.1 MoM as a General Approach to Solving Electromagnetic-Field Problems

The first steps in developing electromagnetic modeling were taken in the mid-1960s. Mei [3–5], Andreasen [6–8], Richmond [9–13], Oshiro [14, 15], Harrington [16, 17], and some other authors presented solutions to various electromagnetic problems. Initial papers dealt with the scattering from cylinders [3, 6, 8, 9, 12] assumed to be infinitely long and excited by a plane wave with the direction of wave incidence orthogonal to the cylinder axis. The field scattered from the cylinder depends only on the two Cartesian coordinates in the plane perpendicular to the cylinder axis. These problems are referred to as two-dimensional (2-D) electromagnetic problems. Solving 2-D problems is often much easier than solving three-dimensional (3-D) problems, in which the fields depend on all three Cartesian coordinates.

Researchers often use 2-D problems to test new concepts. Later, numerical techniques are applied to various 3-D problems. The first class of 3-D problems extensively analyzed included wire antennas and scatterers [4, 5, 10, 11, 16, 17]. The surface current along a wire segment is usually assumed to have only the axial component, which depends only on the axial coordinate. This approximation (referred to as the *thin-wire approximation*) simplified the analysis of the wire structures compared to the analysis of arbitrary metallic structures. In the latter case, two surface-current components must be considered, which

depend on two local coordinates tangential to the surface. A few attempts were made to solve other metallic structures [7, 13–15].

A variety of methods were established for the electromagnetic modeling of composite metallic and dielectric structures. However, all these techniques may be considered as special cases of the MoM, originally described by Kantorovich and Krylov [18] and Kantorovich and Akilov [19]. Harrington systematically investigated the application of the MoM to solve electromagnetic field problems [17]. An extensive report about this investigation was organized in the book *Field Computation by Moment Methods* [20], which represents one of the most frequently cited references on the numerical modeling of electromagnetic fields.

The MoM is a general method for solving linear operator equations.[2] It can be outlined as follows: The unknown quantity (for example, a field or a current density dependent on spatial coordinates) is approximated by a finite series of known functions (referred to as the basis or expansion functions) multiplied by unknown coefficients. This approximation is substituted into the linear operator equation. The left and the right sides of the equation are multiplied by a suitable function (referred to as the test or weighting function) and integrated over the domain in which the test function is defined. Thus, the linear operator equation is transformed into a simple linear algebraic equation. By repeating this procedure for a set of independent test functions whose number should equal the number of the basis functions, a system of linear equations is obtained. A standard numerical procedure (e.g., the Gaussian elimination) solves the linear system to yield the unknown coefficients to give an approximate solution to the operator equation.

In principle, the MoM can be applied to the numerical modeling of arbitrary linear structures. However, this method has limitations primarily governed by the speed and storage capabilities of available digital computers. Using more powerful computers increases the capability of the MoM. Another way to refine the MoM is by choosing proper starting equations, developing flexible basis and testing functions, constructing sophisticated algorithms for the numerical evaluation of integrals encountered in the solution, and designing other specific aids. The objective is to perform the analysis efficiently with a decreased number of unknowns. These refinement techniques are the main topic of this book.

2. By definition, an operator equation is a mapping from one function space to another function space. In the numerical electromagnetic modeling, operator equations are integral or differential equations for the (unknown) distribution of electric and magnetic fields, currents, and charges. These equations can be linear if all the media encountered in the system are linear.

1.2 MoM/SIE, MoM/VIE, and FEM

In its initial applications to electromagnetic field problems, the MoM was primarily used to solve integral equations for the current distribution. In general, these integral equations can be grouped into two classes. The first consists of cases where the unknown currents (and charges) are distributed over boundary surfaces between linear, homogeneous, and isotropic media. These currents are determined by solving integral equations based on the boundary conditions.[3] These integral equations are called *boundary integral equations* (BIEs). If the MoM is applied to solve a BIE, it is abbreviated as MoM/BIE. Wire and metallic structures, assumed to be made of a *perfect electric conductor* (PEC) and situated in a vacuum, belong to this class because the currents are distributed only over their surfaces. Further examples are structures consisting of linear, homogeneous, and isotropic material objects where the fields can be expressed in terms of equivalent electric and magnetic currents on the object surfaces by using the surface equivalence principle [2].

The second class consists of cases where the unknown currents are distributed throughout a linear, homogeneous, and isotropic medium. These currents are determined by solving integral equations based on the constitutive relations.[4] All such integral equations are referred to as *constitutive integral equations* (CIEs). If the MoM is applied to solve a CIE, it is abbreviated as MoM/CIE. Structures made of inhomogeneous, imperfect (real) conductors where the currents are distributed throughout the volumes are an example of this class. By using the volume equivalence theorem [2], the fields inside these objects can be expressed in terms of equivalent volume electric and magnetic currents placed in a vacuum.

In general, BIEs belong to the class of *surface integral equations* (SIEs) because the unknown function is distributed over one or more surfaces. In the particular case of the thin-wire approximation, SIEs degenerate to a *line integral equation* (LIE) where the unknown function is distributed along one or more lines.

CIEs belong to the class of *volume integral equations* (VIEs) where the unknown function is distributed over the volume of the material objects. In some particular cases (e.g., when the skin effect in conductors is pronounced), the volume CIE degenerates into the surface CIE. From the viewpoint of implementation, such surface CIEs resemble BIEs more than CIEs. Hence, the term MoM/SIE will be used for both the MoM/BIE and the MoM applied

3. Boundary conditions are relations for the tangential and normal components of the electromagnetic field vectors at surfaces of discontinuities.
4. Constitutive relations are equations for the field vectors governed by the electric and magnetic properties of the material.

to the surface CIE. The term MoM/VIE will be used for the MoM applied to the volume CIE.

In general, the MoM/VIE, which involves 3-D currents, requires substantially more unknowns than the MoM/SIE, which involves 2-D currents. Exceptions are cases of highly inhomogeneous material objects when the numbers of unknowns are comparable for the two approaches.

Many authors, including Harrington in two chapters of his book [20], have used the MoM to solve differential equations. The unknown function is a field (one or more components of a vector, depending on the problem) or a potential. The application of the MoM to differential equations differs significantly from the application of the MoM to integral equations. Moreover, specialists in integral equations are often not very familiar with electromagnetic modeling using differential equations, and vice versa.

Due to the above and some other historical reasons, the Galerkin method (a special case of the MoM[5]), when applied to differential equations in the frequency domain, is considered to belong to the general method for solving differential equations known as the *finite element method* (FEM) [21]. It is common practice to identify the MoM only with methods for the solution of integral equations[6]. In this book we shall follow this practice and use the terms *MoM* and *FEM* primarily to denote the solution of integral equations and the solution of the differential equations using the Galerkin procedure, respectively.

When solving a numerical electromagnetic modeling problem, it is desirable to select an approach that rapidly yields accurate solutions. Generally, there is a preferable approach (or a set of them) for each class of problems. In particular, the FEM is inherently suited to problems where the fields are bounded to a finite (closed) region (e.g., a region enclosed by a perfectly conducting surface, like metallic cylindrical waveguides). For such problems, a similar number of unknowns is needed by the FEM and the MoM/VIE, but it is usually much greater than that needed by the MoM/SIE. However, the FEM results in a sparse matrix of the system of linear equations, while the matrix obtained by the MoM is always almost full. There exist techniques for manipulations with sparse matrices (some of them are particularly adapted for the FEM) that are much faster and more memory-efficient than manipulations with full matrices of the same order. For the closed-region problems, the FEM is more efficient than the MoM/VIE, but is comparable to the MoM/SIE.

The fields exist in an unbounded (open) region in the majority of practical problems (in particular, antennas and scatterers). To apply the FEM to open

5. In the Galerkin method, the testing functions coincide with the expansion functions.
6. To add to the confusion, FEM practitioners consider the solution of integral equations using the Galerkin procedure as a special case of the FEM.

structures, specific techniques should be implemented that substantially increase the complexity of the method. For open-region problems that involve PEC objects, the MoM is usually the preferred choice. When dielectric or imperfect-conductor objects are present, the decision is not straightforward, particularly if the material is highly inhomogeneous.

In conclusion, the optimal method (MoM/SIE, MoM/VIE, and FEM) depends on the problem to be solved. This book presents an efficient and unified technique for the analysis of general 3-D structures (including both open-region and closed-region problems). Our experience shows the MoM/SIE to be superior to other techniques in the majority of practical cases for analysis of such structures, which has caused the MoM/SIE to be adopted as the dominant method in this book.

1.3 Basic Classes of Composite Metallic and Dielectric Structures (Material Structures)

The complexity of numerical modeling is influenced by the operating principle of the structure and frequency, but most of all by the kind, number, and geometry of the material bodies comprising the structure. For example, if a structure is positively a nonradiating structure (e.g., a waveguide circuit), the method used for the analysis need not take the radiation into account. An electrically small structure may be analyzed using quasistatic equations (i.e., neglecting the propagation effects), whereas an electrically very large structure may be analyzed using *physical optics* (PO). However, the objective of the technique developed in this book is to cover a variety of structures so that the starting equations would be valid for both radiating and nonradiating structures over a wide range of frequencies.

According to complexity of electromagnetic modeling based on solving the SIE, the structures can suitably be grouped into the following three classes:

- Wire structures;
- Metallic structures (i.e., composite wire and plate structures);
- General structures (i.e., composite metallic and dielectric structures).

Note that each class includes the previous one(s). Wherever possible, we will initially apply new methods and concepts to wire structures, then to more general structures. For example, let us consider a *log-periodic dipole array* (LPDA), which is a 3-D electromagnetic problem. The simplest case of such an antenna [corresponding to high-frequency (HF) antennas] is shown in Figure 1.1(a).

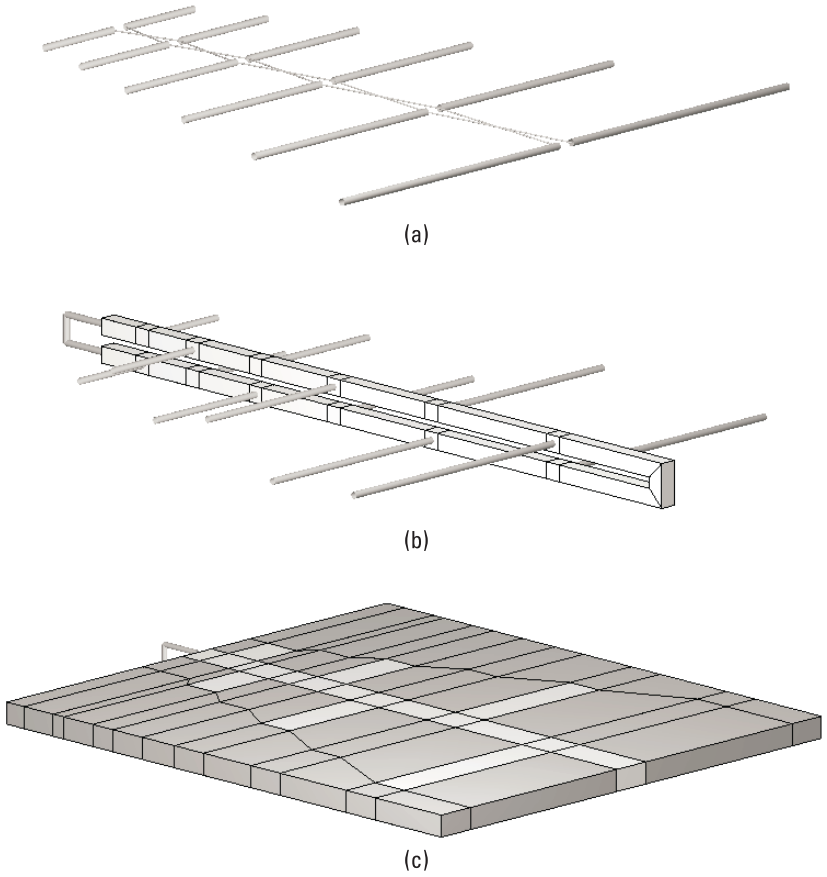

Figure 1.1 Typical models of composite metallic and dielectric structures: (a) log-periodic array of dipole antennas, (b) TV log-periodic antenna, and (c) printed log-periodic antenna.

The antenna consists of thin-wire dipoles interconnected by a two-wire transmission line. Perhaps the simplest way to analyze the antenna is to divide the analysis into two steps. In the first step, we may use any technique suitable for parallel-wire antennas to compute the self and mutual impedances of the dipoles in the array. The second step treats the whole system as a microwave network where the array is a multiport network with the ports interconnected by two-wire lines. An alternative analysis procedure is to combine the feeding line with the dipoles into a unique wire structure. In this approach, we have to analyze nonparallel wires and include wire-to-wire junctions. These junctions can be simple (where two wire segments are interconnected) or multiple (where more than two segments are interconnected).

Figure 1.1(b) shows a more complicated case: a commercial TV receiving antenna. The LPDA model consists of thin-wire dipoles, but the transmission line is rather thick, consisting of two conductors of square cross sections. The feeding region is modeled by a wire "nose." The model is now more complex as the transmission line is modeled by metallic plates to properly take into account the proximity effect and the effect of the line-to-dipole junctions. In the model, besides the wire-to-wire junctions, we can note two new entities: wire-to-plate junctions and plate-to-plate junctions.

The most complicated case considered, shown in Figure 1.1(c), has the LPDA printed on a dielectric substrate (only the top face of the substrate is shown). This antenna is a composite metallic and dielectric structure consisting of metallic wires (in the feeding region at the antenna nose), metallic plates (printed patterns), and a dielectric object (the substrate). We have wire-to-wire, wire-to-plate, and plate-to-plate junctions with metallic plates pressed against the dielectric object and with some edges common to the metallic and dielectric surfaces. We shall consider such edges as metallic-to-dielectric junctions. The analysis of this structure is a complex task, which is precisely within the objective of this book.

Numerous papers have been published on various methods for the electromagnetic modeling of material structures based on the application of the MoM to SIEs. It is virtually impossible to present here a complete review of this material and give proper credit to all of the authors. Hence, the following three sections quote only the most important references relevant to the techniques proposed in this book. The references are grouped according to the complexity of the structures analyzed following the above classification.

1.4 Methods for the Analysis of Thin-Wire Structures

Most often, we base the analysis of thin-wire structures on the solution of either the *electric field integral equation* (EFIE) or the Hallen equation. In both cases, the unknown function is the current distribution, and the integral equation is formulated using the boundary conditions for the electric field. The Hallen equation does not seem to have an extension to general metallic and dielectric structures, so we shall deal only with the EFIE.

The first numerical solution of the EFIE for wires seems to be attributable to Richmond [10, 11]. As the basis functions for the current distribution, he used pulse functions [*piecewise-constant* (PWC) approximation], the Fourier expansion, and several kinds of polynomials, along with the point-matching method (a special case of the MoM[7]) for solving the Pocklington form of the

7. The weighting functions are Dirac's (delta) functions.

EFIE. The pulse functions belong to the class of the so-called subdomain basis functions[8]. The Fourier and the polynomial expansions represent typical entire-domain approximations[9]. However, Richmond's results did not demonstrate any advantage of the entire-domain basis functions over the pulse basis functions. (Some later papers showed that the efficiency of entire-domain basis functions in approximating currents along thin wires could be substantially increased by using the Galerkin method instead of the point-matching method, which was not the case with pulse functions.) The application of subdomain approximations is usually simpler than the application of entire-domain functions, and subdomain basis functions have been favored by most authors ever since.

To solve the EFIE, Harrington combined pulse and triangle basis functions[10] [*piecewise-linear* (PWL) approximation] with the point-matching method and the Galerkin method [16, 20]. He reported two important conclusions. First, for the same accuracy of analysis, the PWL approximation requires two times fewer unknowns than the PWC approximation. Second, the Galerkin method gives more accurate results than the point-matching method when using a small number of unknowns. In addition, together with Mautz, he applied the Galerkin method to wire structures loaded by lumped elements, a case that could not be handled by the point-matching method [17]. Finally, Chao et al. showed that multiple wire junctions could be efficiently analyzed by using overlapping triangles [22]. Thus, a powerful method for solving general wire structures was developed. To obtain acceptable results, this method usually requires about 10 unknowns per wavelength for long wires.

The most efficient method for the thin-wire analysis based on the subdomain basis functions (approximation) was suggested by Richmond [23]. He used a *piecewise-sinusoidal* (PWS) approximation for the wire current. The unknown coefficients are determined by solving the reaction integral equation proposed by Rumsey [24], which is equivalent to the solution of the EFIE by the Galerkin method. This method can give acceptable results with only four unknowns per wavelength for long wires. (To obtain relatively accurate results, about 10 unknowns per wavelength are needed. In that case, the efficiency of the PWS approximation becomes similar to that of the PWL approximation.) An important feature of the PWS approximation is that it results in closed-

8. In a subdomain approximation, the domain in which the unknown function is defined is divided into small cells (wire subsegments for wire structures), and low-order order basis functions are defined over each cell.
9. In an entire-domain approximation, the unknown function is approximated by a unique higher-order series of basis functions over the whole domain (e.g., a whole wire segment for wire structures).
10. Alternatively, they are termed *rooftop basis functions*.

form expressions for the electric field, making this approximation easier to implement than other subdomain basis functions.

Although the subdomain approximation presents no limitation to the complexity of the expansion function used on a subsegment, three-term basis functions seem to be the upper limit found in wide application. Such functions have been used by many authors, almost from the beginning of the numerical analysis of wire structures. Yeh and Mei applied the first trigonometric three-term expansion to solve the Hallen equation [5]. Thereafter, this expansion has often been used to solve the EFIE. Miller et al. introduced the three-term algebraic expansion [25]. Gee et al. suggested a specific extrapolation technique to treat multiple wire junctions [26]. All of these cases use the point-matching method to solve integral equations, which results in a relatively simple and efficient method. The three-term expansion method is described in detail by Miller and Deadrick [27].

Three-term basis functions give more flexibility than one-term basis functions (e.g., pulses) or two-term basis functions (e.g., triangles). However, when combined with the point-matching method, the efficiency of the three-term functions is comparable to the triangle and sinusoidal basis functions combined with the Galerkin method. The following group of papers indirectly explained this paradox. Taylor and Harrison used the same basis functions as Richmond for the analysis of helical antennas, but with the point-matching method [28]. They noticed that the number of unknowns increased rapidly with decreasing wire radius. Pearson and Butler showed that the reason for this behavior came from the discontinuity of the charge distribution [29]. Finally, Butler and Wilton showed that the number of unknowns did not depend on the wire radius if the influence of the charge discontinuity was softened in some way, such as by using the Galerkin method [30]. Considering these papers, it could be expected that the approximation efficiency of the three-term basis functions applied to solve the EFIE would increase by using the Galerkin method instead of the point-matching method.

At the same time, some authors tried to utilize the power of entire-domain approximations. Popovic used the polynomial approximation along thin cylindrical dipoles, which automatically satisfied *Kirchhoff's current law* (KCL) at cylinder ends [31]. To satisfy the KCL, Thiele modified the Fourier expansion used by Richmond [32]. Merewether showed that the number of unknowns needed by such an expansion could further be decreased if the EFIE was solved using the Galerkin method instead of the point-matching technique [33]. However, these two methods were limited to wire structures without junctions. Finally, Silvester and Chan combined a polynomial entire-domain approximation with the Galerkin method to solve wire structures with junctions [34]. The satisfaction of the KCL for wire ends and junctions was forced

by imposing additional equations. Using specific matrix manipulations, these equations were incorporated into the system of linear equations obtained by using the Galerkin method. The resulting method demonstrated high efficiency. However, this technique has not found widespread use, perhaps due to the complicated method of imposing the KCL law.

Finally, the electromagnetic group at the University of Belgrade has expended a lot of effort investigating the possibilities for using the entire-domain approximation in the analysis of thin-wire structures [31, 35–40] based on both the Hallen equation and the EFIE. Regarding the EFIE, in [35] equations written for wire ends and junctions according to the KCL were appended to point-matching equations. For junctions, additional constraints were formulated by using pulse testing functions. However, the required number of unknowns increased as the wire radius decreased, primarily owing to problems in modeling the excitation region. To eliminate this problem and simultaneously increase the speed of the numerical integration, the point-matching method was replaced by the MoM, using subdomain pulse testing functions [37]. An accurate and efficient method was obtained that required six to eight unknowns per wavelength for long wires. The next step incorporated the equations written according to the KCL into the entire-domain expansion and, thus, enabled an easy application of the Galerkin method [38, 39]. This method obtained accurate results with only three to four unknowns per wavelength for long wires. When compared with the Galerkin method combined with triangle basis functions, the method in [38, 39] achieved the same accuracy with three to five times fewer unknowns and in 30 to 100 times shorter analysis times [40]. The approach from [38, 39] is extended in this book to the analysis of general structures.

1.5 Methods for the Analysis of Metallic Structures

Once we have developed an efficient method for the analysis of thin-wire structures, it can be applied to the analysis of arbitrary metallic structures with the metallic surfaces modeled by wire grids, as suggested by Richmond [13]. Often the accuracy of the analysis is unsatisfactory, particularly in the computation of near fields. The number of unknowns needed for the modeling of a wire grid is usually much greater than the number of unknowns required for the direct modeling of the corresponding metallic surface. Investigations of the direct modeling of metallic surfaces date almost from the very beginning of the development of the MoM.

The analysis of arbitrary metallic objects is more complicated than the analysis of wire structures because the current distributed over a metallic surface

has two components[11] that depend on two coordinates, while the current flowing along a thin wire is considered to have one component that depends on one coordinate. The progress in modeling metallic objects was not as rapid as the modeling of wire structures and has left many questions to be answered in the future.

The earliest work on the analysis of 3-D metallic structures based on solving the EFIE seems to be attributable to Andreasen [6]. The solution method was specialized for *bodies of revolution* (BoRs) and inspired many authors to analyze such bodies. Mautz and Harrington [41] proposed the first application of the strict MoM to BoRs in 1969. They used the triangle basis functions along the body generatrix. In subsequent papers, they showed that the EFIE and the *magnetic-field integral equation* (MFIE) failed at interior resonances of the BoRs. They proposed the application of the *combined-field integral equation* (CFIE) [42] and the *combined-source integral equation* (CSIE) [43] to overcome this problem. Finally, Shaeffer and Medgyesi-Mitschang gave the solution for BoRs with attached wires with particular attention paid to wire-to-body junctions [44]. Methods specialized for BoRs enable a substantially more efficient analysis of such bodies than methods intended for the analysis of arbitrary metallic objects, but BoRs represent only one particular class of metallic structures. The analysis of BoRs was often intended only for testing new concepts.

The first method proposed for the modeling of general 3-D structures was suggested by Oshiro in 1965 [14, 15]. Although termed the *source distribution technique*, this method actually represents a special case of the MoM. This method divided a closed surface into small cells and approximated the surface current over each cell by two mutually orthogonal constant components. The current densities are determined by applying the point-matching method to the MFIE. Knepp and Goldhirsh [45] proposed this method by introducing a subdivision of the body into rather sophisticated surface elements (so-called biquadratic spline surfaces), which enabled them to model a body as complicated as a helicopter. Albertsen et al. [46] proposed a method for the analysis of structures composed of both metallic surfaces and wires. For the wires, they formulated the EFIE. They introduced the so-called attachment mode[12] to model wire-to-plate junctions. This gave the first method for the analysis of general closed metallic structures combined with wires, but it did not find a wide application, probably because of the crude approximation used for the currents (pulse functions for both wires and surfaces), which required a relatively large number of unknowns to obtain acceptable results.

11. These components are with respect to a local coordinate system attached to the surface where two coordinates are tangential to the surface.
12. Specific basis functions that model the current flow between a wire and a plate.

In 1975, Wang et al. [47] proposed a PWS approximation for the surface currents over rectangular metallic plates and combined it with the Galerkin method to solve the EFIE. The method demonstrated high accuracy with a relatively small number of unknowns. Newman and Pozar [48] generalized the method by adding wires and introducing a specific attachment mode for handling wire-to-plate junctions. Sing and Adams showed how to define the PWS approximation of surface currents over flat nonrectangular quadrilaterals [49]. Finally, Newman and Tulyathan [50] suggested certain rules for the segmentation of large polygonal plates, which produced an efficient method for the analysis of metallic structures consisting of arbitrarily interconnected polygonal plates. It required only 30 to 40 unknowns per wavelength squared to obtain acceptable results for electrically large plates of simple shapes. Among others, Newman and Alexandropulos used this method for the electromagnetic modeling of large and complicated real structures, such as the Boeing 747 and the Concorde [51]. However, the method is not suitable for curved structures since a flat quadrilateral cannot be defined by four arbitrary points picked on a curved surface. Different attachment modes had to be introduced for various wire-to-plate junctions, as shown by Pozar and Newman [52, 53], which substantially increased the complexity of the code.

In 1980, Glisson and Wilton [54] combined the rooftop basis functions over rectangular surface elements with the Galerkin method for the solution of the EFIE. These basis functions yield a PWL approximation in the direction of the current flow and a PWC approximation in the direction orthogonal to the current flow. Two years later, together with Rao [55], they adopted triangular surface elements as more convenient for the approximation of curved surfaces. The resulting method has been widely used for the last 20 years. The application of this method, however, usually results in a large number of unknowns (e.g., 100–200 per wavelength squared). To minimize the number of unknowns for complicated structures, particular attention should be paid to the generation of an optimal geometrical model (triangular mesh) of the structure analyzed. Virga and Rahmat-Samii adopted the shape-quality factor for triangles from the FEM, along with the corresponding techniques for optimal triangular meshing [56]. More recently, it was shown that an efficient modeling of curved structures could be performed by using quadrilaterals in the form of bilinear surfaces [38, 57]. The rooftop basis functions for such quadrilaterals were defined to satisfy the continuity of the current component normal to the surface edges locally. Establishing the shape-quality factor for bilinear surfaces enabled a further improvement to the modeling by quadrilaterals and reduced the number of unknowns needed in the analysis by more than a half compared with modeling using triangular patches [58].

The first successful attempt to solve a surface problem by an entire-domain approximation was proposed in 1980 [59]. The two components of the surface-

current density on a rectangular scatterer were approximated by a double polynomial series. The unknown coefficients of the series were determined by using the point-matching technique to solve the EFIE. It was demonstrated that 30 unknowns per wavelength squared sufficed to obtain accurate results for the current distribution and the scattered field. Medgyesi-Mitschang and Eftimiu combined a double sinusoidal series with the Galerkin method to solve for the scattering from finite noncircular cylinders [60]. The resulting method was even more efficient than the method reported in [59]. Using the Galerkin test procedure instead of point matching improved the efficiency. However, neither of these two methods could handle general metallic structure with junctions of two or more plates.

After developing basis functions that automatically satisfied the continuity of the current component normal to the plate edges and junctions, it was possible to develop a method for the analysis of general metallic structures [38, 57] that used entire-domain approximation functions and the Galerkin technique. When compared with rooftop basis functions, entire-domain basis functions required 5 to 10 times fewer unknowns per wavelength squared of metallic surfaces and 20 to 500 times shorter CPU time for electrically large plates of simple shapes [61]. In other words, acceptable results could be obtained with even fewer than 10 unknowns per wavelength squared. The applicability of this method was further increased to complicated and electrically large structures, including wire-to-plate junctions. First, the problem of arbitrary wire-to-plate junctions was solved without introducing attachment modes by adopting a unified approach based on a specific segmentation technique [62]. Second, sophisticated automatic procedures were suggested for the segmentation of electrically large plates [63]. All these techniques will be considered in detail in this book.

1.6 Methods for the Analysis of Composite Metallic and Dielectric Structures

The analysis of composite metallic and dielectric structures is much more complicated than the analysis of purely metallic (perfectly conducting) structures. Owing to the complexity of the general problem, the majority of the existing methods have been developed only for cylinders (2-D problems) and BoRs (3-D problems).

Chang and Harrington [64] analyzed material cylinders, while Wu and Tsai [65] analyzed lossy dielectric BoRs by using the formulation that had been originally suggested by Poggio and Miller [66]. The method was later named PMCHW, using the initials of the researchers involved. Solodukhov and

Vasil'ev [67] and Morita [68] analyzed dielectric cylinders, and Vasil'ev and Materikova [69] analyzed dielectric BoRs by the formulation usually attributed to Muller [70]. Mautz and Harrington showed that the PMCHW and Muller's equations were special cases of a more general equation for the analysis of dielectric bodies [71].

Medgyesi-Mitschang and Eftimiu combined the EFIE and the PMCHW to treat metallic BoRs coated by dielectrics [72]. Medgyesi-Mitschang and Putnam analyzed axially inhomogeneous BoRs with multiple metallic and dielectric junctions [73]. To eliminate spurious solutions at internal resonances of coated metallic bodies, Huddleston et al. combined the PMCHW and the CFIE [74]. Putnam and Medgyesi-Mitschang used the same combination of SIEs for the analysis of axially inhomogeneous BoRs [75], paying particular attention to the problem of junctions.

The main effort of many authors has been devoted to the development of an optimal formulation of the SIE that can be used not only for BoRs, but also for arbitrary structures. On the other hand, only moderate attention has been paid to multiple dielectric and metallic junctions.

Only a few published papers are devoted to the analysis of general structures [76–79]. Umashankar et al. developed a method for the analysis of arbitrary homogeneous dielectric objects [76]. The method combined the rooftop basis functions defined over triangles with the Galerkin method for solving the PMCHW. Sarkar et al. used this method to analyze lossy dielectric objects [77]. Rao et al. used the same method with the EFIE for the analysis of conducting bodies coated with lossy materials [78]. Finally, Rao et al. extended the method to handle combined metallic and dielectric structures, where metallic surfaces can be infinitesimally close to dielectric surfaces without their touching each other [79]. This gave rise to the first general method for the analysis of combined metallic and dielectric structures. However, this method requires a relatively large number of unknowns even for simple problems (about 300 unknowns per wavelength squared for dielectric surfaces), which limits its application to simple and electrically relatively small structures.

In the analysis of composite metallic and dielectric structures, the efficiency of the method applied is of a crucial importance. One possible way to achieve such efficiency is to implement entire-domain approximation functions, devoting special care to multiple metallic and dielectric junctions [80]. This concept will be elaborated on in this book.

1.7 What Is New in This Book?

A number of excellent books have appeared during the last three decades that consider MoM solutions of integral equations in the frequency domain [36,

81–93]. Most often, these books treat specific classes of problems (e.g., cylindrical scatterers, wire antennas, BoRs, and waveguides) rather than general 3-D structures. Excluding [91], the techniques applied to general 3-D structures can be summarized as follows:

- The geometry of wires is approximated by right cylinders, while the geometry of metallic and dielectric surfaces is modeled by flat quadrilaterals and triangles. The quality of the geometrical approximation is primarily considered qualitatively, except in the case of modeling by triangles, where the shape-quality factor is adopted from the FEM. Recommendations for meshing electrically large surfaces are given in a rudimentary form. Particularly, in the case of modeling by triangles, the same automatic procedures are used as in the FEM.
- Currents along wires and over metallic and dielectric surfaces are approximated by subdomain basis functions. In particular, the geometrical modeling of surfaces by arbitrary flat quadrilaterals results in line charges at quadrilateral interconnections. A precise inclusion of edge effects is not considered.
- The problem of interconnections of plates and wires is solved by introducing so-called attachment modes, different types of which should be implemented for various types of wire-to-plate junctions. Composite metallic and dielectric junctions are treated in a rudimentary form only for BoRs. Protrusions of wires through dielectric surfaces are considered only for a real ground by using Green's-function techniques (Sommerfeld formulation).
- The excitation of structures can be distributed (e.g., by a uniform plane wave, a TEM wave, or a TE_{10} wave) or localized (e.g., by a point generator or a TEM magnetic current frill).
- Various integral equations are used (e.g., Hallen equation, EFIE, MFIE, CFIE, PMCHW). These equations are solved in most instances by the point-matching method or the Galerkin method, while the method of least squares is seldom used. The Galerkin method is shown to be the most efficient technique for solving integral equations. However, the implementation of the Galerkin method leads to various types of double, triple, and quadruple integrals, which also depend on the integral equation used (e.g., the EFIE and the MFIE).
- Electrically large structures are handled either by using very powerful computers or by hybridization with high-frequency techniques. Antennas above a real ground are solved by hybridization with Green's-function techniques.

In the present book, a unified approach to the electromagnetic modeling of almost-arbitrary composite metallic and dielectric structures in the frequency domain is considered based on the following principal features[13]:

- The modeling of wires is carried out by using so-called generalized wires (in particular, right truncated cones). These wires, as special cases, include cylindrical wires, conical wires, various forms of wire ends, and so forth. Generalized curved quadrilaterals and triangles (in particular, bilinear surfaces and flat triangles) are introduced for the geometrical modeling of metallic and dielectric surfaces, including as special cases flat quadrilaterals and rectangles. The quality of the geometrical approximation is considered quantitatively by introducing the shape-quality factor for generalized quadrilaterals (in particular, for bilinear surfaces).
- Sophisticated automatic techniques for the segmentation of electrically large surfaces are presented, which are applicable for both subdomain and entire-domain approximations for the current distribution.
- Entire-domain basis functions (in particular, polynomials) are adopted for the approximation of currents along wires and over surfaces. For wires, the basis functions automatically satisfy the KCL at wire ends and interconnections. For metallic and dielectric surfaces, the basis functions automatically satisfy the continuity of the current component normal to the surface edges and interconnections. In addition, a precise inclusion of the edge effects and junction effects is investigated.
- The problem of interconnections of plates and wires is solved by a specific segmentation of plates in the vicinity of wires without introducing additional attachment modes. All composite metallic and dielectric junctions, as well as protrusions of wires through dielectric surfaces, are treated in a unified way by introducing multiplet basis functions.
- As an optimal set of excitations, the uniform plane wave, the TEM magnetic current frill (a model of a coaxial-line excitation), and the point generator are stressed. Waveguides are excited by modeling actual feeding circuits and using the point generator.
- Although several integral equations are used, the EFIE is recommended for wires and metallic surfaces, the PMCHW for dielectric surfaces, and the CFIE for the elimination of spurious resonances. Various test procedures are applied in a unified way by using the concept of the

13. Some of these features are partly described in [91], but only for metallic structures. The present book adds a generalization to structures with dielectrics, as well as plenty of new material about metallic structures.

weighted point-matching method, giving priority to the Galerkin method. The implementation of the Galerkin method starts from the generalized scalar formulation of the integral equations, resulting in the same types of integrals for various integral equations.

- Owing to the high efficiency of the method proposed, many electrically large structures are efficiently handled, even using PC computers, without hybridization with high-frequency techniques. Similarly, antennas above a real ground are solved as problems of finite-sized composite metallic and dielectric structures without hybridization with Green's-function techniques.

Thus, a very efficient and flexible method is obtained, which can be applied to almost-arbitrary structures. Although very powerful, the resulting method has inherent restrictions. The basic limitations are related to electrically extremely large objects and to the modeling of anisotropic and bianisotropic media. Hence, guidelines are given in this book for hybridization with the CIE, the FEM, and high-frequency techniques.

1.8 Survey of Chapters

Chapter 2 is devoted to the basic concepts of the numerical electromagnetics. Starting with a simple example, the basic methods for solving integral and differential equations are introduced: the MoM, the Rayleigh-Ritz variational method, the *finite-difference* (FD) method, and the FEM. Some specific topics are also considered, including data postprocessing, memory and analysis time requirements, estimation of solution quality, and adaptive methods.

Chapter 3 exposes the basic theory of time-harmonic fields. It starts from the Maxwell equations, retarded potentials, and expressions for the field vectors. Particular attention is devoted to writing the field vectors in operator notation, which enables writing the integral equations in a compact form. The next part introduces the volume equivalence principle along with the duality relations between the electric and magnetic quantities. The formulation of the general problem in the frequency domain starts with the boundary conditions and the surface equivalence principle. Finally, based on the Poynting theorem and the surface equivalence principle, various forms of the uniqueness theorem are proven and discussed.

Chapter 4 summarizes the field integral equations. The main part is devoted to various types of BIEs: EFIE, MFIE, CFIE, and PMCHW. In particular, the generalized scalar formulation of these equations is presented, enabling a flexible implementation. This summary continues with CIEs, stress-

ing in particular wire and surface CIEs (i.e., including lumped elements and distributed loadings along wires and over plates). Finally, the hybridization of SIEs with VIEs, Green's-function techniques, FEM, and high-frequency techniques is outlined.

Chapter 5 deals with the geometrical modeling of composite structures. First, the modeling of generalized wires by right truncated cones is considered with particular attention paid to wire ends and the piecewise-cylindrical approximation of wires. The main part of this chapter is devoted to the modeling of metallic and dielectric surfaces by generalized quadrilaterals (generalized triangles) and bilinear surfaces (flat triangles). Guidelines are also given for the modeling of dielectric volumes by generalized hexahedrons and tetrahedrons. Particular attention is devoted to the modeling of different types of wire-to-surface junctions. As specific topics, automatic parameterization of 3-D geometries and automatic techniques for the segmentation of electrically large plates are considered.

Chapter 6 deals with approximations of currents and fields. It starts with various types of the approximation of the current along generalized wires. Various conditions for wire ends and multiple junctions are derived and incorporated into the approximation. Thereafter, various approximations of the current distribution over generalized quadrilaterals and triangles are discussed. The advantage of basis functions that automatically satisfy the continuity of the current component normal to surface edges and junctions is demonstrated. After adopting such basis functions, the shape-quality factor for bilinear surfaces is defined, enabling further improvement of the analysis. Similarly, the approximation of the field distribution within generalized hexahedrons and tetrahedrons is elaborated. In particular, it is shown how various types of the current distribution (e.g., those for wires, surfaces, and volumes) can be hybridized. Finally, based on the topological analysis of two canonical problems, the MoM/SIE, MoM/VIE, and FEM are compared. Certain conclusions about the optimal choice of the analysis method are derived.

Chapter 7 deals with modeling the excitation (the uniform plane wave, the point generator, the TEM magnetic current frill, and so forth).

Chapter 8 considers the solution of the field equations by the MoM. First, testing of vector equations in nonorthogonal coordinate systems is given. The next part deals with the weighted point-matching method, which enables a unified approach in the implementation of various test procedures, as well as easy comparison among them. In particular, various topics related to the application of the Galerkin method are considered: the optimal form of the SIE that is tested, evaluation of matrix elements (impedances), and application of the generalized scalar formulation of the SIEs to simplify the testing. Finally, the optimal choice of test procedure is discussed.

Chapter 9 gives complicated numerical examples, including a TV-UHF panel antenna, a reflector antenna together with the waveguide feed and feed struts, and a shielded dielectric horn antenna.

References

[1] Smythe, W. R., *Static and Dynamic Electricity,* New York: McGraw-Hill, 1968.

[2] Harrington, R. F., *Time-Harmonic Electromagnetic Fields,* New York: McGraw-Hill, 1961.

[3] Mei, K. K., and J. G. Van Bladel, "Scattering by Perfectly Conducting Rectangular Cylinders," *IEEE Trans. Antennas Propagat.,* Vol. AP-11, March 1963, pp. 185–192.

[4] Mei, K. K., "On the Integral Equations of Thin Wire Antennas," *IEEE Trans. Antennas Propagat.,* Vol. AP-13, No. 3, May 1965, pp. 374–378.

[5] Yeh, Y. S., and K. K. Mei, "Theory of Conical Equiangular-Spiral Antennas: Part I—Numerical Technique," *IEEE Trans. Antennas Propagat.,* Vol. AP-15, No. 5, September 1967, pp. 634–639.

[6] Andreasen, M. G., "Scattering from Parallel Metallic Cylinders with Arbitrary Cross Section," *IEEE Trans. Antennas Propagat.,* Vol. AP-12, November 1964, pp. 746–754.

[7] Andreasen, M. G., "Scattering from Bodies of Revolution," *IEEE Trans. Antennas Propagat.,* Vol. AP-13, No. 2, March 1965, pp. 303–310.

[8] Andreasen, M. G., "Scattering from Cylinders with Arbitrary Surface Impedance," *Proc. IEEE,* Vol. 53, 1965, pp. 812–817.

[9] Richmond, J. H., "Scattering by a Dielectric Cylinder of Arbitrary Cross-Section Shape," *IEEE Trans. Antennas Propagat.,* Vol. AP-13, May 1965, pp. 338–341.

[10] Richmond, J. H., "Scattering by an Arbitrary Array of Parallel Wires," *IEEE Trans. Microwave Th. and Tech.,* Vol. MTT-13, July 1965, pp. 408–412.

[11] Richmond, J. H., "Digital Computer Solutions of the Rigorous Equations for Scattering Problems," *Proc. IEEE,* Vol. 53, 1965, pp. 796–804.

[12] Richmond, J. H., "TE-Wave Scattering by a Dielectric Cylinder of Arbitrary Cross-Section Shape," *IEEE Trans. Antennas Propagat.,* Vol. AP-14, July 1966, pp. 460–464.

[13] Richmond, J. H., "A Wire Grid-Model for Scattering by Conducting Bodies," *IEEE Trans. Antennas Propagat.,* Vol. AP-14, November 1966, pp. 782–786.

[14] Oshiro, F. K., "Source Distribution Technique for the Solution of General Electromagnetic Scattering Problems," *Proc. of 1st GISAT Symp.,* Vol. 1, Pt. 1, Mitre Corporation, pp. 83–107.

[15] Oshiro, F. K., and K. M. Mitzner, "Digital Computer Solution of Three-Dimensional Scattering Problem," *Digest of IEEE Int. Antennas Propagat. Symp.,* October 1967, pp. 257–263.

[16] Harrington, R. F., "Matrix Methods for Field Problems," *Proc. IEEE,* Vol. 55, No. 2, February 1967, pp. 136–149.

[17] Harrington, R. F., and J. R. Mautz, "Straight Wires with Arbitrary Excitation and Loading," *IEEE Trans. Antennas Propagat.,* Vol. AP-15, July 1967, pp. 502–515.

[18] Kantorovich, L. V., and V. I. Krylov, *Approximate Methods of Higher Analysis,* New York: John Wiley and Sons, 1964.

[19] Kantorovich, L. V., and G. P. Akilov, *Functional Analysis in Normed Spaces,* Oxford, U.K.: Pergamon Press, 1964.

[20] Harrington, R. F., *Field Computation by Moment Methods,* New York: Macmillan, 1968; reprinted by New York: IEEE Press, 1993.

[21] Silvester, P. P., and R. L. Ferrari, *Finite Elements for Electrical Engineers,* 3rd ed., New York: Cambridge University Press, 1996.

[22] Chao, H. H., B. J. Strait, and C. D. Taylor, "Radiation and Scattering by Configurations of Bent Wires with Junctions," *IEEE Trans. Antennas Propagat.,* Vol. AP-19, No. 5, September 1971, pp. 701–702.

[23] Richmond, J. H., *Computer Analysis of Three-Dimensional Wire Antennas,* Ohio State University, ElectroScience Laboratory, Department of Electrical Engineering, Report 2708–4, 1969.

[24] Rumsey, V. H., "Reaction Concept in Electromagnetic Theory," *Physical Review,* Vol. 94, No. 6, 1954, pp. 1483–1491.

[25] Miller, E. K., et al., "Radar Cross Section of a Long Wire," *IEEE Trans. Antennas Propagat.,* Vol. AP-17, No. 3, May 1969, pp. 381–384.

[26] Gee, S., et al., "Computer Techniques for Electromagnetic Scattering and Radiation Analysis," *IEEE Internat. Electromagn. Compat. Symp. Rec.,* 1971, pp. 122–131.

[27] Miller, E. K., and F. J. Deadrick, "Some Computational Aspects of Thin-Wire Modeling," in *Numerical and Asymptotic Techniques in Electromagnetics,* R. Mittra, (ed.), New York: Springer-Verlag, 1975.

[28] Taylor, C. D., and C. W. Harrison, Jr., "On Thin-Wire Multiturn Loop Antennas," *IEEE Trans. Antennas Propagat.,* Vol. AP-22, No. 3, May 1974, pp. 407–413.

[29] Pearson, L. W., and C. M. Butler, "Inadequacies of Collocation Solutions to Pocklington-Type Models of Thin-Wire Structures," *IEEE Trans. Antennas Propagat.,* Vol. AP-23, No. 2, March 1975, pp. 295–298.

[30] Butler, C. M., and D. R. Wilton, "Analysis of Various Numerical Techniques Applied to Thin-Wire Scatterers," *IEEE Trans. Antennas Propagat.,* Vol. AP-23, No. 4, July 1975, pp. 534–540.

[31] Popovic, B. D., "Polynomial Approximation of Current Along Thin-Symmetrical Cylindrical Dipoles," *Proc. IEE,* Vol. 117, 1970, pp. 873–878.

[32] Thiele, G. A., "Wire Antennas," in *Computer Techniques for Electromagnetics,* R. Mittra (ed.), Elmsford, NY: Pergamon Press, 1973.

[33] Merewether, D. E., "The Arbitrarily Driven Long Cylindrical Antenna," *IEEE Trans. Antennas Propagat.,* Vol. AP-23, No. 4, July 1975, pp. 534–540.

[34] Silvester, P. P., and K. K. Chan, "Analysis of Antenna Structures Assembled from Arbitrarily Located Straight Wires," *Proc. IEE,* Vol. 120, No. 1., January 1973, pp. 21–26.

[35] Djordjevic, A. R., B. D. Popovic, and M. B. Dragovic, "A Rapid Method for Analysis of Wire-Antenna Structures," *Archiv fur Electrotechnik,* Vol. 61, 1979, pp. 17–23.

[36] Popovic, B. D., M. B. Dragovic, and A. R. Djordjevic, *Analysis and Synthesis of Wire Antennas,* Chichester, U.K.: Research Studies Press, 1982.

[37] Djordjevic, A. R., et al., *Analysis of Wire Antennas and Scatterers,* Norwood, MA: Artech House, 1990.

[38] Kolundzija, B. M., "General Entire-Domain Galerkin Method for Electromagnetic Modeling of Composite Wire to Plate Structures," *Proc. of 20th EuMc,* Budapest, Hungary, Pt. 1, 1990, pp. 853–858.

[39] Kolundzija, B. M., and B. D. Popovic, "Entire-Domain Galerkin Method for Analysis of Generalized Wire Antennas and Scatterers," *Proc. IEE H,* Vol. 139, No. 1, February 1992, pp. 17–24.

[40] Kolundzija, B. M., "Comparison of a Class of Sub-Domain and Entire-Domain Basis Functions Automatically Satisfying KCL," *IEEE Trans. on Antennas and Propagat.,* Vol. 44, No. 10, October 1996, pp. 1362–1366.

[41] Mautz, J. R., and R. F. Harrington, "Radiation and Scattering from Bodies of Revolution," *J. Appl. Sci. Res.,* Vol. 20, 1969, pp. 405–413.

[42] Mautz, J. R., and R. F. Harrington, "H-Field, E-Field and Combined-Field Solutions for Conducting Bodies of Revolution," *Arch. Elek. Ubertragung,* Vol. 32, No. 4, 1978, pp. 157–164.

[43] Mautz, J. R., and R. F. Harrington, "A Combined-Source Solution for Radiation and Scattering from a Perfectly Conducting Body," *IEEE Trans. Antennas Propagat.,* Vol. AP-27, No. 4, 1979, pp. 445–454.

[44] Shaeffer, J. F., and L. N. Medgyesi-Mitschang, "Radiation from Wire Antenna Attached to Bodies of Revolution: The Junction Problem," *IEEE Trans. Antennas Propagat.,* Vol. AP-29, No. 3, May 1981, pp. 479–487.

[45] Knepp, L., and J. Goldhirsh, "Numerical Analysis of Electromagnetic Radiation Properties of Smooth Conducting Bodies of Arbitrary Shape," *IEEE Trans. Antennas Propagat.,* Vol. AP-29, No. 3, May 1981, pp. 479–487.

[46] Albertsen, N. C., J. E. Hansen, and N. E. Jensen, "Computation of Radiation from Wire Antennas on Conducting Bodies," *IEEE Trans. Antennas Propagat.,* Vol. AP-22, No. 2, March 1974, pp. 479–487.

[47] Wang, N. N., J. H. Richmond, and M. C. Gilreath, "Sinusoidal Reaction Formulation for Radiation and Scattering from Conducting Surfaces," *IEEE Trans. Antennas Propagat.,* Vol. AP-23, No. 3, May 1975, pp. 376–382.

[48] Newman, E. H., and D. M. Pozar, "Electromagnetic Modeling of Composite Wire and Surface Geometries," *IEEE Trans. Antennas Propagat.,* Vol. AP-26, No. 6, November 1978, pp. 784–789.

[49] Singh, J., and A. T. Adams, "A Nonrectangular Patch Model for Scattering from Surfaces," *IEEE Trans. Antennas Propagat.,* Vol. AP-27, No. 4, July 1979, pp. 531–535.

[50] Newman, E. H., and P. Tulyathan, "A Surface Patch Model for Polygonal Plates," *IEEE Trans. Antennas Propagat.,* Vol. AP-30, No. 4, July 1982, pp. 588–593.

[51] Newman, E. H., and P. Alexandropulos, "Polygonal Plate Modeling of Realistic Structures," *IEEE Trans. Antennas Propagat.,* Vol. AP-32, No. 7, July 1984, pp. 742–747.

[52] Pozar, D. M., and E. H. Newman, "Analysis of a Monopole Mounted Near or at the Edge of a Half-Plane," *IEEE Trans. Antennas Propagat.,* Vol. AP-29, No. 3, May 1981, pp. 488–495.

[53] Pozar, D. M., and E. H. Newman, "Analysis of a Monopole Mounted Near an Edge or a Vertex," *IEEE Trans. Antennas Propagat.,* Vol. AP-30, No. 3, May 1982, pp. 401–408.

[54] Glisson, A. W., and D. R. Wilton, "Simple and Efficient Numerical Methods for Problems of Electromagnetic Radiation and Scattering from Surfaces," *IEEE Trans. Antennas Propagat.,* Vol. AP-28, No. 5, September 1980, pp. 593–603.

[55] Rao, S. M., D. R. Wilton, and A. W. Glisson, "Electromagnetic Scattering by Surfaces of Arbitrary Shape," *IEEE Trans. Antennas Propagat.,* Vol. AP-30, No. 3, May 1982, pp. 409–418.

[56] Virga, K. L., and Y. Rahmat-Samii, "RCS Characterization of a Finite Ground Plane with Perforated Apertures: Simulations and Measurements," *IEEE Trans. Antennas Propagat.,* Vol. AP-42, No. 3, July 1994, pp. 1491–1501.

[57] Kolundzija, B. M., and B. D. Popovic, "Entire-Domain Galerkin Method for Analysis of Metallic Antennas and Scatterers," *Proc. IEE,* Pt. H, Vol. 140, No. 1, February 1993, pp. 1–10.

[58] Kolundzija, B. M., "On the Locally Continuous Formulation of Surface Doublets," *IEEE Trans. Antennas Propagat.,* Vol. AP-46, No. 12, December 1998, pp. 1879–1883.

[59] Djordjevic, A. R., and B. M. Kolundzija, "Analysis of Rectangular Scatterers," *Proc. of XXIV conf. of ETAN,* Pristina, Pt. 2, 1980, pp. 463–470 (in Serbian).

[60] Medgyesi-Mitschang, L. N., and C. Eftimiu, "Scattering from Wires and Open Circular Cylinders of Finite Length Using Entire Domain Galerkin Expansions," *IEEE Trans. Antennas Propagat.,* Vol. AP-30, No. 4, July 1982, pp. 628–636.

[61] Kolundzija, B. M., "Accurate Solution of Square Scatterer as Benchmark for Validation of Electromagnetic Modeling of Plate Structures," *IEEE Trans. Antennas Propagat.,* Vol. AP-46, July 1998, pp. 1009–1014.

[62] Kolundzija, B. M., and B. D. Popovic, "General Localized Junction Model in the Analysis of Wire to Plate Junctions," *Proc. IEE,* Pt. H, Vol. 141, No. 1, February 1994, pp. 1–7.

[63] Kolundzija, B. M., "Automatic Mesh Generation Using Single- and Double-Node Segmentation Techniques," *IEEE Antennas and Propagat. Magazine,* Vol. 40, August 1998, pp. 30–38.

[64] Chang, Y., and R. F. Harrington, "A Surface Formulation for Characteristic Modes of Material Bodies," *IEEE Trans. Antennas Propagat.,* Vol. AP-25, No. 6, November 1977, pp. 789–795.

[65] Wu, T. K., and L. L. Tsai, "Scattering from Arbitrarily-Shaped Lossy Dielectric Bodies of Revolution," *Radio Sci.,* Vol. 12, No. 5, 1977, pp. 709–718.

[66] Poggio, J., and E. K. Miller, "Integral Equation Solutions of Three-Dimensional Scattering Problems," in *Computer Techniques for Electromagnetics,* R. Mittra, (ed.), Oxford, U.K.: Pergamon, 1973.

[67] Solodukhov, V. V., and E. N. Vasil'ev, "Diffraction of a Plane Electromagnetic Wave by a Dielectric Cylinder of Arbitrary Cross Section," *Soviet Physics—Tech. Physics,* Vol. 15, 1970, pp. 32–36.

[68] Morita, N., "Analysis of Scattering by a Dielectric Rectangular Cylinder by Means of Integral Equation Formulation," *Electron and Commun. in Japan,* Vol. 57-B, October 1974, pp. 72–80.

[69] Vasil'ev, E. N., and L. B. Materikova, "Excitation of Dielectric Bodies of Revolution," *Soviet Physics—Tech. Physics,* Vol. 10, 1966, pp. 1401–1406.

[70] Muller, C., *Foundations of the Mathematical Theory of Electromagnetic Waves,* Berlin, Germany: Springer-Verlag, 1969.

[71] Mautz, J. R., and R. F. Harrington, "Electromagnetic Scattering from a Homogeneous Material Body of Revolution," *Arch. Elek. Ubertragung,* Vol. 33, 1979, pp. 71–80.

[72] Medgyesi-Mitschang, L. N., and C. Eftimiu, "Scattering from Axisymmetric Obstacles Embedded in Axisymmetric Dielectrics: The MoM Solution," *Appl. Phys.,* Vol. 19, 1979, pp. 275–285.

[73] Medgyesi-Mitschang, L. N., and J. M. Putnam, "Electromagnetic Scattering from Axially Inhomogeneous Bodies of Revolution," *IEEE Trans. Antennas Propagat.,* Vol. AP-32, August 1984, pp. 797–806.

[74] Huddleston, P. L., L. N. Medgyesi-Mitschang, and J. M. Putnam, "Combined Field Integral Equation Formulation for Scattering by Dielectrically Coated Conducting Bodies," *IEEE Trans. Antennas Propagat.,* Vol. AP-34, April 1986, pp. 510–520.

[75] Putnam, J. M., and L. N. Medgyesi-Mitschang, "Combined Field Integral Equation for Inhomogeneous Two- and Three-Dimensional Bodies: The Junction Problem," *IEEE Trans. Antennas Propagat.,* Vol. AP-39, May 1991, pp. 667–672.

[76] Umashankar, K., A. Taflove, and S. M. Rao, "Electromagnetic Scattering by Arbitrary Shaped Three-Dimensional Homogeneous Lossy Dielectric Objects," *IEEE Trans. Antennas Propagat.,* Vol. AP-34, June 1986, pp. 758–766.

[77] Sarkar, T. K., E. Arvas, and S. Ponnapalli, "Electromagnetic Scattering from Dielectric Bodies," *IEEE Trans. Antennas Propagat.,* Vol. AP-37, May 1989, pp. 673–676.

[78] Rao, S. M., et al., "Electromagnetic Scattering from Arbitrary Shaped Conducting Bodies Coated with Lossy Materials of Arbitrary Thickness," *IEEE Trans. Antennas Propagat.,* Vol. AP-39, May 1991, pp. 627–631.

[79] Rao, S. M., et al., "Electromagnetic Radiation and Scattering from Finite Conducting and Dielectric Structures: Surface/Surface Formulation," *IEEE Trans. Antennas Propagat.,* Vol. AP-39, July 1991, pp. 1034–1037.

[80] Kolundzija, B. M., "Electromagnetic Modeling of Composite Metallic and Dielectric Structures," *IEEE Trans. Microwave Theory and Techniques,* Vol. 47, No. 7, July 1999, pp. 1021–1032.

[81] Mittra, R., (ed.), *Computer Techniques for Electromagnetics,* Elmsford, NY: Pergamon Press, 1973.

[82] Mittra, R., (ed.), *Numerical and Asymptotic Techniques in Electromagnetics,* New York: Springer Verlag, 1975.

[83] Strait, B. J., (ed.), *Application of the Method of Moments to Electromagnetic Fields,* St. Cloud, FL: The SCEEE Press, 1980.

[84] Moore, J., and R. Pizer, *Moment Methods in Electromagnetics: Techniques and Applications,* New York: John Wiley and Sons, 1984.

[85] Itoh, T., (ed.), *Numerical Techniques for Microwave and Millimeter-Wave Passive Structures,* New York: John Wiley and Sons, 1990.

[86] Hansen, R. C., (ed.), *Moment Methods in Antennas and Scattering,* Norwood, MA: Artech House, 1990.

[87] Wang, J. H. H., *Generalized Moment Methods in Electromagnetics,* New York: John Wiley and Sons, 1991.

[88] Miller, E. K., L. Medgyesi-Mitschang, and E. H. Newman, (eds.), *Computational Electromagnetics: Frequency-Domain MoM,* Piscataway, NJ: IEEE Press, 1992.

[89] Morita, N., N. Kumagai, and J. Mautz, *Integral Equation Methods for Electromagnetics,* Norwood, MA: Artech House, 1992.

[90] Umashankar, K., *Computational Electromagnetics,* Norwood, MA: Artech House, 1993.

[91] Popovic, B. D., and B. M. Kolundzija, *Analysis of Metallic Antennas and Scatterers,* London: IEE, 1994.

[92] Bancroft, R., *Understanding Electromagnetic Scattering Using the Moment Method,* Norwood, MA: Artech House, 1996.

[93] Peterson, A., S. C. Ray, and R. Mittra, *Computational Methods for Electromagnetics,* New York: IEEE Press, and Oxford, U.K.: Oxford University Press, 1998.

2

MoM

The MoM is a general solution method for linear electromagnetic-field problems based on transformation of field equations into a system of linear equations. The mathematical basis for the MoM has been known for a long time [1, 2]; however, the method was used for solving particular electromagnetic-field problems, such as the analysis of linear antennas in the early 1950s [3, 4], without the knowledge that the same principle could be applied to the solution of arbitrary field problems. The application of the MoM to solving electromagnetic-field problems was first systematically investigated by Harrington [5]. This investigation culminated in 1968 in his book [6], which has been used by many authors as a basic reference.

In this chapter the MoM will be introduced gradually, starting from the formulation of a deterministic field problem and the concept of linear operators (Sections 2.1 and 2.2). Basically, a linear operator equation can be of the integral or differential type, and the application of the MoM significantly differs in these two cases. Consequently, variants of the MoM applied to differential equations are called the FD method and the FEM. The basic concepts of the MoM are explained in Section 2.3, which is devoted to the solution of integral equations. Concepts typical for differential equations are given in Section 2.4. The MoM applied to integral equations is compared with the FD method and the FEM in Section 2.5.

2.1 Formulation of Deterministic-Field Problems

In this book, we analyze electromagnetic-field problems that have a unique solution. For a typical problem of this class, we give the geometry of the structure

(e.g., antenna shape and dimensions) along with the electrical parameters of the media (permittivity, conductivity, and permeability) and the excitation (e.g., an incident plane wave). The objective is to evaluate certain quantities of interest, like the input impedance, radiation pattern, and current distribution.

The first step in solving such a problem is to choose a set of equations (based on Maxwell's equations or derived from them) that properly describes the structure analyzed and yields a unique solution. A problem for which there exists a unique solution is referred to as a deterministic problem.[1]

Usually, the solution can be obtained starting from various sets of equations and boundary conditions. For example, equations can be written either in integral form or in differential form. Each of these sets can be solved by means of various numerical methods. The accuracy and efficiency of the solution depend on both the starting set of equations and the numerical method used for its solution.

Example: The task is to determine the characteristic impedance of a two-conductor line situated in a vacuum (Figure 2.1). The cross sections of conductors are squares of side a, and the distance between their axes is d. The characteristic impedance can be calculated as $Z_c = (c_0 C')^{-1}$, where c_0 is the light velocity in a vacuum and C' is the per-unit-length capacitance of this

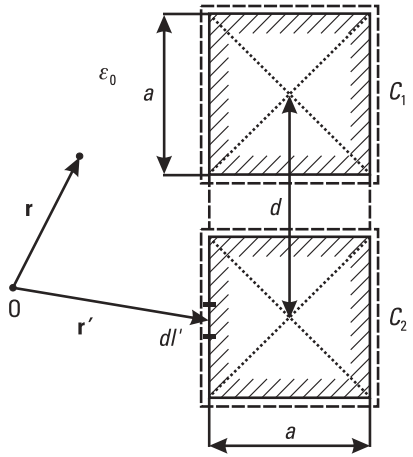

Figure 2.1 Transmission line with square conductors.

1. There exist electromagnetic-field problems that are not deterministic (e.g., modes in a rectangular waveguide) obtained by solving an eigenvalue equation (wave equation). Their solutions are not unique as they include unknown multiplicative constants.

line. The easiest way to determine the capacitance is to solve the corresponding 2-D electrostatic-field problem.[2]

The electric field in a vacuum, **E**, is governed by two differential equations:

$$\nabla \times \mathbf{E}(\mathbf{r}) = 0 \qquad \nabla \cdot \mathbf{E}(\mathbf{r}) = 0 \tag{2.1a,b}$$

where ∇ is "del" operator and **r** is the position vector of an arbitrary field point between conductors.[3] Equation (2.1a) is automatically satisfied if the electric field is expressed in terms of the electrostatic potential V as $\mathbf{E}(\mathbf{r}) = -\nabla V(\mathbf{r})$. After replacing this relation in (2.1b), the Laplace equation is obtained:

$$\Delta V(\mathbf{r}) = 0 \tag{2.2}$$

where $\Delta = \nabla \cdot \nabla$ is the Laplace operator (Laplacian). For a 2-D problem, V is a function of the transverse coordinates only (e.g., the Cartesian x- and y-coordinates).

The same problem can be analyzed utilizing an integral equation, which is obtained if the potential is expressed in terms of surface charges distributed over the conductors. If the total per-unit-length charges of the two conductors are of equal magnitude, but opposite sign (i.e., $Q_1' = -Q_2'$), the integral equation has the form

$$-\frac{1}{2\pi\epsilon_0} \oint_C \rho_s(\mathbf{r}') \ln|\mathbf{r} - \mathbf{r}'| dl' = V(\mathbf{r}) \tag{2.3}$$

where ρ_s is the surface-charge density, $\mathbf{r}'$ is the position vector of the source point,[4] and dl' is a length element along the contour C [7]. The contour C is a union of C_1 and C_2 (the contours of the first and second conductor, respectively). Note that the potential given by (2.3) automatically satisfies the Laplace equation (2.2), which can be proven after lengthy manipulations.

The Laplace equation (2.2) has a unique solution under the conditions defined in the corresponding uniqueness theorem [8]. For example, a unique solution is obtained if the conductor potentials are specified

2. In a 2-D electrostatic-field problem, the charge distribution on the conductor surfaces and the field and potential distributions in the space between conductors are not dependent on the longitudinal coordinate.
3. The field point is the spatial point at which the electromagnetic fields and potentials are evaluated or observed.
4. The source point is the spatial point at which the electromagnetic-field sources (charges and currents) are located.

$$V(\mathbf{r}) = \begin{cases} V_1 \text{ along } C_1 \\ V_2 \text{ along } C_2 \end{cases} \tag{2.4}$$

When this boundary condition is combined with the differential equation (2.2), a unique solution for the potential distribution $V(\mathbf{r})$ is provided. If this boundary condition is substituted into the right-hand side of the integral equation (2.3), the charge distribution $\rho_s(\mathbf{r}')$ is uniquely determined.

Table 2.1 summarizes the above differential and integral formulations used in the analysis of two-conductor lines in a vacuum. Each formulation has certain advantages and drawbacks, which we can address only after providing more insight into the numerical methods.

2.2 Linear Operator Equation

Both equations (integral and differential) given in Table 2.1 are particular cases of linear operator equations whose general form reads

$$L(f) = g \tag{2.5}$$

where L is a linear operator, g is a known function (excitation), and f is the unknown function to be determined (response). To simplify notation, in some cases we shall write the operator equation as $Lf = g$.

Generally, an operator is a mapping of a function f into a function g. For example, the third root, $\sqrt[3]{\ }$, maps the cubic function, x^3, into the linear function, x. However, this is not a linear operator. A linear operator is the one that satisfies the property

$$L(af_1 + bf_2) = aLf_1 + bLf_2 \tag{2.6}$$

Table 2.1
Summary of Basic Equations Used in the Analysis of Two Conductor Lines

	Differential Equation	**Integral Equation**
Formulation	$\Delta V(\mathbf{r}) = 0$ $V(\mathbf{r}) = \begin{cases} V_1 \text{ along } C_1 \\ V_2 \text{ along } C_2 \end{cases}$	$-\frac{1}{2\pi\epsilon_0}\oint_C \rho_s(\mathbf{r}')\ln\lvert\mathbf{r}-\mathbf{r}'\rvert dl' = \begin{cases} V_1 \text{ along } C_1 \\ V_2 \text{ along } C_2 \end{cases}$ $(Q_1' = -Q_2')$
Unknown quantity	$V(\mathbf{r})$	$\rho_s(\mathbf{r}')$

where a and b are arbitrary constants. In the electromagnetic-field analysis, linear operators occur for linear problems (i.e., for systems where the media are linear). Depending on the formulation, the operator can be differential or integral, as is demonstrated in Section 2.1.

Usually, the excitation in electromagnetic-field problem (g) is a known field or potential (e.g., the electrostatic potential in Section 2.1), which is a spatial function of field-point position vector $\mathbf{r}$. In a differential equation, the response (f) is a field or potential (e.g., the electrostatic potential in the Laplace equation). In an integral equation, the response is a field source [e.g., the surface charges in (2.3)]. In both cases, the response is a spatial function of the source-point position vector $\mathbf{r}'$. Hence, the operator equation (2.5) can be written more precisely as $L[f(\mathbf{r}')] = g(\mathbf{r})$. When L is a differential operator, these two position vectors coincide (i.e., $\mathbf{r} = \mathbf{r}'$). When L is an integral operator, these two position vectors are independent.

In electrostatics, the excitation and response are real scalar or vector functions. To further simplify the introduction of basic concepts of the MoM, we shall consider in this chapter only scalar functions (unless stated otherwise). However, in the steady-state analysis of arbitrary electromagnetic structures, both scalar and vector functions can be complex. Hence, most often in other chapters of this book we shall use complex vector linear operator equations.

Generally, an operator cannot act on an arbitrary function; nor can an arbitrary function be obtained as the result of the mapping by the operator. For example, the Laplacian can act only on twice differentiable spatial functions. Thus we define the domain of an operator as the set of all functions f on which it can act. Similarly, we define the range of an operator as the set of all possible functions g resulting from the operation. Hence, the full definition of a linear operator L includes both its mathematical expression that describes the mapping, which satisfies (2.6), and the definition of the domain. (The rigorous mathematical foundation of linear operators and their properties can be found in [9].)

The solution of a linear operator equation consists of determining the response (f) for the given linear operator (L) and excitation (g). This solution procedure is called the analysis.[5] For a deterministic problem, the solution can be written as

$$f(\mathbf{r}') = L^{-1}[g(\mathbf{r})] \tag{2.7}$$

where L^{-1} is the so-called inverse operator of the operator L. If the right side of the above equation is written in a closed-form expression, it is called the analytical solution of the linear operator equation (2.5).

5. The problem of synthesis is defined as determination of the linear operator (L) given response (f) and excitation (g).

Linear operator equations can be solved analytically only in some special cases. Otherwise they must be solved numerically. Moreover, in some cases it is much easier to find a solution numerically than analytically. For example, the capacitance of a square coaxial line can be determined by conformal mapping [10]. However, engineers are usually not familiar with this technique. Instead of learning to do the conformal mapping, it is much easier for them either to solve the problem numerically or use a commercially available software [7].

The excitation is defined at an infinite number of points (i.e., it represents an infinite set of known numbers). The response is also defined at an infinite number of points (i.e., it represents an infinite set of unknown values). Thus, the linear operator equation can be considered as a set of an infinite number of equations in terms of an infinite number of unknowns. However, the "numerical solution of a linear operator equation" means that the solution is obtained after a finite number of numerical operations (i.e., mathematical operations applied to numbers, not to functions). The numerical solution treats both the excitation and the response as finite sets of numbers. Hence, the main problem of the numerical solution of a linear operator equation is to transform this infinite number of equations, which depend on an infinite number of unknowns, into a system of a finite number of equations (possibly linear) that depend on a finite (possibly the same) number of unknowns. Once such a system is obtained, it can be solved by using standard numerical methods.

Example: The integral equation given in Table 2.1 can be shortly written in the linear operator form as

$$V(\rho_s) = V(\mathbf{r}) \qquad V(\rho_s) = -\frac{1}{2\pi\epsilon_0}\oint_C \rho_s(\mathbf{r}')\ln|\mathbf{r}-\mathbf{r}'|dl' \tag{2.8a,b}$$

where V is the linear operator, ρ_s is the response (an unknown function to be determined), and $V(\mathbf{r})$ is the excitation given by (2.4). The domain of the operator consists of all functions ρ_s such that $Q_1' = -Q_2'$.

The efficiency of the analysis, accuracy of the solution, ease of implementation, and so forth, depend a lot on the choice of the linear operator equation used. Methods used to solve integral equations are different from the methods used to solve differential equations. In what follows, the solution of integral equations will be considered first.

2.3 Solution of Integral Equations

2.3.1 Approximation of Unknown Function

Numerical solution of a linear operator equation is always based on some approximation of the unknown function f. Most often the approximation, f_a, is adopted in the form of a finite series,

$$f_a = \sum_{i=1}^{N} a_i f_i \tag{2.9}$$

where N is order of approximation, f_i are known functions, and a_i are unknown coefficients to be determined. Thus, the infinite number of unknown values of the original function, f, is reduced to N unknown coefficients. The functions f_i are termed expansion functions or basis functions.

The definition of a basis function includes the specification of the geometrical element (linear segment, surface patch, or volume brick) over which the function is defined and the mathematical expression of the function. Geometrical elements can be interconnected or overlapped. One or more linearly independent basis functions can be defined over the same element. If only one basis function is defined over each element, it is said that the approximation of the unknown function is performed by zero-order basis functions. (In that case the number of elements is equal to the number of basis functions, N.) If several basis functions are defined over each element, it is said that approximation of unknown function is performed by higher-order basis functions. From the viewpoint of implementation, higher-order basis functions are more complicated than zero-order basis functions. Hence, in this chapter only zero-order basis functions will be discussed. However, higher-order basis functions can significantly increase the efficiency of an analysis, which is extensively exploited later in the book.

The accuracy of the approximation and the accuracy and efficiency of the analysis depend on the approximation order and on the choice of the basis functions. The choice of the basis functions consists of two steps. First, the structure considered is represented as a combination of simple geometrical elements (segments, surfaces, volumes). Second, for each geometrical element the mathematical expression of the basis function is defined. The first step is termed geometrical modeling. The second step represents the approximation of the unknown function in a narrow sense. This choice is significantly simplified if all geometrical elements belong to a limited number of classes and if basis functions of the same form are adopted for all geometrical elements of the same class. Once we define these classes, the problem of choosing the basis

functions is reduced to the problem of the geometrical modeling of the structure considered.

Often in practice the original structure can be defined by a few geometrical elements. However, for the sake of analysis, these elements should be subdivided into a number of smaller geometrical elements. This procedure is termed segmentation or meshing. When the geometrical elements are uniformly subdivided, we talk about uniform segmentation. Most often the accuracy of approximation can be increased if a proper nonuniform segmentation is applied (see Section 2.3.13). However, being based on the experience and prediction, automatic nonuniform segmentation techniques sometimes do not give good enough results. In that case, the analysis should be repeated with a modified starting expansion. It is desirable to have a criterion for this modification. In that case we talk about adaptive segmentation (see Section 2.3.14).

Example: Consider the charge distribution along the transmission line shown in Figure 2.1. The structure that should be modeled consists of two identical square contours. The simplest geometrical element that can be used for approximation of any contour (even curved) is a straight segment. The simplest basis function is constant along the corresponding segment, and zero elsewhere. Each square contour can be modeled by only four straight segments. However, such an approximation is relatively poor. Hence, each square side should be subdivided into subsegments. The simplest segmentation is obtained if each side is uniformly subdivided into equal number of subsegments, n, as shown for $n = 5$ in Figure 2.2. In that case the approximation can be written as

$$\rho_{sa} = \sum_{i=1}^{N} a_i \rho_{si} \qquad \rho_{si} = \begin{cases} 1 & \text{along } i\text{th segment} \\ 0 & \text{elsewhere} \end{cases} \tag{2.10a,b}$$

where $N = 8n$, and ρ_{si} is the unit surface-charge density along the ith segment.

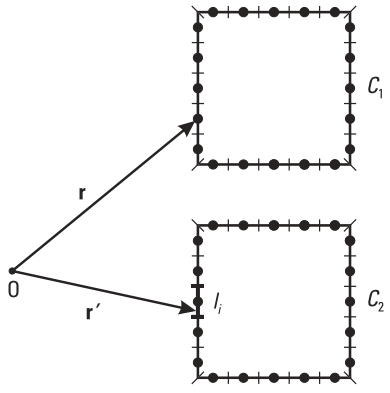

Figure 2.2 Geometrical model of the line shown in Figure 2.1. Dots represent matching points.

2.3.2 Point-Matching Method

After replacing the unknown function f by its approximation f_a given by (2.9), the linear operator equation (2.5) is written in the approximate form

$$\sum_{i=1}^{N} a_i L[f_i(\mathbf{r}')] \approx g(\mathbf{r}) \tag{2.11}$$

Using a finite number of unknown coefficients, this equation cannot be exactly satisfied at an infinite number of spatial points.[6] Several procedures are available to find an approximate solution of equation (2.11). The simplest way is to require the equation be satisfied exactly at N specified points; that is,

$$\sum_{i=1}^{N} a_i L[f_i(\mathbf{r}')] = g(\mathbf{r}_j) \qquad j = 1, \ldots, N \tag{2.12}$$

where $\mathbf{r}_j$ are position vectors of those points. This system of linear equations can be easily solved by using a standard method (e.g., the Gaussian elimination [11]). The points at which the equation is satisfied are called matching points. Hence, this method for solving linear operator equations is known as the point-matching method, or the collocation method [6]. The specific problem related to this method is the optimal choice of matching points [12, 13].

Example: Consider an arbitrary symmetrical two-conductor line. Due to symmetry, the condition $Q_1' = -Q_2'$ is automatically satisfied and $V_1 = -V_2$. Also, V_1 is assumed as given. After replacing (2.10a) into (2.8a), the approximate linear operator equation is obtained as

$$\sum_{i=1}^{N} a_i V_i \approx V(\mathbf{r}) \qquad V_i = V[\rho_{si}(\mathbf{r}')] \tag{2.13a,b}$$

The point-matching equations are obtained by requiring that (2.13) be satisfied at a set of properly adopted points. Using (2.8b), we obtain the system of linear equations

6. If the equation is forced to be satisfied at all points in a space, it can be considered as an overdetermined system of an infinite number of linear equations with a finite number of unknowns, N.

$$\sum_{i=1}^{N} a_i V_{ij} = \begin{cases} V_1 & \mathbf{r}_j \text{ placed on } C_1 \\ -V_1 & \mathbf{r}_j \text{ placed on } C_2 \end{cases} \tag{2.14}$$

$$V_{ij} = -\frac{1}{2\pi\epsilon_0} \oint_{L_i} \ln|\mathbf{r}_j - \mathbf{r}'|\, dl' \qquad j = 1, \ldots, N$$

where L_i denotes the ith segment. By using complex calculus, the integrals V_{ij} are expressed in closed form as

$$V_{ij} = -\frac{1}{2\pi\epsilon_0} \mathrm{Re}\left\{ (z_j - z)\,[\ln(z_j - z) - 1] \Big|_{p_1}^{p_2} \right\} \qquad z_j = p_j + js_j \tag{2.15}$$

where p_j and s_j are coordinates of the jth matching point in the local pOs-coordinate system, such that the ith segment is placed along the p-coordinate axis, from p_1 to p_2 [7]. The approximate surface charge distribution is a discontinuous function along the circumference of a conductor. The potential due to this charge distribution has singularities at segment ends. Hence, it seems reasonable that the matching points are placed in the middle of the segments, as shown by dots in Figure 2.2.

Figure 2.3 shows the charge distribution for a line with square conductors shown in Figure 2.1 ($d/a = 1.368469$), for various numbers of unknowns per square side ($n = 5, 10, 20$). By increasing n, the charge distribution does not change much, except in the vicinity of corners. It can be assumed that a better approximation for the charge distribution is obtained if a nonuniform segmentation is applied.

2.3.3 Data Postprocessing

Very often we are not interested in the unknown function f itself, but in some quantity $C(f)$ that can be calculated from f. Once the unknown coefficients in (2.9) are determined, we can easily calculate the approximate function f_a at any point in space where it is defined. Then we can determine an approximate value of $C(f)$, as

$$C(f) \approx C(f_a) \tag{2.16}$$

Such calculations are referred to as data postprocessing.

Often $C(f)$ can be obtained by various postprocessing procedures, which give the required quantity with a smaller or larger error. Hence, in some cases

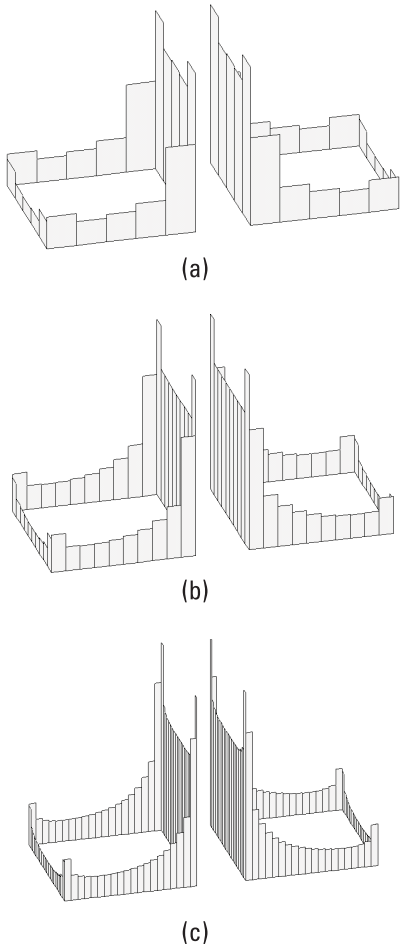

Figure 2.3 Charge distribution at the cross section of the line shown in Figure 2.1 (d/a = 1.368469). Results are obtained by the point-matching method with (a) n = 5, (b) n = 10, and (c) n = 20 unknowns per square side.

the accuracy and efficiency of the method for solving the linear operator equation can be improved by a proper choice of the data postprocessing procedure.

Example: The capacitance of a two-conductor line can be determined using the basic definition (in terms of the charges and voltage) or energy definitions; that is,

$$C' = \frac{Q'}{V_1 - V_2} \qquad C' = \frac{2W'}{(V_1 - V_2)^2} \qquad C' = \frac{Q'^2}{2W'} \qquad (2.17\text{a-c})$$

where Q' is per-unit-length charge of the first conductor, V_1 and V_2 are potentials of the conductors, and W' is the per-unit-length energy of the line. Once the charge distribution is known, Q' and W' are determined as

$$Q'(\rho_s) = \oint_{C_1} \rho_s \, dl \qquad W'(\rho_s) = \frac{1}{2} \oint_{C} \rho_s V(\rho_s) \, dl \tag{2.18a,b}$$

The same formulas can be used if we know an approximate solution, ρ_{sa}, instead of the exact solution, ρ_s. In that case we use (2.18) to evaluate the approximate values of the per-unit-length charge of the first conductor and the approximate per-unit-length energy of the line [i.e., $Q'_a = Q'(\rho_{sa})$ and $W'_a = W'(\rho_{sa})$]. However, the problem in application of definitions (2.17a,b) to approximate solutions is that the approximate potentials, V_{1a} and V_{2a}, are not constant along the respective conductor circumferences. The question is which values of V_{1a} and V_{2a} should be used to evaluate the capacitance. One possibility is to use the given values, V_1 and V_2. Note that V_{1a} and V_{2a} are equal to the given values at the matching points. Another possibility is to use the mean values of the approximate potentials, $\overline{V_{1a}}$ and $\overline{V_{2a}}$, along the respective conductor circumferences,

$$\overline{V_{1a}} = \frac{1}{L_1} \oint_{C_1} V_{1a} \, dl, \qquad \overline{V_{2a}} = \frac{1}{L_2} \oint_{C_2} V_{2a} \, dl \tag{2.19}$$

where L_1 and L_2 are the respective lengths of the contours C_1 and C_2. Definitions that will be used for the numerical example of the line shown in Figure 2.1 are summarized in Table 2.2. For the exact solution, all definitions give the same result. However, for approximate solutions these results differ, and the goal is to choose the definition which provides the most accurate results.

Table 2.2
Summary of Definitions of Per-Unit-Length Capacitance

	C'_1	C'_2	C'_3	C'_4
Formulation	$\dfrac{Q'_a}{V_1 - V_2}$	$\dfrac{Q'_a}{\overline{V_{1a}} - \overline{V_{2a}}}$	$\dfrac{2W'_a}{(V_1 - V_2)^2}$	$\dfrac{(Q'_a)^2}{2W'_a}$

$Q'_a = Q'(\rho_{sa})$: approximate per-unit-length charge of the first conductor.
$\overline{V_{1a}}$, $\overline{V_{2a}}$: mean values of approximate potentials along conductor circumferences.
$W'_a = W'(\rho_{sa})$: approximate per-unit-length energy.

Figure 2.4 shows the relative error of the per-unit-length capacitance of the line with square conductors for $d/a = 1.368469$, versus the number of unknowns per square side, n. The results are obtained using the four definitions of the capacitance given in Table 2.2. The relative error is calculated as

$$\delta_i[\%] = 100\frac{|C_i' - C'|}{C'} \qquad i = 1, 2, 3, 4 \tag{2.20}$$

where C' corresponds to the characteristic impedance, estimated to be $Z = (75 \pm 0.0001)\Omega$ (see Section 2.3.14). To obtain a relative error less than 1%, the numbers of unknowns needed per square side are 8, 6, 10, and 4 for the four definitions of the capacitance, respectively. The best results are obtained by the definition C_4'.

2.3.4 Estimation of the Solution Quality

After obtaining the approximate solution of the linear operator equation, f_a, we are interested in the error with respect to the exact solution, f; that is,

$$E = f_a - f \tag{2.21}$$

Usually, we do not know the exact solution, and this error can only be estimated. If we are not interested in the exact solution f, but in some quantity

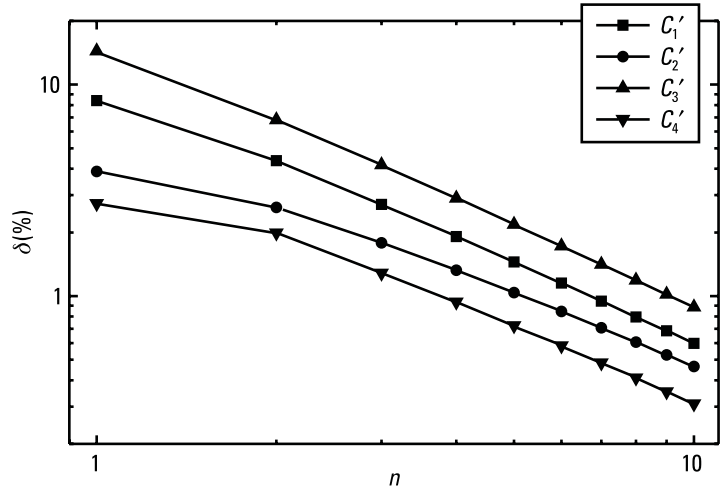

Figure 2.4 Relative error of per-unit-length capacitance of the line shown in Figure 2.1 ($d/a = 1.368469$) versus the number of unknowns per square side, n, obtained by using the four definitions from Table 2.1.

$C(f)$ obtained by data postprocessing, it is desirable that we can estimate the error of $C(f_a)$. In general, such estimations are bounds that can be written in the form

$$[C(f_a)]_{\min} < C(f) < [C(f_a)]_{\max} \tag{2.22}$$

where $[C(f_a)]_{\min}$ and $[C(f_a)]_{\max}$ are the estimated minimum and maximum values of $C(f)$ based on the approximate solution f_a.

Unfortunately, there is no technique that can enable direct estimation of the solution quality in general [14]. Even if the bounds can be evaluated, their calculation often requires much more effort than it takes to obtain the approximate solution. These bounds are usually evaluated with much greater error than the approximate solution. Hence, the practical value of bound-estimation techniques is in making benchmarks, which are useful in determining the accuracy and efficiency of various methods.

Example: The lower and the upper bounds of the per-unit-length capacitance can be directly estimated [15]. Consequently, the characteristic impedance of the two-conductor line is estimated as

$$Z_{\min} < Z < Z_{\max} \qquad Z_{\min} = \frac{V_{1\min} - V_{2\max}}{c_0 Q'_a} \qquad Z_{\max} = \frac{V_{1\max} - V_{2\min}}{c_0 Q'_a} \tag{2.23}$$

where c_0 is the light velocity in a vacuum, and $V_{1\min}$, $V_{1\max}$, $V_{2\min}$, and $V_{2\max}$ represent the minimum and maximum approximate potential distributions along the first and the second conductor, respectively. Let us apply these criteria to the example of Section 2.3.3 with $d/a = 1.368469$, analyzed by the point-matching method. Figure 2.5 shows the lower and the upper bounds of characteristic impedance, $Z_{\min}$ and $Z_{\max}$, together with the approximate value Z_4, versus the number of unknowns per square side, n. Z_4 is evaluated as $(c_0 C'_4)^{-1}$ where C'_4 is given in Table 2.2. The errors of $Z_{\min}$ and $Z_{\max}$ are much greater than the error of Z_4.

Quality of the solution can be indirectly estimated by inspecting the error with which the approximate solution satisfies the linear operator equation. This error, called the residuum, is defined as

$$R = Lf_a - g \tag{2.24}$$

The point-matching method forces the residuum to be zero at matching points. Between the matching points the residuum can be arbitrary. A similar

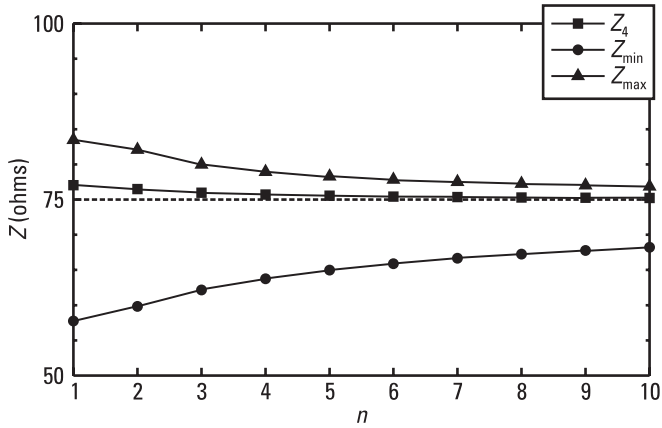

Figure 2.5 The lower and the upper bounds of characteristic impedance of the line shown in Figure 2.1 (d/a = 1.368469), Z_{min} and Z_{max}, together with the approximate value Z_4, versus the number of unknowns per square side, n. Data postprocessing is based on results obtained by the point-matching method.

behavior is noticed if other numerical methods are applied to solve (2.11). Usually, the residuum is larger in some parts of the analyzed structure than in the rest of the structure. This is an indication that the approximation of the function in these parts is poorer than in the rest of the structure.

Example: Consider a transmission line with square conductors for a = 1 mm, d/a = 1.368469, $V_1 = -V_2$ = 1V. Figure 2.6 shows the approximate potential distribution, V_a, versus the local coordinate l along the circumference of the first conductor. The results are obtained by using various numbers of unknowns per square side (n = 5, 10, 20). For n = 5 the linear operator equation is satisfied very well everywhere, except in the vicinity of the corners. By doubling the number of unknowns, performed by bisection of each segment into two equal subsegments, the residuum of the linear operator equation, $R = V(l) - V_1$, generally decreases. In particular, the corner peaks are almost halved. However, to decrease these peaks, it is not necessary to bisect all segments, but only the segments in the vicinity of the corners.

An integral measure of the overall solution quality can be defined by the root mean square (rms) value of the residuum,

$$R_{\text{rms}} = \sqrt{\frac{1}{\Omega}\int_{\Omega} R^2 d\Omega} \tag{2.25}$$

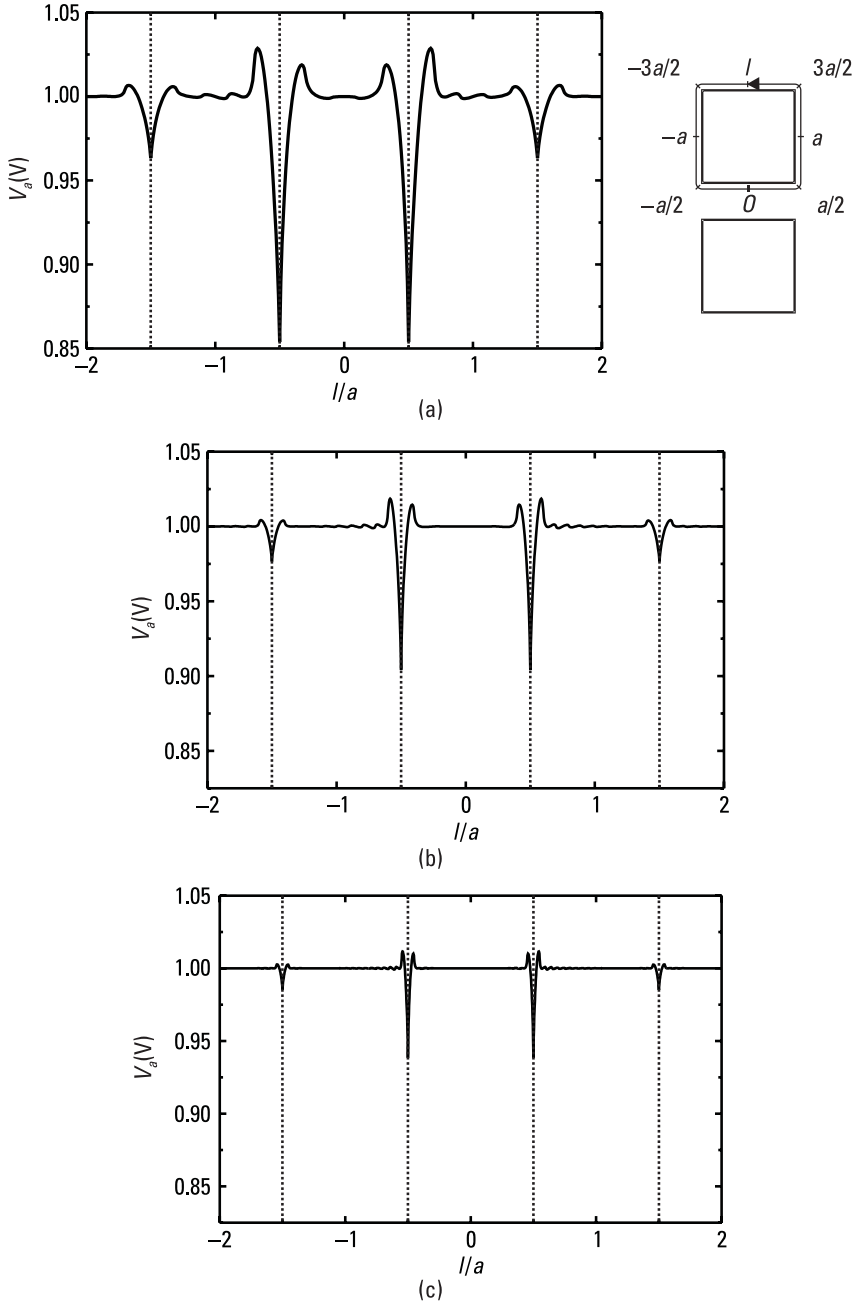

Figure 2.6 Normalized potential distribution along the circumference of one square conductor for various numbers of unknowns per square side: (a) $n = 5$, (b) $n = 10$, and (c) $n = 20$.

where Ω is the spatial region in which the linear operator equation is defined. Most often, by decreasing the R_{rms} the solution quality increases, and by increasing the order of approximation, N, the R_{rms} decreases. However, in general, there is no guarantee that an increase of N and a decrease of the R_{rms} will always result in an increase of the solution quality. (For an example, see Figure 8.7.) To find the correlation between N, R_{rms}, and the solution quality for a specific class of problems and for a numerical method applied, one should perform a lot of numerical experiments. Based on this experience it is possible to establish criteria under which the results are obtained with a prescribed accuracy.

Example: Figure 2.7 shows the relative error of the capacitance per unit length of the line from the previous example, δ_4, versus the rms value of the residuum, R_{rms}. Results are obtained by the point-matching method and δ_4 is calculated according to (2.20). It is seen that δ_4 is approximately a linear function of R_{rms} (i.e., δ_4 [%] = 35 R_{rms}). It can be shown that the same formula is approximately valid for various ratios d/a, as well as for other shapes of the conductor cross sections. Hence, this formula can be used as an experience-based criterion for error estimation for two-conductor lines.

2.3.5 Least-Squares Method

As the approximate solution, f_a, approaches the exact solution, f, the residuum of a linear operator equation, R, defined by (2.24), tends toward zero. It can

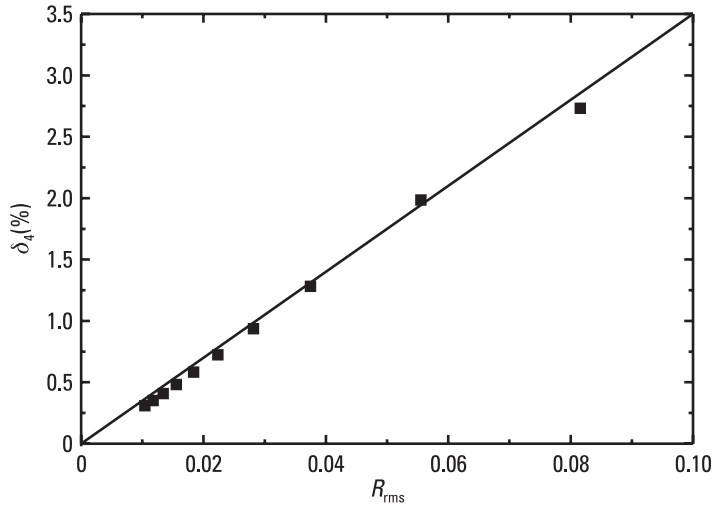

Figure 2.7 Relative error of the characteristic impedance, δ_4, versus the rms value of the residuum, R_{rms}, for the same example as in Figure 2.6.

be expected that by minimizing the residuum in some sense, the error between the approximate and exact solution is also minimized. For example, the residuum can be minimized in the mean-square sense. In that case, we start from the functional[7] written in alternative forms as

$$F = \int_{\Omega} (Lf_a - g)^2 \, d\Omega \qquad F = \int_{\Omega} (Lf_a)^2 \, d\Omega - 2\int_{\Omega} Lf_a g \, d\Omega + \int_{\Omega} g^2 \, d\Omega \tag{2.26a,b}$$

where Ω is the spatial region in which the residuum is defined. After replacing (2.9) into (2.26a), the functional is expressed in terms of unknown coefficients as

$$F = \int_{\Omega} \left(\sum_{i=1}^{N} a_i Lf_i - g \right)^2 d\Omega \tag{2.27}$$

To minimize the functional, its partial derivatives with respect to the unknown coefficients are forced to be zero:

$$\frac{\partial F}{\partial a_j} = 2\int_{\Omega} \left(\sum_{i=1}^{N} a_i Lf_i - g \right) Lf_j \, d\Omega = 0 \qquad j = 1, \ldots, N \tag{2.28}$$

Simple rearrangement results in a system of linear equations:

$$\sum_{i=1}^{N} a_i \int_{\Omega} Lf_j Lf_i \, d\Omega = \int_{\Omega} Lf_j g \, d\Omega \qquad j = 1, \ldots, N \tag{2.29}$$

By solving this system, the unknown coefficients are obtained. This technique is called the *least-squares method*.

Once the unknown coefficients are evaluated, the minimum value of the functional is obtained from (2.27). To obtain an alternative expression for the minimum, multiply (2.29) by the corresponding coefficients (i.e., the jth equation by a_j, $j = 1, \ldots, N$), and sum all equations thus obtained. By using (2.9), the sum is written as

7. A functional is a mapping of a function to a scalar.

$$\int_{\Omega} (Lf_a)^2 \, d\Omega = \int_{\Omega} gLf_a \, d\Omega \tag{2.30}$$

Finally, after replacing (2.30) into (2.26b), the minimum value of the functional is obtained in the form

$$F_{\min} = \int_{\Omega} g^2 \, dv - \int_{\Omega} (Lf_a)^2 \, d\Omega \tag{2.31}$$

Note that Lf_a represents an approximation of the excitation g. Since $F_{\min} > 0$, the mean square value of the approximate excitation is always less than the mean square value of the exact excitation.

Example: Consider an arbitrary symmetrical two-conductor line as in Section 2.3.2. Equations for the least-squares method are obtained by multiplying the left and right side of (2.13a) by the potentials $V_j = V(\rho_{sj})$, $j = 1, \ldots, N$, and integrating along the contour C, resulting in

$$\sum_{i=1}^{N} a_i \int_C V_j V_i \, dl = \int_C V_j V(\mathbf{r}) \, dl \qquad j = 1, \ldots, N \tag{2.32}$$

Once the unknown coefficients are obtained by solving the above matrix equation, the per-unit-length capacitance can be determined according to the definitions in Table 2.2.

Figure 2.8 shows the relative error of the per-unit-length capacitance of a line with square conductors ($d/a = 1.368469$) versus the number of unknowns per square side, n. The relative error is calculated according to (2.20). All four capacitance definitions give very similar results, which means that the accuracy of results does not depend much on data postprocessing. This is in contrast to the point-matching method (see Figure 2.4).

2.3.6 Inner Product

In the previous section we introduced the least-squares method for real scalar linear operator equations. Similarly, we can establish the least-squares method for complex scalar, real vector, and complex vector linear operator equations. In these cases, $(Lf_a - g)^2 = R^2$ in the functional (2.26) is replaced by R^*R, $\mathbf{R} \cdot \mathbf{R}$, and $\mathbf{R}^* \cdot \mathbf{R}$, respectively, where * denotes complex conjugate and $\mathbf{R}$ denotes the residuum of the vector linear operator equation. Generally, the

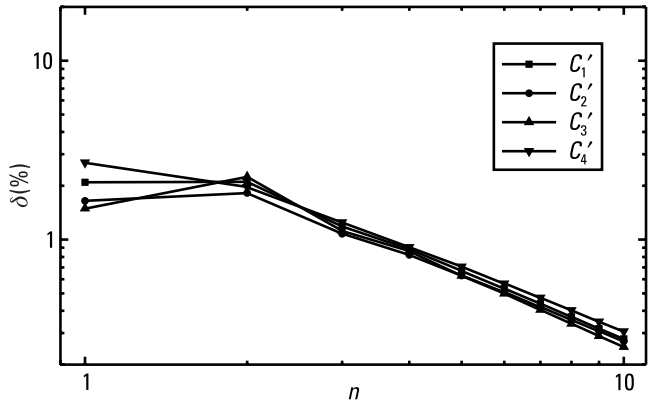

Figure 2.8 Relative error of the per-unit-length capacitance of the line shown in Figure 2.1 (d/a = 1.368469) versus the number of unknowns per square side, n, obtained by using four expressions for per-unit-length capacitance given in Table 2.2.

functional can be defined using the concepts of inner product, natural norm, or distance.

An inner product of two functions p and q, usually denoted by $\langle p, q \rangle$, is a scalar obtained after a set of manipulations on functions p and q. This scalar satisfies the following properties:

$$\langle p, q \rangle = \langle q, p \rangle^* \qquad \langle \alpha p + \beta q, s \rangle = \alpha \langle p, s \rangle + \beta \langle q, s \rangle \qquad \langle p, p \rangle > 0 \tag{2.33}$$

In particular, $\langle p, p \rangle = 0$ if p is a zero function. The natural norm of a function p is defined as

$$\| p \| = \sqrt{\langle p, p \rangle} \tag{2.34}$$

It corresponds to the Euclidean vector concept of length. The metric d of two functions p and q is defined by

$$d(p, q) = \| p - q \| \tag{2.35}$$

and corresponds to the Euclidean vector concepts of distance between two points. (Rigorous mathematical foundation of inner product spaces and their properties can be found in [9].)

Most often in field applications, the inner product is adopted in the form

$$\langle p, q \rangle = \int_{\Omega} p^* \oplus q \, d\Omega \tag{2.36}$$

where the operation $\oplus$ is multiplication (dot product) when p and q are scalar (vector) functions.

Example: In the general form, the least-squares method is based on minimizing the functional that can be represented as the inner product of R with R, or as squared natural norm of R, or as squared distance between Lf_a and g; that is, the functional (2.26) can be written in one of the following three forms:

$$F = \langle R, R \rangle \qquad F = \| R \|^2 \qquad F = d(Lf_a, g)^2 \tag{2.37}$$

In that case, the system of linear equations (2.29) is obtained in the general form

$$\sum_{k=i}^{N} a_i \langle Lf_j, Lf_i \rangle = \langle Lf_j, g \rangle \qquad j = 1, \ldots, N \tag{2.38}$$

and the corresponding minimum value of function (2.29) is obtained in the general alternative forms as

$$F_{\min} = \langle g, g \rangle - \langle Lf_a, Lf_a \rangle \qquad F_{\min} = \| g \|^2 - \| Lf_a \|^2 \tag{2.39a,b}$$

2.3.7 Rayleigh-Ritz Method

The least-squares method is a particular case of the Rayleigh-Ritz variational method [16]. The Rayleigh-Ritz variational method enables solution of an arbitrary deterministic equation. Using this equation, we construct a variational formula F in the form of a functional, which relates the approximate solution f_a and the exact solution f:

$$F = F(f, f_a) \tag{2.40}$$

To be a variational formula, the functional must be stationary with respect to the exact solution. This means that an error of the approximate solution causes a quadratic or higher order variation of the functional.

To examine the stationary character of the functional, the approximate solution is written in the form

$$f_a = f + \theta h \tag{2.41}$$

where θ is a parameter and h is an arbitrary function. The first and second variations of a functional F are

$$\delta F = \left.\frac{dF(f+\theta h)}{d\theta}\right|_{\theta=0} \delta\theta \qquad \delta^2 F = \left.\frac{dF^2(f+\theta h)}{d\theta^2}\right|_{\theta=0} (\delta\theta)^2 \tag{2.42a,b}$$

The stationary character of the functional is expressed by the condition that its first variation is equal to zero. If this condition is satisfied, the second variation determines if the functional has a local extremum around $f_a = f$, or whether the equivalent of saddle point exists.

Once the stationary character of the functional is proven, it can be used as the variational formula for approximate determination of the unknown function (i.e., determination of the approximate solution, f_a). The approximate solution, f_a, is represented as a finite series of known functions multiplied by unknown coefficients according to (2.9). The approximate stationary point of the functional is determined by forcing partial derivatives of the functional with respect to the unknown coefficients to be zero; that is,

$$\frac{\partial F}{\partial a_j} = 0 \qquad j = 1, \ldots, N \tag{2.43}$$

Thus a system of equations is obtained. By solving the system, the unknown coefficients are determined. Solving of the system is particularly easy if the system is linear.

In general, there is no guarantee that increasing the order of approximation N will always result in a decrease of the error of the approximate solution. In particular, consider the case when the second variation (2.42b) is greater than zero, which means that the functional has a minimum at the stationary point. If we increase the order of approximation, by adding new basis functions to the existing set, the approximate minimum value can be decreased, tending to the exact value. (The same situation occurs if each pulse in a PWC approximation is subdivided into two independent pulses.) Most often, when the approximate minimum of the functional approaches the exact one, we improve the overall accuracy of the solution (e.g., see Figure 2.7) Note that the possibility of estimating the solution quality does not exist if the stationary point is a saddle point.

As mentioned, in general, the starting equation need not be linear. Even if the starting equation is linear, the functional need not result in a system of

linear equations. In this book, we always start from a linear operator equation in the form (2.5), and we are interested in functionals that transform it into a system of linear equations. One such functional is given by (2.26). In addition, it is desirable that the second variation of the functional be greater than zero.

Example: In the case of the linear operator equation (2.5), we can define many variational formulas. Three typical formulas, which result in a system of linear equations after the application of the Rayleigh-Ritz method, are

$$F_{RR} = \langle R, R \rangle \qquad F_{ER} = \langle E, R \rangle \qquad F_{EE} = \langle E, E \rangle \tag{2.44a-c}$$

where R is the residuum of the linear operator equation given by (2.24) and E is the error of the approximate solution given by (2.21). To show that these functionals are stationary, represent the approximate function in the form (2.41). The error and the residuum are obtained as

$$E(f + \theta h) = \theta h \qquad R(f + \theta h) = \theta Lh \tag{2.45}$$

After replacing (2.45) into (2.44) and applying (2.33), the functionals are obtained in the form

$$F_{RR}(f + \theta h) = \theta^2 \langle Lh, Lh \rangle \qquad F_{ER}(f + \theta h) = \theta^2 \langle h, Lh \rangle \tag{2.46a,b}$$

$$F_{EE}(f + \theta h) = \theta^2 \langle h, h \rangle \tag{2.46c}$$

According to (2.42a), the first variations of these functionals are zero. It is also seen that the second variations of F_{RR} and F_{EE} are greater than zero. However, F_{ER} and F_{EE} cannot be directly applied, because E is not known.

2.3.8 Galerkin Method

Consider under which condition and how the functional (2.44b) can be used to determine the unknown coefficients in the approximate solution (2.9). After replacing (2.21) and (2.24) into (2.44b), the functional can be written in alternative forms as

$$F = \langle f_a - f, Lf_a - Lf \rangle \qquad F = \langle f_a, Lf_a \rangle - \langle f_a, Lf \rangle - \langle f, Lf_a \rangle + \langle f, Lf \rangle \tag{2.47a,b}$$

The term $\langle f, Lf \rangle$ does not affect the stationary properties of the functional and can be omitted. [Namely, after replacing (2.9) into (2.47b) and differentia-

tion with respect to a_j, the term $\langle f, Lf \rangle$ is lost.] Since $Lf = g$ is known, the only problem in the application of (2.47b) is how to calculate the term $\langle f, Lf_a \rangle$, because f is not known. This problem is avoided if the operator L is self-adjoint.

The operator L^a that is adjoint to the operator L is defined by

$$\langle Lp, q \rangle = \langle p, L^a q \rangle \tag{2.48}$$

where p and q are arbitrary functions in the domain and range of L, respectively. The operator L is said to be self-adjoint if it is equal to its adjoint operator L^a; that is,

$$L = L^a \tag{2.49}$$

If (2.49) is fulfilled, the second and third terms on the right side of (2.47b) are equal, and the functional is reduced to

$$F' = \langle f_a, Lf_a \rangle - 2\langle f_a, Lf \rangle \tag{2.50}$$

After replacing (2.9) in (2.50) and differentiation with respect to a_j, $j = 1, \ldots, N$, we obtain the system of linear equations

$$\sum_{k=i}^{N} a_i \langle f_j, Lf_i \rangle = \langle f_j, Lf \rangle \qquad j = 1, \ldots, N \tag{2.51}$$

The unknown coefficients of the approximate solution are determined by solving this system.

In addition, let the linear operator L be positive definite. The operator is positive definite if

$$\langle p, Lp \rangle > 0 \tag{2.52}$$

for any p in the domain of operator L. In that case the second variation of the functional, which is obtained by applying (2.42b) to (2.46b), is always greater than zero. The stationary point of the functional is a minimum. The exact minimum occurs for $f_a = f$. According to (2.41) and (2.46b), this minimum is zero. When f_a is approximated by (2.9), the solution of the system (2.51) results in the approximate minimum of the functional (2.47), which is always greater than zero.

To obtain alternative expressions for the approximate minimum, multiply (2.51) by the corresponding unknown coefficients (i.e., the jth equation by a_j, $j = 1, \ldots, N$), and sum the resulting equations. By using (2.9), the sum can be written as

$$\langle f_a, Lf_a \rangle = \langle f_a, Lf \rangle \tag{2.53}$$

After replacing (2.53) into (2.47b) and (2.50), the expressions for the approximate minimum are obtain as

$$F_{\min} = \langle f, Lf \rangle - \langle f_a, Lf_a \rangle \qquad F'_{\min} = -\langle f_a, Lf_a \rangle \tag{2.54a,b}$$

Since the approximate minimum $F_{\min}$ is always greater than zero, we conclude that $\langle f_a, Lf_a \rangle$ is always less than $\langle f, Lf \rangle$. In particular, when the order of approximation is increased by adding new basis functions to the existing set, the approximate functional decreases monotonically; that is, the approximate inner product $\langle f_a, Lf_a \rangle$ monotonically approaches (from the lower side) the exact value $\langle f, Lf \rangle$. In general, it is shown that the approximate solution converges to the exact one with increasing the order of approximation [16].

The system (2.51) can be obtained in an alternative way if we take the inner product of f_j, $j = 1, \ldots, N$, and the approximate equation (2.11). This method, which transforms the linear operator equation into a system of linear equations, is known as the Galerkin method. This method can be applied whether the linear operator is self-adjoint or not. When the linear operator is self-adjoint, the Galerkin method is equivalent to the Rayleigh Ritz method applied to the functional F_{ER}. If, in addition, the operator is positive definite, the convergence of the Galerkin method is established. Some operators arising in electromagnetics belong to this class [e.g., electrostatic integral operator defined by (2.8b)]. However, this does not mean that the Galerkin method gives more valuable results than other test procedures. On the other hand, the integral operators L and K occurring in integral equations in the frequency domain (see Chapter 4) do not belong to this class. Nevertheless, the Galerkin method is shown to be the most promising when compared with other test procedures (see Chapter 8). Generally, there is still no theory that can predict which test procedure is the best one. The answer to this question is usually obtained by numerical experiments.

Examples: Consider an arbitrary symmetrical two-conductor line such that $Q'_1 = -Q'_2$ is automatically satisfied and $V_1 = -V_2$ is given. Let the inner product be defined by (2.36), where the operation $\oplus$ is a simple multiplication, and p and q are scalar functions. The Galerkin equations are obtained by taking

the inner product of the basis functions ρ_{sj}, $j = 1, \ldots, N$, and (2.13a), resulting in

$$\sum_{i=1}^{N} a_i \int_{C_j} \rho_{sj} V_i \, dl = \int_{C_j} \rho_{sj} V(\mathbf{r}) \, dl \qquad j = 1, \ldots, N \tag{2.55}$$

where C_j is segment along which the jth basis function is defined. To establish the convergence of the solution, examine the character of the linear operator V.

The expression for the stored per-unit-length energy (2.18b) represents the inner product of ρ_s, and $V(\rho_s)$. Since the energy stored in an electrostatic field is always positive, using (2.52) we conclude that the linear operator $V(\rho_s)$ is positive definite. On the other hand, the electrostatic reciprocity theorem [15] can be written for the two-conductor line as

$$\int_C V(\rho_{s1}) \rho_{s2} \, dl = \int_C \rho_{s1} V(\rho_{s2}) \, dl \tag{2.56}$$

where ρ_{s1} and ρ_{s2} are arbitrary distributions of the surface charges along contour C representing the cross section of the line. Since (2.56) satisfies (2.48) and (2.49), the operator $V(\rho_s)$ is self-adjoint. Hence, application of the Galerkin method to (2.8a) is the same as minimization of the functional

$$F = \int_C (\rho_{sa} - \rho_s)(V_a - V) \, dl \tag{2.57}$$

where ρ_s is an unknown charge distribution to be determined, ρ_{sa} is its approximate value given by (2.10), and V_a and V are potential distributions due to ρ_s and ρ_{sa}, respectively. By minimizing the functional (2.57), we actually minimize the energy of the error charge distribution.

According to (2.18b) and (2.54a) the approximate minimum of functional (2.57) is obtained as

$$F_{\min} = \int_C \rho_s V \, dl - \int_C \rho_{sa} V_a \, dl = 2(W' - W_a') \tag{2.58}$$

where W_a' is the per-unit-length energy stored due to the approximate charge distribution ρ_{sa} [i.e. $W_a' = W'(\rho_{sa})$]. Hence, it is concluded that the energy

of an approximate solution is always smaller than the energy of the exact solution. Considering the C_3' definition of the capacitance (see Table 2.2), minimization of the functional leads to maximization of the approximate capacitance in such a manner that it approaches the exact capacitance from the lower side; that is, the Galerkin method directly minimizes the error of capacitance. Owing to this property, it is possible to provide monotonous convergence of the results for the capacitance (e.g., by subdividing segments).

When combined with the PWC approximation for charges, the Galerkin method has an additional advantage with respect to the point-matching and the least-squares method: All definitions of the capacitance summarized in Table 2.2 give the same value. This can be shown in the following way. According to (2.10), ρ_{sj} is a constant and can be omitted from (2.55), so that the integrals of the approximate and exact potentials along the jth segment are equal. Consequently, the mean values of the approximate potentials along the first and the second conductor are equal to the exact potentials of these conductors, respectively; that is, $\overline{V_{1a}} = V_1$ and $\overline{V_{2a}} = V_2$, resulting in $C_1' = C_2'$. Similarly, starting from the expression for the stored energy (2.18b), it can be shown that $W'(\rho_{sa}) = \frac{1}{2} Q_{1a}(\overline{V_{1a}} - \overline{V_{2a}})$, resulting in $C_1' = C_3' = C_4'$. Since all definitions give the same value, one can choose the definition that is easiest for postprocessing (i.e., the C_1' definition).

For the particular example of a line with square conductors (d/a = 1.368469), Figure 2.9 shows the relative error of the per-unit-length capacitance, δ_4, versus the number of unknowns per square side, n, obtained by three methods: point-matching, least squares, and Galerkin. δ_4 is calculated according to (2.20). The Galerkin method gives a slightly lower relative error of the

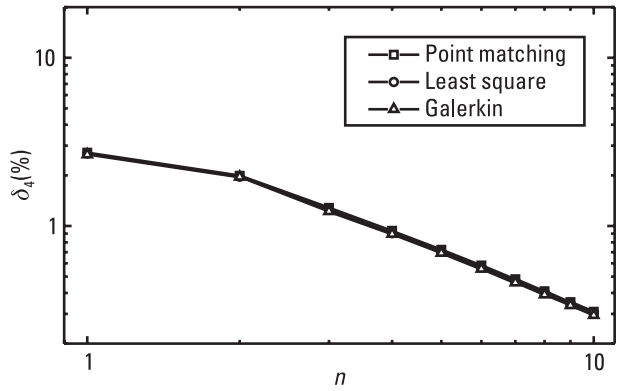

Figure 2.9 Relative error of the per-unit-length capacitance of the line shown in Figure 2.1 (d/a = 1.368469) versus the number of unknowns per square side, n, obtained by various testing procedures.

characteristic impedance, as expected. However, in the case considered, the difference in accuracy among these three methods is practically negligible.

2.3.9 MoM

In Sections 2.3.2, 2.3.5, and 2.3.8, three methods for solving a linear operator equation are reported: the point-matching method, the least-squares method, and the Galerkin method. These methods each transform the linear operator equation into a system of linear equations. In the Galerkin method, the system of linear equations (2.51) is formally obtained if we take the inner product of weighting functions $w_j = f_j$, $j = 1, \ldots, N$, and the approximate equation (2.11). Similarly we obtain the system (2.38) for the least-squares method, and the system (2.12) for the point-matching method, except that the weighting functions are of the form $w_j = Lf_j$ and $w_j = \delta(\mathbf{r} - \mathbf{r}_j)$ (Dirac delta functions), respectively. Many other sets of mutually independent weighting functions can be applied, resulting in various systems of linear equations. All of them can be written in common form as

$$\sum_{i=1}^{N} a_i \langle w_j, Lf_i \rangle = \langle w_j, g \rangle, \qquad j = 1, \ldots, N \tag{2.59}$$

This way of transforming the approximate equation into the system of linear equations is known as testing (test procedure). Hence, the weighting functions w_j, $j = 1, \ldots, N$ are also called test functions.

This general method for solving linear operator equations is known in electromagnetics as the MoM [6]. It consists of the following steps:

1. We start from the linear operator equation (2.5).
2. The unknown function is approximated in the form of a finite series of known basis functions multiplied by unknown coefficient (2.9).
3. The approximation (2.9) is replaced into (2.5), resulting in the approximate equation (2.11).
4. Equation (2.11) is transformed into the system of linear equations (2.59) by using the test procedure.
5. The unknown coefficients in (2.9) are determined by solving the system of linear equations (2.59).

The MoM system of linear equations (2.59) can be shortly written in matrix form as

$$\mathbf{FA} = \mathbf{G} \qquad \mathbf{F} = [\langle w_j, Lf_i \rangle]_{N \times N} \qquad \mathbf{A} = [a_i]_{N \times 1} \qquad \mathbf{G} = [\langle w_j, g \rangle]_{N \times 1} \tag{2.60a-d}$$

where the subscript j stands for rows and the subscript i stands for columns. Weighting functions, matrix elements, and free-term column elements of the three basic variants of the MoM are summarized in Table 2.3.

The name *MoM* is derived from the terminology that $\int x^n f(x)\,dx$ is the nth moment of f [1, 2]. In the MoM, however, the nth moment is evaluated more generally; that is, the function is multiplied by the nth weighting function instead by x^n. Since the unknown functions are determined by equating the moments of Lf_a and g, the method is called the MoM. On the other hand, it can be interpreted that the system of linear equations (2.59) is obtained by forcing the inner product of the residuum (2.24) and the weighting functions to be zero. Hence, the MoM is also named the method of weighted residuals [17]. The third name used for the MoM is the method of projections. It comes from the concept of orthogonal projections onto subspaces in the theory of linear spaces. In some cases, such an interpretation of the MoM enables the solution error to be estimated and the convergence of results to be established [18].

2.3.10 Weighted Point-Matching Method

Suppose that the linear operator equation solved by the MoM is a complex scalar equation. Using inner product in the form (2.36), the matrix elements and free term elements can be written in common form

$$\langle w_j, p \rangle = \int_{\Omega} w_j(\mathbf{r})^* \, p(\mathbf{r})\, d\Omega \qquad j = 1, \ldots, N \tag{2.61}$$

where $\mathbf{r}$ is the field point in domain Ω within which the weighting functions w_j are defined, $p = Lf_i$ for matrix elements, and $p = g$ for free-term elements. In general, the above integral is evaluated numerically as

$$\langle w_j, p \rangle = \sum_{k=1}^{M} b_k w_j(\mathbf{r}_k)^* \, p(\mathbf{r}_k) \qquad j = 1, \ldots, N \tag{2.62}$$

Table 2.3
Weighting Functions, Matrix Elements, and Free-Term Elements for Three Basic Forms of the MoM

	w_j	$\langle w_j, Lf_i \rangle$	$\langle w_j, g \rangle$
Point-matching method	$\delta(\mathbf{r} - \mathbf{r}_j)$	Lf_i	$g(\mathbf{r}_j)$
Least-squares method	Lf_j	$\langle Lf_j, Lf_i \rangle$	$\langle Lf_j, g \rangle$
Galerkin method	f_j	$\langle f_j, Lf_i \rangle$	$\langle f_j, g \rangle$

where $\mathbf{r}_k$ are the integration points and b_k are the weighting coefficients of the formula applied for numerical integration. Consequently, the matrices $\mathbf{F}$ and $\mathbf{G}$ of the MoM matrix equation (2.60a) can be expressed as

$$\mathbf{F} = \mathbf{W} * \mathbf{B}\mathbf{F}_{\mathrm{PM}} \qquad \mathbf{G} = \mathbf{W} * \mathbf{B}\mathbf{G}_{\mathrm{PM}} \tag{2.63}$$

$$\mathbf{W} = [w_j(\mathbf{r}_k)]_{N\times M} \qquad \mathbf{B} = diag(b_1, \ldots, b_M) \tag{2.64}$$
$$\mathbf{F}_{\mathrm{PM}} = [Lf_i(\mathbf{r}_k)]_{M\times N} \qquad \mathbf{G}_{\mathrm{PM}} = [g(\mathbf{r}_k)]_{M\times 1}$$

The subscript i is the column number in the matrices $\mathbf{F}$ and $\mathbf{F}_{\mathrm{PM}}$, the subscript j is the row number in the matrices $\mathbf{F}$ and $\mathbf{W}$, and the subscript k is the column number in the matrix $\mathbf{W}$ and the row number in the matrices $\mathbf{F}_{\mathrm{PM}}$ and $\mathbf{G}_{\mathrm{PM}}$. The matrices $\mathbf{F}_{\mathrm{PM}}$ and $\mathbf{G}_{\mathrm{PM}}$ are labeled with the subscript PM because they have the same form as in the point-matching procedure. Finally, the MoM matrix equation (2.60a) can be written as

$$(\mathbf{W} * \mathbf{B})\,\mathbf{F}_{\mathrm{PM}}\mathbf{A} = (\mathbf{W} * \mathbf{B})\,\mathbf{G}_{\mathrm{PM}} \tag{2.65}$$

The same matrix equation is obtained if we start from the system of linear equations

$$\mathbf{F}_{\mathrm{PM}}\mathbf{A} = \mathbf{G}_{\mathrm{PM}} \tag{2.66}$$

which results from the point-matching method, except that the number of matching points M, in general, is different from the number of unknowns N. (The matching points coincide with the numerical integration points.) After multiplying the left and right sides of the matrix equation (2.66) by the weighting matrix $\mathbf{W} * \mathbf{B}$, we obtain the system (2.65). The point-matching method combined with the multiplication with the weighting matrix is called the weighted point-matching method [19]. Consequently, the MoM based on the inner product in the form (2.36) can be considered as a weighted point-matching method, thus enabling easy implementation of various test procedures (once we develop the point-matching method).

Such interpretation also gives some useful information about testing. For example, it is obvious that for any test procedure the number of integration points M in the domain Ω involved in the inner product (2.36) cannot be smaller than the number N of unknown coefficients; that is, it must be $M \geq N$. If this condition is not satisfied, the point-matching system of (2.66) does not have a solution. Hence, the corresponding MoM system of (2.65) also does not have a solution. If $M = N$, the weighted point-matching method

yields the same results as the ordinary point-matching method. Most often, M should be much greater than N in order that the inner product is evaluated accurately enough for given weighting functions. The optimal M is usually determined by numerical experiments, by seeking the minimal M whose further increase results only in negligible changes of the approximate solution. The number M of integration points may be reduced by an appropriate choice of the formula for numerical integration (e.g., Gaus-Legende formula), but not below N.

Example: Let us show under which condition the overdetermined system of point-matching equations (2.66) solved in the least-squares sense gives the same solution as the least-squares method. After multiplying by $\mathbf{F}_{\mathrm{PM}}^{\mathrm{T}}$, where the superscript T denotes transpose, the overdetermined system (2.66) reduces to the determined system

$$(\mathbf{F}_{\mathrm{PM}}^{\mathrm{T}}\mathbf{F}_{\mathrm{PM}})\mathbf{A} = \mathbf{F}_{\mathrm{PM}}^{\mathrm{T}}\mathbf{G}_{\mathrm{PM}} \tag{2.67}$$

The solution of (2.67) represents the solution of (2.66) in the least-squares sense [11]. If the inner product is defined by (2.61), the least-squares method can be represented by the weighted point-matching system (2.65), where the matrix $\mathbf{W}$ is equal to $\mathbf{F}_{\mathrm{PM}}^{\mathrm{T}}$. After representing the matrix $\mathbf{B}$ in the form

$$\mathbf{B} = \mathbf{B}_{1/2} * \mathbf{B}_{1/2} \qquad \mathbf{B}_{1/2} = \mathrm{diag}(\sqrt{b_1}, \ldots, \sqrt{b_M}) \tag{2.68}$$

the system (2.65) can be written as

$$[(\mathbf{B}_{1/2}\mathbf{F}_{\mathrm{PM}})^{*\mathrm{T}}(\mathbf{B}_{1/2}\mathbf{F}_{\mathrm{PM}})]A = (\mathbf{B}_{1/2}\mathbf{F}_{\mathrm{PM}})^{*\mathrm{T}}(\mathbf{B}_{1/2}\mathbf{G}_{\mathrm{PM}}) \tag{2.69}$$

Equation (2.69) reduces to (2.67) when all coefficients b_k have the same value. If the left and right sides of (2.66) are first multiplied by $\mathbf{B}_{1/2}$, then solved in the least-squares sense, we obtain the same solution as by solving (2.69).

2.3.11 Memory and Analysis-Time Requirements

The number of unknowns used in the analysis can become very high when a complex electromagnetic-field problem is analyzed, if high accuracy is required, or if an inadequate analysis method is applied. In those cases, we should pay attention to memory and analysis-time requirements.

To solve a system of linear equations efficiently, it is necessary that its matrix can be entirely stored in the computer operating memory (RAM). The

memory requirements are determined by the size of the matrix and its sparsity. The number of matrix elements is N^2, where the matrix order N is equal to the number of unknowns. Solving integral equations usually results in full matrices, in which case all N^2 elements should be stored. Since in single precision four bytes are needed to store a real number, the memory needed to store a matrix is

$$\text{MEM [bytes]} = 4N^2 \tag{2.70}$$

For example, if $N = 1{,}000$, we need 4 Mb of RAM. In double precision, eight bytes are used to store a real number and the memory requirements defined by (2.70) are doubled. These requirements are also doubled if matrix elements are complex, since each complex number consists of two real numbers. Finally, the memory requirements are approximately halved for symmetrical matrices.

When a linear operator equation is reduced to a system of linear equations, the analysis time (i.e., time used for the analysis) consists of the matrix fill-in time, the free-term column fill-in time, the time needed to solve the system of linear equations, and the data post-processing time. The free-term column fill-in time and the data post-processing time can often be neglected. The matrix fill-in time is proportional to the squared number of matrix elements, N^2, and to the time needed to evaluate one matrix element, A. The time needed to solve the system of linear equations is proportional to the number of basic operations and to the time needed for one basic operation, B. (One basic operation consists of one addition and one multiplication.) A system of linear equations is usually solved using the so-called direct methods (e.g., Gaussian elimination). In that case, the number of basic operations required to handle full matrices is approximately $N^3/3$. (The time needed to solve the system of linear equations is usually called the matrix inversion time, although this is not quite correct, since the number of basic operations is N^3 for a true inversion.) Finally, the analysis time can be expressed as

$$T = AN^2 + \frac{1}{2}BN^3 \tag{2.71}$$

For a small number of unknowns the matrix fill time is dominant, and the analysis time increases as N^2. For a large N the matrix inversion time is dominant, and the analysis time increases as N^3.

The time needed for evaluation of one matrix element, A, depends on the test procedure applied. Most often the MoM can be regarded as the weighted point-matching method (Section 2.3.10). In that case, the matrix fill-in time

consists of three parts. The first part is imposition of the point-matching matrix equation (2.66). The second part is evaluation of the weighting matrix $\mathbf{W} * \mathbf{B}$. The third part is the multiplication of the matrix equation by $\mathbf{W} * \mathbf{B}$. The time needed for imposition of (2.66) is equal to $A_{\mathrm{PM}} MN$, where A_{PM} represents the evaluation time of one matrix element when the point-matching method is applied. For the least-squares and the Galerkin methods, the evaluation of $\mathbf{W}$ is actually included in imposition of (2.66) and the time needed for multiplication of $\mathbf{W}^*$ and $\mathbf{B}$ is negligible. When the Galerkin method is applied to an integral operator equation, the basis functions are usually defined only in one part of the testing domain and $\mathbf{W}$ is very sparse. When the least-squares method is applied to an integral operator equation, the integral operator acting on the basis functions usually produces a response at all points of testing domain, and $\mathbf{W}$ is full. Hence, the time needed for multiplication of the matrix equation (2.66) by the weighting matrix $\mathbf{W} * \mathbf{B}$ can be neglected for the Galerkin method, while for the least squares-method, this time is equal to BMN^2. Having all this in mind, the analysis time for various methods is summarized in Table 2.4. For the least-squares and Galerkin methods, the matrix can be symmetrical (i.e., $f_{ji} = f_{ij}$), so that the analysis time can be halved.

2.3.12 Choice of Test Procedure

Before applying the MoM to solve a linear operator equation, one must choose a test procedure. Three basic test procedures are the point-matching method, the least-squares method, and the Galerkin method. In rare cases when the equation considered should be solved only once, the optimal test procedure is the point-matching method, which is easiest to implement. However, if a

Table 2.4
Analysis Time of the MoM Due to Different Test Procedures

Test Procedure	Analysis Time
Point-matching	$T = A_{\mathrm{PM}} N^2 + \frac{1}{3} BN^3$
Least-squares	$T = mA_{\mathrm{PM}} N^2 + \left(m + \frac{1}{3} \right) BN^3$
Galerkin	$T = mA_{\mathrm{PM}} N^2 + \frac{1}{3} BN^3$

A_{PM}: time needed for evaluation of matrix element when point-matching method is applied.
B: time needed for one basic operation (addition + multiplication).
m: ratio of number of integration points, M, and number of unknown coefficients, N.

general method should be developed such that the equation considered is to be solved for various problems repeatedly, an optimized test procedure should be designed to provide the desired accuracy with the lowest memory requirements and the shortest analysis time.

Memory requirements are closely related to the number of unknowns used in the analysis (Section 2.3.11). The matrix of the system of linear equations can be symmetrical for the least-squares and the Galerkin method. For the same number of unknowns used in the analysis, the memory requirements for these two methods can be halved when compared with the point-matching method.

The analysis time is also closely related to the number of unknowns used in the analysis, but also to the ratio of the number of integration points and the number of unknowns (see Table 2.4). The analysis time is critical for a large number of unknowns. In that case, for the same number of unknowns, the analysis time is approximately the same for the point-matching and the Galerkin method, and at least four times shorter than for the least-squares method. For symmetrical matrices, the analysis time needed by the Galerkin and the least-squares method is halved.

For the example of a line with two square conductors, all three methods require approximately the same number of unknowns for a specified accuracy. The residuum occurring in this particular problem is a well-behaved function. The experience of many authors shows that similar conclusions are valid for various problems in numerical electromagnetics whenever the residuum is a well-behaved function. When the residuum is not a well-behaved function, the Galerkin method seems to be more robust than the other two methods. Hence, the point-matching method is often used because of its simplicity; however, when efficiency and accuracy are required, the Galerkin method is usually the preferred choice.

2.3.13 Choice of Basis Functions

As illustrated in the previous sections, we can improve the accuracy of analysis by the proper choice of test procedure, but only to a certain level. Moreover, experience shows that the accuracy does not depend much on the test procedure applied when the approximate function enables a good-enough representation of the unknown function. Hence, the choice of the basis functions is even more important than choice of the test procedure.

Theoretically, it is required that the basis functions belong to the domain of the operator. For the Galerkin method, it is also required that the basis functions belong to the range of the operator. If we want the approximate solution to converge with the exact one when increasing the order of the

approximation, the basis functions should form a so-called complete set in the domain of the operator, and so forth. However, in practice, experience and engineering sense usually help more than theory in the choice of the basis functions.

The simplest choice of the basis functions is when these functions have constant values over domains of equal shapes and sizes; that is, when we use a PWC approximation in conjunction with uniform meshing. Regarding the choice of the basis functions, we can improve the accuracy in two ways: using sophisticated basis functions, and performing nonuniform meshing. Various sophisticated approximations can be established. Some of them use functions of higher orders (higher-order basis functions). Some of them are constructed to satisfy specific conditions at the boundaries of the domains of the basis functions. Some of them take into account edge, end, and proximity effects. Regarding the meshing, it is advisable to make the domains smaller in regions where the unknown function is expected to have rapid variations.

Example: Consider nonuniform meshing of a line with two square conductors. The charge distribution in the vicinity of an isolated metallic wedge can be expressed as

$$\rho_s(r) \sim r^{-p} \qquad p = 1 - \frac{\pi}{2\pi - \alpha} \tag{2.72}$$

where r is the normal distance between the source point and the wedge, and α is the angle between two planes forming the wedge [see Figure 2.10(a)]. For the right angle wedge, $\alpha = \pi/2$ and $p = 1/3$. The charge distribution close to

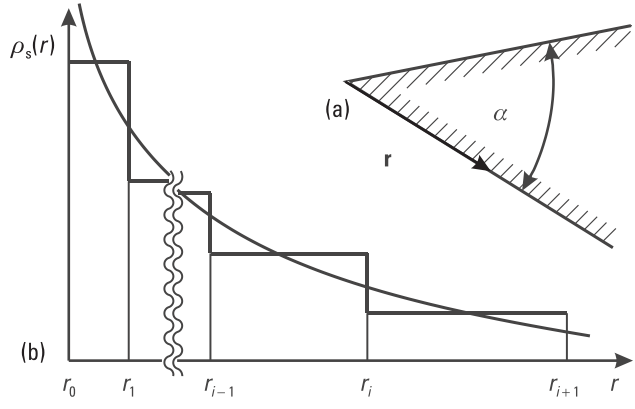

Figure 2.10 (a) Sketch of a metallic wedge, and (b) PWC approximation of charges in its vicinity.

the wedge changes much faster than the charge distribution far away from the wedge. When approaching the wedge, lengths of subsegments should decrease. To obtain an optimal nonuniform segmentation associated with the PWC approximation of charges, ρ_{sa} [see Figure 2.10(b)], we impose the condition that the integrals of the absolute error of the charge distribution are equal for all segments; that is,

$$\int_{r_{i-1}}^{r_i} |\rho_s(r) - \rho_{si}|\, dr = q \qquad i = 1, 2, \ldots \tag{2.73}$$

where r_{i-1} and r_i are the distances from the end points of the ith segment to the wedge, and q is an unknown constant. (For $i = 1$, the first distance is $r_0 = 0$.) In the vicinity of a wedge, the values of r_i depend on the number of segments and their total length, but the ratios r_i/r_1 are constant. Hence, the problem of the nonuniform segmentation is reduced to the problem of finding these ratios. To facilitate the task, let $r_1 = 1$. As the first approximation, $\rho_s(r)$ given by (2.72) is represented by the first two terms of the Taylor series with respect to the middle point of the ith segment, r_{i0}. In that case, the minimum error occurs for $\rho_{sa} = \rho_s(r_{i0})$ and q is easily obtained from (2.73) for $i = 1$. For $i > 1$, we obtain $r_i - r_{i-1} = (r_i + r_{i-1})^a$. Since for almost all segments $r_i - r_{i-1} << 2r_{i-1}$, the right side of this equation can be transformed using the small-argument approximation $(1 + x)^a \cong 1 + ax$. Finally, after simple rearrangements, we obtain the approximate recursive formula for the nonuniform meshing in the vicinity of a wedge

$$r_i = r_{i-1} + \frac{(2r_{i-1})^a}{1 - a(2r_{i-1})^{a-1}} \qquad a = \frac{1}{2}(p + 1) \tag{2.74}$$

Example: Consider a line with two square conductors $a = 1$ mm, $d/a = 1.368469$, $V_1 = -V_2 = 1$V. The analysis is based on the nonuniform segmentation of each square side into $m = 10$ subsegments. Figures 2.11(a) and 2.11(b) show the charge and potential distributions, respectively, along the circumference of one square conductor. The approximations of charges and potential in the vicinity of corners are much better than in the case of uniform segmentation [Figures 2.3(b) and 2.6(b)], while in the central part of each edge these approximations are practically the same. However, the residuum, $R(l) = V(l) - V_1$, is still greater for the left and right square sides than for the upper and lower square sides. This shows that the applied nonuniform segmentation, although much better than the uniform segmentation, is still not optimal. An optimal

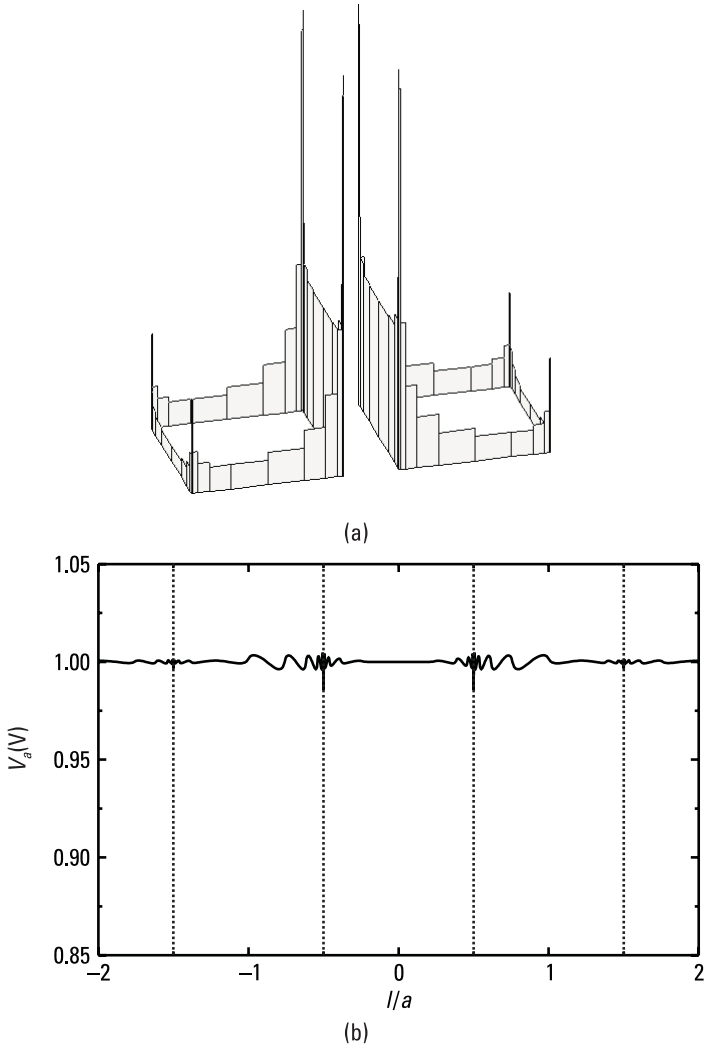

Figure 2.11 (a) Charge distribution, and (b) potential distribution along the circumference of one square conductor obtained by using m = 10 unknowns per square side.

segmentation should provide the same deviation of the approximate potential along any square side.

2.3.14 Adaptive Methods

The accuracy of the PWC approximation is increased if a nonuniform segmentation is applied (see Section 2.3.13). Such a segmentation is always based on

some prediction of the unknown function. For the example of the line with square conductors, the relative distribution of charges in the vicinity of the wedges can be predicted. However, it is not so easy to predict fast variations of charges due to the proximity effect. Hence, in some cases, nonuniform segmentation based on experience and prediction does not give results that are accurate enough. The remedy is to use an adaptive method. Generally, an adaptive method is an iterative method, whose each iteration consists of the following steps:

1. The solution is found for a given approximation of the unknown function.
2. The solution is validated according to some error criterion.
3. If the solution is satisfactory, the analysis is finished. Otherwise, the approximation of the unknown function is modified according to some rule, and the whole procedure is repeated.

The starting approximation of the unknown function can be chosen in various ways. For example, it can be adopted in the form of the simplest possible expansion, for which it is known in advance that many iterations are needed before a satisfactory solution can be obtained. Based on experience, we can also try to adopt a starting expansion that will pass the validation test in the first iteration. In that case, additional iterations are performed only if the initial prediction was not good enough.

Example: Consider an arbitrary two-conductor line. According to (2.23), the exact value of characteristic impedance Z is somewhere between the predicted minimum and maximum values, $Z_{\min}$ and $Z_{\max}$. For some approximate value Z_a in this range, the maximum possible value of the relative error $\delta = |Z_a - Z|/|Z|$ occurs for either $Z = Z_{\min}$ or $Z = Z_{\max}$. This value is minimized if Z_a is equal to the mean harmonic value of $Z_{\min}$ and $Z_{\max}$, in which case δ for $Z = Z_{\min}$ is equal to δ for $Z = Z_{\max}$. Thus, the value obtained for Z_a and the corresponding maximum possible relative error are evaluated as

$$\frac{1}{Z_a} = \frac{1}{2}\left(\frac{1}{Z_{\min}} + \frac{1}{Z_{\max}}\right) \qquad \delta_{\max} = \frac{Z_{\max} - Z_{\min}}{Z_{\max} + Z_{\min}} \qquad (2.75a,b)$$

Based on the above estimation of the relative error, the adaptive method for the analysis of two-conductor lines is established [15]. Let the analysis be performed by using the Galerkin method combined with a PWC approximation

of charges. The main goal of the analysis is to obtain the characteristic impedance with a relative error less than a prescribed value, δ_0 (i.e., $\delta_{\max} < \delta_0$). After replacing (2.23) into (2.75b), this condition can be written in the form

$$\frac{(V_{1\max} - V_{1\min}) + (V_{2\max} - V_{2\min})}{(V_{1\max} + V_{1\min}) - (V_{2\max} + V_{2\min})} < \delta_0 \tag{2.76}$$

The accuracy of the analysis can be improved by increasing the number of segments, for example, by dividing all or some segments into a number of smaller ones. The question is which segments should be further divided.

To decide which segments should be divided, we need a local-error criterion (applicable to segments) equally strict as or stricter than the global criterion given by (2.76). To derive such a criterion is to rewrite (2.76) as

$$\frac{V_{1\max} - V_1}{\Delta V} + \frac{V_1 - V_{1\min}}{\Delta V} + \frac{V_{2\max} - V_2}{\Delta V} + \frac{V_2 - V_{2\min}}{\Delta V} < 4\delta_0 \tag{2.77a}$$

$$\Delta V = \frac{V_{1\max} + V_{1\min}}{4} - \frac{V_{2\max} + V_{2\min}}{4} \tag{2.77b}$$

and require that each of the four terms on the left side of inequality (2.77a) be less than δ_0. If all the segments satisfy this criterion, the relative error for the characteristic impedance is less than δ_0. If not, those segments that do not satisfy the criterion should be further divided and the calculation repeated.

This method enables very accurate analysis of two-conductor lines. For example, for the line with square conductors with $a = 1$ mm, $d/a = 1.368469$, $V_1 = -V_2 = 1$V, the characteristic impedance is estimated to $Z = 75 \pm 0.0001\Omega$, by using about $N = 1000$ segments. If the adaptive analysis starts with two equal segments per square side, the solution is obtained after 24 iterations. The number of iterations can be decreased if we start from more than two equal segments per square side, if the starting solution is obtained using automatic nonuniform segmentation, or if the segments that do not satisfy (2.77a) are nonuniformly subdivided. However, in all these cases, the number of segments required for the prescribed accuracy is about $N = 1{,}000$. Obviously, the number of unknowns needed for the desired accuracy cannot be further decreased by using another segmentation technique, but only by using more sophisticated basis functions.

Figure 2.12 shows the relative error of the characteristic impedance versus the number of unknowns obtained by using three types of segmentation. The accuracy of the results is dramatically increased by using the automatic

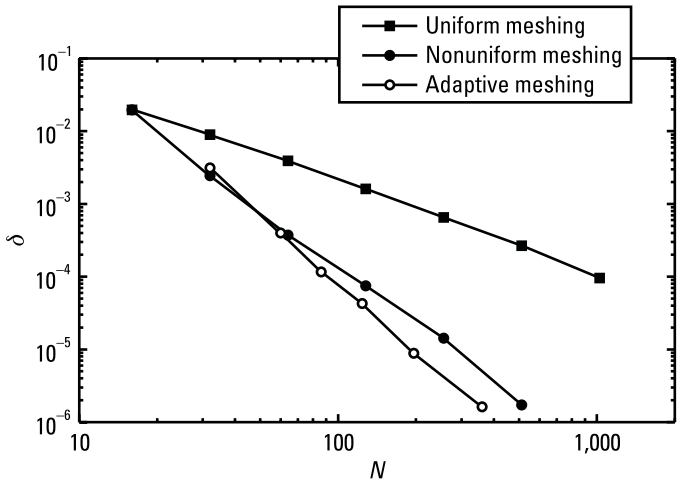

Figure 2.12 Relative error of characteristic impedance of the line with square conductors (d/a = 1.368469) versus the number of unknowns obtained by using various types of segmentation.

nonuniform segmentation instead of the uniform one. The adaptive nonuniform segmentation additionally improves accuracy, approaching the limits of the PWC approximation. However, this additional improvement is relatively small, which means that (2.74) enables very good prediction of the optimal segmentation.

2.4 Solution of Differential Equations

2.4.1 Approximation of Unknown Function

For differential equations, the approximation of unknown function (f) is usually slightly different than for integral equations. Namely, for integral equations the starting approximation of the unknown function is usually given in the form of a series of known basis functions multiplied by unknown coefficients, as in (2.9). For differential equations, the basis functions must satisfy the boundary conditions (the conditions at the boundary of the domain of the unknown function) and must be differentiable at least of the same order as is the order of the differential operator.

To enable easy satisfaction of the boundary conditions when solving differential equations, the domain of the unknown function is usually first discretized by a properly adopted mesh. Then, the approximation of the unknown function is initially represented by a set of discrete, but still unknown,

values of the function itself, which are taken at the mesh nodes. Finally, the discrete values of the unknown function along the boundary are forced to satisfy the boundary conditions.

The problem of differentiability of the basis functions is solved in various ways. Depending on the way, the initial approximation (given in the form of a set of discrete values) is further rearranged. Some methods, like the FD method, use only these discrete unknown function values for analysis. Once these unknowns are determined, other values of the unknown function can be obtained by interpolation. Some other methods, like the FEM, use an approximation of the unknown function at all points of its domain. Hence, it is necessary to perform an interpolation before the analysis is commenced.

The interpolatory approximation can be represented in the form (2.9) where the unknown coefficient a_i is the (unknown) value of the function at the ith node of the mesh. The basis function f_i has unit value at this node and zero values at all other nodes of the mesh. In particular, when the function f is known at the boundary, the expansion (2.9) can be written in the form

$$f_a = \sum_{i=1}^{N} a_i f_i + f_B \qquad f_B = \sum_{i=N+1}^{N+N_B} a_i f_i \tag{2.78a,b}$$

where N is the number of interior nodes (at which the values of the function f, a_i, $i = 1, \ldots, N$, are not known), and N_B is the number of boundary nodes (at which the values of the function f, a_i, $i = N + 1, \ldots, N + N_B$ are known).

Example: Consider the square coaxial line ($b/a = 2$, $V_1 = 1$V, $V_2 = 0$) shown in Figure 2.13. The dielectric is a vacuum. A grid of lines, spaced a distance h apart, defines a mesh of nodes at which the unknown function values should be determined.

The simplest method of interpolation is to subdivide each square cell of the mesh into two triangles as shown by dashed lines in Figure 2.13. Consider one arbitrary triangle in the xOy-plane, determined by the vertices $P_1(x_1, y_1)$, $P_2(x_2, y_2)$, and $P_3(x_3, y_3)$ [Figure 2.14(a)]. Suppose that the function values at these triangle vertices, f_1, f_2, and f_3, are known. The spatial points (x_1, y_1, f_1), (x_2, y_2, f_2), and (x_3, y_3, f_3) uniquely determine a plane as shown by the shaded surface in Figure 2.14(a). The function values at other points of the triangle can be obtained by a linear interpolation; that is,

$$f(x, y) = a + bx + cy \tag{2.79}$$

The coefficients a, b, and c are determined from the condition that the above expression gives values f_1, f_2, and f_3 at vertices P_1, P_2, and P_3, respect-

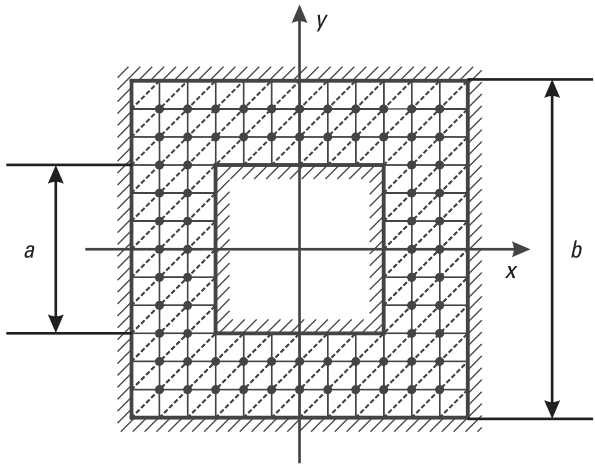

Figure 2.13 Grid of lines spaced distance h apart defines a mesh of points over the cross section of a square coaxial line ($b/a = 2$).

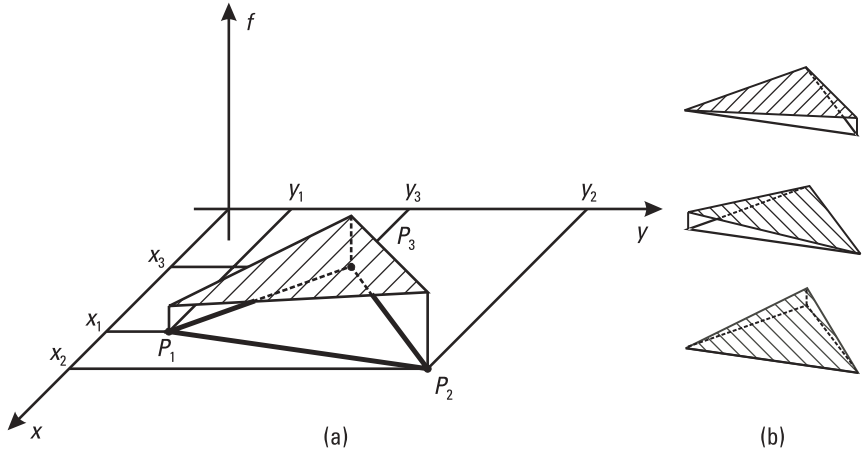

Figure 2.14 Linear approximation of a function over a triangular region: (a) approximation of the complete function, and (b) basis functions having unit value at one node and equal to zero at other two nodes.

ively; that is, these coefficients are determined from the following system of linear equations:

$$a + bx_1 + cy_1 = f_1 \qquad a + bx_2 + cy_2 = f_2 \qquad a + bx_3 + cy_3 = f_3 \tag{2.80}$$

The final solution for these coefficients can be comprehensively written in matrix form as

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{bmatrix}^{-1} \begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix} \tag{2.81}$$

Such an interpolation is most often applied after the analysis is performed. When the interpolatory function is needed for the analysis itself, the function given by (2.79) is usually represented as a linear combination of the three basis functions. Each basis function is equal to unit at one vertex and zero at other two vertices. For example, the first basis function can be written in the form

$$f_1(x, y) = a_1 + b_1 x + c_1 y \tag{2.82}$$

where the coefficients a_1, b_1, and c_1 are determined by (2.81), where f_1 is equal to unit, while f_2 and f_3 are equal to zero. After simple manipulations, these coefficients can be written as

$$a_1 = \frac{x_2 y_3 - x_3 y_2}{2S} \qquad b_1 = \frac{y_2 - y_3}{2S} \qquad c_1 = \frac{x_3 - x_2}{2S} \tag{2.83}$$

where S is the surface area of the triangle [i.e., $2S = (x_2 - x_3)(y_3 - y_1) - (x_3 - x_1)(y_2 - y_3)$]. The other two basis functions can be also represented by (2.82) and (2.83), except that the indices 1, 2, and 3 should be cyclically interchanged. These basis functions are shown by shaded surfaces in Figure 2.14(b).

The decomposition of the interpolatory function into three basis functions enables the continuity of the complete function defined over the triangular mesh to be easily imposed. All basis functions that are different from zero at one node should have the same value at this node (e.g., unit) and are multiplied by the same unknown coefficient. All such basis functions actually make a new multiple basis function (multiplet). Depending on the position of the central multiplet node, various multiplet kinds can occur (Figure 2.15). Expansions consisting of multiplets also represent PWL approximations.

When the outer conductor of the square coaxial line is at a zero potential (i.e., $V_2 = 0$), all multiplets whose central nodes are at the outer boundary can be omitted from the expansion. If the inner conductor is at a unit potential (i.e., $V_1 = 1$V), all multiplets whose central nodes are at the inner boundary

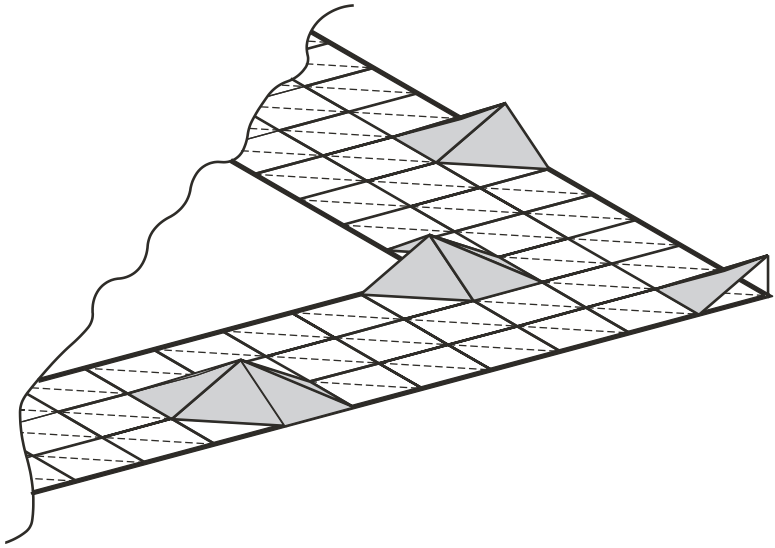

Figure 2.15 Sketch of multiplet basis functions.

should be taken with unit weight. In that case, the approximation for the potential can be written in the form

$$V_a = \sum_{i=1}^{N} a_i V_i + V_B \qquad V_B = \sum_{i=N+1}^{N+N_B} V_i \tag{2.84a,b}$$

where N is the order of the approximation (i.e., the number of interior nodes), V_i are multiplet basis functions, a_i are unknown coefficients (unknown function values at interior nodes) that should be determined, and N_B is the number of boundary nodes.

2.4.2 FD Method

The point-matching method enables the simplest solution of an integral linear equation (see Section 2.3.2). However, for a differential linear equation the point-matching method can be directly applied only if the basis functions are differentiable on at least the same order as that of the differential operator. For example, if the differential equation is of the second order, which is often the case in electromagnetic-field analysis, the basis functions should be at least twice differentiable. Even very sophisticated multiplets do not satisfy this condition.

One possible way to enable using simple basis functions is to apply the FD approximation. This involves replacing all derivatives by finite differences

(i.e., replacing the original operator by an approximate operator). This approximate-operator equation expresses the value of the unknown function at an arbitrary point as a linear combination of its values at several neighboring points. If this equation is imposed at a mesh of points, the finite differences can be defined using only the points belonging to the mesh, so that a system of linear equations is obtained in terms of unknown function values at the points of the mesh. This method is well known as the FD method.

The FD method can be interpreted as the point-matching method applied to a linear differential equation. The differential operator is approximated by the FD operator, the unknown function is represented by an interpolatory function where the function values at the interpolatory nodes are the unknown coefficients to be determined, and the interpolatory nodes are at the same time the matching points.

Example: Consider the square coaxial line shown in Figure 2.13 ($b/a = 2$, $V_1 = 1\text{V}$, $V_2 = 0$). The Laplace equation (2.2) can be written in the xy-coordinate system as

$$\frac{\partial^2 V(x, y)}{\partial x^2} + \frac{\partial^2 V(x, y)}{\partial y^2} = 0 \tag{2.85}$$

The first derivative of the potential $V(x, y)$ with respect to the x-coordinate is approximated as

$$\frac{\partial V(x, y)}{\partial x} \approx \frac{V(x + h, y) - V(x - h, y)}{2h} \tag{2.86}$$

Using (2.86), the second derivative of the potential $V(x, y)$ with respect to the x-coordinate is obtained as

$$\begin{aligned} \frac{\partial^2 V(x, y)}{\partial x^2} &\approx \frac{\dfrac{\partial V(x + h, y)}{\partial x} - \dfrac{\partial V(x - h, y)}{\partial x}}{2h} \\ &\approx \frac{V(x - h, y) - 2V(x, y) + V(x + h, y)}{(2h)^2} \end{aligned} \tag{2.87}$$

Similarly we obtain the second-order difference with respect to the y-coordinate, so that the Laplace equation can be approximated as

$$V(x, y) = \frac{V(x + h, y) + V(x - h, y) + V(x, y + h) + V(x, y - h)}{4} \tag{2.88}$$

After requiring that this approximate-operator equation is satisfied at the interior nodes of the uniform mesh shown in Figure 2.13, the system of linear equations is obtained in the form

$$4V_i - V_{i\text{-Left}} - V_{i\text{-Right}} - V_{i\text{-Down}} - V_{i\text{-Up}} = 0 \qquad i = 1, \ldots, N \tag{2.89}$$

where V_i is the unknown potential at the ith node, and $V_{i\text{-Left}}$, $V_{i\text{-Right}}$, $V_{i\text{-Down}}$, and $V_{i\text{-Up}}$ are potentials at the neighboring nodes (positioned left, right, down, and up from the ith node, respectively). The potentials at neighboring nodes placed inside the coaxial line are also unknown, while the potentials at neighboring nodes placed on the coaxial line conductors are known (V_1 = 1V for the outer conductor, and $V_2 = 0$ for the inner conductor.) The potentials that are known can be moved to the right side of (2.89), thus representing the free terms of the system of linear equations.

Once, the unknown values of potential are determined, all other quantities of interest can be easily calculated by data postprocessing. For example, the electric-field vector

$$\mathbf{E}(x, y) = -\frac{\partial V(x, y)}{\partial x}\mathbf{i}_x - \frac{\partial V(x, y)}{\partial y}\mathbf{i}_y \tag{2.90}$$

can be approximately calculated at the mesh nodes by using the FD approximation given by (2.86). The surface charge density at a node placed at conductor surfaces is equal to the corresponding normal electric-field component multiplied by the free-space permittivity (ϵ_0). The total approximate per-unit-length charge of one conductor is obtained by summing the surface-charge densities for all nodes placed on the conductor surface and multiplying this sum by the distance between the nodes, h.

Finally, the per-unit-length capacitance can be calculated by using various formulations slightly different than those given in Table 2.2. Namely, in the solution of an integral equation, for both conductors the approximate per-unit-length charges are of equal magnitudes, while in the solution of a differential equation, the approximate potential along each conductor is constant. Hence, for a differential equation, the C_2' formulation is equal to the C_1' formulation. However, the C_1' and C_4' formulations based on the per-unit-length charge of the inner conductor are different from the same formulations based on the

charge of the outer conductor. In particular, it is shown that the C_4' formulation for the outer conductor and the C_3' formulation are equal to the C_1' formulation for the outer conductor. So, we actually have only three different formulations: C_1' and C_4' for the inner conductor, and C_3'.

Figure 2.16 shows the characteristic impedance versus the number of segments per inner square side. The results are obtained by using the C_1' formulation applied to the inner conductor and the C_3' formulation, which is the same as the C_1' formulation applied to the outer conductor. In addition, results obtained by the MoM (Galerkin method) are presented. The exact value obtained by the conformal mapping [10] is $Z = 36.786\Omega$. For the Galerkin method all four formulations for capacitance (C_1', $i = 1, 2, 3, 4$) give the same result. The results of solving the Laplace equation based on the C_3' formulation converge much faster than the results based on the C_1' formulation. However, the MoM results converge much faster than any of the FD results.

2.4.3 Sparse Matrices

The solution of a linear differential equation usually results in a sparse matrix equation (i.e., most matrix elements are zero). It is convenient if all matrix elements different from zero are grouped in the vicinity of the main diagonal of the matrix. In that case, the effective bandwidth of the matrix is defined as the maximum distance of nonzero matrix elements from the main diagonal. When the effective bandwidth of the matrix, L, is much less than its order, N, the memory requirements and the matrix inversion time are significantly

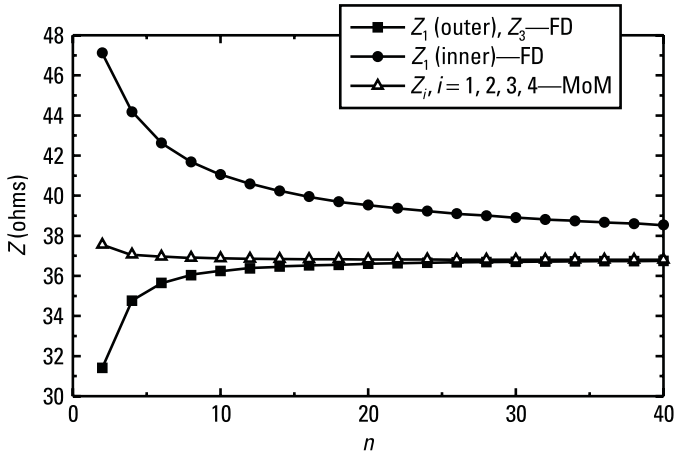

Figure 2.16 Characteristic impedance of square coaxial line ($b/a = 2$) versus the number of unknowns per inner square side, n.

reduced when compared with a full matrix of the same order. If all matrix elements inside the envelope determined by the effective bandwidth of the matrix are different from zero, the size of memory needed for the matrix is

$$\text{MEM [bytes]} = 4[N^2 - (N - L)^2] \approx 8LN \tag{2.91}$$

If the matrix equation is solved by Gaussian elimination, the analysis time, which consists of the matrix fill-in time and the matrix inversion time, can be expressed as

$$T = ANL + BNL^2 \tag{2.92}$$

where A is the time needed for evaluation of one matrix element and B is the time needed for one "basic operation" (addition and multiplication). Usually, A is comparable with B, which means that the matrix inversion time is dominant in the solution of differential equations.

In practice, many elements inside the envelope determined by the effective bandwidth of the matrix are also equal to zero. However, it is difficult to make a general algorithm that can use this type of sparsity to further decrease the memory requirements and the matrix fill-in time given by (2.91) and (2.92). Hence, particular attention in the solution of differential equations is paid to obtaining a matrix with the minimal effective bandwidth. The actual sparsity pattern depends on the ordering of the matrix rows and columns, which, in turn, depends on the numbering scheme used for the mesh nodes. Hence, we can minimize the effective bandwidth of the matrix in two ways: by adopting an optimal numbering scheme for the mesh nodes or developing specialized reordering algorithms [20]. Most often, the optimal numbering scheme minimizes the difference between indices of neighboring nodes.

Example: Consider the square coaxial line shown in Figure 2.13 ($b/a = 2$). If the uniform grid subdivides the inner square side into n segments, the number of unknowns used in the analysis is $N = (2n - 1)^2 - (n + 1)^2 = 3n^2 - 6n$. Let the analysis be performed by the FD method (see Section 2.4.2). If the matrix of the system of linear equations is treated as a full matrix, according to (2.70) and (2.71) the memory requirements and the matrix inversion time for large n are obtained as MEM (bytes) $= 36n^4$ and $T = 9n^6B$.

The efficiency of the method can be improved by using the fact that the system of linear equations is sparse. If we numerate the nodes row by row [as shown in Figure 2.17(a)], the effective bandwidth of the matrix is $L = 2n$. This effective bandwidth can be further decreased to $L = n - 1$ if the nodes are numerated diagonal by diagonal [as shown in Figure 2.17(b)]. According to

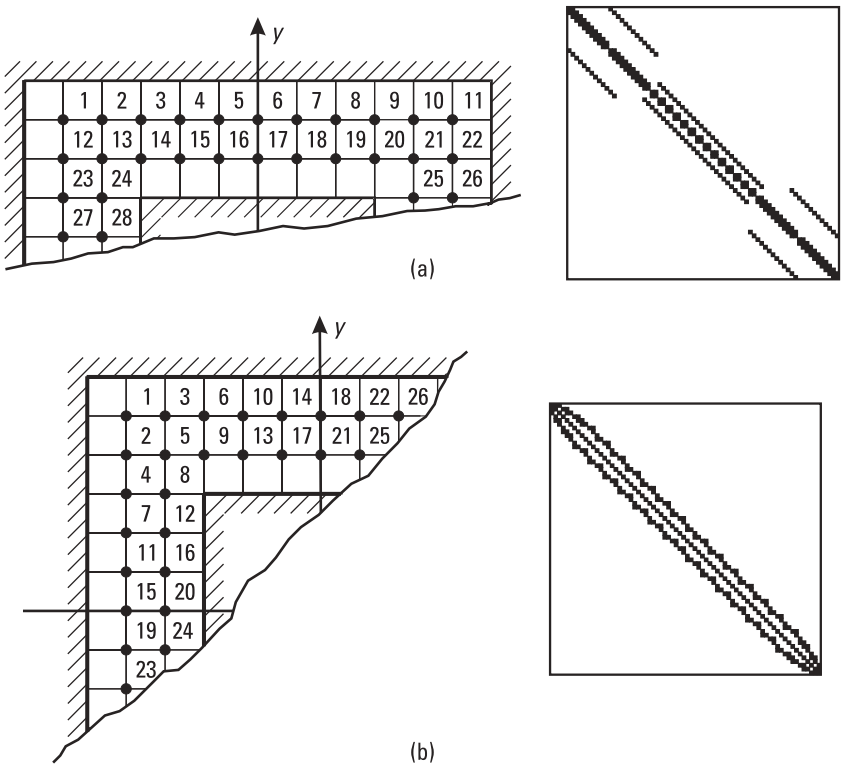

Figure 2.17 Grid and matrix sparsity pattern for two numerations: (a) row by row and (b) diagonal by diagonal.

(2.91) and (2.92), the memory requirements and the matrix inversion time for large n are MEM (bytes) = $24n^3$ and $T = 3n^4B$.

2.4.4 Iterative Procedure

The FD method is sometimes organized as an iterative procedure. However, usually such an iterative procedure is only a particular case of iterative methods applied for the solution of matrix equations. The simplest of them is the Jacobi cyclical iteration method (simultaneous displacement method) [21]. Consider an arbitrary matrix equation $\mathbf{AX} = \mathbf{B}$ of order N, where $\mathbf{A}$ is a square matrix, $\mathbf{X}$ is the column of unknown coefficients, and $\mathbf{B}$ is the free-term column. According to the Jacobi method, the ith unknown coefficient at the $(k+1)$-th iteration is refined as

$$x_i^{(k+1)} = \frac{1}{a_{ii}}\left(b_i - \sum_{j=1,\, j\neq i}^{N} a_{ij} x_j^{(k)}\right) \tag{2.93}$$

The rate of convergence is usually improved if x_j enters (2.93) with its last calculated value; that is,

$$x_i^{(k+1)} = \frac{1}{a_{ii}}\left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^{N} a_{ij} x_j^{(k)} \right) \tag{2.94}$$

This modification is known as the Seidel method (successive displacement method). These two methods belong to the class of linear iterative methods for which the current estimation of the unknown coefficients represents a linear combination of previous estimations. Otherwise, iterative methods are nonlinear. Two major drawbacks to the linear iterative schemes are that the convergence can be quite irregular and very slow.

In attempt to overcome these shortcomings, many nonlinear iterative methods have been developed. Among them the *conjugate-gradient* (CG) method and its variants seem to be the most promising. A detailed derivation of the CG method can be found elsewhere (e.g., in [21]). Here we only summarize the procedure steps. To solve $\mathbf{AX} = \mathbf{B}$, the CG method starts with an initial guess $\mathbf{X}_0$ (e.g., $\mathbf{X}_0 = 0$), giving

$$\mathbf{P}_0 = -\mathbf{A}^{\mathrm{T}}\mathbf{R}_0 \qquad \mathbf{R}_0 = \mathbf{AX}_0 - \mathbf{B} \tag{2.95}$$

Each next iteration (except the last one) consists of the following steps

$$t_n = \frac{\|\mathbf{A}^{\mathrm{T}}\mathbf{R}_n\|^2}{\|\mathbf{AP}_n\|^2} \qquad \mathbf{X}_{n+1} = \mathbf{X}_n + t_n\mathbf{P}_n \qquad \mathbf{R}_{n+1} = \mathbf{R}_n + t_n\mathbf{AP}_n \tag{2.96a-c}$$

$$q_n = \frac{\|\mathbf{A}^{\mathrm{T}}\mathbf{R}_{n+1}\|^2}{\|\mathbf{A}^{\mathrm{T}}\mathbf{R}_n\|^2} \qquad \mathbf{P}_{n+1} = -\mathbf{A}^{\mathrm{T}}\mathbf{R}_{n+1} + q_n\mathbf{P}_n \tag{2.96d,e}$$

Assuming no truncation and roundoff error, the exact solution is obtained after a finite number of steps, which is less than or equal to the matrix order N. (Precisely, the finite number of steps is equal to the number of independent eigenvalues of matrix $\mathbf{A}$.)

The convergence of an iterative method for a particular matrix $\mathbf{A}$ can be predicted by the analysis of the eigenvalues of this matrix. For example, it is found that the number of iterations needed by the CG method to yield a satisfactory result is much lower than the matrix order N if the eigenvalues of the matrix $\mathbf{A}$ are grouped. However, the determination of eigenvalues is much

more complicated than the application of these methods. Hence, from the practical point of view, the simplest way to determine if iterative methods converge regularly and quickly for a specific class of problems is to apply them to this class.

We are particularly interested in applying the iterative methods if they require smaller memory requirements and shorter matrix inversion time than direct methods. Most often, for full matrices both types of methods have the same memory requirements. For sparse matrices, the memory requirements of iterative methods may be reduced $L/L_{\min}$ times compared with direct methods, where $2L_{\min}$ is the mean number of nonzero elements per matrix row. ($L_{\min}$ actually represents the theoretically minimal effective bandwidth of the matrix, which occurs when all matrix elements outside the envelope determined by $L_{\min}$ are zero.)

According to (2.93) and (2.94), the number of basic operations for the Jacobi and Seidel methods is MN^2 for a full matrix and $MN(2L_{\min})$ for a sparse matrix, where M is the number of iterations. Comparing these numbers of basic operations with the second terms in (2.71) and (2.91), it is concluded that the Jacobi and Seidel methods are more efficient than direct methods if $M < N/3$ for a full matrix and if $M < L^2/(2L_{\min})$ for a sparse matrix. According to (2.96), the number of basic operations for the CG method is doubled when compared with the Jacobi and Seidel methods for the same number of iterations.

It can be shown that there is no iterative method that enables more efficient solution than direct methods in most cases. The problem with all iterative methods, unlike direct methods, is how to estimate if the solution obtained at the end of the current iteration is satisfactory or not. In particular, some iterative methods require many more significant digits than direct methods to converge to the same solution. Hence, we recommend direct methods as the basic methods for solving matrix equations.

Example: The FD method applied to the solution of the Laplace equation in Section 2.4.2 can be organized as an iterative procedure. For example, (2.89) can be written in the form

$$V_i^{(k+1)} = \frac{V_{i\text{-Left}}^{(k)} + V_{i\text{-Right}}^{(k)} + V_{i\text{-Down}}^{(k)} + V_{i\text{-Up}}^{(k)}}{4} \qquad i = 1, \ldots, N \tag{2.97}$$

which means that the potential value at the ith node in the $(k+1)$-th iteration is calculated by using potential values at surrounding nodes calculated in the kth iteration. This iterative procedure is actually the Jacobi method applied to the system of linear equations (2.89).

The convergence can be improved if the potentials at the surrounding nodes enter (2.97) with their last calculated values. For example, let us consider evaluation of the potential at the ith node in the $(k + 1)$-th iteration for row-by-row numeration. The last calculated values of the potential at nodes that are left and up are obtained during the $(k + 1)$-th iteration, while those at nodes that are right and down were obtained during the kth iteration. In that sense, (2.97) should be modified to

$$V_i^{(k+1)} = \frac{V_{i\text{-Left}}^{(k+1)} + V_{i\text{-Right}}^{(k)} + V_{i\text{-Down}}^{(k)} + V_{i\text{-Up}}^{(k+1)}}{4} \qquad i = 1, \ldots, N \tag{2.98}$$

This iterative procedure is actually the Seidel method applied to the system of linear equations (2.89).

Figure 2.18 shows the relative error of the capacitance obtained by the two iterative methods versus the number of iterations M, where the number of segments per inner square side is $n = 24$. The relative error is calculated with respect to the value obtained by a direct method. For the same accuracy, the Jacobi method needs twice the iterations of the Seidel method. In this case, the effective matrix bandwidth is $L = n - 1 = 23$, while the mean number of nonzero elements per matrix row is about $2L_{\min} = 4$. In particular, according to (2.98), the basic operation is halved when compared with the Seidel method in general (i.e., it consists of only one addition instead of one addition and one multiplication). Hence, the critical number of iterations, below which the

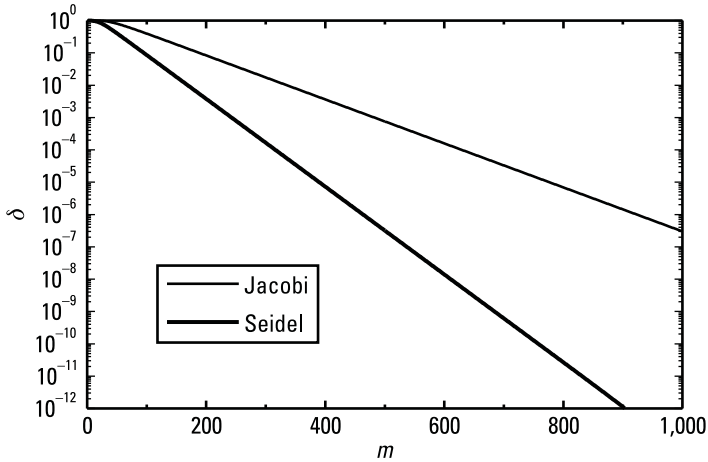

Figure 2.18 Relative error of characteristic impedance of square coaxial line ($b/a = 2$) versus the number of iterations.

Seidel method is more efficient than the direct methods, is $M = L^2/L_{\min} \approx 265$. After $M = 265$ iterations the Seidel method gives capacitance with relative error of about 10^{-3}. For engineers, the relative error of 10^{-3} is usually higher than satisfactory. In this particular case, the iterative method has an advantage over the direct method.

There is a substantial difference between the iterative FD method described above and iterative FD methods applied in the time-domain analysis (FDTD and TLM methods) [22, 23]. In the electrostatic case described here (as well as in any steady-state analysis), we are interested only in one solution of the linear system, which is obtained after the last iteration. In the case of the FDTD method, an iteration represents only one time step and results obtained at the end of this iteration represent the solution of the equation at the considered time instant. Hence, each iteration is a part of the solution. This simple numerical scheme makes the FDTD the preferred method for the time-domain analysis.

2.4.5 FEM Via the Galerkin Method

As mentioned in Section 2.4.2, simple basis functions cannot be directly applied for solution of differential equations [e.g., the Laplace equation (2.2)], and the equation should be transformed appropriately. For example, in the case of the FD method, the differential operator is approximated by the difference operator. Another general well-known technique for the solution of differential equations is the FEM [24–27].

In the FEM, we also start from differential equations and boundary conditions that should be locally satisfied (the so-called strong formulation). Then, we usually redefine the problem in such a manner that these equations and conditions are satisfied in some integral sense (the so-called weak formulation), so that simple basis functions defined within small domains (finite elements) can be applied. Often, the FEM is formulated via the Rayleigh-Ritz method applied to a properly adopted functional, where the functional is constructed from the differential equation to be solved. However, the FEM may be obtained simply by applying the Galerkin method in the following way.

After replacing the approximate expansion (2.78a) into the starting differential equation (2.5) and performing simple rearrangements, the approximate equation is obtained in the form

$$\sum_{i=1}^{N} a_i L f_i \approx g - L f_B \tag{2.99}$$

After applying the Galerkin test procedure and using (2.78b), this approximate equation is transformed into a determined system of linear equations of the form

$$\sum_{i=1}^{N} a_i \langle f_j, Lf_i \rangle = \langle f_j, g \rangle - \sum_{i=N+1}^{N+N_B} a_i \langle f_j, Lf_i \rangle \qquad j = 1, \ldots, N \tag{2.100}$$

However, the differentiability of the basis functions is usually lower than the order of the differential operator (i.e., the basis functions are not in the domain of the operator). In that case, we must transform the inner product $\langle f_j, Lf_i \rangle$ to decrease the order of the differential operator in such a manner that it can be applied to the basis functions. In general, the transformation can be written in the form

$$\langle f_j, Lf_i \rangle = \langle Pf_j, Qf_i \rangle_2 \tag{2.101}$$

where P and Q are differential operators of lower order than L, and the inner product on the right side can be different from the inner product on the left side. This transformation actually extends the domain of the original operator so that the basis function are included into the extended domain. The original operator combined with transformation (2.101) is called the extended operator.

Example: Apply the Galerkin method to the Laplace equation in the case of the line with square conductors shown in Figure 2.13, where the approximate potential is given by (2.84a). Suppose that the basis functions V_i are twice differentiable. After replacing (2.84a) into (2.2), the Laplace equation is obtained in the approximate form

$$\sum_{i=1}^{N} a_i \Delta V_i = -\Delta V_B \tag{2.102}$$

After applying the Galerkin method to this equation and using (2.84b), we obtain a system of linear equations

$$\sum_{i=1}^{N} a_i \int_{S_j} V_j \Delta V_i \, dS = -\sum_{i=N+1}^{N+N_B} \int_{S_j} V_j \Delta V_i \, dS \qquad j = 1, \ldots, N \tag{2.103}$$

where S_j is the surface over which the jth basis function is defined. By using the identity $\nabla(pq) = q\nabla p + p\nabla q$, which is valid for arbitrary scalar functions p and q, the integrals in (2.103) are transformed as

$$\int_{S_j} V_j \Delta V_i \, dS = \int_{S_j} \nabla(V_j \nabla V_i) \, dS - \int_{S_j} \nabla V_j \cdot \nabla V_i \, dS \tag{2.104}$$

By using the divergence theorem, the first integral on the right side of (2.104) is transformed to $\oint_{C_j} V_j \Delta V_i \cdot \mathbf{n} dl$, where $\mathbf{n}$ is the unit normal going out of the contour C_j, which bounds the surface S_j. Since the interior basis function V_j is equal to zero at its boundary contour C_j, this integral is equal to zero. Finally, after substituting (2.104) into (2.103), the system of linear equations is written as

$$\sum_{i=1}^{N} a_i \int_{S_j} \nabla V_j \cdot \nabla V_i \, dS = -\sum_{i=N+1}^{N+N_B} \int_{S_j} \nabla V_j \cdot \nabla V_i \, dS \qquad j = 1, \ldots, N \tag{2.105}$$

Note that the original system of linear equations (2.103) posed by the Galerkin method cannot be applied to once differentiable functions that are not in the domain of the Laplace operator. However, by the transformation (2.104), the domain of the original operator is extended so that the once differentiable functions are included in the domain of the extended Laplace operator.

For multiplets, the matrix elements are easily calculated. Note that ∇V_i is a constant vector over each triangle of the ith multiplet. For the particular case of a uniform mesh consisting of right triangles, shown in Figure 2.13, the matrix elements for $i = 1, \ldots, N + N_B$, and $j = 1, \ldots, N$, are simply evaluated as

$$\int_{S_j} \nabla V_i \cdot \nabla V_j \, dS = \begin{cases} 4/S & i = j \\ 1/S & i \neq j \text{ overlapping multiplets} \\ 0 & i \neq j \text{ nonoverlapping multiplets} \end{cases} \tag{2.106}$$

Generally, a multiplet overlaps with four neighboring multiplets positioned left, right, up, and down. In that case, the FEM system of linear equations

(2.105) is equivalent to the FD system of linear equations (2.89), resulting in the same solution for the approximate potentials at the mesh nodes.

2.5 Choosing the Optimal Method

In the previous sections, we introduced various methods for solution of electromagnetic-field linear equations. Common for all these methods is that the linear equation is transformed into a determined system of linear equations. In all these cases, it can be interpreted that the system of linear equations is obtained by the MoM. If the differential operator is approximated by finite differences, the method is usually called the FD method. If the differential operator is extended or the system is obtained by the Rayleigh-Ritz method, the method is usually called the FEM. Traditionally, the name MoM is applied in narrow sense to denote only solutions of integral equations.

In this book, we are interested primarily in steady-state analysis in the frequency domain. The preferred techniques that can be applied to such problems are the MoM (in the narrow sense) and the FEM. Both methods have certain advantages. The geometrical elements used by the MoM usually have one or two dimensions fewer than those used by the FEM. (For the MoM the unknown quantities are field sources distributed along lines or over surfaces, while for the FEM the unknown quantities are fields or potentials distributed within volumes.) Hence, the geometrical modeling for the MoM is usually easier than for the FEM, and the number of unknowns required by the MoM is substantially lower than the number of unknowns required by the FEM. This disadvantage of the FEM is partly compensated for by the fact that matrices are always sparse, while the MoM matrices are predominantly full. Taking all this into account, experience shows that the MoM is more efficient than the FEM, except in the case of highly inhomogeneous structures [28]. In particular, the FEM is not so well suited for open problems as the MoM is, although a lot of effort has been made to develop absorbing boundary conditions. Hence, in this book the MoM is adopted as the main method for the analysis in the frequency domain.

Example: Let us compare the MoM (Galerkin method) and the FEM in the case of the square coaxial line ($b/a = 2$). Figure 2.19 shows the relative error of the characteristic impedance versus the number of segments, n, into which the side of the inner square is subdivided. For the desired accuracy, the number of segments n needed by FEM is about seven times larger than in the case when the MoM is applied (i.e., $n_{\mathrm{FEM}} \approx 7n_{\mathrm{MoM}}$). This is explained by the fact that the PWC approximation of the charge distribution in the MoM creates

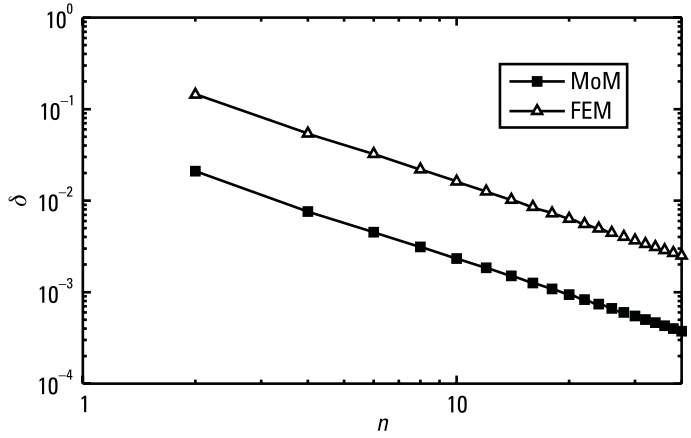

Figure 2.19 Relative error of characteristic impedance obtained by the MoM and the FEM versus the number of segments, n, into which the side of the inner square is subdivided.

continuous field distribution between the conductors, while in the FEM the field distribution between the conductors is PWC. The total number of unknowns used by the MoM, and the FEM in this case is $N_{\text{MoM}} = 12n_{\text{MoM}}$ and $N_{\text{FEM}} = 4(n_{\text{FEM}})^2$. For the desired accuracy, the total number of unknowns used by the FEM is dramatically greater than that used by the MoM; that is, the ratio of these two numbers is about $N_{\text{FEM}}/N_{\text{MoM}} = 16n_{\text{MoM}}$.

For a square coaxial line with $b/a = 2$, the optimal ordering of nodes is the diagonal one [Figure 2.17(b)]. For large n_{FEM}, the number of basic operations needed for solving the resulting sparse system of linear equations is $K_{\text{FEM}} = 3(n_{\text{FEM}})^4$. This number of basic operations is the same as if the line is analyzed by equivalent FD method (see Section 2.4.3). For large N_{MoM}, the number of basic operations needed for solving the full symmetrical system of linear equations is $K_{\text{MoM}} = (N_{\text{MoM}})^3/6$. The ratio of the numbers of basic operations needed by FEM and MoM for the same accuracy is about

$$\frac{K_{\text{FEM}}}{K_{\text{MoM}}} \approx 25n_{\text{MoM}} \tag{2.107}$$

For example, if $n_{\text{MoM}} = 40$, the FEM is 1,000 times slower than the MoM.

2.6 Summary

Before considering the frequency analysis of general composite metallic and dielectric structures, which is the main objective of the book, the basic concepts

of solution methods for electromagnetic-field problems are introduced in this chapter. We focus on linear operator equations and methods for their solution. The general method, which is based on the transformation of linear operator equations into a system of linear equations, is the MoM.

The solution of a field problem based on the MoM consists of the following steps:

- Inspection of field equations and corresponding uniqueness theorems;
- Choice of the starting set of equations;
- Geometrical modeling of the structure;
- Approximation of field sources or fields over or within the geometrical elements;
- Choice of test procedures and evaluation of matrix and free-term elements;
- Solution of the system of linear equations;
- Data postprocessing, including evaluation and validation of output results.

These steps are followed in the rest of the book.

Generally, field problems can be solved starting from integral or differential equations. Application of the MoM to differential equations has some specific features and, hence, specific names: the FD method and the FEM. Considering efficiency, the MoM applied to integral equation is adopted as the main method for the frequency-domain analysis in this book.

References

[1] Kantorovich, L. V., and V. I. Krylov, *Approximate Methods of Higher Analysis,* New York: John Wiley and Sons, 1964.

[2] Kantorovich, L. V., and G. P. Akilov, *Functional Analysis in Normed Spaces,* Oxford, U.K.: Pergamon Press, 1964.

[3] Nomura, Y., and T. Hatta, "The Theory of a Linear Antenna, I," *Tech. Rep. Tohoku Univ.,* Vol. 17, Pt. 1, 1952.

[4] Storm, B., "Investigation into Modern Aerial Theory and a New Solution of Hallen's Integral Equation for a Cylindrical Aerial," Dissertation, Imperial College, London, U.K., 1953.

[5] Harrington, R. F., "Matrix Methods for Field Problems," *Proc. IEEE,* Vol. 55, No. 2, February 1967, pp. 136–149.

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

[7] Djordjevic, A. R., et al., *LINPAR for Windows: Matrix Parameters for Multiconductor Transmission Lines—Software and Users' Manual,* Norwood, MA: Artech House, 1996.

[8] Popovic, B. D., *Introductory Engineering Electromagnetics,* Reading, MA: Addison-Wesley, 1971.

[9] Stakgold, I., *Boundary Value Problems of Mathematical Physics,* Vols. 1 and 2, New York: Macmillan, 1967.

[10] Wadell, B. C., *Transmission Line Design Handbook,* Norwood, MA: Artech House, 1991.

[11] Golub, G. H., and C. F. Van Loan, *Matrix Computation,* Baltimore, MD: Johns Hopkins University Press, 1983.

[12] Sarkar, T. K., "A Note on the Choice of Weighting Functions in the MoM," *IEEE Trans. Antennas Propagat.,* Vol. AP-33, No. 4, April 1985, pp. 436–441.

[13] Sarkar, T. K., A. R. Djordjevic, and E. Arvas, "On the Choice of Expansion and Weighting Functions in the Numerical Solution of Operator Equations," *IEEE Trans. Antennas Propagat.,* Vol. AP-33, No. 9, September 1985, pp. 988–996.

[14] Dudley, D. G., "Error Minimization and Convergence in Numerical Methods," *Electromagnetics,* Vol. 5, 1985, pp. 89–97.

[15] Kolundzija, B. M., et al., "Evaluation of Capacitance with Prescribed Accuracy—Adaptive Method Based on Exact Error Estimation," *IEEE Antennas Propagat. Magazine,* Vol. 40, No. 6, December 1998, pp. 80–84.

[16] Mikhlin, S. G., *Variational Methods in Mathematical Physics,* New York: Macmillan, 1964.

[17] Findlaysen, B. A., *The Method of Weighted Residuals and Variational Principles,* New York: Academic, 1972.

[18] Peterson, A. F., S. L. Ray, and R. Mittra, *Computational Methods for Electromagnetics,* New York: IEEE Press, 1998.

[19] Djordjevic, A. R., and T. K. Sarkar, "A Theorem on the Moment Method," *IEEE Trans. Antennas Propagat.,* Vol. AP-35, No. 3, March 1987, pp. 353–355.

[20] Cuthill, E., "Several Strategies for Reducing the Bandwidth of Matrices," in *Sparse Matrices and Their Applications,* D. J. Rose and R. A. Willoughby, (eds.), New York: Plenum, 1972.

[21] Sarkar, T. K., K. R. Siarkiewicz, and R. F. Stratton, "Survey of Numerical Methods for Solution of Large System of Linear Equation for Electromagnetic Field Problems," *IEEE Trans. Antennas Propagat.,* Vol. AP-29, No. 6, November 1981, pp. 847–856.

[22] Taflove, A., *Finite-Difference Time-Domain Method in Computational Electrodynamics,* Norwood, MA: Artech House, 1995.

[23] Christopoulos, C., *Transmission-Line Modeling Method: TLM,* New York: IEEE Press and Oxford: University Press, 1995.

[24] Zienkiewicz, O. C., and K. Morgan, *Finite Elements and Approximation,* New York: John Wiley and Sons, 1983.

[25] Silvester, P. P., and R. L. Ferrari, *Finite Elements for Electrical Engineers,* 3rd ed., New York: Cambridge University Press, 1996.

[26] Jin, J., *The Finite Element Method in Electromagnetics,* New York: John Wiley and Sons, 1993.

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

[28] Kolundzija, B. M., V. V. Petrovic, and A. R. Djordjevic, "Comparison of MoM/SIE, MoM/VIE and FEM Based on Topological Analysis of Two Canonical Problems," *Digest of IEEE AP-S Int. Symp.,* Vol. 1, June 1998, pp. 274–277.

3

Electromagnetic Theory

This chapter reveals the basic theory of time-harmonic electromagnetic fields. It includes some specific theorems particularly important for the electromagnetic modeling of composite metallic and dielectric structures. More details on this theory can be found, for example, in [1–11]. Readers familiar with this matter can directly proceed to Chapter 4 without a loss of continuity.

3.1 Maxwell's Equations

Macroscopic electromagnetic fields are governed by Maxwell's equations. These equations can be written in integral or differential form. The integral form is more general because the differential form is applicable only to fields that are differentiable functions of the position vector. However, the differential form is more compact and more appropriate for derivations presented in this book. The differential form of Maxwell's equations reads

$$\text{curl } \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \qquad \text{curl } \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t} \qquad \text{div } \mathbf{D} = \rho \qquad \text{div } \mathbf{B} = 0 \tag{3.1a-d}$$

where $\mathbf{E}$ is the electric field intensity, $\mathbf{D}$ is the electric flux density (also referred to as the electric displacement or the electric induction), $\mathbf{B}$ is the magnetic flux density (also referred to as the magnetic induction), $\mathbf{H}$ is the magnetic field intensity, $\mathbf{J}$ is the current density, and ρ is the volume free charge density. The sum $\mathbf{J} + \partial\mathbf{D}/\partial t$ is referred to as the total current density and $\partial\mathbf{D}/\partial t$ as the

displacement current density. Similarly, the term $\partial\mathbf{B}/\partial t$ is referred to as the magnetic displacement current density. All these quantities depend on time (t) and on the position vector ($\mathbf{r}$) or, equivalently, on three spatial coordinates (e.g., the Cartesian coordinates x, y, and z). For simplicity, these dependencies are not explicitly shown. The current density and the charge density are related by the continuity equation:

$$\text{div}\,\mathbf{J} = -\frac{\partial \rho}{\partial t} \tag{3.2}$$

Note that the continuity equation can be derived from Maxwell's equations by taking the divergence of (3.1b) and substituting div$\mathbf{D}$ from (3.1c). Equations (3.1) and (3.2) can be written in a compact form using the Hamiltonian (del operator) ∇, as

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad \nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t} \quad \nabla \cdot \mathbf{D} = \rho \quad \nabla \cdot \mathbf{B} = 0 \quad \nabla \cdot \mathbf{J} = -\frac{\partial \rho}{\partial t} \tag{3.3a-e}$$

The quantities $\mathbf{J}$ and ρ can be regarded as physical sources of any electromagnetic field. However, the charge motion must be originated by a generator (also called an excitation). By analogy with current generators in the circuit theory, the excitation in Maxwell's equations can be modeled by the so-called impressed current, $\mathbf{J}_i$. This current is regarded as a known distribution within a certain spatial domain v. The current density can now be represented as

$$\mathbf{J} = \mathbf{J}_c + \mathbf{J}_i \tag{3.4}$$

where $\mathbf{J}_c$ is the component corresponding to the charge motion caused by the electric field in the conducting medium (referred to as the conduction current). $\mathbf{J}_c$ is a function of $\mathbf{E}$ (more rigorously, it also depends on $\mathbf{B}$, but we shall neglect this dependence) and of the conducting properties of the material (medium) in which the fields are observed, so we can write

$$\mathbf{J}_c = \mathbf{J}_c(\mathbf{E}) \tag{3.5}$$

$\mathbf{D}$ and $\mathbf{E}$ are related by the dielectric properties of the material, and $\mathbf{B}$ and $\mathbf{H}$ are related by the magnetic properties of the material; that is,

$$\mathbf{D} = \mathbf{D}(\mathbf{E}) \qquad \mathbf{B} = \mathbf{B}(\mathbf{H}) \tag{3.6a,b}$$

Equations (3.5) and (3.6) are referred to as the constitutive relations.

Generally, the medium can be linear or nonlinear. For a linear medium, the constitutive relations are

$$\mathbf{J}_c = \sigma\mathbf{E} \qquad \mathbf{D} = \epsilon\mathbf{E} \qquad \mathbf{B} = \mu\mathbf{H} \tag{3.7a-c}$$

where σ is the conductivity, ϵ the permittivity (dielectric constant), and μ the permeability (magnetic constant) of the medium. The permittivity can be expressed as $\epsilon = \epsilon_r \epsilon_0$, where ϵ_r is the relative permittivity of the medium and $\epsilon_0 \approx 8.8542 \cdot 10^{-12}$ F/m is the permittivity of a vacuum. Similarly, $\mu = \mu_r \mu_0$, where μ_r is the relative permeability of the medium and $\mu_0 = 4\pi \cdot 10^{-7}$ H/m is the permeability of a vacuum.

The medium can be homogeneous or inhomogeneous. A homogeneous medium has uniform properties everywhere, whereas the properties of an inhomogeneous medium depend on the spatial coordinates. The parameters σ, ϵ, and μ of a linear homogeneous medium do not depend on the spatial coordinates. For a linear inhomogeneous medium, these parameters do depend on the position.

The medium can also be classified as isotropic, meaning the properties do not depend on the direction of the field vectors, and anisotropic, meaning the properties do depend on the direction.[1] The constitutive relations (3.7) are actually valid for a linear isotropic medium. For a linear anisotropic medium, each Cartesian component of a vector on the left-hand side of (3.7) depends on all three Cartesian components of the vector on the right side.

In this book we shall primarily consider linear, piecewise-homogeneous, and isotropic media.

3.1.1 Maxwell's Equations in the Frequency Domain

A particularly important case for the theoretical analysis and for practical solutions is when the excitation ($\mathbf{J}_i$) is a sinusoidal function of time, referred to as a time-harmonic function. If the medium is linear everywhere, the induced current and charge, as well as all the corresponding field vectors, are also time-harmonic functions.

A time-harmonic scalar function (e.g., a current i in the circuit theory) can be written as

$$i(t) = I_m \cos(\omega t + \psi) \tag{3.8}$$

where I_m is the amplitude, ω the angular frequency ($\omega = 2\pi f$, where f is the frequency), and ψ is the initial phase. In circuit theory, to avoid dealing with

1. A more elaborate classification of the media properties can be found, for example, in [10].

the time dependence, we represent time-harmonic quantities by their complex (phasor) representatives. This reduces the derivatives and integrals with respect to time to algebraic multiplications and divisions, respectively. The analysis is carried out in the complex (frequency) domain, instead of in the time domain. Two definitions of phasors (I) are commonly used. The first definition is

$$i(t) = \mathrm{Re}\left\{ I\sqrt{2}e^{j\omega t} \right\} \qquad I = \frac{I_m}{\sqrt{2}} e^{j\psi} \tag{3.9a,b}$$

where j ($j = \sqrt{-1}$) is the imaginary unit. The module of phasor I ($I_m/\sqrt{2}$) equals the rms of the current $i(t)$. Hence, I in (3.9) is referred to as the complex rms value. By the second definition,

$$i(t) = \mathrm{Re}\left\{ Ie^{j\omega t} \right\} \qquad I = I_m e^{j\psi} \tag{3.10a,b}$$

Hence, I in (3.10) is referred to as the complex amplitude. The choice of (3.9a) or (3.10a) does not affect Maxwell's equations, but it does affect the relations for the energy and power. For example, consider a resistor of resistance R, whose phasor current is I. If the complex rms value is used, the average power of the resistor is given by $P = RII^* = R|I|^2$, where the asterisk denotes complex conjugate. If the complex amplitude is used, the power is evaluated as $P = RII^*/2 = R|I|^2/2$, In this book we shall utilize complex rms values.

A time-harmonic vector quantity is a more complicated issue than a scalar. Let us consider a vector quantity, $\mathbf{A}(t)$, which is a function of time. Let the Cartesian components (projections) of this vector be $A_x(t)$, $A_y(t)$, and $A_z(t)$, such that

$$\mathbf{A}(t) = A_x(t)\mathbf{i}_x + A_y(t)\mathbf{i}_y + A_z(t)\mathbf{i}_z \tag{3.11}$$

where $\mathbf{i}_x$, $\mathbf{i}_y$, and $\mathbf{i}_z$ are the unit vectors of the Cartesian coordinate system. The vector $\mathbf{A}(t)$ is a time-harmonic function if $A_x(t)$, $A_y(t)$, and $A_z(t)$ are time-harmonic functions of the same frequency; that is,

$$\begin{aligned} &A_x(t) = A_{xm}\cos(\omega t + \theta_x) \quad A_y(t) = A_{ym}\cos(\omega t + \theta_y) \\ &A_z(t) = A_{zm}\cos(\omega t + \theta_z) \end{aligned} \tag{3.12}$$

where the amplitudes (A_{xm}, A_{ym}, A_{zm}) and the initial phases (θ_x, θ_y, θ_z) can be arbitrary. Generally, both the magnitude and direction of $\mathbf{A}(t)$ periodically vary as a function of time, so that $\mathbf{A}(t)$ is an elliptically polarized vector. Special cases are the linear and the circular polarization.

For each component in (3.12), the phasor representative can be found (A_x, A_y, A_z, respectively), similar to (3.9a). Finally, the phasor representative of $\mathbf{A}(t)$, and the vector $\mathbf{A}(t)$ in terms of its phasor representative are obtained as

$$\mathbf{A} = A_x\mathbf{i}_x + A_y\mathbf{i}_y + A_z\mathbf{i}_z \qquad \mathbf{A}(t) = \mathrm{Re}\{\mathbf{A}\sqrt{2}\,e^{-j\omega t}\} \qquad (3.13a,b)$$

For simplicity, we shall not introduce a special notation for phasor scalars and vectors to distinguish them from time-domain quantities. This should not be confusing, because further in this book, we shall deal only with the frequency-domain analysis.

As already mentioned, the key advantage of introducing the phasor representation is to facilitate dealing with derivatives with respect to time. The derivative in the time domain is replaced by the multiplication by $j\omega$ in the complex domain. Hence, the Maxwell's equations (3.3a–e) in the complex domain read:

$$\nabla \times \mathbf{E} = -j\omega\mathbf{B} \quad \nabla \times \mathbf{H} = \mathbf{J} + j\omega\mathbf{D} \quad \nabla \cdot \mathbf{D} = \rho \qquad (3.14a\text{-}c)$$

$$\nabla \cdot \mathbf{B} = 0 \quad \nabla \cdot \mathbf{J} = -j\omega\rho \qquad (3.14d,e)$$

where now $\mathbf{E}$, $\mathbf{D}$, $\mathbf{B}$, $\mathbf{H}$, $\mathbf{J}$, and ρ are phasors.

From (3.14e) the charge density can be expressed in terms of the current density. The current density can be separated into the impressed and conduction component according to (3.4). We can associate a part of the charge density to each of the two current components by

$$\nabla \cdot \mathbf{J}_i = -j\omega\rho_i \qquad \nabla \cdot \mathbf{J}_c = -j\omega\rho_c \qquad (3.15a,b)$$

respectively, such that

$$\rho = \rho_i + \rho_c \qquad (3.16)$$

In isotropic media, the constitutive equations in phasor notation are formally identical with (3.7). However, all vector quantities in these equations are phasors now. The permittivity and permeability are also phasors. This corresponds to describing dielectric and magnetic losses by elliptic hysteresis loops. The complex permittivity is usually written as $\epsilon = \epsilon' - j\epsilon''$, where ϵ' and ϵ'' are real positive quantities. The dielectric loss tangent is defined as $\tan\delta_e = \epsilon''/\epsilon'$. Similarly, the complex permeability is written as $\mu = \mu' - j\mu''$ and the magnetic loss tangent is $\tan\delta_m = \mu''/\mu'$.

For a homogeneous medium, two possibilities for simplifying the system (3.14) are often used. The first option is to substitute the constitutive equations (3.7b) and (3.7c) into (3.14a) through (3.14d), yielding

$$\nabla \times \mathbf{E} = -j\omega\mu\mathbf{H} \quad \nabla \times \mathbf{H} = \mathbf{J} + j\omega\epsilon\mathbf{E} \quad \nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon} \quad \nabla \cdot \mathbf{H} = 0 \tag{3.17a-d}$$

The second option is to substitute all three constitutive equations (3.7a) through (3.7c) as explained below.

The current density, $\mathbf{J}$, can formally be separated according to (3.4). By combining (3.7a) and (3.7b) with (3.17b), we obtain

$$\begin{aligned} \nabla \times \mathbf{H} &= \mathbf{J}_i + \sigma\mathbf{E} + j\omega(\epsilon' - j\epsilon'')\mathbf{E} = \mathbf{J}_i + (\sigma + \omega\epsilon'')\mathbf{E} + j\omega\epsilon'\mathbf{E} \\ &= \mathbf{J}_i + j\omega[\epsilon' - j(\sigma/\omega + \epsilon'')]\mathbf{E} \end{aligned} \tag{3.18}$$

At high frequencies it is hard to distinguish between the conducting (Joule) and dielectric losses in a medium. For practical purposes, the loss mechanism is usually irrelevant. Equation (3.18) opens two possibilities for describing these two kinds of looses. The first possibility is to use the equivalent conductivity, $\sigma + \omega\epsilon''$, instead of σ. The second possibility is to use the equivalent complex permittivity, ϵ_e, whose real part is ϵ'; the negative of the imaginary part is $\frac{\sigma}{\omega} + \epsilon''$. In this book, we shall predominantly use the equivalent permittivity, shall not introduce a particular symbol for it, except in this section.

Next, we start from the relation

$$\mathbf{D} = \frac{\epsilon}{\sigma}\mathbf{J}_c \tag{3.19}$$

Taking the divergence of both sides of (3.19) and using (3.14c) and (3.14e) to (3.16) result in

$$\rho_i + \rho_c = -j\omega\frac{\epsilon}{\sigma}\rho_c \qquad \rho = \frac{\epsilon}{\epsilon_e}\rho_i \tag{3.20a,b}$$

which can be substituted into (3.17c). With two Maxwell's equations modified, the resulting system can be written in the final form

$$\nabla \times \mathbf{E} = -j\omega\mu\mathbf{H} \quad \nabla \times \mathbf{H} = \mathbf{J}_i + j\omega\epsilon_e\mathbf{E} \quad \nabla \cdot \mathbf{E} = \frac{\rho_i}{\epsilon_e} \quad \nabla \cdot \mathbf{H} = 0 \tag{3.21a-d}$$

In the resulting system of Maxwell's equations (3.21), there appear the impressed current and the associated charge, both assumed to be known quantities, while the permittivity is replaced by the equivalent permittivity. Note that the systems (3.17) and (3.21) are formally identical.

In this book, we shall use the systems (3.17) and (3.21) interchangeably. Wherever necessary, we shall stress whether the conductive current is included in the displacement current ($j\omega\epsilon_e\mathbf{E}$), resulting in (3.21), or the conductive and impressed currents are considered together, resulting in (3.17). Therefore, dropping the subscript "e" for the equivalent complex permittivity further in this book should not be confusing.

3.2 Retarded Potentials

We consider the electromagnetic field in an unbounded region, filled with a linear, homogeneous, and isotropic medium, of parameters ϵ and μ (which include all kinds of losses). The field vectors can be expressed in terms of the current density, $\mathbf{J}$.

Maxwell's equations (3.17a) and (3.17d) are automatically satisfied if the electric and magnetic field are expressed as

$$\mathbf{E} = -\nabla V_e - j\omega\mathbf{A}_m \qquad \mathbf{H} = \frac{1}{\mu}\nabla \times \mathbf{A}_m \tag{3.22a,b}$$

where V_e and $\mathbf{A}_m$ are auxiliary functions referred to as the electric scalar potential and magnetic vector potential, respectively. After substituting (3.22) into Maxwell's equations (3.17b) and (3.17c), the following two differential equations for the potentials are obtained:

$$\nabla(\nabla \cdot \mathbf{A}_m) - \Delta\mathbf{A}_m = \mu\mathbf{J} - j\omega\epsilon\mu\nabla V_e + \omega^2\epsilon\mu\mathbf{A}_m \quad -\Delta V_e - j\omega\nabla \cdot \mathbf{A}_m = \frac{\rho}{\epsilon} \tag{3.23a,b}$$

We assume the current and charge densities to be known. The problem of solving the four Maxwell's equations for the field vectors is replaced by the problem of solving (3.23a) and (3.23b) for the potentials in terms of the current and charge.

The differential equations (3.23a) and (3.23b) do not give unique solutions for the potentials unless an additional constraint is introduced for the divergence of the magnetic vector potential, referred to as the gauge. We shall use the Lorentz condition:

$$\nabla \cdot \mathbf{A}_m = -j\omega\epsilon\mu V_e \tag{3.24}$$

In that case, (3.23a) and (3.23b) are reduced to a pair of decoupled wave equations:

$$\Delta\mathbf{A}_m - \gamma^2\mathbf{A}_m = -\mu\mathbf{J} \qquad \Delta V_e - \gamma^2 V_e = -\frac{\rho}{\epsilon} \tag{3.25a,b}$$

where $\gamma = j\omega\sqrt{\epsilon\mu}$ is referred to as the propagation constant. In a lossy medium, the permittivity and permeability are complex, so that γ is complex. The real part of γ is the attenuation constant, and the imaginary part is the phase constant.

According to the notation in Figure 3.1, the solutions to (3.25) are

$$\mathbf{A}_m(\mathbf{r}) = \mu \int_v \mathbf{J}(\mathbf{r}')g(R)\,dv' \qquad V_e(\mathbf{r}) = \frac{1}{\epsilon}\int_v \rho(\mathbf{r}')g(R)\,dv' \tag{3.26a,b}$$

where $\mathbf{r}$ is the position vector of the field point (i.e., the point at which the potentials and fields are evaluated), $\mathbf{r}'$ is the position vector of the source point (i.e., the point at which the current and charge are located), v is the source region (i.e., the region occupied by the current and charge), the Green's function for the unbounded region is

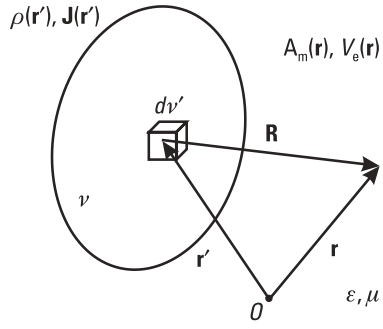

Figure 3.1 Coordinate system for the evaluation of the retarded potentials.

$$g(R) = \frac{e^{-\gamma R}}{4\pi R} \qquad R = |\mathbf{R}| \qquad \mathbf{R} = \mathbf{r} - \mathbf{r}' \tag{3.27a-c}$$

and R is the distance between the field and the source points. The potentials given by (3.26) are referred to as the Lorentz or retarded potentials.

3.3 Field Vectors

3.3.1 Basic Integral Expressions

We obtain integral expressions for the electric and magnetic field by substituting (3.26) into (3.22). The integration in (3.26) is performed with respect to the position of the source point ($\mathbf{r}'$ coordinate), while the differentiation in (3.22) is performed with respect to the position of the field point ($\mathbf{r}$ coordinate). Hence, in ∇V_e and $\nabla \times \mathbf{A}_m$, the order of differentiation and integration can be reversed, yielding

$$\nabla V_e(\mathbf{r}) = \frac{1}{\epsilon} \int_v \nabla[\rho(\mathbf{r}')g(R)]\, dv' \qquad \nabla \times \mathbf{A}_m(\mathbf{r}) = \mu \int_v \nabla \times [\mathbf{J}(\mathbf{r}')g(R)]\, dv' \tag{3.28a,b}$$

Since $\rho(\mathbf{r}')$ is a scalar constant for the del operator, the integrand in (3.28a) is transformed to $\rho(\mathbf{r}')\nabla g(R)$. Using (3.14e), the charge density, $\rho(\mathbf{r}')$, is expressed in terms of the current density, $\mathbf{J}(\mathbf{r}')$. Similarly, in (3.28b) $\mathbf{J}(\mathbf{r}')$ is a vector constant for the del operator, and the integrand is transformed to $\nabla g(R) \times \mathbf{J}(\mathbf{r}')$. After some rearrangements we obtain the final integral expressions for the electric and magnetic field vectors in the form

$$\mathbf{E}(\mathbf{r}) = -\gamma Z \int_v [\mathbf{J}(\mathbf{r}')g(R) - \frac{1}{\gamma^2}\nabla' \cdot \mathbf{J}(\mathbf{r}')\nabla g(R)]\, dv' \tag{3.29a}$$

$$\mathbf{H}(\mathbf{r}) = -\int_v \mathbf{J}(\mathbf{r}') \times \nabla g(R)\, dv' \tag{3.29b}$$

where ∇' is the del operator acting on the source coordinates ($\mathbf{r}'$), and $Z = \sqrt{\mu/\epsilon}$ is the intrinsic (wave) impedance of the medium. The integral expressions (3.29a) and (3.29b) are written in compact form, which may be

convenient for theoretical considerations. However, to actually perform any computations, Green's function and its gradient should be explicitly expressed in terms of R, as shown below.

The Green's-function gradient is easily determined in the Cartesian coordinate system. The position vectors $\mathbf{r}$ and $\mathbf{r}'$ are expressed in terms of the coordinates x, y, and z, and x', y', and z', respectively. According to (3.27b) and (3.27c), the distance R is given by

$$R = \sqrt{(x - x')^2 + (y - y')^2 + (z - z')^2} \tag{3.30}$$

After applying the gradient in the Cartesian coordinate system to an arbitrary function of R, $g(R)$, we obtain

$$\nabla g(R) = \frac{dg(R)}{dR} \nabla R \qquad \nabla R = \mathbf{i}_R \qquad \mathbf{i}_R = \frac{\mathbf{R}}{R} \tag{3.31a-c}$$

Finally, after differentiation of (3.27a) by R, the Green's-function gradient is evaluated as

$$\nabla g(R) = -\frac{\gamma^2}{4\pi} \left[\frac{1}{\gamma R} + \frac{1}{(\gamma R)^2} \right] e^{-\gamma R} \mathbf{i}_R \tag{3.32}$$

After substituting (3.27a) and (3.32) into (3.29), the integral expressions for the electric and magnetic fields are finally obtained as

$$\mathbf{E}(\mathbf{r}) = -\frac{\gamma^2}{4\pi} Z \int_v \left\{ \mathbf{J}(\mathbf{r}') \frac{1}{\gamma R} + \frac{\nabla' \cdot \mathbf{J}(\mathbf{r}')}{\gamma} \mathbf{i}_R \left[\frac{1}{\gamma R} + \frac{1}{(\gamma R)^2} \right] \right\} e^{-\gamma R} dv' \tag{3.33a}$$

$$\mathbf{H}(\mathbf{r}) = \frac{\gamma^2}{4\pi} \int_v \mathbf{J}(\mathbf{r}') \times \mathbf{i}_R \left[\frac{1}{\gamma R} + \frac{1}{(\gamma R)^2} \right] e^{-\gamma R} dv' \tag{3.33b}$$

Equation (3.33b) is applicable to an arbitrary current distribution. Equation (3.33a), however, can be directly applied only to differentiable current distributions, that is, distributions whose divergence can be evaluated. In other words, this equation is applicable to a volume current accompanied by a volume charge.

The integral expressions for the field vectors derived in this section are referred to as the basic expressions because they will be used predominantly in this book.

3.3.2 Alternative Expressions for the Electric Field

An alternative integral expression for the electric field can be obtained if the electric scalar potential in (3.22a) is substituted in terms of the magnetic vector potential by using the Lorentz condition (3.24). This results in an electric field expressed only in terms of the magnetic vector potential as

$$\mathbf{E} = -j\omega\left[\mathbf{A}_m - \frac{1}{\gamma^2}\nabla(\nabla \cdot \mathbf{A}_m)\right] \tag{3.34}$$

The magnetic vector potential is again given by (3.26a). The order of the differentiation and integration in $\nabla(\nabla \cdot \mathbf{A}_m)$ can be reversed; that is,

$$\nabla(\nabla \cdot \mathbf{A}_m) = \mu \int_v \nabla\{\nabla \cdot [\mathbf{J}(\mathbf{r}')g(R)]\}\, dv' \tag{3.35}$$

Since in (3.35) $\mathbf{J}(\mathbf{r}')$ is a vector constant for the del operator, the integrand is transformed into $\nabla[\mathbf{J}(\mathbf{r}') \cdot \nabla g(R)]$. After some rearrangements, we obtain the electric field in the form

$$\mathbf{E}(\mathbf{r}) = -\gamma Z \int_v \left\{\mathbf{J}(\mathbf{r}')g(R) - \frac{1}{\gamma^2}\nabla[\mathbf{J}(\mathbf{r}') \cdot \nabla g(R)]\right\} dv' \tag{3.36}$$

To prepare this equation for the numerical analysis, Green's function and its derivatives should be explicitly expressed in terms of R. Bearing in mind (3.31), the gradient of $\mathbf{J}(\mathbf{r}') \cdot \nabla g(R)$ can be transformed as

$$\begin{aligned}\nabla[\mathbf{J}(\mathbf{r}') \cdot \nabla g(R)] &= \nabla\left\{[\mathbf{J}(\mathbf{r}') \cdot \mathbf{R}]\frac{1}{R}\frac{dg(R)}{dR}\right\} \\ &= \frac{1}{R}\frac{dg(R)}{dR}\nabla[\mathbf{J}(\mathbf{r}') \cdot \mathbf{R}] + [\mathbf{J}(\mathbf{r}') \cdot \mathbf{R}]\nabla\left(\frac{1}{R}\frac{dg(R)}{dR}\right)\end{aligned} \tag{3.37}$$

Applying the expression for the gradient in the Cartesian coordinate system to $\mathbf{J}(\mathbf{r}') \cdot \mathbf{R}$, we easily prove that $\nabla[\mathbf{J}(\mathbf{r}') \cdot \mathbf{R}] = \mathbf{J}(\mathbf{r}')$. The gradient in the

last term on the right-hand side of (3.37) is evaluated according to (3.31), where $g(R)$ is replaced by $dg(R)/dR/R$. Finally, after some manipulations the electric field is obtained as

$$\mathbf{E}(\mathbf{r}) = -\frac{\gamma^2}{4\pi} Z \int_v \left\{ \mathbf{J}_T(\mathbf{r}') \left[\frac{1}{\gamma R} + \frac{1}{(\gamma R)^2} + \frac{1}{(\gamma R)^3} \right] - 2\mathbf{J}_R(\mathbf{r}') \left[\frac{1}{(\gamma R)^2} + \frac{1}{(\gamma R)^3} \right] \right\} e^{-\gamma R}\, dv' \tag{3.38}$$

where $\mathbf{J}_R(\mathbf{r}')$ and $\mathbf{J}_T(\mathbf{r}')$ are components of $\mathbf{J}$ that are collinear and perpendicular, respectively, to the unit vector, $\mathbf{i}_R$, directed from the source to the field point; that is,

$$\mathbf{J}_R = (\mathbf{J} \cdot \mathbf{i}_R)\mathbf{i}_R \qquad \mathbf{J}_T = \mathbf{J} - (\mathbf{J} \cdot \mathbf{i}_R)\mathbf{i}_R = \mathbf{i}_R \times (\mathbf{J} \times \mathbf{i}_R) \tag{3.39a,b}$$

The advantage of (3.38) over (3.33a) is that (3.38) can be applied to an arbitrary current distribution, which need not be derivable. The disadvantage of (3.38) is that it contains terms proportional to $1/R$, $1/R^2$, and $1/R^3$, whereas (3.33a) contains terms proportional only to $1/R$ and $1/R^2$.

The problem occurs when the field point is inside the source region, v, because $1/R^i$, $i = 1, 2, 3$, tends to infinity when $\mathbf{r}'$ tends to $\mathbf{r}$. In other words, $1/R^i$ is singular at $\mathbf{r}' = \mathbf{r}$. When the field point is outside the source region, but very close to it, the functions $1/R^i$ change very rapidly when $\mathbf{r}'$ approaches $\mathbf{r}$, reaching large values. We refer to those functions as quasisingular functions. We shall refer to the largest exponent i in an expression as the singularity order of that expression. The numerical integration of singular and quasisingular functions is a hard task. Increasing the singularity order makes the numerical integration more difficult. Hence, it is desirable that integral expressions for the field vectors contain terms with the lowest possible order of singularity. For (3.38), the singularity order is $i = 3$, while for (3.33a), this order is $i = 2$.

Starting from (3.29a), an integral expression for the electric field with the singularity order of only $i = 1$ is obtained in the following manner. Bearing in mind that $\nabla' R = -\nabla R$ and $\nabla' g(R) = -\nabla g(R)$, the second term in brackets on the right side of (3.29a) is transformed to

$$\begin{aligned} \nabla' \cdot \mathbf{J}(\mathbf{r}')\nabla g(R) &= -\nabla' \cdot \mathbf{J}(\mathbf{r}')\nabla' g(R) \\ &= \nabla'[\nabla' \cdot \mathbf{J}(\mathbf{r}')]g(R) - \nabla'[\nabla' \cdot \mathbf{J}(\mathbf{r}')g(R)] \end{aligned} \tag{3.40}$$

According to the divergence theorem, the integral of the second term on the right side of (3.40) over the source region v can be represented as

$$\int_v \nabla'[\nabla' \cdot \mathbf{J}(\mathbf{r}')g(R)]\,dv' = \oint_S \nabla' \cdot \mathbf{J}(\mathbf{r}')g(R)\,d\mathbf{S}' \tag{3.41}$$

where $d\mathbf{S}' = \mathbf{n}dS'$, S is the surface bounding the source region v, dS' is an element of this surface, and $\mathbf{n}$ is the outward unit vector on S. After substituting (3.40) and (3.41) into (3.29a), the electric field is written as

$$\mathbf{E}(\mathbf{r}) = -\gamma Z \int_v \left\{ \mathbf{J}(\mathbf{r}') - \frac{1}{\gamma^2} \nabla [\nabla' \cdot \mathbf{J}(\mathbf{r}')] \right\} g(R)\,dv' \tag{3.42}$$

$$- \frac{Z}{\gamma} \oint_S \nabla' \cdot \mathbf{J}(\mathbf{r}')g(R)\,d\mathbf{S}'$$

Taking into account the expression for Green's function (3.27a), it is obvious that (3.42) contains only terms proportional to $1/R$. However, compared with the two previous integral expressions for the electric field, (3.42) contains a surface integral in addition to a volume integral, which somewhat complicates the numerical evaluation. Equation (3.42) can be applied only if the charge density is a derivable function of coordinates, that is, if the current density is a twice-derivable function.

Depending on the application, one of the equations (3.33a), (3.38), and (3.42) may be advantageous over the remaining two. For example, in the numerical analysis of wire antennas, these expressions are involved in the derivation of the so-called two-potential equation, Pocklington's equation, and Schelkunoff's equation, respectively. Pocklington's equation is hardest for the numerical evaluation of integrals, but it does not require the approximation for the current distribution to be derivable. The integrals in Schelkunoff's equation, on the other hand, are easiest for the numerical evaluation, but the approximation for the current must be a twice derivable function of the coordinate along the wire, and the equation is convenient practically only for parallel wires. The two-potential equation is between the other two equations, and it has found the widest application. A more elaborate comparison of these three equations can be found in [12]. Similarly, in this book, (3.33a) is used predominantly in the implementation of integral equations for analyzing arbitrary composite metallic and dielectric structures, as is explained in Chapter 8.

3.3.3 Definition of *L* and *K* Operators

As a compilation of the integral expressions (3.29), (3.33), (3.36), (3.38), and (3.42), the electric and magnetic fields can be compactly written in operator form[2] as

$$\mathbf{E} = -ZL(\mathbf{J}) \qquad \mathbf{H} = -K(\mathbf{J}) \tag{3.43a,b}$$

There are three different expressions for the L operator and one for the K operator. Each expression is usually written in two forms: in terms of the Green's function and its derivatives, and in terms of the distance R. The first form is convenient for theoretical considerations, while the second form is necessary for numerical evaluations. All these possibilities are summarized in Table 3.1. For each expression, Table 3.1 also shows the singularity order.

3.3.4 Expressions for Fields Due to Surface Current

Starting from the integral expressions for the fields due to a volume current, we easily obtain integral expressions for fields due to a surface current. The volume current density, $\mathbf{J}$, the volume element dv, and the volume del operator, ∇', should be replaced by the surface current density, $\mathbf{J}_s$, the surface element dS, and the surface del operator, ∇'_s, respectively. Let us consider the operators L (the first form, which has the singularity order $i = 2$) and K, given in Table 3.1. When applied to the surface current, these operators are written in the form

$$L(\mathbf{J}_s) = \gamma \int_S \left[\mathbf{J}_s(\mathbf{r}')g(R) - \frac{1}{\gamma^2}\nabla'_s \cdot \mathbf{J}_s(\mathbf{r}')\nabla g(R) \right] dS' \tag{3.44a}$$

$$= \frac{\gamma^2}{4\pi} \int_S \left\{ \mathbf{J}_s(\mathbf{r}')\frac{1}{\gamma R} + \frac{\nabla'_s \cdot \mathbf{J}_s(\mathbf{r}')}{\gamma}\mathbf{i}_R \left[\frac{1}{\gamma R} + \frac{1}{(\gamma R)^2} \right] \right\} e^{-\gamma R} dS'$$

$$K(\mathbf{J}_s) = \int_S \mathbf{J}_s(\mathbf{r}') \times \nabla g(R)\, dS' = -\frac{\gamma^2}{4\pi} \int_S \mathbf{J}_s(\mathbf{r}') \times \mathbf{i}_R \left[\frac{1}{\gamma R} + \frac{1}{(\gamma R)^2} \right] e^{-\gamma R} dS' \tag{3.44b}$$

2. In the literature, the intrinsic impedance is usually included in the operator L. Equation (3.43), however, results in dual expressions for the electric and magnetic fields in terms of the electric and magnetic currents.

Table 3.1
Summary of Expressions for the Operators L and K Written in Terms of Green's Function and Its Derivatives and in Terms of the Distance R

	$g(R)$	R
$L(\mathbf{J})$ $(i = 2)$	$\gamma \int_v \left[\mathbf{J}(\mathbf{r}')\, g(R) - \frac{1}{\gamma^2} \nabla' \cdot \mathbf{J}(\mathbf{r}') \nabla g(R) \right] dv'$	$\frac{\gamma^2}{4\pi} \int_v \left\{ \mathbf{J}(\mathbf{r}') \frac{1}{\gamma R} + \frac{\nabla' \cdot \mathbf{J}(\mathbf{r}')}{\gamma} \mathbf{i}_R \left[\frac{1}{\gamma R} + \frac{1}{(\gamma R)^2} \right] \right\} e^{-\gamma R} dv'$
$L(\mathbf{J})$ $(i = 3)$	$\gamma \int_v \left\{ \mathbf{J}(\mathbf{r}')\, g(R) - \frac{1}{\gamma^2} \nabla [\mathbf{J}(\mathbf{r}') \cdot \nabla g(R)] \right\} dv'$	$\frac{\gamma^2}{4\pi} \int_v \left\{ \mathbf{J}_T(\mathbf{r}') \left[\frac{1}{\gamma R} + \frac{1}{(\gamma R)^2} + \frac{1}{(\gamma R)^3} \right] - 2\mathbf{J}_R(\mathbf{r}') \left[\frac{1}{(\gamma R)^2} + \frac{1}{(\gamma R)^3} \right] \right\} e^{-\gamma R} dv'$ $\mathbf{J}_R = (\mathbf{J} \cdot \mathbf{i}_R)\mathbf{i}_R \qquad \mathbf{J}_T = \mathbf{i}_R \times (\mathbf{J} \times \mathbf{i}_R)$
$L(\mathbf{J})$ $(i = 1)$	$\gamma \int_v \left\{ \mathbf{J}(\mathbf{r}') - \frac{1}{\gamma^2} \nabla [\nabla' \cdot \mathbf{J}(\mathbf{r}')] \right\} g(R)\, dv' + \frac{1}{\gamma} \oint_S \nabla' \cdot \mathbf{J}(\mathbf{r}')\, g(R)\, d\mathbf{S}'$	$\frac{\gamma^2}{4\pi} \int_v \left\{ \mathbf{J}(\mathbf{r}') - \frac{1}{\gamma^2} \nabla [\nabla' \cdot \mathbf{J}(\mathbf{r}')] \right\} \frac{1}{\gamma R} e^{-\gamma R} dv' + \frac{1}{4\pi} \oint_S \nabla' \cdot \mathbf{J}(\mathbf{r}') \frac{1}{\gamma R} e^{-\gamma R} d\mathbf{S}'$
$K(\mathbf{J})$ $(i = 2)$	$\int_v \mathbf{J}(\mathbf{r}') \times \nabla g(R)\, dv'$	$-\frac{\gamma^2}{4\pi} \int_v \mathbf{J}(\mathbf{r}') \times \mathbf{i}_R \left[\frac{1}{\gamma R} + \frac{1}{(\gamma R)^2} \right] e^{-\gamma R} dv'$

Note: The index i is the singularity order.

A delicate case is when the field point is placed on the surface S, as the integrands are singular. When $\mathbf{r}'$ approaches $\mathbf{r}$, the distance R tends to zero, leading to infinite values of the terms $1/R^i$, $i = 1, 2, 3$. Further, when $\mathbf{r}' = \mathbf{r}$, the unit vector of R, $\mathbf{i}_R$, is undefined. Although, physically, the fields must be finite on a smooth part of the surface S, both the electric and the magnetic field are discontinuous at the surface, that is, $\mathbf{E}$ and $\mathbf{H}$ have a jump at the surface. Hence, the results of (3.44) depend on the assumed location of the field point, that is, if it is on one face of the surface, on the other face, or, mathematically, right on the surface. We may visualize the last case by assuming the surface to have the form of a very thin shell and taking the field point to be right in the middle of the shell thickness.

To further clarify this issue, let us subdivide the surface S into two parts, a circular principal area with a small radius, δ, centered at the field point ($\mathbf{r}$), denoted by S_δ, and the remainder, denoted by $S \setminus S_\delta$ [Figure 3.2(a)]. If in the integration we exclude the principal area and let δ approach zero, we evaluate the integrals in the principal value sense. This corresponds to assuming the field point to be right on the surface (i.e., in the middle of the shell thickness). For this case, we shall say that the operators L and K are applied in the principal value sense and denote them by L_0 and K_0, respectively.

We shall further denote the integrals over the surface S_δ by $L_\delta(\mathbf{J}_s)$ and $K_\delta(\mathbf{J}_s)$. If the radius δ is small enough, the terms proportional to $1/R$ can be neglected and the current and charge densities on S_δ can be assumed practically uniform. In that case, we also have $e^{-\gamma R} \approx 1$ (i.e., the retardation can be neglected due to the small distance between the source and field points), and the integrals $L_\delta(\mathbf{J}_s)$ and $K_\delta(\mathbf{J}_s)$ can be written in the form

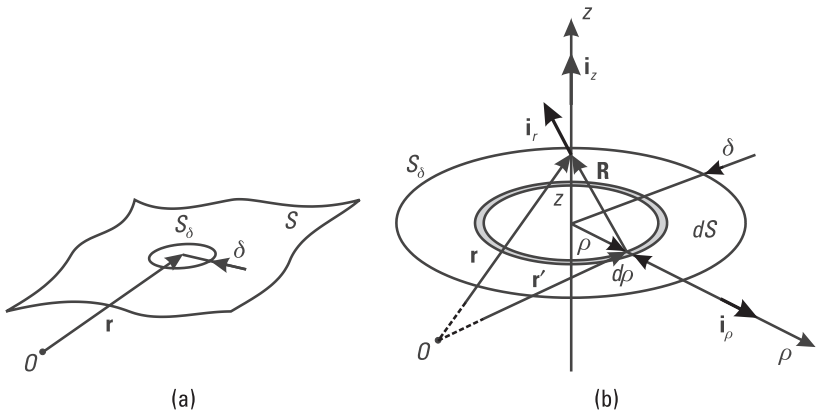

Figure 3.2 Division of the source surface when the field point is on the surface: (a).

$$L_\delta(\mathbf{J}_s) = \frac{1}{4\pi}\frac{\nabla'_s \cdot \mathbf{J}_s(\mathbf{r}')}{\gamma}\int\limits_{S_\delta}\frac{\mathbf{i}_R}{R^2}\,dS' \qquad K_\delta(\mathbf{J}_s) = -\frac{1}{4\pi}\mathbf{J}_s(\mathbf{r}')\times\int\limits_{S_\delta}\frac{\mathbf{i}_R}{R^2}\,dS' \tag{3.45a,b}$$

which can be interpreted as quasistatic fields of a circular disk (carrying a uniform charge and a uniform current). In both cases, we have to evaluate the same integral.

Let us adopt a local cylindrical coordinate system centered at the field point with the z-axis perpendicular to the surface S_δ as shown in Figure 3.2(b). Let the field point be on the z-axis. The vector $\mathbf{R}$, its modulus, R, and the unit vector, $\mathbf{i}_R$, can be expressed respectively, in the cylindrical coordinate system, as

$$\mathbf{R} = -\rho\mathbf{i}_\rho + z\mathbf{i}_z \qquad R = \sqrt{\rho^2 + z^2} \qquad \mathbf{i}_R = -\frac{\rho}{R}\mathbf{i}_\rho + \frac{z}{R}\mathbf{i}_z \tag{3.46a-c}$$

where ρ is the radial coordinate (not to be confused with the symbol for the charge density), and $\mathbf{i}_\rho$ is the corresponding unit vector. Due to the rotational symmetry, the result of the integration in (3.45) has only the z-component. Since the derivative of R with respect to ρ is $dR/d\rho = \rho/R$, the surface element can be expressed as $dS = 2\pi\rho d\rho = 2\pi R\,dR$. So, the integral can be evaluated for arbitrary z as

$$\int\limits_{S_\delta}\frac{\mathbf{i}_R}{R^2}\,dS = 2\pi z\left(\int\limits_0^\delta\frac{dR}{R^2}\right)\mathbf{i}_z = 2\pi\left(\frac{z}{|z|} - \frac{z}{\sqrt{\delta^2 + z^2}}\right)\mathbf{i}_z \tag{3.47}$$

If z approaches zero from positive values, (i.e., from the upper side of the surface S_δ), the integral in (3.47) tends to $+2\pi\mathbf{i}_z$. If z approaches zero from the negative values (i.e., from the lower side of the surface), the integral tends to $-2\pi\mathbf{i}_z$. Hence, when $z \to \pm 0$, we obtain

$$\lim_{z\to\pm 0}[L_\delta(\mathbf{J}_s)] = \mp\frac{1}{2}c\rho_s\mathbf{i}_z \qquad \lim_{z\to\pm 0}[K_\delta(\mathbf{J}_s)] = \mp\frac{1}{2}\mathbf{J}_s\times\mathbf{i}_z \tag{3.48}$$

where the surface charge density, ρ_s, is related to the surface current density, $\mathbf{J}_s$, through the continuity equation, specialized for surfaces,

$$\rho_s = \frac{j}{\omega} \nabla_s \cdot \mathbf{J}_s \tag{3.49}$$

We can interpret (3.48) as the operators $L_\delta(\mathbf{J}_s)$ and $K_\delta(\mathbf{J}_s)$ evaluated at points $z = 0^+$ (upper signs) and $z = 0^-$ (lower signs). If, however, we take the field point to be at $z = 0$, we have $L_\delta(\mathbf{J}_s) = 0$ and $K_\delta(\mathbf{J}_s) = 0$.

To summarize, when $L_\delta(\mathbf{J}_s)$ and $K_\delta(\mathbf{J}_s)$ are added to $L_0(\mathbf{J}_s)$ and $K_0(\mathbf{J}_s)$, respectively, we obtain the total fields at any point. When the field point is away from the surface S, the contributions of $L_\delta(\mathbf{J}_s)$ and $K_\delta(\mathbf{J}_s)$ are negligibly small, and the fields due to the surface current are given by

$$\mathbf{E} = -ZL(\mathbf{J}_s) \qquad \mathbf{H} = -K(\mathbf{J}_s) \tag{3.50}$$

where L and K denote the operators given by (3.44). If the field point is placed on a face of the surface S, the operators L and K can be written as

$$L(\mathbf{J}_s) = \begin{cases} L_0(\mathbf{J}_s) - \dfrac{1}{2} c\rho_s \mathbf{n} & z = 0^+ \\ L_0(\mathbf{J}_s) & z = 0 \\ L_0(\mathbf{J}_s) + \dfrac{1}{2} c\rho_s \mathbf{n} & z = 0^- \end{cases} \tag{3.51}$$

$$K(\mathbf{J}_s) = \begin{cases} K_0(\mathbf{J}_s) - \dfrac{1}{2} \mathbf{J}_s \times \mathbf{n} & z = 0^+ \\ K_0(\mathbf{J}_s) & z = 0 \\ K_0(\mathbf{J}_s) + \dfrac{1}{2} \mathbf{J}_s \times \mathbf{n} & z = 0^- \end{cases}$$

where L_0 and K_0 denote the operators K and L given by (3.44) applied in the principal value sense, while $\mathbf{n}$ is the unit normal on S (which originates from the positive face of the surface S).

3.3.5 Far-Field Expressions

Consider a field point far away from the source region. We assume the coordinate origin is located within the source region or close to it. If the distance of the field point from the source point, R, is much greater than the operating wavelength, the terms proportional to $1/R^2$ and $1/R^3$ in (3.38) and (3.33b) can be neglected. After some rearrangements, the fields can be expressed as

$$\mathbf{E}(\mathbf{r}) = -Z\frac{\gamma}{4\pi}\int_v \mathbf{J}_T(\mathbf{r}')\frac{e^{-\gamma R}}{R}dv' \qquad \mathbf{H}(\mathbf{r}) = \frac{\gamma}{4\pi}\int_v \mathbf{J}(\mathbf{r}')\times\mathbf{i}_R\frac{e^{-\gamma R}}{R}dv' \tag{3.52}$$

If the distance R is much greater than the maximal dimension of the source region, the vector $\mathbf{R}$ is practically parallel to the position vector of the field point, $\mathbf{r}$ (Figure 3.3). Equations (3.27b,c) can be approximated by $R = r - \mathbf{r}' \cdot \mathbf{i}_r$, where $r = |\mathbf{r}|$ and $\mathbf{i}_r = \mathbf{r}/r$ is the unit vector of the field-point direction, which coincides with the unit vector $\mathbf{i}_R$. The term $1/R$ in (3.52) can be approximated by the term $1/r$. However, the approximation must not be applied to the term $e^{-\gamma R}$, as it would introduce a significant phase error. Finally, after some rearrangements, the far fields can be written as

$$\mathbf{E}(\mathbf{r}) = -Z\mathbf{i}_r \times \left\{ \frac{\gamma e^{-\gamma r}}{4\pi r}\left[\int_v \mathbf{J}(\mathbf{r}')\, e^{\gamma \mathbf{r}'\cdot\mathbf{i}_r} dv'\right]\times\mathbf{i}_r \right\} \tag{3.53}$$

$$\mathbf{H}(\mathbf{r}) = \frac{\gamma e^{-\gamma r}}{4\pi r}\left[\int_v \mathbf{J}(\mathbf{r}')\, e^{\gamma \mathbf{r}'\cdot\mathbf{i}_r} dv'\right]\times\mathbf{i}_r$$

3.4 Volume Equivalence Principle

In Section 3.3 it was shown that field vectors in a linear, homogenous, and isotropic medium can be expressed in closed integral forms in terms of the

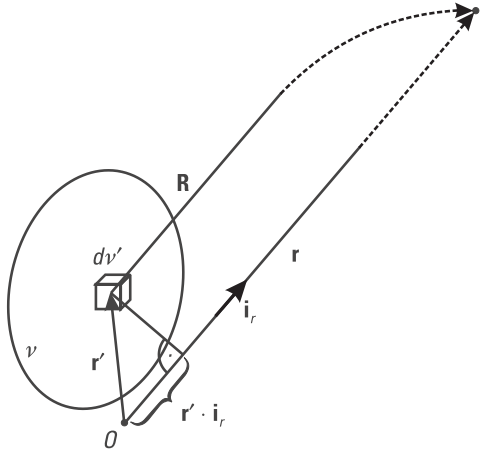

Figure 3.3 Coordinate system for the evaluation of far fields.

electric current ($\mathbf{J}$). These expressions can also be applied to inhomogenous and anisotropic media by artificially making them homogeneous and isotropic.

According to (3.17b) and (3.18), we have two options as to where to include the conduction current, $\mathbf{J}_c = \sigma\mathbf{E}$, mathematically. One option is to add it to the impressed current $\mathbf{J}_i$. Alternatively, the conduction current can be included in the displacement current, $j\omega\epsilon\mathbf{E}$, by using the equivalent complex permittivity of the medium. Similarly, it is possible to go in the opposite direction and include a part of the displacement current into the conduction current. We can adopt an arbitrary permittivity ϵ and split the electric flux density into two parts: $\epsilon\mathbf{E}$ and $\mathbf{D} - \epsilon\mathbf{E}$. We can make the medium homogeneous and isotropic with respect to this permittivity by adding the displacement current, $j\omega(\mathbf{D} - \epsilon\mathbf{E})$, to the impressed and conduction currents as

$$\mathbf{J} = \mathbf{J}_i + \mathbf{J}_c + j\omega(\mathbf{D} - \epsilon\mathbf{E}) \tag{3.54}$$

In an analogous way, we can arbitrarily adopt a permeability μ and split the magnetic flux density into two parts, $\mu\mathbf{H}$ and $\mathbf{B} - \mu\mathbf{H}$. The medium is made homogeneous and isotropic with respect to this permeability by introducing a fictitious magnetic current

$$\mathbf{M} = j\omega(\mathbf{B} - \mu\mathbf{H}) \tag{3.55}$$

Using (3.54) and (3.55), Maxwell's equations (3.14a-d) can be rearranged as

$$\nabla \times \mathbf{E} = -\mathbf{M} - j\omega\mu\mathbf{H} \quad \nabla \times \mathbf{H} = \mathbf{J} + j\omega\epsilon\mathbf{E} \quad \nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon} \quad \nabla \cdot \mathbf{H} = \frac{\tau}{\mu} \tag{3.56a-d}$$

where τ is the density of the fictitious magnetic charge, which is a counterpart of the electric charge. In the system (3.56), the density of the electric current, $\mathbf{J}$, and the density of the electric charge, ρ, are related through the continuity equation (3.14e). The magnetic current, $\mathbf{M}$, and the magnetic charge, τ, are related through the following continuity equation[3]:

$$\nabla \cdot \mathbf{M} = -j\omega\tau \tag{3.57}$$

The system (3.56) can be interpreted in the following way. $\mathbf{E}$ and $\mathbf{H}$ in the original linear, inhomogeneous, and anisotropic medium are identical with

3. Continuity equations (3.14e) and (3.57) can be derived by taking the divergence of (3.56a) and (3.56b) and using (3.56c) and (3.56d), respectively.

E and **H** in the linear, homogeneous, and isotropic medium (of parameters ϵ and μ) in which there exist equivalent volume electric and magnetic currents and charges. These currents and charges mathematically substitute the modified properties of the original medium. This substitution is called the volume equivalence principle.

By setting $\mathbf{M} = 0$ (hence, $\tau = 0$), the system of Maxwell's equations (3.56) is reduced to the system (3.17). The latter system was used in Section 3.3 to obtain the integral expressions for **E** and **H**. However, to include the contribution of the magnetic current and charge properly, appropriate integral expressions must be derived relating **E** and **H** to the magnetic current and charge. These expressions can be derived in a straightforward manner by using the duality between the electric and magnetic quantities.

3.5 Duality Relations Between Electric and Magnetic Quantities

After setting $\mathbf{J} = 0$ in Maxwell's equations (3.56) and reordering, this system can be written in the form

$$\nabla \times \mathbf{H} = j\omega\epsilon\mathbf{E} \quad \nabla \times \mathbf{E} = -\mathbf{M} - j\omega\mu\mathbf{H} \quad \nabla \cdot \mathbf{H} = \frac{\tau}{\mu} \quad \nabla \cdot \mathbf{E} = 0 \tag{3.58a-d}$$

Equations (3.58a) and (3.58d) are automatically satisfied if the electric and magnetic fields are expressed as

$$\mathbf{E} = -\frac{1}{\epsilon}\nabla \times \mathbf{A}_e \qquad \mathbf{H} = -\nabla V_m - j\omega\mathbf{A}_e \tag{3.59a,b}$$

where magnetic scalar potential (V_m) and electric vector potential ($\mathbf{A}_e$) are auxiliary functions.

The system (3.58) can formally be obtained from the system (3.17) by using the duality relations summarized in Table 3.2.[4] Similarly, (3.59) can be obtained from (3.22). Consequently, closed-form integral expressions for the potentials and fields due to the magnetic current and charge can be obtained by applying these duality relations to the closed-form integral expressions for the potentials and fields due to the electric current and charge. The results are

4. Two equations of the same mathematical form are called dual equations. Quantities occupying the same position in dual equations are called dual quantities [3].

Table 3.2
Duality Relations Between Electric and Magnetic Quantities

$\mathbf{E} \rightarrow \mathbf{H}$	$\mathbf{A}_m \rightarrow \mathbf{A}_e$	$V_e \rightarrow V_m$	$\mathbf{J} \rightarrow \mathbf{M}$	$\rho \rightarrow \tau$	$\epsilon \rightarrow \mu$
$\mathbf{H} \rightarrow -\mathbf{E}$	$\mathbf{A}_e \rightarrow -\mathbf{A}_m$	$V_m \rightarrow -V_e$	$\mathbf{M} \rightarrow -\mathbf{J}$	$\tau \rightarrow -\rho$	$\mu \rightarrow \epsilon$

summarized in Table 3.3 along with the Lorentz condition and the continuity equation.

The application of the duality relations to the system (3.56) results in the same system of equations. The application of the duality relations to the integral expressions for electric field due to the electric and magnetic currents results in the expression for the magnetic field, and vice versa. Hence, the electric and magnetic fields due to the electric and magnetic currents can be expressed in a compact operator notation

$$\mathbf{E} = -ZL(\mathbf{J}) + K(\mathbf{M}) \qquad \mathbf{H} = -K(\mathbf{J}) - YL(\mathbf{M}) \qquad (3.60a,b)$$

where the operators L and K are given in Table 3.1, and $Y = 1/Z$ is the intrinsic admittance of the medium. Equations (3.60a,b) are dual, too.

3.6 Boundary Conditions

Maxwell's equations in differential form cannot be written at points in space where the fields are discontinuous. An important location of the field discontinu-

Table 3.3
Summary of Electric and Magnetic Quantities and Relations

	Electric	Magnetic
Field	$\mathbf{E} = -\nabla V_e - j\omega\mathbf{A}_m - \frac{1}{\epsilon}\nabla \times \mathbf{A}_e$	$\mathbf{H} = -\nabla V_m - j\omega\mathbf{A}_e + \frac{1}{\mu}\nabla \times \mathbf{A}_m$
Scalar potential	$V_e(\mathbf{r}) = \frac{1}{\epsilon}\int_v \rho(\mathbf{r}')\, g(R)\, dv$	$V_m(\mathbf{r}) = \frac{1}{\mu}\int_v \tau(\mathbf{r}')\, g(R)\, dv$
Vector potential	$\mathbf{A}_e(\mathbf{r}) = \epsilon\int_v \mathbf{M}(\mathbf{r}')\, g(R)\, dv$	$\mathbf{A}_m(\mathbf{r}) = \mu\int_v \mathbf{J}(\mathbf{r}')\, g(R)\, dv$
Lorentz condition	$V_e = \frac{j}{\omega\epsilon\mu}\nabla \cdot \mathbf{A}_m$	$V_m = \frac{j}{\omega\epsilon\mu}\nabla \cdot \mathbf{A}_e$
Continuity equation	$\rho = \frac{j}{\omega}\nabla \cdot \mathbf{J}$	$\tau = \frac{j}{\omega}\nabla \cdot \mathbf{M}$

ities is a boundary surface between two different media. Consider the boundary surface between two regions, denoted by i and j, whose parameters are $\epsilon^{(i)}$, $\mu^{(i)}$ and $\epsilon^{(j)}$, $\mu^{(j)}$, respectively (Figure 3.4). The permittivities are assumed to take into account the conduction mechanism. The fields at the boundary surface are governed by the following boundary conditions:

$$\mathbf{n}_{ij} \times (\mathbf{E}^{(i)} - \mathbf{E}^{(j)}) = -\mathbf{M}_{sij} \quad \mathbf{n}_{ij} \times (\mathbf{H}^{(i)} - \mathbf{H}^{(j)}) = \mathbf{J}_{sij} \tag{3.61a,b}$$

$$\mathbf{n}_{ij} \cdot (\epsilon^{(i)}\mathbf{E}^{(i)} - \epsilon^{(j)}\mathbf{E}^{(j)}) = \rho_{sij} \quad \mathbf{n}_{ij} \cdot (\mu^{(i)}\mathbf{H}^{(i)} - \mu^{(j)}\mathbf{H}^{(j)}) = \tau_{sij} \tag{3.61c,d}$$

$$\mathbf{n}_{ij} \cdot (\mathbf{J}^{(i)} - \mathbf{J}^{(j)}) + \nabla_s \cdot \mathbf{J}_{sij} = -j\omega\rho_{sij} \tag{3.61e}$$

$$\mathbf{n}_{ij} \cdot (\mathbf{M}^{(i)} - \mathbf{M}^{(j)}) + \nabla_s \cdot \mathbf{M}_{sij} = -j\omega\tau_{sij} \tag{3.61f}$$

In the equations above, $\mathbf{n}_{ij}$ is the unit vector perpendicular to the boundary surface, directed from region j to region i; $\mathbf{E}^{(i)}$, $\mathbf{H}^{(i)}$, $\mathbf{J}^{(i)}$, and $\mathbf{M}^{(i)}$ ($\mathbf{E}^{(j)}$, $\mathbf{H}^{(j)}$, $\mathbf{J}^{(j)}$, and $\mathbf{M}^{(j)}$) are the fields and the volume currents densities at the boundary surface inside region $i(j)$; $\mathbf{J}_{sij}$ and $\mathbf{M}_{sij}$ are the densities of the surface electric and magnetic currents located at the boundary surface; while ρ_s and τ_s are the densities of the surface electric and magnetic charges. Finally, ∇_s denotes the surface divergence.

Starting from (3.61), various special boundary conditions can be obtained. The first case is the boundary surface between two real perfect dielectrics. In such media, there cannot exist electric and magnetic currents and time-varying

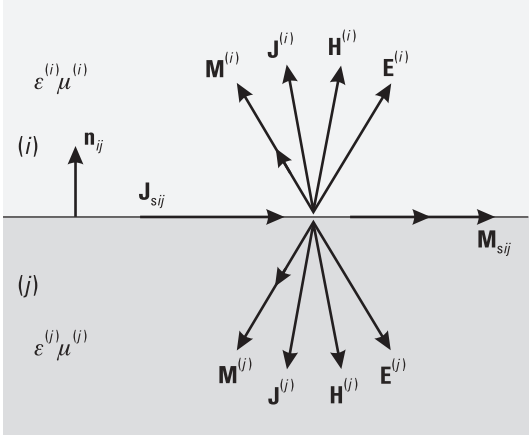

Figure 3.4 Field and current vectors at a boundary surface.

charges. In addition, there are no surface currents and charges at the boundary surface. Hence, (3.61e) and (3.61f) should be dropped out, and the right sides of the remaining equations are zero. The second case is the boundary surface between two real imperfect (lossy) dielectrics. In these media, there cannot exist the magnetic current and charge, but there are the volume electric current and charge. At the boundary surface, there exists the surface electric charge, but no surface electric current.

The third case, again involving realistic media, is when region j is filled with a good conductor. Such a material can be mathematically described by a large conductivity (σ). Assuming the frequency is high enough that the skin effect is pronounced, the fields penetrate into the medium only a short distance. We denote the region by the superscript 0 instead of j. Locally, the penetrated fields can be approximated by a plane wave propagating perpendicularly to the boundary surface [Figure 3.5(a)]. Thus, the electric and magnetic fields in region 0, $\mathbf{E}^{(0)}$, and $\mathbf{H}^{(0)}$, can be expressed as

$$\mathbf{E}^{(0)} = \mathbf{E}^{(0)}\big|_{z=0}\, e^{-\gamma_0 z} \qquad \mathbf{H}^{(0)} = Y^{(0)}\mathbf{i}_z \times \mathbf{E}^{(0)} \tag{3.62a,b}$$

where $\gamma^{(0)}$ is the propagation constant in region 0, $Y^{(0)}$ is the intrinsic admittance of this region ($Y^{(0)} = 1/Z^{(0)}$, where $Z^{(0)}$ is the intrinsic impedance), z is a local coordinate directed perpendicularly from the boundary surface into region 0, and $\mathbf{i}_z$ is the corresponding unit vector. The propagation constant and the intrinsic impedance of the conductor are

$$\gamma^{(0)} = \alpha^{(0)} + j\beta^{(0)} \approx (1+j)\sqrt{\frac{\omega\mu\sigma}{2}} \qquad Z^{(0)} = R^{(0)} + jX^{(0)} \approx (1+j)\sqrt{\frac{\omega\mu}{2\sigma}} \tag{3.63a,b}$$

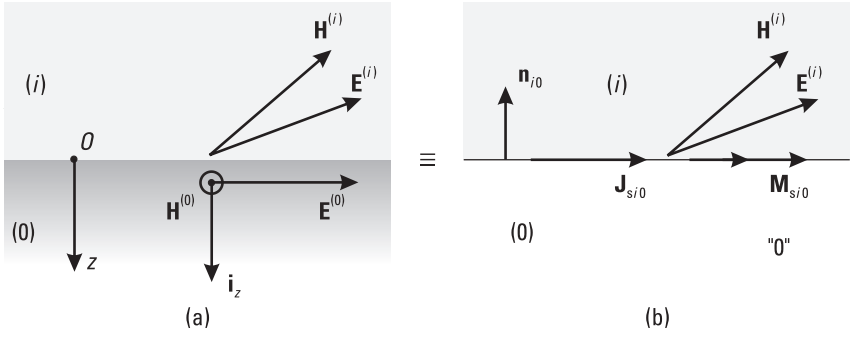

Figure 3.5 Impedance boundary surface: (a) real case and (b) idealized case.

where μ is the permeability and σ the conductivity of the material. Note that real and imaginary parts of each constant are equal. The wave that penetrates into the conductor is attenuated due to the Joule losses. The reciprocal of the attenuation constant is the skin depth.[5] The magnitude of the intrinsic impedance of a good conductor is much smaller than the intrinsic impedance of a vacuum ($\sqrt{\mu_0/\epsilon_0} \approx 376.73\Omega$). Its real part is referred to as the surface resistance of the conductor.

A generalization of the third case is when region j is filled with a medium with a large loss tangent (electric or magnetic). Under the conditions that the wavelength in the region is much smaller than the wavelength in the surrounding regions, the field is attenuated fast enough so that the skin depth is much smaller than the dimensions of the region. The real part of the intrinsic impedance or admittance of such a medium is always positive.

In the third case there are no surface electric and magnetic currents and magnetic charges at the boundary surface. As a consequence, the tangential components of the electric and magnetic fields in region i ($\mathbf{E}^{(i)}$ and $\mathbf{H}^{(i)}$), are equal to the respective tangential components in region 0 ($\mathbf{E}^{(0)}$ and $\mathbf{H}^{(0)}$). Finally, using (3.62b), the tangential components of $\mathbf{E}^{(i)}$ and $\mathbf{H}^{(i)}$ can be interrelated by the following dual relations:

$$\mathbf{H}^{(i)}\big|_{\tan} = -Y^{(0)}\mathbf{n}_{i0} \times \mathbf{E}^{(i)} \qquad \mathbf{E}^{(i)}\big|_{\tan} = Z^{(0)}\mathbf{n}_{i0} \times \mathbf{H}^{(i)} \qquad (3.64a,b)$$

where $\mathbf{n}_{i0}$ is the unit vector perpendicular to the boundary surface, directed from region 0 towards region i.

Instead of the real boundary problem depicted in Figure 3.5(a), we can consider an approximately equivalent, simpler problem shown in Figure 3.5(b). We reduce the depth of field penetration into region 0 to zero by artificially increasing the real part of the propagation constant in region 0, $\gamma^{(0)}$, to infinity. The field distribution in region i practically does not change provided the intrinsic impedance of region 0, $Z^{(0)}$, remains constant. Now, the fields within region 0 do not exist any more. However, to preserve the situation in region i, equivalent electric and magnetic currents must be placed at the boundary surface to satisfy the boundary conditions (3.61a) and (3.61b). The densities of these currents, respectively, are

$$\mathbf{J}_{si0} = \mathbf{n}_{i0} \times \mathbf{H}^{(i)} \qquad \mathbf{M}_{si0} = -\mathbf{n}_{i0} \times \mathbf{E}^{(i)} \qquad (3.65a,b)$$

5. The skin effect in a good conductor is pronounced if the skin depth is smaller than any dimension of the conductor.

The tangential components of $\mathbf{E}^{(i)}$ and $\mathbf{H}^{(i)}$ are related by (3.64). Hence, the surface electric and magnetic currents are related by the following dual relations:

$$\mathbf{M}_{si0} = -Z^{(0)}\mathbf{n}_{i0} \times \mathbf{J}_{si0} \qquad \mathbf{J}_{si0} = Y^{(0)}\mathbf{n}_{i0} \times \mathbf{M}_{si0} \qquad (3.66a,b)$$

Relations (3.64) and (3.66) are collectively referred to as the *impedance boundary conditions* (IBCs). Region 0 is referred to as the *impedance boundary region* (IBR) and its surface as the *impedance boundary surface* (IBS). For the above-mentioned dissipative materials, the real parts of $Z^{(0)}$ and $Y^{(0)}$ are always positive.[6]

The losses in region 0 are usually relatively small if the magnitude of $Z^{(0)}$ is much smaller than the intrinsic impedance of region i, $Z^{(i)}$. If we assume $|Z^{(0)}| << |Z^{(i)}|$, we can approximately take $Z^{(0)} = 0$. This eliminates the losses in region 0, the magnetic current at boundary surface is zero due to (3.66a), and the boundary conditions (3.61a)–(3.61f) reduce to

$$\mathbf{n}_{i0} \times \mathbf{E}^{(i)} = 0 \quad \mathbf{n}_{i0} \times \mathbf{H}^{(i)} = \mathbf{J}_{si0} \quad \mathbf{n}_{i0} \cdot \mathbf{E}^{(i)} = \frac{\rho_{si0}}{\epsilon} \qquad (3.67a\text{-}c)$$

$$\mathbf{n}_{i0} \cdot \mathbf{H}^{(i)} = 0 \quad \mathbf{n}_{i0} \cdot \mathbf{J}^{(i)} = -j\omega\rho_{si0} - \nabla_s \cdot \mathbf{J}_{si0} \qquad (3.67d,e)$$

The medium for which $Z^{(0)} = 0$ is referred to as a PEC.

We can now assume a fictitious dual situation when $|Y^{(0)}| << |Y^{(i)}|$ (i.e., $|Z^{(0)}| >> |Z^{(i)}|$) where we can approximately take $Y^{(0)} = 0$. In that case, again, there are no losses in region 0, the electric current at the boundary surface is zero due to (3.66b), and the boundary conditions (3.61) reduce to

$$\mathbf{n}_{i0} \times \mathbf{E}^{(i)} = -\mathbf{M}_{si0} \quad \mathbf{n}_{i0} \times \mathbf{H}^{(i)} = 0 \quad \mathbf{n}_{i0} \cdot \mathbf{E}^{(i)} = 0 \qquad (3.68a\text{-}c)$$

$$\mathbf{n}_{i0} \cdot \mathbf{H}^{(i)} = \frac{\tau_{si0}}{\mu} \quad \mathbf{n}_{i0} \cdot \mathbf{M}^{(i)} = -j\omega\tau_{si0} - \nabla_s \cdot \mathbf{M}_{si0} \qquad (3.68d,e)$$

The medium for which $Y^{(0)} = 0$ is referred to as a *perfect magnetic conductor* (PMC).

When region i is an exterior region, we can mathematically consider this region to be bounded by a sphere (referred to as the radiation sphere) whose center is at the coordinate origin (Figure 3.3) and whose radius (r) tends to

6. The concept of the IBC can be extended to other cases of purely reactive surfaces (e.g., corrugated surfaces), which is not within our scope. An extensive treatment of the boundary conditions can be found in [13].

infinity. Provided the conditions are fulfilled for the validity of the far-field expressions (3.53), there exist simple relations between the electric and magnetic fields at the radiation sphere,

$$\mathbf{E}^{(i)} = -Z^{(i)}\mathbf{i}_r \times \mathbf{H}^{(i)} \qquad \mathbf{H}^{(i)} = Y^{(i)}\mathbf{i}_r \times \mathbf{E}^{(i)} \tag{3.69a,b}$$

where $\mathbf{i}_r$ is the unit vector directed from the coordinate origin toward the far-field point. This vector coincides with the unit normal on the radiation sphere. The equations in (3.69) are referred to as the radiation boundary conditions. Table 3.4 shows a summary of some important specific boundary conditions.

3.7 Formulation of the Basic Field Problem in the Frequency Domain

In this section, we shall formulate the basic electromagnetic-field problem treated in this book. As mentioned in Section 3.1, linear media can be inhomogeneous and anisotropic. We shall deal primarily with isotropic media and approximate any inhomogeneous medium by a piecewise-homogeneous medium. The analysis will be performed in the frequency domain.

Figure 3.6 shows a sketch of the basic problem considered here. The electromagnetic system consists of a finite number of finite-size linear, homogeneous, and isotropic regions situated in an unbounded linear, homogeneous, and isotropic environment. Most often, this environment is a vacuum, but it also can be another medium (e.g., water). The finite-size regions will be referred to as interior regions, while the environment (infinite) region will be referred

Table 3.4
Summary of Specific Boundary Conditions

Impedance	PEC	PMC	Radiation
$\mathbf{E}^{(i)}\|_{\tan} = Z^{(i)}\mathbf{n}_{i0} \times \mathbf{H}^{(i)}$	$\mathbf{n}_{i0} \times \mathbf{E}^{(i)} = 0$	$\mathbf{n}_{i0} \times \mathbf{E}^{(i)} = -\mathbf{M}_{si0}$	$\mathbf{E}^{(i)} = -Z^{(i)}\mathbf{i}_r \times \mathbf{H}^{(i)}$
$\mathbf{H}^{(i)}\|_{\tan} = -Y^{(i)}\mathbf{n}_{i0} \times \mathbf{E}^{(i)}$	$\mathbf{n}_{i0} \times \mathbf{H}^{(i)} = \mathbf{J}_{si0}$	$\mathbf{n}_{i0} \times \mathbf{H}^{(i)} = 0$	$\mathbf{H}^{(i)} = Y^{(i)}\mathbf{i}_r \times \mathbf{E}^{(i)}$
$\mathbf{M}_{si0} = -Z^{(i)}\mathbf{n}_{i0} \times \mathbf{J}_{si0}$	$\mathbf{n}_{i0} \times \mathbf{E}^{(i)} = \rho_{si0}/\epsilon^{(i)}$	$\mathbf{n}_{i0} \cdot \mathbf{E}^{(i)} = 0$	
$\mathbf{J}_{si0} = Y^{(i)}\mathbf{n}_{i0} \times \mathbf{M}_{si0}$	$\mathbf{n}_{i0} \cdot \mathbf{H}^{(i)} = 0$	$\mathbf{n}_{i0} \cdot \mathbf{H}^{(i)} = \tau_{si0}/\mu^{(i)}$	
	$\mathbf{n}_{i0} \cdot \mathbf{J}^{(i)} = -j\omega\rho_{si0} - \nabla_s \cdot \mathbf{J}_{si0}$	$\mathbf{n}_{i0} \cdot \mathbf{M}^{(i)} = -j\omega\tau_{si0} - \nabla_s \cdot \mathbf{M}_{si0}$	

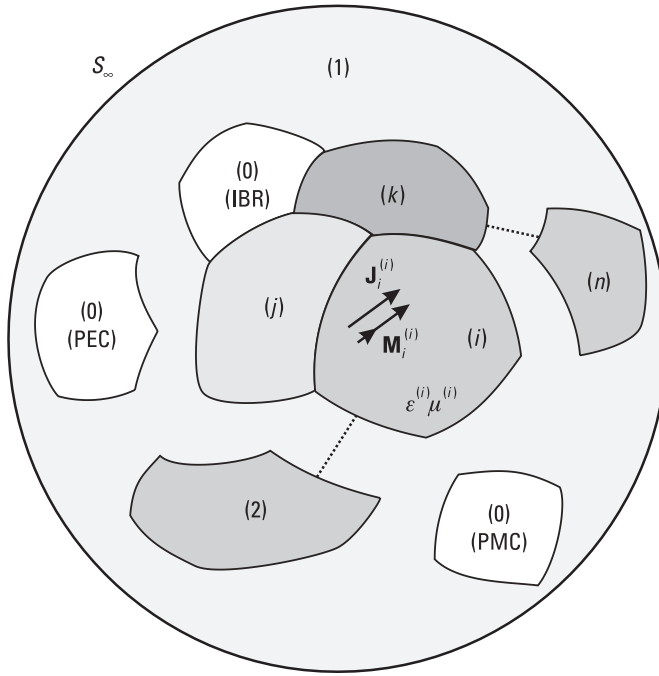

Figure 3.6 Basic electromagnetic field problem treated in this book.

to as the exterior region. In Figure 3.6 the exterior region is denoted as region 1, although this numbering scheme is not necessary.

Some of the interior regions can be PECs, PMCs, or IBRs. In the interior of any such region, the electromagnetic field is always zero. Hence, all of them will collectively be denoted as region 0 and referred to as zero-field regions. In general, on the boundary surface of region 0, the intrinsic impedance, $Z^{(0)}$, can be a function of the position on the surface.

As special cases of zero-field regions, we shall consider metallic wires and plates. These entities can be made of a PEC, or they can be IBRs. Plates forming an open surface can be considered as a degenerate case of a zero-field region, where one dimension (thickness) of the region is zero.

In all the remaining regions (including the exterior region) generally there exist electromagnetic fields. These regions will collectively be referred to as non-zero-field regions. We denote the total number of such regions by n. The medium filling out region i is assumed to be described by the complex permittivity, $\boldsymbol{\epsilon}^{(i)}$, and permeability, $\boldsymbol{\mu}^{(i)}$, $i = 1, \ldots, n$, which include all losses. In any of these regions, there may exist impressed electric and magnetic currents, $\mathbf{M}_i^{(i)}$ and $\mathbf{J}_i^{(i)}$, $i = 1, \ldots, n$, whose angular frequency is ω. As a particular case, these electric and magnetic currents can be boundary surface currents.

The main objective of the frequency-domain analysis is to evaluate the distribution of the electric and magnetic fields or their sources in the non-zero-field regions.

3.8 Poynting Theorem

The Poynting theorem, a consequence of the energy conservation law, is usually formulated only for the electric current. However, for the purpose of this book, we need a form of this theorem that includes both the electric and magnetic currents. To derive the required form in the complex domain, we integrate the dot product of **E** with the complex-conjugate of (3.56b) over the volume v,

$$\int_v \mathbf{E} \cdot \nabla \times \mathbf{H}^* \, dv = \int_v \mathbf{E} \cdot \mathbf{J}^* \, dv - j\omega\epsilon^* \int_v |\mathbf{E}|^2 \, dv \tag{3.70}$$

Using identity $\nabla \cdot (\mathbf{a} \times \mathbf{b}) = \mathbf{b} \cdot (\nabla \times \mathbf{a}) - \mathbf{a} \cdot (\nabla \times \mathbf{b})$ for arbitrary vectors **a** and **b**, the left side of (3.58b) is written as

$$\int_v \mathbf{E} \cdot \nabla \times \mathbf{H}^* \, dv = \int_v \mathbf{H}^* \cdot \nabla \times \mathbf{E} dv - \int_v \nabla \cdot (\mathbf{E} \times \mathbf{H}^*) \, dv \tag{3.71}$$

The term $\nabla \times \mathbf{E}$ on the right side of (3.71) is expressed as $-\mathbf{M} - j\omega\mu\mathbf{H}$ according to (3.56a), while the second integral on the right side of (3.71) is transformed using the divergence theorem as

$$\int_v \nabla \cdot (\mathbf{E} \times \mathbf{H}^*) \, dv = \oint_S (\mathbf{E} \times \mathbf{H}^*) \cdot d\mathbf{S} \tag{3.72}$$

where S is the boundary surface of volume v, and **n** is the unit normal vector directed outwards from the volume. Finally, after some rearrangements of (3.70), the Poynting theorem becomes

$$j\omega \int_v (\mu |\mathbf{H}|^2 - \epsilon^* |\mathbf{E}|^2) \, dv = -\int_v (\mathbf{J}^* \cdot \mathbf{E} + \mathbf{M} \cdot \mathbf{H}^*) \, dv - \oint_S (\mathbf{E} \times \mathbf{H}^*) \cdot d\mathbf{S} \tag{3.73}$$

Equation (3.73) is a power-balance equation. The left side represents the complex power used to maintain the electromagnetic field inside volume v.

The first term on the right side is the complex power supplied by the sources (electric and magnetic currents) located in the region. The remaining terms on the right side give the complex power entering the volume from its environment.

Starting from (3.73), we easily formulate the Poynting theorem for region i depicted in Figure 3.6. The boundary surface of region i, S_i, can be represented as a union of boundary surfaces between region i and all the surrounding regions. When region i is the exterior region, its outer boundary surface, S_∞, should be included into this union. In that general case, the Poynting theorem is expressed as

$$j\omega \int_{v_i} [\mu^{(i)} |\mathbf{H}^{(i)}|^2 - (\epsilon^{(i)})^* |\mathbf{E}^{(i)}|^2] \, dv = -\int_{v_i} [(\mathbf{J}^{(i)})^* \cdot \mathbf{E}^{(i)} + \mathbf{M}^{(i)} \cdot (\mathbf{H}^{(i)})^*] \, dv$$

$$+ \sum_{\substack{j=0 \\ j \neq i}}^{n} \int_{S_{ij}} [\mathbf{E}^{(i)} \times (\mathbf{H}^{(i)})^*] \cdot d\mathbf{S}_{ij} + \int_{S_\infty} [\mathbf{E}^{(i)} \times (\mathbf{H}^{(i)})^*] \cdot (-\mathbf{i}_r) \, dS \tag{3.74}$$

where $d\mathbf{S}_{ij} = \mathbf{n}_{ij} \, dS$, $\mathbf{n}_{ij}$ is the unit normal vector directed from region j towards region i, and $\mathbf{i}_r$ is the radial unit vector. The jth term exists in the sum only if regions i and j have a common boundary surface. The last term in (3.74) exists only when region i is the exterior region.

3.9 Surface Equivalence Principle

The closed-form integral expressions for the potentials and fields, given in Sections 3.3 and 3.5, can be written for the currents situated in an unbounded linear, homogeneous, and isotropic medium. These fields are governed by Maxwell's equations (3.17) and the constitutive equations (3.7). However, in many cases it is necessary to have closed-form expression for the potentials and fields in a bounded linear, homogeneous, and isotropic medium, surrounded by an arbitrary linear medium.

Consider, now, any interior region in Figure 3.6. Let us denote the volume of the region by v and its boundary (closed) surface by S. For such a region, the integral expressions of Sections 3.3 and 3.5 cannot be applied directly. However, using the surface equivalence principle, revealed in this section, the fields and potentials in v can be expressed in terms of the actual sources located in v and equivalent sources located on S.

To derive the surface equivalence principle, we first consider the case when the impressed field sources are located only outside S [Figure 3.7(a)]. We start from the following (dual) equations[7]

$$\mathbf{E} = -\gamma Z \oint_S \left\{ (-\mathbf{n} \times \mathbf{H}_s) g(R) - \frac{1}{\gamma^2} \nabla [(-\mathbf{n} \times \mathbf{H}_s) \cdot \nabla g(R)] \right\} dS' \qquad (3.75a)$$
$$+ \oint_S (\mathbf{n} \times \mathbf{E}_s) \times \nabla g(R)\, dS'$$

$$\mathbf{H} = -\gamma Y \oint_S \left\{ (\mathbf{n} \times \mathbf{E}_s) g(R) - \frac{1}{\gamma^2} \nabla [(\mathbf{n} \times \mathbf{E}_s) \cdot \nabla g(R)] \right\} dS' \qquad (3.75b)$$
$$- \oint_S (-\mathbf{n} \times \mathbf{H}_s) \times \nabla g(R)\, dS'$$

where $\mathbf{n}$ is the unit normal on S, directed outwards, and $\mathbf{E}_s$ and $\mathbf{H}_s$ are the electric and magnetic field at the boundary surface, just inside the region. According to (3.75), the electric and magnetic field within v, $\mathbf{E}$ and $\mathbf{H}$, whose sources are only in the exterior region, can uniquely be evaluated by knowing the components of both the electric field and the magnetic field tangential to

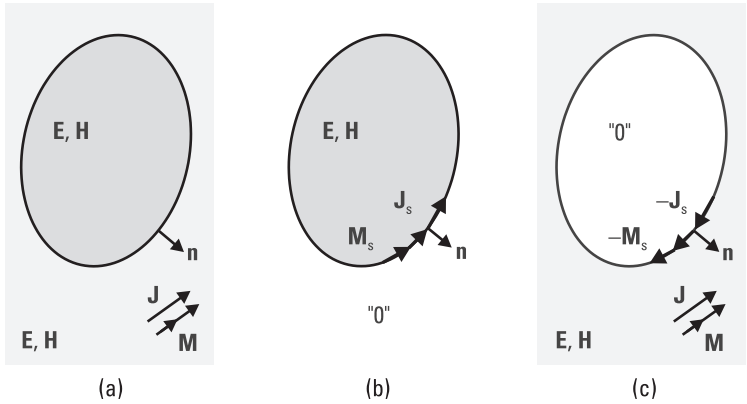

Figure 3.7 Surface equivalence principle for exterior sources: (a) original problem, (b) equivalent exterior problem, and (c) equivalent interior problem.

7. Detailed derivation of these equations can be found in [14]. These equations can be considered as variants of the Kirchhoff-Huygens formulas [1].

the region boundary. Similarly, it may be shown that (3.75) can be used for the evaluation of the electric and magnetic fields in the exterior region if this field has impressed sources only in the interior region [Figure 3.8(a)]. In that case, the unit normal vector **n** is directed from the exterior region into the interior region. As an important consequence, if the components of both $\mathbf{E}_s$ and $\mathbf{H}_s$ tangential to the boundary of a source-free region are zero, the fields within this region are zero.

Equations (3.75a) and (3.75b) are identical with the integral expressions for the electric and magnetic field due to electric and magnetic currents placed at the boundary surface S, whose densities, respectively, are

$$\mathbf{J}_s = -\mathbf{n} \times \mathbf{H}_s \qquad \mathbf{M}_s = \mathbf{n} \times \mathbf{E}_s \tag{3.76a,b}$$

where $\mathbf{E}_s$ and $\mathbf{H}_s$ are the fields at the boundary surface. This can be shown using (3.60), when the operators L and K, given in Table 3.1, are applied to these surface currents.

As consequence, if there exist impressed sources only in the exterior region [Figure 3.7(a)], the fields in the interior region will not change if these external sources are replaced by the equivalent surface electric and magnetic currents located on the region boundary [Figure 3.7(b)].

On the interior face of surface S, the tangential components of **E** and **H** due to the equivalent currents $\mathbf{J}_s$ and $\mathbf{M}_s$ are identical to the corresponding components involved in (3.76). Starting from the boundary conditions (3.61), it can be concluded that the components of the electric and magnetic fields, produced by these currents, tangential to S in the exterior region, are equal to zero. In other words, the equivalent currents, $\mathbf{J}_s$ and $\mathbf{M}_s$, do not produce fields in the exterior region.

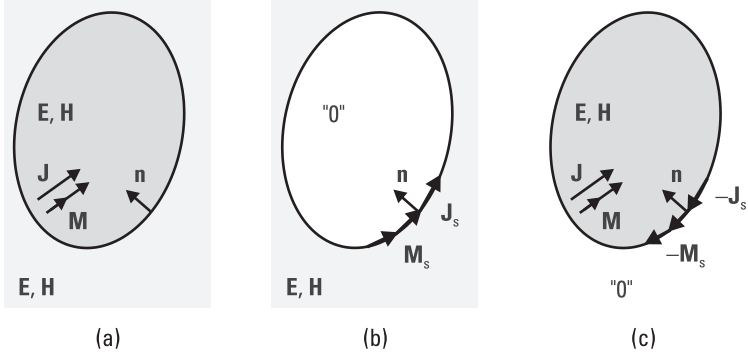

Figure 3.8 Surface equivalence principle for interior sources: (a) original problem, (b) equivalent interior problem, and (c) equivalent exterior problem.

Subtract the equivalent currents and their respective fields from the original problem [Figure 3.7(c)]. As a result, the fields in the exterior region are not changed, while the fields in the interior region are annihilated. Thus, the exterior source problem, shown in Figure 3.7(a), can be represented as a superposition of the interior- and exterior-region problems, shown in Figure 3.7(b) and 3.7(c), respectively. This conclusion is referred to as the surface equivalence principle for exterior sources.

In a similar way, the surface equivalence principle can be established for the interior sources [Figure 3.8(a)]. The field in the exterior region remains unchanged when we replace the interior sources by equivalent electric and magnetic currents placed on the region boundary [Figure 3.8(b)]. We evaluate the equivalent currents, $\mathbf{J}_s$ and $\mathbf{M}_s$, by using (3.76), except that now the unit normal, $\mathbf{n}$, is directed from the exterior toward the interior region. In this case, the fields in the interior region are zero. If we subtract the equivalent currents and their fields from the original problem [Figure 3.8(c)], the fields in the interior region are unchanged, while the fields in the exterior region are annihilated. Hence, the interior source problem, shown at Figure 3.8(a), can be represented as a superposition of the interior and exterior region problems, shown in Figure 3.8(b) and 3.8(c), respectively.

In general, the field has impressed sources in both the interior and exterior regions. Such a problem can be always decomposed into an interior- and an exterior-source problem. The electric and magnetic fields in the interior region can be evaluated in terms of the interior impressed sources and the equivalent surface currents, which satisfy (3.76). The fields in the exterior region can similarly be evaluated in terms of the exterior sources and the equivalent currents that are the negative of the equivalent currents for the interior problem.

3.10 Uniqueness Theorem

The first step in solving a field problem consists of choosing a set of starting equations. To choose this set properly, it is necessary to know the conditions under which the solution exists and is unique. We are interested in the uniqueness theorem for the general field problem in the frequency domain shown in Figure 3.6 and described in Section 3.7. The theorem will be given in two forms: for a single region and for a multiple region.

3.10.1 Single Region

Consider region i as filled with a lossy, linear, homogenous, and isotropic medium of parameters $\epsilon^{(i)}$ and $\mu^{(i)}$, excited by known impressed electric and

magnetic currents (Figure 3.6). The region can be either interior or exterior. Let us suppose there exist two solutions of Maxwell's equations (3.56) for the electromagnetic field inside region i (e.g., due to different sources outside region i). Inside region i we assume the impressed currents to be identical for the two solutions. These solutions can be represented by the field vectors $\mathbf{E}_1^{(i)}$, $\mathbf{H}_1^{(i)}$, and $\mathbf{E}_2^{(i)}$, $\mathbf{H}_2^{(i)}$, respectively. The difference between these two solutions can be represented by the difference fields as

$$\delta\mathbf{E}^{(i)} = \mathbf{E}_1^{(i)} - \mathbf{E}_2^{(i)} \qquad \delta\mathbf{H}^{(i)} = \mathbf{H}_1^{(i)} - \mathbf{H}_2^{(i)} \tag{3.77a,b}$$

If the difference fields are forced to zero, a unique solution for the field inside region i is obtained. Let us consider under which conditions the difference fields are forced to zero.

The difference fields have no sources inside region i. Hence, the real part of (3.74) can be written as

$$\begin{aligned}\omega \int_{v_i} [\mu''^{(i)} |\delta\mathbf{H}^{(i)}|^2 + \epsilon''^{(i)} |\delta\mathbf{E}^{(i)}|^2]\, dv \\ = \sum_{\substack{j=0 \\ j\neq i}}^{n} \mathrm{Re}\left\{ \oint_{S_{ij}} [\delta\mathbf{E}^{(i)} \times (\delta\mathbf{H}^{(i)})^*] \cdot \mathbf{n}_{ij}\, dS \right\} \\ + \mathrm{Re}\left\{ \oint_{S_\infty} [\delta\mathbf{E}^{(i)} \times (\delta\mathbf{H}^{(i)})^*] \cdot (-\mathbf{i}_r)\, dS \right\}\end{aligned} \tag{3.78}$$

Since the medium within region i is lossy, the left side of (3.78) represents the average power dissipated in the region. This power is always nonnegative. It can be zero only if the difference fields are zero everywhere within region i. Hence, the right side of (3.78), which is the average power entering region i, can also only be nonnegative. Obviously, if the average power entering region i is zero, the difference fields are automatically zero.

In order that the average power entering region i be zero, it is sufficient to prove that mixed products on the right side of (3.78) are zero or nonpositive. The value of $[\delta\mathbf{E}^{(i)} \times (\delta\mathbf{H}^{(i)})^*] \cdot \mathbf{n}_{ij}$, $j = 0, \ldots, n$, $j \neq i$, does not change if the vectors cyclically change their places; that is,

$$\begin{aligned}[\delta\mathbf{E}^{(i)} \times (\delta\mathbf{H}^{(i)})^*] \cdot \mathbf{n}_{ij} &= (\mathbf{n}_{ij} \times \delta\mathbf{E}^{(i)}) \cdot (\delta\mathbf{H}^{(i)})^* \\ &= -[\mathbf{n}_{ij} \times (\delta\mathbf{H}^{(i)})^*] \cdot \delta\mathbf{E}^{(i)}\end{aligned} \tag{3.79}$$

Hence, this mixed product is zero if at least one of the following conditions is fulfilled:

$$\mathbf{n}_{ij} \times \delta\mathbf{E}^{(i)} = 0 \qquad \mathbf{n}_{ij} \times (\delta\mathbf{H}^{(i)})^* = 0 \tag{3.80a,b}$$

These conditions can be fulfilled, respectively, if the tangential components of the electric and the magnetic field are known at the boundary surface between regions i and j (for $j = 0, \ldots, n, j \neq i$).

If region 0 is a PEC, from the boundary condition (3.67a) it follows that the tangential component of the difference field automatically satisfies (3.80a). Similarly, if region 0 is a PMC, (3.68b) implies that (3.80b) is automatically satisfied. If, however, region 0 is an IBR, the mixed product $[\delta\mathbf{E}^{(i)} \times (\delta\mathbf{H}^{(i)})^*] \cdot \mathbf{n}_{i0}$ can be transformed according to the IBC (3.46a)

$$[\delta\mathbf{E}^{(i)} \times (\delta\mathbf{H}^{(i)})^*] \cdot \mathbf{n}_{i0} = (\mathbf{n}_{i0} \times \delta\mathbf{E}^{(i)}) \cdot (\delta\mathbf{H}^{(i)})^* = -(Y^{(0)})^* \left|\mathbf{n}_{i0} \times \delta\mathbf{E}^{(i)}\right|^2 \tag{3.81}$$

The real part of the admittance $Y^{(0)}$ is nonnegative (as the media are assumed to be passive everywhere), so that the real part of the mixed product in (3.81) is nonpositive. Similarly, using the radiation boundary condition (3.69a) we can show that the mixed product $[\delta\mathbf{E}^{(i)} \times (\delta\mathbf{H}^{(i)})^*] \cdot (-\mathbf{i}_r)$ is nonpositive at the outer boundary surface in infinity.

In particular, consider a lossless interior region i.[8] Assume that at each point of its boundary surface one and only one of the following conditions is fulfilled:

- The tangential component of the electric field is known [implying (3.80a) is satisfied].
- The tangential component of the magnetic field is known [implying (3.80b) is satisfied].

As a consequence, there is no power flow between region i and its environment at any point on the boundary surface. The difference fields inside region i ($\delta\mathbf{E}^{(i)}$ and $\delta\mathbf{H}^{(i)}$) will not change if the following are true:

- The part of the boundary surface where (3.80a) is valid is covered by a PEC.
- The part of the boundary surface where (3.80b) is valid is covered by a PMC.

8. To establish uniqueness in this case, we shall follow an intuitive approach, rather than a mathematically rigorous proof.

Region i can now be considered a lossless cavity that can support resonant fields (cavity modes) at a set of discrete frequencies (resonant frequencies). The size, shape, and material properties of the cavity determine the resonant frequencies. These cavity modes are not the target of our analysis, and they will be regarded as spurious solutions. The solution can be represented as a sum of the original field and a spurious cavity mode of an arbitrary magnitude. Hence, the solution for the field inside region i is not unique at the resonant frequencies if only one of tangential field components is known at its boundary surface. Away from the resonant frequencies, the solution is unique. Note that a lossless exterior region does not support cavity modes, so that the solution is unique under the same conditions as for a lossy exterior region.

To obtain a unique solution at resonant frequencies, we must know the tangential components of both the electric and magnetic fields. Now, the difference fields satisfy both conditions (3.80a) and (3.80b). After plugging these conditions into (3.75), we obtain that difference fields in region i are zero, even at the cavity resonances.

The unique solution is provided even if we know tangential components of both fields on a part of the boundary surface, ΔS, while over the remaining part, $S - \Delta S$, only one of these components is known. In that case, the difference fields satisfy both conditions (3.80a) and (3.80b) on ΔS and one of these conditions on $S - \Delta S$. If there were a resonant mode, which satisfied such boundary conditions, it could also exist in a cavity with a hole over ΔS. However, according to common engineering sense, the energy of the cavity mode would leak through the hole, no matter how small it was. Hence, there is no such a resonant mode and the solution is unique.

Instead of imposing both conditions (3.80a) and (3.80b) on ΔS, we can apply the IBC. In that case the corresponding zeroth term on the right side of (3.78) is equal to zero. According to (3.81) this is possible only if the difference field satisfies both conditions (3.80a) and (3.80b).

In conclusion, the electric and magnetic fields inside a region are uniquely determined by the known sources in it if, at every point of its boundary surface, one of the following conditions is fulfilled, as applicable:

- The tangential component of the electric field is known.
- The tangential component of the magnetic field is known.
- The IBC is satisfied.
- The radiation boundary condition is satisfied.

The exception is a lossless interior region, whose boundary surface is not an IBS, at resonant frequencies. In that case, we should know the tangential

components of both fields at least on a part of the boundary surface and the tangential component of either electric or magnetic field on the rest of the surface.

3.10.2 Multiple Region

Consider the multiple-region problem shown in Figure 3.6. Suppose there are two solutions to Maxwell's equations for the electromagnetic field inside each of the n regions for a given set of impressed sources. (The field is zero inside region 0, and this region will be excluded from the considered multiple region.) In particular, in region i these two solutions can be represented by the field vectors $\mathbf{E}_1^{(i)}$, $\mathbf{H}_1^{(i)}$, and $\mathbf{E}_2^{(i)}$, $\mathbf{H}_2^{(i)}$, respectively. The difference of these two solutions is represented by the difference field vectors, $\delta\mathbf{E}^{(i)}$ and $\delta\mathbf{H}^{(i)}$, as defined by (3.77). The real part of (3.74) for the difference field can be written for region i in the form (3.78). After summing these equations for $i = 1, \ldots, n$, we obtain

$$\begin{aligned}\omega \sum_{i=1}^{n} \int_{v_i} [\mu''^{(i)} |\delta\mathbf{H}^{(i)}|^2 + \epsilon''^{(i)} |\delta\mathbf{E}^{(i)}|^2 \, dv] \\ = \sum_{i=1}^{n} \mathrm{Re}\left\{ \oint_{S_{i0}} [\delta\mathbf{E}^{(i)} \times (\delta\mathbf{H}^{(i)})^*] \cdot \mathbf{n}_{i0} \, dS \right\} \\ + \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} \mathrm{Re}\left\{ \oint_{S_{ij}} [\delta\mathbf{E}^{(i)} \times (\delta\mathbf{H}^{(i)})^*] \cdot \mathbf{n}_{ij} \, dS \right\} \\ + \mathrm{Re}\left\{ \oint_{S_\infty} [\delta\mathbf{E}^{(i)} \times (\delta\mathbf{H}^{(i)})^*] \cdot (-\mathbf{i}_r) \, dS \right\}\end{aligned} \tag{3.82}$$

If all media are lossy, the left side of this equation represents the average power dissipated in the multiple region. This side is always nonnegative. It can be zero only if the difference field is zero at all points of all n regions. In all other cases, it is positive. Hence, the right side of (3.82), which represents the average power entering the multiple region, can be only nonnegative. In all cases when the average power entering the multiple region is zero, the difference field in each particular region is also zero.

Let us consider under which conditions the average power entering the multiple region is zero or nonpositive. This power is represented by the three

terms on the right side of (3.82). The first term represents the average power entering the multiple region from region 0, through a PEC, PMC, or impedance boundary surface. For a PEC surface, this power is zero if condition (3.67a) is applied. For a PMC surface, this power is zero if condition (3.67b) is applied. Finally, in cases when an IBC is taken into account, this power is always nonpositive as explained in Section 3.10.1.

The second term on the right side of (3.82) represents the sum of powers entering into each of the n regions from all its neighboring regions. Bearing in mind that the surfaces S_{ij} and S_{ji} geometrically coincide, and that the unit normal vectors $\mathbf{n}_{ij}$ and $\mathbf{n}_{ji}$ are parallel, but of opposite orientations, this term can be written as

$$\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n}\mathrm{Re}\left\{\int_{S_{ij}}[\delta\mathbf{E}^{(i)}\times(\delta\mathbf{H}^{(i)})^{*}]\cdot d\mathbf{S}_{ij}\right\}= \tag{3.83}$$

$$\sum_{i=1}^{n}\sum_{j=i+1}^{n}\mathrm{Re}\left\{\int_{S_{ij}}[\delta\mathbf{E}^{(i)}\times(\delta\mathbf{H}^{(i)})^{*}-\delta\mathbf{E}^{(j)}\times(\delta\mathbf{H}^{(j)})^{*}]\cdot d\mathbf{S}_{ij}\right\}$$

Using (3.61a) and (3.61b) and noting that the surface currents on the right-hand sides can only be impressed sources, the boundary conditions for the difference fields can be written in the form

$$\mathbf{n}_{ij}\times(\delta\mathbf{E}^{(i)}-\delta\mathbf{E}^{(j)})=0 \qquad \mathbf{n}_{ij}\times(\delta\mathbf{H}^{(i)}-\delta\mathbf{H}^{(j)})=0 \tag{3.84}$$

Hence, the right side of (3.83) is zero.

The third term on the right side of (3.82) represents the average power entering the multiple region from the radiation boundary surface. If the radiation boundary condition is satisfied, this term is always nonpositive.

In conclusion, the multiple-region problem involving lossy media in all regions considered has a unique solution for a given set of impressed sources if the following conditions are fulfilled, as applicable:

- The tangential component of the electric field at the PEC boundary surface is zero.
- The tangential component of the magnetic field at the PMC boundary surface is zero.
- The IBC is satisfied at the impedance boundary surface.

- The radiation boundary condition is satisfied at the radiation boundary surface.
- The tangential components of the electric and magnetic fields satisfy the boundary conditions given by (3.61a) and (3.61b) at the boundary surface between two nonzero-field regions.

Following similar reasoning as in Section 3.10.1, it can be found that the same set of conditions is valid if some or all regions are lossless. The exception is one lossless interior region, or a set of such regions mutually connected, whose boundary surface is completely PEC or PMC at resonant frequencies. In that case, we should know the tangential components of both fields at least on a part of the boundary surface and the tangential component of either electric or magnetic field on the rest of the surface.

3.11 Summary

This chapter primarily presents the classic theory of the electromagnetic field. This field is governed by the four Maxwell's equations and three constitutive relations. If time-harmonic field sources are impressed in linear media, the resulting fields are also time-harmonic. The time dependence of such fields can be eliminated from Maxwell's equations if they are written in the frequency domain instead of the time domain, which simplifies the equations. The analysis can further be facilitated if the field vectors are expressed in terms of potentials. If the medium is infinite, linear, homogeneous, and isotropic, the potentials and fields can be expressed in the form of integrals in terms of the electric and magnetic currents and charges. In this chapter, operators L and K are defined and aimed at writing field integral equations in the following chapters in a compact form. The electromagnetic field satisfies many theorems, which can be used to facilitate the analysis. In this chapter a particular stress was placed on equivalence and uniqueness theorems essential for the remainder of this book.

An elementary problem of the electromagnetic field analysis is when known field sources are located in a linear, homogeneous, and isotropic medium. In that case, the fields can be evaluated by integration. However, even in this simple case, the integration must be carried out numerically as the solution is seldom known in a closed form.

In real problems the distribution of the field sources is not known a priori. The distribution of the fields throughout the space or the distribution of the field sources (within certain finite or infinite volumes, over surfaces, or along lines) can be determined only numerically by solving appropriate integral

or differential equations. Along these guidelines is the general problem of composite metallic and dielectric structures treated in this book, defined in Section 3.7.

References

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

[2] Morse, P. N., and H. Feshbach, *Methods of Theoretical Physics,* Vols. 1 and 2, New York: McGraw-Hill, 1953.

[3] Harrington, R. F., *Time-Harmonic Electromagnetic Fields,* New York: McGraw-Hill, 1961.

[4] Jackson, J. D., *Classical Electrodynamics,* New York: John Wiley and Sons, 1962.

[5] Jordan, E. C., and K. G. Balmain, *Electromagnetic Waves and Radiating Systems,* Englewood Cliffs, NJ: Prentice Hall, 1968.

[6] Popovic, B. D., *Introductory Engineering Electromagnetics,* Reading, MA: Addison-Wesley, 1971.

[7] Van Bladel, J., *Electromagnetic Fields,* New York: Hemisphere, 1985.

[8] Balanis, C. A., *Advanced Engineering Electromagnetics,* New York: John Wiley and Sons, 1989.

[9] Collin, R. E., *Field Theory of Guided Waves,* New York: IEEE Press, 1991.

[10] Lindell, V., *Methods for Electromagnetic Field Analysis,* New York: IEEE Press, 1992.

[11] Ramo, S., J. R. Whinnery, and T. Van Duzer, *Fields and Waves in Communication Electronics,* New York: John Wiley and Sons, 1994.

[12] Popovic, B. D., M. B. Dragovic, and A. R. Djordjevic, *Analysis and Synthesis of Wire Antennas,* Chichester, U.K.: Research Studies Press, 1982.

[13] Senior, T. B. A., and J. L. Volakis, *Approximate Boundary Conditions in Electromagnetics,* London: IEE Press, 1995.

[14] Peterson, A. F., S. L. Ray, and R. Mittra, *Computational Methods for Electromagnetics,* New York: IEEE Press, 1998.

4

Field Integral Equations

Computer modeling of electromagnetic field is based on numerical solutions of appropriate field equations. These equations can be classified into two basic groups according to the choice of the unknown quantity to be solved.

The unknown quantity for the first group is the distribution of the electromagnetic field (or related potentials) within a spatial region. Corresponding field equations are differential Maxwell's equations or equations derived from them (e.g., wave equations). In that case, boundary conditions and constitutive relations are usually built into the approximation for the field distribution.

The unknown quantity for the second group of equations is the distribution of field sources. Most often, these sources are distributed within a smaller spatial region than the electromagnetic field itself. The field sources can be distributed along lines, over surfaces, or within volumes. For this group of equations, Maxwell's equations are automatically satisfied everywhere except at region boundaries. Corresponding equations are formulated based on boundary conditions or constitutive equations. In both cases, the electromagnetic field is expressed as an integral of field sources, and the resulting equations are integral equations. According to the classification of Section 1.2, these equations are BIEs and CIEs.

In the first section of this chapter, the BIEs for metallic structures are elaborated. We start with the EFIEs and MFIEs and their scalar formulation. Then alternative forms of these equations are given, which are based on equivalent sources and *extended boundary conditions* (EBCs). Special care is devoted to elimination of spurious fields that can occur at internal resonances of closed bodies. First, a survey of various techniques that can be used in conjunction with the EFIE is made (e.g., minimum-norm solution, higher-order basis func-

tions, and insertion of losses). Then, the *augmented field integral equation* (AFIE), *combined field integral equation* (CFIE), and *combined source integral equation* (CSIE) are presented. Further, it is shown how the analysis based on the EFIE can be facilitated for open bodies and wire structures. Particularly, for wire structures, the Hallén equation is given. To include losses in a metallic body due to the skin effect, its boundary is treated as an ideal impedance boundary surface. Hence, various types of the *impedance boundary condition integral equations* (IBC-IE) are considered. Finally, an attempt is made to choose optimal BIEs for analysis of composite metallic structures.

A survey of the BIEs for combined metallic and dielectric structures is given in the second section. We start with the EFIE, MFIE, and CSIE for multiple-region problems. However, attention is paid mainly to *combined region integral equations* (CRIEs), and to Muller and PMCHW equations as their particular forms. Further, treatment of thin plates at an interface between two regions is discussed. Finally, an attempt is made to choose optimal BIEs for analysis of combined metallic and dielectric structures.

In the third section, the CIEs are presented. First, VIEs, which represent a general form of CIEs, are considered. Then, degenerate forms are shown, enabling analysis of metallic surfaces with distributed loadings, wires with distributed loadings, and wires with concentrated loadings.

BIEs and CIEs are usually solved by the MoM. Sometimes, the MoM applied to BIEs and CIEs is hybridized with the FEM, Green's-function techniques, and asymptotic high-frequency techniques. A short overview of hybrid methods is given in the fourth section.

4.1 BIEs for Metallic Structures

Most often, losses in metallic structures are neglected and metallic structures are considered as PEC bodies (i.e., bodies made of a perfect electric conductor). This section is primarily devoted to the analysis of such bodies immersed into a homogeneous, linear, and isotropic medium. In some cases, however, the losses cannot be neglected. If in a lossy metallic structure the skin effect is pronounced, the structure can be considered as a body with an impedance boundary surface (Section 3.6). However, if the skin effect is not fully pronounced, the metallic structure must be treated in the same way as a lossy dielectric structure (Section 4.2).

Consider a PEC body situated in a homogeneous, linear, and isotropic medium [Figure 4.1(a)]. According to notation from Section 3.7, the space occupied by the PEC body is designated as region 0. Since the surrounding space can be a part of a multiple region, it is designated as region i, whose

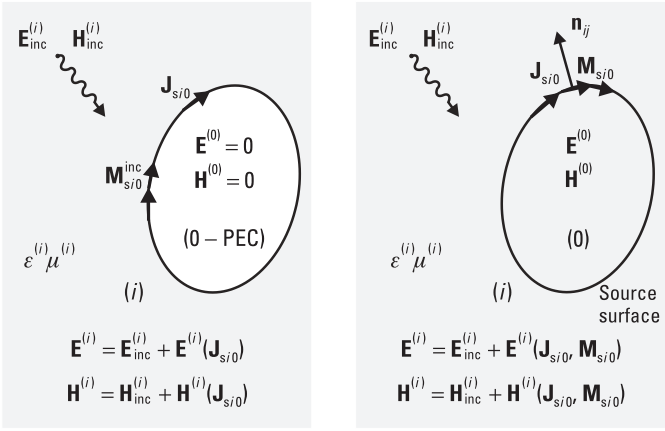

Figure 4.1 PEC body placed in time-harmonic incident electromagnetic field: (a) original problem, and (b) equivalent problem.

medium parameters are $\epsilon^{(i)}$ and $\mu^{(i)}$. In this section, region i is assumed to be infinite.

The body is placed in a time-harmonic incident electromagnetic field of angular frequency ω. The incident fields in regions i and 0 are denoted $\mathbf{E}^{(i)}_{\text{inc}}$, $\mathbf{H}^{(i)}_{\text{inc}}$, and $\mathbf{E}^{(0)}_{\text{inc}}$, $\mathbf{H}^{(0)}_{\text{inc}}$, respectively. The sources of the incident field are, generally, located in region i, as well as on the body surface in the form of magnetic currents,[1] $\mathbf{M}^{\text{inc}}_{si0}$. As a result, electric currents are induced over the body surface of density $\mathbf{J}_{si0}$. (The indices "$i0$" stress that the currents are on the boundary surface between regions i and 0.) These currents produce a scattered electromagnetic field in both regions given by vectors $\mathbf{E}^{(i)}(\mathbf{J}_{si0})$, $\mathbf{H}^{(i)}(\mathbf{J}_{si0})$, and $\mathbf{E}^{(0)}(\mathbf{J}_{si0})$, $\mathbf{H}^{(0)}(\mathbf{J}_{si0})$, respectively. The total fields in these regions (which are the sums of the corresponding incident and scattered fields) are denoted by $\mathbf{E}^{(i)}$, $\mathbf{H}^{(i)}$, $\mathbf{E}^{(0)}$, $\mathbf{H}^{(0)}$.

Once the induced electric currents $\mathbf{J}_{si0}$ are determined, all other quantities of interest can be easily evaluated. Namely, the total field inside the PEC body is zero (i.e., $\mathbf{E}^{(0)} = 0$ and $\mathbf{H}^{(0)} = 0$). Hence, the field in region i will not change if the PEC body is removed and region 0 is filled by the same medium as region i. The induced currents, $\mathbf{J}_{si0}$, are now located in an infinite homogeneous medium. Field scattered from the PEC body can be evaluated by closed-form integral expressions in terms of $\mathbf{J}_{si0}$ (Sections 3.3.1 and 3.3.2).

1. The influence of impressed sources in the form of electric currents placed at the PEC body surface is completely annihilated by induced currents. Hence, these kinds of sources are omitted.

Most often, the main goal of an analysis of PEC bodies is to determine a distribution of electric currents induced over the body surface. However, we can artificially introduce equivalent sources (electric and magnetic currents) on the body surface, which differ from the original induced currents, but produce the exact scattered field. The reason for introducing the equivalent sources is to provide a general approach. Consequently, consider for the moment a PEC cavity, complementary to the PEC body. The medium within the cavity is homogeneous, of parameters $\epsilon^{(i)}$ and $\mu^{(i)}$. The cavity is excited by arbitrary impressed magnetic currents placed on the boundary surface ($\mathbf{M}_{si0}$). As a result, electric currents are induced on this surface, such that the total field outside the cavity is zero. If these magnetic and electric currents are placed on the surface bounding the volume occupied by the original PEC body [Figure 4.1(b)], the scattered field in region i will not change. Now, there is a part of the scattered field due to magnetic currents given by vectors $\mathbf{E}^{(i)}(\mathbf{M}_{si0})$ and $\mathbf{H}^{(i)}(\mathbf{M}_{si0})$ in region i and vectors $\mathbf{E}^{(0)}(\mathbf{M}_{si0})$ and $\mathbf{H}^{(0)}(\mathbf{M}_{si0})$ in region 0. However, total electric currents on the surface now are not identical to the induced currents in the original problem. In addition, the field inside region 0 is not zero any more (i.e., in general, $\mathbf{E}^{(0)} \neq 0$ and $\mathbf{H}^{(0)} = 0$).

To formulate a BIE for evaluation of the scattered field starting from the above concept of equivalent currents, we consider the equivalent problem shown in Figure 4.1(b). Impose boundary conditions for the surface between regions i and 0. After setting $j = 0$ in (3.61a) and (3.61b) and adding the term $\mathbf{M}_{si0}^{\text{inc}}$ on the right side of (3.61a), we obtain

$$\mathbf{n}_{i0} \times (\mathbf{E}^{(i)} - \mathbf{E}^{(0)}) = -\mathbf{M}_{si0} - \mathbf{M}_{si0}^{\text{inc}} \qquad \mathbf{n}_{i0} \times (\mathbf{H}^{(i)} - \mathbf{H}^{(0)}) = \mathbf{J}_{si0} \tag{4.1a,b}$$

where $\mathbf{n}_{i0}$ is unit normal directed from region 0 to region i.

The total electric and magnetic field in both regions can be expressed in terms of the scattered field due to induced electric and magnetic currents and the incident field as

$$\mathbf{E}^{(i)} = \mathbf{E}^{(i)}(\mathbf{J}_{si0}) + \mathbf{E}^{(i)}(\mathbf{M}_{si0}) + \mathbf{E}_{\text{inc}}^{(i)} \tag{4.2a}$$

$$\mathbf{E}^{(0)} = \mathbf{E}^{(0)}(\mathbf{J}_{si0}) + \mathbf{E}^{(0)}(\mathbf{M}_{si0}) + \mathbf{E}_{\text{inc}}^{(0)} \tag{4.2b}$$

$$\mathbf{H}^{(i)} = \mathbf{H}^{(i)}(\mathbf{J}_{si0}) + \mathbf{H}^{(i)}(\mathbf{M}_{si0}) + \mathbf{H}_{\text{inc}}^{(i)} \tag{4.3a}$$

$$\mathbf{H}^{(0)} = \mathbf{H}^{(0)}(\mathbf{J}_{si0}) + \mathbf{H}^{(0)}(\mathbf{M}_{si0}) + \mathbf{H}_{\text{inc}}^{(0)} \tag{4.3b}$$

According to (3.50), (3.51), and (3.60), the scattered electric and magnetic fields on both faces of the boundary surface (the face in region i and the face in region 0) due to electric and magnetic currents are

$$\mathbf{E}^{(i)}(\mathbf{J}_{si0}) = -Z^{(i)} L_0^{(i)}(\mathbf{J}_{si0}) + \frac{\rho_{si0}}{2\epsilon^{(i)}}\mathbf{n}_{i0} \tag{4.4a}$$

$$\mathbf{E}^{(0)}(\mathbf{J}_{si0}) = -Z^{(i)} L_0^{(i)}(\mathbf{J}_{si0}) - \frac{\rho_{si0}}{2\epsilon^{(i)}}\mathbf{n}_{i0} \tag{4.4b}$$

$$\mathbf{E}^{(i)}(\mathbf{M}_{si0}) = K_0^{(i)}(\mathbf{M}_{si0}) - \frac{1}{2}\mathbf{M}_{si0} \times \mathbf{n}_{i0} \tag{4.5a}$$

$$\mathbf{E}^{(0)}(\mathbf{M}_{si0}) = K_0^{(i)}(\mathbf{M}_{si0}) + \frac{1}{2}\mathbf{M}_{si0} \times \mathbf{n}_{i0} \tag{4.5b}$$

$$\mathbf{H}^{(i)}(\mathbf{J}_{si0}) = -K_0^{(i)}(\mathbf{J}_{si0}) + \frac{1}{2}\mathbf{J}_{si0} \times \mathbf{n}_{i0} \tag{4.6a}$$

$$\mathbf{H}^{(0)}(\mathbf{J}_{si0}) = -K_0^{(i)}(\mathbf{J}_{si0}) - \frac{1}{2}\mathbf{J}_{si0} \times \mathbf{n}_{i0} \tag{4.6b}$$

$$\mathbf{H}^{(i)}(\mathbf{M}_{si0}) = -Y^{(i)} L_0^{(i)}(\mathbf{M}_{si0}) + \frac{\tau_{si0}}{2\mu^{(i)}}\mathbf{n}_{i0} \tag{4.7a}$$

$$\mathbf{H}^{(0)}(\mathbf{M}_{si0}) = -Y^{(i)} L_0^{(i)}(\mathbf{M}_{si0}) - \frac{\tau_{si0}}{2\mu^{(i)}}\mathbf{n}_{i0} \tag{4.7b}$$

where $Z^{(i)}$ and $Y^{(i)}$ are the intrinsic impedance and admittance of the medium in region i, $L_0^{(i)}$ and $K_0^{(i)}$ are the operators defined by (3.44) applied in the principal value sense, and ρ_{si0} and τ_{si0} are the densities of electric and magnetic charges, respectively, placed over the boundary surface. Incident electric and magnetic fields on both faces of the boundary surface are related through the boundary conditions

$$\mathbf{n}_{i0} \times (\mathbf{E}_{\text{inc}}^{(i)} - \mathbf{E}_{\text{inc}}^{(0)}) = -\mathbf{M}_{si0}^{\text{inc}} \qquad \mathbf{H}_{\text{inc}}^{(i)} = \mathbf{H}_{\text{inc}}^{(0)} \tag{4.8a,b}$$

Starting from boundary condition (4.1), various integral equations for induced currents $\mathbf{J}_{si0}$ or $\mathbf{M}_{si0}$ can be obtained. By solving these equations, the electric and magnetic currents become known, and the scattered field can be easily evaluated.

4.1.1 EFIEs and MFIEs

The basic BIEs are the EFIE and the MFIE. These equations are ascribed originally to Maue [1]. His paper, published in 1949, can be regarded as a generalization of the earlier work by Pocklington [2], Hallén [3], and others. However, the application of these equations to real numerical solutions of scattering problems began only 15 years later with the work of Andreasen [4], who used the EFIE, and Oshiro and Mitzner [5], who used the MFIE. The EFIE and the MFIE written for PEC bodies represent the starting point in the development of most BIEs.

The EFIE is based on the boundary condition (4.1a). Since the total electric field is equal to zero inside the PEC body ($\mathbf{E}^{(0)} = 0$) and there are no magnetic currents induced on its surface ($\mathbf{M}_{si0} = 0$), we can write (4.1a) in alternative forms:

$$\mathbf{n}_{i0} \times \mathbf{E}^{(i)} = -\mathbf{M}_{si0}^{\text{inc}} \qquad \mathbf{n}_{i0} \times \mathbf{E}^{(0)} = 0 \tag{4.9a,b}$$

After replacing (4.4a) into (4.2a) [with ($\mathbf{M}_{si0} = 0$) and (4.2a) into (4.9a)], and after substituting (4.4b) into (4.2b) [with ($\mathbf{M}_{si0} = 0$) and (4.2b) into (4.9b)], the EFIE for induced currents $\mathbf{J}_{si0}$ is written in two equivalent ways[2]:

$$Z^{(i)}\mathbf{n}_{i0} \times L_0^{(i)}(\mathbf{J}_{si0}) = \mathbf{n}_{i0} \times \mathbf{E}_{\text{inc}}^{(i)} + \mathbf{M}_{si0}^{\text{inc}} \tag{4.10a}$$

$$Z^{(i)}\mathbf{n}_{i0} \times L_0^{(i)}(\mathbf{J}_{si0}) = \mathbf{n}_{i0} \times \mathbf{E}_{\text{inc}}^{(0)} \tag{4.10b}$$

So, by solving the EFIE we force both conditions (4.9a,b) [i.e., that the tangential component of electric field be equal to known $\mathbf{n}_{i0} \times \mathbf{M}_{si0}^{\text{inc}}$ (or to zero if $\mathbf{M}_{si0}^{\text{inc}} = 0$) just outside the body, and to zero just inside the body]. According to the uniqueness theorem (Section 3.10.1), the EFIE provides a unique solution at all frequencies for the scattered field outside the PEC body. A zero-field inside it and unique induced currents are obtained only away from the resonant frequencies.

The MFIE is based on the boundary condition (4.1b). Since the total magnetic field is equal to zero inside the PEC body ($\mathbf{H}^{(0)} = 0$), we can write (4.1b) in alternative forms:

$$\mathbf{n}_{i0} \times \mathbf{H}^{(i)} = \mathbf{J}_{si0} \qquad \mathbf{n}_{i0} \times \mathbf{H}^{(0)} = 0 \tag{4.11a,b}$$

After substituting (4.6b) into (4.3b) and (4.3b) into (4.11b), or after substituting (4.6a) into (4.3a) and (4.3a) into (4.11a), the MFIE is obtained in the form

2. Equation (4.10b) is obtained from (4.10a), and vice versa, by using (4.8a).

$$\mathbf{n}_{i0} \times K_0^{(i)}(\mathbf{J}_{si0}) + \frac{1}{2}\mathbf{J}_{si0} = \mathbf{n}_{i0} \times \mathbf{H}_{\text{inc}}^{(0)} \tag{4.12}$$

However, induced electric currents $\mathbf{J}_{si0}$ are unknown. Hence, by solving the MFIE, we only force [through condition (4.11b)] the tangential component of the magnetic field to be equal to zero just inside the body. Consequently, electric and magnetic fields ($\mathbf{E}^{(0)}$ and $\mathbf{H}^{(0)}$) are zero at all points inside the body, except at resonant frequencies (see Section 3.10.1). By satisfying (4.11b), we indirectly force (4.9b), but only away from the resonant frequencies. The MFIE provides a unique solution for all quantities (unique scattered field outside the PEC body and induced currents, and a zero field inside) only away from the resonant frequencies.

For example, consider a spherical scatterer of radius a = 1.41m excited by a plane wave. The scatterer is modeled by m = 24 bilinear surface patches,[3] as is shown in the inset of Figure 4.2. The model is inscribed into the original sphere, which causes a frequency shift of numerical results towards higher frequencies. The first-order approximation (rooftop)[4] is used for currents over each patch. The unknown coefficients of this approximation are found by using the Galerkin method[5] to solve the EFIE and the MFIE. Figure 4.2 shows the

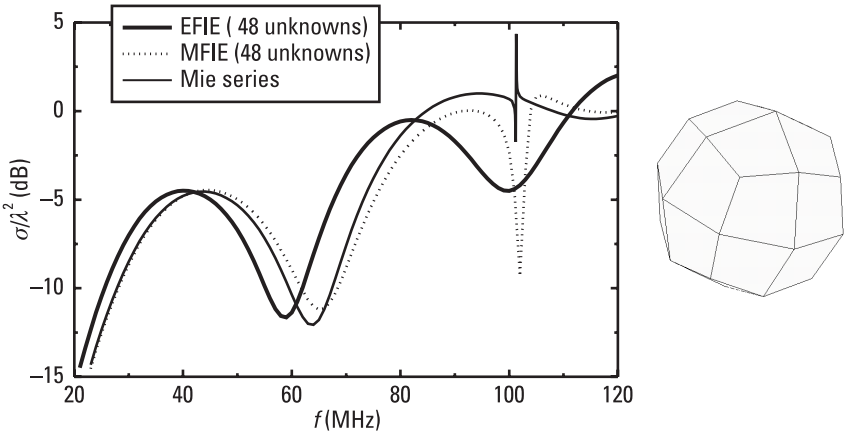

Figure 4.2 Monostatic RCS of a sphere scatterer shown in the inset a = 1.41m, normalized by λ^2 versus frequency obtained by solving the EFIE and the MFIE.

3. Bilinear surfaces are explained in Chapter 5. This kind of patch is used in all examples in this book.
4. Approximations for the current distribution are explained in Chapter 6.
5. The Galerkin method and other testing procedures are explained in Chapter 7. The Galerkin method is used by default in all examples in this book.

monostatic radar cross section (RCS) normalized by λ^2 versus frequency. At lower frequencies, the numerical results show an acceptable agreement with the exact RCS obtained in the form of the Mie series [6]. At higher frequencies, the agreement worsens. In particular, in the vicinity of the first interior resonant frequency of the PEC sphere, a spurious response can be observed in the RCS.

An exact solution of the EFIE and MFIE would have a spurious field only exactly at the resonant frequency. In particular, the spurious solution would not influence the scattered field (outside the PEC body) in the case of the EFIE. However, both numerical solutions fail not only at the resonant frequency, but in a wide range near the resonance.

4.1.2 Generalized Scalar Formulation of the EFIE and MFIE

The EFIE and MFIE can be derived starting from the boundary conditions (4.9b) and (4.11b), respectively, written for points just inside the PEC body. These boundary conditions can be written in two common forms

$$\mathbf{n}_{i0} \times \mathbf{F} = 0 \qquad \mathbf{F}_{\tan} = 0 \tag{4.13a,b}$$

where $\mathbf{F}$ stands for $\mathbf{E}^{(0)}$ (for the EFIE) or $\mathbf{H}^{(0)}$ (for the MFIE), and $\mathbf{F}_{\tan}$ is the tangential component of $\mathbf{F}$. [The cross product of (4.13a) and unit normal $\mathbf{n}_{i0}$ gives (4.13b).] From the theoretical point of view, these two forms are identical. However, when the corresponding integral equations are solved by the MoM, the choice of the optimal form depends on the test procedure adopted, as will be shown in Chapter 8. In some cases, none of these forms is convenient for implementation. Hence, in what follows we shall consider how various forms of (4.13) can be obtained.

A numerical solution of a particular problem requires that the vector formulations of the basic field integral equations be replaced by the corresponding scalar formulations. As the first step, a local coordinate system (with coordinates p and s) is placed on the boundary surface [e.g., see Figure 5.8(a)]. The tangential component of $\mathbf{F}$ can be decomposed in the local ps-coordinate system as

$$\mathbf{F}_{\tan} = F_p \mathbf{i}_p + F_s \mathbf{i}_s \tag{4.14}$$

where $\mathbf{i}_p$ and $\mathbf{i}_s$ are unit vectors along the p- and s-coordinate lines. Thus, the vector equation (4.13b) is imposed by requiring each component of $\mathbf{F}_{\tan}$ to be zero; that is,

$$F_p = 0 \qquad F_s = 0 \tag{4.15a,b}$$

The pair (4.15a,b) represents only one possible scalar formulation of (4.13b). Although this scalar formulation looks simple, most often it is not the optimal one.

In general, the local coordinates p and s need not be lengths (for example, they can be angular coordinates or, more generally, parametric coordinates), and the unit vectors, $\mathbf{i}_p$ and $\mathbf{i}_s$, need not be orthogonal. Hence, the components F_p and F_s are not easily obtainable, but the p and s-projections of the vector $\mathbf{F}_{\tan}$ are easily calculated as

$$F'_p = \mathbf{i}_p \cdot \mathbf{F}_{\tan} = \mathbf{i}_p \cdot \mathbf{F} \qquad F'_s = \mathbf{i}_s \cdot \mathbf{F}_{\tan} = \mathbf{i}_s \cdot \mathbf{F} \tag{4.16a,b}$$

After replacing (4.14) into (4.16), we obtain the system of two linear equations in terms of F_p and F_s.

$$F'_p = F_p + (\mathbf{i}_p \cdot \mathbf{i}_s)F_s \qquad F'_s = (\mathbf{i}_p \cdot \mathbf{i}_s)F_p + F_s \tag{4.17a,b}$$

By solving the system, F_p and F_s are expressed in terms of F'_p and F'_s as

$$F_p = \frac{F'_p - (\mathbf{i}_p \cdot \mathbf{i}_s)F'_s}{1 - (\mathbf{i}_p \cdot \mathbf{i}_s)^2} \qquad F_s = \frac{F'_s - (\mathbf{i}_p \cdot \mathbf{i}_s)F'_p}{1 - (\mathbf{i}_p \cdot \mathbf{i}_s)^2} \tag{4.18a,b}$$

After replacing (4.18) into (4.15), the scalar formulation of (4.13b) is obtained in a more practical form:

$$\alpha_{pp}F'_p + \alpha_{ps}F'_s = 0 \qquad \alpha_{sp}F'_p + \alpha_{ss}F'_s = 0 \tag{4.19a,b}$$

where the weighting coefficients α_{pp}, α_{ps}, α_{sp}, and α_{ss} are given by

$$\alpha_{pp} = \alpha_{ss} = \frac{1}{1 - (\mathbf{i}_p \cdot \mathbf{i}_s)^2} \qquad \alpha_{ps} = \alpha_{sp} = -\frac{(\mathbf{i}_p \cdot \mathbf{i}_s)}{1 - (\mathbf{i}_p \cdot \mathbf{i}_s)^2} \tag{4.20a,b}$$

Similarly, we can require that p and s-component of (4.13a) are equal to zero. The unit normal $\mathbf{n}_{i0}$ is expressed in terms of $\mathbf{i}_p$ and $\mathbf{i}_s$ as

$$\mathbf{n}_{i0} = \frac{\mathbf{i}_p \times \mathbf{i}_s}{|\mathbf{i}_p \times \mathbf{i}_s|} \tag{4.21}$$

Using (4.14), (4.18), and (4.21) and after some manipulations, we obtain the p- and s-component of (4.13a) in the form (4.19), where the weighting coefficients are

$$\alpha_{pp} = \alpha_{ss} = 0 \qquad \alpha_{ps} = \alpha_{sp} = \frac{1}{|\mathbf{i}_p \times \mathbf{i}_s|} \tag{4.22a,b}$$

The scalar formulations of the boundary condition, based on the two vector conditions (4.13a,b), are particular cases of (4.19). In general, the weighting coefficients (α_{pp}, α_{ps}, α_{sp}, and α_{ss}) are arbitrary functions of the local p- and s-coordinates, such that equations (4.19a,b) are linearly independent; that is,

$$\det\begin{vmatrix} \alpha_{pp} & \alpha_{ps} \\ \alpha_{sp} & \alpha_{ss} \end{vmatrix} = \alpha_{pp}\,\alpha_{ss} - \alpha_{ps}\,\alpha_{sp} \neq 0 \tag{4.23}$$

Hence, (4.19), with coefficients α_{pp}, α_{ps}, α_{sp}, and α_{ss} satisfying (4.23), can be regarded as a generalized scalar form of the basic field integral equations (EFIE and MFIE).

Various choices of these coefficients are possible. The simplest form of equations (4.19a,b) is obtained by adopting $\alpha_{pp} = \alpha_{ss} = 1$ and $\alpha_{ps} = \alpha_{sp} = 0$. However, as mentioned, the optimal choice of these coefficients depends on the method used for solution of the integral equation considered, as will be shown in Chapter 8.

4.1.3 Equivalent Sources and EBCs

One of the most difficult problems in the application of the basic field integral equations is the evaluation of singular integrals.[6] Such integrals occur because the source surface S' (i.e., the surface over which the unknown currents are distributed) coincides with the field surface S (i.e., the surface at which the field due to these currents is evaluated). The evaluation of the singular integrals can be avoided only if S' and S have no common points. One possibility is to introduce equivalent sources, and the other is to introduce the concept of EBCs.

We shall first consider the concept of equivalent sources. According to the uniqueness theorem (Section 3.10.1), the field scattered from a PEC body is uniquely determined for a given incident field if the component of the electric field tangential to the body surface (which in this section coincides with the field surface, S) just outside the body is zero. As long as this condition is

6. These integrals are hidden in operators L_0 and K_0 in (4.10b) and (4.12).

satisfied, the field in region i will not change if the original induced currents (located on the body surface) are replaced by arbitrary electric or magnetic current distributions inside the region 0. These fictitious currents are referred to as equivalent sources. An example is shown in Figure 4.3(a), where the equivalent sources are located on a surface S', enclosed by the body surface S.

Starting from the boundary condition (4.9a), the EFIE for equivalent sources is easily derived. The original problem [Figure 4.1(a)] and the equivalent problem have identical fields only in region i. The fields inside region 0 are different. Hence, integral equations for equivalent sources cannot be derived starting from the boundary condition (4.11b), which is written for the face of S just inside the body (i.e., in region 0). Consequently, a MFIE for equivalent sources cannot be established.

If equivalent sources are not placed too close to S, the field integrals are neither singular nor quasisingular. Generally, the equivalent sources can be distributed within a part of the volume of region 0. In many numerical approaches, however, they are distributed over a closed surface (S') inside region 0 [Figure 4.3(a)]. (For a problem with a rotational symmetry, the equivalent sources can be distributed along the axis of rotation of S.) Finally, these sources can have the form of point sources placed on S' [7]. Point sources may be convenient as their field can be evaluated analytically, while the evaluation of the field due to the surface sources requires a surface (dual) integration.

An important problem is the choice of the number and position of the point sources. Suppose the surface S is smooth enough and the component of the incident field tangential to S is a continuous and slowly varying function.

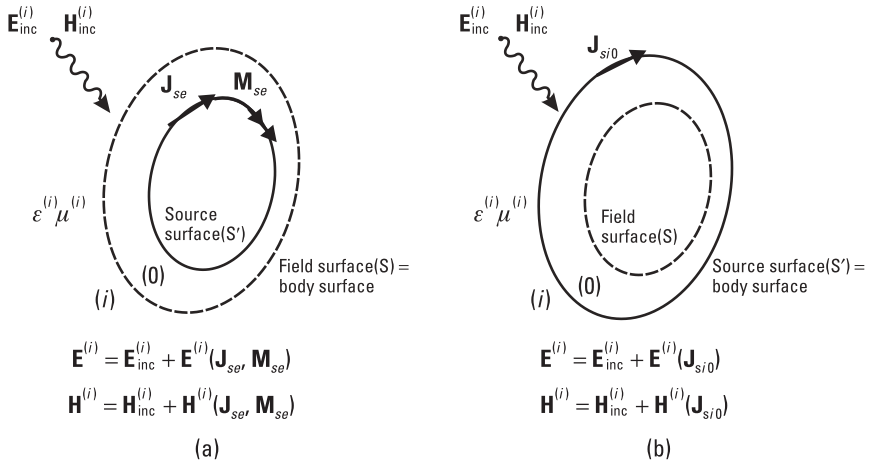

Figure 4.3 (a) Equivalent sources with respect to region i and (b) the EBC.

Then the component of the actual scattered field tangential to S should also be a continuous and slowly varying function. Such a scattered field can be provided by the equivalent sources if the distance between neighboring point sources is comparable or smaller than the distance between the sources and S. For a given number of point sources, this condition defines the maximum size of S'. Theoretically, S' can be arbitrarily small. However, by decreasing its size, contributions of various equivalent sources to the field at any point on S become similar. If the EFIE is solved by the MoM, all unknown coefficients enter into a linear equation with similar weights and the system of linear equations becomes less stable. Hence, the surface S' must not be too small. If the number of point sources is increased, S' should approach S, and vice versa, if S' approaches S, the number of unknown point sources should be increased.

A particular problem occurs when the boundary surface S is not smooth enough (e.g., when it has sharp wedges). At some parts of S, the tangential component of the scattered field is rapidly varying function. To approximate such variations, the equivalent point sources should be placed very close to the critical parts of S, leading to a dramatic increase of the number of point sources. Finally, point sources cannot be applied to open metallic surfaces.

The concept of equivalent sources is simple and straightforward. It can be generalized to multiple-region problems [8]. However, the number of unknowns needed for analysis of nonsmooth bodies seems to be much greater than if the original sources are used. In what follows, the concept of equivalent sources will be used only in the form of the image theory to reduce the number of unknowns used in the analysis.

In contrast to the equivalent sources, we can retain the original sources on the body surface (which now coincides with the source surface, S') and adopt the field surface S to be enclosed by S' [Figure 4.3(b)], thus again separating S and S'. Equations (4.9a) and (4.11b) imposed for such a field surface can be regarded as extensions of the classic boundary conditions. For electrostatics, an EBC was proposed by Smythe [9]. However, the EBCs have been widely used in the numerical modeling of time-harmonic electromagnetic fields only after the work of Waterman [10].

Consider the homogeneous region bounded by S' and the part of that region bounded by S. It is theoretically impossible to annihilate an incident field only inside S without annihilating it within the whole region bounded by S'. By imposing an EBC on S, it is required that the tangential component of the electric or magnetic field is zero. This ensures the field inside S is zero, except at resonant frequencies of the volume encapsulated by S. Hence, the field is zero in the whole volume encapsulated by S'. In conclusion, the EBCs provide uniqueness of the field solution in all cases, except at internal resonant frequencies of S.

By increasing the distance between the source and field points, the evaluation of field integrals becomes easier. Theoretically, the surface S can be arbitrarily small. However, by shrinking S, the field due to a source element becomes more similar at various points of S. In a MoM solution of a field integral equation, an unknown coefficient enters into various linear equations with similar weights and the system of linear equations becomes less stable. Hence, it is desirable that S is relatively close to the source surface, S'.

The main disadvantage of EBC, when compared with the exact boundary conditions, is that there is no general technique that provides an optimal choice of the field surface. For open metallic surfaces, the EBC cannot be used. However, the EBC is particularly suitable for the thin-wire analysis [11], where the field surface, S, can degenerate to the wire axis, as will be shown in Section 4.1.10.

4.1.4 Spurious Resonances

If the basic field integral equations are exactly satisfied and the medium is lossless, a spurious field can theoretically occur in region 0 only at discrete (resonant) frequencies (f_r). This spurious field is related to the cavity mode that can exist in region 0 enclosed by a PEC. In that case, region 0 represents an ideal cavity. It is perfectly shielded without losses in the conductor or dielectric, and it does not have a feed or load.

In a numerical solution, the basic field integral equations are not exactly satisfied. The components of the electric and magnetic field tangential to the boundary surface, on the face in region 0, deviate from zero, and there is a power flow between regions i and 0 (described by the Poynting vector). Region 0 now resembles a real cavity. Parts of the boundary surface with an average power flow into the cavity can be considered as a feed, and parts with an average power flow outwards can be considered as a load. Such a flow can significantly modify the near field from both faces of the boundary surface and significantly influences the scattered field at interior resonances.

If the boundary conditions are satisfied relatively well, the coupling of the cavity with the feed and the load is relatively weak. If the cavity is filled by a medium with relatively small losses, the quality factor (Q) of the cavity is relatively high. As a result, a strong spurious field exists in region 0 not only at a resonant frequency, but also in a narrow nearby frequency range. With a better satisfaction of the boundary conditions, the coupling of the cavity with the feed and the load decreases, and the quality factor increases. The frequency range with a strong spurious field decreases. In the limiting case when the tangential component of either the electric or magnetic field (on the face of

the boundary surface in region 0) tends to zero, this frequency range reduces to the single resonant frequency.

For example, consider a cubical scatterer of side a = 2m, excited by an incident plane wave of intensity E = 1 V/m, (Figure 4.4, inset [12]). The analysis is based on the EFIE solved by the Galerkin method. The cube is modeled by 24 patches and the second order approximation is used for currents over each patch, resulting in 192 unknown coefficients. Figure 4.4(a) and 4.4(b) shows the magnitudes of the z-component of the electric field and the y-component of the magnetic field, respectively, at f = 105 MHz, in the xOy-plane, for $-1.5a < x, y < 1.5a$. Both field components are small inside the cube ($a < x, y < a$). Figure 4.4(c) and 4.4(d) shows the same field components at f = 106 MHz. A spurious field, very close to the TE_{101} cavity mode, can be observed. The maximum of the z-component of the electric field is in the center of the cube (i.e., at the coordinate origin).

Figure 4.5(a) shows the electric-field strength in the center of the cube as a function of frequency for various geometrical models and numbers of unknowns used for the current approximation. The results are obtained by using the geometrical model of the cube made of m = 24, 54, and 96 patches (Figure 4.5, inset). A rooftop approximation of currents results in N = 48,

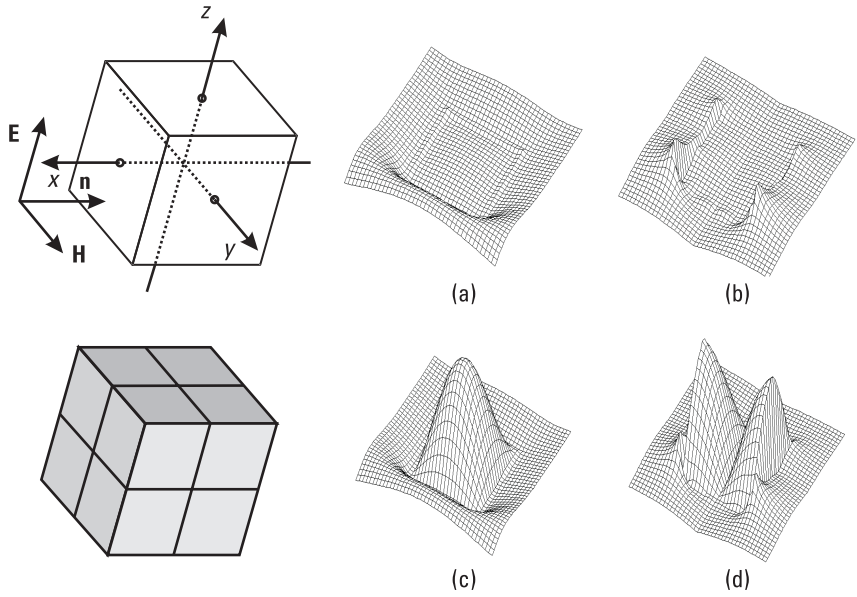

Figure 4.4 Distribution of electromagnetic field over xOy-plane ($-1.5a < x, y < 1.5a$) for the cube scatterer of side a shown in the inset: (a) E_z and (b) H_y at f = 105 MHz, respectively; (c) E_z and (d) H_y at f = 106 MHz, respectively.

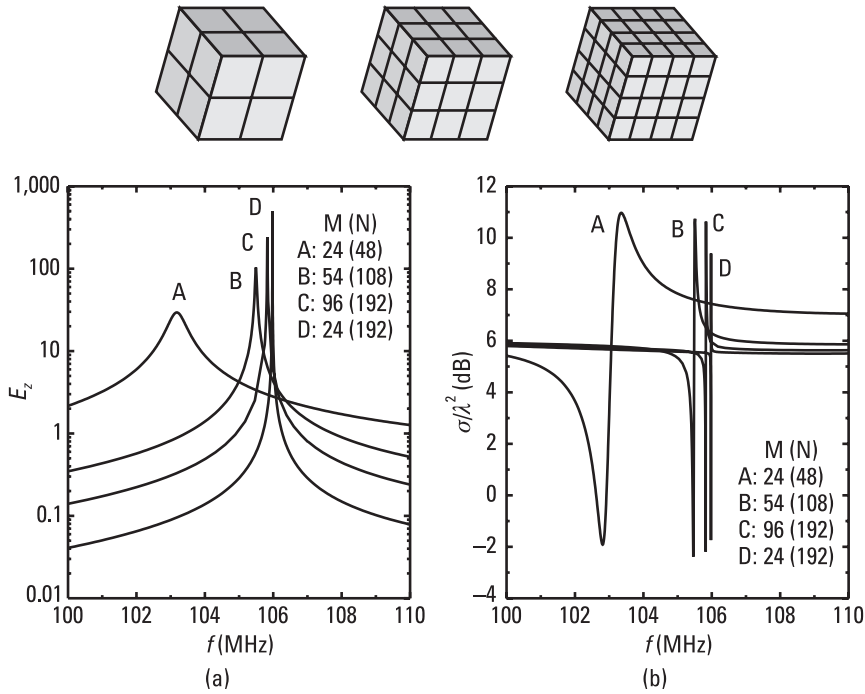

Figure 4.5 (a) Maximum electric-field strength of spurious electric field and (b) monostatic RCS of the cubical scatterer for various numbers of unknowns and geometrical models shown in the inset.

108, and 192 unknowns, respectively. A significant spurious field appears in a frequency range near the resonant frequency. By increasing the number of unknowns, the width of the frequency range decreases and the position of the resonant frequency is moved toward the theoretical value. At the same time, the maximum field value at the resonant frequency increases. This frequency range can be further decreased if the second-order approximation is used for currents for the model with $m = 24$ patches, resulting in $N = 192$ unknowns. When the cube is modeled by $m = 24$ patches and the third-order approximation is used for currents over each patch, resulting in $N = 108$ unknowns, this frequency range is so narrow that it can be hardly detected. Hence, the corresponding results are not shown in Figure 4.5(a). Figure 4.5(b) shows the monostatic RCS normalized by λ^2 versus frequency, for the same cases as in Figure 4.5(a). The scattered field is maximally disturbed at the same frequencies at which the spurious field has the maximum strength, as expected.

There exist several techniques for eliminating spurious resonances. One way is to introduce losses in region 0 (e.g., by inserting a lossy dielectric object

into this region [13]). The coupling of such a cavity with the numerically introduced spurious feed and load is practically unaffected, whereas the quality factor can be significantly decreased. For example, insert a lossy dielectric cube of side $b = a/4$ and relative permittivity $\epsilon = 1 - j\epsilon''$ into the cubical scatterer (Figure 4.6, inset). Figure 4.6 shows the monostatic RCS versus frequency for various values of ϵ''. For small ϵ'' (i.e., for small losses), the spurious resonance is not completely suppressed. So, in a numerical solution of the EFIE, techniques for suppression of the spurious resonances should be applied even if the medium surrounding the PEC body has small losses. For large enough ϵ'', the spurious resonance is completely suppressed. The shortcoming of the present technique is that the analysis of the inserted body is much more complicated than the analysis of the PEC body. In addition to the EFIE written for the PEC body, a specific field integral equation for the dielectric boundary surface should be written (see Section 4.2). The solution results in a significantly greater number of unknowns than for the original problem.

The spurious resonances can be suppressed without introducing additional equations and unknowns. Sarkar and Rao proposed the minimum-norm solution of the EFIE [14]. This method starts from the system of linear equations regularly obtained by the MoM applied to the EFIE. The least-squares error, with which the system is satisfied, is minimized together with the norm of the unknown coefficients. The problem in the application of this method is the choice of weights with which the least-squares error and the norm are taken into account.

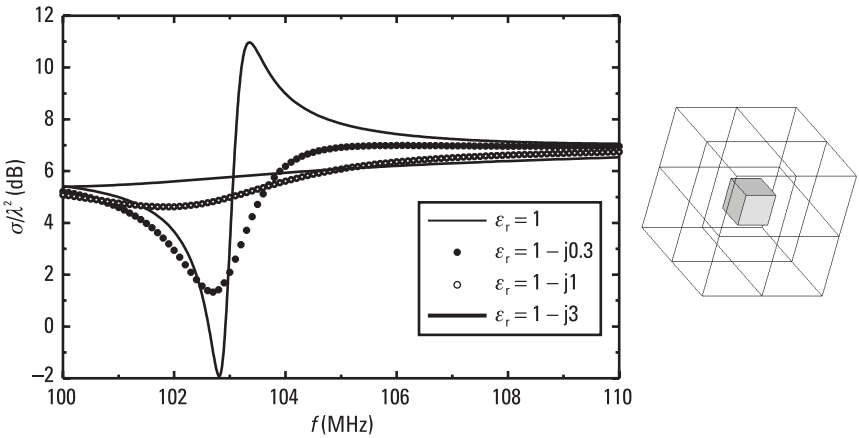

Figure 4.6 Monostatic RCS of metallic cubical scatterer with inserted dielectric cube (a = 2m, $b = a/4$, m = 24), normalized by λ^2, versus frequency.

Other techniques for the elimination of the spurious resonances do not change the given structure, but modify the set of equations used for their analysis, as is explained in the following sections.

4.1.5 AFIE

The basic system of linear equations obtained by starting from either the EFIE or MFIE, can be extended by equations aimed at eliminating resonant modes. An overdetermined system of equations is thus obtained, which can be solved by least-squares techniques. A particular problem is the choice of the weights of individual equations in the overdetermined system.

For example, matching points can be adopted not only on the surface of the body, but also inside the body at places where the resonant mode is known to be nonzero [15]. This makes the choice of the matching points somewhat difficult. The method is sometimes called the augmented boundary condition (ABC) method [16]. However, it belongs to the same group as the method described below.

To eliminate the spurious resonances from the solution based on the EFIE, Yaghjian [17, 18] adds to the EFIE the equation based on the boundary condition for the component of the electric field normal to the boundary surface. This boundary condition can be imposed either just outside the PEC body or just inside it as

$$\mathbf{n}_{i0} \cdot \mathbf{E}^{(i)} = \frac{\rho_{si0}}{\epsilon^{(i)}} \qquad \mathbf{n}_{i0} \cdot \mathbf{E}^{(0)} = 0 \tag{4.24a,b}$$

Both boundary conditions lead to the same field integral equation. For example, after substituting (4.4b) into (4.2b) (with $\mathbf{M}_{si0} = 0$), and resulting (4.2b) into condition (4.24b), this integral equation can be written as

$$Z^{(i)}\mathbf{n}_{i0} \cdot L_0^{(i)}(\mathbf{J}_{si0}) + \frac{\rho_{si0}}{2\epsilon^{(i)}} = \mathbf{n}_{i0} \cdot \mathbf{E}_{\text{inc}}^{(0)} \tag{4.25}$$

This equation and the EFIE given by (4.10b) represent two components of the integral equation obtained from the condition that the total electric field just inside the PEC body is zero. Hence, this integral equation is referred to as the *augmented electric field integral equation* (AEFIE). In a similar way, the MFIE can be augmented with the equation based on the boundary condition for the component of magnetic field normal to the boundary surface, resulting in the *augmented magnetic field integral equation* (AMFIE).

When these equations are solved by the MoM, an overdetermined system of linear algebraic equations results with theoretically 50% more equations than the number of unknowns. At any matching point, two equations are obtained corresponding to the standard EFIE or MFIE and one resulting from the boundary condition for the normal component of the electric or magnetic field, respectively.

It is similarly possible to start from the simultaneous system of equations resulting from both the EFIE and MFIE. When these equations are solved by the MoM, an overdetermined system of linear algebraic equations results with theoretically 100% more equations than the number of unknowns. At any matching point, two equations are obtained corresponding to the EFIE and two equations corresponding to the MFIE.

The shortcoming of all methods mentioned in this section is that they result in an overdetermined system of linear equations, which is usually solved in the least-squares sense. The number of basic operations needed for the system solution using direct techniques (e.g., the Gaussian elimination) is $K_o = MN^2 + N^3/3$, where N is number of unknowns and M is number of equations ($M > N$). On the other hand, the solution of a regular system of N equations, occurring in the case of the EFIE and the MFIE, requires only $K_r = N^3/3$ basic operations. The number of basic operations needed for the solution of an overdetermined system is, in the best case ($M \approx N$), four times greater than for a regular system. In particular, for the overdetermined system obtained by solving the AEFIE, $M = 1.5N$ (in the best case) and the ratio of the basic operations needed for solution of the overdetermined and the corresponding regular system is $K_o/K_r = 5.5$. In addition, the overdetermined systems are much less stable than the determined systems obtained by the EFIE. Conclusively, the methods resulting in overdetermined systems of linear equations are not recommended in practice.

4.1.6 CFIE

Most often the spurious resonances are eliminated from the solution by using the CFIE. This equation is established in different forms through the works of Oshiro and Mitzner [5], Poggio and Miller [19], and Mautz and Harrington [20].

The CFIE is based on combination of basic boundary conditions, (4.9b) and (4.11b), which can be written in alternate forms as

$$\alpha(\mathbf{E}^{(0)})_{\tan} + \mathbf{n}_{i0} \times \mathbf{H}^{(0)} = 0 \qquad \alpha(\mathbf{H}^{(0)})_{\tan} - \mathbf{n}_{i0} \times \mathbf{E}^{(0)} = 0 \tag{4.26a,b}$$

[The cross-product of unit normal vector $\mathbf{n}_{i0}$ and (4.26a), in which α is replaced by its reciprocal value, results in (4.26b).] The same equation can be obtained after some manipulation from the basic boundary conditions (4.9a) and (4.11c). Hence, in what follows we shall consider only the first form. Let us examine if this combined boundary condition provides a unique solution for fields inside and outside region i.

There are no sources in region 0, so that the real part of the complex power P flowing into the region should be zero or greater; that is,

$$\mathrm{Re}\{P\} \geq 0 \qquad P = -\oint_{S_{i0}} [\mathbf{E}^{(0)} \times (\mathbf{H}^{(0)})^*] \cdot \mathbf{n}_{i0}\, dS \qquad (4.27a,b)$$

The mixed product in (4.27b) can be transformed to $[(\mathbf{H}^{(0)})^* \times \mathbf{n}_{i0}] \cdot \mathbf{E}^{(0)}$. Using the combined boundary condition (4.26a), the complex power P can be written as

$$P = -\alpha^* \oint_{S_{i0}} |\mathbf{E}_{\tan}^{(0)}|^2\, dS \qquad (4.28)$$

Assuming the real part of α is greater than zero, the real part of the complex power given by this expression can be only nonpositive. However, according to (4.27a), this real part can be only nonnegative. Hence, by imposing (4.26a), this real part is forced to be zero, and the tangential components of both $\mathbf{E}^{(0)}$ and $\mathbf{H}^{(0)}$ are also forced to be zero. In conclusion, imposing the combined boundary condition yields a unique solution for the field inside both regions i and 0, even at resonant frequencies.

After replacing (4.4b) into (4.2b), (4.6b) into (4.3b), and resulting (4.2b) and (4.3b) (with $\mathbf{M}_{si0} = 0$) into (4.26a), the CFIE is obtained in the form

$$\alpha Z^{(i)} [L_0^{(i)}(\mathbf{J}_{si0})]_{\tan} + \mathbf{n}_{i0} \times K_0^{(i)}(\mathbf{J}_{si0}) + \frac{1}{2}\mathbf{J}_{si0} = \alpha(\mathbf{E}_{\mathrm{inc}}^{(0)})_{\tan} + \mathbf{n}_{i0} \times \mathbf{H}_{\mathrm{inc}}^{(0)} \qquad (4.29)$$

A particular problem in applying the CFIE is the choice of α. The role of this constant, normalized by the intrinsic impedance, is to enter the electric and magnetic fields into the equation with approximately equal weights. Obviously, for a plane wave, $\alpha = 1/Z^{(i)}$ should be adopted. Fortunately, numerical experiments show that results are relatively insensitive to the choice of α. In

contrast to the AFIE, the CFIE leads to a determined system of linear equations, as in the case of the EFIE and the MFIE.

For example, consider a spherical scatterer of radius a = 1.41m, excited by a plane wave. The sphere is modeled by m = 54 bilinear surfaces (Figure 4.7, inset). The first-order approximation is used for currents over each patch (rooftop approximation). Figure 4.7 shows the monostatic RCS normalized by πa^2 versus $2\pi a/\lambda$. The results are compared with those obtained by solving the EFIE. Obviously, the CFIE successfully eliminates the spurious resonances from the solution.

4.1.7 Generalized Scalar Formulation of CFIE

As stated in Section 4.1.2, generalized scalar formulations of a vector integral equation may have certain advantages in a MoM solution in comparison with just finding scalar projections of the vector equation. Hence, we derive here the generalized scalar formulation of CFIE.

The boundary conditions (4.26a,b) can be written in common form as

$$\alpha \mathbf{F}_{\tan} + \mathbf{n}_{i0} \times \mathbf{G} = 0 \tag{4.30}$$

where $\mathbf{F}$ and $\mathbf{G}$ stand for $\mathbf{E}^{(0)}$ and $\mathbf{H}^{(0)}$ in the case of (4.26a) and for $\mathbf{H}^{(0)}$ and $-\mathbf{E}^{(0)}$ in the case of (4.26b).

To obtain a scalar formulation of (4.30), adopt a local *ps*-coordinate system placed at the boundary surface [e.g., see Figure 5.8(a)]. The tangential

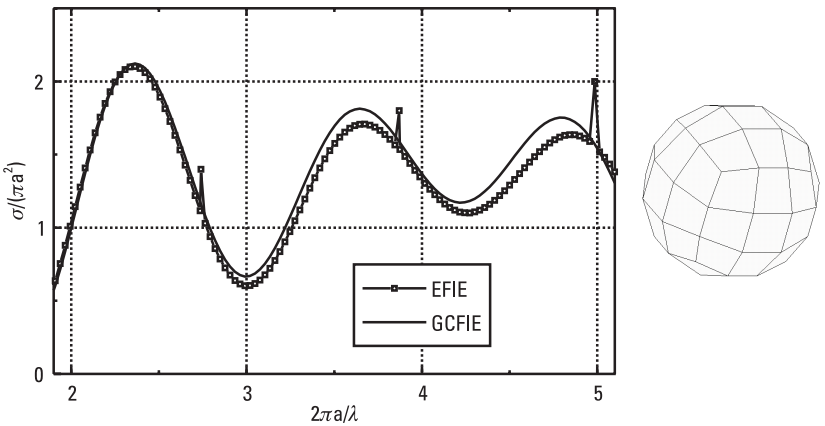

Figure 4.7 Monostatic RCS of a spherical scatterer (a = 1.41m), normalized by πa^2, versus $2\pi a/\lambda$. The geometrical model of the sphere consists of m = 54 patches as shown in the inset. (After: [21].)

components of vectors **F** and **G** can be decomposed in the local *ps*-coordinates according to (4.14). Having in mind (4.21), (4.30) is decomposed into its *p*- and *s*-components as

$$\alpha F_p - \frac{1}{|\mathbf{i}_p \times \mathbf{i}_s|}[(\mathbf{i}_p \cdot \mathbf{i}_s) G_p + G_s] = 0 \tag{4.31a}$$

$$\alpha F_s + \frac{1}{|\mathbf{i}_p \times \mathbf{i}_s|}[G_p + (\mathbf{i}_p \cdot \mathbf{i}_s) G_s] = 0 \tag{4.31b}$$

That is, the *p*- and *s*-components of **F** are expressed as specific linear combinations of the *p*- and *s*-components of **G**. Suppose that some other linear combinations of these components can be used instead of the classic CFIE; that is, that the *p*- and *s*-components of **F** can be written as

$$F_p = \alpha_{pp} G_p + \alpha_{ps} G_s \qquad F_s = \alpha_{sp} G_p + \alpha_{ss} G_s \tag{4.32a,b}$$

where α_{pp}, α_{ps}, α_{sp}, and α_{ss} are complex coefficients. Let us determine these coefficients so that (4.32) implies that the components of **F** and **G**, tangential to the boundary surface, are zero. (In that case, the spurious resonances are eliminated from the solution.) We follow a similar procedure as in Section 4.1.6.

There are no sources inside region 0, and the real part of the complex power (i.e., the average power) flowing into the region should be zero or greater, as expressed by (4.27a). The average power is expressed in terms of **F** and **G** as

$$\mathrm{Re}\{P\} = -\oint_{S_{i0}} \mathrm{Re}\{(\mathbf{F} \times \mathbf{G}^*) \cdot \mathbf{n}_{i0}\}\, dS \tag{4.33}$$

F and **G** can be decomposed in the local *ps*-coordinate system according to (4.14). Using (4.21), $\mathbf{n}_{i0}$ is expressed in terms of $\mathbf{i}_p$ and $\mathbf{i}_s$. The average power flowing into region 0 is obtained as

$$\mathrm{Re}\{P\} = -\oint_{S_{i0}} \mathrm{Re}\{F_p G_s^* - F_s G_p^*\} |\mathbf{i}_p \times \mathbf{i}_s|\, dS \tag{4.34}$$

Suppose, for the moment, that all coefficients α_{pp}, α_{ps}, α_{sp}, and α_{ss} are real. According to (4.32), the first term under the integral on the right side of (4.34) is written as

$$\mathrm{Re}\{F_p G_s^* - F_s G_p^*\} = \alpha_{ps}|G_s|^2 - \alpha_{sp}|G_p|^2 + (\alpha_{pp} - \alpha_{ss})\,\mathrm{Re}\{G_p G_s^*\} \tag{4.35}$$

The coefficients α_{pp}, α_{ps}, α_{sp}, and α_{ss} should be adopted so that this term will always be nonnegative. The signs of the first and second terms on the right side of (4.35) depend only on the signs of the coefficients α_{ps} and α_{sp}, respectively, while the sign of the third term cannot be forced by choosing the coefficients α_{pp} and α_{ss}. To make the sum of the first two terms greater than zero, the coefficient α_{ps} should also be greater than zero, but the coefficient α_{sp} should be less than zero. Using the identity

$$|\sqrt{\alpha_{ps}}\,G_s \pm \sqrt{-\alpha_{sp}}\,G_p|^2 = \alpha_{ps}|G_s|^2 - \alpha_{sp}|G_p|^2 \pm 2\sqrt{-\alpha_{ps}\alpha_{sp}}\,\mathrm{Re}\{G_p G_s^*\} \tag{4.36}$$

the first two terms on the right side of (4.36) are related to the third term through the inequality

$$\alpha_{ps}|G_s|^2 - \alpha_{sp}|G_p|^2 \geq 2\sqrt{-\alpha_{ps}\alpha_{sp}}\,|\mathrm{Re}\{G_p G_s^*\}| \tag{4.37}$$

Finally, it is concluded that right side of (4.35) is nonnegative if the coefficients α_{pp}, α_{ps}, α_{sp}, and α_{ss} satisfy the conditions

$$\alpha_{ps} > 0 \qquad \alpha_{sp} < 0 \qquad \sqrt{-2\alpha_{ps}\alpha_{sp}} \geq |\alpha_{pp} - \alpha_{ss}| \tag{4.38a-c}$$

Consequently, the average power flowing into region 0 is forced to be nonpositive. However, according to (4.27a), this power can be only nonnegative. By imposing (4.32) with the coefficients α_{pp}, α_{ps}, α_{sp}, and α_{ss} satisfying the above conditions, the average power is forced to be zero. The tangential components of both **F** and **G** are also forced to be zero. Equation (4.32a,b) then represents the generalized scalar formulation of the CFIE.

If the coefficients α_{pp}, α_{ps}, α_{sp}, and α_{ss} are complex, the above conditions stay valid for real parts of these coefficients. It can be shown that imaginary parts of α_{ps} and α_{sp} can be arbitrary, while the sum of the imaginary parts of α_{pp} and α_{ss} should be zero. These conditions represent a generalization of conditions given in [21]. It can be shown that same conditions stand if the p- and s-projections of **F** are expressed in terms of the p- and s-projections of **G**.

There are various choices of the coefficients α_{pp}, α_{ps}, α_{sp}, and α_{ss}, starting from the conditions given by (4.38). The simplest form of (4.32) is obtained by adopting $\alpha_{pp} = \alpha_{ss} = 0$ and $\alpha_{ps} = -\alpha_{sp} = 1$. However, the optimal

choice depends on the method used for the solution of these equations (see Chapter 8).

4.1.8 CSIE

Consider the scattered field inside region i due to currents induced on the boundary surface between regions i and 0 [Figure 4.1(a)]. The scattered field will not change if the original sources are replaced by an arbitrary combination of equivalent electric and magnetic currents distributed over the same surface [see Figure 4.1(b)] such that the boundary condition for the tangential component of the electric field (on the face in region i) is maintained. [In general, this condition is given by (4.9a).] Various combinations of equivalent electric and magnetic currents can fulfill this condition, resulting in different fields inside region 0.

To eliminate a spurious resonant field that can occur inside region 0 at a resonant frequency, these current distributions should be related by

$$\mathbf{M}_{si0} = \alpha \mathbf{n}_{i0} \times \mathbf{J}_{si0} \tag{4.39}$$

where α is an arbitrary complex constant whose real part is positive. Suppose the opposite—that there are two sets of electric and magnetic currents, $\mathbf{J}_{si0}^{(1)}$, $\mathbf{M}_{si0}^{(1)}$, and $\mathbf{J}_{si0}^{(2)}$, $\mathbf{M}_{si0}^{(2)}$, which satisfy conditions (4.9a) and (4.39). The corresponding total electric and magnetic fields in region i are identical, that is, $\delta\mathbf{E}^{(i)} = \mathbf{E}_2^{(i)} - \mathbf{E}_1^{(i)} = 0$ and $\delta\mathbf{H}^{(i)} = \mathbf{H}_2^{(i)} - \mathbf{H}_1^{(i)} = 0$, while in region 0 these fields differ. The boundary conditions (4.1a,b) can be written for the total difference fields in the form

$$\mathbf{n}_{i0} \times \delta\mathbf{E}^{(0)} = \delta\mathbf{M}_{si0} \qquad \mathbf{n}_{i0} \times \delta\mathbf{H}^{(0)} = -\delta\mathbf{J}_{si0} \tag{4.40a,b}$$

where $\delta\mathbf{E}^{(0)} = \mathbf{E}_2^{(0)} - \mathbf{E}_1^{(0)} = 0$ and $\delta\mathbf{H}^{(0)} = \mathbf{H}_2^{(0)} - \mathbf{H}_1^{(0)} = 0$ are the total difference fields in region 0, and $\delta\mathbf{J}_{si0} = \mathbf{J}_{si0}^{(2)} - \mathbf{J}_{si0}^{(1)}$ and $\delta\mathbf{M}_{si0} = \mathbf{M}_{si0}^{(2)} - \mathbf{M}_{si0}^{(1)}$ are difference currents. Obviously, the difference currents also obey (4.39).

The complex power flowing into region 0 due to the difference fields can be written as

$$P = -\oint_{S_{i0}} (\delta\mathbf{E}^{(0)} \times \delta\mathbf{H}^{(0)}) \cdot \mathbf{n}_{i0}\, dS = -\oint_{S_{i0}} (\mathbf{n}_{i0} \times \delta\mathbf{E}^{(0)}) \cdot (\delta\mathbf{H}^{(0)})^*\, dS \tag{4.41}$$

After replacing (4.39) and (4.40) into (4.41), we obtain

$$P = -\oint_{S_{i0}} \delta\mathbf{M}_{si0} \cdot [\mathbf{n}_{i0} \times (\delta\mathbf{J}_{si0})^*]\, dS = -\frac{1}{\alpha^*} \oint_{S_{i0}} |\delta\mathbf{M}_{si0}|^2\, dS \quad (4.42)$$

The real part of this complex power is always nonpositive. However, according to (4.27a), this real part can be only nonnegative. Hence, by imposing condition (4.39), this real part is forced to be zero. As a result, the difference currents are also forced to be zero, which means that any spurious resonant field is suppressed.

The integral equation is derived in the following way. Equation (4.39) is replaced into (4.5a), the resulting (4.5a) and (4.4a) are replaced into (4.2a), and the resulting (4.2a) is replaced into (4.9a). The equation obtained can be simplified using (4.8a). Finally, the cross product of the equation and the unit vector normal gives

$$[Z^{(i)} L_0^{(i)}(\mathbf{J}_{si0}) - \alpha K_0^{(i)}(\mathbf{n}_{i0} \times \mathbf{J}_{si0})]_{\tan} + \frac{\alpha}{2}\mathbf{J}_{si0} = (\mathbf{E}_{\text{inc}}^{(0)})_{\tan} \quad (4.43)$$

Based on a combination of electric and magnetic currents, this equation is referred to as the CSIE. This technique for eliminating spurious resonances was theoretically established for 3-D cases by Brakhage and Werner [22]. However, it was first applied to 2-D cases by Greenspan and Werner [23] and Bolomey and Tabbara [24]. Thereafter, it was applied to 3-D cases by Mautz and Harrington [25].

A particular problem in applying (4.43) is the choice of α. The role of this constant is to make the electric and magnetic currents enter into the equation with approximately the same weights. Obviously, for a plane wave (which is created by a combined electric and magnetic current sheet) this constant should be equal to the intrinsic impedance of the medium (i.e., $\alpha = Z^{(i)}$). The same value is generally recommended by Mautz and Harrington [25].

Compared with the EFIE, the implementation of the CSIE is equally complicated as for the CFIE. Namely, the CFIE requires an additional evaluation of the magnetic field due to the electric currents, while the CSIE requires an additional evaluation of the electric field due to the magnetic currents. The shortcoming of the CSIE is that it does not provide a correct solution for the current distribution over the boundary surface. The combined sources obtained as the solution have no physical meaning. However, the actual currents can be evaluated indirectly in the following way. First, the near field is evaluated on the face of the boundary surface in region i. Then, having in mind that the field due to the actual currents is zero in region 0, the actual currents are easily

obtained using the boundary conditions. Another disadvantage of the method is that it does not provide a zero field inside region 0. This field is parasitic and it is often used as a measure of the quality of the numerical solution. Hence, the CFIE is advantageous when compared with the CSIE.

4.1.9 Thin-Plate Approximation

Consider a PEC body in the form of a thin plate [Figure 4.8(a)]. When the thickness of the body is taken into account, the analysis can be performed in the same way as for any solid PEC body: The integral equation is imposed for each face of the plate. When the plate thickness tends to zero, the two opposite faces of the plate merge into one open surface as shown in Figure 4.8(b). Such a body is sometimes called an open body, while the solid PEC body is called a closed body. Hence, integral equations derived for closed bodies may not be valid any more. In this section, we shall examine the validity of the EFIE and MFIE for open bodies.

We start by considering the plate to be very thin, but still to have a finite thickness. Let us denote the current densities on the opposite faces of the thin plate by $\mathbf{J}'_{si0}$ and $\mathbf{J}''_{si0}$, and the corresponding unit normals directed from region 0 to region i by $\mathbf{n}'_{i0}$ and $\mathbf{n}''_{i0}$. The electric field inside the PEC body (i.e., inside region 0), $\mathbf{E}^{(0)}(\mathbf{J}_{si0})$, can be expressed as

$$\mathbf{E}^{(0)}(\mathbf{J}_{si0}) = \mathbf{E}^{(0)}(\mathbf{J}'_{si0}) + \mathbf{E}^{(0)}(\mathbf{J}''_{si0}) \tag{4.44}$$

The vectors $\mathbf{E}^{(0)}(\mathbf{J}'_{si0})$ and $\mathbf{E}^{(0)}(\mathbf{J}''_{si0})$ can be further expressed according to (4.4b) as

$$\mathbf{E}^{(0)}(\mathbf{J}'_{si0}) = -Z^{(i)} L_0^{(i)}(\mathbf{J}'_{si0}) - \frac{\rho'_{si0}}{2\epsilon^{(i)}}\,\mathbf{n}'_{i0} \tag{4.45a}$$

$$\mathbf{E}^{(0)}(\mathbf{J}''_{si0}) = -Z^{(i)} L_0^{(i)}(\mathbf{J}''_{si0}) - \frac{\rho''_{si0}}{2\epsilon^{(i)}}\,\mathbf{n}''_{i0} \tag{4.45b}$$

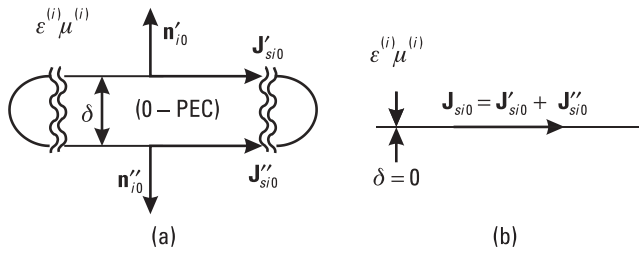

Figure 4.8 Thin-plate approximation: (a) original case and (b) infinitesimally thin plate.

On the other hand, the boundary condition (4.9b) can be written for both faces of the plate as

$$\mathbf{n}'_{i0} \times \mathbf{E}^{(0)} = 0 \qquad \mathbf{n}''_{i0} \times \mathbf{E}^{(0)} = 0 \tag{4.46a,b}$$

Both forms lead to the same integral equation. For example, after replacing (4.45) into (4.44), the resulting (4.44) into (4.2b), and the resulting (4.2b) into (4.46a), we obtain

$$Z^{(i)} \mathbf{n}'_{i0} \times L_0^{(i)}(\mathbf{J}'_{si0} + \mathbf{J}''_{si0}) = \mathbf{n}'_{i0} \times \mathbf{E}^{(0)}_{\text{inc}} \tag{4.47}$$

This equation is the EFIE for finite-thickness PEC plates.

We now assume the plate thickness to tend to zero. In this case, $\mathbf{J}'_{si0}$ and $\mathbf{J}''_{si0}$ in (4.47) merge into a unique current sheet of density $\mathbf{J}_{si0} = \mathbf{J}'_{si0} + \mathbf{J}''_{si0}$. Equation (4.47) now obtains the same form as (4.10b), which is valid for closed bodies. The solution of (4.10b), however, can give only the sum of currents existing on the opposite faces of the plate. This solution for the currents gives a complete solution for the fields in region i.

We repeat the same procedure for the MFIE. Assuming the plate to have a finite thickness, the magnetic field inside the body, $\mathbf{H}^{(0)}(\mathbf{J}_{si0})$, can be expressed as

$$\mathbf{H}^{(0)}(\mathbf{J}_{si0}) = \mathbf{H}^{(0)}(\mathbf{J}'_{si0}) + \mathbf{H}^{(0)}(\mathbf{J}''_{si0}) \tag{4.48}$$

The vectors $\mathbf{H}^{(0)}(\mathbf{J}'_{si0})$ and $\mathbf{H}^{(0)}(\mathbf{J}''_{si0})$ can be further expressed according to (4.6b) as

$$\mathbf{H}^{(0)}(\mathbf{J}'_{si0}) = -K_0^{(i)}(\mathbf{J}'_{si0}) - \frac{1}{2}\mathbf{J}'_{si0} \times \mathbf{n}'_{i0} \tag{4.49a}$$

$$\mathbf{H}^{(0)}(\mathbf{J}''_{si0}) = -K_0^{(i)}(\mathbf{J}''_{si0}) - \frac{1}{2}\mathbf{J}''_{si0} \times \mathbf{n}''_{i0} \tag{4.49b}$$

On the other hand, the boundary condition (4.11b) can be written for both faces of the plate as

$$\mathbf{n}'_{i0} \times \mathbf{H}^{(0)} = 0 \qquad \mathbf{n}''_{i0} \times \mathbf{H}^{(0)} = 0 \tag{4.50a,b}$$

Both forms lead to the same integral equation. For example, replacing (4.49) into (4.48), the resulting (4.48) into (4.3b), and the resulting (4.3b) into (4.50a), the MFIE for finite-thickness PEC plates is obtained in the form

$$\mathbf{n}'_{i0} \times K_0^{(i)}(\mathbf{J}'_{si0} + \mathbf{J}''_{si0}) + \frac{1}{2}(\mathbf{J}'_{si0} - \mathbf{J}''_{si0}) = \mathbf{n}'_{i0} \times \mathbf{H}_{\text{inc}}^{(0)} \tag{4.51}$$

This MFIE depends on both the sum and difference of the currents existing on opposite faces of the plate. Hence, if we assume the plate thickness to tend to zero, and these two current sheets merge into a single one, we cannot formulate the MFIE in terms of the resulting current density ($\mathbf{J}_{si0}$). Consequently, we cannot use the MFIE (4.12) for field solution in region i.

Once the EFIE has been solved for the sum of currents, the MFIE (4.51) (assuming the plate thickness tends to zero) can be used to obtain the difference of the currents. From these two solutions, the currents on each face of the plate are easily determined. However, the currents can be evaluated in a simpler way. The EFIE provides a complete field solution inside region i, including the component of the magnetic field tangential to the plate surface. By applying the boundary condition for the magnetic field, taking into account that the magnetic field inside the plate is equal to zero, the current distributions on both faces of the plate are directly determined.

Another method was suggested by Newman and Schrote [26]. Starting from the reaction concept of Rumsey [27], they derived two equations, one dependent on the sum of currents and another dependent on the difference of the currents. Both equations are solved using the MoM. However, such a procedure is more complicated than the previous one, based on the calculation of the near magnetic field.

For example, let as consider a square-plate scatterer of side $a = 1\lambda$ and thickness δ excited by a plane wave of intensity $E = 1$ V/m [Figure 4.9(a)]. The scatterer can be modeled as is (a closed body with six flat surfaces) or as an infinitesimally thin plate. In the first case, the currents on both faces are independently modeled, while in the second case only the sum of the currents is modeled. The model of the infinitesimally thin plate consists of five patches to precisely take into account the edge effect [Figure 4.9(b)]. The first-order current approximation is used along short edges and the eighth-order along long edges, resulting in $N = 320$ unknowns for the first model and $N = 176$ for the second model. The unknowns are determined by the Galerkin method applied to the EFIE. For the second case, once the sum of currents is determined, the currents on the opposite faces of the scatterer are evaluated from the near magnetic field.

Figure 4.9(c) shows the magnitude of the z-component of the current at both faces of the scatterer versus the y-coordinate for two plate thicknesses: $\delta = 0$ (infinitesimally thin plate) and $\delta = 0.1\lambda$ (closed-body model). The current distributions in the two cases are almost identical. Figure 4.9(d) shows the bistatic RCS normalized by λ^2 versus the angle θ in the xOz-plane, for three

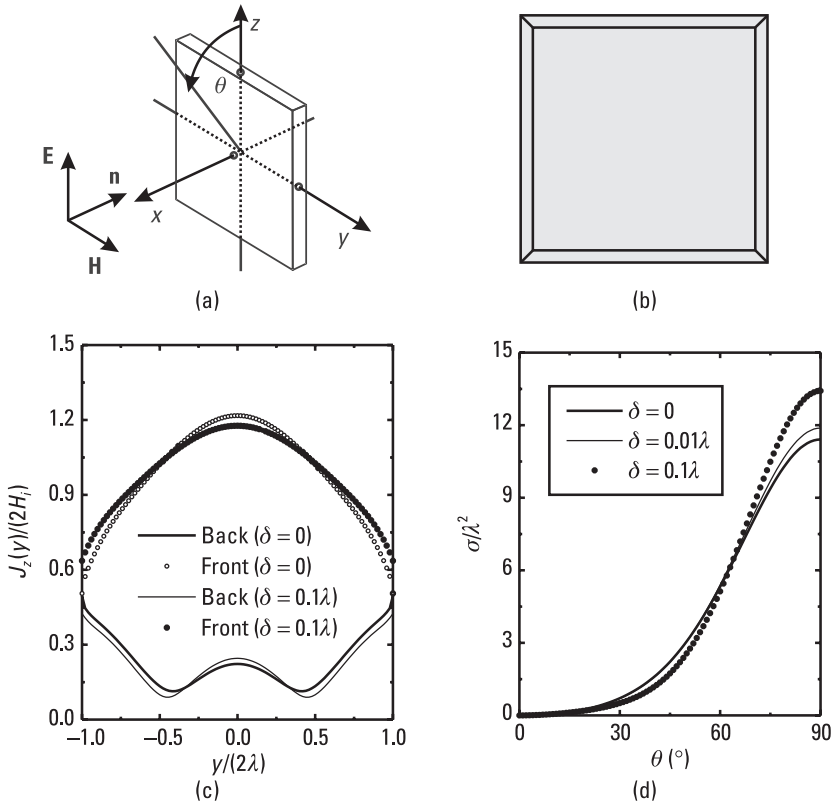

Figure 4.9 Thin-plate scatterer: (a) original structure, (b) approximate model, (c) magnitude of the *z*-component of the current at scatterer faces, and (d) bistatic RCS versus the angle θ in the *xOz*-plane.

different plate thicknesses: $\delta = 0$, $\delta = 0.01\lambda$, and $\delta = 0.1\lambda$. By increasing the plate thickness, the bistatic RCS increases. When the thickness is less than 0.01λ, the maximal values of the bistatic RCS for the closed body and the infinitesimally thin plate differ less than 5%. Even for very thick plates (e.g., $\delta = 0.1\lambda$), results can be acceptable. Particularly for the far-field evaluation, it is not necessary to take the edge effect precisely into account. The same results for scattering are obtained if the infinitesimally thin scatterer is modeled by using only one patch, resulting in only $N = 112$ unknowns. The thin-plate approximation can be widely used for modeling of real plates in order to simplify the analysis.

4.1.10 Thin-Wire Approximation

Wires are metallic bodies in the form of curvilinear cylinders of circular cross sections (of variable radii). [An arbitrary wire in the vicinity of a PEC body is

shown in Figure 4.10(a).] A wire is assumed to be thin if its radius is simultaneously much smaller than the wavelength, the length of its curvilinear axis, and the curvature radius of the axis. For a thin wire having cross section of a constant radius, the circumferential component of the surface current can usually be neglected when compared with the axial component, whereas the axial component practically does not depend on the circumferential coordinate. These assumptions are referred to as the thin-wire approximation. When the thin-wire approximation is applied to a thin wire having a cross section of variable radius, the surface current density vector is directed along wire generatrices, and the circumferential component is still neglected. (Generatrices of a wire are at each point perpendicular to the wire circumferences.)

The thin-wire approximation enables a simplification of integral equations used for their numerical evaluation. The current along a cylindrical wire produces only the axial component of the electric field. Hence, this current can be determined from the condition that the axial component of the total electric field on the wire surface is zero. To define an analogous condition for a generally shaped wire, we introduce a local s-coordinate line along an arbitrary wire generatrix [Figure 4.10(b)]. Then the condition (4.9b) is written in the form

$$\mathbf{i}_s \cdot \mathbf{E}^{(0)} = 0 \tag{4.52}$$

where $\mathbf{i}_s$ is the unit vector along the local s-coordinate line, and $\mathbf{E}^{(0)}$ is the total electric field just inside the wire. (According to the notation introduced in the previous sections, the wire represents region 0, surrounded by region i.) After expressing $\mathbf{E}^{(0)}$ in terms of surface currents, an integral equation is obtained that represents a particular form of the general EFIE, given by (4.9b). (While the general EFIE is decomposed into two scalar components, each dependent

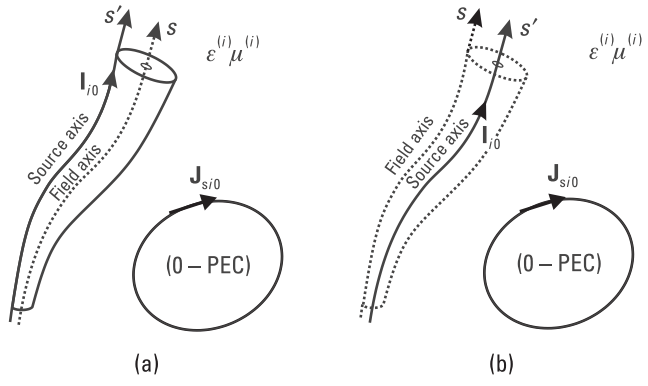

Figure 4.10 Thin-wire approximation based on (a) EBC and (b) equivalent sources.

on two current components, this particular form is a single scalar equation dependent on one current component.) However, solving this particular form is not a much easier task than solving the general EFIE. Namely, in both cases, the most difficult problem is the evaluation of dual integrals contained in the expressions for the electric field (Section 3.3.2).

The electric field in the vicinity of the wire and on the surface of the wire will suffer acceptably small changes if the current distributed over the wire surface is replaced by an equivalent current along the wire axis, denoted by s' in Figure 4.10(b). However, the expressions for the electric field vector will contain only single integrals. In the original double integrals, the first integration is performed around the circumference. This circumferential integral represents the so-called exact kernel for the second integration. In the single integrals, the integration around the circumference is omitted, and the kernel of the integral is "reduced" when compared with the exact kernel. The thin-wire approximation combined with the reduced kernel enables very efficient analysis of wire structures.

The reduced kernel can be alternatively introduced using the EBC instead of the concept of equivalent sources. The EBC for thin wires is usually written in the form that the axial component of the total electric field is zero along the wire axis. If the local s-coordinate line coincides with the wire axis [Figure 4.10(a)], this boundary condition has the same form as (4.52). The axial component of the electric field along the wire axis produced by the actual sources at the wire surface will not practically change if the current flowing along the wire surface is replaced by a line (filamental) current of the same intensity, flowing along an arbitrary wire generatrix. [We introduce a local s'-coordinate line along that generatrix; see Figure 4.10(a).] As a result, expressions for the electric field will contain only single integrals with the reduced kernel.

For a single (straight) cylindrical dipole, both approaches that lead to the reduced kernel give identical solutions. For more complex structures, the results differ slightly, but the differences can usually be neglected. Both approaches lead to formally the same integral equation. The vector $\mathbf{E}^{(0)}$ along the s-axis can be written as

$$\mathbf{E}^{(0)} = \mathbf{E}^{(0)}(I_{i0}) + \mathbf{E}^{(0)}(\mathbf{J}_{si0}) + \mathbf{E}^{(0)}_{\text{inc}} \tag{4.53}$$

where $\mathbf{E}^{(0)}(I_{i0})$ is the electric field due to the filamental current I_{i0} flowing along the s'-axis, $\mathbf{E}^{(0)}(\mathbf{J}_{si0})$ is the electric field due to surface currents of other PEC bodies surrounding the wire considered, and $\mathbf{E}^{(0)}_{\text{inc}}$ is the incident electric field. According to (3.43a), $\mathbf{E}^{(0)}(I_{i0})$ and $\mathbf{E}^{(0)}(\mathbf{J}_{si0})$ can be written in linear operator form as

$$\mathbf{E}^{(0)}(I_{i0}) = -Z^{(i)}L(I_{i0}) \qquad \mathbf{E}^{(0)}(\mathbf{J}_{si0}) = -Z^{(i)}L(\mathbf{J}_{si0}) \quad (4.54a,b)$$

After substituting (4.54) into (4.53), and the resulting (4.53) into (4.52), the EFIE is obtained in the form

$$Z^{(i)}\mathbf{i}_s \cdot [L(I_{i0}) + L(\mathbf{J}_{si0})] = \mathbf{i}_s \cdot \mathbf{E}^{(0)}_{\text{inc}} \quad (4.55)$$

The currents $\mathbf{J}_{si0}$ are assumed to be relatively far away from the wire, so that the main problem in imposing (4.55) is the evaluation of $L(I_{i0})$ along the s-axis. We can use any of the integral forms of operator L listed in Table 3.1. These expressions differ in the singularity order (i) as explained in Section 3.3.2. As a result, three specific EFIEs are obtained: the two-potential equation (for $i = 2$), the Pocklington equation (for $i = 3$), and the Schelkunoff equation (for $i = 1$). The original formulations of these equations were derived in [2] and [28] for an analytical solution of the current distribution along cylindrical wire dipoles. These three equations differ only in the numerical difficulties encountered with the evaluation of the quasisingular integrals. If they are solved numerically using the same method, they will give practically identical solutions.

For straight parallel wires, the Schelkunoff equation has the lowest singularity order. This is the easiest case for implementation, in particular with the point-matching testing procedure, as shown by Dragovic [29]. However, this advantage of the Schelkunoff equation does not exist for an arbitrary wire structure. If the two-potential equation is combined with any test procedure except the point-matching, the singularity order of integrals can be reduced (Chapter 8). Hence, the two-potential EFIE is recommended.

The thin-wire approximation is efficient even when the actual surface current has a circumferential dependence. For example, consider an electrically short wire scatterer of length $l = 0.01\lambda$ and radius $a = l/10$, excited by an axially polarized plane wave of intensity $E = 1$ V/m. First, we analyze the scatterer as a PEC metallic body. The cylindrical part of the scatterer surface is modeled by 16 plates, and each base is modeled by 16 plates [Figure 4.11(a)]. The current flowing over a patch has two components dependent on two local coordinates. The first order of the current approximation is used for short edges, and the fourth order for long edges, resulting in $N = 192$ unknowns. The unknown coefficients are determined using the Galerkin method applied to the EFIE. Figure 4.11(b) shows the real and imaginary parts of the axial current component along the scatterer circumference. The real part depends strongly on the circumferential coordinate; the imaginary part is practically constant.

Next, we apply the thin-wire approximation to the same scatterer. The assumed current now has only one component dependent on one coordinate.

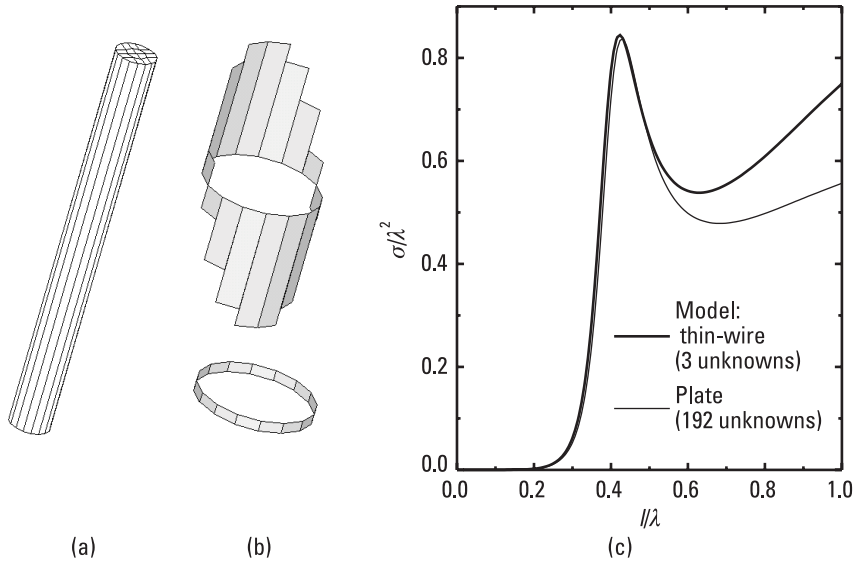

Figure 4.11 Wire scatterer of radius $a = l/10$ excited by an axially polarized plane wave: (a) plate model, (b) real part (up) and imaginary part (down) of the axial current component along the middle circumference of the scatterer of length $l = 0.01\lambda$, and (c) monostatic RCS versus scatterer electric length.

The fourth order of approximation for this current is used, resulting in $N = 3$ unknowns. Figure 4.11(c) shows the monostatic RCS obtained from the two models versus the electric length of the scatterer [12]. The difference between these results is negligible for wire lengths up to half a wavelength long. The main reason for an increased discrepancy for longer wires is the large electrical size of the wire diameter, which becomes greater than 0.1λ. In conclusion, the thin-wire approximation can be applied even when the current is strongly dependent on the circumferential coordinate, provided the wire radius is small compared with the wavelength.

The MFIE cannot be formulated either for equivalent sources (see Section 4.1.3) or for thin-wire analyses based on EBC (along the wire axis). The later formulation is not possible because an arbitrary current that has only the axial component along a cylindrical wire, depending only on the axial coordinate, automatically creates a zero magnetic field along the wire axis. Besides the EFIE, which is predominantly used in thin-wire analysis, an alternative choice is the Hallén equation.

4.1.11 Hallén Equation

The vector $\mathbf{E}^{(0)}(I_{i0})$, given by (4.54a), can be expressed according to (3.34) in terms of the corresponding magnetic vector potential, $\mathbf{A}_m^{(0)}(I_{i0})$, as

$$\mathbf{E}^{(0)}(I_{i0}) = -j\omega\left\{\mathbf{A}_m^{(0)}(I_{i0}) - \frac{1}{(\gamma^{(i)})^2}\nabla[\nabla \cdot \mathbf{A}_m^{(0)}(I_{i0})]\right\} \tag{4.56}$$

Substituting (4.56) into (4.53), and the resulting (4.53) into (4.52), after some manipulation, we obtain

$$\frac{\partial^2 \mathbf{A}_{ms}^{(0)}(I_{i0})}{\partial s^2} - (\gamma^{(i)})^2 \mathbf{A}_{ms}^{(0)}(I_{i0}) = \frac{j}{\omega}(\gamma^{(i)})^2 \mathbf{i}_s \cdot (\mathbf{E}^{(0)}(\mathbf{J}_{si0}) + \mathbf{E}_{\text{inc}}^{(0)}) \tag{4.57}$$

where $\mathbf{A}_{ms}^{(0)}(I_{i0})$ is the s-component of $\mathbf{A}_m^{(0)}(I_{i0})$. This equation is actually a differential equation in terms of $\mathbf{A}_{ms}^{(0)}(I_{i0})$ versus the s-coordinate. The solution can be written as

$$\mathbf{A}_{ms}^{(0)}(I_{i0}) = C_1 \cos(\gamma^{(i)} s) + C_2 \sin(\gamma^{(i)} s) + \frac{j}{\omega}\gamma^{(i)} \int_0^s \mathbf{i}_s \cdot [\mathbf{E}^{(0)}(\mathbf{J}_{si0}) + \mathbf{E}_{\text{inc}}^{(0)}] \sin[\gamma^{(i)}(s-t)]\, dt \tag{4.58}$$

where C_1 and C_2 are unknown constants. Equation (4.58) is referred to as the Hallén equation. The original formulation was derived in 1937 by Hallén, who used it to obtain an analytical solution for cylindrical wire dipoles [3].

Hallén's method was extensively used by King, culminating in his classical monograph in 1956 [30]. After the paper of Mei in 1965 [31], the Hallén equation was predominantly solved numerically by the MoM. High-speed digital computers enabled solutions of complex wire structures. The problem of multiple junctions was solved by Taylor in 1969 [32]. Thereafter, the main problem in the implementation of the Hallén equation was evaluation of the integral on the right side of (4.58). Finally, in 1976, Kominami and Rokushima showed that the order of the integration can be reduced by one [33], which enabled a very efficient analysis of arbitrary thin-wire structures [34].

In contrast to the EFIE, the Hallén equation is almost insensitive to the choice of the testing procedure [35]. For example, consider a symmetrical dipole antenna of arm length h and radius a excited by a TEM magnetic current frill of radii a and b (Figure 4.12, inset). The analysis is performed using the following testing procedures: the point-matching, the Galerkin method, and the least-squares method. Figure 4.12 shows the relative error (with respect to a high-order numerical solution) in the input admittance versus the order of the polynomial approximation of current along one dipole arm. All test proce-

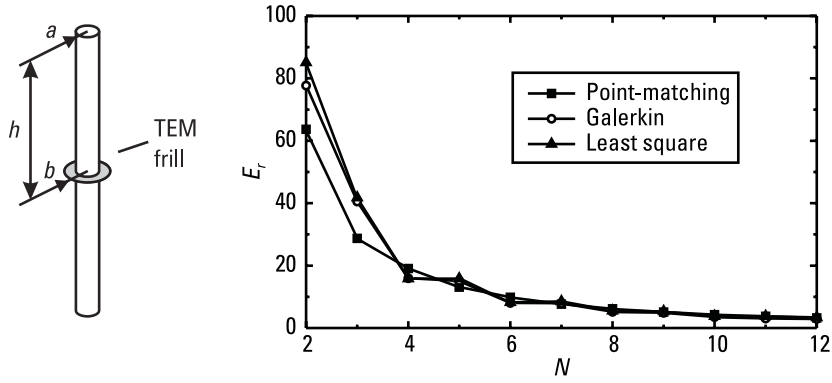

Figure 4.12 Relative error in the admittance of a halfwave dipole shown in the inset (h/a = 100, b/a = 2.3) versus the order of approximation of current along one dipole arm.

dures yield almost the same accuracy. Hence, the point-matching method, being the simplest test procedure, is almost always applied.

The main shortcoming of the Hallén equation is that it cannot easily be generalized to the analysis of plate structures and combined metallic and dielectric structures. It was shown that the Hallén equation solved by the point-matching method is equivalent to the EFIE tested by sine dipoles [36]. Also, with a proper choice of testing functions, the EFIE can be at least as efficient as the Hallén equation [37]. Hence, the EFIE is recommended, at least in cases when wires are combined with plates and dielectrics.

4.1.12 IBC-IE

Losses in a metallic body due to the skin effect can be taken into account if its boundary is considered to be an ideal impedance boundary surface (see Section 3.6). In that case, the metallic body is treated in the same way as a PEC body shown in Figure 4.1, except that there are, in addition, magnetic currents over the boundary surface, which are related to electric currents through the IBC (3.66).

As for PEC bodies, the induced electric and magnetic currents can be determined from the basic field integral equations, the EFIE and the MFIE. Since the resulting equations include the IBC, they are denoted by IBC-EFIE and IBC-MFIE [38]. However, the basic field integral equations fail at interior resonant frequencies. To eliminate the spurious resonances, these two equations can be combined into an IBC-CFIE [39, 40]. As in the case of the CFIE for PEC bodies, the formulation of the IBC-CFIE requires the introduction of an arbitrary weighting constant.

An integral equation alternative to the IBC-CFIE can be derived, keeping in mind that the field in region i is uniquely determined by the IBC (for fields) (Section 3.10.1). This condition can be written in two forms: (3.64a,b) (Section 3.6). Here, we repeat for convenience the IBC for fields (3.64a) and the IBC for currents (3.66a)

$$(\mathbf{E}^{(i)})_{\tan} = Z^{(0)}\mathbf{n}_{i0} \times \mathbf{H}^{(i)} \qquad \mathbf{M}_{si0} = -Z^{(0)}\mathbf{n}_{i0} \times \mathbf{J}_{si0} \quad (4.59a,b)$$

where $Z^{(0)}$ is the intrinsic impedance of the metallic body at its boundary. The integral equation is derived in the following way. First, (4.59b) is replaced into (4.5a) and (4.7a). Then (4.4a) and (4.5a) are replaced into (4.2a), and (4.6a) and (4.7a) are replaced into (4.3a). Finally, (4.2a) and (4.3a) are replaced into (4.59a) resulting in

$$\begin{aligned} &\left[\frac{Z^{(i)}}{Z^{(0)}} L_0^{(i)}(\mathbf{J}_{si0}) + K_0^{(i)}(\mathbf{n}_{i0} \times \mathbf{J}_{si0})\right]_{\tan} \\ &+ \mathbf{n}_{i0} \times \left[\frac{Z^{(0)}}{Z^{(i)}} L_0^{(i)}(\mathbf{n}_{i0} \times \mathbf{J}_{si0}) - K_0^{(i)}(\mathbf{J}_{si0})\right] \\ &= \frac{1}{Z^{(0)}}(\mathbf{E}_{\text{inc}}^{(i)})_{\tan} - \mathbf{n}_{i0} \times \mathbf{H}_{\text{inc}}^{(i)} \end{aligned} \quad (4.60)$$

This integral equation is known as the IBC-IE [41].

4.1.13 Optimal Choice of BIEs for Analysis of Metallic Structures

Composite metallic structures consist of wires, thin plates (open structures), and solid metallic bodies (closed structures). When losses due to the skin effect can be neglected, there are six general types of BIEs that can be used for analysis: EFIE, MFIE, AEFIE, AMFIE, CFIE, and CSIE.

An efficient analysis of wires and thin plates is based on the thin-wire and thin-plate approximations in conjunction with the EFIE. For closed bodies, all six equations can be used. If a closed body is analyzed by the EFIE or the MFIE, the solution contains a spurious field in the vicinity of each interior resonant frequency. This spurious field can be partly eliminated by increasing the quality of approximation, or completely eliminated by insertion of losses inside the body. In conjunction with the EFIE, the elimination can be done by solving the resulting system of linear equations in the minimum norm sense. Solutions of the AEFIE, AMFIE, CFIE, and CSIE do not suffer from spurious

fields. By making all meaningful combinations, we can define nine methods, which are summarized in Table 4.1.

All the methods from Table 4.1 differ regarding their accuracy, efficiency, and ease of implementation. In contrast to other methods, the CSIE does not yield physical solutions for currents. The EFIE and the MFIE combined with inserted losses require additional unknowns as compared with other methods. The AEFIE, the AMFIE, and the minimum norm solution of the EFIE result in an overdetermined system of linear equations, while a determined system is obtained in all other cases.

All integral equations from Table 4.1 are expressed in terms of two operators, L and K. If the same testing procedure is applied to all equations, the implementation is easier with equations based on a single operator (the EFIE, the EFIE solved in the minimum norm sense, and the MFIE). Note that the AEFIE and the AMFIE are also based on a single operator.

There is another property very important for developing general codes, which is often overlooked. Consider an arbitrarily shaped body, such as a metallic donut modeled by bilinear patches (Figure 4.13). For all BIEs except the EFIE, it is necessary to know the direction of the outward normal unit vector $\mathbf{n}$ on the body surface. The algorithm for automatic determination of this direction, which starts only from node coordinates of the patches, is relatively complicated and time consuming. An alternative is to request the user to properly order the four nodes for a patch, obeying the right-hand rule with respect to the outward normal. This forces the user to go to additional, very unpleasant effort in defining the structure. It is, hence, desirable that the direction of the unit normal vector does not influence the BIE. The EFIE is the only choice from this standpoint.

Table 4.1 summarizes all seven properties discussed above. In most cases, the simple EFIE is superior to other equations. If sophisticated basis functions are used for the current approximation, the range in the vicinity of a resonant frequency in which spurious fields occur becomes so narrow that it is very difficult to detect. If the spurious fields must positively be suppressed, the CFIE is probably the best candidate. If losses due to the skin effect cannot be neglected, one can apply the IBC-IE. However, the same problem can be solved more easily by using the concept of distributed loadings for metallic surfaces, as will be shown in Section 4.3.2.

4.2 BIEs for Combined Metallic and Dielectric Structures

Analysis of a combined metallic and dielectric structure can be defined as a multiple-region problem, which is described in Section 3.7 and shown in Figure

Table 4.1
Summary of Properties of BIEs Used for PEC Structures

Equation	Handles Thin Wires and Plates	Suppresses Spurious Fields	Includes Physical Currents	Without Additional Unknowns	Determined System of Equations	Uses Single Operator	Without Outward Normal
EFIE	+		+	+	+	+	+
EFIE+losses	+	+	+		+		+
EFIE+m.n.s.*	+	+	+	+		+	+
MFIE			+	+	+	+	
MFIE+losses		+	+		+		
AEFIE		+	+	+		+	
AMFIE		+	+	+		+	
CFIE		+	+	+	+		
CSIE		+		+	+		

* Minimum norm solution.

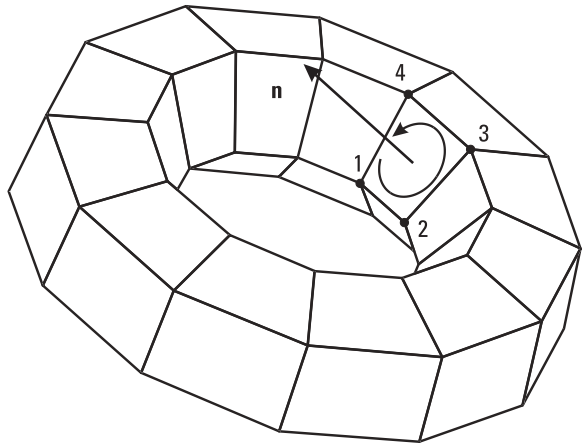

Figure 4.13 Determination of outward unit normal vector.

3.6. Consider an arbitrary region with a nonzero electromagnetic field, such as region i shown in Figure 4.14(a). According to the surface equivalence theorem (Section 3.9), the influence of all sources outside region i can be replaced by equivalent currents placed at the boundary surface of region i [Figure 4.14(b)], in which case the field outside region i becomes zero. Hence, the region outside region i is denoted as region $0 - i$ (i.e., region 0 with respect to region i in the equivalent problem for region i).

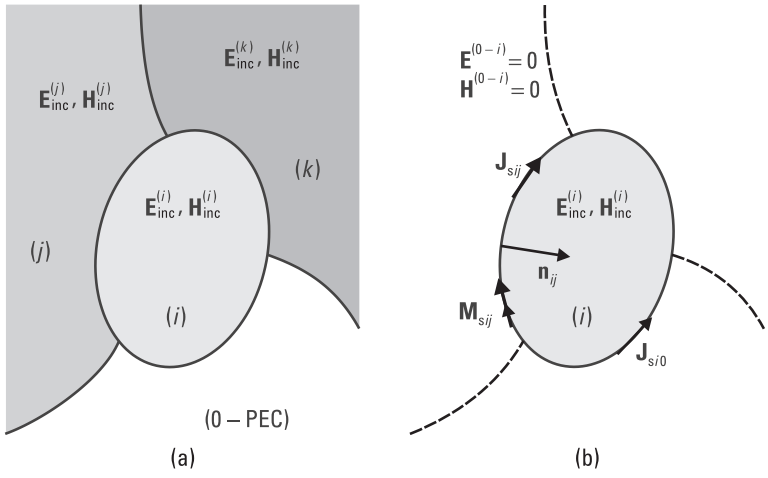

Figure 4.14 Decomposition of a multiple-region problem into single-region problems: (a) original problem and (b) equivalent problem for region i.

The densities of equivalent currents at the boundary surface between regions i and j are

$$\mathbf{J}_{sij} = \mathbf{n}_{ij} \times \mathbf{H}^{(i)} \qquad \mathbf{M}_{sij} = -\mathbf{n}_{ij} \times \mathbf{E}^{(i)} \tag{4.61a,b}$$

where $\mathbf{n}_{ij}$ is unit normal directed from region j to region i, and $\mathbf{E}^{(i)}$ and $\mathbf{H}^{(i)}$ are the electric and magnetic fields at the boundary surface just inside region i. If we consider the equivalent currents for region j, $\mathbf{J}_{sji}$ and $\mathbf{M}_{sji}$, they are evaluated according to (4.61), but with interchanged indices i and j; that is,

$$\mathbf{J}_{sji} = \mathbf{n}_{ji} \times \mathbf{H}^{(j)} \qquad \mathbf{M}_{sji} = -\mathbf{n}_{ji} \times \mathbf{E}^{(j)} \tag{4.62a,b}$$

$\mathbf{H}^{(i)}$ and $\mathbf{H}^{(j)}$ satisfy the boundary condition (3.61b), and $\mathbf{E}^{(i)}$ and $\mathbf{E}^{(j)}$ satisfy (3.61a). Since there are no electric and magnetic currents between regions i and j, these boundary conditions are written as

$$\mathbf{n}_{ij} \times (\mathbf{H}^{(i)} - \mathbf{H}^{(j)}) = 0 \qquad \mathbf{n}_{ij} \times (\mathbf{E}^{(i)} - \mathbf{E}^{(j)}) = 0 \tag{4.63a,b}$$

After expressing the field vectors in terms of the equivalent currents and using $\mathbf{n}_{ij} = -\mathbf{n}_{ji}$, the equivalent currents are related as

$$\mathbf{J}_{sij} = -\mathbf{J}_{sji} \qquad \mathbf{M}_{sij} = -\mathbf{M}_{sji} \tag{4.64a,b}$$

Region i can be partly bounded by PEC, PMC, and IBC surfaces. On a PEC surface there exist only electric currents, $\mathbf{J}_{si0}$, on a PMC surface there are only magnetic currents, $\mathbf{M}_{si0}$, and on an IBC surface there exist both types of equivalent currents, $\mathbf{J}_{si0}$ and $\mathbf{M}_{si0}$, related by the IBC.

Following this reasoning, a multiple-region problem, consisting of n nonzero regions, is decomposed into n single-region problems. Each single-region problem consists of solving the field inside that region surrounded by a corresponding zero-field region. All single-region problems are coupled through the equivalent currents placed on their common boundaries. These currents are mathematically related by (4.64), except for PEC, PMC, and IBC surfaces.

Such a decomposed multiple-region problem has a unique solution determined by the incident field (or impressed sources), which can exist in any region, provided the boundary conditions for the tangential electric and magnetic field components are fulfilled for each single-region problem. If these boundary conditions are imposed in region i, they have the form given by (4.61).

Let us prove that the solution is unique:

1. Equations (4.61a,b) combined with (4.64a,b) provide the boundary conditions (4.63a,b) to be satisfied; that is, the tangential components of the electric and magnetic fields are continuous across the source-free boundary surface between two non-zero-field regions.
2. For a PEC surface, $\mathbf{M}_{si0} = 0$ and (4.61b) provide that the tangential component of the electric field at the PEC boundary surface is zero.
3. For a PMC surface, $\mathbf{J}_{si0} = 0$ and (4.61a) provide that the tangential component of the magnetic field at the PMC boundary surface is zero.
4. For an IBC, $\mathbf{J}_{si0}$ and $\mathbf{M}_{si0}$ are related by the IBC for currents, so that either (4.61a) or (4.61b) provides the boundary conditions for the fields to be satisfied.
5. According to the theorem given in Section 3.10.2, if the above conditions are satisfied, the solution inside a multiple region is unique.

The boundary conditions (3.61a,b) for the boundary surface S_{ij} for the single-region problem shown in Figure 4.14(b) are

$$\mathbf{n}_{ij} \times (\mathbf{H}^{(i)} - \mathbf{H}^{(0-i)}) = \mathbf{J}_{sij} \qquad \mathbf{n}_{ij} \times (\mathbf{E}^{(i)} - \mathbf{E}^{(0-i)}) = -\mathbf{M}_{sij} \tag{4.65a,b}$$

where $\mathbf{E}^{(0-i)}$ and $\mathbf{H}^{(0-i)}$ are the total electric and magnetic fields, respectively, inside region $0 - i$. Combining these equations with (4.61), an alternative set of boundary conditions is obtained as

$$\mathbf{n}_{ij} \times \mathbf{H}^{(0-i)} = 0 \qquad \mathbf{n}_{ij} \times \mathbf{E}^{(0-i)} = 0 \tag{4.66a,b}$$

Satisfaction of both conditions (4.66a,b) is provided even if only one of these conditions is satisfied, except at interior resonances of a zero-field region. The first equation leads to the MFIE, while the second leads to the EFIE. It is interesting that these equations have not been widely used in the past on a stand-alone basis, although Kishk and Shafai showed their efficiency to be comparable with other frequently used integral equations [42].

At interior resonances, the satisfaction of both equations (4.66a,b) can be provided by using techniques that are usually applied for the elimination of spurious resonances in the analysis of scattering from a PEC body, resulting in various types of integral equations (e.g., AFIE and CFIE). Being optimal for scattering from the PEC body, only the CFIE for multiple-region problems

will be considered. This equation has not been used in the past, except very recently by Sheng et al. [43], who showed that this equation is convenient when the BIE is hybridized with the FEM. At the boundary surface of two adjacent regions (e.g., regions i and j), the boundary conditions imposed for the corresponding single-region problems can be combined in various ways, as shown by Mautz and Harrington [44]. Particular attention will be devoted to Muller and PMCHW formulation.

4.2.1 EFIE, MFIE, and CFIE for Multiple-Region Problems

The EFIE can be derived either from (4.61b) or (4.66b). (In both cases the same equation is obtained.) Here, we start with (4.66b). The total electric field in region $0 - i$ can be expressed as

$$\mathbf{E}^{(0-i)} = \sum_{\substack{k=0 \\ k\neq i}}^{n} \mathbf{E}^{(0-i)}(\mathbf{J}_{sik}, \mathbf{M}_{sik}) + \mathbf{E}_{\text{inc}}^{(0-i)} \tag{4.67}$$

where $\mathbf{E}^{(0-i)}(\mathbf{J}_{sik}, \mathbf{M}_{sik})$ represents the scattered field outside region i, which is produced by currents placed on the boundary surface between regions i and k, and $\mathbf{E}_{\text{inc}}^{(0-i)}$ is the corresponding incident field. The kth term exists in the above sum only if regions i and k have a common boundary surface. We suppose the incident electric field continuous across the boundary surface. Using (3.51) and (3.60a), the scattered field at the boundary surface between regions i and j, just outside region i, due to the currents placed on this boundary surface, can be expressed as

$$\mathbf{E}^{(0-i)}(\mathbf{J}_{sij}, \mathbf{M}_{sij}) = -Z^{(i)} L_0^{(i)}(\mathbf{J}_{sij}) + K_0^{(i)}(\mathbf{M}_{sij}) - \frac{\rho_{sij}}{2\epsilon^{(i)}}\mathbf{n}_{ij} + \frac{1}{2}\mathbf{M}_{sij} \times \mathbf{n}_{ij} \tag{4.68}$$

where $Z^{(i)}$ is the intrinsic impedance of the medium filling region i, $L_0^{(i)}$ and $K_0^{(i)}$ are operators defined by (3.44) applied in the principal value sense, and ρ_{sij} is the density of charges placed on the boundary surface. Similarly, the scattered field at the boundary surface of regions i and j, just outside region i, due to the currents placed on the boundary surface of regions i and k, can be expressed as

$$\mathbf{E}^{(0-i)}(\mathbf{J}_{sik}, \mathbf{M}_{sik}) = -Z^{(i)} L_0^{(i)}(\mathbf{J}_{sik}) + K_0^{(i)}(\mathbf{M}_{sik}) \tag{4.69}$$

Finally, after substituting (4.68) and (4.69) into (4.67), and the resulting equation into (4.66b), the general form of the EFIE for the boundary surface of regions i and j is obtained as

$$\sum_{\substack{k=0\\k\neq i}}^{n} \mathbf{n}_{ij} \times [Z^{(i)} L_0^{(i)}(\mathbf{J}_{sik}) - K_0^{(i)}(\mathbf{M}_{sik})] - \frac{1}{2}\mathbf{M}_{sij} = \mathbf{n}_{ij} \times \mathbf{E}_{\text{inc}}^{(0-i)} \tag{4.70}$$

The MFIE can be derived either from (4.61a) or (4.66a). On the other hand, the conditions (4.66a,b) are dual, and the expressions for the electric and magnetic fields due to the electric and magnetic currents are also dual [see (3.60)]. It means that the EFIE and the MFIE for multiple-region problems are dual equations. After applying the duality relations given in Table 3.2 to the EFIE (4.70), the MFIE is obtained in the form

$$\sum_{\substack{k=0\\k\neq i}}^{n} \mathbf{n}_{ij} \times [K_0^{(i)}(\mathbf{J}_{sik}) + Y^{(i)} L_0^{(i)}(\mathbf{M}_{sik})] + \frac{1}{2}\mathbf{J}_{sij} = \mathbf{n}_{ij} \times \mathbf{H}_{\text{inc}}^{(0-i)} \tag{4.71}$$

The CFIE for a multiple-region problem can be derived starting from the combined boundary condition

$$\alpha(\mathbf{E}^{(0-i)})_{\tan} + \mathbf{n}_{ij} \times \mathbf{H}^{(0-i)} = 0 \tag{4.72}$$

This equation is identical to (4.26a), except that it is written for the boundary surface between regions i and j. It represents a linear combination of conditions (4.66a,b), which result in the EFIE and the MFIE given by (4.70) and (4.71), respectively. Combining these EFIE and MFIE, we obtain the CFIE as

$$\begin{aligned} &\alpha \sum_{\substack{k=0\\k\neq i}}^{n} [Z^{(i)} L_0^{(i)}(\mathbf{J}_{sik}) - K_0^{(i)}(\mathbf{M}_{sik})]_{\tan} \\ &+ \sum_{\substack{k=0\\k\neq i}}^{n} \mathbf{n}_{ij} \times [K_0^{(i)}(\mathbf{J}_{sik}) + Y^{(i)} L_0^{(i)}(\mathbf{M}_{sik})] \\ &+ \frac{1}{2}(\mathbf{J}_{sij} + \alpha \mathbf{n}_{ij} \times \mathbf{M}_{sij}) = \alpha(\mathbf{E}_{\text{inc}}^{(0-i)})_{\tan} + \mathbf{n}_{ij} \times \mathbf{H}_{\text{inc}}^{(0-i)} \end{aligned} \tag{4.73}$$

Equations (4.70), (4.71), and (4.73) are also valid if $j = 0$. For a PEC surface, we should set $\mathbf{M}_{si0} = 0$; for a PMC surface $\mathbf{J}_{si0} = 0$; and for an IBC surface $\mathbf{J}_{si0}$ and $\mathbf{M}_{si0}$ are related through the IBC for currents. For $n = 1$ these

equations written for a PEC surface reduce to the single-region expressions (4.10b), (4.12), and (4.29).

4.2.2 CRIEs

In this section we simultaneously consider single-region problems for two neighboring regions (e.g., regions i and j in Figure 4.14). For each region we can impose a set of two boundary conditions (4.66a,b), which results in a total of four boundary conditions. As shown in [44], these four conditions can be merged into two equations. Boundary conditions (4.66a,b) written for region j are multiplied by arbitrary constant α and β, respectively, and added to conditions (4.66a,b) written for region i, resulting in

$$\mathbf{n}_{ij} \times \mathbf{E}^{(0-i)} + \alpha \mathbf{n}_{ji} \times \mathbf{E}^{(0-j)} = 0 \qquad \mathbf{n}_{ij} \times \mathbf{H}^{(0-i)} + \beta \mathbf{n}_{ji} \times \mathbf{H}^{(0-j)} = 0 \tag{4.74a,b}$$

We shall refer to (4.74a,b) as combined-region boundary conditions.

Let us see when conditions (4.74a,b) provide a unique solution for a multiple-region problem. Suppose that the four basic boundary conditions are satisfied at all boundary surfaces except the surface between regions i and j. The complex powers entering the zero-field regions surrounding regions i and j are

$$P_{ij} = -\oint_{S_{ij}} [\mathbf{E}^{(0-i)} \times (\mathbf{H}^{(0-i)})^*] \cdot \mathbf{n}_{ij}\, dS \tag{4.75a}$$

$$P_{ji} = -\oint_{S_{ji}} [\mathbf{E}^{(0-j)} \times (\mathbf{H}^{(0-j)})^*] \cdot \mathbf{n}_{ji}\, dS \tag{4.75b}$$

Since there are no sources in the zero-field regions, the real parts of these powers should be nonnegative; that is,

$$\mathrm{Re}\{P_{ij}\} \geq 0 \qquad \mathrm{Re}\{P_{ji}\} \geq 0 \tag{4.76a,b}$$

From (4.74), $\mathbf{E}^{(0-j)}$ and $\mathbf{H}^{(0-j)}$ can be expressed in terms of $\mathbf{E}^{(0-i)}$ and $\mathbf{H}^{(0-i)}$, and then substituted into (4.75b). Since $\mathbf{n}_{ij} = -\mathbf{n}_{ji}$, the complex power entering the zero-field region surrounding region j is obtained as

$$P_{ji} = \frac{1}{\alpha\beta^*} \oint_{S_{ji}} [\mathbf{E}^{(0-i)} \times (\mathbf{H}^{(0-i)})^*] \cdot \mathbf{n}_{ij}\, dS = \frac{-1}{\alpha\beta^*} P_{ij} \tag{4.77}$$

Adopt that the product $\alpha\beta^*$ is real and positive. If one of $\mathrm{Re}\{P_{ij}\}$ and $\mathrm{Re}\{P_{ji}\}$ is nonnegative, another one must be nonpositive. However, according to (4.76) both powers must be nonnegative. This is possible only if both P_{ij} and P_{ji} are zero. Hence, the satisfaction of (4.74) provides a unique solution for the fields if $\alpha\beta^*$ is real and positive.

The general form of integral equations based on (4.74) is obtained by writing the EFIE (4.70) and the MFIE (4.71) separately for regions i and j and combining the expressions according to (4.74):

$$\mathbf{n}_{ij} \times \left\{ \sum_{k=0,k\neq i}^{n} [Z^{(i)} L_0^{(i)}(\mathbf{J}_{sik}) - K_0^{(i)}(\mathbf{M}_{sik})] - \alpha \sum_{k=0,k\neq j}^{n} [Z^{(j)} L_0^{(j)}(\mathbf{J}_{sjk}) - K_0^{(j)}(\mathbf{M}_{sjk})] \right\} - \frac{1}{2}(1-\alpha)\mathbf{M}_{sij} = \mathbf{n}_{ij} \times (\mathbf{E}_{\text{inc}}^{(0-i)} - \alpha \mathbf{E}_{\text{inc}}^{(0-j)}) \tag{4.78a}$$

$$\mathbf{n}_{ij} \times \left\{ \sum_{k=0,k\neq i}^{n} [K_0^{(i)}(\mathbf{J}_{sik}) + Y^{(i)} L_0^{(i)}(\mathbf{M}_{sik})] - \beta \sum_{k=0,k\neq j}^{n} [K_0^{(j)}(\mathbf{J}_{sjk}) + Y^{(j)} L_0^{(j)}(\mathbf{M}_{sjk})] \right\} + \frac{1}{2}(1-\beta)\mathbf{J}_{sij} = \mathbf{n}_{ij} \times (\mathbf{H}_{\text{inc}}^{(0-i)} - \beta \mathbf{H}_{\text{inc}}^{(0-j)}) \tag{4.78b}$$

We shall refer to (4.78) as the CRIE. A variety of CRIEs is obtained for various choices of α and β.

4.2.3 Muller Formulation

This formulation is usually attributed to Muller [45]. It was applied to dielectric cylinders by Solodukhov and Vasil'ev [46] and Morita [47], and to bodies of revolution by Vasil'ev and Materikova [48]. The Muller equations are a particular form of CRIE (4.74) obtained for

$$\alpha = -\frac{\epsilon^{(j)}}{\epsilon^{(i)}} \qquad \beta = -\frac{\mu^{(j)}}{\mu^{(i)}} \tag{4.79}$$

This choice of α (β) eliminates the quasistatic electric (magnetic) field due to the electric (magnetic) charges associated with $\mathbf{J}_{sij}$ ($\mathbf{M}_{sij}$) on the left side of (4.78a) [(4.78b)]. As a result, the numerical evaluation of integrals occurring in the CRIEs is slightly facilitated.

The Muller CRIE is suitable for boundary surfaces between two regions filled by materials of similar properties. For example, consider an isolated dielectric body of a unit relative permittivity, placed in a vacuum. In that case, the number of regions is two. (According to the convention in Section 3.7, the exterior region filled by vacuum is denoted by 1, which means that the region occupied by dielectric body is denoted by 2.) The CRIEs (4.78a,b) reduce to

$$\mathbf{M}_{s12} = -\mathbf{n}_{12} \times (\mathbf{E}_{\text{inc}}^{(0-1)} + \mathbf{E}_{\text{inc}}^{(0-2)}) \qquad \mathbf{J}_{s12} = \mathbf{n}_{12} \times (\mathbf{H}_{\text{inc}}^{(0-1)} + \mathbf{H}_{\text{inc}}^{(0-2)}) \tag{4.80a,b}$$

When these equations are solved by the MoM, the unknown quantities ($\mathbf{M}_{s12}$ and $\mathbf{J}_{s12}$) are directly tested (see Chapter 8). Hence, the error of solution for these currents is directly minimized, which is not the case with other integral equations used for solution of multiple-region problems. If the relative permittivity of the dielectric body is close to 1, the terms $\mathbf{M}_{sij}$ and $\mathbf{J}_{sij}$ are dominant in the Muller CRIE. According to [42, 44], in such cases, the Muller formulation gives more accurate results for currents than other integral equations.

4.2.4 PMCHW Formulation

The PMCHW formulation was introduced by Poggio and Miller [19]. Cheng and Harrington applied this formulation to cylinders [49], and Wu and Tsai applied it to bodies of revolution [50]. Using the initials of the authors, Mautz and Harrington introduced the acronym for this formulation [44].

The PMCHW equations are a particular form of CRIE (4.74) obtained for

$$\alpha = 1 \qquad \beta = 1 \tag{4.81a,b}$$

Such a choice eliminates the terms $\mathbf{M}_{sij}$ and $\mathbf{J}_{sij}$ on the left sides of CRIE (4.78), resulting in [51]

$$\mathbf{n}_{ij} \times \left\{ \sum_{k=0,k\neq i}^{n} [Z^{(i)} L_0^{(i)}(\mathbf{J}_{sik}) - K_0^{(i)}(\mathbf{M}_{sik})] \right.$$

$$\left. - \sum_{k=0,k\neq j}^{n} [Z^{(j)} L_0^{(j)}(\mathbf{J}_{sjk}) - K_0^{(j)}(\mathbf{M}_{sjk})] \right\} \tag{4.82a}$$

$$= \mathbf{n}_{ij} \times (\mathbf{E}_{\text{inc}}^{(0-i)} - \mathbf{E}_{\text{inc}}^{(0-j)})$$

$$\mathbf{n}_{ij} \times \left\{ \sum_{k=0,k\neq i}^{n} [K_0^{(i)}(\mathbf{J}_{sik}) + Y^{(i)} L_0^{(i)}(\mathbf{M}_{sik})] \right.$$

$$\left. - \sum_{k=0,k\neq j}^{n} [K_0^{(j)}(\mathbf{J}_{sjk}) + Y^{(j)} L_0^{(j)}(\mathbf{M}_{sjk})] \right\} \tag{4.82b}$$

$$= \mathbf{n}_{ij} \times (\mathbf{H}_{\text{inc}}^{(0-i)} - \mathbf{H}_{\text{inc}}^{(0-j)})$$

These two equations are the most frequently used variant of CRIEs [42, 51–53].

Using the boundary conditions (4.65a,b) written for single-region problems in regions i and j, the boundary conditions (4.74a,b) with $\alpha = 1$ and $\beta = 1$ can be transformed into (4.63a,b), which are the boundary conditions for a multiple-region problem. An alternative form of these boundary conditions can be written as

$$(\mathbf{E}^{(i)})_{\tan} = (\mathbf{E}^{(j)})_{\tan} \qquad (\mathbf{H}^{(i)})_{\tan} = (\mathbf{H}^{(j)})_{\tan} \tag{4.83a,b}$$

Hence, the PMCHW formulation (4.82) can be directly obtained by imposing the boundary conditions for a multiple-region problem, without decomposing it into single-region problems.

The PMCHW is often considered unsuitable for the analysis of dielectric bodies whose permittivity is very close to the permittivity of the surrounding medium. Mautz and Harrington found that the Muller solution was more accurate than the PMCHW solution in such cases [44]. It is also reported that the PMCHW cannot handle successfully materials with high permittivity. Sarkar et al. reported that a triangular patch surface code based on the PMCHW formulation yielded unstable results for $\epsilon_r = 1 - j1{,}000$ [54]. The following examples show that PMCHW, if properly implemented, can handle very accurately materials from very low to very high permittivities.

First, consider a cubical scatterer of side a placed in a vacuum and excited by a plane wave [Figure 4.4(a)]. If we assume that the cube is made of a PEC,

we obtain the normalized RCS shown by the solid line in Figure 4.15. Let us now assume that the cube is made of a dielectric material of parameters $\epsilon_r = 1$ and $\mu_r = 1$ (i.e., the cube is empty). In this case, the scattered field should theoretically vanish. However, a numerical solution involves approximate electric and magnetic equivalent currents placed over the cube surface. The difference between the approximate and the exact equivalent currents produces an erroneous radiated field. The ratio of this erroneous field and the field scattered from the metallic cube can be considered as a figure of merit of the applied approximation. Figure 4.15 shows the erroneous monostatic RCS versus the electric length of the cube side for various orders of current approximation. By increasing this order, the erroneous field decreases up to some value. The maximal difference between the monostatic RCS of the metallic cube and the erroneous monostatic RCS is about 60 dB for electrically small cubes and about 45 dB for electrically large cubes. For most applications, such suppression of the erroneous field is satisfactory.

A further suppression of the erroneous field cannot be achieved by a further increase of the approximation order. It can be shown that the limit observed in Figure 4.15 is related to the stability of MoM matrices. These results were obtained in single precision (i.e., by using about seven significant digits in computation). This limit for the erroneous field can be lowered only by increasing the number of significant digits.

A second example considers a cubical scatterer placed in a vacuum and excited by a plane wave [Figure 4.4(a)]. The side length is $a = 0.2\lambda_0$, where

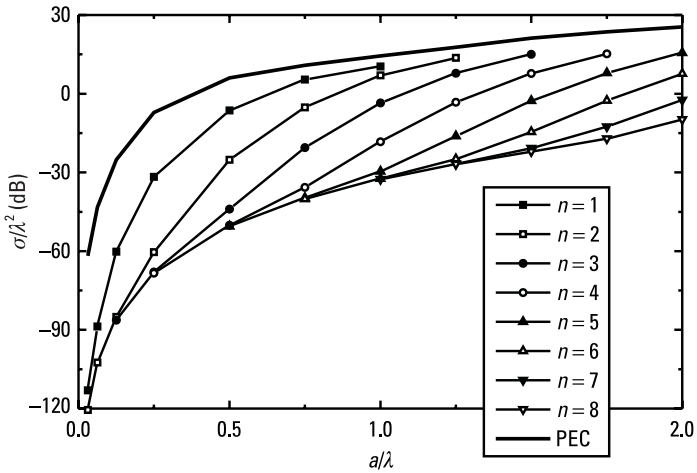

Figure 4.15 Erroneous monostatic RCS of an empty cube, shown in Figure 4.4(a), versus electric length of cube side for various orders of current approximation. Results for a metallic cube are given for comparison. (© 2002 ACES [12].)

λ_0 is the wavelength in a vacuum. The cube is made of a dielectric, whose relative permittivity is $\epsilon_r = 1 - j10^K$, $K = 2, 3, 4$. The second-order approximation is used for currents over each cube side, resulting in $N = 96$ unknowns. Figure 4.16 shows the bistatic RCS versus the angle θ. With increases in the coefficient K, the results tend to those obtained for a PEC cube, as they should. Stable and accurate results are obtained even for $K = 4$ (i.e., $\epsilon_r = 1 - j10{,}000$).

4.2.5 Thin Plate at an Interface Between Two Regions

Consider a PEC body in the form of a thin plate at an interface between two regions, such as regions i and j [see Figure 4.17(a)]. When the actual thickness

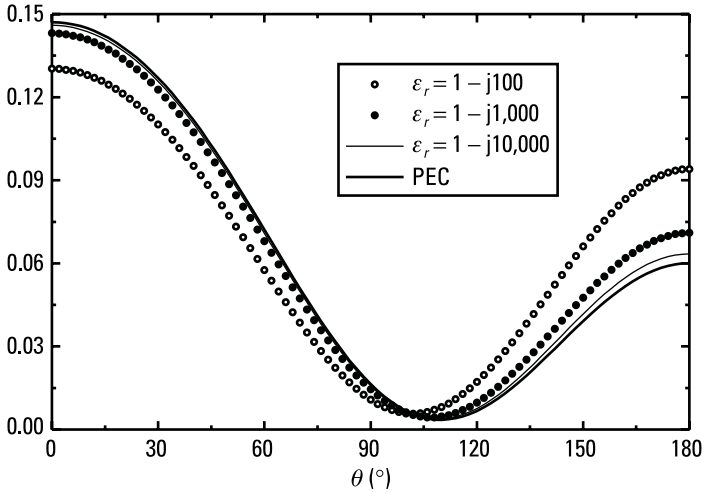

Figure 4.16 Bistatic RCS of the cubical scatterer of side $a = 0.2\lambda_0$, versus the angle θ. The cube is made either of a PEC or a dielectric whose relative permittivity is $\epsilon_r = 1 - j10^K$, $K = 2, 3, 4$. (© 2002 ACES [12].)

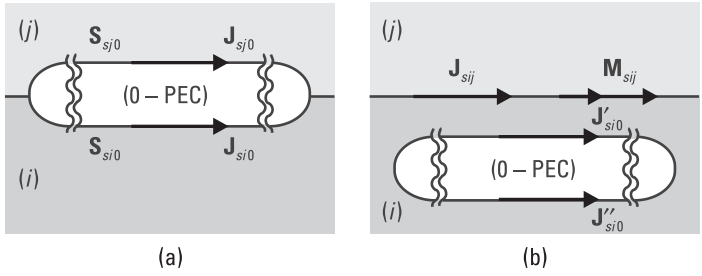

Figure 4.17 Thin plate at interface of two regions: (a) original problem, and (b) equivalent problem.

of the plate is taken into account, the analysis can be performed as for any solid PEC body: The integral equation is imposed for each face of the plate.

When the plate thickness tends to zero, the two opposite faces of the plate merge into one surface. The question is if it is possible to merge the currents placed on the opposite sides of such an infinitesimally thin plate, $\mathbf{J}_{si0}$ and $\mathbf{J}_{sj0}$, into a single current sheet. The answer is negative. This can be seen from the following consideration. When an integral equation is imposed for the surface S_{i0} only currents belonging to region i enter into the equation, including $\mathbf{J}_{si0}$. Since $\mathbf{J}_{sj0}$ does not belong to region i, it is not included into this equation. Similarly, $\mathbf{J}_{sj0}$ enters into the integral equation imposed for S_{j0}, while $\mathbf{J}_{si0}$ is not included. Both currents enter into the integral equation imposed for S_{ij}, but $\mathbf{J}_{si0}$ is under operators $L_0^{(i)}$ and $K_0^{(i)}$, while $\mathbf{J}_{sj0}$ is under operators $L_0^{(j)}$ and $K_0^{(j)}$. Since in each of these equations currents $\mathbf{J}_{si0}$ and $\mathbf{J}_{sj0}$ appear independently, they cannot be merged into a single current.

The field in the vicinity of the plate will not practically change if we slightly immerse the plate into one of the regions, such as into region i [see Figure 4.17(b)]. In that case, the currents located on the opposite faces of the plate ($\mathbf{J}'_{si0}$ and $\mathbf{J}''_{si0}$) can be merged into a single current sheet (Section 4.1.9). However, we now have an additional boundary surface between the two regions and additional electric and magnetic currents $\mathbf{J}_{sij}$ and $\mathbf{M}_{sij}$. In other words, the currents $\mathbf{J}_{si0}$ and $\mathbf{J}_{sj0}$ are replaced by currents $\mathbf{J}'_{si0} + \mathbf{J}''_{si0}$, $\mathbf{J}_{sij}$, and $\mathbf{M}_{sij}$, resulting in 50% more unknowns, as shown by Rao et al. [55].

4.2.6 Optimal Choice of BIEs for Analysis of Multiple-Region Problems

For pure dielectric structures (not combined with PEC, PMC, and IBC bodies), the five most important general types of BIEs are the EFIE, MFIE, CFIE, PMCHW, and Muller equation. If the analysis is performed by EFIE and the MFIE, the spurious fields can occur under specific conditions inside the dielectric bodies in the vicinity of their interior resonant frequencies. Other three equations give solutions that are free of spurious fields. The first three equations are imposed for one side of the region boundary, while the last two combine fields from both sides of the boundary between two regions. Hence, the PMCHW and Muller equation cannot be incorporated into a hybrid method in which one region is treated by BIEs and another region is treated by some other type of equations [e.g., differential equations (like the FEM)].

All five equations are based on both operators L and K. Hence, their application could seem equally difficult. However, as explained in Section 4.1.13, a big problem in the implementation is to determine the direction of the outward normal unit vector on the body surface. This problem is avoided

only for the PMCHW because the direction of the unit normal vector does not influence this equation.

In general, pure dielectric structures are combined with PEC, PMC, and IBC bodies. (IBC bodies are used when the skin effect is pronounced. However, the same problem can be solved more easily by using the concept of distributed loadings for metallic surfaces, as will be shown in Section 4.3.2. Hence, in what follows, the IBC bodies will not be discussed.) All five equations, except the Muller one, can be used for PEC and PMC surfaces. The PMCHW degenerates into the EFIE at PEC surfaces and into the MFIE at PMC surfaces. Hence, the PMCHW, like the EFIE, can be used for the PEC thin-wire and thin-plate approximations. Similarly, like the MFIE, the PMCHW can be used for the PMC thin-wire and thin-plate approximations. In such cases, the EFIE and MFIE can be also considered as degenerate forms of the CFIE.

The case when the PMCHW is used for both PEC and dielectric surfaces is sometimes termed the EFIE + PMCHW method. Beginning with the paper of Medgyesi-Mitschang and Eftimiu [56], this method has been frequently used for the analysis of composite metallic and dielectric structures.

Only the CFIE inherently suppresses the spurious resonances inside PEC bodies. (With other equations, the suppression can be performed by using higher-order basis functions, minimum norm solution, and losses.) Another possibility is to hybridize the CFIE for PEC bodies with the PMCHW for dielectric bodies, as shown by Huddleston et al. [57] and Putnam and Medgyesi-Mitschang [58].

All of the above properties are summarized in Table 4.2. The CFIE and the PMCHW are obviously superior to other equations. If the spurious fields must positively be suppressed, or if the equation should be hybridized with the FEM, the CFIE is probably the best candidate. However, the simplest for implementation is the PMCHW.

4.3 CIEs

Any composite metallic and dielectric structure can be considered as a specific inhomogeneous medium. By using the volume equivalence theorem, the medium can be made homogeneous in the whole infinite space by introducing equivalent volume electric and magnetic currents, as explained in Section 3.4. Once these volume currents are determined, all other quantities of interest can be easily evaluated.

Any current distribution in such an infinite homogeneous medium automatically satisfies Maxwell's equations because the current density appears on the right side of these equations as the source term. From Maxwell's equations

Table 4.2
Summary of Properties of BIEs Used for Composite PEC, PMC, and Dielectric Structures

Equation	Inherently Suppresses Spurious Fields	Single-Side Equation	Without Outward Normal Direction	PEC	Thin PEC	PMC	Thin PMC
EFIE		+		+	+	+	
MFIE		+		+		+	+
CFIE	+	+		[+]*		[+]	
Muller PMCHW			+	[+]	[+]	[+]	[+]

* Brackets signify that degenerate forms of the equation satisfy the property.

for such an infinite space, we can uniquely determine the electromagnetic field due to the current distribution, but not the distribution itself. However, there is a unique current distribution that satisfies the constitutive relations resulting from the volume equivalence theorem. Hence, in determining these currents, we start from the constitutive relations.

Most often the inhomogeneous region i can be represented as a body made of an inhomogeneous medium of parameters ϵ_m, μ_m, and σ_m, surrounded by a homogenous medium of parameters ϵ, μ, and σ. After replacing the constitutive relations (3.7a-c) for the body into the expressions for equivalent volume currents (3.54) and (3.55), the constitutive relations are obtained in the form

$$\mathbf{J} = \sigma_E \mathbf{E} \qquad \sigma_E = \sigma_m - \sigma + j\omega(\epsilon_m - \epsilon) \tag{4.84a,b}$$

$$\mathbf{M} = \sigma_H \mathbf{H} \qquad \sigma_H = j\omega(\mu_m - \mu) \tag{4.85a,b}$$

where σ_E and σ_H are equivalent electric and magnetic conductivities, respectively. Similarly, constitutive relations can be derived for anisotropic and bianisotropic media. Starting from various constitutive relations, and expressing field vectors $\mathbf{E}$ and $\mathbf{H}$ in terms of equivalent volume currents $\mathbf{J}$ and $\mathbf{M}$, different CIEs are obtained for the volume currents.

The CIE was first applied by Richmond in 1965 to solve the problem of 2-D scattering from infinite dielectric cylinders [59]. Thereafter many papers have been published, an exhaustive list of which is made by Notaros [60]. According to this list, the CIEs for isotropic media are most often applied. After Livesay and Chen [61], some authors dealt with the anisotropic dielectric problem, usually interested in the field penetration into biological bodies. The CIEs can be written for volumes, surfaces, lines, and even points. In what follows, all these cases will be elaborated for isotropic media.

4.3.1 VIEs

Using (3.60), the total electric and magnetic fields, $\mathbf{E}$ and $\mathbf{H}$, are expressed in terms of equivalent volume currents, $\mathbf{J}$ and $\mathbf{M}$, as

$$\mathbf{E} = -ZL(\mathbf{J}) + K(\mathbf{M}) + \mathbf{E}_{\text{inc}} \qquad \mathbf{H} = -K(\mathbf{J}) - YL(\mathbf{M}) + \mathbf{H}_{\text{inc}} \tag{4.86a,b}$$

where L and K are the integral operators defined in Table 3.1, and $\mathbf{E}_{\text{inc}}$ and $\mathbf{H}_{\text{inc}}$ are the incident fields.[7] After substituting (4.86) into (4.84a) and (4.85a),

7. Any set of currents plugged into (4.86) results in vectors $\mathbf{E}$ and $\mathbf{H}$, which satisfy Maxwell's equations (3.56). To force these currents to represent the solution for the given incident field, they must satisfy constitutive equations (4.84) and (4.85).

two coupled integral equations for the unknown electric and magnetic currents are obtained as

$$\frac{\mathbf{J}}{\sigma_E} + ZL(\mathbf{J}) - K(\mathbf{M}) = \mathbf{E}_{\text{inc}} \qquad \frac{\mathbf{M}}{\sigma_H} + YL(\mathbf{M}) + K(\mathbf{J}) = \mathbf{H}_{\text{inc}} \tag{4.87a,b}$$

These integral equations are referred to as the VIEs. When $\mu_m = \mu$, the magnetic currents are zero and (4.87a) suffices. When $\epsilon_m = \epsilon$, the electric currents are zero and (4.87b) suffices. The CIEs are most often applied to dielectric structures [62, 63] and rarely to magnetic structures [62].

When compared with the surface formulation based on the BIEs, an advantage of the volume formulation is that it does not suffer from spurious resonances. On the other side, the VIEs usually require many more unknowns than BIEs [64]. The solution of large VIE systems can be significantly accelerated if the *conjugate gradient method* (CGM) is combined with the *fast Fourier transform* (FFT) [65]. It seems that the VIEs are not suitable for analysis of thin metallic wires and plates. Therefore, some authors combine the VIE for dielectrics with the surface formulation of the BIEs to treat composite metallic and dielectric structures [66]. However, in that case, the FFT cannot be easily combined with the CGM. Hence, the authors recommend the VIEs primarily for anisotropic and bianisotropic media, which cannot be handled by the BIEs.

4.3.2 Metallic Surfaces with Distributed Loadings

Consider a thin[8] layer of thickness δ, made of a homogeneous medium of parameters ϵ_m, μ_0, and σ_m, surrounded by a nonconducting homogenous medium of parameters ϵ and μ_0. We can homogenize the medium inside the layer with respect to the surrounding medium by introducing equivalent volume electric currents. If $|\sigma_m + j\omega\epsilon_m| >> |j\omega\epsilon|$, currents normal to the layer surface can be neglected[9] (see Figure 4.18). If the skin effect is not pronounced, the currents tangential to the layer surface are approximately uniformly distributed along the transverse direction. The tangential component of the electric field inside the layer and in its vicinity, $\mathbf{E}_{\text{tan}}$, will not practically change if the

8. A layer is assumed to be thin based on the criteria discussed in Section 4.9. As a rule of thumb, δ should be less than 0.1λ.
9. For this particular case, boundary conditions (3.61a) and (3.61c) reduce to $\mathbf{E}_{\text{tan}}^{(i)} = \mathbf{E}_{\text{tan}}^{(j)}$ and $(\sigma^{(i)} + j\omega\epsilon^{(i)})\mathbf{E}_{\text{n}}^{(i)} = (\sigma^{(j)} + j\omega\epsilon^{(j)})\mathbf{E}_{\text{n}}^{(j)}$. If $|\sigma^{(i)} + j\omega\epsilon^{(i)}| << |\sigma^{(j)} + j\omega\epsilon^{(j)}|$, the electric field in medium j is practically tangential to the boundary surface. In our case, the index i corresponds to the layer and j to the surrounding medium.

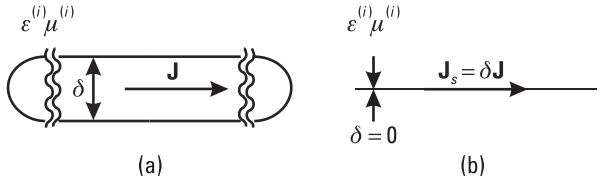

Figure 4.18 Equivalenting of a thin-material layer by (a) volume electric current and (b) electric-current sheet.

volume currents are replaced by surface currents $\mathbf{J}_s = \mathbf{J}\delta$, [Figure 4.18(b)]. The constitutive equation (4.84a) reduces to

$$\mathbf{E}_{\tan} = Z_s\mathbf{J}_s \qquad Z_s = R_s + jX_s = \frac{1}{\sigma_E\delta} \tag{4.88a,b}$$

where Z_s is the surface impedance, and σ_E is given by (4.84b). When σ_m tends to infinity, Z_s tends to zero, and the thin-layer becomes a PEC. When Z_s differs from zero, a nonmagnetic layer can be considered as a metallic surface (plate) with a distributed loading.

Suppose that a loaded metallic plate is placed inside an infinite region i. According to the notation in Section 4.1.1, the region occupied by the metallic plate is region 0, $\mathbf{E}_{\tan}$ is written as $(\mathbf{E}^{(0)})_{\tan}$, and $\mathbf{J}_s$ is written as $\mathbf{J}_{si0}$. For $Z_s = 0$, (4.88a) reduces to the boundary condition (4.9b), which results in the EFIE (4.10b). For $Z_s \neq 0$, the CIE is easily obtained by adding the term $n_{i0}Z_s\mathbf{J}_{si0}$ to the left side of (4.10b); that is,

$$Z^{(i)}[L_0^{(0)}(\mathbf{J}_{si0})]_{\tan} + Z_s\mathbf{J}_{si0} = (\mathbf{E}_{\text{inc}}^{(0)})_{\tan} \tag{4.89}$$

In what follows this equation is referred to as the surface (thin-plate) CIE.

The concept of a resistive surface ($Z_s = R_s$) was introduced by Lang [67]. Medgyesi-Mitschang and Putnam derived the integral equation for the resistive surface in the same paper in which they derived the IBC-CFIE [41]. However, they gave numerical results only for the IBC-CFIE, probably considering the other two integral equations less important. Thereafter, many authors have applied the thin-plate CIE.

There are three important cases when the surface CIEs are much more efficient than the BIEs: (a) a resistive layer, (b) a dielectric layer, and (c) a skin-effect surface. For the resistive and the dielectric layer, the surface impedance is obtained directly from (4.88b) and (4.84b), as given in Table 4.3. When such layers are analyzed by the BIEs, different equivalent electric and magnetic currents are obtained for the opposite layer faces. When the surface CIE is

Table 4.3
Summary of Most Important Surface Impedances

	Resistive Layer	**Dielectric Layer**	**Skin-Effect Surface**
Surface impedance	$Z_s = \dfrac{1}{\sigma_m \delta}$	$Z_s = \dfrac{1}{j\omega(\epsilon_m - \epsilon^{(i)})\delta}$	$Z_s = \sqrt{\dfrac{\pi \mu_0 f}{\sigma_m}}$

used, these four current sheets are replaced by only one electric current sheet, thus reducing the number of unknowns by a factor of four.

For example, consider a symmetrical dipole antenna of arm h = 0.25m and radius r/h = 100, placed in a vacuum in the middle of a spherical radome of inner and outer radii $a - \delta/2$ and $a + \delta/2$, (a = 0.35m, δ = 0.02m), respectively, and relative permittivity $\epsilon_m = 4$. If the radome is analyzed by the PMCHW formulation, each face is modeled by M = 96 patches (Figure 4.19, inset). The rooftop approximation of currents results in N = 768 unknowns. If the radome is analyzed by the surface CIE, it is replaced by a loaded metallic sphere of radius a, modeled by M = 96 patches and only N = 192 unknowns. Figure 4.19 shows the input admittance of the dipole antenna versus frequency. The difference between the results obtained by the two approaches is relatively small compared with the difference between these results and the antenna admittance without the radome. According to the expression for Z_s given in Table 4.3, the results obtained by the surface CIE are identical for all radomes for which $(\epsilon_m/\epsilon_0 - 1)\delta$ = 0.06m. It can be shown that by decreasing δ, but

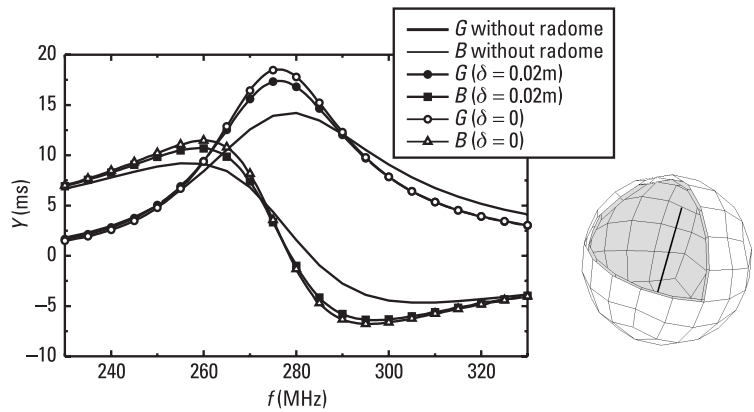

Figure 4.19 Input admittance of a dipole antenna (h = 0.25m, r/h = 100) placed in the center of a spherical radome (a = 0.35m, δ = 0.02m, and ϵ_m = 4). (One-eighth of the radome is not drawn to show its interior.)

keeping $(\epsilon_m/\epsilon_0 - 1)\delta = 0.06$m, the results obtained by the PMCHW tend to the results obtained by the surface CIEs, as expected. For thick layers, the approximation by the loaded surface is not accurate enough, and the PMCHW method should be used.

When a closed metallic body with skin-effect losses is analyzed by the PMCHW, both equivalent electric and magnetic currents should be placed on its surface. If the skin depth is much smaller than the curvature radii of the body surface, the magnetic currents are related to the electric currents through the IBC. Hence, the number of unknowns with the IBC-IE is twice as small as with the PMCHW. The analysis can be further simplified if the impedance boundary surface is replaced by an equivalent resistively loaded metallic surface when the analysis involves only electric currents. However, this is always an approximate equivalence. Namely, the space inside a closed loaded metallic surface is not perfectly isolated from the outer electromagnetic field. If the field that leaks into such a surface is much smaller than the incident field, the approximate equivalence can be established by requiring both surfaces to have identical Joule losses per unit area, that is, that the surface resistance of the distributed loading Z_s be equal to the surface resistance of the skin-effect surface given by the real part of (3.63b).

For example, consider a metallic cubical scatterer of side $a = 2$m and conductivity σ, excited by a plane wave polarized along the z-axis, whose the electric field is $E = 1$ V/m at frequency $f = 150$ MHz. The fifth-order approximation is used for the current distribution over each cube side. The unknown coefficients are determined by using the Galerkin method applied to the EFIE. Figure 4.20 shows the z-component of the electric field along the x-axis. Results are presented for two different conductivities. When $\sigma \to 0$, the field is not exactly zero, which can be explained by numerical errors. Almost identical results (not shown in Figure 4.20) are obtained for $\sigma = 1$ MS/m. If we decrease the conductivity to $\sigma = 0.001$ MS/m, the field penetration through the loaded metallic surfaces cannot be neglected. Since in most applications the conductivity of metals is greater than $\sigma = 1$ MS/m, the simplest way to take into account the skin-effect losses is by using the concept of resistively loaded metallic surfaces.

4.3.3 Wires with Distributed Loadings

Consider a thin rod of radius a made of a homogeneous material of parameters ϵ_m, μ_0, and σ_m, surrounded by a nonconducting homogenous medium of parameters ϵ and μ_0. We can homogenize the medium inside the rod with respect to the surrounding medium by introducing equivalent volume electric currents. If $|\sigma_m + j\omega\epsilon_m| >> |j\omega\epsilon|$, the radial and circular current components

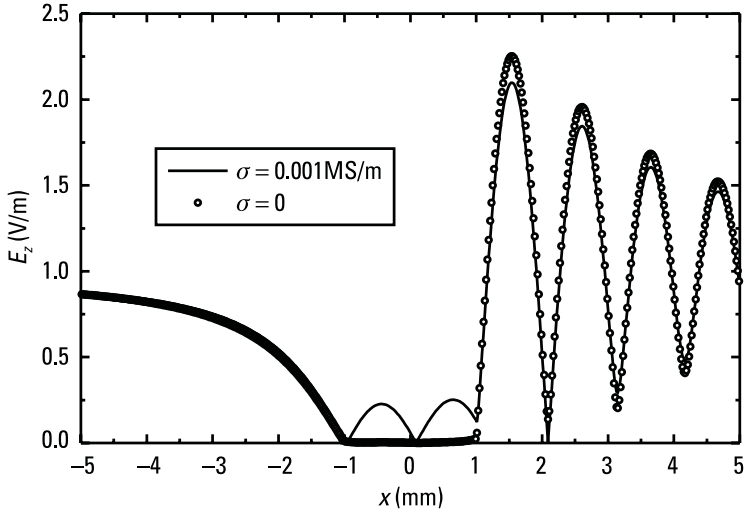

Figure 4.20 The z-component of the electric field along the x-axis inside a metallic cubical scatterer (a = 2m) excited by a z-polarized plane wave (E = 1 V/m). The scatterer is modeled as a closed, loaded metallic surface.

(defined with respect to the rod axis) can be neglected [Figure 4.21(a)]. If the skin effect in the rod is not pronounced, the axial current component is approximately constant over the wire cross section. The axial component of the electric field along the rod axis and the electric field at the points not quite close to the rod practically will not change if the volume currents are replaced by line electric currents of intensity $I = \mathbf{i}_s \cdot \mathbf{J} a^2 \pi$, positioned along an arbitrary wire generatrix [Figure 4.21(b)]. The constitutive equation (4.84a) reduces to

$$\mathbf{i}_s \cdot \mathbf{E} = Z' I \qquad Z' = R' + jX' = \frac{1}{\sigma_E a^2 \pi} \tag{4.90a,b}$$

where $\mathbf{i}_s$ and $\mathbf{E}$ are the unit vector and electric field along the wire axis, Z' is the per-unit-length impedance of the rod, and σ_E is given by (4.84b). When

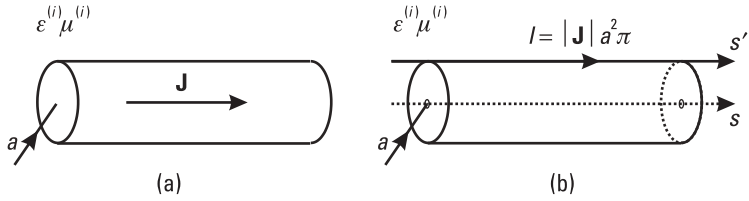

Figure 4.21 Equivalenting a thin-material rod by (a) volume electric currents and (b) line electric currents along a wire generatrix.

σ_m tends to infinity, Z' tends to zero, and the thin rod becomes a PEC. When Z' differs from zero, a nonmagnetic rod can be considered as a wire with a distributed loading.

Suppose that a loaded wire is placed in the vicinity of a PEC body inside an infinite region i [Figure 4.18(a)]. According to the notation in Section 4.1.10, the region occupied by the wire is region 0, $\mathbf{E}$ is written as $\mathbf{E}^{(0)}$, and I is written as I_{i0}. For $Z' = 0$, (4.90a) reduces to the boundary condition (4.52) that results in the EFIE (4.55). For $Z' \neq 0$, the CIE is easily obtained by adding the term $Z'I_{i0}$ to the left side of (4.55); that is,

$$Z^{(i)}\mathbf{i}_s \cdot [L(I_{i0}) + L(\mathbf{J}_{si0})] + Z'I_{i0} = \mathbf{i}_s \cdot \mathbf{E}^{(0)}_{\text{inc}} \tag{4.91}$$

In what follows this equation is referred to as the line (thin-wire) CIE.

If $X' = 0$, the loading is resistive. For $R' = 0$, the loading can be capacitive ($X' < 0$) or inductive ($X' > 0$). A resistive loading that is used in traveling-wave antennas is usually made in the form of a thin resistive layer (of thickness δ) over a dielectric rod [68]. The wire with a pronounced skin-effect can be modeled by an equivalent thin resistive layer. Uniform capacitive loading is made in the form of a dielectric rod. A summary of the corresponding per-unit-length impedances, Z', is given in Table 4.4. A variable capacitive loading can be realized as a series of short wires with dielectric disks inserted between them [69] or as a series of short metallic rings placed along a dielectric rod [11]. An inductive loading is easily realized in the form of a coil.

For example, consider a capacitively loaded dipole antenna in a vacuum excited by a point voltage generator. The arm height is $h = 0.25$m, and its radius is $a = h/10$. The middle third of each arm is replaced by a dielectric rod of relative permittivity $\epsilon_m = 9$. When the analysis is based on the PMCHW formulation, the dipole is modeled by $M = 8$ patches around its circumference (Figure 4.22, inset). The first-order approximation for the currents is used for short edges, and the second-order for long edges, resulting in $N = 165$ unknowns. When the antenna is analyzed by the thin-wire CIE, each arm is made of five

Table 4.4
Summary of Most Important Distributed Loadings Along Wire of Radius a

	Resistive Layer of Thickness δ	**Dielectric Rod**	**Skin-Effect**
Per-unit-length impedance	$Z' = \dfrac{1}{2\pi a\delta\sigma_m}$	$Z' = \dfrac{1}{j\omega(\epsilon_m - \epsilon^{(i)})a^2\pi}$	$Z' = \dfrac{1}{2\pi a}\sqrt{\dfrac{\pi\mu_0 f}{\sigma_m}}$

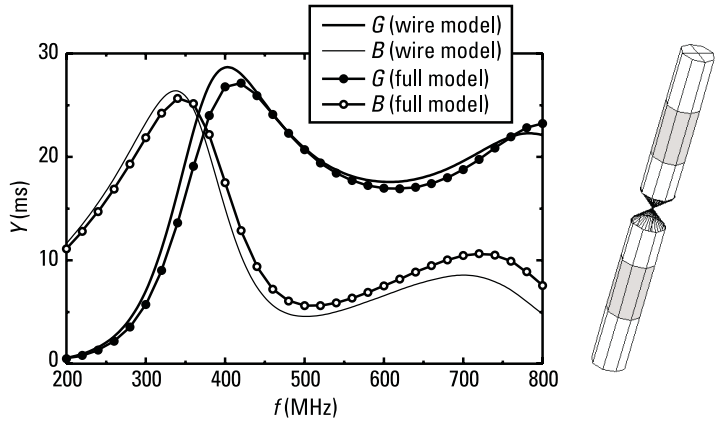

Figure 4.22 Input admittance of a capacitively loaded dipole antenna shown in the inset (h = 0.25m, $h_1 = h/3$, $h_2 = 2h/3$, $r = h/10$, $\epsilon_m = 9$) versus frequency.

wires, one of which is loaded. (The feed area and the wire end are modeled by short conical wires, the second of them being almost flat.) The first-order approximation for currents is used for short wires and the second-order for long wires, resulting in $N = 13$ unknowns. Figure 4.22 shows the antenna input admittance versus frequency. A good agreement between the results obtained by the two models is observed. An even better agreement is obtained when the antenna radius decreases.

4.3.4 Wires with Concentrated Loadings

Consider a wire, a part of which is loaded (Figure 4.23). If this part is very short compared with the wavelength, the total current along this part is approximately constant along it, and the circuit theory can be applied. The voltage between the ends of the loaded part, V, and the total current along the loaded part, I, are related by integrating the left and right sides of the thin-wire CIE (4.90a) along the loaded part; that is,

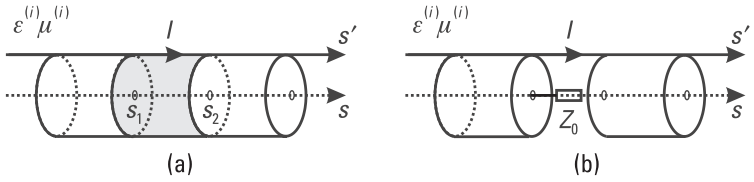

Figure 4.23 (a) Wire with a short distributed loading, and (b) the equivalent concentrated (lumped) loading.

$$V = Z_0 I \qquad V = \int_{s_1}^{s_2} \mathbf{i}_s \cdot \mathbf{E}(s)\, ds \qquad Z_0 = \int_{s_1}^{s_2} Z'(s)\, ds \qquad (4.92a\text{-}c)$$

where Z_0 is the impedance, and s_1 and s_2 are the starting and ending s-coordinates of the loaded part, respectively. The field outside the loaded part practically will not change if the length of the loaded part tends to zero, but the impedance Z_0 remains constant. The resulting degenerate form of the distributed loading is referred to as a concentrated (lumped) loading.

When the length of the loaded part tends to zero, the magnitude of the s-component of the electric field tends to infinity. The constitutive equation for the loaded wire (4.90a) reduces to

$$\mathbf{i}_s \cdot \mathbf{E}(s) = Z_0 I \delta(s - s_0) \qquad (4.93)$$

where s_0 is the s-coordinate where the loading is positioned, and $\delta(s_0 - s)$ is the delta function centered at s_0.[10] The thin-wire CIE (4.91) reduces to

$$Z^{(i)} \mathbf{i}_s \cdot [L(I_{i0}) + L(\mathbf{J}_{si0})] + Z_0 I_{i0} \delta(s - s_0) = \mathbf{i}_s \cdot \mathbf{E}^{(0)}_{\text{inc}} \qquad (4.94)$$

In what follows, this equation is referred to as the thin-wire CIE for lumped loadings.

At lower frequencies concentrated loadings are technically realized by inserting small resistors, capacitors, and coils at desired positions along the wire. For example, resistive and capacitive concentrated loadings are used in the design of broadband and traveling-wave antennas [69–72]. By using the thin-wire CIE for lumped loadings, the analysis of such antennas becomes very efficient when compared with methods that model more precisely the field in the region of loadings. However, by increasing the frequency, the analysis based on the thin-wire CIE for lumped loadings becomes less accurate. This is the same problem as associated with modeling the wire-antenna excitation by a delta-function (point) generator [11]. One source of errors is the parasitic capacitance that appears across the load.

For example, consider a simple circuit consisting of a resistor excited by a point voltage generator (see Figure 4.24). The resistor is made in the form of a resistive cylindrical rod of height h = 10 mm, radius a = 1 mm, and conductivity σ = 63.66 S/m. The cylinder bases are covered by thin metallic

10. Integration of the left and the right side of (4.93) along the wire axis in the vicinity of s_0 results in (4.92a).

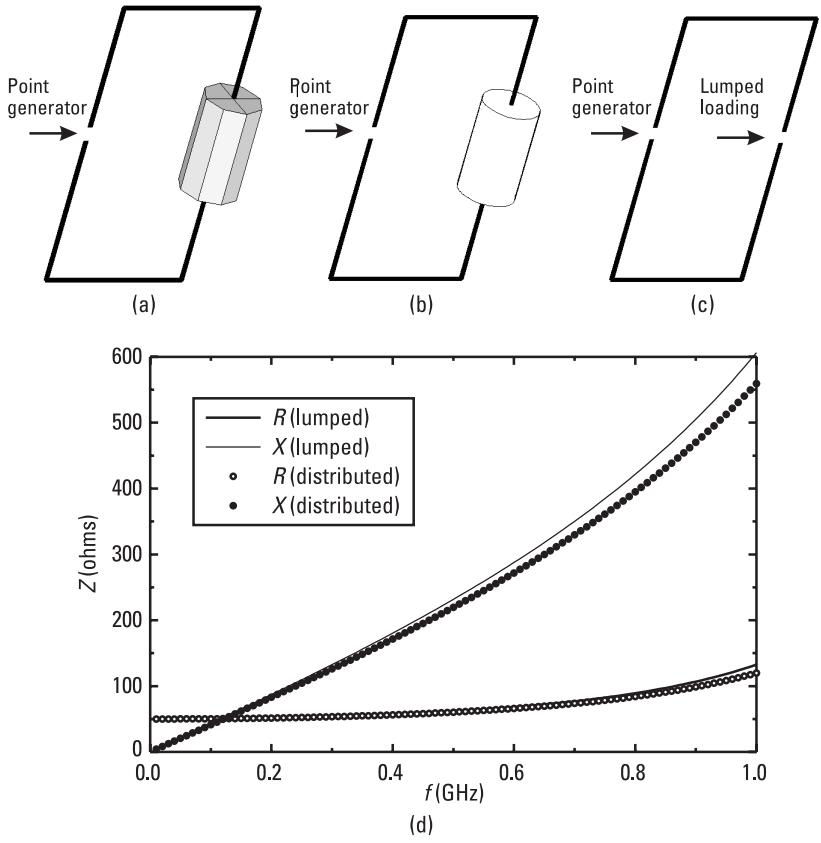

Figure 4.24 Electromagnetic modeling of a resistor: (a) composite metallic and dielectric structure, (b) wire with distributed loading, (c) concentrated (lumped) loading in wire, and (d) input impedance versus frequency as seen by the point generator.

layers (electrodes), which are connected to the generator by thin wires of radius $r = 0.2$ mm. The circuit contour is rectangular of side lengths $l_1 = 15$ mm and $l_2 = 30$mm. The circuit is solved using the EFIE for the thin wires that connect the electrodes with the generator. The resistor is modeled in three ways: (a) the PMCHW is applied to a composite metallic and dielectric structure using $M = 8$ patches around its circumference [Figure 4.24(a)]; (b) the thin-wire CIE is applied to a loaded wire of per-unit-length impedance $Z' = 5\ \text{k}\Omega/\text{m}$ [Figure 4.24(b)]; (c) the CIE for lumped loadings is applied to a lumped loading of impedance $Z_0 = 50\Omega$ [Figure 4.24(c)]. The first order approximation is used for currents along all wires and edges, resulting in $N = 99$, 7, and 6 unknowns, respectively, for the complete circuit. Figure 4.24(d) shows the input impedance versus frequency as seen by the point generator for the second and third model.

(The results obtained by the first two models practically coincide.) A good agreement can be observed. The discrepancy increases at higher frequencies.

4.4 Hybrid Methods

In this book we adopt the BIEs as the basic equations for the analysis of composite metallic and dielectric structures. In many cases these equations enable very efficient solutions. The efficiency may be improved if the BIEs are combined with each other or with other techniques, resulting in hybrid methods. In addition, hybrid techniques can solve some problems that cannot be handled by the BIEs.

All BIEs are essentially SIEs, even if they are degenerate (for wires). Thin-plate and thin-wire CIEs are very similar to BIEs. Hence, various BIEs and CIEs are easily combined. For example, the EFIE for wires was combined with the MFIE for closed metallic bodies by Albertsen et al. [73]. The CFIE for PEC bodies was combined with the PMCHW for dielectric surfaces by Huddleston et al. [57]. The EFIE for wires was combined with the thin-wire CIEs by Popovic et al. [11]. The necessity for a hybridization with other techniques usually occurs when the problem involves (a) highly inhomogeneous regions, (b) anisotropic or bianisotropic regions, (c) antennas in the vicinity of a real ground or a multilayered substrate, or (d) electrically large objects.

4.4.1 Surface/Volume Integral Formulation

In contrast to the BIEs, the analysis of anisotropic and bianisotropic objects can be handled by the VIEs (volume CIEs). However, the VIEs are not suitable for metallic bodies, thin metallic plates, and wires. One possible way to solve a general problem is to combine the BIEs with VIEs, where the VIEs are used for the anisotropic and bianisotropic regions, while the BIEs are used for the remaining regions. Sarkar and Arvas proposed such a method [66], although they applied it only to isotropic dielectrics. For highly inhomogeneous dielectrics, such a hybrid method can be more efficient than a method based only on BIEs [64].

In the hybrid methods, both the BIEs and VIEs involve both surface and volume currents. Hence, the hybrid method is called a surface/volume integral formulation. Conceptually, the hybrid method represents a mixture of two basic methods. However, the implementation of a VIE into an existing code based on a BIE requires a lot of effort. For example, if the BIE is solved by the Galerkin method, double surface integrals should be evaluated. The hybrid surface/volume formulation requires an additional evaluation of double volume integrals, as well as combined surface and volume integrals.

4.4.2 MoM and FEM

Computer modeling of electromagnetic fields in the frequency domain, which is based on numerical solutions of differential equation, is usually performed by the FEM [74–77]. For time-harmonic electromagnetic fields, the FEM can be interpreted as the Galerkin method applied to the solution of wave equations. Most often the FEM is applied to isotropic materials that can be inhomogeneous. For such cases, the first two Maxwell equations, (3.56a) and (3.56b), can be written as

$$\mu_r^{-1}\nabla\times\mathbf{E} = -\mu_r^{-1}\mathbf{M}_i - j\omega\mu_0\mathbf{H} \qquad \epsilon_r^{-1}\nabla\times\mathbf{H} = \epsilon_r^{-1}\mathbf{J}_i + j\omega\epsilon_0\mathbf{E} \tag{4.95a,b}$$

where the relative permittivity (ϵ_r) and permeability (μ_r) depend on spatial coordinates, while $\mathbf{J}_i$ and $\mathbf{M}_i$ are impressed electric and magnetic currents. After applying the curl operator on the left and right sides of (4.95a), this equation can be written as

$$\nabla\times(\mu_r^{-1}\nabla\times\mathbf{E}) = -\nabla\times(\mu_r^{-1}\mathbf{M}_i) - j\omega\mu_0\nabla\times\mathbf{H} \tag{4.96}$$

The term $\nabla\times\mathbf{H}$ on the right side can be replaced from (4.95b). As a result, the wave equation for the electric field is obtained,

$$\nabla\times(\mu_r^{-1}\nabla\times\mathbf{E}) + \gamma_0^2\epsilon_r\mathbf{E} + \gamma_0 Z_0\mathbf{J}_i + \nabla\times(\mu_r^{-1}\mathbf{M}_i) = 0 \tag{4.97}$$

where γ_0 and Z_0 are the propagation coefficient and the intrinsic impedance for a vacuum, respectively. Similarly, if we apply the curl operator on the left and right sides of (4.95b), the wave equation for the magnetic field is obtained:

$$\nabla\times(\epsilon_r^{-1}\nabla\times\mathbf{H}) + \gamma_0^2\mu_r\mathbf{H} + \gamma_0 Y_0\mathbf{M}_i + \nabla\times(\epsilon_r^{-1}\mathbf{J}_i) = 0 \tag{4.98}$$

In a similar way we can derive a wave equation for bianisotropic media.

In general, the electric (or magnetic) wave equation, if considered alone, does not provide a unique solution for the electric (magnetic) field. In addition, an equation for the divergence of the electric (magnetic) flux density should be defined. If the wave equation is valid only inside a bounded region, certain boundary conditions should also be taken into account. Particularly, if a wave equation is solved by using the FEM, the additional divergence equation can be avoided [77].

However, the FEM is not suitable for open problems (i.e., the problems that treat electromagnetic fields in an infinite unbounded region). One possible

way to overcome this problem is to combine the FEM with the BIEs. The equivalent currents, which enter into the BIE, are related to the electric and magnetic fields that enter into the FEM, by the following boundary conditions:

$$\mathbf{J}_s^{(\mathrm{BIE})} = -\mathbf{n} \times \mathbf{H}^{(\mathrm{FEM})} \qquad \mathbf{M}_s^{(\mathrm{BIE})} = \mathbf{n} \times \mathbf{E}^{(\mathrm{FEM})} \tag{4.99}$$

where $\mathbf{n}$ is unit normal going outside the FEM region. The surface and volume basis functions used for the approximation of the equivalent currents and fields should satisfy the above boundary conditions.

Such hybrid methods were first applied to cylinders by Silvester and Hsieh [78] and McDonald and Wexler [79]. A pioneering work on the application of the finite elements methods to the problem of radiation and scattering was done by Volakis and his collaborators, as summarized in [80]. In all these cases, the FEM was upgraded, since it was not suitable for open problems.

The FEM is also not suitable for analyses of thin metallic plates and wires. This is one of the reasons that in this book we adopt the BIEs as the basic equations for the analysis of composite metallic and dielectric structures. However, there is also the need to upgrade the BIEs by the FEM. Namely, the BIEs cannot be applied to anisotropic or bianisotropic objects. In addition, highly inhomogeneous structures can be analyzed more efficiently by the FEM [64].

4.4.3 MoM and Green's-Function Techniques

Consider a current-source element, $\mathbf{J}dv$, near an arbitrary body placed in a homogenous and isotropic medium (see Figure 4.25). This current element induces surface or volume currents in the body. The current-source element and its field, $d\mathbf{E}(\mathbf{r})$, are linearly related. This relation can be expressed in matrix form as

$$[d\mathbf{E}(\mathbf{r})] = [g(\mathbf{r}, \mathbf{r}')][\mathbf{J}(\mathbf{r}')\, dv] \tag{4.100}$$

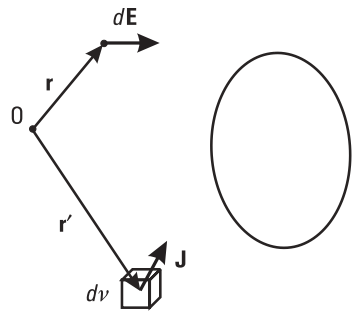

Figure 4.25 A current element near a body.

where $[d\mathbf{E}(\mathbf{r})]$ and $[\mathbf{J}(\mathbf{r}')\,dv]$ are column matrices containing the components of the vectors $d\mathbf{E}(\mathbf{r})$ and $\mathbf{J}(\mathbf{r}')\,dv$, and $[g(\mathbf{r}, \mathbf{r}')]$ is a square matrix representing the Green's function. The Green's function $[g(\mathbf{r}, \mathbf{r}')]$ depends on the size and shape of the body and the electrical properties of material from which the body is made.

In some special cases, the Green's function can be obtained in a closed (analytical) form. An example is a current element located in a free space. The Green's function also has a relatively simple form when the currents induced on a PEC body can be replaced by a finite number of images of the current-source element. Examples are a current-source element above an infinite PEC plane or in a corner. In all other cases, even a simple PEC sphere, the evaluation of the Green's function is much more complicated.

Composite material structures often include bodies whose scattered field can be taken into account through the Green's function; for example, a satellite on which a monopole antenna is mounted can be modeled by a PEC sphere. The analysis of this structure can be performed by a pure MoM (i.e., a MoM that is not combined with Green's-function techniques). In that case, the largest number of unknown coefficients is wasted for approximation of currents over the sphere. The number of unknowns can be drastically reduced if the MoM is combined with the Green's-function technique, as suggested by Tesche and Neureuther [81]. In that case, the MoM equations are imposed only for the wire. Each current element of the wire induces currents on the PEC sphere. The field along the wire due to these induced currents is included through the Green's function of a current element in the presence of a PEC sphere.

For example, consider a monopole antenna of height $h = 0.25\lambda$ mounted on a sphere of radius $a = h$. If the pure MoM is used for the analysis, the sphere is modeled by $M = 96$ bilinear patches (Figure 4.26, inset). The first-order approximation is used for the currents over each patch (rooftop approximation), and the second-order approximation is used for the current along the monopole antenna, resulting in a total of $N = 205$ unknowns. Out of this number, only two unknowns are required for the wire. Figure 4.26 shows the relative power radiation pattern by the antenna versus the angle θ, which is measured from the wire axis. A very good agreement between the results obtained by the pure MoM and those obtained by the hybrid method [81] is observed.

The hybrid method proposed by Tesche and Neureuther is very powerful if a single PEC sphere is combined with simple structures. If the number of unknowns required by the PEC sphere does not represent the major number used in the pure MoM analysis, the MoM can be more efficient than the hybrid method. Namely, the evaluation of the Green's function that includes the presence of the sphere is a few times longer than the evaluation of the free-space Green's function. Note that the Green's function used by Tesche and

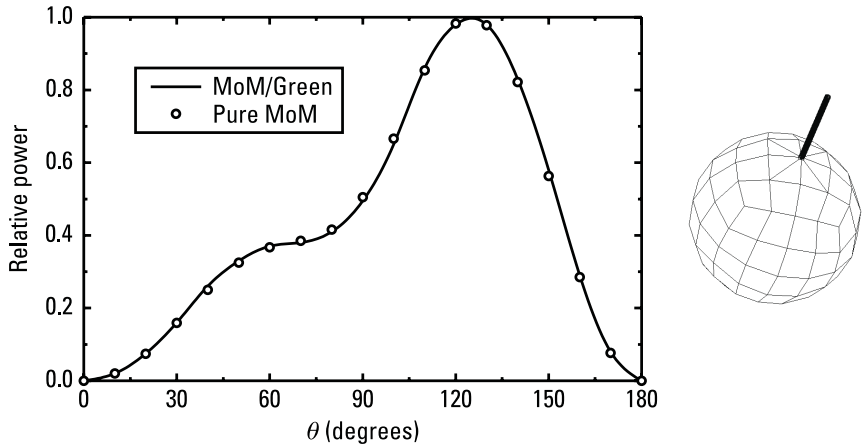

Figure 4.26 Relative power radiation pattern of a monopole antenna mounted on a PEC sphere, shown in the inset, versus the angle θ.

Neureuther is valid only for a whole sphere. Hence, their hybrid technique cannot be used for bodies that include parts of sphere (e.g., a rocket with a hemispherical cap). For a problem involving several spheres, the MoM equations for only one of them can be avoided. Generally, the implementation of the Green's-function technique into a pure MoM code is a tedious job. As a result, the efficiency of the pure MoM code is improved only in a limited number of cases.

However, there are two important situations when Green's-function techniques cannot easily be bypassed by a pure MoM code: antennas in the presence of a real ground plane (soil), and multilayered structures (e.g., a dielectric substrate for microwave circuits or patch antennas). The real ground is usually modeled as lossy homogeneous isotropic and linear half-space. The derivation of the Green's function for the current element above such space is a difficult task. The first solution was derived by Sommerfeld in 1909 [82], and his work was followed by many authors. The expressions that are very often used today were derived by Banos [83]. In the last two decades, this Green's-function technique has been often combined with the MoM. For example, a hybrid method for arbitrarily oriented thin-wire antennas over an imperfect ground was given by Sarkar [84]. Multilayered structures became interesting relatively recently with the development of microstrip antennas. A general formulation for microstrip structures was proposed by Mosig and Gardiol [85]. A PEC body of an arbitrary shape immersed in an arbitrary multilayered structure was considered by Michalski and Zheng [86, 87].

4.4.4 MoM and Asymptotic High-Frequency Techniques

In principle, the MoM can be applied to the solution of arbitrary composite metallic and dielectric structures (with linear media). In practice, the application of the MoM is limited by memory and CPU time requirements of the problem considered. By increasing the electric size and complexity of the structure, the number of unknowns needed for the analysis rapidly increases. Hence, the application of the MoM for given computer resources is limited by the electrical size and complexity of the structure. Under some conditions, the analysis of large objects can be successfully performed using asymptotic high-frequency techniques [88].

Most high-frequency techniques are classified as ray-based techniques and current-based techniques. In the ray-based techniques, the electromagnetic waves are treated as optical rays. The basic technique is the *geometrical optics* (GO) [89]. For large bodies, bounded by smooth surfaces, the GO provides a field solution in the form of reflected and transmitted rays. For example, when a ray comes to a flat boundary surface between two regions, it is reflected and transmitted according to the Snell law. Incident rays are assumed to interact only with a part of the object that is visible from the source point. Hence, the GO rays do not exist in the shadow region behind the object (i.e., in the region where the rays emanating from the source are blocked by the obstacle). In addition, the GO does not take into account discontinuities of the obstacle surface (e.g., edges and vertices).

These shortcomings of the GO are overcome by a systematic inclusion of Keller's diffracted rays, resulting in the *geometrical theory of diffraction* (GTD) [89, 90]. The diffracted rays are excited, for example, by an incident ray which strikes a discontinuity of the object surface, or when it grazes a smooth convex surface. In its original form, the GTD is a purely ray-optical technique, and it does not provide a total field that is continuous across shadow boundaries. This problem was solved by the uniform GTD (or UTD) [91]. For all three techniques (GO, GTD, and UTD), the reflected, transmitted, and diffracted rays become incident rays if they strike the body again. Hence, the solution for complex and large bodies is obtained after multiple reflections, transmissions, and diffractions. The ray-based methods are suitable for the analysis of electrically large objects, particularly of relatively simple shapes.

Consider a case when electrically small and large objects are combined [Figure 4.27(a)]. (An example is a dipole antenna placed near the airplane surface.) The analysis of this composite structure can be performed by imposing the MoM equations for the small body, where the influence of the large object is taken into account through the Green's function for current element in the

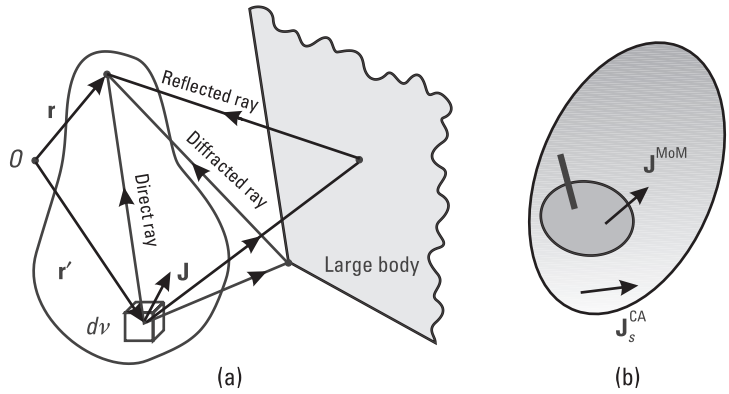

Figure 4.27 A hybrid method involving the MoM and (a) ray-based techniques (GO, GTD, and UTD) and (b) current-based techniques (PO, Fock, and fringe currents).

presence of the large body. The part of the Green's function due to the large body is evaluated by the ray tracing. Such a hybrid method was first applied by Thiele and Newhouse in 1975 [92]. Thereafter, in a short time, many papers were published, summarized in detail by Burnside and Pathak [93]. However, the ray-tracing techniques can readily handle a limited menu of typical discontinuities. Any new specific discontinuity [e.g., the junction of a monopole to a metallic surface shown in Figure 4.27(b)] must be treated separately (e.g., using the MoM). However, when a combination of ray-based methods and the MoM is applied, it is necessary to provide a continuous current flow on the body surface, which is not easily achievable.

The current-based techniques take into account the scattering of electromagnetic waves through currents induced on the body surface. The basic technique is PO [89], which provides a field solution for large objects bounded by smooth surfaces. For example, for a PEC body, the incident field at the visible surface is locally treated as a plane wave, while at the nonvisible surface the incident field is neglected. The surface currents are evaluated as

$$\mathbf{J}_s^{PO} = \begin{cases} 2\mathbf{n} \times \mathbf{H}_{\text{inc}} & \text{illuminated region} \\ 0 & \text{shadowed region} \end{cases} \tag{4.101}$$

where $\mathbf{H}_{\text{inc}}$ is the incident magnetic field. This solution can be further improved if the above currents are replaced by Fock currents, which represent a much more accurate solution for the diffraction of the plane wave by smooth convex bodies [94]. The Fock currents extend into the shadow region (where the PO currents are assumed to be zero), and they show a continuous field behavior across the shadow boundary. When compared with the ray-based techniques,

the extensive ray tracing is avoided, but evaluation of fields involves double integration.

Consider a large object with a specific discontinuity region that cannot be handled accurately enough by the Fock currents [Figure 4.27(b)]. The analysis of such an object can be performed by imposing the MoM equations for the discontinuity region (e.g., the region near a monopole antenna mounted on the body). The total currents over the body surface can be written as

$$\mathbf{J}_s = \mathbf{J}_s^{\text{MoM}} + \mathbf{J}_s^{\text{CA}} \tag{4.102}$$

when $\mathbf{J}_s^{\text{MoM}}$ represents unknown currents in the discontinuity region that should be determined, and $\mathbf{J}_s^{\text{CA}}$ represents currents over the complete body surface that are determined by a current-based asymptotic method. Such a hybrid method was first applied by Medgyesi-Mitschang and Wang [95].

However, bodies of interest usually are not smooth. The diffraction of waves from edges results in additional fringe currents that are not included in the PO or Fock currents. For example, consider a large square PEC scatterer (Figure 4.9) of side $a = 6\lambda$, excited normally by a z-polarized plane wave of magnitude $E = 1$ V/m. The analysis is performed using the Galerkin method to solve the EFIE for the unknown current distribution. Figure 4.28 shows

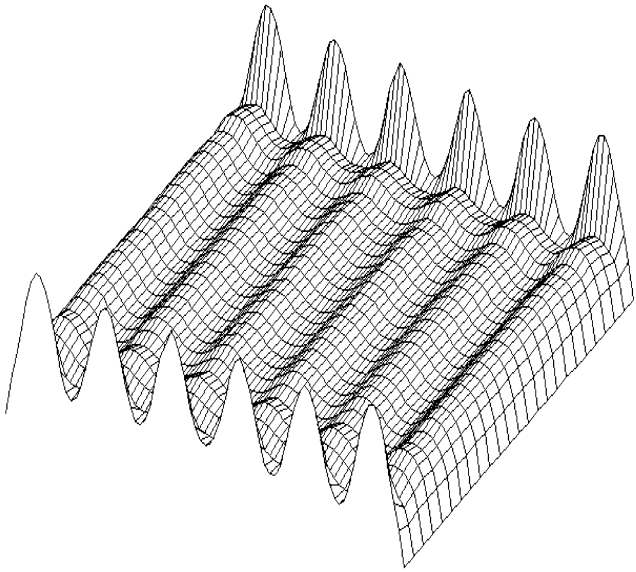

Figure 4.28 The z-component of the current over a square scatterer of side $a = 6\lambda$, shown in Figure 4.9.

the z-component of the current over the scatterer surface. According to the PO solution, this current component is constant over the scatterer surface. It is seen that a precise MoM solution oscillates around the PO solution. In addition to the PO currents, there are resonant currents caused by reflections from the opposite plate edges. These currents are known as fringe currents. Although the generation of these currents is attributed to the edges, they exist all over the surface.

To exploit advantages of the current-based asymptotic techniques fully, the fringe currents should be included in an asymptotic solution, as shown by Jakobus and Landstorfer [96]. They include correction terms in (4.102), which they derived from the exact solution of the half-plane scattering problem. Thus, an efficient method was obtained requiring a very small number of unknowns compared with the pure MoM solution.

4.5 Summary

Most often, a composite metallic and dielectric structure can be regarded as a multiple-region problem where each region is filled by a homogeneous, isotropic, and linear medium. The field at any point can be calculated in terms of surface electric currents on metallic surfaces and equivalent electric and magnetic currents on dielectric surfaces. These currents can be determined by solving integral equations based on various kinds of boundary conditions, collectively referred to as the BIEs.

For PEC structures probably the optimal choice is the EFIE, which is based on the boundary condition for the component of electric field vector tangential to the PEC surface. This formulation can be used for thin wires, plates, and solid bodies. The major shortcoming of this equation is spurious fields in the solution for frequencies near internal resonances of solid bodies. However, by using sophisticated basis functions, the frequency range in which such spurious fields occur becomes so narrow that it can be hardly detected.

For dielectric bodies probably the optimal choice is the pair of PMCHW equations, which are based on the boundary conditions for the components of the electric and magnetic fields tangential to the dielectric surface. For PEC surfaces, the PMCHW equations reduce to the EFIE. Hence, the PMCHW formulation with the EFIE as its degenerate form represents the optimal base for the analysis of composite metallic and dielectric structures.

The power of the EFIE formulation can be further increased by combining it with the surface CIEs, resulting in a model of PEC plates with distributed loadings. Such a model enables an efficient analysis of dielectric radomes and

metallic plates with skin-effect losses. Degenerate forms of the PEC plates with distributed loadings are wires with distributed and lumped loadings.

The PMCHW formulation cannot be used if the composite structure contains anisotropic and bianisotropic parts. In that case the PMCHW should be hybridized with the volume CIE. Note that the original form of the PMCHW cannot be hybridized with the FEM.

Finally, the efficiency of the PMCHW formulation can be further improved if it is hybridized with Green's-function techniques for antennas above a real ground or a multilayered substrate and high-frequency techniques for electrically large structures.

References

[1] Maue, A. W., "Toward Formulation of a General Diffraction Problem via an Integral Equation," *J. Phys.,* Vol. 126, No. 7, 1949, pp. 601–618.

[2] Pocklington, H. E., "Electrical Oscillations in Wires," *Cambridge Phil. Soc. Proc.,* Vol. 9, 1897, pp. 324–332.

[3] Hallen, E., "Theoretical Investigation into the Transmitting and Receiving Qualities of Antennae," *Nova Acta Soc. Sci. Upsal.,* 1938, pp. 1–44.

[4] Andreasen, M. G., "Scattering from Bodies of Revolution," *IEEE Trans. Antennas Propagat.,* Vol. AP-13, No. 2, March 1965, pp. 303–310.

[5] Oshiro, F. K., and K. M. Mitzner, "Digital Computer Solution of Three-Dimensional Scattering Problem," *Digest of IEEE Int. Antennas Propagat. Symp.,* October 1967, Ann Arbor, MI, pp. 257–263.

[6] Mie, G., "Beitrage zur Optik Truber Medien, Speziell Koloider Metallosungen," *Ann. Phys.,* Vol. 25, 1908, p. 377.

[7] Popovic, B. D., and B. M. Kolundzija, "Approximate Analysis of Electromagnetic Scatterers by Optimisation of Equivalent Sources," *Proc. 4th Bosnian Hercegovian Symp. Informatics,* Jahorina, Yugoslavia, 1981 (in Serbian).

[8] Leviatan, Y., A. Boag, and A. Boag, "Generalized Formulations for Electromagnetic Scattering from Perfectly Conducting and Homogeneous Material Bodies-Theory and Numerical Solution," *IEEE Trans. Antennas Propagat.,* Vol. AP-36, No. 12, December 1988, pp. 1722–1734.

[9] Smythe, W. R., "Charged Right Conducting Cylinder," *J. Appl. Phys.,* Vol. 27, 1956, pp. 917–920.

[10] Waterman, P. C., "Matrix Formulation of Electromagnetic Scattering," *Proc. IEEE,* Vol. 53, 1965, pp. 805–812.

[11] Popovic, B. D., M. B. Dragovic, and A. R. Djordjevic, *Analysis and Synthesis of Wire Antennas,* Chichester, U.K.: Research Studies Press, 1982.

[12] Kolundzija, B., J. Ognjanovic, and T. Sarkar, "On the Limit of WIPL-D Code," *Proc. 18th Annual Review of Progress in Applied Computational Electromagnetics, ACES Conf.,* Monterey, CA, 2002, pp. 615–622.

[13] Monson, J. C., and N. J. Damaskos, "A Scheme for Eliminating Internal Resonances: The Parasitic Body Technique," *IEEE Trans. Antennas Propagat.,* Vol. AP-42, No. 8, November 1982, pp. 1089–1096.

[14] Sarkar, T. K., and S. M. Rao, "A Simple Technique for Solving E-Field Integral Equations for Conducting Bodies at Internal Resonances," *IEEE Trans. Antennas Propagat.,* Vol. AP-30, No. 6, November 1982, pp. 1250–1254.

[15] Mittra, R., and C. A. Klein, "Stability and Convergence of Moment Solutions," in *Numerical and Asymptotic Techniques in Electromagnetics,* R. Mittra, (ed.), New York: Springer-Verlag, 1975.

[16] Medgyesi-Mitschang, L. N., and D. S. Wang, "Hybrid Solutions at Internal Resonances," *IEEE Trans. Antennas Propagat.,* Vol. AP-33, 1985, pp. 671–674.

[17] Yaghjian, A. D., "Augmented Electric- and Magnetic-Field Integral Equations," *Radio Sci.,* Vol. 16, 1981, pp. 987–1001.

[18] Yaghjian, A. D., and R. V. McGahan, "Broadside Radar Cross Section of the Perfectly Conducting Cube," *IEEE Trans. Antennas Propagat.,* Vol. AP-33, No. 3, 1985, pp. 321–329.

[19] Poggio, A. J., and E. K. Miller, "Integral Equation Solutions of Three-Dimensional Scattering Problems," in *Computer Techniques for Electromagnetics,* R. Mittra, (ed.), Oxford, U.K.: Pergamon Press, 1973.

[20] Mautz, J. R., and R. F. Harrington, "H-Field, E-Field and Combined-Field Solutions for Conducting Bodies of Revolution," *Arch. Elek. Ubertragung,* Vol. 32, No. 4, 1978, pp. 157–164.

[21] Kolundzija, B. M., "Generalized Combined Field Integral Equation," *Digest IEEE AP-S Int. Symp.,* Vol. 2, July 1996, pp. 852–855.

[22] Brakhage, H., and P. Werner, "Uber das Dirichlestsche Aussenraum Problem fur die Helmholtzsche Schwingungsgleichung," *Archiv. Math.,* Vol. 16, 1965, pp. 325–329.

[23] Greenspan, D., and P. Werner, "A Numerical Method for the Exterior Dirichlet Problem for the Reduced Wave Equation," *Arch. Ration. Mech. Anal.,* Vol. 23, 1966, pp. 288–316.

[24] Bolomey, J., and W. Tabbara, "Numerical Aspects on Coupling Between Complementary Boundary Value Problems," *IEEE Trans. Antennas Propagat.,* Vol. AP-21, No. 3, 1973, pp. 356–363.

[25] Mautz, J. R., and R. F. Harrington, "A Combined-Source Solution for Radiation and Scattering from a Perfectly Conducting Body," *IEEE Trans. Antennas Propagat.,* Vol. AP-27, No. 4, 1979, pp. 445–454.

[26] Newman, E. H., and M. R. Schrote, "On the Current Distribution for Open Surfaces," *IEEE Trans. Antennas Propagat.,* Vol. AP-31, No. 3, May 1983, pp. 515–518.

[27] Rumsey, V. H., "Reaction Concept in Electromagnetic Theory," *Phys. Rev.,* Vol. 94, June 1954, pp. 1483–1491.

[28] Schelkunoff, S. A., *Advanced Antenna Theory,* New York: John Wiley and Sons, 1952.

[29] Dragovic, M. B., "On Synthesis of Cylindrical Antenna with Concentrated Loadings," D. Sc. Thesis, University of Belgrade, Department of Electrical Engineering, 1979 (in Serbian).

[30] King, R. W. P., *The Theory of Linear Antennas,* Boston, MA: Harvard University Press, 1956.

[31] Mei, K. K., "On the Integral Equations of Thin Wire Antennas," *IEEE Trans. Antennas Propagat.,* Vol. AP-13, No. 3, May 1965, pp. 374–378.

[32] Taylor, C. D., "Electromagnetic Scattering from Arbitrary Configuration of Wires," *IEEE Trans. Antennas Propagat.,* Vol. AP-17, No. 5, September 1969, pp. 662–663.

[33] Kominami, M., and K. Rokushima, "A Simplified Integral Equation for Thin-Wire Antennas or Scatterers," *Trans. IECE Japan,* Vol. 59-B, No. 5, May 1976, pp. 267–273.

[34] Popovic, B. D., *CAD of Wire Antennas and Related Radiating Structures,* Taunton, MA: Research Studies Press, Ltd., 1991.

[35] Kolundzija, B. M., "Electromagnetic Modeling of Wire to Plate Structures," D. Sc. Thesis, University of Belgrade, Department of Electrical Engineering, 1990 (in Serbian).

[36] Wilton, D. R., and C. M. Butler, "Efficient Numerical Techniques for Solving Pocklington's Equation and Their Relationships to Other Methods," *IEEE Trans. Antennas Propagat.,* Vol. AP-24, No. 1, January 1976, pp. 83–86.

[37] Kolundzija, B. M., and B. D. Popovic, "Entire-Domain Galerkin Method for Analysis of Generalized Wire Antennas and Scatterers," *Proc. IEE H,* Vol. 139, No. 1, February 1992, pp. 17–24.

[38] Mitzner, K. M., "An Integral Equation Approach to Scattering from a Body of Finite Conductivity," *Radio Sci.,* Vol. 2, 1967, p. 1459.

[39] Jones, D. S., *Methods in Electromagnetic Wave Propagation,* Oxford, U.K.: Clarendon, 1979.

[40] Rogers, J. R., *Numerical Solutions to Ill-Posed and Well-Posed Impedance Boundary Condition Integral Equations,* Lincoln Lab. (MIT), Tech. Rep. 641, November 1983.

[41] Medgyesi-Mitschang, L. N., and J. M. Putnam, "Integral Equation Formulations for Imperfectly Conducting Scatterers," *IEEE Trans. Antennas Propagat.,* Vol. AP-33, No. 2, February 1985, pp. 206–214.

[42] Kishk, A., and L. Shafai, "Different Formulations for Numerical Solution of Single Multibodies of Revolution with Mixed Boundary Conditions," *IEEE Trans. Antennas Propagat.,* Vol. AP-34, May 1986, pp. 666–673.

[43] Sheng, X. -Q., et al., "On the Formulation of Hybrid Finite-Element and Boundary-Integral Methods for 3-D Scattering," *IEEE Trans. Antennas Propagat.,* Vol. AP-46, No. 3, March 1998, pp. 303–311.

[44] Mautz, J. R., and R. F. Harrington, "Electromagnetic Scattering from a Homogenous Material Body of Revolution," *Arch. Elek. Ubertragung,* Vol. 33, 1979, pp. 71–80.

[45] Muller, C., *Foundations of the Mathematical Theory of Electromagnetic Waves,* Berlin, Germany: Springer-Verlag, 1969.

[46] Solodukhov, V. V., and E. N. Vasil'ev, "Diffraction of a Plane Electromagnetic Wave by a Dielectric Cylinder of Arbitrary Cross Section," *Soviet Physics—Tech. Physics,* Vol. 15, 1970, pp. 32–36.

[47] Morita, N., "Analysis of Scattering by a Dielectric Rectangular Cylinder by Means of Integral Equation Formulation," *Electron. Commun. Japan,* Vol. 57-B, October 1974, pp. 72–80.

[48] Vasil'ev, E. N., and L. B. Materikova, "Excitation of Dielectric Bodies of Revolution," *Soviet Physics—Tech. Physics,* Vol. 10, 1966, pp. 1401–1406.

[49] Chang, Y., and R. F. Harrington, "A Surface Formulation for Characteristic Modes of Material Bodies," *IEEE Trans. Antennas Propagat.,* Vol. AP-25, No. 6, November 1977, pp. 789–795.

[50] Wu, T. K., and L. L. Tsai, "Scattering from Arbitrarily-Shaped Lossy Dielectric Bodies of Revolution," *Radio Sci.,* Vol. 12, No. 5, 1977, pp. 709–718.

[51] Kolundzija, B. M., "Electromagnetic Modeling of Composite Metallic and Dielectric Structures," *IEEE Trans. Microwave Theory Tech.,* Vol. 47, No. 7, July 1999, pp. 1021–1032.

[52] Umashankar, K., A. Taflove, and S. M. Rao, "Electromagnetic Scattering by Arbitrary Shaped Three-Dimensional Homogenous Lossy Dielectric Objects," *IEEE Trans. Antennas Propagat.,* Vol. AP-34, June 1986, pp. 758–766.

[53] Rao, S. M., et al., "Electromagnetic Scattering from Arbitrary Shaped Conducting Bodies Coated with Lossy Materials of Arbitrary Thickness," *IEEE Trans. Antennas Propagat.,* Vol. AP-39, May 1991, pp. 627–631.

[54] Sarkar, T. K., E. Arvas, and S. Ponnapalli, "Electromagnetic Scattering from Dielectric Bodies," *IEEE Trans. Antennas Propagat.,* Vol. AP-37, No. 5, May 1989, pp. 673–676.

[55] Rao, S. M., et al., "Electromagnetic Radiation and Scattering from Finite Conducting and Dielectric Structures: Surface/Surface Formulation," *IEEE Trans. Antennas Propagat.,* Vol. AP-39, No. 7, July 1991, pp. 1034–1037.

[56] Medgyesi-Mitschang, L. N., and C. Eftimiu, "Scattering from Axisymmetric Obstacles Embedded in Axisymmetric Dielectrics: The MoM Solution," *Appl. Phys.,* Vol. 19, 1979, pp. 275–285.

[57] Huddleston, P. L., L. N. Medgyesi-Mitschang, and J. M. Putnam, "Combined Field Integral Equation Formulation for Scattering by Dielectrically Coated Conducting Bodies," *IEEE Trans. Antennas Propagat.,* Vol. AP-34, April 1986, pp. 510–520.

[58] Putnam, J. M., and L. N. Medgyesi-Mitschang, "Combined Field Integral Equation for Inhomogeneous Two- and Three-Dimensional Bodies: The Junction Problem," *IEEE Trans. Antennas Propagat.,* Vol. AP-39, May 1991, pp. 667–672.

[59] Richmond, J. H., "Scattering by a Dielectric Cylinder of Arbitrary Cross Section Shape," *IEEE Trans. Antennas Propagat.,* Vol. AP-13, March 1965, pp. 334–341.

[60] Notaros, B. M., "Numerical Analysis of Dielectric Bodies of Arbitrary Inhomogeneity in the Electromagnetic Field," Ph.D. Thesis, University of Belgrade, 1995 (in Serbian).

[61] Livesay, D. E., and K. M. Chen, "Electromagnetic Fields Induced Inside Arbitrarily Shaped Biological Bodies," *IEEE Trans. Microwave Theory Tech.,* Vol. MTT-22, December 1974, pp. 1273–1280.

[62] Newman, E. H., and P. Tulyathan, "Wire Antennas in the Presence of a Dielectric/Ferrite Inhomogeneity," *IEEE Trans. Antennas Propagat.,* Vol. AP-26, No. 4, July 1978, pp. 587–593.

[63] Schaubert, H., D. R. Wilton, and A. W. Glisson, "A Tetrahedral Modeling Method for Electromagnetic Scattering by Arbitrarily Shaped Inhomogeneous Dielectric Bodies," *IEEE Trans. Antennas Propagat.,* Vol. AP-32, January 1984, pp. 77–85.

[64] Kolundzija, B. M., and V. V. Petrovic, "Comparison of MoM/SIE, MoM/VIE and FEM Based on Topological Analysis of Two Canonical Problems," *Digest IEEE AP-S Int. Symp.,* Vol. 1, June 1998, pp. 274–277.

[65] Catedra, M. F., E. Gago, and L. Nuno, "A Numerical Scheme to Obtain the RCS of Three-Dimensional Bodies of Resonant Size Using the Conjugate Gradient Method and the Fast Fourier Transform," *IEEE Trans. Antennas Propagat.,* Vol. AP-37, No. 5, May, 1989, pp. 528–537.

[66] Sarkar, T. K., and E. Arvas, "An Integral Equation Approach to the Analysis of Finite Microstrip Antennas: Volume/Surface Formulation," *IEEE Trans. Antennas Propagat.,* Vol. AP-38, No. 3, March 1990, pp. 305–312.

[67] Lang, K. C., "Babinet's Principle for a Perfectly Conducting Screen with Aperture Covered by Resistive Sheet," *IEEE Trans. Antennas Propagat.,* Vol. AP-21, No. 5, September 1973, pp. 738–740.

[68] Senior, T. B. A., "Some Extensions of Babinet's Principle in Electromagnetic Theory," *IEEE Trans. Antennas Propagat.,* Vol. AP-25, No. 3, May 1977, pp. 417–420.

[69] Shen, L. C., "An Experimental Study of Imperfectly Conducting Dipoles," *IEEE Trans. Antennas Propagat.,* Vol. AP-15, No. 5, 1967, pp. 606–611.

[70] Rao, B. L. J., J. E. Ferris, and W. E. Zimmerman, "Broadband Characteristics of Cylindrical Antennas with Exponentially Tapered Capacitive Loading," *IEEE Trans. Antennas Propagat.,* Vol. AP-17, No. 5, 1969, pp. 145–151.

[71] Altshuler, E. E., "The Traveling-Wave Linear Antenna," *IRE Trans.,* Vol. AP-9, July 1961, pp. 324–329.

[72] Nyquist, D. P., and K. M. Chen, "The Traveling Wave Linear Antenna with Nondissipative Loading," *IEEE Trans. Antennas Propagat.,* Vol. AP-16, No. 1, 1968, pp. 21–31.

[73] Albertsen, N. C., J. E. Hansen, and N. E. Jensen, "Computation of Radiation from Wire Antennas on Conducting Bodies," *IEEE Trans. Antennas Propagat.,* Vol. AP-22, No. 2, March 1974, pp. 200–206.

[74] Silvester, P. P., and R. L. Ferrari, *Finite Elements for Electrical Engineers,* 3rd ed., New York: Cambridge University Press, 1996.

[75] Jin, J., *The Finite Element Method in Electromagnetics,* New York: John Wiley and Sons, 1993.

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

[77] Salazar-Palma, M., et al., *Iterative and Self-Adaptive Finite-Elements in Electromagnetic Modeling,* Norwood, MA: Artech House, 1998.

[78] Silvester, P., and S. Hsieh, "Finite-Element Solution of 2-Dimensional Exterior Field Problems," *Proc. IEE,* Vol. 118, December 1971, pp. 1743–1747.

[79] McDonald, B. H., and A. Vexler, "Finite-Element Solution of Unbounded Field Problems," *IEEE Trans. Microwave Theory Tech.,* Vol. MTT-20, December 1972, pp. 841–847.

[80] Jin, J.-M., J. L. Volakis, and J. D. Collins, "A Finite-Element–Boundary-Integral Method for Scattering and Radiation by Two- and Three-Dimensional Structures," *IEEE Antennas Propagat. Magazine,* Vol. 33, 1991, pp. 22–32.

[81] Tesche, F. M., and A. R. Neureuther, "Radiation Patterns for Two Monopoles on a Perfectly Conducting Sphere," *IEEE Trans. Antennas Propagat.,* Vol. AP-18, No. 5, September 1970, pp. 692–694.

[82] Sommerfeld, A. N., "Uber die Ausbreitung der Wellen in der Drahtlosen Telegraphie," *Ann. Physic,* Vol. 28, 1909, pp. 665–736.

[83] Banos, A., *Dipole Radiation in the Presence of a Conducting Half-Space,* New York: Pergamon Press, 1966.

[84] Sarkar, T. K., "Analysis of Arbitrarily Oriented Thin Wire Antennas over a Plane Imperfect Ground," *Arch. Elec. Ubertrag.,* Vol. 31, 1977, pp. 449–457.

[85] Mosig, J. R., and F. E. Gardiol, "General Integral Equation Formulation for Microstrip Antennas and Scatterers," *Proc. IEEH,* Vol. 132, No. 7, December 1985, pp. 424–432.

[86] Michalski, K. A., and D. Zheng, "Electromagnetic Scattering and Radiation by Surfaces of Arbitrary Shape in Layered Media, Part I: Theory," *IEEE Trans. Antennas Propagat.,* Vol. AP-38, No. 3, March 1990, pp. 335–344.

[87] Michalski, K. A., and D. Zheng, "Electromagnetic Scattering and Radiation by Surfaces of Arbitrary Shape in Layered Media, Part II: Results for Continuous Half-Spaces," *IEEE Trans. Antennas Propagat.,* Vol. AP-38, No. 3, March 1990, pp. 345–352.

[88] Kouyoumjian, R. G., "Asymptotic High Frequency Methods," *Proc. IEEE,* Vol. 53, August 1965, pp. 864–876.

[89] James, G. L., *Geometrical Theory of Diffraction for Electromagnetic Waves,* Stevenage, England: Peter Peregrinus Ltd., 1976.

[90] Keller, J. B., "Geometrical Theory of Diffraction," *J. Opt. Soc. Amer.,* Vol. 52, February 1962, pp. 116–130.

[91] Kouyoumjian, R. G., and P. H. Pathak, "A Uniform Geometrical Theory of Diffraction for an Edge in a Perfectly Conducting Surface," *Proc. IEEE,* Vol. 62, 1975, pp. 1448–1461.

[92] Thiele, G. A., and T. H. Newhouse, "A Hybrid Technique for Combining Moment Methods with the Geometrical Theory of Diffraction," *IEEE Trans. Antennas Propagat.,* Vol. AP-23, No. 1, 1975, pp. 62–69.

[93] Burnside, W. D., and P. H. Pathak, "A Summary of Hybrid Solutions Involving Moment Methods and GTD," in *Applications of the MoM to Electromagnetic Fields,*" St. Cloud, MN: SCEEE Press, 1980.

[94] Fock, V. A., "The Distribution of Currents Induced by a Plane Wave on the Surface of a Conductor," *J. Phys. USSR,* Vol. 10, 1946, pp. 130–136.

[95] Medgyesi-Mitschang, L. N., and D.-S. Wang, "Hybrid Solutions for Scattering from Perfectly Conducting Bodies of Revolution," *IEEE Trans. Antennas Propagat.,* Vol. AP-31, No. 4, July 1983, pp. 570–583.

[96] Jakobus, U., and F. M. Landstorfer, "Improved PO-MM Hybrid Formulation for Scattering from Three-Dimensional Perfectly Conducting Bodies of Arbitrary Shape," *IEEE Trans. Antennas Propagat.,* Vol. AP-43, No. 2, February 1995, pp. 162–169.

5

Geometrical Modeling

The task of geometrical modeling[1] consists of representing a composite metallic and dielectric structure by a proper combination of building elements (linear segments, surface patches, and volume bricks) for which approximating functions can be easily defined. Such modeling requires the definition of at least two coordinate systems. The first, the global coordinate system, serves for the definition of the structure as a whole, for the description of the near-field points, far-field points, and so forth. The second is a local coordinate system (or possibly systems) associated with building elements and is used for the description of the corresponding unknown quantity.

The geometrical modeling can be exact or approximate. The exact modeling has the purpose of exactly representing the geometry of an analyzed structure. However, such representation requires the development of various types of building elements for various structures. Approximate modeling implies designing a sufficiently accurate representation of a given geometry using a limited menu of building elements. Hence, approximate modeling is more convenient for the general purpose. However, exact modeling is still very important. Very often it represents an intermediate step in approximate modeling. Finally, the analysis based on exact modeling serves as check of accuracy of the analysis based on approximate modeling.

In practically all available numerical techniques for electromagnetic modeling, the linear, surface, and volume building elements can be treated as particular cases of generalized wires, quadrilaterals (or triangles), and hexahedrons (or tetrahedrons), which can be used for either exact or approximate

1. Mathematical description (modeling) of the structure geometry is called geometrical modeling.

modeling of composite structures. Most often, the geometrical modeling by these building elements can be performed independently for wires, metallic and dielectric surfaces, and dielectric volumes. Hence, the first three sections of this chapter are devoted to these three basic types of building elements, respectively. However, in the case of wire-to-plate junctions and protrusions of wires through dielectric surfaces, the geometrical modeling of wires and metallic and dielectric surfaces is strongly dependent on the particular junction in question. Hence, treatment of these junctions and protrusions requires particular attention. Section 5.4 elaborates two general approaches (attachment modes and localized junctions) to this problem.

Efficiency of the electromagnetic modeling includes the effort required to describe the problem in a form understandable to the computer. As an aid to the user of the computer code, it is necessary to develop algorithms for automatic modeling of such structures (Section 5.5). In particular, it is desirable that electrically large elements be automatically subdivided into sufficiently small elements (Section 5.6).

5.1 Wire Structures

Most authors model wires by cylinders [1–7]. In many cases, such a simple building element enables (almost) exact modeling of practical structures (e.g., arrays of thin dipole antennas, Yagi-Uda arrays, and log-periodic arrays [3–7]). However, such a building element shows deficiencies in the analysis of fat wires, wire ends, abrupt and continuous changes of wire radii, curved wires, and so forth [4, 6, 7]. In all these cases we need more sophisticated building elements.

5.1.1 Generalized Wires

A generalized wire is a metallic body in the form of a curvilinear cylinder of circular cross sections of variable radius to which the thin-wire approximation is applied [8, 9]. We assume that the axis of the generalized wire is a continuous and smooth curve and that the wire radius changes continuously and smoothly along the axis. The geometry of the generalized wire is completely determined by parametric equations of its axis and radius:

$$\mathbf{r}_a = \mathbf{r}_a(s) \qquad a = a(s) \qquad -1 \leq s \leq 1 \tag{5.1}$$

where s is a parametric coordinate along the wire axis, and $s = -1$ and $s = 1$ are coordinates of the wire beginning and the wire end, respectively (Figure 5.1).

Sometimes, a model of the generalized wire can be applied to wires that are not thin or even to structures that are usually not considered to be wires.

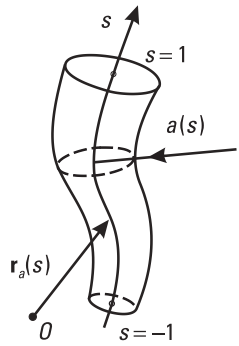

Figure 5.1 Definition of generalized wire.

For example, we can use generalized wires to model electrically thick antennas of revolution [10, 11], regions of axially symmetrical wire ends and abrupt changes in the wire diameter [12], or a frill with radial electric surface currents used for the treatment of wire-to-plate junctions [13]. The generalized wire can also be used for analysis of a wire of noncircular cross section. The distribution of the total current along such a wire does not change significantly if an equivalent circular cross section is used instead of the original one. The equivalent radius is determined from the requirement that the quasistatic per-unit-length energy in the vicinity of the wire is equal in the original and equivalent cases [14]. (For example, for an infinitely thin strip of width w, the equivalent radius is known to be $a = w/4$.)

Wire segments are most often straight and have a constant radius. Circular and helical wires belong to another important class. For these two and many other classes of wires, there exists an exact form of (5.1). (For example, Yeh and Mei used exact equations for the axis and radius of conical equiangular spiral antennas [15].) However, the evaluation of potential and field integrals for each new class of wires requires a new implementation of (5.1). On the other hand, very good results for currents along wires can be obtained even if geometry of these wires is not modeled very precisely. Moreover, the main part of the analysis error can be due to neglecting the circular current component and the circumferential dependence of the axial component. Hence, it is suitable to adopt (5.1) in a general form that can enable an efficient approximation of arbitrary wire geometries (e.g., in the form of splines).

5.1.2 Approximation of Wires by Spline Curves

The approximation of the generalized wire consists of the approximation of the parametric equations for its radius and axis. These equations can be approximated by a polynomial in the standard manner:

$$\mathbf{r}_a(s) = \sum_{i=0}^{n} \mathbf{B}_i s^i \qquad a(s) = \sum_{i=0}^{n} b_i s^i \tag{5.2}$$

where n is order of approximation, and the scalar coefficients, b_i, and vector coefficients, $\mathbf{B}_i$, can be determined in various ways. The simplest way to determine these coefficients is to equate these polynomials with known values of the position vectors of the wire axis and radius at $n + 1$ interpolation points, resulting in

$$\mathbf{r}_a(s) = \sum_{i=1}^{n+1} L_i^n(s)\mathbf{r}_i \qquad a(s) = \sum_{i=1}^{n+1} L_i^n(s) a_i \qquad L_i^n(s) = \prod_{k=1, k\neq i}^{n+1} \frac{s - s_k}{s_i - s_k} \tag{5.3a-c}$$

where $L_i^n(s)$ is known as the Lagrange interpolation polynomial of the nth order. Note that $L_i^n(s)$ has unit value at the ith interpolation point, while it is equal to zero at all other interpolation points [16].

Many other interpolation polynomials are available. Some take into account not only tabulated values of the function being approximated, but also tabulated values of the first- and higher-order derivatives of this function. Others are characterized by specific distributions of points used for the interpolation. In all instances, the goal is to obtain the best approximation (in some sense) of the original curve. However, all interpolation polynomials show undesirable oscillations if higher-order approximations are used. To avoid this, the segments are usually subdivided into a set of shorter subsegments, each described by a separate low-order ($n \leq 3$) interpolation polynomial called a spline.

The Lagrange spline, with the first tabulated point at the beginning and the last at the end of the interpolation segment, can provide continuity of the wire local axis and its radius. Better modeling of the geometry can be obtained by using other splines, such as the B-spline, or beta-spline, used frequently in computer graphics [17]. These splines not only provide continuity of the wire axis and its radius, but also eliminate the possibility of a broken (although continuous) line approximating the wire axis. However, this improvement in modeling of geometry is usually not necessary.

As indicated in the introduction to Section 5.1, wire structures are very often assembled from cylinders, represented by the Lagrange splines of the lowest possible degree ($n = 0$ for the radius and $n = 1$ for the axis). When the Lagrange splines of degree $n = 1$ are used for both the radius and the axis, the right truncated cone is obtained. This type of element was rarely used in the past [8, 11, 12]. However, it is shown that a right truncated cone enables in some cases a significantly more accurate modeling of generalized wires than

a cylinder [9]. Except for a small additional analytical effort, the determination of current distribution along truncated cones does not require longer computing time than that required for cylindrical wires. On the other hand, the approximation of the generalized wires by higher-order splines increases considerably both the analytical effort and the computing time for the same number of unknowns (e.g., for quadratic splines [18]). Therefore, right truncated cones seem to be the optimal basic elements for the modeling of geometry and the analysis of structures assembled from generalized wires.

5.1.3 Right Truncated Cones

A right truncated cone is determined by the position vectors and radii of its beginning and its end, $\mathbf{r}_1$ and a_1, and $\mathbf{r}_2$ and a_2, respectively (Figure 5.2). Starting from (5.3), the parametric equations of the cone axis and its local radius can be written in two forms:

$$\mathbf{r}_a(s) = \frac{1-s}{2}\mathbf{r}_1 + \frac{1+s}{2}\mathbf{r}_2 \qquad a(s) = \frac{1-s}{2}a_1 + \frac{1+s}{2}a_2 \qquad -1 \le s \le 1 \tag{5.4}$$

$$\mathbf{r}_a(s) = \mathbf{r}_c + \mathbf{r}_s s \qquad a(s) = a_c + a_s s \qquad -1 \le s \le 1 \tag{5.5}$$

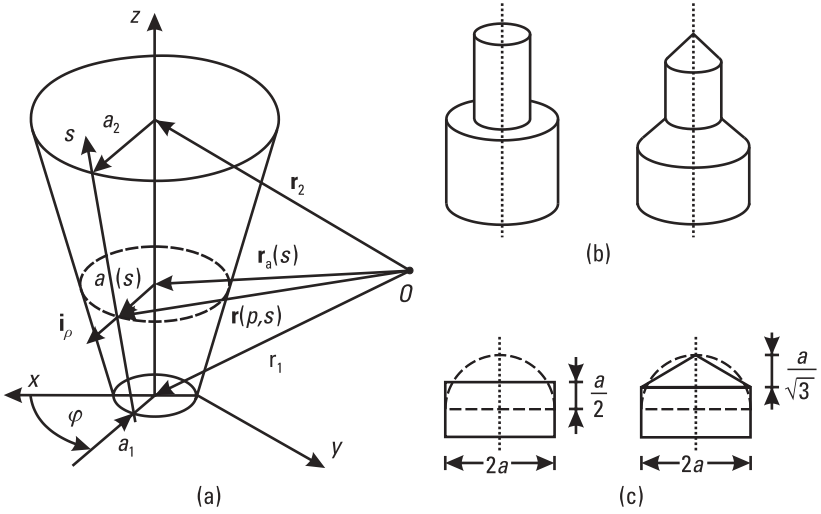

Figure 5.2 A right truncated cone: (a) general form; (b) flat (frill-like) change of wire radius and flat end, and conical change of wire radius and conical end (feed); and (c) flat and conical end equivalent to a hemispherical end.

The s-axis is moved from the cone axis to a cone generatrix, so that the surface current flowing along a truncated cone has only the component along the generatrix.

As special cases, the truncated cone degenerates into a right cylinder ($a_1 = a_2$), an ordinary cone ($a_2 = 0$), a flat disc ($a_2 = 0$, $\mathbf{r}_1 = \mathbf{r}_2$), and a frill ($\mathbf{r}_1 = \mathbf{r}_2$). The right truncated cone and its degenerate forms can be used for modeling of cylindrical wires with flat (frill-like) or conical changes of the wire radius, as well as of flat and conical wire ends and feeds, as indicated in Figure 5.2(b).

The treatment of wire ends and feeds can have a significant influence on the solution of the current distribution along wires [7, 12]. This effect is particularly pronounced for relatively thick wires (approximately, $h/a < 30$, where h is the wire length and a is its radius) of approximately resonant lengths or wires representing a part of a resonant structure. A typical example is a relatively thick resonant vertical monopole antenna above a perfectly conducting ground plane base-driven by a TEM magnetic current frill. The monopole antenna and its symmetrical equivalent are shown in the inset of Figure 5.3.

The antenna was analyzed using two geometrical models. In the first case, the flat end was not taken into account, while in the second case, it was modeled by a flat disc. The dipole current was approximated by basis functions that satisfy quasistatic relation at the coaxial line opening (see Table 6.1) in addition to current continuity at wire ends and junctions. The current distribution along

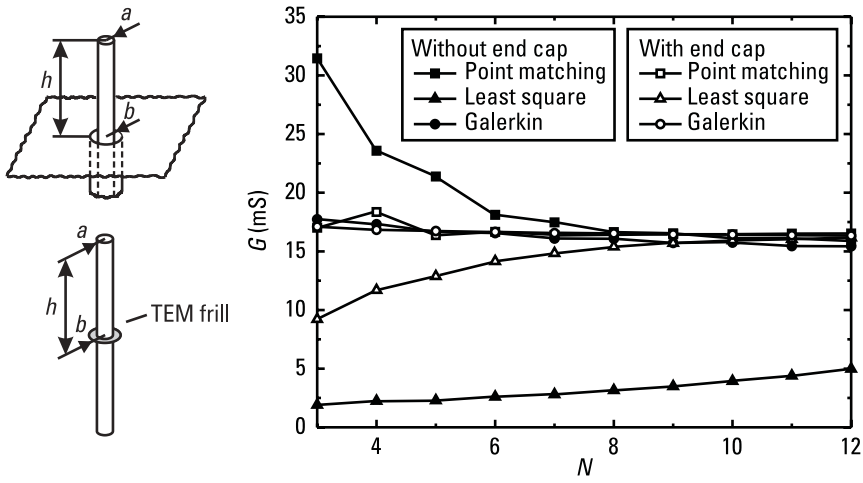

Figure 5.3 Conductance of a quarter-wavelength monopole antenna ($h/a = 25$, $b/a = 2.3$) against the degree N of the polynomial approximation along the cylindrical monopole part. The monopole and its symmetrical equivalent are sketched in the inset. (After: [7].)

the monopole was adopted as a polynomial of degree N and along the radius of the flat end as a second-degree polynomial. The solution of the EFIE was performed by three testing procedures: the point-matching procedure, the least-squares technique, and the Galerkin method. Figure 5.3 shows the conductance of an antenna with $h = \lambda/4$, $h/a = 25$, and $b/a = 2.3$, plotted against the degree N of the polynomial. The accuracy and convergence are significantly improved by taking the end effect into account.

Wire ends of simpler shapes can be substituted by approximately equivalent flat or conical ends [12], as is shown for the hemispherical end cap in Figure 5.2(c). The equivalent end is obtained by slightly changing the length of the straight part of the wire and replacing the original wire end by a flat or conical one. The size of the equivalent end is obtained by requiring that the surfaces of the original and equivalent ends have equal areas. Hence, approximately the same amount of charge is located on both ends, so that in both cases the field along the wire due to these charges is also approximately the same. In addition, for the equivalent conical end, it is required that total length of the wire be unchanged.

This equivalency works even in the case of electrically large wire ends. For example, let us consider a fat rotationally symmetrical antenna with a hemispherical end cap (Figure 5.4, inset). The height of the antenna is $H = 250$ mm, the height of the feeding cone is $h = 50$ mm, and the sphere radius is $a = 50$ mm. The antenna is modeled in three ways (shown in the inset of Figure 5.4): by 80 surface patches (plates), by three degenerate right truncated cones, one of them representing the flat wire end, and by three

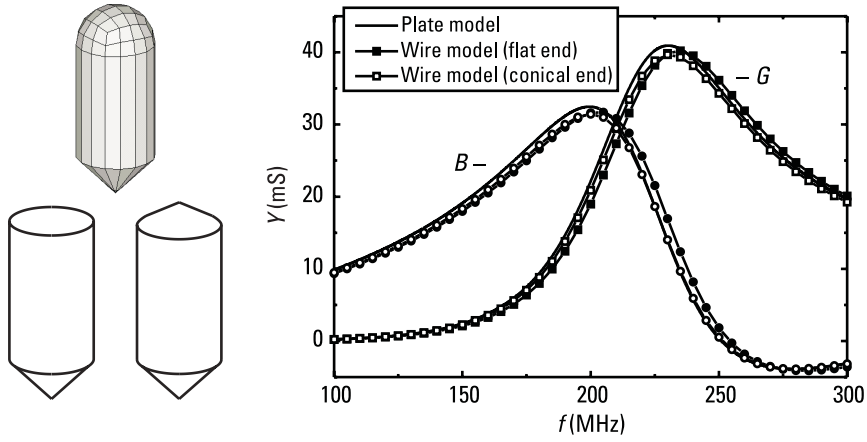

Figure 5.4 Input admittance of a fat, rotationally symmetrical antenna with a hemispherical cap modeled by surface patches, an equivalent wire with an almost flat end, and an equivalent wire with a conical end.

degenerate right truncated cones, one of them representing the conical wire end. The plate model results in 260 unknowns, while each of the wire models results in 6 unknowns. Figure 5.4 shows the input admittance of the antenna in the vicinity of its first resonance. A good agreement among the results is observed.

5.1.4 Piecewise Cylindrical (Conical) Approximation of Wires

A representation of a generalized wire by two or more cylinders (right truncated cones) is referred to as the piecewise-cylindrical (piecewise-conical) approximation. The segmentation of generalized wires is needed in three cases: for wires of variable radii, for curved wires, and for electrically long wires. A general technique for obtaining such an approximation is as follows. A wire structure is considered as a combination of generalized wires. If these generalized wires are not straight enough (or their radii do not change almost linearly along their axes), they are subdivided into smaller generalized wires. The generalized wires are replaced by right truncated cones having the same ending nodes and radii. Finally, the truncated cones, whose electrical lengths are greater than the maximal allowed value, are further subdivided into shorter truncated cones, whose electrical lengths are smaller than maximal allowed value.

For example, let us consider a circular-loop antenna of radius a = 1m made of thin wire of radius r = 1 mm. The circle is modeled by an inscribed n-sided polygon, where n = 3, 4, 5, 6, 8, and 32. The second-order approximation for currents is used along each of these sides. Figure 5.5(a) shows the input admittance of this antenna in the vicinity of the first resonance. By increasing n, the results converge with those for n = 32, which are found to be very accurate.

In general, satisfactory results for the antenna impedance, current distribution, radiation pattern, and so forth, can be obtained only if the deviation of the segmented wire axis from the original axis is small enough. However, the general techniques for obtaining piecewise-cylindrical and conical approximations of generalized wires, explained above, do not provide the minimal deviation of the segmented wire axis. To provide this feature for some geometries, we can implement the concept of equivalent radius of wire curvature.

5.1.5 Equivalent Radius of Wire Curvature

Let us illustrate the concept of the equivalent radius on the example of a circle of radius a (Figure 5.6). The circle is modeled by an n-sided polygon inscribed into a circle of an equivalent radius a_n. One way of determining the equivalent radius is to minimize the maximal deviation of this polygon from the original circle, $d_{n\max}$. This deviation is determined as

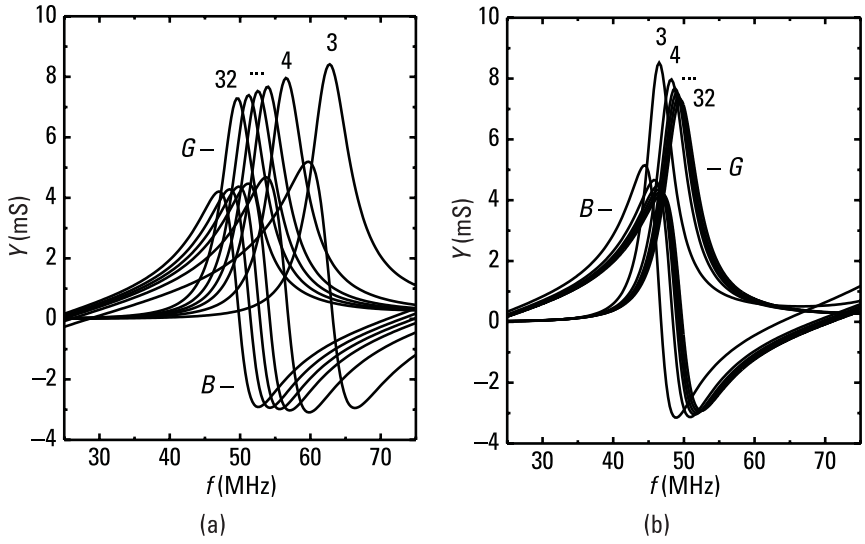

Figure 5.5 The input impedance of a circular-loop antenna of radius a = 1m, made of thin wire of radius r = 1 mm, obtained for various number of segments n of the polygon inscribed in (a) a circle of radius a, and (b) a circle of equivalent radius.

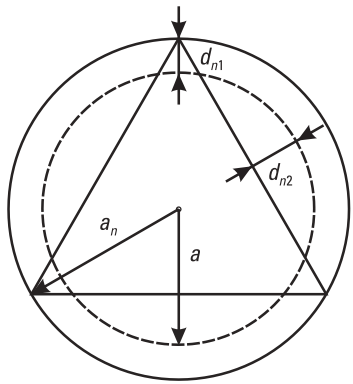

Figure 5.6 The concept of equivalent radius.

$$d_{n\max} = \max(d_{n1}, d_{n2}) \qquad d_{n1} = a_n - a \qquad d_{n2} = a - a_n \cos\phi_n \tag{5.6}$$

where d_{n1} is the smallest distance between a polygon vertex and the circle, d_{n2} is the smallest distance between the midpoint of a polygon side and the circle, and $\phi_n = \pi/n$ is the angle under which a half of the polygon side is seen from the circle center. The maximum deviation $d_{n\max}$ is minimized, requiring that

distances d_{n1} and d_{n2} be equal. In that case, the equivalent radius is obtained in the normalized form as

$$\frac{a_n}{a} = \frac{2}{1 + \cos\phi_n}\left(\approx 1 + \frac{\pi^2}{4n^2}\right) \tag{5.7}$$

where the approximation valid for small angles is given in brackets. For large n the calculated radius tends to the arithmetic mean of the equivalent radii, which correspond to the polygons inscribed and circumscribed with respect to the original circle. Values of equivalent radius for various n are given in Table 5.1.

If the concept of equivalent radius is applied to the loop antenna in the previous section, all curves for $n > 3$ almost coincide [Figure 5.5(b)]. Very accurate results are obtained even with $n = 3$ (the circle is approximated by a triangle). The question is how to choose the number of segments, n, used for modeling the circle, when the electrical size of the circle increases. One possibility is to increase this number in such a manner that the maximal deviation normalized by the wavelength remains constant. Let the maximal deviation be $d_{\max}$. After replacing $d_{n\max} = d_{\max}$ and the small-angle approximation of (5.7) into (5.6), the number of segments needed for modeling the circle with the prescribed maximal deviation $d_{\max}$ can be written as

$$n = n_0\sqrt{\frac{a}{\lambda}} \qquad n_0 = \sqrt{\frac{\pi^2\lambda}{4d_{\max}}} \tag{5.8}$$

where n_0 is the number of segments needed for modeling the circle whose radius equals one wavelength. According to this expression, the number of segments increases as a square root of radius. For example, when $d_{\max} = 1/25\lambda$, the modeling of the circle is performed with $n_0 = 8$ segments per wavelength.

Let us consider a circular-loop antenna of radius $a = 1$m made of thin wire of radius $r = 1$ mm placed in the xOy-plane (Figure 5.7, inset). The antenna is fed by a point generator placed at the x-axis. The circle is modeled

Table 5.1
Equivalent Radius a_n of a Circular Loop of Radius a Modeled by an N-Sided Polygon

n	3	4	5	6	7	8	9	10	11	12
a_n/a	1.333	1.172	1.106	1.072	1.052	1.040	1.031	1.025	1.021	1.017

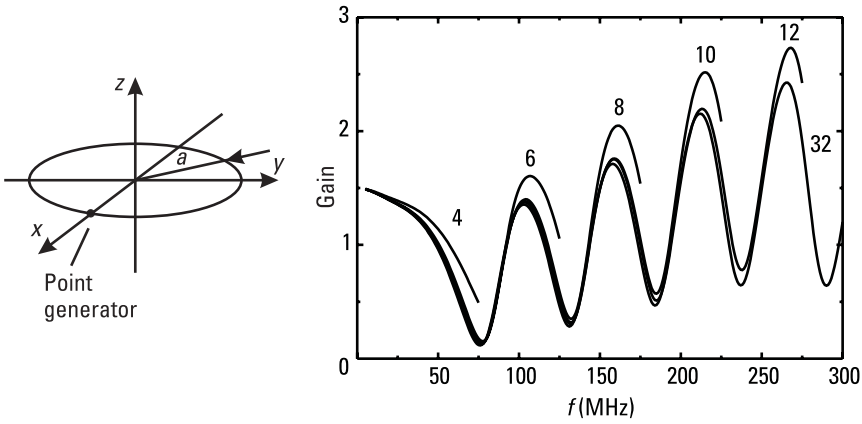

Figure 5.7 Gain of a circular-loop antenna of radius a = 1m, made of a thin wire of radius r = 1 mm, versus frequency obtained for various numbers of segments n of a polygon inscribed into circle of an equivalent radius. (© 2002 ACES [19].)

by an n-sided polygon inscribed in a circle of the optimal equivalent radius, where n = 4, 6, 8, 10, 12, and 32. The second (third)-order approximation for currents is used along each side. Figure 5.7 shows the gain in the direction of the x-axis up to f = 300 MHz. At this frequency, the loop radius is equal to one wavelength. When increasing the frequency, it is necessary to increase the number of polygon sides in order to maintain the accuracy of results, as should be done. However, for higher frequencies this number increases linearly in terms of the electrical length of the loop radius, which is faster than predicted by (5.8). Hence, by increasing the electrical length of a loop, the maximal deviation normalized by the wavelength must be decreased.

Keeping in mind the example above, the rule for choosing the minimum number of segments, n, needed for modeling the circle can be simplified as follows:

1. For circles whose circumference is shorter than one wavelength, n = 4;
2. For each increase of the circumference for up to half wavelength (more precisely, 0.55λ), the number of segments, n, should be increased by one.

The electrical length of one straight wire segment is limited to about 0.5λ, requiring typically four to six unknowns per wavelength. The same number of unknowns is needed in the analysis of linear wire antennas. Hence, the analysis of loops and related structures based on straight wire segments and

the concept of equivalent radius is equally efficient as the analysis based on curved wire segments whenever the segments are shorter than about 0.5λ. If we want to model electrically long circles with wire segments whose lengths are greater than 0.5λ, we must use higher-order splines for geometrical modeling and higher-order approximations for currents. In that case, the efficiency of the analysis increases slightly. For example, if the current is approximated by polynomials of order four to seven along segments 1 to 2λ long, accurate results are obtained with only three to four unknowns per wavelength [9, 20].

5.2 Metallic and Dielectric Surfaces

Many authors model metallic and dielectric surfaces by flat rectangular and triangular patches [13, 21–28]. In some cases, such simple building elements enable exact geometrical modeling (e.g., rectangular and corner scatterers [22] and corner reflectors [24]), so that the analysis based on flat rectangular patches was found to be more efficient than the analysis based on flat triangular patches. On the other hand, flat triangles are convenient for approximate modeling of arbitrary curved structures (e.g., circular disks and spheres [25]), which is not the case with flat quadrilaterals. More sophisticated building elements enable a more accurate geometrical modeling and a more efficient corresponding analysis, but require additional analytical preparation and programming effort. In what follows, all these aspects of the geometrical modeling of surfaces are elaborated.

5.2.1 Generalized Quadrilaterals and Triangles

A generalized quadrilateral (in the wide sense) is a curved curvilinear quadrilateral. A generalized quadrilateral (in the narrow sense) is defined by the parametric equation

$$\mathbf{r} = \mathbf{r}(p, s) \qquad -1 \leq p, s \leq 1 \tag{5.9}$$

where p and s are arbitrary local coordinates associated with the surface, and coordinates of the quadrilateral sides are $p, s = \pm 1$ [Figure 5.8(a)] [8, 29]. This parametric equation performs the mapping of a square of side $a = 2$ in a local ps-coordinate system into a curved quadrilateral in the 3-D global coordinate system. Any side of the generalized quadrilateral requires only one surface-current component (the one not tangential to the side considered) in formulating the current continuity equation.

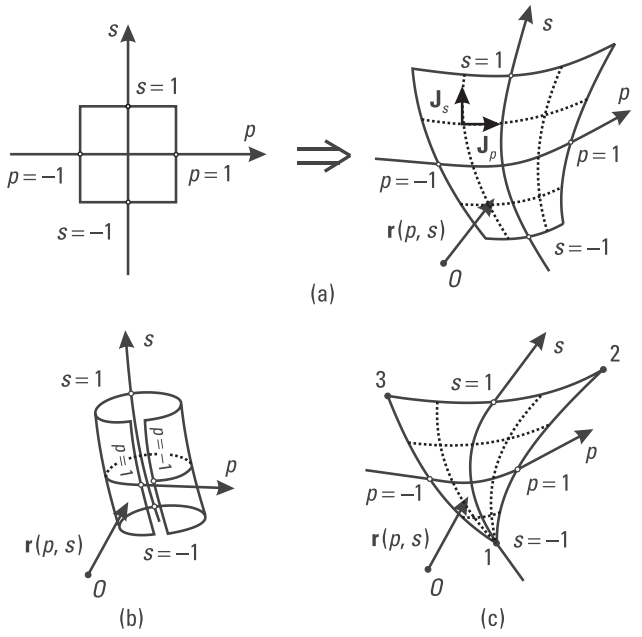

Figure 5.8 (a) Generalized quadrilateral defined in a local *ps*-coordinate system and its degenerate forms: (b) generalized wire and (c) generalized triangle.

The generalized wire can be represented as a degenerate form of the generalized quadrilateral, as illustrated in Figure 5.8(b). The parametric equation of the generalized wire surface is

$$\mathbf{r}(p, s) = \mathbf{r}_a(s) + a(s)\mathbf{i}_\rho(p, s) \qquad -1 \le p, s \le 1 \tag{5.10}$$

where p is related to the circumferential angle, ϕ, as $p = \phi/\pi$, and $\mathbf{i}_\rho(p, s)$ is the unit vector, radial with respect the wire axis. Therefore, any theory that includes the electromagnetic analysis of generalized quadrilaterals can be applied to generalized wires. However, the current distribution along the generalized wires is 1-D (one current component depending on one coordinate only), and the current distribution over the generalized quadrilaterals is 2-D (two current components depending on two coordinates). This fact can be used to make the analysis of the generalized wires formally the same, but significantly simpler and more rapid than the analysis of the generalized quadrilaterals. Therefore, in what follows we shall also consider specific methods for the analysis of the generalized wires.

If two adjacent nodes of a generalized quadrilateral coincide, it degenerates into a curved curvilinear triangle (i.e., a generalized triangle [29]). In what

follows, we shall consider the generalized triangle obtained by merging nodes defined by (−1, −1) and (1, −1) in the *ps*-coordinate system [Figure 5.8(c)]. The surface-current density cannot be represented at merged nodes in terms of its *p*- and *s*-components alone.[2] A remedy is to represent the surface-current density vector over triangles by means of three components. The problem is how to adopt the third component. The solution of the problem lies in the possibility that the same triangle can be parameterized in three alternative local *ps*-coordinate systems. In the coordinate system shown in Figure 5.8(c), all *s*-coordinate lines bunch into node 1 while *p*-coordinate lines are quasiparallel to the opposite side. Alternatively, all *s*-coordinate lines can bunch into node 2 (node 3), while *p*-coordinate lines are again quasiparallel to the corresponding opposite side.

Another possibility is to use simplex coordinates (surface area coordinates). Let us consider the equilateral triangle $\Delta 123$ of unit height shown in Figure 5.9. The simplex coordinates of a point P inside the triangle are defined as the shortest distances of the point from the triangle sides. Two simplex coordinates uniquely determine the position of the point, which means that the third coordinate is dependent on the first two. Having in mind that triangle $\Delta 123$ consists of triangles $\Delta 12P$, $\Delta 23P$, and $\Delta 31P$, and that the surface areas of these triangles are $1/\sqrt{3}$, $t_1/\sqrt{3}$, $t_2/\sqrt{3}$, and $t_3/\sqrt{3}$, respectively, this dependence is obtained in the form

$$t_1 + t_2 + t_3 = 1 \tag{5.11}$$

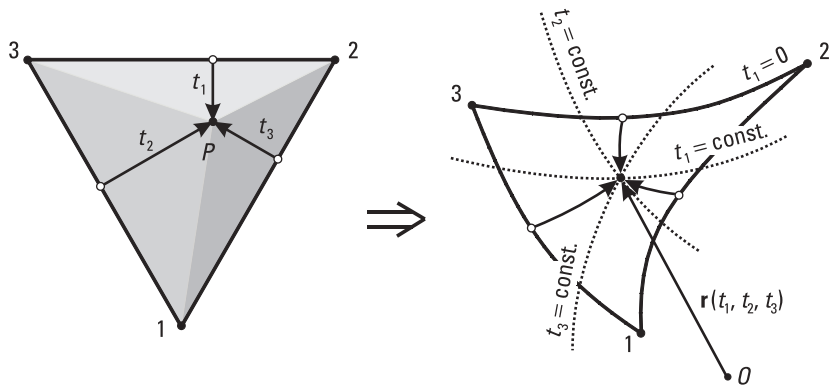

Figure 5.9 Generalized triangle defined in local simplex coordinate system.

2. The coordinate unit vector $\mathbf{i}_s$ is not uniquely defined at points where the *s*-coordinate lines intersect. Hence, the *s*-current component can be only zero. The *p*-current component cannot represent the total current, except in special cases.

Now, the generalized triangle can be defined by the parametric equation

$$\mathbf{r} = \mathbf{r}(t_1, t_2, t_3) \qquad 0 \le t_1, t_2, t_3 \le 1 \tag{5.12}$$

This equation performs mapping of an equilateral triangle in a local simplex coordinate system into a curvilinear triangle in the 3-D global coordinate system. Now, the current can be decomposed into three components (e.g., along lines for which t_1-, t_2-, and t_3-coordinates are constant.

Obviously, a part of the theory valid for the generalized quadrilaterals is not applicable to the generalized triangles. In what follows, the necessary modifications of the general theory to generalized triangles will therefore be discussed.

5.2.2 Unitary Vectors and Related Quantities

Starting from the parametric equation (5.9), all geometrical quantities of interest, such as length elements, unit vectors, surface area elements, and so forth, can be easily obtained. It is suitable to express these quantities in terms of the so-called unitary vectors $\mathbf{a}_p$ and $\mathbf{a}_s$, defined as

$$\mathbf{a}_p = \frac{d\mathbf{r}(p, s)}{dp} \qquad \mathbf{a}_s = \frac{d\mathbf{r}(p, s)}{ds} \tag{5.13a,b}$$

Generally, the p- and s-coordinates are not length coordinates. If they are not, the length elements dl_p and dl_s along the coordinate lines, corresponding to the increments dp and ds of the p- and s-coordinates, respectively, are obtained as [30]

$$dl_p = e_p\, dp \qquad dl_s = e_s\, ds \qquad e_p = |\mathbf{a}_p| \qquad e_s = |\mathbf{a}_s| \tag{5.14}$$

where e_p and e_s are the so-called Lame coefficients. The unit vectors of the local coordinate system at any point of the surface element are obtained as

$$\mathbf{i}_p = \frac{\mathbf{a}_p}{e_p} \qquad \mathbf{i}_s = \frac{\mathbf{a}_s}{e_s} \tag{5.15}$$

Finally, the area of the surface element dS can be expressed as

$$dS = dl_p\, dl_s \sin \alpha_{ps} = |\mathbf{a}_p \times \mathbf{a}_s|\, dp\, ds \tag{5.16}$$

where α_{ps} is the local angle between the p- and s-coordinate lines.

These geometrical quantities can be also determined starting from (5.12) by expressing one simplex coordinate in terms of the other two, which can be considered as the p- and s-coordinates.

5.2.3 Exact Modeling of Surfaces by Generalized Quadrilaterals

Exact modeling of a surface by the generalized quadrilaterals is possible only if the surface can be described exactly by one or several analytical expressions. If these expressions, at the same time, represent the parametric equations of the generalized quadrilaterals, the modeling does not need any further procedure. If they do not, which is usually the case, the parametric equations of the generalized quadrilaterals are obtained by using specific mapping techniques.

For example, let us show how to define generalized quadrilaterals that exactly conform to a sphere of radius a [29]. First, imagine a cube of side $2b$ placed within the sphere so that the centers of the sphere and the cube coincide (Figure 5.10). For all six cube faces (i.e., for each square), the local (p, s) coordinate systems are defined next, with $-1 \leq p, s \leq 1$. For example, for the upper face the parametric equation is

$$\mathbf{r}_b(p, s) = b(\mathbf{i}_x p + \mathbf{i}_y s + \mathbf{i}_z) \tag{5.17}$$

Finally, the cube faces together with the coordinate systems are projected onto the sphere surface, with the common center as the projection center. For example, the parametric equation of the generalized quadrilateral obtained by projecting the upper cube face onto the sphere is

$$\mathbf{r}_a(p, s) = a\frac{\mathbf{r}_b(p, s)}{|\mathbf{r}_b(p, s)|} = a\frac{\mathbf{i}_x p + \mathbf{i}_y s + \mathbf{i}_z}{\sqrt{1 + p^2 + s^2}} \tag{5.18}$$

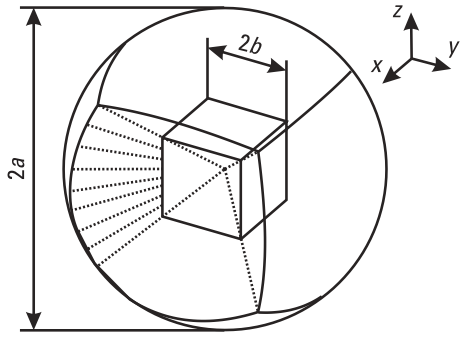

Figure 5.10 Exact modeling of a sphere by a generalized quadrilateral. (After: [29].)

If instead of the cube we start with a tetrahedron, the sphere is modeled by generalized triangles.

Using (5.18) it is not difficult to compute other geometrical quantities of interest (e.g., unitary vectors). Similarly, we can derive parametric equations of generalized quadrilaterals for various geometries (e.g., circles and cylinders) (see Section 5.5). In each particular case, the unitary vectors and the related geometrical quantities, current distributions, field integrals, and so forth, are evaluated in a different way. Hence, making an algorithm for electromagnetic modeling based on exact modeling of a large set of various geometries would be very tedious job. On the other hand, there are many structures that cannot be exactly modeled by this set. As the result, it is desirable to adopt one or two flexible geometrical shapes that can be used for approximate, but accurate enough, modeling of almost any structure.

5.2.4 Approximations of Surfaces by Spline Quadrilaterals

In analogy with the approximate modeling of the generalized wires, the parametric equation of the generalized quadrilateral can be approximated by a double vector polynomial as

$$\mathbf{r}(p, s) = \sum_{i=0}^{n_p} \sum_{j=0}^{n_s} \mathbf{B}_{ij} p^i s^j \tag{5.19}$$

where $\mathbf{B}_{ij}$ are vector coefficients that can be determined in various ways. The simplest way is to start with a mesh of p- and s-coordinate lines, $p = p_i$, $i = 1, \ldots, (n_p + 1)$, and $s = s_j$, $j = 1, \ldots, (n_s + 1)$ [e.g., as shown for $n_p = n_s = 4$ in Figure 5.11(a)]. Then the vector polynomial is equated with the known values of the position vectors at the nodes of the mesh, $\mathbf{r}_{ij}$, which results in the system of linear equations $\mathbf{r}(p_i, s_j) = \mathbf{r}_{ij}$, $i = 1, \ldots, (n_p + 1)$, $j = 1, \ldots, (n_s + 1)$. Values of the position vectors $\mathbf{r}_{ij}$ can be evaluated by using analytical expressions for surfaces, or they can be measured. Finally, the system is solved for the unknown coefficients $\mathbf{B}_{ij}$. The polynomial (5.19) can be written in compact form as

$$\mathbf{r}(p, s) = \sum_{i=1}^{n_p+1} \sum_{j=1}^{n_s+1} L_i^n(p) L_j^n(s) \mathbf{r}_{ij} \tag{5.20}$$

where $L_i^n(p)$ and $L_j^n(s)$ are Lagrange polynomials given by (5.3c). The product $L_i^n(p) L_j^n(s)$ has unit value at the interpolation node whose position vector is $\mathbf{r}_{ij}$, while it is equal to zero at all other interpolation nodes.

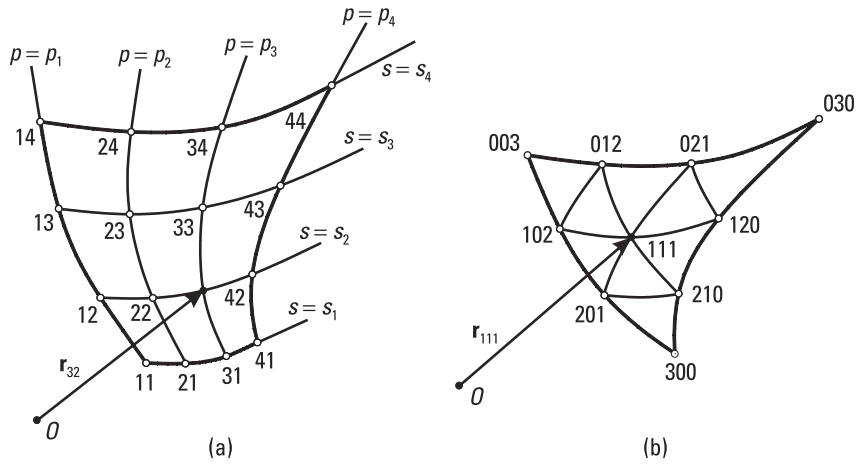

Figure 5.11 Polynomial approximation of (a) a generalized quadrilateral and (b) a generalized triangle.

If the generalized quadrilateral shown in Figure 5.11(a) degenerates into a generalized triangle, so that its side $s = -1$ degenerates into a point, all nodes along this side merge into one node. In order that (5.20) can be applied, all position vectors $\mathbf{r}_{i1}$, $i = 2, \ldots, 5$, should be replaced by the position vector $\mathbf{r}_{11}$. Such a network of nodes does not enable an approximation of the generalized triangle in the other two ps-coordinate systems associated with the other two vertices of the triangle (discussed in Section 5.2.1). An exception is a flat triangle, for which the interpolation is performed using three corner nodes.

Simplex coordinates enable unique interpolation of a generalized triangle for all three current components. We start from the coordinate lines $t_1 = i/n$, $t_2 = j/n$, and $t_3 = k/n$, $i, j, k = 1, \ldots, n$. Coordinate lines, such that $i + j + k = n$, intersect at a single node $\mathbf{r}_{ijk}$ [Figure 5.11(b)]. Using these nodes as interpolation nodes, the generalized triangle is approximately represented as

$$\mathbf{r}(t_1, t_2, t_3) = \sum_{\substack{i,j,k=0 \\ i+j+k=n}}^{n} \mathbf{r}_{ijk} R_i^n(t_1) R_j^n(t_2) R_k^n(t_3) \qquad R_i^n(t_1) = \prod_{k=0}^{i-1} \frac{t_1 - \frac{k}{n}}{\frac{i}{n} - \frac{k}{n}} \tag{5.21a,b}$$

$R_i^n(t_1)$ is a polynomial having unit value for $t_1 = i/n$ and equal to zero at nodes $t_1 = k/n$, $k = 1, \ldots, i-1$. $R_j^n(t_2)$ and $R_k^n(t_3)$ are also given by (5.21b), except that i and t_1 are replaced by j and t_2, and k and t_3, respectively.

In analogy with the choice of splines for modeling of generalized wires, splines given by (5.20) and (5.21) appear to be a good choice for modeling of geometry of generalized quadrilaterals and triangles, respectively [29, 31–33]. Possible orders of these splines are one (bilinear surface and flat triangle), two (biquadratic surfaces), and three (bicubic surfaces). The application of higher-order splines enables more accurate geometrical modeling at the expense of a significant increase of both the analytical effort and computing time. However, higher-order spline approximations for the geometrical modeling of curved surfaces may yield a more efficient analysis only if combined with a higher-order approximation for the current distribution.[3] Therefore, the bilinear surfaces and flat triangles seem to be the optimal basic elements for geometrical modeling and the analysis of structures assembled from surfaces.

5.2.5 Bilinear Surfaces and Flat Triangles

A bilinear surface is, in general, a nonplanar quadrilateral, defined uniquely by four vertices [8, 29]. The parametric equation of such a quadrilateral can be written in alternative forms

$$\mathbf{r}(p, s) = \frac{1}{4} \sum_{i=1}^{2} \sum_{j=1}^{2} \mathbf{r}_{ij} [1 + (-1)^i p] [1 + (-1)^j s] \qquad -1 \le p, s \le 1 \tag{5.22}$$

$$\mathbf{r}(p, s) = \mathbf{r}_c + \mathbf{r}_p p + \mathbf{r}_s s + \mathbf{r}_{ps} ps \qquad -1 \le p, s \le 1 \tag{5.23}$$

where $\mathbf{r}_{ij}$, $i, j = 1, 2$, are position vectors of its vertices [see Figure 5.12(a)], and $\mathbf{r}_c$, $\mathbf{r}_p$, $\mathbf{r}_s$, and $\mathbf{r}_{ps}$ represent linear combinations of $\mathbf{r}_{ij}$.

Starting from (5.23) we can easily compute any other geometrical quantity of interest. For example, unitary vectors are obtained as

$$\mathbf{a}_p = \mathbf{r}_p + \mathbf{r}_{ps} s \qquad \mathbf{a}_s = \mathbf{r}_s + \mathbf{r}_{ps} p \tag{5.24}$$

The unitary vector $\mathbf{a}_p$ and the corresponding unit vector $\mathbf{i}_p$ do not depend on the p-coordinate. Similarly, the unitary vector $\mathbf{a}_s$ and the corresponding unit vector $\mathbf{i}_s$ do not depend on the s-coordinate. Hence, although a bilinear

3. Knepp and Goldhirsh use biquadratic quadrilateral surfaces combined with pulse basis [31]. Wandzura uses biquadratic triangular surfaces combined with rooftop basis functions [32]. Valle et al. use *nonuniform rational B-splines* (NURBS) expanded in terms of Bezier patches combined with basis functions of the rooftop type [33]. For all examples they present, the same accuracy can be achieved if higher-order spline surfaces are replaced by bilinear surfaces or flat triangles.

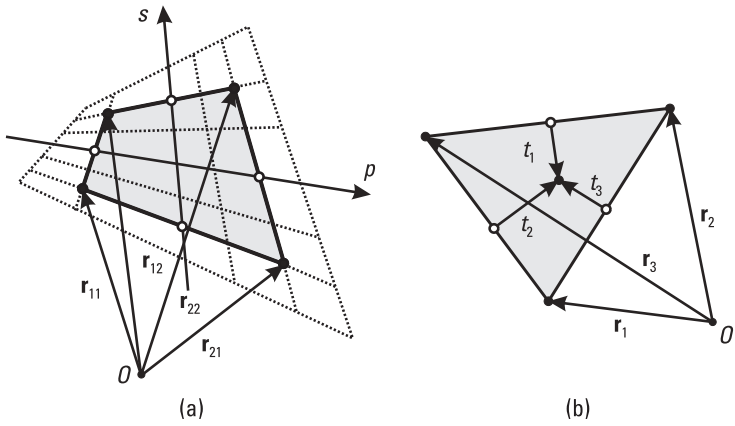

Figure 5.12 (a) Bilinear surface, and (b) flat triangle.

surface is curved in the general case, all the p- and s-coordinate lines are straight lines. This is why such surfaces are termed bilinear. Owing to this property, a bilinear surface cannot be concave or convex, but only inflected (or planar). Therefore it can be used for efficient approximation only of those curved surfaces, or parts of surfaces, which do not have pronounced concave or convex properties.

Depending on the values of the vectors $\mathbf{r}_p$, $\mathbf{r}_s$, and $\mathbf{r}_{ps}$, a bilinear surface takes different degenerate forms. (The vector $\mathbf{r}_c$ is the position vector of the origin of the local ps-coordinate system, and it does not influence the shape of the surface.) Only one of the vectors $\mathbf{r}_p$, $\mathbf{r}_s$, and $\mathbf{r}_{ps}$ can be zero. If $\mathbf{r}_{ps}$ is zero, the bilinear surface degenerates into a rhomboid. If, in addition, $\mathbf{r}_p$ and $\mathbf{r}_s$ are mutually perpendicular, the rhomboid degenerates into a rectangle. As another particular case, if $\mathbf{r}_{ps}$ is nonzero and it is in the plane defined by the vectors $\mathbf{r}_p$ and $\mathbf{r}_s$, the bilinear surface degenerates into a flat quadrilateral.

It can be easily shown that if $\mathbf{r}_p = \pm\mathbf{r}_{ps}$ or $\mathbf{r}_s = \pm\mathbf{r}_{ps}$, the bilinear surface degenerates into a flat triangle. Alternatively, using (5.21) and $\mathbf{r}_{100} = \mathbf{r}_1$, $\mathbf{r}_{010} = \mathbf{r}_2$, and $\mathbf{r}_{001} = \mathbf{r}_3$, a flat triangle can be represented in simplex coordinates as

$$\mathbf{r}(t_1, t_2, t_3) = \mathbf{r}_1 t_1 + \mathbf{r}_2 t_2 + \mathbf{r}_3 t_3 \tag{5.25}$$

where $\mathbf{r}_i$, $i = 1, 2, 3$, are the position vectors of its vertices [Figure 5.12(b)].

Some authors use arbitrary flat quadrilaterals as building elements [34–36]. In contrast to bilinear surfaces, they are not suitable for modeling curved surfaces because they cannot be defined by four arbitrary points in space, unless the points are in one plane. It will be shown in Chapter 8 that determination of

current distribution over bilinear surfaces does not require longer computing time than that for rectangles. The analysis based on bilinear surfaces was found to be at least two times more efficient than the analysis based on flat triangular patches [37]. Hence, we give advantage to bilinear surfaces.

5.2.6 Piecewise Almost-Flat Approximation of Surfaces

Representation of metallic or dielectric surfaces by two or more bilinear surfaces is referred to as the piecewise almost-flat approximation. (A bilinear surface is generally a curved surface. However, in the geometrical modeling it usually resembles a flat surface—hence the term *almost-flat*.) The basic technique for obtaining such an approximation is as follows. Each surface is represented (exactly or approximately) as a combination of generalized quadrilaterals and triangles. If surfaces (edges) of these quadrilaterals and triangles are not sufficiently flat (straight), they are subdivided, possibly along the p- and s- or simplex coordinate lines, into smaller generalized quadrilaterals and triangles. The obtained generalized quadrilaterals and triangles are replaced by bilinear surfaces and flat triangles, respectively, having the same ending nodes. When the continuity of currents should be satisfied at the junction of two adjacent quadrilaterals or triangles, the geometrical modeling must provide that each pair of adjacent quadrilaterals or triangles have a common side with two common nodes.

For example, let us make a piecewise almost-flat approximation of a sphere. First, the sphere is exactly modeled by six generalized quadrilaterals [Figure 5.13(a)]. (The procedure is explained in detail in Section 5.2.3.) Then each such quadrilateral is subdivided along the basic p- and s-coordinate lines into four smaller generalized quadrilaterals [Figure 5.13(b)]. Finally, each smaller quadrilateral is replaced by a bilinear surface [Figure 5.13(c)]. Instead of using $n = 2$ bilinear surfaces along a side of a generalized quadrilateral, we can use $n = 3$ or more bilinear surfaces along this side [Figure 5.13(d–f)] For arbitrary n, the total number of bilinear surfaces is $m = 6n^2$. By increasing n, the number of bilinear surfaces and the corresponding number of unknowns increases dramatically. The question is what minimal n gives sufficiently accurate results.

In general, satisfactory results can be obtained only if the deviation of the approximate surfaces from the original one is small enough. Besides that, it is desirable that the original and the approximate surfaces have equal surface areas, and that the original and the approximate bodies have the same volumes. To provide these properties, the basic technique for obtaining the piecewise almost-flat approximation of surfaces, explained above, is combined with the concept of equivalent radius for surfaces.

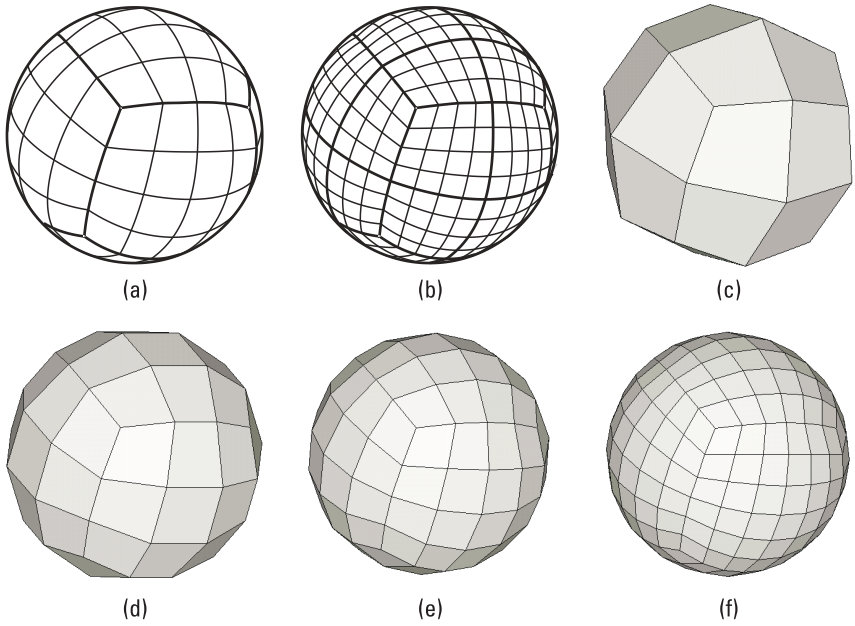

Figure 5.13 Piecewise almost-flat approximation of a sphere. Sphere is modeled by (a) 6 and (b) 24 generalized quadrilaterals; and (c) 24, (d) 54, (e) 96, and (f) 216 bilinear surfaces.

5.2.7 Concept of Equivalent Radius for Surfaces

For flat surfaces that have a curved contour (e.g., a flat circular disc) and surfaces curved in one direction only (e.g., a cylinder of circular cross section), the equivalent radius is introduced in the same way as the equivalent radius for a circular loop (Section 5.1.5).

In what follows, let us consider surfaces curved in two directions (i.e., at each point, two different radii of curvature can be defined). The concept of the equivalent radii for such surfaces will be illustrated using a sphere of radius a. The sphere is modeled by $6n^2$ bilinear surfaces as explained in the Section 5.2.6. However, the model is not inscribed into the original sphere of radius a, but into the equivalent sphere of radius a_n. The equivalent radius a_n is determined from the condition that the maximal deviation of the approximate surface with respect to the original surface is minimal. The equivalent radii can be approximately expressed as

$$a_n \cong a\left(1 + \frac{\pi^2}{32n^2}\right) \tag{5.26}$$

Values for different n are listed in Table 5.2.

For example, let us consider scattering from a metallic sphere of radius a = 1m. The sphere is modeled using n = 2, 4, 6, and 8, which results in m = 24, 96, 216, and 396 bilinear surfaces, and the concept of equivalent radius is applied. Figure 5.14 shows the RCS normalized by λ^2 versus frequency up to f = 600 MHz. Results for n = 6 practically coincide with the exact values obtained in the form of the Mie series [38]. All results almost coincide up to about f = 150 MHz. After f = 150 MHz, results for n = 2 quickly diverge, while the other results almost coincide up to about f = 330 MHz. After f = 330 MHz, the results for n = 4 quickly diverge, and so forth.

The example above gives the rule for choosing the minimal number of segments n needed for an accurate analysis:

1. For spheres whose radius is smaller than half wavelength, n = 2;
2. For each quarter-wavelength increment of the radius, n should be increased by one.

Table 5.2
Equivalent Radius a_n of a Sphere of Radius a Modeled by $6n^2$ Bilinear Surfaces

n	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
a_n/a	1.3084	1.0771	1.0343	1.0193	1.0123	1.0086	1.0063	1.0048	1.0038	1.0031

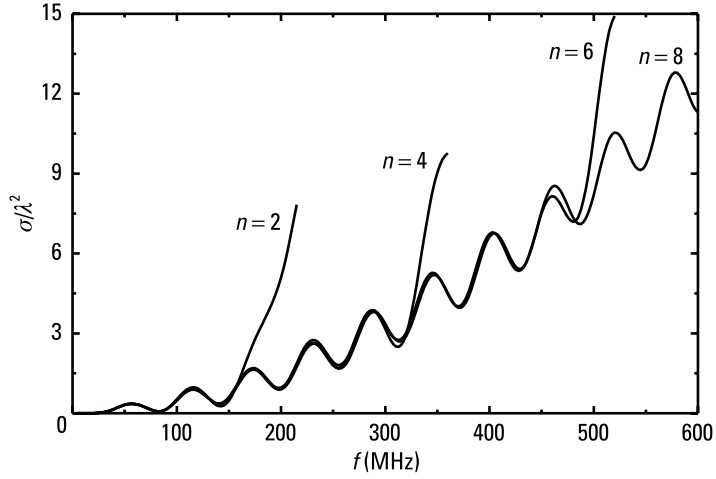

Figure 5.14 Monostatic RCS of a spherical scatterer (a = 1m) normalized by λ^2 versus frequency. The geometrical model of the sphere consists of m = 24, 96, 216, and 396. The model is inscribed into the sphere of equivalent radius. (© 2002 ACES [19].)

The electrical length of one side of the bilinear surface is limited to 0.4λ, so that the second-order approximation for currents must be used, resulting in 48 unknowns per wavelength squared. The same number of unknowns is needed in the analysis of flat-plate scatterers made of bilinear surfaces of the same size. Hence, the analysis of spheres and related structures based on the bilinear surfaces and the concept of equivalent radius is equally efficient as the analysis based on curved surfaces, whenever the sides of the surfaces are shorter than 0.4λ. If we want to model electrically large spheres by surface patches whose sides are longer than 0.4λ, we must use higher-order splines for the geometrical modeling and higher-order approximations for the currents. This can increase the efficiency of the analysis. For example, if the currents are approximated by polynomials of order four to seven for patch sides $1–2\lambda$ long, accurate results are obtained with only about 24 to 32 unknowns per wavelength squared.

5.3 Dielectric Volumes

If the analysis of dielectric structures is performed by using the MoM to solve VIEs or the FEM to solve differential equations (see Chapter 4), the unknown equivalent sources (charges and currents) or unknown fields are distributed throughout volumes of dielectric bodies or surrounding media. To model these volumes, the first published papers used simple building elements, like cubes and tetrahedrons. For example, when the analysis is based on VIEs, Livesay and Chen used cubes [39], while Schaubert et al. introduced tetrahedrons [40]. Later, Notaros et al. applied trilinear volumes (hexahedrons) [41]. Approximately at the same time, these elements were used for the analysis (at low frequencies) performed by the FEM. For example, Barton and Cendes performed modeling by tetrahedrons [42], van Wellij introduced trilinear volumes [43], and Crowley et al. proposed curvilinear hexahedrons [44].

Volume building elements can be easily defined by a generalization of surface building elements given in Section 5.2. A generalized hexahedron is a curvilinear hexahedron defined by the parametric equation

$$\mathbf{r} = \mathbf{r}(p, q, s) \qquad -1 \leq p, q, s \leq 1 \tag{5.27}$$

where p, q, and s are local parametric coordinates. Note that its boundary faces coincide with the local coordinate planes $p, q, s = \pm 1$. Consequently, any of these faces requires only one volume current (field) component (the one not tangential to the face considered) in formulating the continuity condition for the normal component.

Exact modeling of structures by the generalized hexahedrons is possible in a limited number of cases, when these structures can be described exactly

by analytical expressions. For example, by mapping a cube volume to a sphere volume, the sphere is modeled by one hexahedron only. In analogy with the approximate modeling of generalized wires and quadrilaterals, the parametric equation for the generalized hexahedron is approximated by a triple vector spline of the Lagrange type. For the spline of order one along each dimension, we obtain a trilinear volume, which represents a curved hexahedron defined uniquely by eight arbitrarily placed vertices. The parametric equation for a trilinear volume is

$$\mathbf{r}(p, q, s) = \frac{1}{8} \sum_{i=1}^{2} \sum_{j=1}^{2} \sum_{k=1}^{2} \mathbf{r}_{ijk} [1 - (-1)^i p] [1 - (-1)^j q] [1 - (-1)^k s] \quad (5.28)$$

$$-1 \leq p, q, s \leq 1$$

where $\mathbf{r}_{ijk}$, $i, j, k = 1, 2$, are the position vectors of its vertices, as sketched in Figure 5.15(a).

Special cases of trilinear volumes are a cube, regular parallelepiped, oblique parallelepiped, prism of quadrilateral cross section, prism of triangular cross section, pyramid, tetrahedron, and so forth. In last three cases, the current (field) cannot be decomposed into three components along p-, q-, and s-coordinate lines. For example, consider the case when trilinear volume degenerates into tetrahedron [Figure 5.15(b)]. At the point ($p = -1$, $q = -1$, $s = -1$), the s-coordinate lines intersect, so that the unit vector $\mathbf{i}_s$ is not uniquely defined there. The p- and q-components of the volume current cannot represent the total current, except in special cases. A remedy is to represent the volume

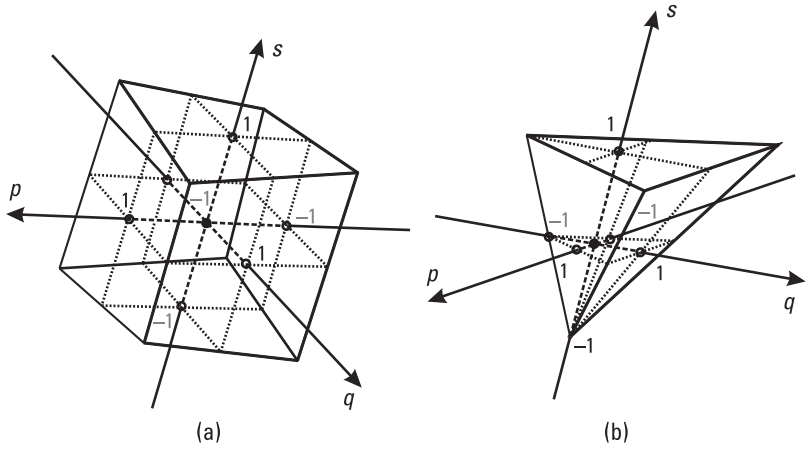

Figure 5.15 Parametric volume elements: (a) trilinear volume and (b) tetrahedron.

current density vector within tetrahedrons by means of four components. The problem is how to adopt the fourth component. One possibility is to parameterize the same tetrahedron in four local pqs-coordinate systems. In the coordinate system shown in Figure 5.15(b), all s-coordinate lines bunch into the node ($p = -1$, $q = -1$, $s = -1$), while the pq-coordinate surfaces are quasiparallel to the opposite face. Alternatively, all s-coordinate lines can bunch into one of the other three nodes, while the pq-coordinate planes are again quasiparallel to the opposite face.

Another possibility is to use simplex coordinates. Let us imagine an equilateral tetrahedron ABCD of unit height. The simplex coordinates of a point P inside the tetrahedron are defined as the shortest distances of the point from the tetrahedron faces. Three simplex coordinates uniquely determine the position of the point, which means that the fourth coordinate is dependent on the first three. The tetrahedron ABCD consists of four tetrahedrons, ABCP, ABDP, BCDP, and CADP. The volume of these tetrahedrons are $1/\sqrt{12}$, $t_1/\sqrt{12}$, $t_2/\sqrt{12}$, $t_3/\sqrt{12}$, and $t_4/\sqrt{12}$, respectively, this dependence is obtained in the form

$$t_1 + t_2 + t_3 + t_4 = 1 \tag{5.29}$$

Now, the generalized tetrahedron (curvilinear tetrahedron) can be defined by the parametric equation

$$\mathbf{r} = \mathbf{r}(t_1, t_2, t_3, t_4) \qquad 0 \le t_1, t_2, t_3, t_4 \le 1 \tag{5.30}$$

This equation performs mapping of the equilateral tetrahedron in the local simplex coordinate system into a curvilinear tetrahedron in the 3-D global coordinate system. By analogy with the approximate modeling of generalized triangles, the parametric equation of the generalized hexahedron is approximated by a quadruple vector spline in terms of the simplex coordinates. For the spline of order one along each simplex coordinate, we obtain a regular tetrahedron whose parametric equation is

$$\mathbf{r}(t_1, t_2, t_3, t_4) = \mathbf{r}_1 t_1 + \mathbf{r}_2 t_2 + \mathbf{r}_3 t_3 + \mathbf{r}_4 t_4 \tag{5.31}$$

where $\mathbf{r}_i$, $i = 1, 2, 3, 4$, are the position vectors of the tetrahedron's vertices.

Starting from the parametric equation (5.27), all geometrical quantities of interest can be easily obtained using the unitary vector concept (Section 5.2.2). The only difference is that now we have additional quantities: the coordinate q, the unitary vector $\mathbf{a}_q$, the Lame coefficient e_q, and the unit vector $\mathbf{i}_q$. Particularly, the volume element dv can be expressed as

$$dv = (\mathbf{a}_p \times \mathbf{a}_q) \cdot \mathbf{a}_s \, dp \, dq \, ds \tag{5.32}$$

These geometrical quantities can be also determined starting from (5.30) by expressing one simplex coordinate in terms of the other three, which can be then considered as the p-, q-, and s-coordinates.

The basic technique for approximate modeling of volumes by trilinear volumes and regular tetrahedrons represents a generalization of the technique for approximate modeling of surfaces by bilinear surfaces and flat triangles. A structure or its surrounding media are first represented as an exact or approximate combination of generalized hexahedrons and tetrahedrons. If faces (edges) of these hexahedrons and tetrahedrons are not sufficiently flat (straight), they are subdivided, possibly over the p-, q-, and s-planes, into smaller generalized hexahedrons and tetrahedrons. Trilinear volumes and regular tetrahedrons replace generalized hexahedrons and tetrahedrons having the same ending nodes. Finally, the trilinear volumes and regular tetrahedrons, whose edges are electrically longer than the maximal allowed value, are further subdivided into smaller trilinear volumes and regular tetrahedrons, whose edges are electrically shorter than the maximum allowed value.

The geometrical modeling of surfaces is shown to be much more complicated than the geometrical modeling of wires. Similarly, the geometrical modeling of volumes is much complicated than the geometrical modeling of surfaces. This is one of the reasons why in this book the advantage for modeling of composite structures is given to the MoM applied to SIE when compared with either the MoM applied to VIE or the FEM applied to differential equations.

5.4 Wire-to-Plate Junctions and Related Structures

Geometrical modeling of wire-to-plate junctions (and related structures, such as protrusions of wires through dielectric surfaces) requires particular attention. Basis functions usually employed for approximation of currents are continuous functions along wires and over plates (see Chapter 6). However, if a wire's arbitrary part is connected to an arbitrary part of a plate, the current distribution is a continuous function neither along the wire, nor over the plate. Hence, the geometrical modeling of wire-to-plate junctions cannot be performed as independent geometrical modeling of wires and plates. This problem can be avoided if wires are modeled by plates. In that case, the geometry of a wire-to-plate junction is modeled in the same way as the geometry of a plate-to-plate junction. However, this type of modeling is impractical, because it unnecessarily requires a large number of building elements and unknowns to be introduced for wires.

For example, consider straight cylindrical wire connected to the center of a square plate. If the wire is electrically thin, then in the vicinity of the junction the radial surface current component is dominant on the plate. The intensity of this component is inversely proportional to the distance from the junction. Obviously, such a current distribution cannot be approximated by independent continuous basis functions along the cylinder and over the square. Very good approximation of these currents can be obtained if the cylinder and the square are modeled by bilinear surfaces. However, instead of one wire segment, at least five patches are used, and the number of unknowns for wires is increased at least 10 times [45].

To take advantage of the existing efficient techniques for thin-wire analysis, specific techniques must be applied. For example, Albertsen et al. introduce additional basis function over quadrilaterals in the form of hyperbolic paraboloid [21]. (These basis functions represent examples of the more general wedge functions proposed by Wachspress [46].) However, these basis functions cannot be used in conjunction with sophisticated basis functions that satisfy the continuity equation at interconnections of surface patches. Two general approaches that support the sophisticated basis functions are elaborated in what follows.

5.4.1 Attachment Modes

The current in the vicinity of a wire-to-plate junction can be decomposed into two components: the radial component flowing from the wire over a frill and vanishing at the outer frill edge, and a slowly varying component distributed over the plate. The frill is a new geometrical element that should be added to the joined wire and plate [Figure 5.16(a)]. Such a frill, together with the associated basis function for the approximation of the surface current, is referred to as the attachment mode [13, 47].

Numerical experiments have shown that the optimum outer radius of the simple frill is between 0.2 and 0.25 wavelengths [13]. If it is significantly less than 0.2 wavelengths, it turns out to be quite difficult to obtain satisfactory results for the current distribution. Therefore, for a junction close to an edge, a corner, a wedge, or a vertex, new attachment modes are needed [48, 49] [Figure 5.16(b–e)]. These attachment modes consist of four types of simpler geometrical elements: a complete and a truncated frill segment, and a complete and a truncated frill sector. Neither of these attachment modes can be applied if the radius of the plate curvature is much smaller than the wavelength. In that case, other attachment modes are needed.

Such an approach complicates the analysis significantly. For example, the minimum number of classes of field and impedance integrals that must be evaluated in the analysis is n and n^2, respectively, where n is the number of

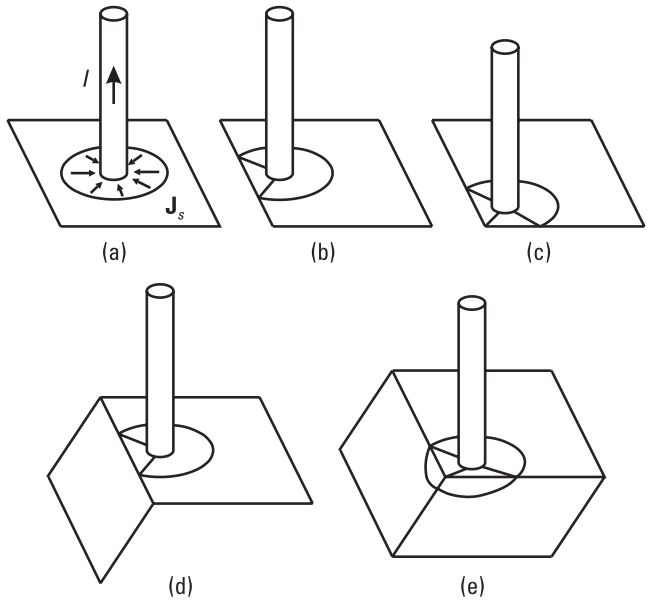

Figure 5.16 Attachment modes: wire is (a) in the middle of a plate, (b) close to an edge, (c) close to a corner, (d) close to a wedge, and (e) close to a vertex.

types of basic building elements used for modeling the geometry. For structures modeled exclusively by means of truncated cones and bilinear surfaces, $n = 2$ and $n^2 = 4$. However, if attachment modes are used consisting, for example, of five new types of building elements, these numbers become $n = 7$ and $n^2 = 49$. Instead of using a number of complicated additional attachment modes dealing with various wire-to-plate junctions, we recommend a specific segmentation technique based on the so-called general localized junction model.

5.4.2 General Localized Junction Model

Consider the junction of generalized wires and generalized quadrilaterals [Figure 5.17(a)]. One end of each wire and one side of each quadrilateral are situated in an electrically small domain, which will be referred to as the junction domain. All these ends are assumed to be interconnected by electrically small building elements (not shown) situated in the junction domain.

From the quasistatic analysis, the total current flowing out of the junction domain is found to be approximately zero. The partial currents flowing through the ends of the wires and the sides of the quadrilaterals in the junction domain do not depend excessively on the shape of the small building elements interconnecting them. It can, therefore, be assumed that a satisfactory analysis of such

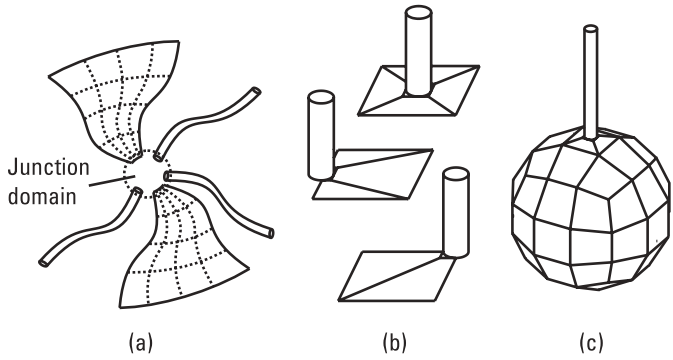

Figure 5.17 Localized junction: (a) general case, (b) a wire connected to the plate center, plate edge, and plate corner, and (c) a sphere with attached wire.

a junction can be obtained if the interconnecting elements are omitted from the geometrical model, provided that the current expansion for the wires and quadrilaterals is adopted so that the total current flowing out from the junction domain is zero. In other words, the interconnecting elements provide only quasistatic current connections between the ends of wires and sides of quadrilateral plates in the junction domain, and the field due to currents in these elements can be neglected.

This simple reasoning has significant consequences on the philosophy of modeling the geometry of metallic antennas and scatterers. No strict modeling of interconnections is indispensable, provided that the current continuity equation is stipulated properly. Specifically, two or more interconnected elements in a model need not be in physical contact at all, or they may be in contact at a single point or along only a part of their sides. This useful conclusion greatly simplifies modeling the geometry of junction domains. The junction domain is assumed to be electrically small. Hence, such a junction model will be referred to as the general localized junction model [50].

Using this simplified model of a junction, it is possible to analyze efficiently and accurately any wire-to-plate junction. For example, the junction of a wire and a central part of the plate (or a plate edge or a metallic corner) can be represented as a junction of a truncated cone and four (three, two) bilinear surfaces [Figure 5.17(b)]. Starting from the last two examples, it is simple to model the junction of a wire and a metallic wedge and the junction of a wire and a metallic vertex. Finally, the junction of a wire and a curved plate can be also modeled using this approach [Figure 5.17(c)].

For example, let us consider a monopole antenna at the center of a square plate [Figure 5.18(a)] with h = 421 mm, r = 0.8 mm, and l = 914 mm. The square plate is divided into M = 4 or M = 16 patches [Figure 5.18(b) and

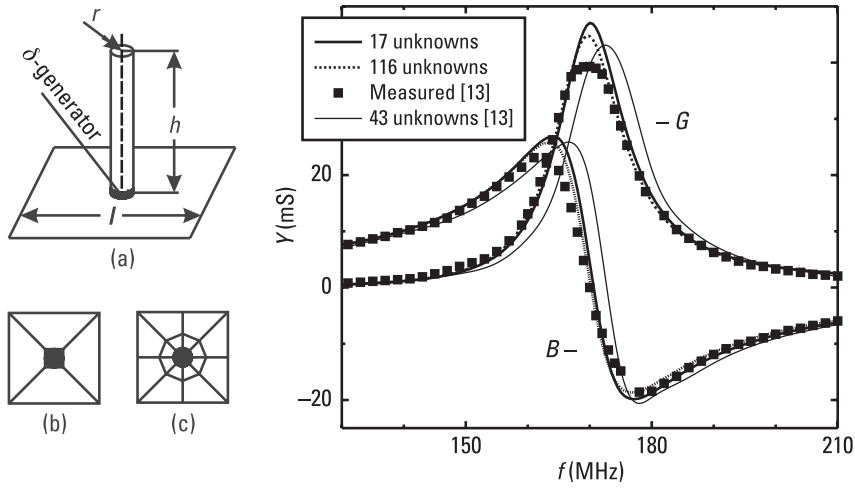

Figure 5.18 Monopole antenna at a center of square plate: (a) original structure, (b) segmentation scheme (M = 4 patches), (c) segmentation scheme (M = 16 patches), and (d) input admittance versus frequency. (After: [50].)

5.18(c)]. Figure 5.18(d) shows the admittance of the antenna versus frequency. Theoretical results obtained by the present method with N = 17 unknowns (M = 4 patches) and N = 116 unknowns (M = 16 patches) are compared with theoretical results obtained by using attachment modes with N = 43 unknowns and with experimental results [13]. Very good agreement of the results obtained by the present method with the experimental results is observed. (Theoretical results based on attachment modes are less accurate.) The difference between the results obtained by the present method with M = 4 patches and M = 16 patches is negligible. Hence, the simple approximation of the plate indicated in Figure 5.18(b) suffices.

5.4.3 Protrusion of a Wire Through a Dielectric Surface

Let us consider a wire protruding from a dielectric surface [Figure 5.19(a)]. If the analysis is based on the SIEs solved by the MoM, the problem is decomposed into two single-region problems with the cross sections shown in Figure 5.19(b). The field inside each region is produced by electric currents along the part of the wire situated in the region, and by equivalent electric and magnetic currents over the dielectric surface. Hence, there is a continuous flow of electric currents from the part of the wire to the dielectric surface in the first region, and a continuous flow of electric currents from the dielectric surface to the part of the wire in the second region [Figure 5.19(b)]. Keeping in mind the equivalence theorem (on which the decomposition is based), equivalent currents on opposite

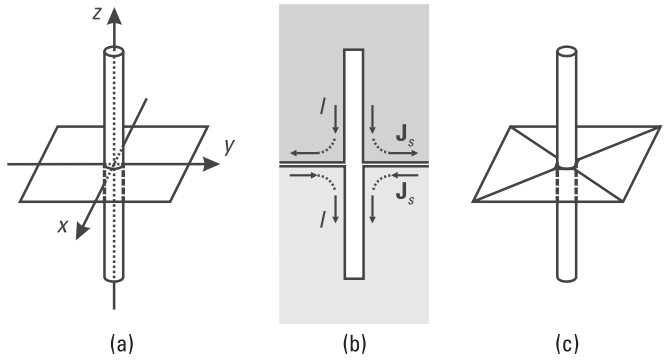

Figure 5.19 Protrusion of a wire through a dielectric surface: (a) original problem, (b) decomposition into two single-region problems, and (c) geometrical modeling.

faces of the dielectric surface have opposite directions, but the same intensity. Consequently, the continuity of currents along the wire is also satisfied. (The component of magnetic currents normal to the contour that represents a wire-to-surface junction vanishes at the contour.) The geometry of the protrusion and the distribution of electric currents resemble those of a wire-to-plate junction. Hence, the protrusion can be handled by any of two above-mentioned approaches (i.e., either by introducing attachment modes or performing segmentation based on the concept of general localized junction). We recommend the second approach, when the protrusion area is partitioned as shown in Figure 5.19(c).

For example, consider a symmetrical dipole antenna (h = 0.25m, r = $h/10$) protruding from a dielectric cube of side a = 0.6m (Figure 5.20, inset). The antenna is fed by a point voltage generator positioned exactly at the protrusion. Figure 5.20 shows admittance of the antenna versus frequency in the vicinity of the first resonance. Results are given for two values of the relative permittivity ϵ_r = 1 and ϵ_r = 2. In the first case, the cube is actually filled by a vacuum. The figure also gives the results for a dipole antenna without the cube. The results for the empty cube agree well with results for a stand-alone dipole antenna. This agreement is expected from the physical point of view, but not necessary from the numerical point of view, since two different numerical models are used to obtain the results. The good agreement justifies the proposed model. On the other hand, the increase of the dielectric constant (from ϵ_r = 1 to ϵ_r = 2) causes dramatic changes in the admittance, as is expected.

5.5 Automatic Parameterization of 3-D Geometries

The geometrical modeling of composite structures can be facilitated by introducing an automated segmentation (parameterization) of various 3-D geometries.

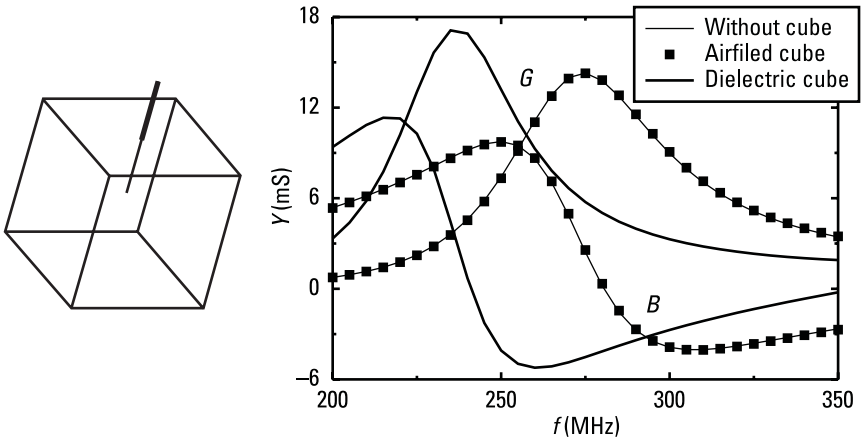

Figure 5.20 Admittance of a symmetrical dipole antenna (h = 0.25m, r = h/10) protruding from a dielectric cube of side a = 0.6m, versus frequency.

Many of them belong to five general classes that we term here the generalized circle, the generalized sphere, the generalized reflector, the two-generatrix body (2G body), and the generalized helix.

The generalized sphere is defined by the 3-D superquadric equation [51]:

$$\left(\frac{x}{a}\right)^{2/t} + \left(\frac{y}{b}\right)^{2/t} + \left(\frac{z}{c}\right)^{2/t} = 1 \tag{5.33}$$

where a, b, and c are lengths of half-axes along the x-, y-, and z-coordinates, respectively, and t is the shape factor. To obtain an exact model (made of six generalized quadrilaterals), we map a parallelepiped of sides $2a$, $2b$, and $2c$ onto the inscribed generalized sphere (Section 5.2.3). In the approximate model, each generalized quadrilateral is represented by $n \times n$ bilinear surfaces (Section 5.2.6). Depending on the values of the parameters a, b, c, and t, we obtain various forms of the generalized sphere. For t = 0, 1, and 2, the body degenerates to a parallelepiped, an ellipsoid, and an octahedron, respectively [Figure 5.21(a)]. In particular, for $a = b = c$ the ellipsoid degenerates into a sphere.

The generalized circle is defined by the 2-D superquadric equation [51], which is obtained from (5.33) for z = 0. Mapping a rectangle of sides $2a$ and $2b$ onto the inscribed generalized circle results in the exact model made of only one generalized quadrilateral. This mapping can be done in many ways, resulting in various parametric equations. Two parametric equations convenient for the approximate modeling are

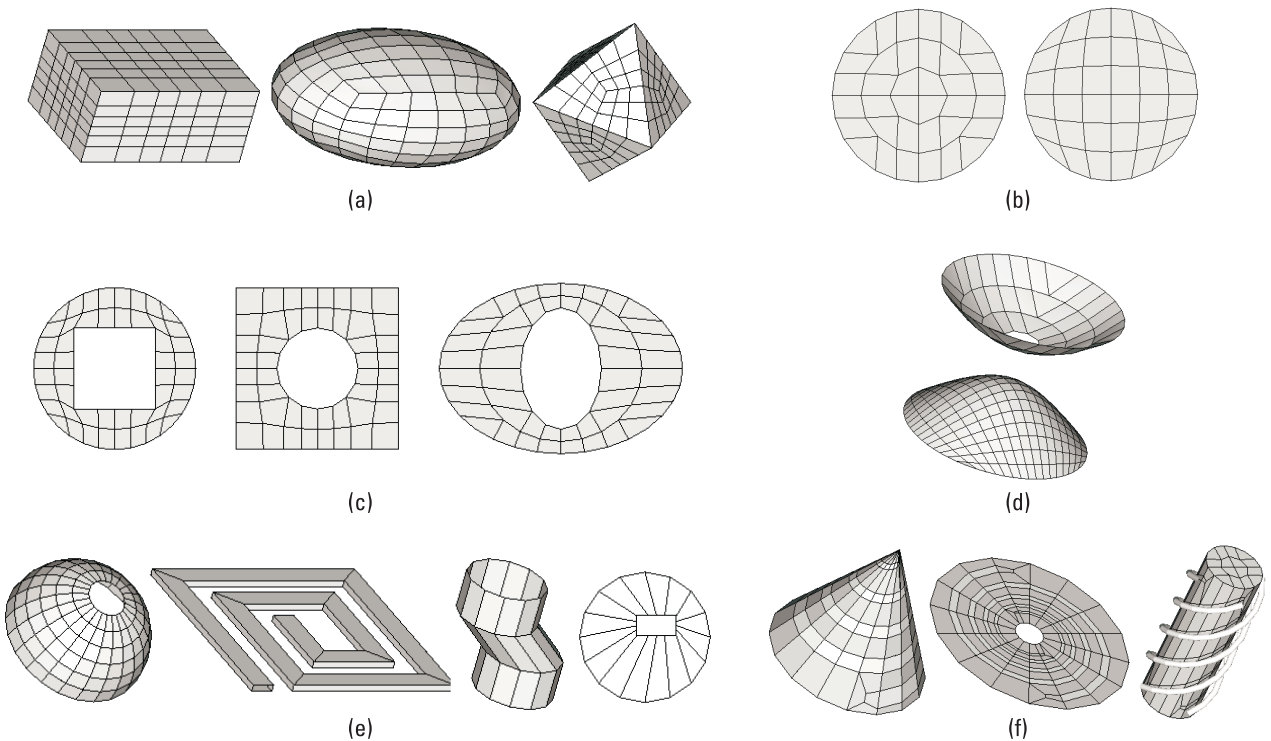

Figure 5.21 Automatic parameterization of 3-D geometries: (a) generalized sphere, (b) generalized circle without a hole, (c) generalized circle with a hole, (d) generalized reflectors, (e) 2G bodies, and (f) generalized helices.

$$\mathbf{r}(p, s) = \frac{(ap)\mathbf{i}_x + (bs)\mathbf{i}_y}{\sqrt[2/t]{p^{2/t} + s^{2/t}}} \max(p, s) \tag{5.34a}$$

$$\mathbf{r}(p, s) = \sqrt[2/t]{1 - \frac{1}{2}s^{2/t}}(ap)\mathbf{i}_x + \sqrt[2/t]{1 - \frac{1}{2}p^{2/t}}(bs)\mathbf{i}_y \tag{5.34b}$$

For example, consider two approximate models of a circle ($a = b$) [Figure 5.21(b)]. By omitting the four bilinear surfaces from the central part of the first model, a geometrical model of a frill is obtained. The parametric equation (5.34a) can be generalized to be valid for a generalized circle with a hole where the hole contour is defined by another 2-D superquadric equation. Thus, a variety of new shapes can be parameterized. As an example, Figure 5.21(c) shows models of a circle with a square hole, a square with a circular hole, and an ellipse with an elliptical hole.

A generalized reflector is defined as a part of a generic surface (paraboloid, ellipsoid, hyperboloid), whose main axis coincides with the z-axis. The projection of the reflector onto the xOy–plane defines its aperture, which can have different shapes (e.g., circle, ellipse, or square) and can be symmetrical with respect to the x- and y-axes or offset. The aperture bound is defined by a 2-D superquadric function. The surface of a generalized reflector is given by

$$z = \begin{cases} \dfrac{\rho^2}{4f} & \text{paraboloidal} \\ d\sqrt{1 + \dfrac{\rho^2}{f^2 - d^2}} & \begin{array}{l}\text{hyperboloidal } (d < f) \\ \text{elliptical } (d > f)\end{array} \end{cases} \qquad \rho = \sqrt{(x - x_0)^2 + (y - y_0)^2} \tag{5.35}$$

where f is the focal distance (i.e., the distance from the focus to the origin) and d is the minimal distance from the surface to the xOy-plane. Approximation of the generalized reflector by bilinear surfaces is performed in the same way as the approximation of a generalized circle, except that z-coordinate is evaluated according to (5.35) [51, 52]. As an example, Figure 5.21(d) shows models for a paraboloidal reflector with a circular hole and a hyperboloidal reflector.

Bodies of two generatrices (2G-bodies) are obtained from two generatrices by applying a certain rule [53]:

- *BoR:* The first generatrix is rotated about the second one, which is a circle.

- *Body of translation (BoT):* The first generatrix is translated along the second generatrix.
- *Body of constant cut (BoCC):* The first generatrix is moved along the second generatrix in such manner that a plane containing the first generatrix is normal to the second generatrix.
- *Body of connected generatrices (BoCG):* The first generatrix is connected to the second generatrix.

Generatrices can be defined in various ways, for instance, as a set of nodes placed in a plane, which are mutually connected by straight lines or by a 2-D super-quadric function. As an example, Figure 5.21(e) shows a model of a half sphere with a circular hole (BoR), a strip coil (BoCC), a broken pipe (BoT), and a circle with an arbitrarily positioned circular hole (BoCG).

Helical structures represent a wide class of structures. Basically, a helix consists of an arbitrary number of turns of a constant radius. However, the radius can also change with the circumferential angle in a linear or exponential way. In particular, a helix antenna can degenerate into a flat spiral. Helix antennas can be made of wires or of strips. The wire radius or the strip width can be constant or vary with the angle. They can stand alone in the air or be wound around a dielectric rod. They can be unifilar, bifilar, or quadrifilar. Finally, simple helices can be combined into a composite helix. All these structures can be completely specified by a limited number of common parameters. All these structures can be segmented using a unique general algorithm. Hence, these structures are grouped into a class named the generalized helix. Details of the algorithm for automated parameterization of the generalized helix will be omitted here. As an example, Figure 5.21(f) shows models of a unifilar strip helix placed on a dielectric cone, a bifilar exponential strip spiral placed on a flat dielectric surface, and a quadrifilar wire helix wound around a dielectric cylinder.

5.6 Automatic Segmentation of Electrically Large Surface Patches

The segmentation of bilinear surfaces and flat triangles (from which the geometrical model is made) is needed if sides of these elements are electrically longer than a critical electrical length (e.g., 0.1–0.25λ for subdomain basis functions and 1–2λ for entire-domain basis functions). Let us denote this length as L_c. The edge equal or longer (shorter) than the critical length (L_c) will be referred to as a long (short) edge. Any plate containing at least one long edge will be

referred as a large plate. The segmentation technique should be applied to the structure if it contains at least one large plate (i.e., at least one long edge).

Triangular and quadrilateral meshing schemes are widely used in the FEM analysis of 2-D problems [54–56]. Hence, it is natural to use experience from the FEM and implement it in the MoM. For example, the definition of the quality shape factor [54] and automatic triangulation routines [55] can be directly applied in the MoM if geometrical models consist of plates placed in a plane [57]. However, when the SIE is solved by the MoM, the surfaces are placed in 3-D space, and routines from the FEM need to be modified or new routines developed. There are only few papers devoted to the specific segmentation problems occurring in the case when the SIE is solved by the MoM [35, 58]. For quadrilateral meshing it is found that as many elements as possible should be of the rectangular type [59]. Based on these conclusions, two segmentation techniques for polygonal plates are developed [35]. However, these techniques do not enable automatic mesh generation. Various desirable properties of the segmentation from the viewpoint of solving the SIE by the MoM are discussed in [58], and accordingly, two segmentation techniques are suggested. The discussion and the proposed techniques will be briefly presented in this section.

In general, it is desirable that the segmentation scheme enables efficient analysis, that the segmentation be performed automatically, and that the segmentation procedure be relatively simple and easily implemented. For the subdomain approximation, a plate should be subdivided into as few patches as possible, with sides shorter than the critical length. In the case of the quadrilateral (triangular) meshing, the optimal shape of a patch is a square (equilateral triangle).[4] In particular, quadrilateral meshing is at least twice more efficient than triangular modeling [37]. The efficiency of the quadrilateral meshing can be further improved if the entire-domain approximation is applied. In this case, the order of approximation along a pair of non-neighboring quadrilateral edges is dependent on the electrical length of the edges, and the optimal shape of the patch is a rectangle.

The efficiency of the analysis depends also on the mutual position of patches. For rectangular patches, the numerical integration can be facilitated if they are parallel [35]. The evaluation of the mutual impedance of two patches is much more difficult if they overlap. It is known that the basis functions, satisfying the continuity equation at the interconnection of patches, are much more efficient than the basis functions not satisfying this equation. The basis functions satisfying this equation are usually defined over two neighboring patches with two common nodes determining the common edge.

4. For a given maximal edge length, the square (equilateral triangle) is a quadrilateral (triangle) with the largest surface area.

The example shown in Figure 5.22(a) suggests an additional requirement. Let only the edges AB and AC be long edges. In that case, only the plate ABCD is a large plate. The segmentation marked by dashed lines in Figure 5.22(b) is not correct. Namely, new subplates do not have two common nodes with neighboring plates (i.e., they are not electrically connected with the neighboring plates). A possible correct segmentation is marked by dashed lines in Figure 5.22(c). The segmentation of the large plate ABCD causes an undesirable segmentation of all other small plates of the structure. The segmentation of small plates can be avoided if the large plate is subdivided as shown in Figure 5.22(d). Finally, it is concluded that the optimal segmentation technique should partition large plates in a way that does not cause the partition of neighboring small plates.

We can summarize the desirable properties of the segmentation from the viewpoint of solving the SIE by the MoM:

- Shapes of patches should be as regular as possible [i.e., as close as possible to a square (rectangle) or an equilateral triangle].
- The final model should be made mostly of quadrilaterals, possibly without triangles.
- Electrically connected neighboring patches should have a common edge defined by two common nodes.
- The neighboring patches should not overlap.
- The segmentation technique should enable a partition of large plates in manner that does not cause the partition of neighboring small plates.

Starting from these properties, two techniques for automatic meshing are proposed.

For the moment, let us consider only plates whose edges are shorter than $2L_c$. Let any edge longer than L_c be subdivided into two short edges. Depending

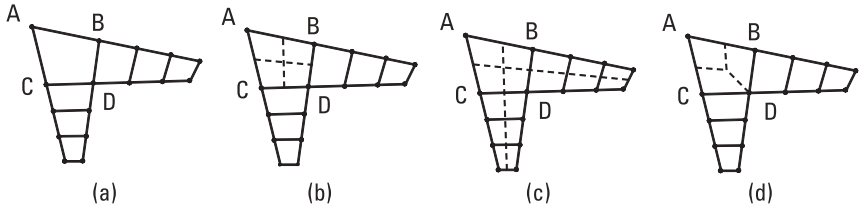

Figure 5.22 The partitioning of a large plate connected to small plates: (a) the initial structure, (b) an incorrect partitioning, (c) partitioning of the small plates caused by partitioning of the large plate, and (d) partitioning of the large plate that does not cause partitioning of the small plates.

on the number and the position of the subdivided edges, there are six basic situations and corresponding segmentation schemes [Figure 5.23(a–f)]. The situations given in Figure 5.23(a, d, and e) cannot be segmented into quadrilaterals without introducing a small triangular hole. However, the hole of the maximal dimension of 0.01 of the wavelength practically does not affect the analysis.

In general, plate edges can be equal or longer than $2L_c$. Then, the segmentation is performed by the iteration procedure. Each iteration consists of two basic steps:

1. Each long edge is subdivided into the minimum number of short edges possible by defining new (interior) equispaced nodes;
2. Each large plate is subdivided into smaller plates according to one of the schemes given in Figure 5.23(a–f), using only one interior node per long edge.

The remaining interior nodes will be used in the next iteration. As a result of this partial segmentation, new long edges can be created. The iterative procedure is finished when all interior nodes are taken into account and all long edges are subdivided into short edges.

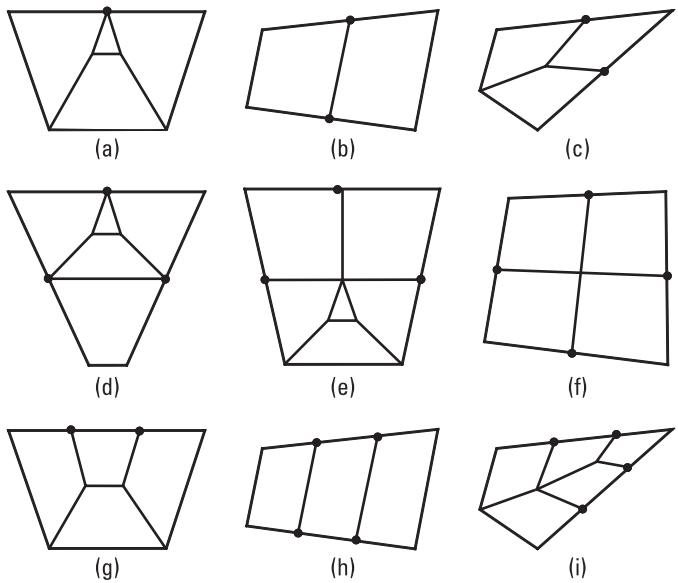

Figure 5.23 Basic schemes for the partitioning of large plates: (a) to (f) the single-node technique, and (g) to (i) the double-node technique.

The basic segmentation schemes given in Figure 5.23(d–f) can be obtained in two iteration steps, combining the segmentation schemes given at Figure 5.23(a, b). Thus, a simple segmentation algorithm is obtained. Considering that only one interior node per long edge is used in one iteration, the segmentation technique will be termed the *single-node* (SN) technique. The corresponding mesh will be referred to as the SN mesh.

The small holes introduced in the basic schemes given in Figure 5.23(a–f) can be avoided if each interior node placed in the middle of long edges is replaced by two equispaced nodes. The first three basic schemes are modified as shown in Figure 5.23(g–i). This segmentation technique and the corresponding mesh will be referred to as the *double-node* (DN) technique and the DN mesh, respectively

For example, let us consider a square plate of side $a = 3L_c$, with two triangular holes. Figure 5.24 shows the starting model and the corresponding SN and DN meshes. (The holes introduced by the SN technique are very small, and therefore not visible in the figure.) The SN mesh consists of fewer quadrialterals than the DN mesh, but the quadrialterals in DN mesh have more regular shapes than quadrilaterals in the SN mesh. When a standard accuracy is required, both techniques combined with the entire-domain approximation need approximately the same number of unknowns [58]. However, when a high accuracy is required from the entire-domain solution, the DN technique is superior to the SN technique, as will be shown in the next example.

Let us consider a 20-dB standard-gain horn antenna [Figure 5.25(a)]. The antenna dimensions are $a = 0.9''$, $b = 0.4''$, $l_a = 4.87''$, $l_b = 3.62''$, $l_1 + l_2 = 1.57''$, and $l_3 = 4.87''$. The horn is excited by a waveguide fed by a coaxial line. The antenna is made of a perfect conductor. The plates of a finite thickness are modeled by infinitesimally thin plates, resulting in the surface currents representing the sum of interior and exterior antenna currents. If only the radiation pattern is evaluated, the excitation by the coaxial line can be

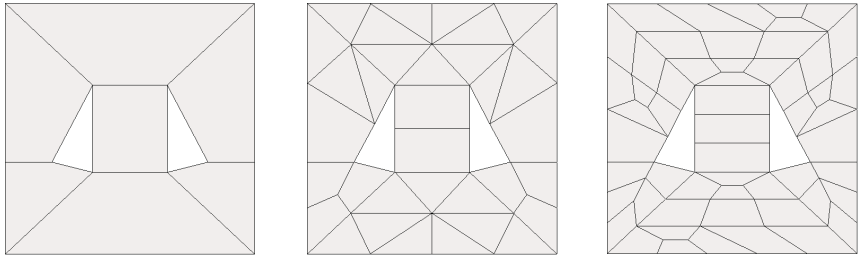

Figure 5.24 Square plate with two triangular holes: (a) starting structure, (b) SN mesh, and (c) DN mesh.

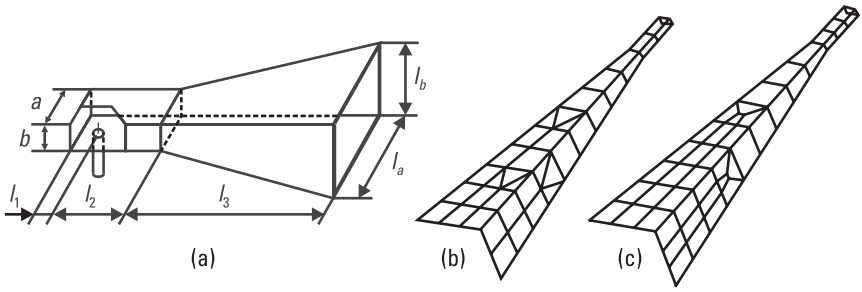

Figure 5.25 Horn antenna: (a) sketch, (b) SN model, and (c) DN model.

replaced by a dipole antenna situated in the waveguide in such a manner that two symmetry planes can be used to facilitate the analysis. The corresponding geometrical models based on SN and DN meshing are shown in Figures 5.25(b) and 5.25(c), respectively. A good agreement between the theoretical results obtained by using the DN mesh (891 unknowns) and the experimental results can be observed (see Section 9.2).

Similar results are obtained if the SN mesh (922 unknowns) is used instead of the DN mesh (Figure 5.26). By continuously increasing the orders of approximation, the number of unknowns takes the following values: 922, 945, 975, 983, 997, 1,005, 1,013, 1,021, 1,030, 1,038, 1,166, and so forth. The results obtained for 922 to 1,166 unknowns are shown in Figure 5.26(b). A very good agreement among all the results can be observed for levels that are no more than 25 dB below the directive gain in the main direction. For lower levels, a significant "noise" can be observed. By continuously increasing the orders of approximation for the DN mesh, the number of unknowns takes the following values: 891, 901, 909, 917, 941, 949, 973, 999, 1,107, and so forth. The results obtained for 891 to 1,107 unknowns are shown in Figure 5.26(a). A very good agreement among all these results can also be observed. The "noise" occurring for very low levels is negligible.

5.7 Summary

The task of the geometrical modeling is to represent the structure by a proper combination of building elements (linear segments, surface patches, and volume bricks) where approximating functions can be easily defined. Such modeling requires at least two coordinate systems. The global coordinate system serves for the definition of the structure as a whole, and for the description of the near-field points, far-field points, and so forth. The local coordinate system (or possibly systems) is associated with building elements, and it is used for the

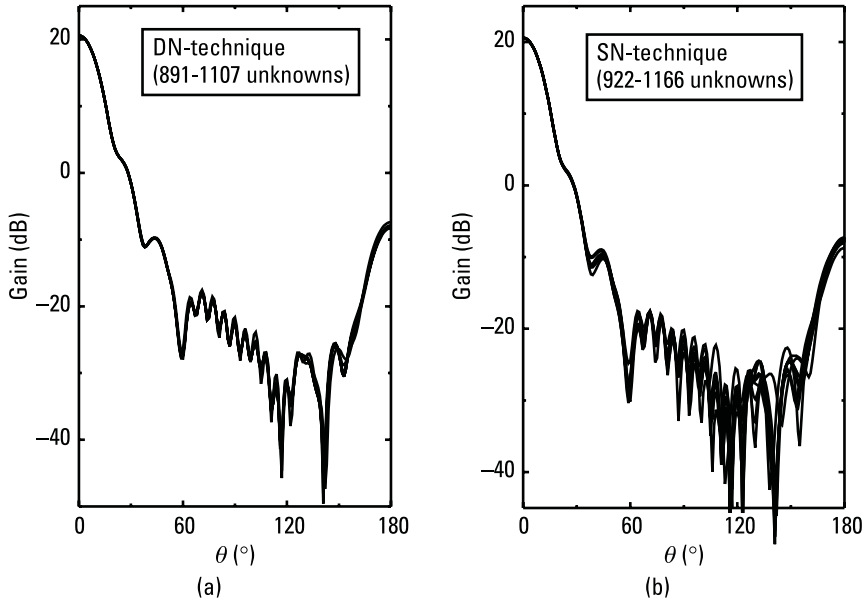

Figure 5.26 The *H*-plane pattern of the horn antenna shown in Figure 5.25 at 10 GHz: (a) results obtained for the DN mesh for various numbers of unknowns (891–1,107), and (b) results obtained for the SN mesh for various numbers of unknowns (922–1,166).

description of the corresponding unknown quantity. The linear, surface, and volume building elements can be treated as particular cases of generalized building elements (wires, quadrilaterals, triangles, hexahedrons, and tetrahedrons), which can be used for either exact or approximate modeling of composite structures.

Exact modeling requires the development of specific building elements for each new structure. Hence, it is not convenient from the viewpoint of making a general algorithm for electromagnetic modeling. However, there are two reasons why exact modeling is still very important. First, it often represents an intermediate step in approximate modeling. Second, analysis based on exact geometrical modeling serves as a check of the accuracy of the analysis based on approximate modeling.

The generalized building elements can be efficiently approximated by scalar and vector splines. The building elements defined by the first-order splines (right truncated cones, bilinear surfaces, flat triangles, trilinear volumes, and tetrahedrons), combined with the concept of the equivalent radius, enable a very efficient modeling of curved wires, curved surfaces, and curved bodies. The building elements defined by higher-order splines can improve the efficiency

of the analysis of curved structures only if at least one element dimension is greater than some critical length (e.g., 0.5 wavelength), and if higher order basis functions are used.

The geometrical modeling of dielectric volumes (required by the MoM/VIE or FEM) is a much more complex task than geometrical modeling of metallic and dielectric surfaces (required by the MoM/SIE). In particular, the MoM/SIE analysis based on bilinear surfaces seems to be at least twice as efficient than that based on flat triangles. Hence, the authors give advantage to the MoM/SIE in the electromagnetic modeling of composite metallic and dielectric structures and recommend right truncated cones and bilinear surfaces as basic elements for the geometrical modeling.

Special care should be devoted to the modeling of wire-to-plate junctions and protrusions of wires through dielectric surfaces. There are two general approaches when the modeling is based on truncated cones and bilinear surfaces. According to the first one, each class of junctions (protrusions) is modeled by a specific attachment mode. Instead of using a number of relatively complicated attachment modes, a specific segmentation technique, which is based on the general localized junction model, is applied to the junction (protrusion) area. We recommend the second approach.

As an aid to the user of the computer code, it is necessary to develop algorithms for automatic modeling of complex structures or their parts. Many of them belong to five general classes of 3-D geometries: the generalized sphere, the generalized circle, the generalized reflector, the two-generatrix body (2G body), and the generalized helix. In particular, there are various types of the 2G body, like a BoR, BoT, BoCC, and BoCG. Algorithms for the automatic modeling of these classes by bilinear surfaces are given in this chapter.

A geometrical model cannot provide satisfactory numerical results if the building elements are electrically larger than allowed (in terms of the wavelength). Two methods for the automatic segmentation of electrically large bilinear surfaces are presented in this chapter: the single-node and the double-node technique. The latter technique produces more regular shapes of surface patches, and is thus recommended.

References

[1] Richmond, J. H., "Digital Solutions of the Rigorous Equations for Scattering Problems," *Proc. IEEE,* Vol. 53, 1965, pp. 796–804.

[2] Harrington, R. F., and J. R. Mautz, "Straight Wires with Arbitrary Excitation and Loading," *IEEE Trans. Antennas Propagat.,* Vol. AP-15, July 1967, pp. 502–515.

[3] Strait, B. J., and K. Hirasawa, "Array Design for a Specified Pattern by Matrix Methods," *IEEE Trans. Antennas Propagat.,* Vol. AP-17, No. 2, March 1969, pp. 237–239.

[4] Thiele, G. A., "Wire Antennas," in *Numerical Techniques for Electromagnetics,* R. Mittra, (ed.), London: Pergamon, 1973.

[5] Imbriale, W. A., "Applications of the Method of Moments to Thin Wire Elements and Arrays," in *Numerical and Asymptotic Techniques in Electromagnetics,* R. Mittra, (ed.), Berlin, Germany: Springer-Verlag, 1975.

[6] Miller, E. K., and F. J. Deadrick, "Some Computational Aspects of Thin-Wire Modeling," in *Numerical and Asymptotic Techniques in Electromagnetics,* R. Mittra, (ed.), New York: Springer-Verlag, 1975.

[7] Popovic, B. D., and B. M. Kolundzija, *Analysis of Metallic Antennas and Scatterers,* London: IEE, 1994.

[8] Kolundzija, B. M., "Electromagnetic Modeling of Wire-to-Plate Structures," D.Sc. Dissertation, University of Belgrade, 1990 (in Serbian).

[9] Kolundzija, B. M., and B. D. Popovic, "Entire-Domain Galerkin Method for Analysis of Generalized Wire Antennas and Scatterers," *Proc. IEE H,* Vol. 139, No. 1, February 1992, pp. 17–24.

[10] Kalafus, R. M., "Broadband Band Dipole Design Using the Method of Moments," *IEEE Trans. Antennas Propagat.,* Vol. AP-19, No. 6, November 1971, pp. 771–773.

[11] Djordjevic, A. R., B. D. Popovic, and M. B. Dragovic, "Analysis of Electrically Thick Antennas of Revolution," *Proc. 3rd ICAP,* Norwich, U.K., IEE Conf. Publ. 219, 1983, pp. 390–394.

[12] Kolundzija, B. M., "Effect of a Wire End in Thin-Wire Analysis," *Proc. IEEE AP-S Symposium,* Syracuse, NY, 1988, pp. 843–846.

[13] Newman, E. H., and D. M. Pozar, "Electromagnetic Modeling of Composite Wire and Surface Geometries," *IEEE Trans. Antennas Propagat.,* Vol. AP-26, No. 6, November 1978, pp. 784–789.

[14] Popovic, B. D., and A. Nesic, "Generalization of the Concept of Equivalent Radius of Thin Cylindrical Antennas," *Proc. IEE H,* Vol. 131, 1984, pp. 153–158.

[15] Yeh, Y. S., and K. K. Mei, "Theory of Conical Equiangular-Spiral Antennas," *IEEE Trans. Antennas Propagat.,* Vol. AP-15, No. 5, September 1967, pp. 634–639.

[16] Froeberg, C.-E., *Introduction to Numerical Analysis,* Reading, MA: Addison-Wesley, 1965.

[17] Barsky, B. A., *Computer Graphics and Geometric Modeling Using Beta-Spline,* New York: Springer-Verlag, 1988.

[18] Champagne, N. J., J. T. Williams, and D. R. Wilton, "The Use of Curved Segments for Numerically Modeling Thin Wire Antennas and Scatterers," *IEEE Trans. Antennas Propagat.,* Vol. 40, No. 6, June 1992, pp. 682–689.

[19] Kolundzija, B., J. Ognjanovic, and T. Sarkar, "On the Limit of WIPL-D Code," *Proc. of 18th Annual Review of Progress in Applied Computational Electromagnetics, ACES Conference,* Monterey, CA, 2002, pp. 615–622.

[20] Kolundzija, B. M., "Comparison of a Class of Sub-Domain and Entire-Domain Basis Functions Automatically Satisfying KCL," *IEEE Trans. Antennas Propagat.,* Vol. 44, No. 10, October 1996, pp. 1362–1366.

[21] Albertsen, N. C., J. E. Hansen, and N. E. Jensen, "Computation of Radiation from Wire Antennas on Conducting Bodies," *IEEE Trans. Antennas Propagat.,* Vol. AP-22, No. 2, March 1974, pp. 479–487.

[22] Wang, N. N., J. H. Richmond, and M. C. Gilreath, "Sinusoidal Reaction Formulation for Radiation and Scattering from Conducting Surfaces," *IEEE Trans. Antennas Propagat.,* Vol. AP-23, No. 3, May 1975, pp. 376–382.

[23] Glisson, A. W., and D. R. Wilton, "Simple and Efficient Numerical Methods for Problems of Electromagnetic Radiation and Scattering from Surfaces," *IEEE Trans. Antennas Propagat.,* Vol. AP-28, No. 5, September 1980, pp. 593–603.

[24] Kolundzija, B. M., and A. R. Djordjevic, "Analysis of Dipole Antenna with Corner Reflector," *Proc. 7th Colloquium on Microwave Communication,* Budapest, Hungary, 1982, pp. 319–322.

[25] Rao, S. M., D. R. Wilton, and A. W. Glisson, "Electromagnetic Scattering by Surfaces of Arbitrary Shape," *IEEE Trans. Antennas Propagat.,* Vol. AP-30, No. 3, May 1982, pp. 409–418.

[26] Umashankar, K., A. Taflove, and S. M. Rao, "Electromagnetic Scattering by Arbitrary Shaped Three-Dimensional Homogeneous Lossy Dielectric Objects," *IEEE Trans. Antennas Propagat.,* Vol. AP-34, June 1986, pp. 758–766.

[27] Sarkar, T. K., E. Arvas, and S. Ponnapalli, "Electromagnetic Scattering from Dielectric Bodies," *IEEE Trans. Antennas Propagat.,* Vol. AP-37, May 1989, pp. 673–676.

[28] Rao, S. M., et al., "Electromagnetic Radiation and Scattering from Finite Conducting and Dielectric Structures: Surface/Surface Formulation," *IEEE Trans. Antennas Propagat.,* Vol. AP-39, July 1991, pp. 1034–1037.

[29] Kolundzija, B. M., and B. D. Popovic, "Entire-Domain Galerkin Method for Analysis of Metallic Antennas and Scatterers," *Proc. IEE,* Pt. H, Vol. 140, No. 1, February 1993, pp. 1–10.

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

[31] Knepp, D. L., and J. Goldhirsh, "Numerical Analysis of Electromagnetic Radiation Properties of Smooth Conducting Bodies of Arbitrary Shape," *IEEE Trans. Antennas Propagat.,* Vol. 20, No. 3, May 1972, pp. 383–388.

[32] Wandzura, S., "Electric Current Basis Functions for Curved Surfaces," *Electromagnetics,* Vol. 12, No. 1, 1992, pp. 77–91.

[33] Valle, L., F. Rivas, and M. F. Catedra, "Combining the Moment Method with Geometrical Modeling by NURBS Surfaces and Bezier Patches," *IEEE Trans. Antennas Propagat.,* Vol. 42, No. 3, March 1994, pp. 373–381.

[34] Singh, J., and A. T. Adams, "A Nonrectangular Patch Model for Scattering from Surfaces," *IEEE Trans. Antennas Propagat.,* Vol. AP-27, No. 4, July 1979, pp. 531–535.

[35] Newman, E. H., and P. Tulyathan, "A Surface Patch Model for Polygonal Plates," *IEEE Trans. Antennas Propagat.,* Vol. AP-30, No. 4, July 1982, pp. 588–593.

[36] Newman, E. H., and P. Alexandropulos, "Polyonal Plate Modeling of Realistic Structures," *IEEE Trans. Antennas Propagat.,* Vol. AP-32, No. 7, July 1984, pp. 742–747.

[37] Kolundzija, B. M., "On the Locally Continuous Formulation of Surface Doublets," *IEEE Trans. Antennas Propagat.,* Vol. 46, No. 12, December 1998, pp. 1879–1883.

[38] Mie, G., "Beitrage zur Optik Truber Medien, Speziell Koloider Metallosungen," *Ann. Phys.,* Vol. 25, 1908, p. 377.

[39] Livesay, D. E., and K-M. Chen, "Electromagnetic Fields Induced Inside Arbitrarily Shaped Biological Bodies," *IEEE Trans. Microwave Theory Tech.,* Vol. 22, No. 12, December 1974, pp. 1273–1280.

[40] Schaubert, D. H., D. R. Wilton, and A. W. Glisson, "A Tetrahedral Modeling Method for Electromagnetic Scattering by Arbitrarily Shaped Inhomogeneous Dielectric Bodies," *IEEE Trans. Antennas Propagat.,* Vol. 32, January 1984, pp. 77–85.

[41] Notaros, B. M., B. D. Popovic, and B. M. Kolundzija, "Electro-Magnetic Modeling of Inhomogeneous Dielectric Bodies by Trilinear Hexahedrons," *Proc. of XXXVIII Conf. ETRAN,* Nis, Pt. II, June 1994, pp. 119–120.

[42] Barton, M. L., and Z. J. Cendes, "New Vector Finite Elements for Three-Dimensional Magnetic Field Computation," *J. Appl. Phys.,* Vol. 61, No. 8, April 1987, pp. 3919–3921.

[43] Van Welij, J. S., "Calculation of Eddy Currents in Terms of H in Hexahedra," *IEEE Trans. Magnetics,* Vol. 21, 1985, pp. 2239–2241.

[44] Crowly, C. W., P. P. Silvester, and H. Hurwitz, "Covariant Projection Elements for 3-D Vector Field Problems," *IEEE Trans. Magnetics,* Vol. 24, 1988, pp. 397–400.

[45] Kolundzija, B., and B. Reljic, "Plate Modeling of Wire Structures," *Proc. of IEEE AP-S Int. Symp.,* Montreal, Quebec, 1997, pp. 1798–1801.

[46] Wachpress, E. L., "A Rational Basis for Function Approximation," *Publ. Conf. Applications of Numerical Analysis,* Berlin, Germany: Springer-Verlag, 1971.

[47] Shaeffer, J. F., and L. N. Medgyesi-Mitschang, "Radiation from Wire Antenna Attached to Bodies of Revolution: The Junction Problem," *IEEE Trans. Antennas Propagat.,* Vol. 29, 1981, pp. 479–487.

[48] Pozar, D. M., and E. H. Newman, "Analysis of a Monopole Mounted Near an Edge or a Vertex," *IEEE Trans. Antennas Propagat.,* Vol. 29, 1981, pp. 488–495.

[49] Pozar, D. M., and E. H. Newman, "Analysis of a Monopole Mounted Near or at the Edge of a Half Plane," *IEEE Trans. Antennas Propagat.,* Vol. 30, 1982, pp. 401–408.

[50] Kolundzija, B. M., and B. D. Popovic, "General Localized Junction Model in the Analysis of Wire-to-Plate Junctions," *Proc. IEE H,* Vol. 141, No. 1, February 1994, pp. 1–7.

[51] Kolundzija, B. M., J. S. Ognjanovic, and T. K. Sarkar, *WIPL-D—Electromagnetic Modeling of Composite Metallic and Dielectric Structures, Software and User's Manual,* Norwood, MA: Artech House, 1995.

[52] Kolundzija, B. M., and T. K. Sarkar, "Analysis of Prime-Focus Symmetrical Reflector Antennas by WIPL Code," *Proc. of 26th EuMc,* Prague, Czech Republic, 1996, pp. 781–785.

[53] Kolundzija, B. M., and A. R. Djordjevic, "Spiral Super-Quadric Generatrix and Bodies of Two Generatrices in Automated Parameterization of 3-D Geometries," *IEEE Transaction on Microwave Theory and Techniques,* Vol. 45, No. 5, May 1997, pp. 864–866.

[54] Lindholm, D. A., "Automatic Triangular Mesh Generation on Surface of Polyhedra," *IEEE Trans. Magnetics.,* Vol. MAG-19, November 1983, pp. 2539–2542.

[55] Segerlind, L. J., *Applied Finite Element Analysis,* New York: John Wiley and Sons, 1984.

[56] Cendes, Z. J., D. N. Shenton, and H. Shahnasser, "Magnetic Field Computation Using Delaunay Triangulation and Complementary Finite Element Methods," *IEEE Trans. Magnetics.,* Vol. MAG-19, November 1983, pp. 2551–2554.

[57] Virga, K. L., and Y. Rahmat-Samii, "RCS Characterization of a Finite Ground Plane with Perforated Apertures: Simulations and Measurements," *IEEE Trans. Antennas Propagat.,* Vol. AP-42, July 1994, pp. 1491–1501.

[58] Kolundzija, B. M., "Automatic Mesh Generation Using Single- and Double-Node Segmentation Techniques," *IEEE Antennas Propagat. M.,* Vol. 40, No. 4, August 1998, pp. 30–38.

[59] Newman, E. H., and D. M. Pozar, "Considerations for Efficient Wire/Surface Modeling," *IEEE Trans. Antennas Propagat.,* Vol. AP-28, January 1980, pp. 121–125.

6

Approximation of Currents and Fields

The approximation of the unknown current (field) distribution is always adopted in the form of a series of known functions (referred to as basis or expansion functions), which are multiplied by unknown coefficients that should be determined. The approximation is tightly connected with the equation and the method that is used for the analysis.

If we use the MoM to solve the SIEs, the unknown quantities are surface currents. For a generalized wire, we assume that the surface currents have only the component along the wire that depends on the local coordinate along the wire. For a quadrilateral, the currents are split into two components dependent on two local coordinates. For a generalized triangle, the surface currents are decomposed into three components dependent on three local coordinates. Hence, the approximations of currents along generalized wires, quadrilaterals, and triangles are treated separately in Sections, 6.1, 6.2, and 6.3, respectively.

If we use the MoM to solve the VIEs, the unknown quantities are volume currents. Generally, the volume currents have three components dependent on three coordinates. This is the way the currents are approximated within generalized hexahedrons. However, for a generalized tetrahedron the current is decomposed into four components dependent on four local coordinates. (An analogous situation is if we use the FEM, except that the unknowns are field or potential distributions.) The approximations of currents for the generalized tetrahedrons and hexahedrons are given in Sections 6.4 and 6.5.

Although approximations of currents for various building elements differ in the number of components and coordinates on which they depend, there are many similarities between them. Regarding the ease of implementation, it is extremely convenient to describe the approximate currents in local parametric

coordinate systems, the same ones with respect to which the building elements are specified. The stability, accuracy, and efficiency of the results are significantly improved if the approximations are forced to satisfy the continuity equation or quasistatic relations. The efficiency and accuracy are improved if higher-order expansions are used for electrically larger building elements. In particular, the efficiency and accuracy are improved if the concept of the entire-domain approximation is used instead of the subdomain approximation.

The concepts of the entire-domain and subdomain approximations evolved from the early beginning of numerical modeling in electromagnetics. Originally, basis functions are classified as entire-domain functions, defined over the entire (continuous) domain of the integral operator, and as subdomain functions, defined over a subsection of the entire domain. However, such classification makes sense only if the surface currents are approximated over simple structures (e.g., a single wire, a metallic plate, or a dielectric cube), when each basis function can be easily defined over the entire domain of the structure. In the case of complex, real-life examples, it is very difficult and impractical to define each basis function over the entire domain of the structure. In that sense, we consider the basis functions as entire-domain functions if they are defined over the entire domain of a building element. We consider them to be subdomain functions if they are defined over a part of the domain of a building element. Usually, the subdomain expansions use the lowest-order basis functions over a building element, although this is not necessary. One should distinguish the classification of subdomain and entire-domain basis functions from the classification of low- and high-order basis functions. Finally, subdomain basis functions are easier to implement than entire-domain ones. Sections 6.1 to 6.5 consider these issues in detail.

A particular problem is how to make an electrical connection between basically incompatible building elements (e.g., a wire to a plate, a triangle to a quadrilateral, and so forth). These situations are considered in Section 6.6.

Because the same composite metallic and dielectric structure can be analyzed in three different ways (MoM/SIE, MoM/VIE, and FEM), it is very important to consider the efficiency of these approaches. A comparison of the three methods is made in Section 6.7, based on the topological analysis of two canonical problems.

6.1 Approximation of Currents Along Wires

Most often the approximation of currents along generalized wires is expressed in terms of length coordinates. However, we shall use the representation in

terms of parametric coordinates, which leads to much simpler expressions.[1] The total current along a wire is approximated as

$$I(s) = \sum_{i=0}^{n} a_i f_i^n(s) \quad -1 \le s \le 1 \tag{6.1}$$

where $f_i^n(s)$ are known basis functions, a_i are unknown coefficients to be determined, and n is the approximation order. The per-unit-length charge density, which corresponds to $I(s)$, is determined by

$$Q'(s) = \frac{j}{\omega} \frac{dI(s)}{dl_s} = \frac{j}{\omega e_s(s)} \frac{dI(s)}{ds} \tag{6.2}$$

where dl_s is the length element along the generatrix of the wire and $e_s(s)$ is the corresponding Lame coefficient.

The superscript n in $f_i^n(s)$ means that, generally, the ith basis function is different for various n. If, however, the ith basis function does not depend on n, the superscript can be omitted. In that case we obtain the so-called hierarchical expansion: By increasing the approximation order, we only add new (higher-order) basis functions to the existing expansion.

The accuracy and efficiency of the analysis depend on the choice of the basis functions. Hence, we consider various kinds of basis functions used in the thin-wire analysis. In Section 6.1.1 we present the most common subdomain expansions, which are simple and easy to implement. However, the entire-domain expansions, whose basic forms are defined in Section 6.1.2, have higher approximating capabilities. In order to exploit these potentials fully, the continuity of currents or quasistatic relations at wire free ends and junctions should be incorporated into the expansions (Sections 6.1.3 and 6.1.5). Among various approximations, the combined trigonometric and polynomial expansion is found to be particularly efficient (Section 6.1.4). Finally, in Section 6.1.6 we show how the basis functions along wires can be defined in terms of simplex coordinates, which will be important later, for the definition of basis functions for generalized triangles and tetrahedrons.

6.1.1 Subdomain Basis Functions

To construct a subdomain approximation, the wire segment is subdivided into subsegments. Then, the basis functions $f_i^n(s)$ in (6.1) are defined along one or more neighboring subsegments.

1. It can be achieved by proper normalization of the length coordinates.

The simplest subdomain basis function is a pulse, which results in the PWC approximation of currents [1]. The pulse is defined as a constant function along one subsegment and zero elsewhere; that is,

$$f_i^n(s) = \begin{cases} 1 & s_i \le s \le s_{i+1} \\ 0 & \text{elsewhere} \end{cases} \qquad i = 0, \ldots, n \tag{6.3}$$

The simplest subdivision occurs when points s_i are uniformly distributed [Figure 6.1(a)]. The PWC approximation results in point charges (delta-function charges) at both endpoints of a subsegment [Figure 6.1(b)]. These charges produce singular local fields, which do not exist in reality. Therefore, the PWC approximation yields inferior accuracy and efficiency compared with other

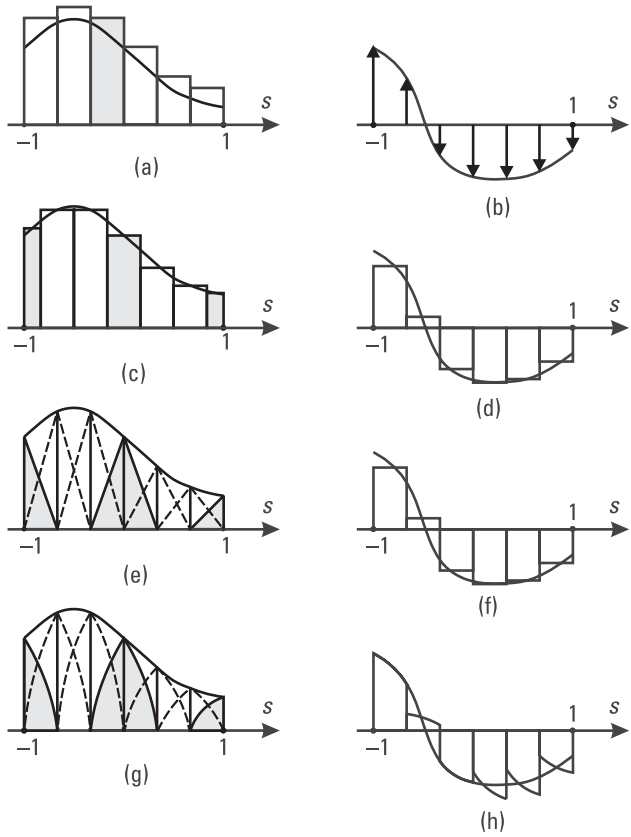

Figure 6.1 Subdomain approximation of the current along a generalized wire: (a) and (c) PWC, (e) PWL, and (g) PWS. The corresponding charges are given in (b), (d), (f), and (h), respectively.

methods [2]. For example, as many as 126 unknowns are needed for one-wavelength helical antenna [3].

The analysis based on pulse basis functions for the current distribution can be improved if the charge distribution is also approximated by a PWC function, although the two distributions do not strictly satisfy the continuity equation [2, 4]. The pulses for the charges are interleaved with the pulses for the current. The endpoints for the charge pulses are defined by s_i', $i = 0, \ldots, n$, such that $s_i \leq s_i' \leq s_{i+1}$. The amplitude of the charge pulses located between s_i' and s_{i+1}' is then determined from (6.2), where the derivative of the current is obtained using the FD approximation [e.g., see (2.86)]. We usually adopt that $s_i = -1 + 2i/(n+1)$ and $s_i' = -1 + (2i+1)/(n+1)$. In that case, all current and charge pulses have the same width (in terms of the local coordinate s), except the end current pulses, whose width is halved [Figure 6.1(c) and 6.1(d)]. The combination of the PWC approximations for the currents and charges results in a relatively accurate and efficient analysis. Acceptable results for long wire scatterers are obtained with 10 unknowns per wavelength. Few times more unknowns are needed for antenna problems (e.g., 32 segments for a half-wavelength dipole) [4].

The PWL approximation is obtained if each basis function is defined in the form of an increasing and a decreasing linear function along two neighboring subsegments; that is,

$$f_i^n(s) = \begin{cases} \dfrac{s - s_{i-1}}{s_i - s_{i-1}} & s_{i-1} \leq s \leq s_i \\ \dfrac{s_{i+1} - s}{s_{i+1} - s_i} & s_i \leq s \leq s_{i+1} \end{cases} \qquad i = 1, \ldots, n-1 \tag{6.4}$$

These basis functions are called triangle or rooftop basis functions. The simplest subdivision is when the points s_i are uniformly distributed [Figure 6.1(e)]. The corresponding charges are represented by pulse functions [Figure 6.1(f)]. The basis functions for $i = 0$ and $i = n$ are defined only on one subsegment [Figure 6.1(e)]. They enable satisfaction of the continuity equation [more precisely, *Kirchhoff's current law* (KCL)] for currents at wire ends and wire junctions. The zero current is forced at a free wire end by omitting the corresponding end basis function. At the junction of two wires, the continuity equation is forced by equating the unknown coefficients of the corresponding end basis functions (i.e., combining two end basis functions into one triangle basis function). At a junction of m wires, the continuity equation is satisfied by overlapping $m-1$ triangle basis functions (Figure 6.2) [5]. Flexible expansions for currents along wires are thus obtained, halving the number of unknowns

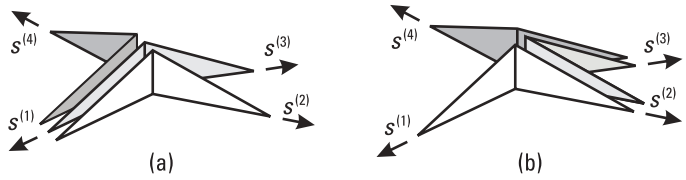

Figure 6.2 Satisfaction of continuity equation at multiple junction of wires by overlapping triangle basis functions: (a) basic way and (b) alternative way.

when compared with the PWC approximation of currents and charges [4]. Complicated structures have been efficiently analyzed using this technique, such as dipole arrays [6], log-periodic antennas [7], and so forth.

The PWS approximation is obtained if basis functions are defined in the form of an increasing and a decreasing sine function along two neighboring subsegments [8]; that is,

$$f_i^n(s) = \begin{cases} \dfrac{\sin[\beta L(s - s_{i-1})]}{\sin[\beta L(s_i - s_{i-1})]} & s_{i-1} \le s \le s_i \\ \dfrac{\sin[\beta L(s_{i+1} - s)]}{\sin[\beta L(s_{i+1} - s_i)]} & s_i \le s \le s_{i+1} \end{cases} \qquad i = 1, \ldots, n-1 \tag{6.5}$$

where β is the phase coefficient in the surrounding medium, L is the half-length of the given wire segment, and the segment is subdivided in the same way as for triangle basis functions. These basis functions are called sine doublets or sine dipoles. For uniformly distributed points s_i, the sine dipoles and the corresponding charges are shown in Figure 6.1(g) and 6.1(h). The end basis functions defined along the first and last subsegments, respectively, satisfy the continuity equation for currents at wire ends and wire junctions in the same way as explained for triangle basis functions [8]. The PWS approximation is more efficient than the PWL one. Acceptable results can be obtained with only four unknowns per wavelength of long wires. This can be explained by the fact that the dominant part of the current distribution along thin wires is sinusoidal. However, an accurate analysis requires more unknowns. For example, an analysis of a quarter-wavelength monopole used 10 unknowns [9]. Efficiency of the method is also enabled by explicit evaluation of the field integrals, whereas the potential and impedance integrals are expressed in terms of special functions [10, 11]. Accurate results for complicated structures, like Yagi-Uda and log-periodic dipole arrays were reported in [12].

The triangle and sine basis functions can alternatively be considered as two-term interpolation functions for current between two neighboring nodes.

Some authors use three-term expansions, which enable interpolation for current involving three neighboring nodes [13–16]. At two neighboring segments, which have the ith node in common, three-term algebraic and sinusoidal expansions are written as

$$I_i(s) = A_i + B_i s + C_i s^2 \tag{6.6}$$

$$I_i(s) = A_i + B_i \sin(\beta s) + C_i \cos(\beta s) \qquad s_{i-1} < s < s_{i+1}$$

The coefficients A_i, B_i, and C_i are determined from the conditions that $I(s)$ equals I_{i-1}, I_i, and I_{i+1}, which represent the values of the current at nodes $i - 1$, i, and $i + 1$, respectively. The three-term basis functions easily satisfy the continuity equation for currents at wire ends and wire junctions. However, the charge distribution along wires is still discontinuous. When compared with two-term expansions (triangles and sine doublets), three-term expansions produce better approximations. However, the order of approximation cannot be increased in steps of one, but in steps of two. Since the minimal number of unknowns required for the analysis is proportional to electrical lengths of wires, it is desirable that the order of approximation can be increased in steps of one. The highest efficiency can be achieved if a unique expansion scheme is adopted for all wires and the order of the expansion is proportional to the electrical length of the corresponding wire.

For example, consider a quarter-wave monopole antenna above a perfectly conducting ground plane, base-driven by a TEM magnetic current frill [17]. Figure 6.3 shows the relative error in admittance of the antenna (h/a = 100,

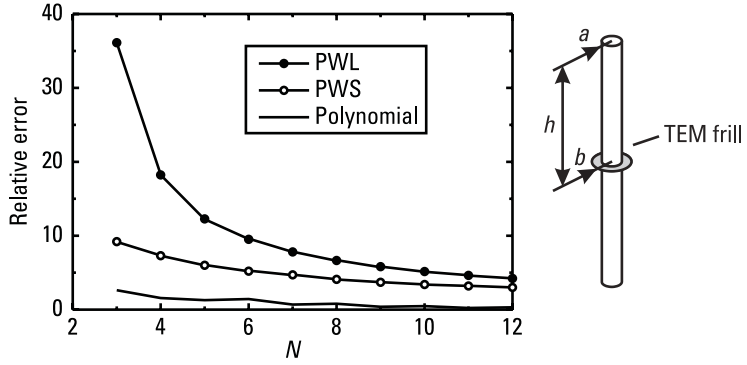

Figure 6.3 Relative error in the admittance of a quarter-wave monopole above a perfectly conducting ground plane (h/a = 100, b/a = 2.3) versus the number of unknowns in the analysis for various types of basis functions. The symmetrical dipole equivalent to the antenna is shown in the inset. (After: [17].)

$b/a = 2.3$) versus the number of unknowns used in the analysis. The relative error is defined as $E_r = |Y - Y_0|/|Y_0|$, where Y_0 is the "exact" value obtained by several methods with a sufficiently high degree of the current approximation. The results are obtained by three variants of the Galerkin method. It is seen that the entire-domain approximation is much more accurate than the PWS approximation, and the PWS approximation is slightly more accurate than the PWL approximation. By increasing the number of unknowns, both subdomain approximations result in about the same accuracy. This is explained by the fact that for electrically shorter segments, the sine dipoles become very similar to the triangle dipoles.

6.1.2 Entire-Domain Basis Functions

To obtain the entire-domain approximation of currents along wires, the basis functions in (6.1) are defined along the entire wire segment. These functions should be mutually independent. Physical distributions of currents and charges along wires are smooth functions. Hence, it is desirable that the basis functions be continuous and have continuous at least the first derivative. From the viewpoint of implementation, it is important that the basis functions be simple and easily evaluated.

The most common entire-domain approximations are polynomial and trigonometric (Fourier) functions. The basic polynomial and trigonometric expansions use simple power and cosine (sine) basis functions:

$$f_i(s) = s^i \qquad f_i(s) = \begin{cases} \cos(kCs) & i = 2k \\ \sin(kCs) & i = 2k-1 \end{cases} \qquad i = 0, \ldots, n \tag{6.7a-c}$$

where the constant C is an arbitrary real number ($0 < C < \pi$). Similar expansions are suggested in [18]. We call these expansions basic because any other basis function of the polynomial or trigonometric type can be obtained as a linear combination of power and cosine (sine) basis functions. The basis function of the ith order does not change with increasing the total expansion order n (i.e., the basic entire-domain expansions belong to the group of hierarchical expansions). Hence, we denote the basis functions as $f_i(s)$ instead of $f_i^n(s)$.

Trigonometric basis functions posses the orthogonality property if $C = \pi$. The integral of the product $f_i f_j$ from $s = -1$ to $s = 1$ is equal to zero for any indices i and j, except for $i = j$. This property may facilitate the solution of integral and differential equations in some cases. Some authors suggest that the orthogonal Legendre polynomials of the ith order be used instead of the power functions of the ith order [18]. However, in the thin-wire analysis, using

orthogonal basis functions does not facilitate the analysis because the evaluation of the potential, field, and impedance integrals does not contain a simple integration of $f_i f_j$ along the s-coordinate. Hence, none of these integrals can be annihilated due to the orthogonality property of the basis functions. On the other hand, implementation of Legendre polynomials is much more complicated than that for power functions. Hence, we give advantage to simple power functions over the Legendre polynomials or any other orthogonal polynomial (i.e., Chebyshev and Hermite).

Basic entire-domain expansions do not automatically satisfy the continuity equation at wire ends and junctions. For wires, the continuity equation reduces to the KCL. Numerical experiments show that satisfaction of the KCL significantly improves the accuracy and stability of results [2, 19, 20]. For an arbitrary junction of m wires, the KCL can be written in the form

$$\sum_{k=1}^{m} \pm I^{(k)}(s = \pm 1) = 0 \qquad I^{(k)}(s) = \sum_{i=0}^{n} a_i^{(k)} s^i \tag{6.8a,b}$$

where $I^{(k)}(s)$ approximates the current distribution along the kth wire in the junction. The + sign is taken for currents going out from the junction, while the − sign is taken for currents going into the junction. If the total number of basis functions is N and the number of wire junctions (including free wire ends) is N_{jun}, the number of equations that can be written according to the MoM is $N - N_{\text{jun}}$. The number of unknowns can be decreased by N_{jun} if the KCL is incorporated into the entire-domain expansion, when equations of the form (6.8) can be omitted. More importantly, this incorporation enables application of the Galerkin test procedure, which requires that the number of test functions be equal to the number of basis functions.

To satisfy the KCL at wire ends and junctions automatically, the basic entire-domain expansions should be rearranged. This can be done in a variety of ways. One way is to combine the starting basis functions to generate interpolation functions. For example, for the power functions, the interpolation functions are the Lagrange polynomials given by (5.3c) and the expansion is written in the form

$$I(s) = \sum_{i=1}^{n+1} I_i L_i^n(s) \qquad -1 \le s \le 1 \tag{6.9}$$

where I_i is the value of the current at the ith node. However, by increasing the order of approximation, a completely new set of interpolation functions is obtained. To cover all orders from $n = 0$ to $n = N$, we must operate with

$(N + 1)(N + 2)/2$ different basis functions instead of $N + 1$ functions for the basic entire-domain expansion. Hence, we give advantage to rearrangements that lead to hierarchical expansions, which automatically satisfy the KCL. Historically, the first hierarchical expansions automatically satisfied the KCL only for ends [21, 22]. The next section shows how the expansion is developed for arbitrary structures to include ends and single and multiple wire junctions [17].

6.1.3 Inclusion of KCL into Basis Functions

In order that the KCL can easily be satisfied at a wire end that is free or joined with other wires, the expansion along the wire should be rearranged so that all basis functions are equal to zero at the wire end, except for one function that has a unit value. This procedure will be illustrated on the example of the polynomial expansion [23]

$$I(s) = \sum_{i=0}^{n} a_i s^i \qquad -1 \le s \le 1 \tag{6.10}$$

The values of the current $I(s)$ at the beginning and end of the wire segment, I_1 and I_2, are expressed in terms of the unknown coefficients as

$$I_1 = I(-1) = a_0 - a_1 + \sum_{i=2}^{n} a_i (-1)^i \qquad I_2 = I(+1) = a_0 + a_1 + \sum_{i=2}^{n} a_i \tag{6.11}$$

The unknown coefficients a_0 and a_1 are expressed in terms of I_1, I_2, and other unknown coefficients. After replacing these expressions into (6.10) and simple rearrangements, we obtain

$$I(s) = I_1 N(s) + I_2 N(-s) + \sum_{i=2}^{n} a_i S_i(s) \qquad -1 \le s \le 1 \tag{6.12}$$

$$N(s) = \frac{1 - s}{2} \qquad S_i(s) = \begin{cases} s^i - 1 & i = 2k \\ s^i - s & i = 2k + 1 \end{cases} \tag{6.13a,b}$$

The resulting basis functions are equal to zero at wire ends, except that $N(s)$ has unit value at $s = -1$ and $N(-s)$ has unit value at $s = 1$. Hence, $N(s)$ and $N(-s)$ are called node basis functions. Obviously, only the node basis

functions are used to represent the current at a junction of wires. If a wire segment is not connected to other wires, its current is approximated only by the $S_i(s)$ basis functions. Hence, we call them segment basis functions.

Further, consider how the KCL is automatically satisfied at an arbitrary junction of m wires. For convenience, all wire segments are assumed to start at the junction with $s = -1$. All node basis functions that are used for satisfying the KCL are grouped into the junction expansion of the form

$$I_{\text{jun}} = \sum_{k=1}^{m} I^{(k)} N^{(k)}(s) \tag{6.14}$$

where $I^{(k)}$ is the current going out the junction and entering the kth wire. According to (6.8a), the current going out of the junction and entering the first wire, $I^{(1)}$, is expressed as

$$I^{(1)} = -\sum_{k=2}^{m} I^{(k)} \tag{6.15}$$

After substituting this expression into (6.14), the junction expansion is written in the form

$$I_{\text{jun}} = \sum_{k=2}^{m} I^{(k)} D^{(1,k)} \qquad D^{(i,j)} = N^{(j)} - N^{(i)} \tag{6.16}$$

where $D^{(i,j)}$ are new basis functions composed of two node basis functions. In what follows they will be called doublets (or dipoles). (Similarly, the segment basis functions defined along one segment are called singlets.) The doublet provides current flow between the ith and jth wires and automatically satisfies the continuity of the current at the junction. The junction expansion consists of $m - 1$ doublets. All of them overlap along the first wire in the junction. In the particular case where the starting basis functions are simple power functions, the doublets are triangle basis functions [Figure 6.2(a)].

Node basis functions can be combined in other ways to satisfy the continuity equation (6.8a). For example, the junction expansion can be written as

$$I_{\text{jun}} = \sum_{k=1}^{m-1} I^{(k,k+1)} D^{(k,k+1)} \qquad I^{(k,k+1)} = -\sum_{j=1}^{k} I^{(j)} \tag{6.17}$$

where $I^{(k,k+1)}$ is the intensity at the junction of the current flowing from wire k to wire $(k + 1)$. Again, the junction expansion consists of $m - 1$ overlapping

doublets. However, in this case, at most two doublets overlap along each wire, as shown for the triangle basis functions in Figure 6.2(b).

Thus, a flexible approximation of currents is obtained. The current distribution over a complex structure is approximated by segment expansions consisting of singlets and junction expansions consisting of doublets. When all wires in the structure are short enough in the electrical sense, the lowest order of approximation suffices. In that case, all singlets are omitted. By increasing the electrical length of wires, the order of approximation is increased, and the singlets are included into the approximation. For example, consider a chain of four aligned and interconnected wires. Figure 6.4 shows doublets and quadratic and cubic singlets of the polynomial type along the chain.

To illustrate the efficiency of the proposed entire-domain approximation of the polynomial type, let us consider a wire scatterer of length $l = 10.5\lambda$ and radius $l/a = 10^4$ excited by a normally incident plane wave, whose electric field vector is collinear with the wire axis. The scatterer is modeled by m segments of equal length. The same order of approximation, n, is used for currents along these segments. The number of unknowns used in the analysis is $N = mn - 1$. The analysis is performed for various values of m and n. The estimation of the solution accuracy is based on the relative error defined as

$$E_{\text{rms}} = \sqrt{\frac{\int_C |I - I_0|^2 \, dl}{\int_C |I_0|^2 \, dl}} \tag{6.18}$$

where I is the approximate current distribution along the wire scatterer and I_0 is the current distribution along the wire scatterer that is considered to be an accurate reference. This reference is obtained by increasing the number of segments above a value when the approximate current distribution practically does not change.

Figure 6.5 shows E_{rms} (solid curve) and the analysis time (dashed curves),[2] versus the number of unknowns, N. Each curve is obtained by using the same order of approximation, n. The increase of N is achieved by increasing the number of segments, m. The curves obtained for $n = 1$ correspond to the subdomain PWL approximation. The E_{rms} curves for $n = 6$, 7, and 8 are not drawn, being very close to the curve for $n = 9$. Analysis time curves for $n = 4, \ldots, 9$ are not drawn, being very close to the curve for $n = 3$. Increasing the

2. All results are obtained by the WIPL program running on IBM AT-486/33 MHz [24].

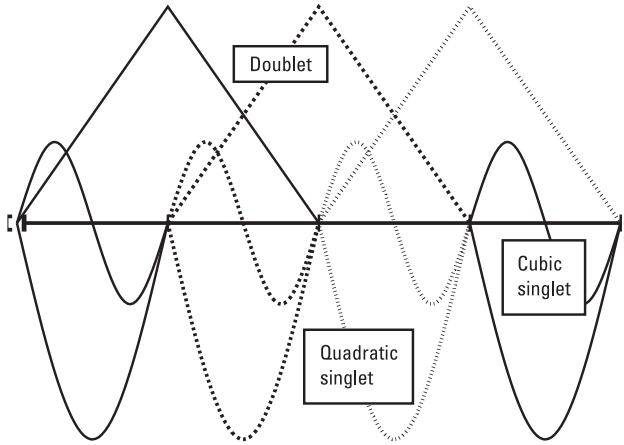

Figure 6.4 Doublets and quadratic and cubic singlets along the chain of four interconnected wires.

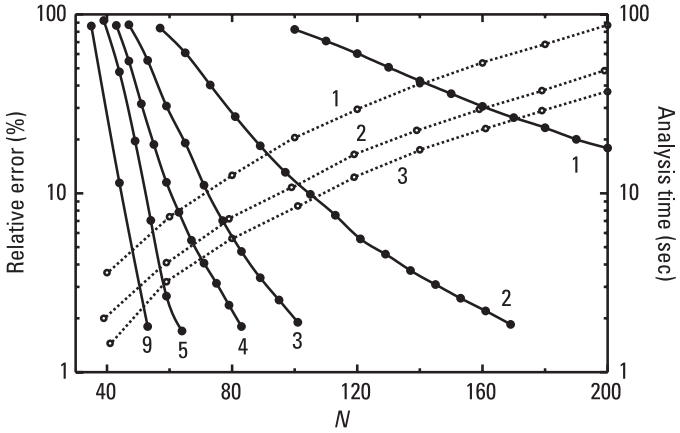

Figure 6.5 Relative error, E_{rms}, (solid curve) and analysis time (dashed curve) for thin electrically long scatterer ($l = 10.5\lambda$, $l/a = 10^4$) in terms of number of unknowns, N, for various orders of polynomial expansions, n. (After: [23].)

order of expansion, n, decreases the total number of unknowns, N, required to obtain a given relative error. This phenomenon is more pronounced for lower relative errors (higher accuracy). For $n > 5$, the rate of decrease of N slows down. For the same number of unknowns, the time used for a subdomain analysis is greater than the time used for an entire-domain analysis. For a given relative error, the entire-domain approximations require three to five times fewer unknowns and 30 to 100 times shorter analysis time than the subdomain

approximations. In general, acceptable results for long wires can be obtained using entire-domain approximations with only three to four unknowns per wavelength.

Numerical experiments show that stable results can be obtained for high-order approximations, provided the integrals involved in the analysis are evaluated with sufficient accuracy. This accuracy is limited by the number of significant digits used in computations. For example, if the single precision is used (seven significant digits), stable results are achieved if the order of the polynomial expansion does not exceed $n = 10$. If more than 10 unknowns are needed for the required accuracy, the wire segment must be divided into two or more shorter subsegments onto which the expansions of order not exceeding 10 are applied.

6.1.4 Combined Polynomial and Trigonometric Expansions

The trigonometric basis functions are more time consuming to evaluate and more complicated to implement than the polynomial basis functions. Hence, using trigonometric expansions is justified only if they require fewer unknowns than polynomial expansions.

In particular, consider a trigonometric expansion whose basis functions are defined by (6.7b) such that $C = \beta L = 2\pi L/\lambda$, where β is the phase coefficient and $2L$ is the length of the wire segment. For $k = 1$, the basis functions (6.7b) are $\cos(\beta Ls)$ and $\sin(\beta Ls)$. A linear combination of these functions represents the solution of the homogeneous part of the equation for the current distribution along a straight, infinitely thin wire. Obviously, these two functions represent the natural dominant part of the current distribution along arbitrarily shaped thin wires. However, to provide $C \leq \pi$, the segment length is limited to one wavelength ($L \leq \lambda/2$). We can apply (6.7b) to arbitrarily long wires (keeping $C \leq \pi$ by taking a smaller β), but the approximating potential of the expansion is decreased because the basis functions for $k = 1$ do not represent the dominant part of the current distribution any more.

The question is how to use functions $\cos(\beta Ls)$ and $\sin(\beta Ls)$ for segments longer than one wavelength. One possibility is to combine these functions with the power functions (6.7a). Since the leading terms in the McLorain expansions of $\cos(\beta Ls)$ and $\sin(\beta Ls)$ are the constant and linear function, respectively, the constant and linear functions can be omitted from the combined expansion. However, for specific values of β and L, functions $\cos(\beta Ls)$ and $\sin(\beta Ls)$ are equal to zero at wire ends; therefore, they cannot be considered as functions f_0 and f_1 in the entire-domain expansion. Hence, the basis functions of the combined trigonometric and polynomial expansion are initially adopted in the form

$$f_i(s) = s^{i+2} \quad (i = 0, \ldots, n-2) \tag{6.19}$$
$$f_{n-1}(s) = \sin(\beta L s) \qquad f_n(s) = \cos(\beta L s)$$

The form of the node and segment basis functions is different for different values of n:

$$N(s) = \begin{cases} \dfrac{\sin[\beta L(1-s)]}{\sin(2\beta L)} & n = 2 \\ \dfrac{s^2}{2} - \dfrac{\sin(\beta L s)}{2\sin(\beta L)} & n = 3 \\ \dfrac{s^2 - s^3}{2} & n \geq 4 \end{cases} \tag{6.20a}$$

$$S_i(s) = \begin{cases} s^i - s^2 & i = 2, 4, \ldots \\ s^i - s^3 & i = 3, 5, \ldots \\ \sin(\beta L s) - s^3 \sin(\beta L) & i = n-1 \\ \cos(\beta L s) - s^2 \cos(\beta L) & i = n \end{cases} \tag{6.20b}$$

Since $\sin(2\beta L) = 0$ for $2L = \lambda/2$, we can apply $n = 2$ for wire lengths up to half a wavelength. Since $\sin(\beta L) = 0$ for $2L = \lambda$, we can apply $n = 3$ for wire lengths up to one wavelength. There is no restriction to using higher orders ($n \geq 4$). The doublets made of the node basis functions for $n = 2$ are of the same form as the sine doublets (6.5) (i.e., for $n = 2$, the combined trigonometric and polynomial approximation reduces to the PWS approximation).

For electrically long wires, the combined expansion permits accurate current representation by a fairly small number of terms. Acceptable results are obtained with only 1 or 2 unknowns per wavelength of long wires, as shown by the following example. Consider an electrically long monopole wire antenna ($a = 0.021\lambda$ and $b/a = 2.3$) whose dipole counterpart is sketched in the inset of Figure 6.6. Figure 6.6 shows the admittance of the antenna versus its electrical length. The results obtained with the entire-domain polynomial and combined trigonometric and polynomial expansions, with $N = 10$ unknowns in both cases (one of which is used to model the end effect), are compared with the experimental results [25]. The results obtained with the polynomial expansions exhibit good agreement with experimental results up to about $\beta h = 15$, while those obtained with the combined polynomial/trigonometric expansion are accurate for the entire range of βh considered.

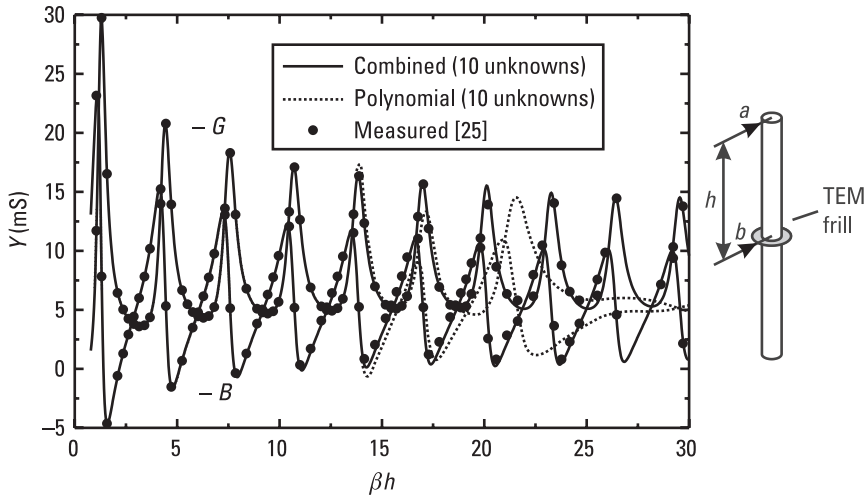

Figure 6.6 Admittance of electrically long monopole wire antenna ($a = 0.021\lambda$ and $b/a = 2.3$) versus its electrical length. (Dipole equivalent of the antenna is shown in the inset.) (After: [17].)

6.1.5 Quasistatic Treatment of Wire Ends and Interconnections

Wire ends and junctions represent discontinuities of wire structures. These discontinuities occupy electrically small regions. Hence, the charge distribution at these discontinuities can be determined by a quasistatic analysis. Inclusion of the quasistatic behavior of charges into basis functions may contribute significantly to the overall accuracy of the analysis [2, 17, 26]. It is difficult to model such behavior precisely. However, the main effect is achieved if we establish relations for the charges per unit length at wire ends and junctions. The most important relations are given in Table 6.1.

Table 6.1
The Most Important Quasistatic Relations for the Charges Per Unit Length at Wire Ends and Junctions

Type of Wire Junction (End)	Quasistatic Relation for Per-Unit-Length Charges
Cone apex [26, 27]	$Q'^{(1)} = 0$
Coaxial line opening	$Q'^{(1)} = 2\pi\epsilon U/\ln(b/a)$
Two wires of the same radius	$Q'^{(1)} = Q'^{(2)}$
m thin wires of the same radius [28]	$Q'^{(1)} = \ldots = Q'^{(k)} = \ldots = Q'^{(m)}$

These relations are included into the basis functions in a similar way as the continuity equation for currents (see Section 6.1.3). First, define two new unknowns, I_1' and I_2', as

$$I_1' = \left.\frac{dI(s)}{ds}\right|_{s=s_1} \qquad I_2' = \left.\frac{dI(s)}{ds}\right|_{s=s_2} \tag{6.21a,b}$$

Then, two unknown coefficients in (6.12) are expressed in terms of I_1', I_2', and the remaining unknown coefficients. After rearranging, (6.12) is written as

$$\begin{aligned} I(s) = {} & I_1 N(s) + I_2 N(-s) + I_1' N_Q(s) - I_2' N_Q(-s) \\ & + \sum_{i=4}^{n} a_i S_i(s) \qquad -1 \le s \le 1 \end{aligned} \tag{6.22}$$

where the node basis functions for currents and charges and the segment basis functions are defined as

$$N(s) = \frac{2 - 3s + s^3}{4} \quad N_Q(s) = \frac{1 - s - s^2 + s^3}{4} \tag{6.23}$$

$$S_i(s) = \begin{cases} s^i - \dfrac{i}{2}s^2 + \dfrac{i-2}{2} & i = 2k \\ s^i - \dfrac{i-1}{2}s^3 + \dfrac{i-3}{2}s & i = 2k+1 \end{cases}$$

All the basis functions and their first derivatives are equal to zero at both wire ends, except that $N(s)$ and $N(-s)$ have a unit value at $s = -1$ and $s = 1$, respectively, and the first derivatives of $N_Q(s)$ and $-N_Q(-s)$ have a unit value at $s = -1$ and $s = 1$, respectively. Obviously, the segment basis functions do not influence the satisfaction of either the continuity equation or the quasistatic relations. To satisfy the continuity equation at junctions automatically, the node basis functions for currents are grouped into doublets as explained in Section 6.1.3. The node basis functions for charges are the only basis functions that take a part in the satisfaction of the quasistatic relations.

The quasistatic relation at a free conical (or flat) wire end is automatically satisfied by omitting the corresponding node basis function for charges (i.e., $N_Q(s)$ at $s = -1$). If a wire above a ground plane is excited by the TEM magnetic current frill at $s = -1$, the unknown coefficient I_1' is determined using

(6.2), (6.21a), and the quasistatic relation for coaxial line opening from Table 6.1.

The quasistatic relation at a junction of m wires is automatically satisfied in the following way. Suppose that the s-axes of all wires start at the junction with $s = -1$. First, all node basis functions that are used for the satisfaction of the quasistatic equation at this junction are grouped in the junction expansion (for charges) of the form

$$I'_{\text{jun}} = \sum_{k=1}^{m} I'^{(k)} N_Q^{(k)}(s) \tag{6.24}$$

where $I'^{(k)}$ is the first derivative of the current going out the junction and entering the kth wire. According to (6.2), $I'^{(k)}$ can be expressed in terms of the accompanied charges as $I'^{(k)} = -j\omega e_s^{(k)} Q'^{(k)}$. Finally, using the quasistatic relation from Table 6.1, the junction expansion is expressed in terms of one basis function only; that is,

$$I'_{\text{jun}} = -j\omega Q' M^{(1,m)} \qquad M^{(1,m)} = \sum_{k=1}^{m} e_s^{(k)} N_Q^{(k)}(s) \tag{6.25}$$

where the new basis function, $M^{(1,m)}$, is composed of m node basis functions for charges. This function will be called a multiplet (or multipole). This basis function alone satisfies the quasistatic relation at the multiple wire junction. The only unknown in the junction expansion for charges is Q'.

Thus a flexible approximation for currents is obtained. The current distribution over a complex structure is approximated by segment expansions consisting of singlets, junction expansions for currents consisting of doublets, and junction expansions for charges consisting of multiplets. When all wires in the structure are electrically short enough, the lowest-order approximation suffices. In that case, all singlets are omitted, and the approximation is reduced to a subdomain type consisting of doublets and multiplets. By increasing the electrical length of wires, the order of approximation is increased and the singlets are added to the approximation. For example, consider a chain of four aligned interconnected wires. Figure 6.7 shows doublets, multiplets, and singlets of the fourth and fifth order along the chain.

To illustrate the efficiency of the polynomial expansions that includes both the continuity equation and the quasistatic relation, consider the same electrically long wire scatterer as in Section 6.1.3. The scatterer is modeled again by m segments of equal length, and the same order of approximation, n, is used for currents along these segments. However, by including the quasistatic

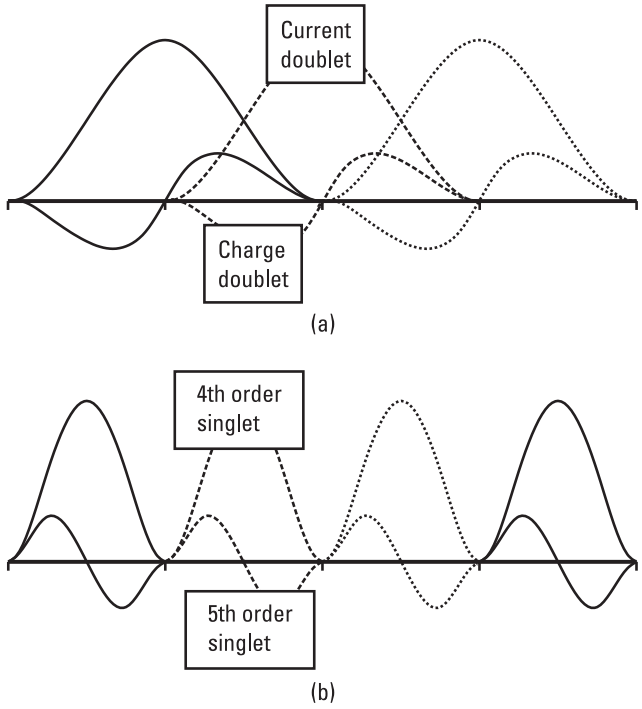

Figure 6.7 (a) Doublets and (b) singlets order along the chain of four interconnected wires.

relation in the polynomial expansions, in addition to the KCL, the number of unknowns is reduced to $N = m(n - 1)$. Figure 6.8 shows the relative error, E_{rms}, versus the number of unknowns, N. Each curve is obtained by using the same order of approximation, n, where the increase of N is achieved by increasing the number of segments, m. Curves for $n = 3$ and 9 are drawn when the quasistatic relation is included, and curves for $n = 1$, 3, and 9 are drawn when the quasistatic relation is not included. The inclusion of the quasistatic relation decreases the number of unknowns, and the accuracy does not depend much on the order of expansion, n. Acceptable results can be obtained using only two or three unknowns per wavelength of long wires.

6.1.6 Basis Functions in Terms of Simplex Coordinates

Usually, basis functions for wires are not expressed in terms of simplex coordinates. However, using these coordinates, we can easily derive basis function for triangles and tetrahedrons. To prepare the background for this generalization, we shall express the basis functions for wires in terms of the simplex coordinates.

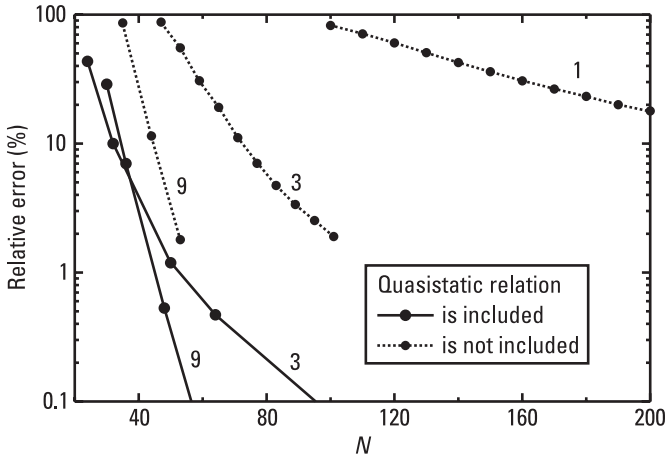

Figure 6.8 Relative error, E_{rms}, for thin electrically long scatterer ($l = 10.5\lambda$, $l/a = 10^4$) in terms of number of unknowns, N, for various orders of polynomial expansions, n, with and without inclusion of the quasistatic relation.

One way to obtain such a representation is to start from the basis functions given in terms of the parametric s-coordinate and to express the s-coordinate in terms of the simplex coordinates. For example, if $-\leq s \leq 1$, the s-coordinate and the simplex coordinates are related as

$$s = 1 - 2t_1 = 2t_2 - 1 = t_2 - t_1 \qquad t_1 = \frac{1-s}{2} \qquad t_2 = \frac{1+s}{2} \tag{6.26a-c}$$

In particular, since $t_1 + t_2 = 1$, (6.12) and (6.13) can be written as

$$I(t_1, t_2) = I_1 N(t_1) + I_2 N(t_2) + \sum_{i=2}^{n} a_i S_i(t_1, t_2) \qquad 0 \leq t_1, t_2 \leq 1 \tag{6.27}$$

$$N(t_1) = t_1 \qquad S_i(t_1, t_2) = \begin{cases} (t_2 - t_1)^i - 1 & i = 2k \\ (t_2 - t_1)^i - (t_2 - t_1) & i = 2k+1 \end{cases} \tag{6.28a,b}$$

Another way to obtain the basis functions in terms of the simplex coordinates, which automatically satisfy the continuity of currents at wire free ends and junctions, is to start from double power functions of t_1 and t_2; that is

from $t_1^p t_2^q$, where p and q are arbitrary integers. If t_1 and t_2 are expressed in terms of s according to (6.26b) and (6.26c), we obtain a polynomial in terms of s, whose order is $i = p + q$. Now, let us choose a set of basis functions $f_i(t_1, t_2) = t_1^p t_2^q$, $i = 0, \ldots, n$, such that $i = p + q$. Since the ith basis function is a polynomial of the ith order in terms of s, the set is complete, and its order is n. For the ith basis function, either p or q can be chosen arbitrarily. If both p and q are greater than zero, the basis function is equal to zero at both wire ends, and all basis functions for $i \geq 2$ represent the segments basis functions. In particular, it is convenient that the set of basis functions consists of, alternatively, odd and even functions with respect to the central point of the wire segment. Even functions are obtained by adopting $p = q = k$ [i.e., they have the form $(t_1 t_2)^k$]. However, there exists no such choice of integers p and q that makes $t_1^p t_2^q$ odd with respect to the central segment point. The simplest way to obtain the odd functions is to combine two functions $t_1^p t_2^q$ of the same order $i = 2k + 1$ [e.g., in the form $(t_1 t_2)^k (t_2 - t_1)$]. In that case, the first two basis functions are $f_0 = 1$ and $f_1 = t_2 - t_1$. Having in mind that $t_1 + t_2 = 1$, the zeroth basis function can be written as $f_0 = t_1 + t_2$. Now, f_0 and f_1 are easily combined into the node basis functions given by (6.28a) (i.e., $(f_0 + f_1)/2 = t_1$ and $(f_0 - f_1)/2 = t_2$). The resulting expansion has the same form as given by (6.27) and (6.28), except that the segment basis functions are now defined as

$$S_i(t_1, t_2) = \begin{cases} (t_1 t_2)^k & i = 2k \\ (t_1 t_2)^k (t_2 - t_1) & i = 2k + 1 \end{cases} \tag{6.29}$$

It can easily be shown that S_2 and S_3 in (6.29) are equal to $-S_2$ and $-S_3$ in (6.28b), respectively. For example, let us consider a chain of four aligned interconnected wires. Doublets, quadratic, and cubic singlets look the same as shown in Figure 6.4, except that the singlets should be taken with a minus sign. Differences occur for singlets of higher orders.

6.2 Approximation of Currents over Generalized Quadrilaterals

As elaborated in Chapter 5, generalized quadrilaterals are used for modeling metallic and dielectric surfaces accompanied by surface electric and equivalent (electric and magnetic) currents, respectively. Both types of currents, electric and magnetic, are approximated in the same way. Hence, in what follows, approximations of currents will be given for electric currents only. A generalized

quadrilateral is defined in a local ps-coordinate system, so it is natural to expand the currents in the same coordinate system into p- and s-components as

$$\mathbf{J}_s(p, s) = J_{sp}(p, s)\mathbf{i}_p(p, s) + J_{ss}(p, s)\mathbf{i}_s(p, s) \tag{6.30}$$

where $\mathbf{i}_p$ and $\mathbf{i}_s$ are unit vectors along p- and s-coordinate lines (see Figure 5.8a). However, the p-current component can be treated as the s-current component defined over the same generalized quadrilateral with interchanged p- and s-coordinates. Thus, in general, the distribution of surface currents can be represented as a sum of s-components defined over generalized quadrilaterals that overlap or are interconnected. Hence, in what follows, $\mathbf{J}_s$ will designate only the s-component of the current. The density of the corresponding surface charges is determined as

$$\rho_s(p, s) = \frac{j}{\omega}\operatorname{div}_s \mathbf{J}_s = \frac{j}{\omega|\mathbf{a}_p \times \mathbf{a}_s|}\frac{\partial}{\partial s}\left\{J_s(p, s)\frac{|\mathbf{a}_p \times \mathbf{a}_s|}{|\mathbf{a}_s|}\right\} \tag{6.31}$$

where ω is the angular frequency, and $\mathbf{a}_p$ and $\mathbf{a}_s$ are unitary vectors of the generalized quadrilateral (see Section 5.2.2).

The definition of the basis functions for surface currents over generalized quadrilaterals is similar to that for currents along generalized wires (see Section 6.1). In Section 6.2.1 we start with the subdomain approximation of surface currents. However, while for electrically thin wires it was possible to consider the current distribution as a function of the longitudinal wire coordinate only, this is not the case for generalized quadrilaterals. This complicates significantly the equations expressing the continuity of currents along the interconnections and requires special attention. In that sense, we consider approximate and exact formulations of surface doublets in Section 6.2.2 and generalized rooftop basis functions (exact formulation) in Section 6.2.3. Starting from the entire-domain approximation given in Section 6.2.4, Sections 6.2.5 and 6.2.9 show how the continuity equation and quasistatic relations can be included into the basis functions. In particular, we consider single and multiple metallic junctions, single and multiple dielectric junctions, and composite metallic and dielectric junctions in Sections 6.2.6, 6.2.7, and 6.2.8, respectively. Finally, Section 6.2.10 presents a square scatterer benchmark for the validation of the electromagnetic modeling of plate structures.

6.2.1 Subdomain Approximation

A given patch is subdivided (e.g., by a mesh of local p- and s-coordinate lines) into subpatches. The basis functions are defined over one or more neighboring

subpatches. Once the patch is subdivided into subpatches, it is convenient to adopt a local parametric coordinate system over each subpatch such that $-1 \leq p, s \leq 1$. The subdomain functions over all subpatches are then given by the same mathematical expressions in terms of the p- and s-coordinates. Hence, in what follows we shall consider that $-1 \leq p, s \leq 1$ for each subdomain basis function unless stated otherwise.

A basis function defined over one patch is called a singlet. The simplest singlet is a pulse. It is usually defined as the s-component of a current having a constant value over one subpatch. These subpatches are accompanied by line charges along two opposite edges perpendicular to the current flow at $s = -1$ and $s = 1$. Therefore, the pulses are not convenient for solving the EFIE. However, since charges do not enter into the MFIE, these basis functions can successfully be used for the solution [29]. This method requires about 200 unknowns per wavelength squared for metallic surfaces. In contrast to the EFIE, the MFIE can be applied only to closed bodies. In order for basis function to be applied for solution of the EFIE successfully, they must satisfy the continuity of currents at patch free ends and junctions. This condition can be fulfilled by doublets (basis functions defined over two neighboring patches).

Let us consider two neighboring patches in the form of generalized quadrilaterals, defined in local $ps^{(1)}$ and $ps^{(2)}$ parametric coordinate systems [Figure 6.9(a)]. The patches are interconnected along the common p-coordinate line

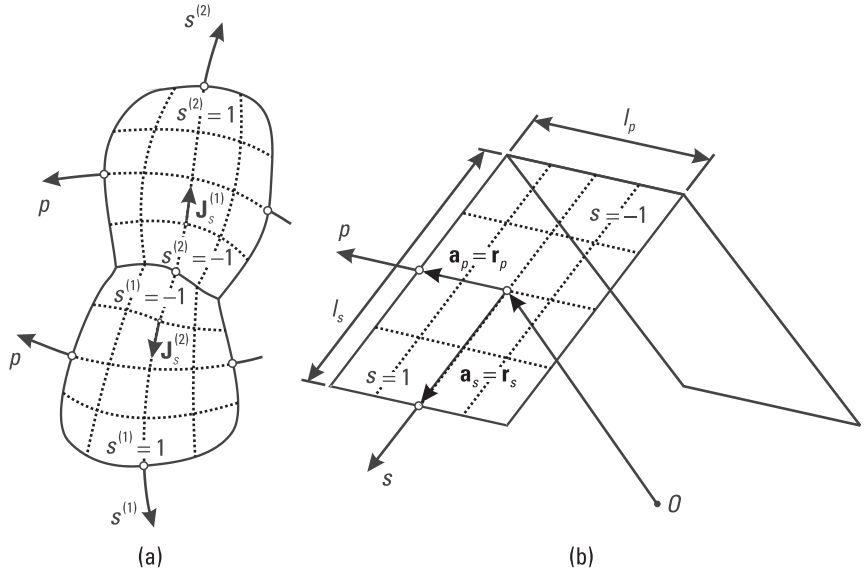

Figure 6.9 Doublets defined over (a) two generalized quadrilaterals and (b) two rectangles.

(i.e., the p-coordinate line for which $s^{(1)} = s^{(2)} = -1$). A doublet basis function can be written as

$$\mathbf{J}_s(\mathbf{r}) = \begin{cases} \mathbf{J}_s^{(1)}(p, s^{(1)}) & \mathbf{r} \in S^{(1)} \\ -\mathbf{J}_s^{(2)}(p, s^{(2)}) & \mathbf{r} \in S^{(2)} \end{cases} \tag{6.32}$$

where $\mathbf{J}_s^{(1)}(p, s^{(1)})$ and $\mathbf{J}_s^{(2)}(p, s^{(2)})$ are currents defined over the first and second patch, respectively, and $S^{(1)}$ and $S^{(2)}$ are surfaces of these patches. Usually, expressions for these currents are adopted in the same mathematical form.

Basically, doublets are defined over rectangular patches [Figure 6.9(b)]. Rectangular doublets are easily obtained by the generalization of wire (segment) doublets. The current density vector and the corresponding charge density over a rectangular doublet arm can be written as

$$\mathbf{J}_s(p, s) = \frac{N(s)}{l_p}\mathbf{i}_s \qquad \rho_s(p, s) = \frac{j}{\omega S}\frac{\partial N(s)}{\partial s} \tag{6.33}$$

where l_p is the width of the doublet along its p-coordinate line, and $N(s)$ is an arbitrary node basis function defined for wires [e.g., by (6.13a)]. In particular, the rooftop basis function and sine doublet are obtained if $N(s)$ is given by (6.13a) and (6.20a), respectively. The above current density vector is normalized so that the total current flowing through the interconnection has a unit magnitude. Such basis functions satisfy the continuity equation at the junction not only in the global, but also in the local way.

Rectangular doublets lead to relatively simple potential and field integrals that occur in the analysis (see Chapter 8). When a structure is modeled by rectangles, the distributions of currents and charges over each rectangle are modeled by one or more overlapping doublets. For sufficiently small doublets, these distributions should be approximately constant. Such distributions can be successfully modeled by doublets known as rooftop basis functions or sine doublets, which are obtained if $N(s)$ is adopted in the form (6.13a) and (6.20a). All good properties of a rectangular doublet, which are also desirable in the case of any other doublet types, are summarized in Table 6.2.

Owing to these good properties, rectangular doublets are shown to be very efficient in the analysis of structures that can be easily modeled by rectangles. For rooftop basis functions, acceptable results are obtained by using 6×6 patches and 72 unknowns per wavelength squared [30]. For sine doublets, acceptable results can be obtained with only 4×4 patches and 32 unknowns per wavelength squared [31, 32].

Table 6.2
List of Desired Properties of Doublets

Number	Short Name	Description
P1	Global continuity equation	Total current flowing out the first arm is equal to the total current flowing into the second arm.
P2	Local continuity equation	The normal current components are equal at each point of the junction of two arms.
P3	Simple potential (field) integrals	Potential (field) integrals do not contain Lame coefficients.
P4	Constant current distribution	Possibility of approximating constant current density vector.
P5	Constant charge distribution	Possibility of approximating constant charge distribution.

For example, consider a square scatterer of size $\lambda \times \lambda$ (Figure 6.10, inset). The scatterer is excited by a normally incident plane wave, which is polarized in the direction of the s-axis. The scatterer is subdivided into $m \times m$ patches. The analysis is performed by using $N = 2m(m - 1)$ unknowns. Figure 6.10 shows 3-D graphs of the magnitude of the s-current component over the scatterer obtained for $m = 3, 4, 5, 6, 7, 8$.

6.2.2 Approximate and Exact Formulation of Surface Doublets

Let us consider a doublet defined over two flat quadrilaterals (Figure 6.11). The expression for the current distribution of a flat quadrilateral doublet can

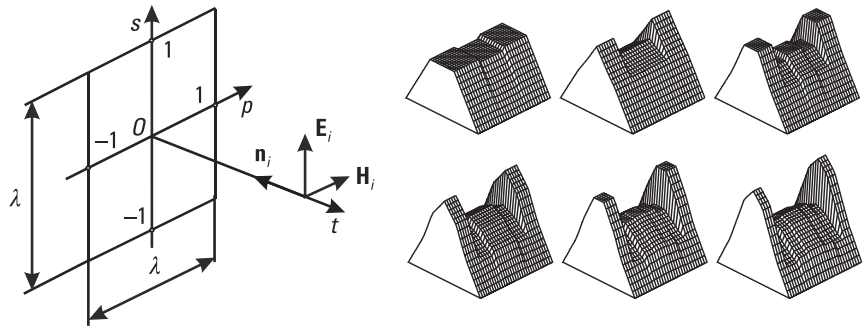

Figure 6.10 Magnitude of the s-current component (normalized with respect to the magnetic-field intensity of the incident plane wave) over the scatterer, shown in the inset, obtained for $m = 3, 4, 5, 6, 7, 8$.

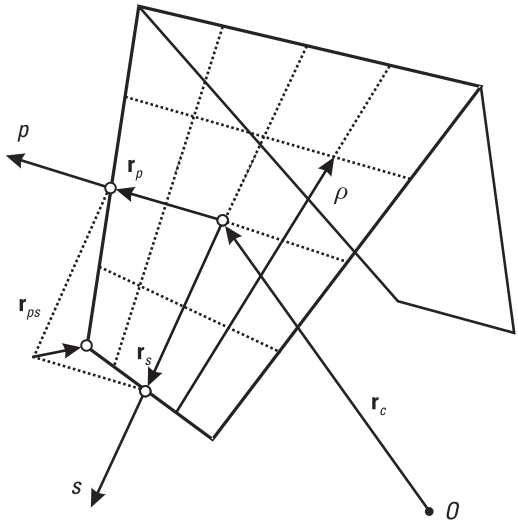

Figure 6.11 Doublets defined over flat quadrilaterals.

be written in the same form as for the rectangular doublet, except that l_p is a function of the s-coordinate and $\mathbf{i}_s$ is a function of the p-coordinate [33, 34]. Such a basis function does not satisfy any of the six desired properties. To satisfy the P1 property (global continuity equation), a normalization constant C is introduced by imposing that the total current of unit magnitude flows through the interconnection. By expressing l_p and $\mathbf{i}_s$ in terms of the unitary vectors, $\mathbf{a}_p$ and $\mathbf{a}_s$, the current distribution over one doublet arm is written in the form

$$\mathbf{J}_s(p, s) = C\frac{N(s)}{2|\mathbf{a}_p|\,|\mathbf{a}_s|}\mathbf{a}_s \qquad C^{-1} = \frac{1}{2}\int_{-1}^{1} \sin\alpha_{ps}(p, -1)\,dp \tag{6.34a,b}$$

where $\sin\alpha_{ps}(p, -1)$ is the angle between p- and s-coordinate lines along the interconnection. Since such a formulation of the doublet satisfies the continuity equation only globally, it is referred to as the approximate formulation.

In order to satisfy the property P2 (local continuity equation), the approximate formulation given by (6.34a) is slightly modified [35], resulting in

$$\mathbf{J}_s(p, s) = \frac{N(s)}{2|\mathbf{a}_p \times \mathbf{a}_s|}\mathbf{a}_s \qquad \rho_s(p, s) = \frac{j}{2\omega|\mathbf{a}_p \times \mathbf{a}_s|}\frac{dN(s)}{ds} \tag{6.35a,b}$$

Satisfaction of the P2 property implies satisfaction of the P1 property. Since such formulation of doublets also locally satisfies the continuity equation, it is referred to as the exact formulation.[3] Such formulation of doublets satisfies the P3 property (simple potential integrals). By proper adoption of the function $N(s)$, it is possible to satisfy the P4 property and, under some conditions, the P5 property.

6.2.3 Rooftop Basis Functions (Exact Formulation)

Rooftop basis functions (exact formulation) over generalized quadrilaterals are defined by (6.32), where the current density vector over one arm is defined by (6.35a), and the node function is adopted according to (6.13a). In what follows, the rooftop basis functions over flat quadrilaterals will be considered.

Let us define the position vector with respect to the free edge as $\boldsymbol{\rho} = -(1 - s)\mathbf{a}_s$ (Figure 6.11). A rooftop basis function is [36]

$$\mathbf{J}_s(p, s) = \frac{1}{4|\mathbf{a}_p \times \mathbf{a}_s|}\boldsymbol{\rho} \tag{6.36}$$

In particular, for a flat quadrilateral (represented as a bilinear surface), the vector $\mathbf{r}_{ps}$ can be expressed as a linear combination of the vectors $\mathbf{r}_p$ and $\mathbf{r}_s$ in the form

$$\mathbf{r}_{ps} = \alpha\mathbf{r}_p + \beta\mathbf{r}_s \qquad -1 < \alpha, \beta < 1 \tag{6.37}$$

Keeping in mind that

$$|\mathbf{a}_p \times \mathbf{a}_s| = (1 + \beta p + \alpha s)|\mathbf{r}_p \times \mathbf{r}_s| \qquad S = 4|\mathbf{r}_p \times \mathbf{r}_s| \tag{6.38}$$

where S is the surface area of a flat quadrilateral, the current and charge distributions over one doublet arm can be written as

$$\mathbf{J}_s(p, s) = \frac{1}{1 + \beta p + \alpha s}\frac{1}{S}\boldsymbol{\rho} \qquad \rho_s(p, s) = \frac{j}{\omega}\frac{1}{1 + \beta p + \alpha s}\frac{1}{S} \tag{6.39a,b}$$

Using (6.39a), it can be shown that the P4 property (constant current distribution) is satisfied [36]. It is seen from (6.39b) that the P5 property

3. Expression (6.35) can be applied not only to flat quadrilateral doublets, but also to doublets defined over generalized quadrilaterals.

(constant charge distribution) is satisfied only for rectangles and rhomboids (i.e., for $\alpha, \beta = 0$). For $\alpha, \beta \neq 0$, undesirable variations of the charge distribution are obtained. Since the quadrilaterals that produce greater deviations of the charge distribution are less desirable, the shape-quality factor, Q, can be defined as the ratio of the minimal and maximal charge densities; that is,

$$Q = \frac{\min\{\rho_s(p, s)\}}{\max\{\rho_s(p, s)\}} = \frac{1 - |\alpha| - |\beta|}{1 + |\alpha| + |\beta|} \tag{6.40}$$

To examine the influence of the properties P2 and P5 on the overall solution, consider a strip scatterer of length $l = \lambda/2$ and width $w = l/5$. The scatterer is excited normally by an incident plane wave, which is polarized along the scatterer length. The scatterer is subdivided into $n = 12$ patches, so that the shape-quality factor is $Q = 1$, $Q = 1/2$, and $Q = 1/10$ (Figure 6.12, inset). Figure 6.12 shows 3-D graphs of the current distributions and the corresponding charge distributions along these scatterers. The results are obtained by using quadrilateral rooftop basis functions (exact formulation). In particular, Figure 6.12(d) shows the same results as Figure 6.12(c), except that the order of the current approximation along each patch coordinate is increased by one. The current distribution practically does not depend on the quadrilateral shape, which is the consequence of the P2 property. The charge distributions corresponding to a lower Q show larger variations, as is expected according to (6.39b). Finally, Figure 6.12(d) shows that the undesirable local variations of the charge distribution can be avoided by using higher orders of approximation.

Acceptable results can be obtained even by using quadrilaterals of very a low Q (e.g., $Q = 1/100$). However, the number of unknowns needed for accurate analysis increases for lower Q, especially if the approximate formulation is used instead of the exact one. When the exact formulation is used, the optimal geometrical model should contain patches for which $Q > 1/2$ [36].

6.2.4 Entire-Domain Basis Functions

Keeping in mind the exact formulation of the surface doublets, it is convenient to adopt the initial entire-domain approximation for the s-current component over a generalized quadrilateral as [35, 37]

$$\mathbf{J}_s(p, s) = \sum_{i=0}^{n_p} \sum_{j=0}^{n_s} a_{ij} \mathbf{F}_{ij}(p, s) \tag{6.41a}$$

$$\mathbf{F}_{ij}(p, s) = \frac{f_i(p) g_j(s)}{2|\mathbf{a}_p \times \mathbf{a}_s|} \mathbf{a}_s \qquad -1 \leq p, s \leq 1 \tag{6.41b,c}$$

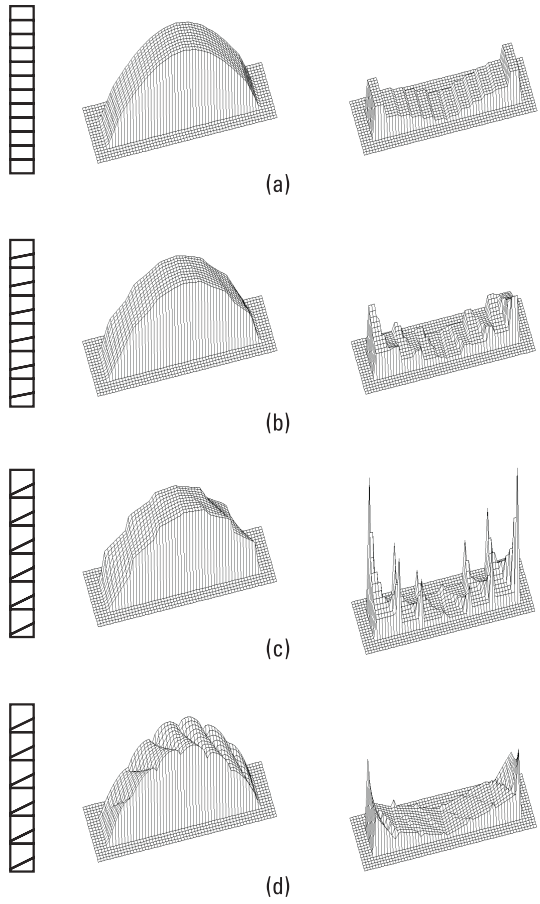

Figure 6.12 Current distribution and corresponding charge distribution over strip scatterers shown in the inset ($l = \lambda/2$, $w = l/5$). The scatterer is excited normally by an incident plane wave, which is polarized along the scatterer length. Rooftop basis functions are applied for (a) $Q = 1$, (b) $Q = 1/2$, and (c) $Q = 1/10$. Higher-order approximation is applied in case (d) $Q = 1/10$. (After: [36].)

where n_p and n_s are the orders of approximation along the p- and s-coordinates, a_{ij} are unknown coefficients to be determined, $\mathbf{F}_{ij}(p, s)$ are known vector basis functions, $f_i(p)$ and $g_j(s)$ are arbitrary known functions, and $\mathbf{a}_p$ and $\mathbf{a}_s$ are unitary vectors.

In general, the basis function of indices i and j depends on orders n_p and n_s [e.g., in $\mathbf{F}_{ij}$, $f_i(p)$ and $g_j(s)$ can be interpolation polynomials of the Lagrange type]. However, as noticed for the thin-wire analysis, this degree of freedom only complicates the analysis and does not make it more efficient. Hence, in

what follows we consider only hierarchical expansions (i.e., such basis functions that for given indices i and j do not depend on the overall orders of approximation n_p and n_s). Obviously, the functions $f_i(p)$, $i = 0, \ldots, n_p$ should be mutually independent, and functions $g_j(s)$, $j = 0, \ldots, n_s$ should also be mutually independent. Since the distributions of currents and charges over generalized quadrilaterals are smooth functions, it is desirable that the basis functions be continuous and that at least the first derivative be continuous. From the viewpoint of implementation, it is important that the basis functions be simple and easily evaluated. Similarly, as for the thin wire-analysis, the orthogonality properties of trigonometric basis functions and Legendre polynomials cannot be used to facilitate the analysis. Hence, we recommend that functions $f_i(p)$ and $g_j(s)$ be adopted in the form of simple power functions; that is,

$$f_i(p) = p^i \ (i = 0, \ldots, n_p) \qquad g_j(s) = s^j \ (j = 0, \ldots, n_s) \tag{6.42}$$

The basic entire-domain expansions do not automatically satisfy the continuity equation at patch free edges and junctions. Similarly, as for the thin-wire structures, the accuracy and efficiency of the analysis can be significantly improved if the basis functions automatically satisfy the continuity equation.

6.2.5 Inclusion of Continuity Equation into Basis Functions

In order that the continuity equation can be easily satisfied at patch free edges or junctions, the initial entire-domain expansion over a patch should be rearranged. The rearrangement procedure will be illustrated with the example of the polynomial expansion [35, 37], which is written here for convenience as

$$\mathbf{J}_s = \frac{\mathbf{a}_s}{2|\mathbf{a}_p \times \mathbf{a}_s|} \sum_{i=0}^{n_p} \left[\sum_{j=0}^{n_s} a_{ij} s^j \right] p^i \qquad -1 \le p, s \le 1 \tag{6.43}$$

Let us introduce new unknowns c_{i1} and c_{i2}, $i = 0, \ldots, n_p$, defined as

$$c_{i1} = \sum_{j=0}^{n_s} a_{ij}(-1)^j \qquad c_{i2} = -\sum_{j=0}^{n_s} a_{ij} \tag{6.44}$$

Starting from these equations, two unknown coefficients for any i, a_{i0}, and a_{i1}, are expressed in terms of other unknown coefficients, a_{ij}

($j = 2, \ldots, n_s$), c_{i1}, and c_{i2}. Substituting the resulting expressions for a_{i0} and a_{i1} into (6.43), the expression for $\mathbf{J}_s$ is rearranged in the form

$$\mathbf{J}_s(p, s) = \sum_{i=0}^{n_p} \left\{ c_{i1}\mathbf{E}_i(p, s) + c_{i2}\mathbf{E}_i(p, -s) + \sum_{j=2}^{n_s} a_{ij}\mathbf{P}_{ij}(p, s) \right\} \quad (6.45)$$

$$-1 \le p, s \le 1$$

$$\mathbf{E}_i(p, s) = \frac{\mathbf{a}_s}{2|\mathbf{a}_p \times \mathbf{a}_s|} p^i N(s) \quad (6.46)$$

$$\mathbf{P}_{ij}(p, s) = \frac{\mathbf{a}_s}{2|\mathbf{a}_p \times \mathbf{a}_s|} p^i S_j(s) \quad (j = 2, \ldots, n_s)$$

where $N(s)$ and $S_j(s)$ are the node and segment functions defined in the thin-wire analysis by (6.13).

All basis functions are equal to zero at the edges defined by $s = \pm 1$, except functions $\mathbf{E}_i(p, s)$ at the edge $s = -1$ and functions $\mathbf{E}_i(p, -s)$ at the edge $s = 1$. If the first (second) edge is free, the continuity equation is automatically satisfied by omitting functions $\mathbf{E}_i(p, s)[\mathbf{E}_i(p, -s)]$. If a surface patch is not connected to other patches at all, its current is approximated only by $\mathbf{P}_{ij}(p, s)$ basis functions. In what follows, such a current expansion (basis functions) will be called a patch expansion (patch basis functions). A few typical patch basis functions are sketched in Figure 6.13.

For example, consider again a square scatterer of size $\lambda \times \lambda$ excited by a normally incident plane wave as in Section 6.2.1. Figure 6.14 shows the magnitude of the s-component along the p-coordinate axis obtained by using a polynomial expansion of order $n = 8$, consisting of patch basis functions only ($N = 112$ unknowns) and rooftop basis functions defined over $M = 8 \times 8$ patches ($N = 112$ unknowns) [38]. The results are compared with an accurate benchmark solution (see Section 6.2.10). The polynomial basis functions produce a more accurate approximation than the rooftop basis functions.

Only $\mathbf{E}_i(p, s)$ and $\mathbf{E}_i(p, -s)$ take part in satisfying the continuity equation along the interconnected first and second edge, respectively. Therefore, they will be referred to as edge basis functions. All edge basis functions used to satisfy the continuity equation at a junction will be grouped in the so-called junction expansions. Hence, for complex structures, the current distribution is approximated by patch and junction expansions.

The edge basis functions $\mathbf{E}_i(p, s)$ and $\mathbf{E}_i(p, -s)$ have the same form with respect to the pertinent edge. Namely, if the s-coordinate is replaced by $-s$, the basis function $\mathbf{E}_i(p, s)$ is transformed into $\mathbf{E}_i(p, -s)$, and vice versa.

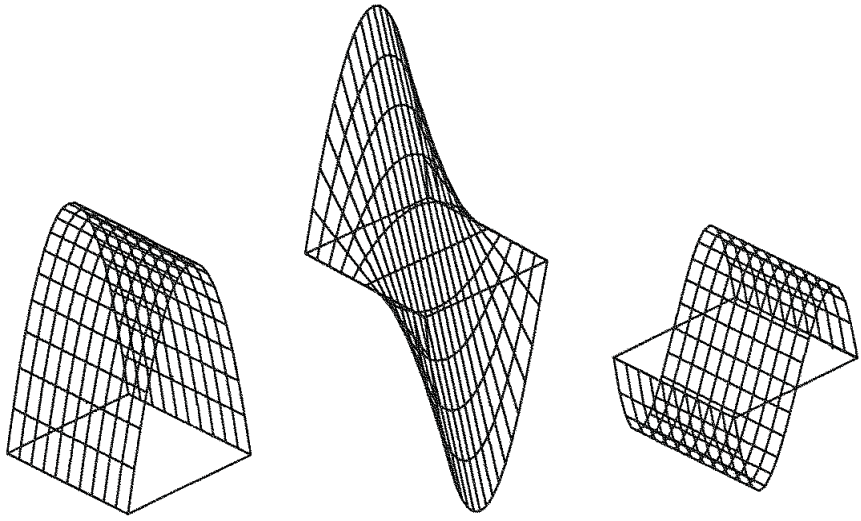

Figure 6.13 Patch basis functions. (After: [35].)

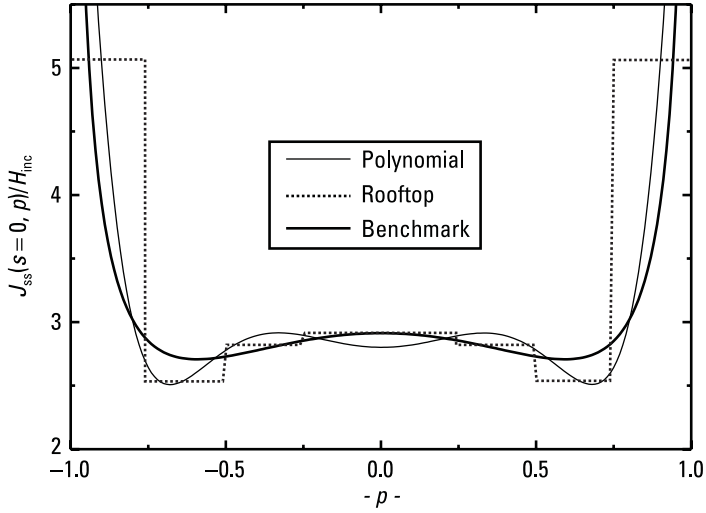

Figure 6.14 Magnitude of the s-component along the p-coordinate axis of a square scatterer (shown in the inset of Figure 6.10) of size $\lambda \times \lambda$. Results obtained by polynomial expansion of order $n = 8$ and rooftop basis functions defined over $M = 8 \times 8$ patches are compared with accurate benchmark solution. (After: [38].)

Hence, any junction can be treated as consisting of patches interconnected along the side $s = -1$. The corresponding junction expansion consists of edge basis functions of the form $\mathbf{E}_i(p, s)$ multiplied by unknown coefficients of the form c_{i1}. To simplify notation, the index 1 will be omitted from c_{i1}.

The normal component of the ith edge basis function going out of the junction can be evaluated as

$$\{\mathbf{E}_i(p, -1)\}_{\text{norm.}} = (-1)^i \left|\mathbf{i}_p(p, -1) \times \mathbf{E}_i(p, -1)\right| = \frac{p^i}{\left|\mathbf{a}_p(p, -1)\right|} \tag{6.47}$$

where $\mathbf{i}_p(p, -1)$ is the unit vector along the junction. If all interconnected generalized quadrilaterals are described by the same parametric equation along the junction, which is most often the case, they have the same p-unitary vectors along the junction; that is, the expressions $\left|\mathbf{a}_p(p, -1)\right|$ are equal for all interconnected generalized quadrilaterals. In that case, all edge basis functions of index i defined over the interconnected elements have the same normal component going out of the junction. Owing to this property, it is possible to satisfy exactly the continuity of the normal current component at the junction by a proper combination of edge basis functions. The continuity equation along this edge must be satisfied independently for each order i. The proper combinations can be found by using general rules that are different for three general classes of junctions: metallic junctions, dielectric junctions, and composite metallic and dielectric junctions.

6.2.6 Single and Multiple Metallic Junctions

Let us consider a junction of m open metallic surfaces, the cross section of which is shown in Figure 6.15 [35, 37]. Surfaces are designated so that the kth surface is followed by the $(k+1)$-th surface, and the mth surface is followed by the first surface. The junction expansion can be written in the form

$$\mathbf{J}_{\text{jun}} = \sum_{i=0}^{n_p} \left[\sum_{k=1}^{m} c_i^{(k)} \mathbf{E}_i^{(k)} \right] \tag{6.48}$$

where $\mathbf{E}_i^{(k)}$ are edge basis functions defined over the kth surface.

According to the continuity equation, the total current flowing out of the junction, per unit length of the junction, $J_{\text{out}}(p)$, should be equal to zero. Starting from (6.47) and (6.48), this condition can be written as

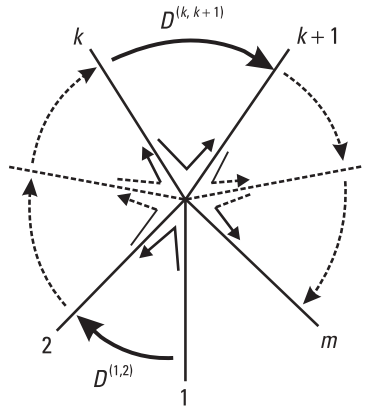

Figure 6.15 Cross section of a junction made of m metallic surfaces.

$$J_{\text{out}}(p) = \sum_{i=0}^{n_p} \left[\sum_{k=1}^{m} c_i^{(k)} \{\mathbf{E}_i^{(k)}(p, -1)\}_{\text{norm}} \right] \tag{6.49}$$

$$= \sum_{i=0}^{n_p} \left[\sum_{k=1}^{m} c_i^{(k)} \right] \frac{p^i}{|\mathbf{a}_p^{(k)}(p, -1)|} = 0$$

To force the above condition for each value of the p-coordinate, we must satisfy the set of equations

$$\sum_{k=1}^{m} c_i^{(k)} = 0 \qquad i = 0, \ldots, n_p \tag{6.50}$$

After rearranging (6.48) and including (6.50), we obtain

$$\mathbf{J}_{\text{jun}} = \sum_{i=0}^{n_p} \left\{ \sum_{k=1}^{m-1} d_i^{(k)} \mathbf{D}_i^{(k,k+1)} \right\} \qquad d_i^{(k)} = \sum_{j=1}^{k} - c_i^{(j)} \tag{6.51}$$

$$\mathbf{D}_i^{(j,k)} = \mathbf{E}_i^{(k)} - \mathbf{E}_i^{(j)} \tag{6.52}$$

where $\mathbf{D}_i^{(j,k)}$ are new basis functions composed of two edge basis functions. In what follows they will be called doublets (or dipoles). This junction expansion (for an arbitrary order i) is schematically shown in Figure 6.15.

For the lowest-order expansion given by (6.45) (i.e., for $n_p = 0$ and $n_s = 1$), the patch basis functions drop out, and the junction expansion consists only

of the doublets $\mathbf{D}_0^{(k,k+1)}$ ($k = 1, \ldots, m-1$). The doublet $\mathbf{D}_0^{(k,k+1)}$ actually represents a rooftop basis function (exact formulation), presented in Section 6.2.3. Considering that an edge is most often shared by two patches (i.e. $m = 2$), the minimal number of the p- and s-unknowns per surface patch is about $N_{\text{patch}} = 2$. This number is doubled if the edge basis functions are not combined into doublets. A few typical doublets of various orders are sketched in Figure 6.16.

For example, consider again a square scatterer of size $\lambda \times \lambda$ excited by a normally incident plane wave, as in Section 6.2.1. Figure 6.17 shows the relative error of the approximation of the current versus the number of unknowns, N. The relative error is evaluated according to (6.62) with respect to the benchmark solutions given in Section 6.2.10. Each curve is obtained by using the same order of approximation, $n = 1, 2, 3$, and 4, along each patch coordinate. The increase of N is achieved by increasing the number of patches, $M = m \times m \times m$, $m = 1, 2, 3, 4$. By increasing the order of the entire-domain expansion, n, the same accuracy can be achieved by decreasing the total number of unknowns, N. Good results for scattering are obtained if the relative error is less that 20%; that is, by using the expansion of fourth order, which means 32 unknowns per wavelength squared. The power of the entire-domain approximation is more pronounced for larger patches (e.g., 2×2 wavelengths squared), when acceptable results can be obtained with only 10 unknowns per wavelength squared.

6.2.7 Single and Multiple Dielectric Junctions

Consider a junction of m dielectric surfaces bounding m dielectric regions, whose cross section is shown in Figure 6.18 [37]. Surfaces and regions are designated in a manner that the kth region is followed by the $(k+1)$-th region,

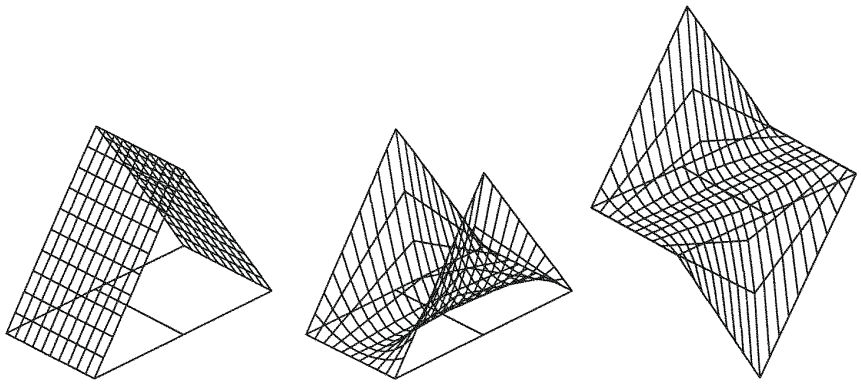

Figure 6.16 Doublet basis functions. (After: [35].)

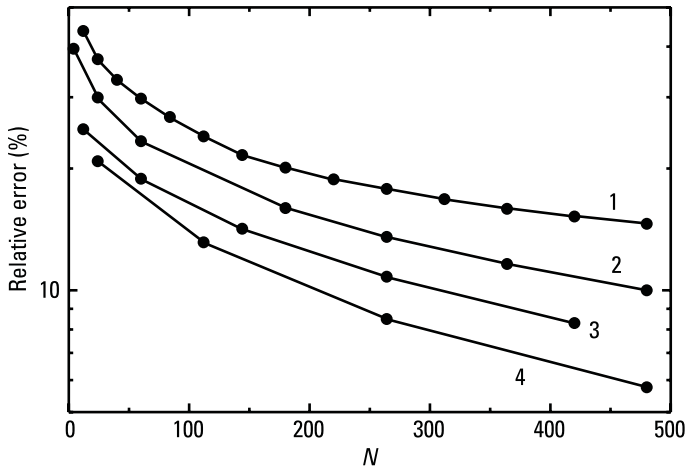

Figure 6.17 Relative error, E, for currents over a square scatterer of size $\lambda \times \lambda$ (shown in inset of Figure 6.10) in terms of number of unknowns, N, for various orders of approximation, n.

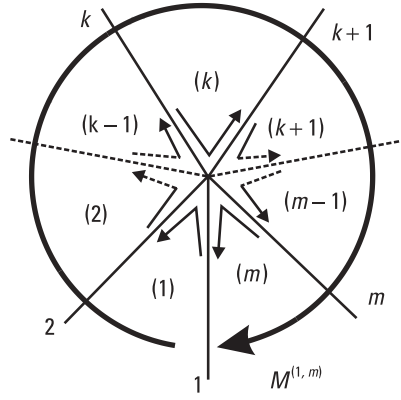

Figure 6.18 Cross section of a junction made of m dielectric surfaces.

and the kth region is bounded by the kth and the $(k + 1)$-th surface. [To simplify notation, the first surface (region) is also designated as the $(m + 1)$-th surface (region), and the mth region is also designated as the zeroth region.] The kth surface is the boundary surface between the $(k - 1)$-th and kth region.

Let $\mathbf{E}_i^{(k)}$ be an edge basis function defined along the kth surface in the $(k - 1)$-th region. The equivalent surface currents have the same magnitude, but opposite reference directions, on the opposite sides of the boundary surface between two regions. Hence, the edge basis function defined along the kth

surface in the kth region is $(-1)\mathbf{E}_i^{(k)}$. The electric junction current can be written as

$$J_{\text{jun}} = \sum_{i=0}^{n_p} \left\{ \sum_{k=1}^{m} [c_i^{(k+1)} \mathbf{E}_i^{(k+1)} - c_i^{(k)} \mathbf{E}_i^{(k)}] \right\} \tag{6.53}$$

where the expression given in brackets is defined over the boundary surfaces of the kth region. Hence, the summation with respect to k cannot be performed algebraically. This summation means that currents belonging to different regions are considered together. According to the continuity equation, the total current flowing out of the junction in the kth region per unit length of the junction, $J_{\text{out}}^{(k)}(p)$, should be equal to zero. Starting from (6.47) and (6.53), this condition can be written as

$$\begin{aligned} J_{\text{out}}^{(k)}(p) &= \sum_{i=0}^{n_p} [c_i^{(k+1)} \{\mathbf{E}_i^{(k+1)}(p, -1)\}_{\text{norm}} - c_i^{(k)} \{\mathbf{E}_i^{(k)}(p, -1)\}_{\text{norm}}] \qquad (6.54) \\ &= \sum_{i=0}^{n_p} [c_i^{(k+1)} - c_i^{(k)}] \frac{p^i}{|\mathbf{a}_p^{(k)}(p, -1)|} = 0 \end{aligned}$$

To satisfy the above condition for each value of the p-coordinate, the difference of the unknown coefficients given in the brackets should be equal to zero for each value of indices i and k, resulting in

$$c_i^{(1)} = c_i^{(2)} = \ldots = c_i^{(k)} = \ldots = c_i^{(m)} = c_i^{(m+1)} \tag{6.55}$$

After substituting (6.55) into (6.53), the electric junction current can be written in the form

$$\mathbf{J}_{\text{jun}} = \sum_{i=0}^{n_p} c_i^{(1)} \mathbf{M}_i^{(1,m)} \qquad \mathbf{M}_i^{(1,m)} = \sum_{k=1}^{m} \mathbf{D}_i^{(k,k+1)} \tag{6.56}$$

where $\mathbf{M}_i^{(1,m)}$, $i = 0, \ldots, n_p$ are new basis functions obtained as combinations of doublet basis functions, given by (6.52). These doublets belong to different regions. Since each basis function is defined over m interconnected elements, it will be referred to as a multiplet (or a multipole). The junction current is schematically shown in Figure 6.18.

For the lowest order of expansions given by (6.45) (i.e., for $n_p = 0$ and $n_s = 1$), patch expansions vanish and the junction expansion consists only of

one multiplet, $\mathbf{M}_0^{(1,m)}$. There is only one electric and one magnetic multiplet per junction. The simplest region bounded by flat quadrilaterals is a cube. Considering that an edge (i.e., a multiplet basis function) is most often shared by four cubes, the minimal number of unknowns per cube region is about $N_{\text{domain}} = 6$. This number is four times greater when edge basis functions are not combined into multiplets.

For example, let us consider a homogeneous lossless dielectric cube in a vacuum and excited by a plane electromagnetic wave [37]. The cube edge length is $a = \lambda_0/5 = 3\lambda_d/5$, where λ_d represents the wavelength in the cube dielectric of relative permittivity $\epsilon_r = 9$. Three cube edges coincide with the x-, y-, and z-axes. The electric field intensity of the plane wave is $\mathbf{E}_i = \exp(j\beta_0 z)\mathbf{i}_x$ V/m. The cube is analyzed in three ways (Figure 6.19, inset). First, the cube surface is modeled by $M = 6$ patches and the entire-domain approximation for currents is applied, resulting in $N = 96$ unknowns. Run time required for the analysis on a Pentium at 133 MHz is $T = 2.1$ seconds. When the cube surface is modeled by $M = 54$ patches and the subdomain approximation for currents (rooftop basis functions) is applied, it results in

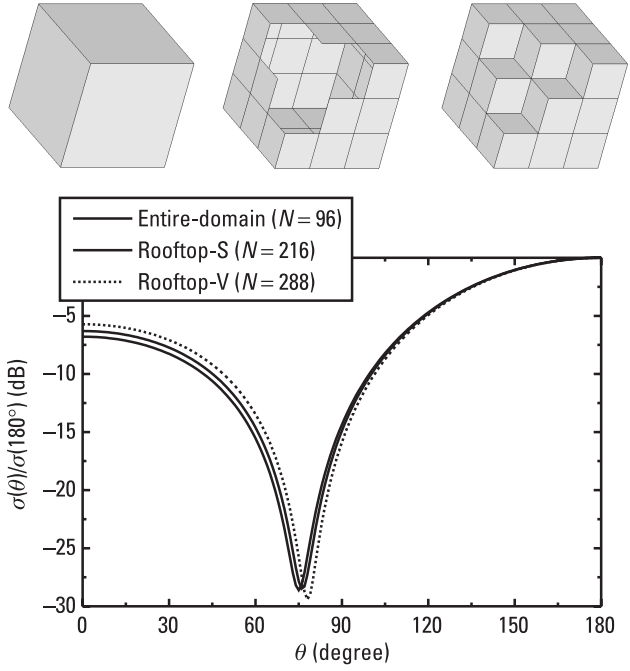

Figure 6.19 Bistatic RCS of the cube versus the angle θ. Entire-domain, rooftop-S, and rooftop-V approximations are applied to geometrical models, shown in the inset from left to right, respectively. (After: [37].)

$N = 216$ unknowns. The analysis time is $T = 17.1$ seconds. Finally, the cube is subdivided into $L = 3 \times 3 \times 3 = 27$ small cubes, each of which is modeled by $M = 6$ patches, resulting in a total of $M = 108$ patches. When the subdomain approximation for currents is applied, it yields $N = 288$ unknowns. The analysis time is $T = 18.7$ seconds. Since the surface rooftop basis functions are distributed over the cube surface in the second case, and over the cube volume in the third case, these two approximations will be designated rooftop-S and rooftop-V approximations, respectively. For the rooftop-V approximation, the number of patches is increased by 100% when compared with the rooftop-S approximation. However, owing to the definition of multiplet basis functions, the number of unknowns is increased by only 33%. Obviously, using the rooftop basis functions inside the homogeneous dielectric cube is unnecessary. However, the rooftop-V approximation is interesting because of its potential application in modeling inhomogeneous bodies.

Figure 6.19 shows the bistatic RCS in the plane $\phi = 0$ (xOz-plane) normalized with respect to its maximal value. The results obtained by the proposed method are compared with those based on the approximation of currents by $N = 596$ rooftop basis functions defined over triangles [39]. Good agreement between the results can be observed. Regarding the number of unknowns used in the analysis, the quadrilateral modeling and the entire-domain approximation are much more efficient than the triangular modeling and the subdomain approximation.

6.2.8 Composite Metallic and Dielectric Junctions

Let us consider an arbitrary junction of metallic and dielectric surfaces. The tangential component of electric field is equal to zero along the junction. Hence, there are no magnetic currents flowing through the junction, and the junction expansion consists of electric currents only. By using the previously mentioned classes of junctions, the following set of rules is established for the construction of electric junction expansions, consisting of edge basis functions of the same order along the p-coordinate [37]:

1. Current continuity between two neighboring surfaces is realized by a doublet basis function.
2. There are always two independent doublets placed on the opposite faces of an open metallic surface.
3. There is always only one doublet placed over a closed metallic surface.
4. There are always two doublets placed on the opposite faces of a dielectric surface, having the same magnitude and opposite reference directions.
5. Multiplets are always defined between two neighboring metallic surfaces, open or closed, with at least one dielectric surface between them.

The only exception to rule 2 occurs for junctions of open metallic surfaces, given in Figure 6.15. Namely, there is only one doublet belonging to the first and the mth surfaces. The second doublet, connecting these surfaces, can be represented as a linear combination of all other doublets at the junction and is therefore omitted.

A variety of composite junctions can occur. Three typical examples are used to illustrate the above rules [37]. First, let us consider a junction made of one metallic and three dielectric wedges whose cross section is shown in Figure 6.20(a). The junction expansion is given by (6.56) with $m = 3$.

Further, let us consider an open metallic surface immersed in a vacuum and connected with the metallic and dielectric wedges [Figure 6.20(b)]. The junction expansion is given by

$$\mathbf{J}_{\text{jun}} = \sum_{i=0}^{n_p} [c_i^{(1)} \mathbf{D}_i^{(1,2)} + c_i^{(2)} \mathbf{M}_i^{(2,3)}] \tag{6.57}$$

Finally, let us consider an open metallic surface protruding through the dielectric surface. Such a composite junction is modeled by two open metallic surfaces and two dielectric surfaces [Figure 6.20(c)]. The junction expansion is written in the form

$$\mathbf{J}_{\text{jun}} = \sum_{i=0}^{n_p} [c_i^{(1)} \mathbf{M}_i^{(1,2)} + c_i^{(3)} \mathbf{M}_i^{(3,4)}] \tag{6.58}$$

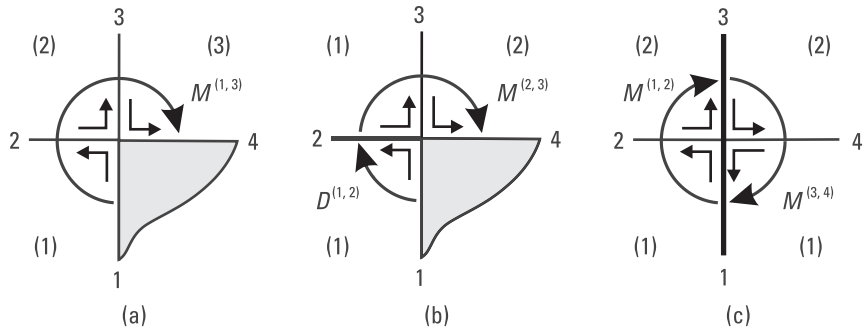

Figure 6.20 Cross sections of three typical composite metallic and dielectric junctions: (a) metallic wedge (shaded region) in vacuum (region 1) connected to two dielectric wedges (regions 2 and 3); (b) open metallic surface (surface 2) immersed in vacuum (region 1) connected to metallic wedge (shaded region) and dielectric wedge (region 2); and (c) open metallic surface (made of surfaces 1 and 3) protruding from vacuum (region 1) through the dielectric surface (region 2).

The multiplets $\mathbf{M}_i^{(1,2)}$ and $\mathbf{M}_i^{(3,4)}$ are different. The first consists of doublets $\mathbf{D}_i^{(1,2)}$ and $\mathbf{D}_i^{(2,3)}$, while the second consists of doublets $\mathbf{D}_i^{(3,4)}$ and $\mathbf{D}_i^{(4,1)}$. In general, a junction expansion for a composite metallic and dielectric structure is written in the form of a linear combination of doublets and multiplets.

6.2.9 Inclusion of Quasistatic Relation (Edge Effect) into Basis Functions

Let us consider the edge of a metallic wedge. The distributions of the surfaces currents and charges in the vicinity of the edge can be expressed as

$$\mathbf{J}_s(p \approx \pm 1) \sim (1 \pm p)^{b-1} \qquad \rho_s(s \approx \pm 1) \sim (1 \pm s)^{b-1} \qquad b = \frac{\pi}{2\pi - \alpha} \tag{6.59}$$

where α is the angle between the two planes forming the wedge [40]. The same expressions are valid for a composite metallic and dielectric junction, except that the edge coefficient b is calculated in a more complicated way. For analysis based on the MoM solution of SIEs, it is more important to have a good approximation for currents than to have a good approximation for charges. The main part of the edge effects is taken into account if only the first condition in (6.59) is satisfied. This can be done if the node and patch basis functions given by (6.46) are modified as [41]

$$\mathbf{E}_i(p, s) = \frac{\mathbf{a}_s}{2|\mathbf{a}_p \times \mathbf{a}_s|} \frac{p^i N(s)}{(1+p)^{1-b_p^-}(1-p)^{1-b_p^+}} \tag{6.60}$$

$$\mathbf{P}_{ij}(p, s) = \frac{\mathbf{a}_s}{2|\mathbf{a}_p \times \mathbf{a}_s|} \frac{p^i S_j(s)}{(1+p)^{1-b_p^-}(1-p)^{1-b_p^+}}$$

where b_p^- and b_p^+ are edge coefficients at edges $p = -1$ and $p = 1$, respectively. The edge effect associated with charges can be completely taken into account if (6.60) is modified as

$$\mathbf{E}_i(p, s) = \frac{\mathbf{a}_s}{2|\mathbf{a}_p \times \mathbf{a}_s|} \frac{p^i}{(1+p)^{1-b_p^-}(1-p)^{1-b_p^+}} \frac{N(s)}{(1-s)^{1-b_s^+}} \tag{6.61}$$

$$\mathbf{P}_{ij}(p, s) = \frac{\mathbf{a}_s}{2|\mathbf{a}_p \times \mathbf{a}_s|} \frac{p^i}{(1+p)^{1-b_p^-}(1-p)^{1-b_p^+}} \frac{Sj(s)}{(1+s)^{1-b_s^-}(1-s)^{1-b_s^+}}$$

where b_s^- and b_s^+ are edge coefficients at edges $s = -1$ and $s = 1$, respectively. The basis functions (6.60) are more difficult to implement than the basis functions (6.46), and the basis functions (6.61) are more difficult than the basis functions (6.60). The impedance matrix elements due to the basis functions (6.46), (6.60), and (6.61) can be expressed in terms of $m = 1$, $m = 4$, and $m = 36$ classes of integrals, respectively. Hence, for the basis functions (6.61), the matrix fill time is about $m = 36$ times longer than for the basis functions (6.46).

For example, let us consider a corner scatterer modeled by $M = 2$ patches, respectively (Figure 6.21, inset). The corresponding edge coefficients are $b = 1/2$ (free edge) and $b = 2/3$ (right angle between plates). The incidence angle of a plane wave is $\phi = 30°$. Figure 6.21 shows the magnitude of the s-current component along the p-axis of the scatterer obtained without and with the edge effect included, and for two orders of the current approximation ($N = 4$ and $N = 6$ along each patch edge [41]). Including the edge effect

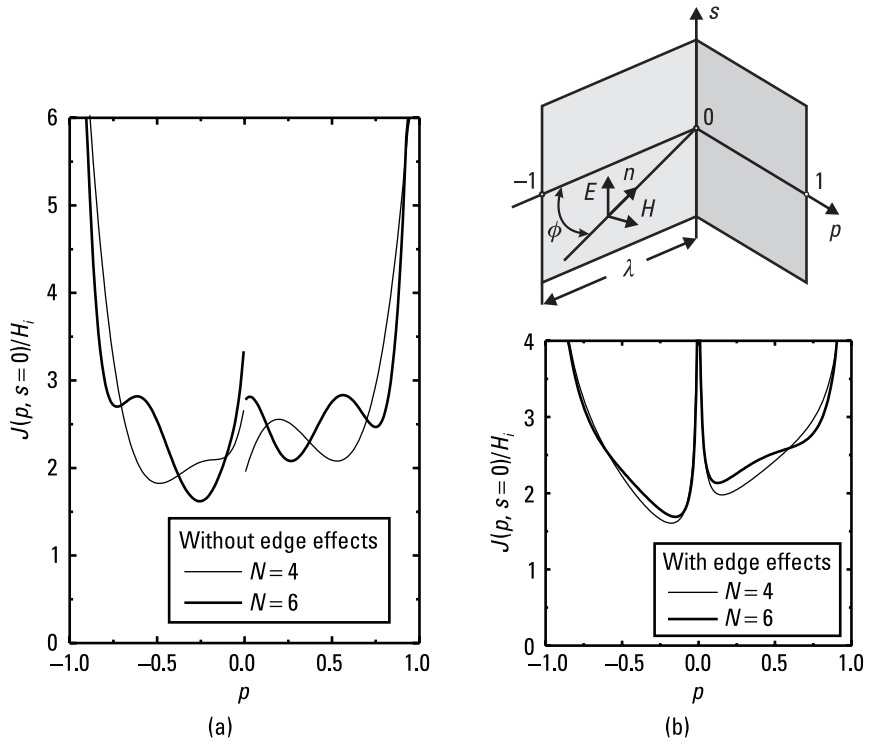

Figure 6.21 Magnitude of s-current component along p-coordinate axis of corner scatterer shown in the inset: (a) edge effect is not included, and (b) edge effect is included. (After: [41].)

substantially increases the accuracy and stability of the solution for the local variations of currents. On the other hand, inclusion of the edge effect practically does not affect the radiation pattern.

6.2.10 Square Scatterer Benchmark

Most often, the accuracy of approximate results obtained by the MoM solution of SIEs is subjectively discussed by comparing theoretical and experimental results. In order to treat the error quantitatively, we need a precise benchmark to which we may apply a metric, and we need the metric with which to measure the error.

The structure that is used as a benchmark should be a simple one that reflects most of the analysis problems. In what follows, let us concentrate on pure metallic structures. The analyzed structures usually contain edges. The current distribution exhibits quasisingular behavior in the vicinity of metallic edges, which is more pronounced if the edges are not connected. Hence, the structure used for comparison of analysis methods should contain free edges. The simplest structure is a square-plate scatterer. The scatterer should be sufficiently large in order that the accuracy of the analysis can be investigated in the large range of the number of unknowns. Hence, the benchmark structure is adopted in the form of an infinitesimally thin square scatterer of size $\lambda \times \lambda$ excited by a normally incident plane wave. The scatterer is situated in a local ps-coordinate system such that the plate edges coincide with the coordinate lines $p = \pm 1$ and $s = \pm 1$. The wave is polarized in the direction of the s-axis. The effective value of the wave electric field is $E = 1$ V/m [38].

The main goal of the electromagnetic modeling of plate structures based on the MoM is to determine the current distribution. Therefore, the error metric should be directly connected with the current distribution. The error metric cannot be based on the rms value of the current deviation. Namely, the surface current component parallel to the free plate edge tends to infinity as a reciprocal value of the square root of the distance from the edge, and the rms value of the deviation is infinitely large. Hence, it is convenient that the error metric is based on the mean absolute value of the current deviation. The relative error is evaluated as

$$E[\%] = 100 \frac{\int_S |\mathbf{J}_s - \mathbf{J}_{s0}| \, dS}{\int_S |\mathbf{J}_{s0}| \, dS} \tag{6.62}$$

where $\mathbf{J}_s$ is the approximate and $\mathbf{J}_{s0}$ is the accurate (reference) value of the surface current density vector over the scatterer. This expression can be applied either to the total current or the current components.

Use of the above error metric requires knowledge of the accurate solution. However, there is no exact solution for the current distribution over the square plate scatterer. Hence, only the approximate solution, which is estimated to be sufficiently close to the accurate solution, can be used as the reference. The approximate solution that converges quickly to the exact one is obtained by means of basis functions that include the quasistatic behavior. The p-current component is an odd function of the local p- and s-coordinates, equal to zero at edges $p = \pm 1$. The s-current component is an even function of the p- and s-coordinates equal to zero at the edges $s = \pm 1$. Hence, the expansions for the p- and s-current components are written in the form

$$J_{sp}(p, s) \cong \sum_{i=1(2)}^{n-1} \left\{ \sum_{j=3(2)}^{n} a_{pij}(p^j - p) \right\} \frac{s^i}{\sqrt{1 - s^2}} \tag{6.63a}$$

$$J_{ss}(p, s) \cong \sum_{i=0(2)}^{n-1} \left\{ \sum_{j=2(2)}^{n} a_{sij}(s^j - 1) \right\} \frac{p^i}{\sqrt{1 - p^2}} \tag{6.63b}$$

where n is the order of the approximation. The unknown coefficients in the above expansions are determined by the Galerkin method with special care devoted to the evaluation of potential, field, and impedance integrals.

Let us consider how the approximate solution changes with increases in the order of the current approximation n. Figure 6.22(a) shows the real and imaginary parts of the s-current component along the p-coordinate line [38]. The current is normalized with respect to the magnetic field of the incident plane wave. The results are given for $n = 2$, 3, 4, and 8. Figure 6.22(b) shows the same results, but along the s-coordinate line. The fast convergence of the results with increasing the order of approximation n is observed. The results obtained for $n = 4$ are very close to those obtained for $n = 8$. The results obtained for $n = 5$, 6, and 7 (not shown in the figure) are between the results for $n = 4$ and $n = 8$, while the results for $n = 9$ and 10 (not shown in the figure) practically coincide with those for $n = 8$. Hence, the order of the current approximation $n = 8$ is adopted for the benchmark solution. Table 6.3 gives the corresponding coefficients of the p-current component given by (6.63a). Table 6.4 gives the corresponding coefficients of the s-current component given by (6.63b).

Figure 6.23(a) shows the magnitude of the p-current component (normalized with respect to the magnetic field of the incident plane wave) over the scatterer [38]. Since this component is infinite along the edges $s = \pm 1$, it is

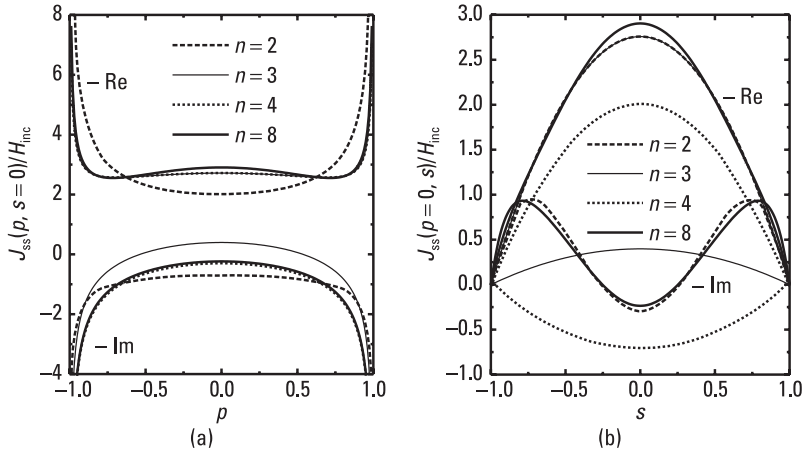

Figure 6.22 Real and imaginary parts of the *s*-current component (normalized with respect to the magnetic field of the incident plane wave) along (a) the *p*-coordinate line and (b) the *s*-coordinate line for various orders of approximation, *n*. (After: [38].)

Table 6.3
Coefficients of *p*-Current Component Given by (6.63a) for $n = 8$ (After: [38])

$a_{pij} \cdot 10^3$	$j = 3$	$j = 5$	$j = 7$
$i = 1$	$-0.412 - j2.259$	$0.365 + j2.666$	$-0.372 - j1.657$
$i = 3$	$-0.172 + j1.931$	$1.818 - j4.995$	$-0.911 + j3.624$
$i = 5$	$2.804 - j2.101$	$-6.718 + j6.730$	$3.823 - j4.860$
$i = 7$	$-2.217 + j0.859$	$5.085 - j3.261$	$-2.955 + j2.391$

Table 6.4
Coefficients of *s*-Current Component Given by (6.63b) for $n = 8$ (After: [38])

$a_{sij} \cdot 10^3$	$j = 2$	$j = 4$	$j = 6$	$j = 8$
$i = 0$	$9.622 - j10.649$	$-1.149 + j14.536$	$-6.534 - j15.652$	$5.763 + j11.140$
$i = 2$	$-7.492 + j0.105$	$-0.021 - j5.201$	$6.415 + j9.338$	$-5.686 - j7.146$
$i = 4$	$2.398 + j1.720$	$-1.509 + j0.547$	$0.545 - j2.875$	$0.502 + j1.075$
$i = 6$	$-0.265 - j1.055$	$1.003 - j1.728$	$-1.062 + j5.947$	$0.261 - j3.714$

shown only for $-0.95 \le s \le 0.95$. Figure 6.23(b) shows the normalized magnitude of the *s*-current component over the scatterer. This component is infinite along the edges $p = \pm 1$, so it is shown only for $-0.95 \le p \le 0.95$.

Finally, let us compare the efficiency of three approximations: rooftop, polynomial without taking into account the edge effect, and polynomial with

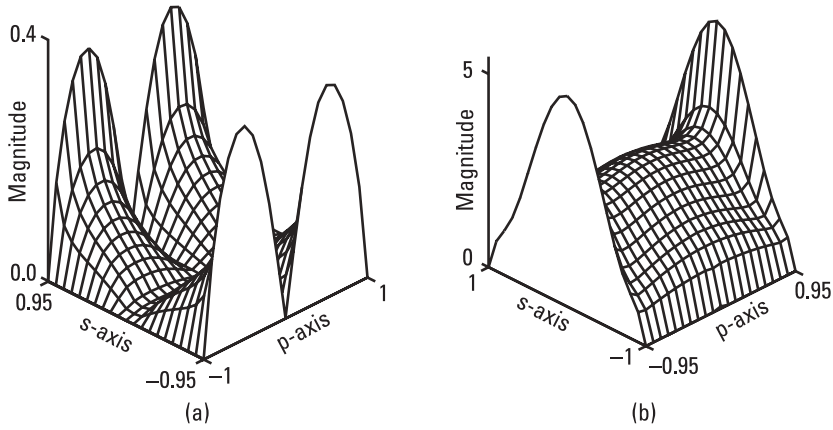

Figure 6.23 Magnitude of (a) the *p*-current component and (b) the *s*-current component over the scatterer. Both components are normalized with respect to the magnetic field of the incident plane wave. The results are shown for the benchmark solution of order $n = 8$. (After: [38].)

the edge effect (for $n < 7$). For the rooftop basis functions, the scatterer is uniformly divided into patches. The patches are squares of equal size. The number of patches along one scatterer side is n. The number of unknowns N corresponding to the order of the polynomial approximation n is equal to the number of the rooftop basis functions if the number of patches along one scatterer side is also n. The relation between N and n is given in Table 6.5.

Figure 6.24 shows the relative error of the approximations for the *s*-current components, respectively, versus the order of approximation, n [38]. Since the magnitude of the *s*-current component is much greater than the magnitude of the *p*-current component, the relative error of the approximation for the total current is almost the same as the relative error of the approximation for the *s*-current component. For the same accuracy, the rooftop solution needs a few times greater number of unknowns than the polynomial solution without the

Table 6.5
The Number of Unknowns, N, Corresponding to the Order of Polynomial Approximation, n*

n	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
N	4	12	24	40	60	84	112	144	180	220	264	312	364	420	480

* The same relation exists between the number of rooftop basis functions and the number of corresponding patches along the scatterer side.

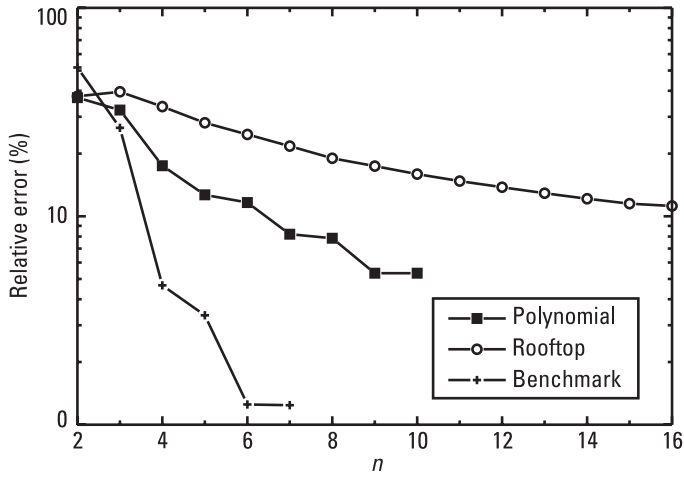

Figure 6.24 Relative error of approximations for the s-current component versus the order of approximation n. The relative error is evaluated for polynomial approximation with edge effect (for $n < 7$), polynomial approximation without edge effect, and rooftop approximation.

edge effect, which, in turn, requires a few times greater number of unknowns than the polynomial solution with the edge effect included.

6.3 Approximation of Currents over (Generalized) Triangles

6.3.1 Doublets and Rooftop Basis Functions

When two nodes of a generalized quadrilateral coincide, it reduces to a generalized triangle. However, the basis function valid for the quadrilateral reduces to a form that cannot be applied to the triangle. For example, let us consider a rooftop basis function (exact formulation) defined over a pair of flat quadrilaterals (Figure 6.11). If a flat quadrilateral degenerates into a flat triangle, the coefficients α and β in (6.37) reduce to $\alpha = -1$ and $\beta = 0$ (i.e. $\mathbf{r}_{ps} = -\mathbf{r}_p$) (Figure 6.25). The current distribution over the degenerate quadrilateral and the corresponding charge distribution, given by (6.39), are not defined for $s = 1$ (i.e., the direction of the current is not defined, and the charge density is infinite). Hence, such a basis function is inconvenient for the approximation of currents and charges over a triangle. This is confirmed by the shape-quality factor given by (6.40), which is $Q = 0$.

To define a proper doublet over a pair of generalized triangles, the node basis function in (6.35) should be squared, resulting in

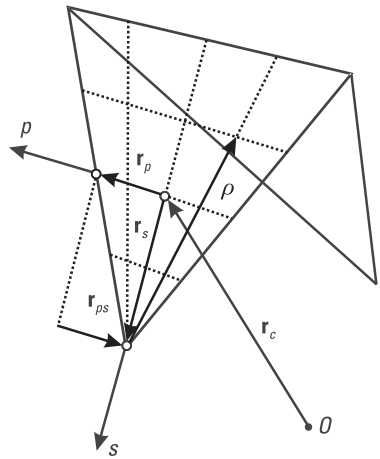

Figure 6.25 Rooftop basis function over a pair of triangles.

$$\mathbf{J}_s(p, s) = \frac{[N(s)]^2}{2|\mathbf{a}_p \times \mathbf{a}_s|}\mathbf{a}_s \qquad \rho_s(p, s) = \frac{jN(s)}{\omega|\mathbf{a}_p \times \mathbf{a}_s|}\frac{dN(s)}{ds} \tag{6.64}$$

These current and charge distributions are well defined for $s = 1$ (i.e., the current has zero value and the charge density is finite). The current distribution over a generalized triangle is represented by three such overlapped doublets.

In particular, let us consider a doublet defined over a pair of flat triangles and the node basis function defined by (6.13a). The position vector with respect to the free node is defined as $\boldsymbol{\rho} = -(1 - s)\mathbf{a}_s = 2N(s)\mathbf{a}_s$. After substituting $\alpha = -1$ and $\beta = 0$ into (6.38), the denominator of (6.64) becomes $2|\mathbf{a}_p \times \mathbf{a}_s| = N(s)S$. Finally, the current and charge distributions over one doublet arm are obtained in the form

$$\mathbf{J}_s(p, s) = \frac{1}{2S}\boldsymbol{\rho} \qquad \rho_s(p, s) = \frac{j}{\omega}\frac{1}{S} \tag{6.65}$$

This doublet represents a rooftop basis function for triangles [42]. Such doublets are also known as Rao-Wilton-Glisson (RWG) basis functions. Such doublets satisfy all five of the properties desirable for doublets listed in Table 6.2. According to [42], acceptable results are obtained with fewer than 100 unknowns per square wavelength. However, the following example demonstrates that triangular modeling requires about twice as many unknowns as quadrilateral modeling.

Consider a square plate scatterer of size $3\lambda \times 3\lambda$ with two triangular holes (Figure 6.26) [36]. The scatterer is situated in the xOy-plane with the y-axis along

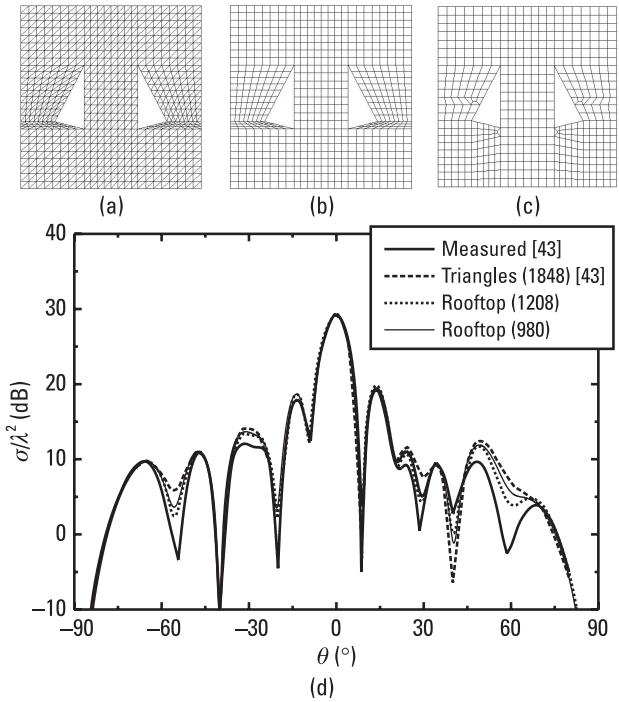

Figure 6.26 Square plate scatterer of size $3\lambda \times 3\lambda$ with two triangular holes: (a) to (c) geometrical models; (d) monostatic RCS in the plane $\phi = 90°$ versus angle θ. The scatterer is situated in xOy-plane with y-axis along the longest hole edge. Incident electric field has only θ-component. (After: [36].)

the longest hole edge. The incident electric field has only the θ-component. The scatterer is analyzed by using three geometrical models: a triangular model (N = 1,848) [43], a quadrilateral model obtained by merging pairs of triangles in the triangular model (N = 1,208 unknowns), and a quadrilateral model obtained by the double-node technique (the starting model is marked by bold lines, N = 980 unknowns). These three models are shown in Figure 6.26(a–c), respectively. Figure 6.26(d) shows the monostatic RCS in the plane $\phi = 90°$ versus the angle θ. Results obtained by using the three models are compared with measured results [43]. The results obtained by the quadrilateral modeling are closer to the experimental results than the results obtained by the triangular modeling.

6.3.2 Entire-Domain Approximation in Simplex Coordinates

Entire-domain approximations for currents over a generalized triangle can be easily developed in terms of simplex coordinates. We follow an approach similar to 2-D scalar and vector basis functions in the FEM literature (e.g., in [44]).

As the first step, let us develop an entire-domain expansion that enables the continuity of a scalar function over a triangle and along its edges. Such an expansion consists of triple power functions of simplex coordinates t_1, t_2, and t_3; that is, functions $t_1^p t_2^q t_3^r$, where p, q, and r are arbitrary integers. Since $t_1 + t_2 + t_3 = 1$, the simplex coordinates are mutually dependent. If t_3 is expressed in terms of t_1 and t_2, the function $t_1^p t_2^q t_3^r$ becomes a double polynomial in terms of independent coordinates t_1 and t_2. Let us consider a particular direction in the $t_1 t_2$ plane, such as $t_2 = at_1 + b$, where a and b are arbitrary constants. The function $t_1^p t_2^q t_3^r$ becomes a single polynomial in terms of t_1 only, whose order is $p + q + r = m$. We require that the expansion consist of independent functions $t_1^p t_2^q t_3^r$, such that $p + q + r = m \leq n$, where n is the order of approximation.

For the lowest-order approximation ($n = 0$), all coefficients p, q, and r must be zero, the function $t_1^p t_2^q t_3^r$ reduces to one, and the expansion consists of one term only. For the first order approximation ($n = 1$), only one of the coefficients p, q, and r can be different from zero and equal to one. The function $t_1^p t_2^q t_3^r$ reduces to t_1, t_2, or t_3. Since t_3 can be expressed in terms of t_1, and t_2, the set of independent functions is 1, t_1, and t_2. (An alternative set of independent functions is t_1, t_2, and t_3.) The expansion of order $n + 1$ can be obtained from the expansion of order n by adding $n + 2$ independent functions $t_1^p t_2^q t_3^r$ or their linear combinations, such that $p + q + r = n + 1$.

If two of the integers p, q, and r are equal to zero, the function $t_1^p t_2^q t_3^r$ is zero at one edge and two corresponding nodes. If one of these integers is equal to zero, the function is zero at two edges and all three nodes. Finally, if none of these integers is equal to zero, the function is zero at all three edges. Hence, only the functions t_1^p (t_2^q, t_3^r) are not equal to zero at the node $t_1 = 1$ ($t_2 = 1$, $t_3 = 1$). However, we would like to have only functions t_1, t_2, and t_3 equal to one at these nodes. Hence, only the expansion consisting of functions $t_1^p t_2^q$ ($t_2^q t_3^r$, $t_3^r t_1^p$) is not equal to zero along the edge $t_3 = 0$ ($t_1 = 0$, $t_2 = 0$). To easily match these expansions at the junction of two triangles, it is convenient that they consist of alternatively odd and even basis functions along the edge with respect to the center of the edge. The simplest form of such a set of independent functions is given by (6.27), (6.28a), and (6.29), as explained in Section 6.1.6.

All functions $t_1^p t_2^q t_3^r$ for integers p, q, and r different from zero can be written as $(t_1 t_2 t_3)^s t_1^{p-s} t_2^{q-s} t_3^{r-s}$, where $s = \min(p, q, r)$. In $t_1^{p-s} t_2^{q-s} t_3^{r-s}$ at least one of the integers $p - s$, $q - s$, and $r - s$ is equal to zero. In order that functions $(t_1 t_2 t_3)^s t_1^{p-s} t_2^{q-s} t_3^{r-s}$ be independent for a fixed s, the functions $t_1^{p-s} t_2^{q-s} t_3^{r-s}$ must be adopted in the same way as the set of independent

functions $t_1^p t_2^q$, $t_2^q t_3^r$, and $t_3^r t_1^p$, which enables the continuity of the scalar function along the edges.

Finally, basis functions making a set that enables the continuity of a scalar function over a triangle and along its edges can be written in the form

$$N(t_1) = t_1 \qquad E_i(t_1, t_2) = \begin{cases} (t_1 t_2)^k & i = 2k \\ (t_1 t_2)^k (t_2 - t_1) & i = 2k + 1 \end{cases} \tag{6.66a,b}$$

$$P_{ij}(t_1, t_2, t_3) = (t_1 t_2 t_3)^l \begin{cases} N(t_1) & i = 1 \\ E_i(t_1, t_2) & i \geq 2 \end{cases} \qquad j = 3l \geq 3 \qquad i + j \leq n \tag{6.66c}$$

where $N(t_1)$ is a node basis function, $E_i(t_1, t_2)$ are edge basis functions, and $P_{ij}(t_1, t_2, t_3)$ are patch basis functions. The remaining basis functions are obtained by a cyclical permutation of indices 1, 2, and 3. Grouping node basis functions into multiplets and edge basis functions into doublets, the continuity of the scalar function is automatically satisfied.

Generally, the surface current can be decomposed into two components on the surface. Each component can be approximated by a scalar expansion (in terms of known basis functions and unknown coefficients) multiplied by a known vector function, which defines the reference direction of the component at each point of the surface. These known vector functions for two current components must not be collinear at any point of the surface. Hence, the approximation of currents in terms of simplex coordinates can be obtained if the expansion, which consists of basis functions (6.66a-c) and their cyclical permutations, is multiplied by two properly adopted noncolinear vector functions. Moreover, various basis functions can be multiplied by different pairs of known vector functions. The pair of known vector functions should be adopted in such a manner that the continuity of currents flowing across the junction of two triangles can be easily satisfied.

First, let us consider the basis function $N(t_1)$. It is convenient to express t_1 in terms of t_2 and t_3, so that the parametric coordinate lines t_2 and t_3 coincide with two edges of the triangle [Figure 6.27(a)]. This triangle represents a half of the quadrilateral (obtained by splitting the quadrilateral along one diagonal) whose edges are placed along the coordinate lines $t_2 = 0$, $t_2 = 1$, $t_3 = 0$, and $t_3 = 1$. Similarly to Sections 6.2.2 and 6.2.4, the continuity of currents flowing across the coordinate lines $t_2 = 0$ and $t_3 = 0$ can be provided if $N(t_1)$ is multiplied by $\mathbf{a}_{t_2}^{(1)} \Big/ \left| \mathbf{a}_{t_2}^{(1)} \times \mathbf{a}_{t_3}^{(1)} \right|$ and $\mathbf{a}_{t_3}^{(1)} \Big/ \left| \mathbf{a}_{t_2}^{(1)} \times \mathbf{a}_{t_3}^{(1)} \right|$, where $\mathbf{a}_{t_2}^{(1)}$

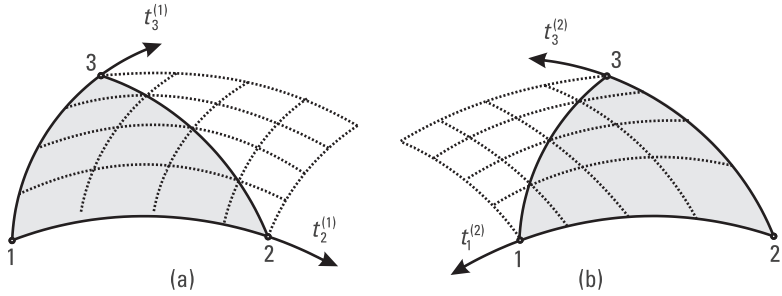

Figure 6.27 Equivalence of simplex coordinates of a triangle and parametric coordinates of a quadrilateral when (a) t_1 and (b) t_2 are expressed in terms of t_1 and t_3.

and $\mathbf{a}_{t_3}^{(1)}$ are unitary vectors along the t_2- and t_3-coordinate lines. (The superscript (1) means that the coordinate lines are obtained starting from the simplex coordinates when t_1 is expressed in terms of t_2 and t_3.) On the other hand, according to the formula for the gradient of a scalar function expressed in terms of parametric coordinates [45], the gradients of parametric coordinates t_2 and t_3 are obtained as

$$\nabla t_2 = \frac{\mathbf{a}_{t_3}^{(1)} \times \mathbf{n}}{\left|\mathbf{a}_{t_2}^{(1)} \times \mathbf{a}_{t_3}^{(1)}\right|} \qquad \nabla t_3 = \frac{\mathbf{n} \times \mathbf{a}_{t_2}^{(1)}}{\left|\mathbf{a}_{t_2}^{(1)} \times \mathbf{a}_{t_3}^{(1)}\right|} \tag{6.67}$$

In that case, $\mathbf{a}_{t_3}^{(1)} \Big/ \left|\mathbf{a}_{t_2}^{(1)} \times \mathbf{a}_{t_3}^{(1)}\right|$ and $\mathbf{a}_{t_2}^{(1)} \Big/ \left|\mathbf{a}_{t_2}^{(1)} \times \mathbf{a}_{t_3}^{(1)}\right|$ can be expressed as $\mathbf{n} \times \nabla t_2$ and $-\mathbf{n} \times \nabla t_3$, respectively. Thus, two vector functions are obtained in the form $\mathbf{n} \times (t_1 \nabla t_2)$ and $-\mathbf{n} \times (t_1 \nabla t_3)$. In a similar way, starting from the scalar node basis functions $N(t_2)$ and $N(t_3)$, we obtain vector basis functions in the form $\mathbf{n} \times (t_2 \nabla t_3)$, $-\mathbf{n} \times (t_2 \nabla t_1)$, $\mathbf{n} \times (t_3 \nabla t_1)$, and $-\mathbf{n} \times (t_3 \nabla t_2)$. At each edge only two basis functions have the component normal to the edge. For example, at the edge 1-2 these components are $\mathbf{n} \times (t_1 \nabla t_2)$ and $-\mathbf{n} \times (t_2 \nabla t_1)$. To provide symmetry with respect to the central point of the edge, these basis functions are combined into $\mathbf{n} \times (t_1 \nabla t_2 - t_2 \nabla t_1)$ and $\mathbf{n} \times (t_1 \nabla t_2 + t_2 \nabla t_1)$.

Further, let us consider the scalar edge basis functions given by (6.66b). If t_1 is expressed in terms of t_2 and t_3, these basis functions should be multiplied by $\mathbf{n} \times \nabla t_2$ and $-\mathbf{n} \times \nabla t_3$, resulting in one set of vector edge basis functions. Alternatively, if t_2 is expressed in terms of t_1 and t_3 [Figure 6.27(b)], these basis functions should be multiplied by $-\mathbf{n} \times \nabla t_1$ and $\mathbf{n} \times \nabla t_3$, resulting in another set of vector edge basis functions. To preserve the symmetry in terms of the simplex coordinates, these two sets can be combined [i.e., the edge basis

functions should be multiplied by $\mathbf{n} \times (\nabla t_1 - \nabla t_2)$ and $\mathbf{n} \times \nabla t_3$]. All vector basis functions related to edge 1-2 can be written in the form

$$\mathbf{E}_i^{\mathrm{N}}(t_1, t_2) = \mathbf{n} \times \begin{cases} t_1 \nabla t_2 - t_2 \nabla t_1 & i = 0 \\ t_1 \nabla t_2 + t_2 \nabla t_1 & i = 1 \\ (t_1 t_2)^k (\nabla t_2 - \nabla t_1) & i = 2k \\ (t_1 t_2)^k (t_2 - t_1)(\nabla t_2 - \nabla t_1) & i = 2k + 1 \end{cases} \tag{6.68}$$

$$\mathbf{E}_i^{\mathrm{T}}(t_1, t_2) = \mathbf{n} \times \begin{cases} (t_1 t_2)^k \nabla t_3 & i = 2k \\ (t_1 t_2)^k (t_2 - t_1) \nabla t_3 & i = 2k + 1 \end{cases}$$

The functions $\mathbf{E}_i^{\mathrm{N}}(t_1, t_2)$ have the normal component only to the edge 1-2. The functions $\mathbf{E}_i^{\mathrm{T}}(t_1, t_2)$ have no normal component to any edge and have the tangential component only to the edge 1-2.

Finally, let us consider the scalar patch basis functions given by (6.66c). If we multiply these functions by $\mathbf{n} \times \nabla t_1$, $\mathbf{n} \times \nabla t_2$, and $\mathbf{n} \times \nabla t_3$, and then combine them in the same way as the scalar node and edge basis functions from which they are composed, we obtain vector basis functions in the form

$$\mathbf{P}_{ij}(t_1, t_2, t_3) = (t_1 t_2 t_3)^l \begin{cases} \mathbf{E}_i^{\mathrm{T}}(t_1, t_2) & i \geq 0 \\ \mathbf{E}_i^{\mathrm{N}}(t_1, t_2) & i > 1 \end{cases} \qquad j = 3l \geq 3 \tag{6.69}$$

The basis functions $\mathbf{E}_i^{\mathrm{N}}$ and $\mathbf{E}_i^{\mathrm{T}}$, which are related to the edges 2-3 and 3-1, and the corresponding basis functions $\mathbf{P}_{ij}$ are obtained by a cyclical permutation of indices 1, 2, and 3 in (6.68) and (6.69). Only the basis functions $\mathbf{E}_i^{\mathrm{N}}$ take part in the satisfaction of the continuity of currents flowing across the junction of two triangles.

In the FEM literature, basis functions that maintain the continuity between cells are known as conforming functions. The vector basis functions that only impose the continuity of the normal component (briefly referred to as the normal continuity) between cells, like those given by (6.68) and (6.69), are designated div-conforming. Similarly, vector basis functions that only impose the continuity of the tangential component (the tangential continuity) are referred to as curl-conforming [46]. The curl-conforming and div-conforming basis functions are related as

$$\mathbf{J}_{\text{curl-conforming}} = \mathbf{J}_{\text{div-conforming}} \times \mathbf{n} \tag{6.70}$$

Hence, by replacing (6.68) and (6.69) into (6.70), the curl-conforming basis functions are obtained. Such basis functions have application in FEM solutions of 2-D problems.

6.4 Generalized Hexahedrons

As elaborated in Chapter 5, generalized hexahedrons are used for modeling dielectric domains with associated equivalent volume electric (magnetic) currents or electric (magnetic) fields. The equivalent volume currents in dielectric domains are determined by applying the MoM to the VIEs. It is convenient to express the electric (magnetic) currents in terms of the electric (magnetic) flux density, which has a continuous normal component at media interfaces [47]. Such div-conforming basis functions are briefly described in Section 6.4.1. Alternatively, the electric (magnetic) field in dielectric domains is used as the unknown quantity when differential equations are solved. If these equations are solved by the FEM, it is required that basis functions satisfy the tangential continuity of the field [48]. Such curl-conforming basis functions are briefly described in Section 6.4.2.

6.4.1 Basis Functions That Maintain Normal Continuity (VIE)

Since equivalent volume currents, electric and magnetic, are approximated in the same way, this kind of approximation will be illustrated with the example of electric currents. If a region filled by a dielectric of complex permittivity ϵ is homogenized with respect to a vacuum, the electric currents and the corresponding charges are expressed in terms of electric flux density as

$$\mathbf{J} = j\omega\kappa\mathbf{D} \qquad \kappa = \frac{\epsilon - \epsilon_0}{\epsilon} \qquad \rho = -\text{div}(\kappa\mathbf{D}) \tag{6.71}$$

where ω is the angular frequency and κ is the so-called contrast ratio.

Since a generalized hexahedron is defined in a local pqs-coordinate system, it is natural to expand the flux density in the same coordinate system as

$$\mathbf{D} = D_p\mathbf{i}_p + D_q\mathbf{i}_q + D_s\mathbf{i}_s \tag{6.72}$$

where $\mathbf{i}_p$, $\mathbf{i}_q$, and $\mathbf{i}_s$ are the unit vectors along p-, q-, and s-coordinate lines (see Figure 5.15). However, the p- and q-components can be treated as the s-components defined within the same generalized hexahedron whose parametric coordinates are cyclically interchanged. Hence, in what follows, **D** will

designate only the s-component of the flux density vector. The corresponding volume charge density is determined in terms of the parametric coordinates as

$$\rho(p, q, s) = \frac{-1}{(\mathbf{a}_p \times \mathbf{a}_q) \cdot \mathbf{a}_s} \frac{\partial}{\partial s} \left\{ \kappa D(p, q, s) \frac{(\mathbf{a}_p \times \mathbf{a}_q) \cdot \mathbf{a}_s}{|\mathbf{a}_s|} \right\} \quad (6.73)$$

where $\mathbf{a}_p$, $\mathbf{a}_q$, and $\mathbf{a}_s$ are unitary vectors of the generalized hexahedron (see Section 5.3).

In order that the normal continuity of the flux density at the interface between two hexahedrons can be easily satisfied, the initial entire-domain approximation for the s-component of this vector is adopted as [49]

$$\mathbf{D}(p, q, s) = \sum_{i=0}^{n_p} \sum_{j=0}^{n_q} \sum_{k=0}^{n_s} a_{ijk} \mathbf{D}_{ijk}(p, q, s) \quad (6.74)$$

$$\mathbf{D}_{ijk}(p, q, s) = \frac{f_i(p) g_j(q) h_k(s)}{4(\mathbf{a}_p \times \mathbf{a}_q) \cdot \mathbf{a}_s} \mathbf{a}_s$$

where n_p, n_q, and n_s are orders of approximation along the p-, q-, and s-coordinates; a_{ijk} are unknown coefficients to be determined; $\mathbf{D}_{ijk}(p, q, s)$ are known vector basis functions; $f_i(p)$, $g_j(q)$, and $h_k(s)$ are arbitrary known functions; and $\mathbf{a}_p$, $\mathbf{a}_q$, and $\mathbf{a}_s$ are unitary vectors. To match the normal components in two hexahedrons at the common interface, the initial expansion is rearranged in a similar way as for surface currents in Section 6.2.5. In particular, if the basis functions $f_i(p)$, $g_j(q)$, and $h_k(s)$ are expressed in the form of simple power functions p^i, q^j, and s^k, the rearranged expansion is obtained in the form

$$\mathbf{D}(p, q, s) = \sum_{i=0}^{n_p} \sum_{j=0}^{n_q} \left\{ c_{ij1} \mathbf{F}_{ij}(p, q, s) + c_{ij2} \mathbf{F}_{ij}(p, q, -s) + \sum_{k=2}^{n_s} a_{ijk} \mathbf{B}_{ijk}(p, q, s) \right\} \quad (6.75)$$

$$\mathbf{F}_{ij}(p, q, s) = \frac{\mathbf{a}_s}{4(\mathbf{a}_p \times \mathbf{a}_q) \cdot \mathbf{a}_s} p^i q^j N(s) \quad (6.76a)$$

$$\mathbf{B}_{ijk}(p, q, s) = \frac{\mathbf{a}_s}{4(\mathbf{a}_p \times \mathbf{a}_q) \cdot \mathbf{a}_s} p^i q^j S_k(s) \quad (6.76b)$$

where $N(s)$ and $S_k(s)$ are node and segment functions defined in the thin-wire analysis by (6.13).

Since only the basis functions $\mathbf{F}_{ij}(p, q, \pm s)$ are not equal to zero at faces $s = \pm 1$, they will be referred to as face basis functions. The normal-continuity at the interface between two hexahedrons is automatically satisfied if the face basis functions of the same indices i and j are grouped into doublets. All these doublets are grouped into the face expansion. Since the basis functions $\mathbf{B}_{ijk}(p, q, s)$ approximate the flux density inside a hexahedron brick, we call them brick basis functions. All brick basis functions of one hexahedron are grouped into brick expansions. For complex structures, the flux density vector is approximated by brick and face expansions. Since the brick and face basis functions maintain the normal continuity between adjacent hexahedron bricks, they belong to the class of div-conforming basis functions.

For the lowest order of expansion (6.75) ($n_p = n_q = 0$, $n_s = 1$), the brick basis functions drop out, and the face expansion reduces to a doublet of the lowest order. According to (6.76a) the flux density within one doublet arm and corresponding charges are given by

$$\mathbf{D}(p, q, s) = \frac{N(s)}{4(\mathbf{a}_p \times \mathbf{a}_q) \cdot \mathbf{a}_s} \mathbf{a}_s \qquad \rho(p, q, s) = \frac{-1}{4(\mathbf{a}_p \times \mathbf{a}_q) \cdot \mathbf{a}_s} \frac{\partial [\kappa N(s)]}{\partial s} \tag{6.77}$$

Considering that a face is most often shared by two bricks, the minimal number of unknowns per brick is $N_{\text{brick}} = 3$. This number is doubled if the face basis functions are not combined into doublets.

When a generalized hexahedron reduces to a trilinear volume, the position vector with respect to the free s-face can be defined as $\boldsymbol{\rho} = -(1 - s)\mathbf{a}_s$ (Figure 6.28). If the contrast ratio does not depend on the spatial coordinates, the flux density within one doublet arm and the corresponding charges are written in the form

$$\mathbf{D}(p, q, s) = \frac{1}{8(\mathbf{a}_p \times \mathbf{a}_q) \cdot \mathbf{a}_s} \boldsymbol{\rho} \qquad \rho(p, q, s) = \frac{\kappa}{8(\mathbf{a}_p \times \mathbf{a}_q) \cdot \mathbf{a}_s} \tag{6.78a,b}$$

This doublet represents a volume rooftop basis function (exact formulation).

We can define the first five desired properties of volume doublets in a similar way as for the surface doublets. The volume rooftop basis functions satisfy properties P1 (global continuity), P2 (local continuity), and P3 (simple integrals). However, the P4 property (constant field distribution) can be satisfied only for some degenerate hexahedrons (regular and oblique parallelepipeds and

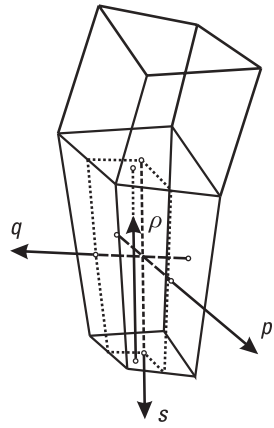

Figure 6.28 Volume rooftop basis function.

prisms). In that case, $\mathbf{D}(p, q, s) = \boldsymbol{\rho}/V$ and $\rho(p, q, s) = \kappa/V$, where V is the volume of the doublet arm.

The fifth desired property is somewhat different than that for the surface doublets. Namely, in a homogenous dielectric there are no volume charges (i.e., it is desirable that basis functions approximating the electric flux density not be accompanied by volume charges). Hence, the P5 property can be succinctly called "zero charge distribution." Obviously, a single volume rooftop basis function does not satisfy the P5 property. However, a linear combination of six rooftop basis functions can easily satisfy this property.

For example, let us consider a homogeneous lossless dielectric sphere of relative permittivity $\epsilon_r = 36$ situated in an incident plane electromagnetic wave of the electric field intensity $\mathbf{E}_i = \exp(j\beta_0 z)\mathbf{i}_x$ V/m (Figure 6.29). Let the sphere radius a be such that $\beta_0 a = 0.408$. The sphere is modeled by $M = 54$ patches, and the subdomain (rooftop-S) approximation is applied ($N = 6$ unknowns). Figure 6.29 shows the magnitude of the x-electric-field component along the z-axis. These results are compared with results obtained by solving the VIE combined with hexahedral modeling ($N = 344$) [49], with results obtained by solving the VIE combined with tetrahedral modeling ($N = 304$) [47], and with the analytical solution in the form of Mie's series [47]. In all numerical solutions, two symmetry planes are taken into account to reduce the number of unknowns. The results obtained by the surface doublets are closer to the analytical solution than the results obtained by the volume doublets, requiring at the same time a significantly smaller number of unknowns.

6.4.2 Basis Functions That Maintain Tangential Continuity (FEM)

When differential equations in terms of the electric or magnetic field are solved by the FEM, it is required that basis functions satisfy the tangential continuity

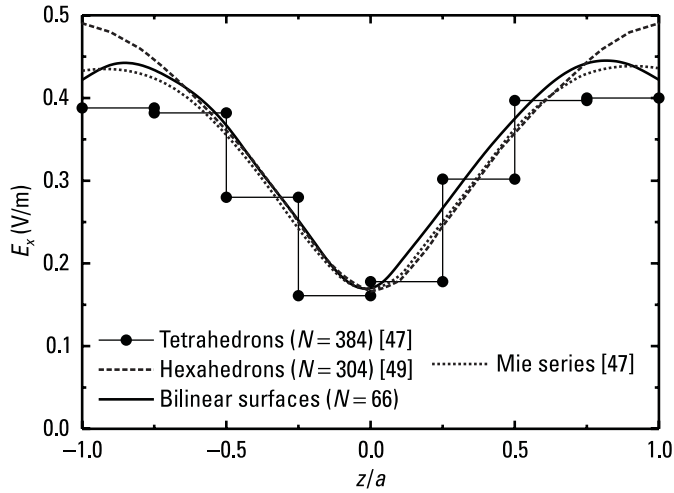

Figure 6.29 Magnitude of the x-electric-field component along the z-axis of a homogeneous lossless dielectric sphere of radius a and relative permittivity $\epsilon_r = 36$ ($\beta_0 a = 0.408$) excited by plane wave $\mathbf{E}_i = \exp(j\beta_0 z)\mathbf{i}_x$ V/m. Results are obtained using different surface and volume doublets. (After: [37].)

of the field [48]. Since both the electric and magnetic fields are approximated in the same way, this type of approximation will be illustrated with an example of the electric field.

Since a generalized hexahedron is defined in a local pqs-coordinate system, it is natural to expand the electric field vector in the same coordinate system. However, to obtain curl-conforming basis functions, the decomposition is not done in terms of the unit vector $\mathbf{i}_p$, $\mathbf{i}_q$, and $\mathbf{i}_s$, or the corresponding unitary vectors $\mathbf{a}_p$, $\mathbf{a}_q$, and $\mathbf{a}_s$, but in terms of the inverse unitary vectors $\mathbf{a}^p$, $\mathbf{a}^q$, and $\mathbf{a}^s$, defined as

$$\mathbf{a}^p = \frac{\mathbf{a}_q \times \mathbf{a}_s}{(\mathbf{a}_q \times \mathbf{a}_s) \cdot \mathbf{a}_p} \qquad \mathbf{a}^q = \frac{\mathbf{a}_s \times \mathbf{a}_p}{(\mathbf{a}_s \times \mathbf{a}_p) \cdot \mathbf{a}_q} \qquad \mathbf{a}^s = \frac{\mathbf{a}_p \times \mathbf{a}_q}{(\mathbf{a}_p \times \mathbf{a}_q) \cdot \mathbf{a}_s} \tag{6.79}$$

[The name *inverse unitary vector* is due to the fact that the product of a unitary vector and its inverse vector is equal to one (e.g., $\mathbf{a}_p \cdot \mathbf{a}^p = 1$). The product of the unitary vector and other inverse vectors is equal to zero (e.g., $\mathbf{a}_p \cdot \mathbf{a}^q = 0$).] The decomposition of a vector along inverse unitary vectors is also referred to as decomposition into covariant components [45, 50]. Similarly, the decomposition of a vector along unitary vectors is referred to as decomposition into contravariant components.

The $\mathbf{a}^p$- and $\mathbf{a}^q$-components can be treated as the $\mathbf{a}^s$-components defined within the same generalized hexahedron, but with the parametric coordinates cyclically interchanged. Hence, in what follows, $\mathbf{E}$ will designate only the $\mathbf{a}^s$-component of the electric field. In order that the tangential continuity of the electric field at the interface between two hexahedrons can be satisfied, the initial entire-domain approximation for the $\mathbf{a}^s$-component of this vector is adopted similarly to [50, 51] as

$$\mathbf{E}(p, q, s) = \sum_{i=0}^{n_p} \sum_{j=0}^{n_q} \sum_{k=0}^{n_s} a_{ijk} \mathbf{E}_{ijk}(p, q, s) \tag{6.80a}$$

$$\mathbf{E}_{ijk}(p, q, s) = f_i(p) g_j(q) h_k(s) \mathbf{a}^s \tag{6.80b}$$

where n_p, n_q, and n_s are the orders of approximation along the p-, q-, and s-coordinates, a_{ijk} are unknown coefficients to be determined, $\mathbf{E}_{ijk}(p, q, s)$ are known vector basis functions, and $f_i(p)$, $g_j(q)$, and $h_k(s)$ are arbitrary known functions. Since the scalar product of (6.80b) by $\mathbf{a}^p$ or $\mathbf{a}^q$ is zero, the expansion $\mathbf{E}(p, q, s)$ has the tangential components only at faces $p = \pm 1$ and $q = \pm 1$. In order that the tangential components in two hexahedrons can be matched at the common interface, the initial expansion is rearranged in a similar way as for the surface currents in Section 6.2.5. In particular, if the basis functions $f_i(p)$, $g_j(q)$, and $h_k(s)$ are expressed in the form of simple power functions p^i, q^j, and s^k, the rearranged expansion is obtained in the form

$$\begin{aligned} \mathbf{E}(p, q, s) = \sum_{k=0}^{n_s} \Bigg\{ & \sum_{j=1}^{2} \sum_{i=1}^{2} d_{ijk} \mathbf{W}_k[(-1)^i p, (-1)^j q, s] \\ & + \sum_{j=2}^{n_q} \sum_{i=1}^{2} c_{ijk} \mathbf{F}_{jk}[(-1)^i p, q, s] \\ & + \sum_{j=1}^{2} \sum_{i=2}^{n_p} b_{ijk} \mathbf{F}_{ik}[(-1)^j q, p, s] + \sum_{j=2}^{n_q} \sum_{i=2}^{n_p} a_{ijk} \mathbf{B}_{ijk}(p, q, s) \Bigg\} \end{aligned} \tag{6.81}$$

$$\begin{aligned} \mathbf{W}_k(p, q, s) &= \mathbf{a}^s N(p) N(q) s^k \\ \mathbf{F}_{jk}(p, q, s) &= \mathbf{a}^s N(p) S_j(q) s^k \\ \mathbf{B}_{ijk}(p, q, s) &= \mathbf{a}^s S_i(p) S_j(q) s^k \end{aligned} \tag{6.82}$$

where $N(s)$ and $S(s)$ are the node and segment functions defined in the thin-wire analysis by (6.13).

Only the tangential components of the basis functions $\mathbf{W}_k(\pm p, \pm q, s)$ are not equal to zero along edges $(p, q) = \pm 1, \pm 1$. In addition, the tangential components of these basis functions are not equal to zero only along the wedges corresponding to these edges. Hence, we call them wedge basis functions. After combining the wedge basis functions of the same order k into multiplets, the tangential continuity along the edges is automatically satisfied. The tangential components of the basis functions $\mathbf{F}_{jk}(\pm p, q, s)$ and $\mathbf{F}_{ik}(\pm q, p, s)$ are not equal to zero only at faces $p = \pm 1$ and $q = \pm 1$. Hence, we call them face basis functions. The face basis functions grouped in doublets, together with the wedge basis functions grouped in multiplets, automatically satisfy the tangential continuity over faces. Since the basis functions $\mathbf{B}_{ijk}(p, q, s)$ approximate the electric field inside a hexahedron brick, we call them brick basis functions. For complex structures, the electric field vector is approximated by brick, face, and wedge expansions. Since the brick, face, and wedge basis functions maintain the tangential continuity between the adjacent hexahedron bricks, they belong to the class of curl-conforming basis functions. For the lowest-order expansion $(n_p = 1, n_q = 1, n_s = 0)$, the brick and face basis functions drop out, and the wedge expansion reduces to a multiplet of the lowest order. Considering that a wedge is most often shared by four bricks, the minimal number of unknowns per brick is $N_{\text{brick}} = 3$.

6.5 Generalized Tetrahedrons

When a generalized hexahedron degenerates into a generalized tetrahedron, the basis function valid for the hexahedron reduces to a form that cannot be applied to the tetrahedron. For example, let us consider a volume rooftop basis function (exact formulation) defined over a pair of trilinear volumes, shown in Figure 6.28. Let the trilinear volume degenerate into a regular tetrahedron such that the face $s = 1$ degenerates into a vertex and the face $q = 1$ degenerates into an edge (Figure 6.30). The direction of the electric flux density given by (6.78a) is not defined at the vertex, and the corresponding charge density is infinite.

To define a proper doublet basis function over the pair of generalized tetrahedrons, the flux density over one arm and corresponding charges given by (6.77) should be modified as

$$\mathbf{D}(p, q, s) = \frac{N(q)\,[N(s)]^3}{4(\mathbf{a}_p \times \mathbf{a}_q) \cdot \mathbf{a}_s}\,\mathbf{a}_s \qquad \rho(p, q, s) = \frac{-N(q)}{4(\mathbf{a}_p \times \mathbf{a}_q) \cdot \mathbf{a}_s}\,\frac{\partial[\kappa N(s)^3]}{\partial s} \tag{6.83}$$

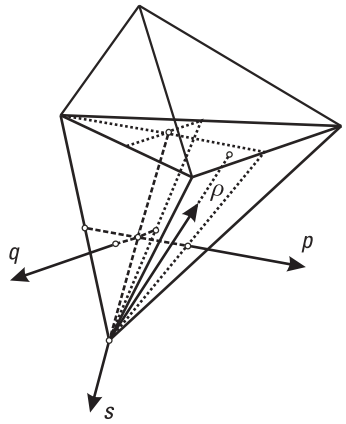

Figure 6.30 Volume rooftop basis functions for a pair of tetrahedrons.

In particular, let us consider the node basis function defined by (6.13a). For an ordinary tetrahedron (with flat faces), the position vector with respect to the free node is defined as $\boldsymbol{\rho} = -(1 - s)\mathbf{a}_s = 2N(s)\mathbf{a}_s$. The denominator in (6.83) becomes $8(\mathbf{a}_p \times \mathbf{a}_q) \cdot \mathbf{a}_s = 3N(q)[N(s)]^2 V$, where V is the volume of the tetrahedron. Finally, if the contrast ratio does not depend on spatial coordinates, the flux density in one arm of the doublet and the corresponding charges are obtained as

$$\mathbf{D}(p, q, s) = \frac{1}{3V}\boldsymbol{\rho} \qquad \rho(p, q, s) = \frac{\kappa}{V} \tag{6.84}$$

The resulting doublet represents a volume rooftop basis function (exact formulation) for ordinary tetrahedrons [47].

An entire-domain approximation for currents and fields in the generalized tetrahedron, which satisfies either the normal or the tangential continuity, can be developed in terms of simplex coordinates, following an approach similar to that presented in Section 6.3.2 or in [52].

6.6 Approximation of Currents and Fields Across Junctions of Incompatible Building Elements

Consider a junction of a small and a large quadrilateral [Figure 6.31(a)]. The order of approximation of currents in one direction along a quadrilateral is directly proportional to the maximal electrical length in this direction. Hence, the number of edge basis functions is equal on the opposite sides of the

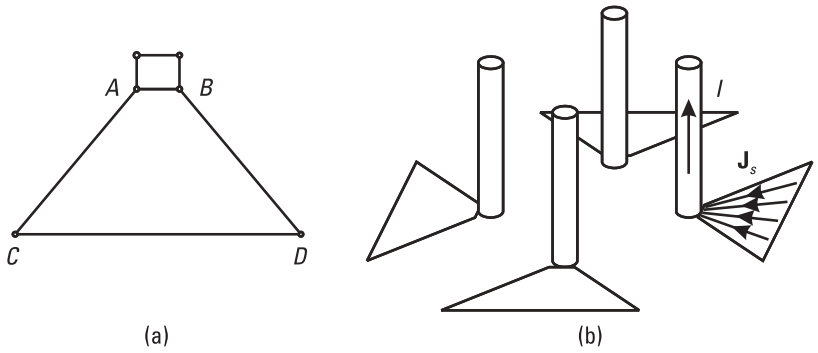

Figure 6.31 Approximations of currents and fields across junctions of incompatible building elements: (a) junction of small and large quadrilaterals, and (b) wire-to-plate junction.

quadrilateral, even if these sides are of very different lengths. As a result, the numbers of edge basis functions along the junction can differ for the interconnected quadrilaterals. To satisfy the normal continuity of currents along the junction, the edge basis functions in the large quadrilateral that have no counterpart in the small quadrilateral should be omitted. For example, let the electrical lengths of sides AB and CD be 0.15λ and 1λ, respectively. The order of approximation in the direction of the interconnecting edge is 1 for the small quadrilateral and 4 for the large quadrilateral. There is only one edge basis function (of the zeroth order) along the junction in the small quadrilateral, while there are four edge basis functions (of orders zero to three) in the large quadrilateral. Obviously, only the edge basis functions of the zeroth order can be grouped into a doublet of the zeroth order. Since the edge basis functions of orders one to three in the large quadrilateral have no the counterpart in the small quadrilateral, they should be omitted. In a similar way, we handle the junction of a small and a large triangle, the interface between a small and a large hexahedron, and the interface of a small and a large tetrahedron.

Consider next a wire-to-plate junction. Basis functions that take part in the satisfaction of the KCL for the junction domain are node basis functions associated with the wire ends and edge basis functions associated with the short plate edges. The edge basis functions of higher orders represent an unnecessary degree of freedom along short edges. Hence, these basis functions are omitted from the plate expansions. The KCL is automatically satisfied if the node and edge basis functions are grouped into doublets, along which the continuity is satisfied in a global way (i.e., the total current flowing from one doublet arm is equal to the total current flowing into another doublet arm). For example, consider the junction of a wire and a central part of the plate, which is represented

by a junction of a wire and four plates [Figure 6.17(b)]. The node basis function associated with the wire end is grouped with four edge basis functions associated with the short plate edges into four doublets [Figure 6.31(b)]. Each doublet consists of the node basis function and an edge basis function. The total current of the edge basis function flowing out from the plate is equal to the current given by the node basis function flowing into the wire. Similarly, we satisfy the KCL for the junction domain of a wire protruding out of a dielectric surface.

In order that normal continuity can be imposed along a junction (common edge) of a quadrilateral and a triangle, the normal components of edge functions defined over the quadrilateral and triangle should vary in the same way along the edge. The normal component of the ith edge basis function defined over the quadrilateral varies along the edge $s = \pm 1$ as p^i (normalized with the Lame coefficient along the edge). The normal component of the ith edge basis function defined over a triangle varies along edge $t_3 = 0$ as $(t_1 t_2)^k$ $(i = 2k)$ and $(t_1 t_2)^k (t_1 - t_2)$ $(i = 2k + 1)$ (normalized with the Lame coefficient along the edge). The simplex coordinates t_1 and t_2 along the edge can be expressed in terms of a single parametric coordinate along the edge, as explained for segments in Section 6.1.4. In that case, $(t_1 t_2)^k = (1 - p^2)^k$ and $(t_1 t_2)^k (t_1 - t_2) = (1 - p^2)^k p$. Only the first two edge basis functions (i.e., for $i = 0$ and $i = 1$) defined over the quadrilateral and the triangle can be directly matched and grouped into doublets to satisfy the continuity in a local way. For example, for $i = 0$ we obtain a rooftop basis function. If we want to match the edge basis functions of higher orders, we must first rearrange them.

Current (field) expansions in a tetrahedron and a hexahedron can be matched only by introducing a new building element between them in the form of a prism. The current (field) expansion in the prism is obtained by combining the elements of the expansions in the hexahedron and the tetrahedron [53].

A combination of volume and surface building elements is used when the VIE is combined with the SIE, and when the FEM is combined with the SIE. In the first case, expansions for the electric flux in volume building elements need not be related in any way to expansions for currents over surface building elements, even when a surface building element touches a volume building element. In the second case, surface building elements represent the boundary surface of the region occupied by volume building elements. The electric (magnetic) field inside the volume is related to equivalent magnetic (electric) currents over the volume boundary through the boundary conditions (3.61a-f). In order that the boundary conditions are automatically satisfied, surface basis functions, $\mathbf{F}_S$, and volume basis functions, $\mathbf{F}_V$, should be related as

$$\mathbf{F}_S = \mathbf{n} \times \mathbf{F}_V \text{ (at surface } S) \tag{6.85}$$

It can be shown that face basis functions defined for generalized hexahedrons and edge and patch basis functions defined for generalized quadrilaterals automatically satisfy this relation. The same can be shown for vector basis functions defined for generalized triangles and tetrahedrons in Sections 6.3 and 6.5.

6.7 Comparison of MoM/SIE, MoM/VIE, and FEM Based on Topological Analysis

It is well known that there exists no unique method that most efficiently covers all kinds of electromagnetic problems. The following statements are generally considered to be true:

- The MoM is more convenient for open problems and the FEM is more convenient for closed problems.
- The MoM/SIE is well suited to wire and plate structures, while the MoM/VIE and the FEM are well suited to dielectric structures.
- The MoM/SIE requires fewer unknowns than the MoM/VIE and the FEM.
- In the FEM (sparse matrices), the system of linear equations is solved much faster than in the MoM/SIE and the MoM/VIE (full matrices).
- The MoM is more efficient for small problems, while the FEM is more efficient for large problems.

However, these conclusions are not always correct. Moreover, they do not give quantitative measures that can be used for deciding which method is optimal for the solution of a specific problem. A quantitative measure can be obtained from the topological analysis of two canonical problems [54].

Let us consider a cubical body of side a, consisting of $m_d \times m_d \times m_d$ smaller cubes made of different linear isotropic and homogenous dielectrics (Figure 6.32). The first canonical problem is when the body is completely covered by the PEC. This is a closed problem. Otherwise, we have the second canonical problem, which is an open problem. For the closed problem, the MoM/VIE (applied to the dielectric) must be combined with the MoM/SIE (applied to the PEC). Similarly, for the open problem, it is supposed that the weak FEM formulation is combined with the MoM/SIE. In what follows, these hybrid methods will be simply called MoM/VIE and FEM.

When the analysis is performed by the MoM/SIE, each cube face of edge $b = a/m_d$ is subdivided into $m_p \times m_p$ patches. When the analysis is performed

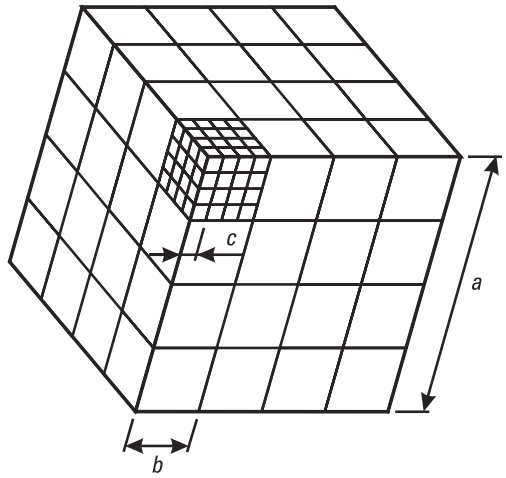

Figure 6.32 Cubical body of side a consisting of $m_d \times m_d \times m_d$ smaller cubes made of different linear isotropic and homogenous dielectrics.

by the MoM/VIE or the FEM, each cube of edge $b = a/m_d$ is subdivided into $m_p \times m_p \times m_p$ cubical bricks. In both cases, m_p depends on the electrical length of the edge b and the order of approximation used along this edge.

Approximations used in the MoM/SIE, the MoM/VIE, and the FEM satisfy the normal continuity at the patch edges, the normal continuity at the brick faces, and the tangential continuity at the brick faces, respectively. In general, let the expansion of order n_p be used in each direction of the square patch (cube) of edge $c = b/m_p = a/(m_p m_d)$. Finally, let us suppose that, for a given accuracy, the same order n_p of approximation is required by all methods considered.

After simple manipulations [54], we obtain the number of unknowns needed to solve open and closed problems by the MoM/SIE, the MoM/VIE, and the FEM as

$$N_{\text{open}}^{(\text{SIE})} = 12n^2 m_d + 12n^2 - 6nm_d^2 + 6n \qquad N_{\text{close}}^{(\text{SIE})} = 12n^2 m_d - 6nm_d^2 + 6n \tag{6.86}$$

$$N_{\text{open}}^{(\text{VIE})} = 3n^3 + 3n^2 \qquad N_{\text{close}}^{(\text{VIE})} = 3n^3 + 15n^2 \tag{6.87}$$

$$N_{\text{open}}^{(\text{FEM})} = 3n^3 + 18n^2 + 3n \qquad N_{\text{close}}^{(\text{FEM})} = 3n^3 + 6n^2 + 3n \tag{6.88}$$

where $n = n_p m_p m_d$. By comparing above formulas, we see that

$$N_{\text{open}}^{(\text{SIE})} < N_{\text{open}}^{(\text{FEM})} \text{ for } n_d > 3.4 - \frac{2.4}{m_d} \tag{6.89}$$

$$N_{\text{close}}^{(\text{SIE})} < N_{\text{close}}^{(\text{FEM})} \text{ for } n_d > 3.4 + \frac{2.4}{m_d} \tag{6.90}$$

$$N_{\text{open}}^{(\text{SIE})} < N_{\text{open}}^{(\text{VIE})} \text{ for } n_d > 3.4 + \frac{3.6}{m_d} \tag{6.91}$$

$$N_{\text{close}}^{(\text{SIE})} < N_{\text{close}}^{(\text{VIE})} \text{ for } n_d > 3.4 - \frac{6}{m_d} \tag{6.92}$$

where $n_d = n/m_d$ is the number of unknowns per edge b. Finally, the following can be concluded:

- The MoM/VIE requires fewer unknowns than the FEM for open problems.
- The FEM requires fewer unknowns than the MoM/VIE for closed problems.
- The MoM/SIE requires fewer unknowns than the MoM/VIE and FEM if n_d is large enough. In particular, for large m_d, this occurs if $n_d > 3$. The MoM/VIE and the FEM require fewer unknowns than the MoM/SIE only for highly inhomogeneous structures.

However, the efficiency of the solution depends not only on the number of unknowns, but also on the total number of arithmetic operations. For complex and electrically large problems, the main part of the number of operations is due to the solution of the system of linear equations. For full and sparse matrices, the number of basic operations, K, can be estimated as

$$K_{\text{full}} = \frac{1}{3}N^3 \qquad K_{\text{sparse}} = NL^2 \tag{6.93}$$

where N is the number of unknowns, L is the half-bandwidth of the sparse matrix, and a basic operation consists of one summation and one multiplication. For complex ($m_d >> 1$) and electrically large problems ($n >> 1$), the total number of unknowns can be written in approximate form as

$$N^{(\text{SIE})} = 12n^2 m_d - 6nm_d^2 \qquad N^{(\text{FEM})} = N^{(\text{VIE})} = 3n^3 \tag{6.94}$$

For the MoM/VIE, matrices are always full. However, for both the FEM and the MoM/SIE, the unknown coefficients can be numerated so that matrices are sparse. For closed problems, the half-bandwidths for the FEM and the MoM/SIE are

$$L_{\text{close}}^{(\text{FEM})} = \frac{N}{m_p m_d} \qquad L_{\text{close}}^{(\text{SIE})} = 2\frac{N}{m_d} \tag{6.95}$$

Obviously, K for sparse matrices is lower than K for full matrices only if $L > N/\sqrt{3}$. For large problems, $m = m_p m_d$ is always much greater than 1. Hence, the FEM matrices should be always treated as sparse. However, for the MoM/SIE, the matrices should be treated as sparse only if $m_d \geq 2$, which is always satisfied for complex problems.

Comparing (6.93) to (6.95), the total number of operations used by the MoM/SIE can be related to the total number of operations used by the FEM as

$$K^{(\text{SIE})} = K^{(\text{FEM})} \frac{256}{n_p^2 n_d}\left(1 - \frac{1}{2n_d}\right)^3 \tag{6.96}$$

The same relation is approximately valid for open structures. It is seen from (6.96) that, for large enough n_d, the MoM/SIE is more efficient than the FEM; that is,

$$K^{(\text{SIE})} < K^{(\text{FEM})} \qquad n_d \geq \max\left(n_p, \frac{256}{n_p^2}\right) \tag{6.97a,b}$$

In particular, the MoM/SIE is more efficient than the FEM if higher-order expansions such that $n_p > 6$ are applied. Such expansions should be used only if homogenous parts of bodies are electrically large enough. Usually, orders $n_p > 6$ are used for dimensions greater than 0.5 to 1λ. (For $n_p = 1, 2, 3$, $n_d \geq 256, 64, 28$, respectively.)

However, for a given accuracy, the MoM/SIE usually requires a lower order of approximation n_p than the FEM. In that case (6.97a) is satisfied for lower n_p and n_d. This can be explained by the fact that the field inside a homogeneous material body, which is approximated by volume basis functions, is discontinuous (the normal continuity is not satisfied), while a discontinuous approximation of currents over the surface of the body results in a continuous field inside the body. Hence, we find an advantage to the MoM/SIE with

respect to the MoM/VIE and the FEM, except for highly inhomogeneous structures.

6.8 Summary

The approximation of the currents (fields) is tightly connected with the equation and the method that is used for the analysis. For the MoM/SIE, the geometry of the structure is defined by surfaces of generalized wires, quadrilaterals, and triangles. The corresponding surface currents are assumed to have one, two, and three surface components, respectively, expressed in terms of one, two, and three local coordinates. For the MoM/VIE and the FEM, the geometry of the structure is defined by domains of generalized hexahedrons and tetrahedrons. The corresponding volume currents (fields) are assumed to have three and four components, respectively, expressed in terms of three and four parametric coordinates.

For all these expansions, it is desirable that they be easily implemented and lead to sufficiently accurate, stable, and efficient solutions. To achieve these goals, some general rules must be obeyed. Regarding the ease of implementation, it is convenient to describe the approximate currents in local parametric coordinate systems, the same ones with respect to which the building elements are specified. In particular, it is recommended that using orthogonal and interpolation polynomials (Legendre, Chebyshev, Lagrange, and so forth) be avoided. The orthogonality properties of the polynomial basis functions cannot be exploited to facilitate the analysis. The interpolation polynomials lead to nonhierarchical basis functions. In both cases, the expansions are more complex than those based on power functions. Beside power functions, we recommend using simple trigonometric functions. If properly combined with the power functions, they enable extremely efficient analysis of thin-wire structures. However, such a combination is much more complicated to implement than expansions based on power functions only. In addition, there is no adequate combination of power and trigonometric functions for surfaces and volumes. Hence, in this chapter we recommend using simple expansions made of power functions.

The stability, accuracy, and efficiency of the results are significantly improved if the approximations are forced to satisfy the normal continuity (MoM/SIE, MoM/VIE) or the tangential continuity (FEM). To satisfy the normal continuity, the unknown quantity should be expanded along unitary vectors in a local parametric coordinate system. To satisfy the tangential continuity, it should be expanded along inverse unitary vectors. In the next step, the expansions are rearranged, resulting in node, edge (wedge), patch (face), and brick basis functions. By proper grouping of the basis functions into doublets

and multiplets, either the normal or tangential continuity are automatically satisfied. In addition to the normal (tangential) continuity, we can incorporate quasistatic relations into the expansions. However, this inclusion makes the expansions much more difficult to implement, often without substantial improvement in the accuracy of practically important quantities like impedance, gain, and so forth.

In particular, the efficiency of such expansions is increased if higher-order basis functions are used for electrically larger building elements, and if the concept of entire-domain expansions is used instead of the concept of subdomain expansions. According to the entire-domain approach, currents (fields) are approximated by a single expansion throughout a building element. The order of approximation in each coordinate direction of a building element is proportional to the maximal electrical dimension of the building element in this direction. The accuracy of results is increased by increasing the order of approximation.

It is desirable that the building element has as regular shape as possible, when it occupies the maximal volume for the given orders of the approximation. According to this criterion, we need more unknowns and operations for electromagnetic modeling based on volume building elements (MoM/VIE and FEM) than the one based on surface building elements (MoM/SIE), except for highly inhomogeneous structures. According to this criterion, we need more unknowns for surfaces (volumes) modeled by regular triangles (tetrahedrons) than by regular quadrilaterals (hexahedrons). However, the geometrical modeling of surfaces (volumes) by triangles (tetrahedrons) is easier than that by quadrilaterals (hexahedrons). Hence, there is a necessity to combine various kinds of building elements (i.e., to hybridize various methods) to provide flexible geometrical modeling and maximally efficient analysis. In that case, special care should be devoted to ensure the compatibility of various building elements. Hybrid methods that use various kinds of building elements have difficult implementation. Hence, in this book we propose the MoM/SIE in conjunction with surface modeling by quadrilaterals (and wire modeling by right truncated cones) as the basic method for electromagnetic modeling in the frequency domain.

References

[1] Richmond, J. H., "A Wire Grid-Model for Scattering by Conducting Bodies," *IEEE Trans. Antennas Propagat.,* Vol. AP-14, November 1966, pp. 782–786.

[2] Miller, E. K., and F. J. Deadrick, "Some Computational Aspects of Thin-Wire Modeling," in *Numerical and Asymptotic Techniques in Electromagnetics,* R. Mittra, (ed.), New York: Springer-Verlag, 1975.

[3] Thiele, G. A., "Wire Antennas," in *Computer Techniques for Electromagnetics,* R. Mittra, (ed.), Elmsford, NY: Pergamon Press, 1973.

[4] Harrington, R. F., *Field Computation by Moment Methods,* New York: Macmillan, 1968; reprinted by New York: IEEE Press, 1993.

[5] Chao, H. H., B. J. Strait, and C. D. Taylor, "Radiation and Scattering by Configurations of Bent Wires with Junctions," *IEEE Trans. Antennas Propagat.,* Vol. AP-19, No. 5, September 1971, pp. 701–702.

[6] Strait, B. J., and K. Hirasawa, "Array Design for a Specified Pattern by Matrix Methods," *IEEE Trans. Antennas Propagat.,* Vol. AP-17, No. 2, March 1969, pp. 237–239.

[7] Kyle, R. H., "Mutual Coupling Between Log-Periodic Antennas," *IEEE Trans. Antennas Propagat.,* Vol. AP-18, No. 1, September 1971, pp. 15–22.

[8] Richmond, J. H., *Computer Analysis of Three-Dimensional Wire Antennas,* Ohio State University, ElectroScience Laboratory, Department of Electrical Engineering, Report 2708–4, 1969.

[9] Richmond, J. H., "Monopole Antenna on Circular Disc," *IEEE Trans. Antennas Propagat.,* Vol. AP-32, No. 12, December 1984, pp. 1282–1287.

[10] Otto, D. V., and J. H. Richmond, "Rigorous Field Expression for Piecewise Sinusoidal Line Source," *IEEE Trans. Antennas Propagat.,* Vol. AP-17, No. 1, January 1969, p. 98.

[11] Richmond, J. H., and H. Geary, "Mutual Impedance of Non-Planar-Skew Sinusoidal Dipoles," *IEEE Trans. Antennas Propagat.,* Vol. AP-23, No. 1, May 1975, pp. 412–414.

[12] Imbriale, W. A., "Applications of the Method of Moments to Thin Wire Elements and Arrays," in *Numerical and Asymptotic Techniques in Electromagnetics,* R. Mittra, (ed.), Berlin, Germany: Springer-Verlag, 1975.

[13] Neureuther, A. R., et al., "A Comparison of Numerical Methods for Thin Wire Antennas," *Digest of the 1968 URSI Fall Meeting,* Berkeley, CA, 1968.

[14] Miller, E. K., et al., "Radar Cross Section of a Long Wire," *IEEE Trans. Antennas Propagat.,* Vol. AP-17, No. 3, May 1969, pp. 381–384.

[15] Yeh, Y. S., and K. K. Mei, "Theory of Conical Equiangular-Spiral Antennas: Part I—Numerical Technique," *IEEE Trans. Antennas Propagat.,* Vol. AP-15, No. 5, September 1967, pp. 634–639.

[16] Gee, S., et al., "Computer Techniques for Electromagnetic Scattering and Radiation Analysis," *IEEE Internat. Electromagn. Compat. Symp. Rec.,* 1971, pp. 122–131.

[17] Kolundzija, B. M., and B. D. Popovic, "Entire-Domain Galerkin Method for Analysis of Generalized Wire Antennas and Scatterers," *Proc. IEE H,* Vol. 139, No. 1, February 1992, pp. 17–24.

[18] Richmond, J. H., "Digital Computer Solutions of the Rigorous Equations for Scattering Problems," *Proc. IEEE,* Vol. 53, 1965, pp. 796–804.

[19] Popovic, B. D., M. B. Dragovic, and A. R. Djordjevic, *Analysis and Synthesis of Wire Antennas,* Chichester, U.K.: Research Studies Press, 1982.

[20] Silvester, P. P., and K. K. Chan, "Analysis of Antenna Structures Assembled from Arbitrarily Located Straight Wires," *Proc. IEE,* Vol. 120, No. 1., January 1973, pp. 21–26.

[21] Popovic, B. D., "Polynomial Approximation of Current Along Thin-Symmetrical Cylindrical Dipoles," *Proc. IEE,* Vol. 117, 1970, pp. 873–878.

[22] Merewether, D. E., "The Arbitrarily Driven Long Cylindrical Antenna," *IEEE Trans. Antennas Propagat.,* Vol. AP-23, No. 4, July 1975, pp. 534–540.

[23] Kolundzija, B. M., "Comparison of a Class of Sub-Domain and Entire-Domain Basis Functions Automatically Satisfying KCL," *IEEE Trans. Antennas Propagat.,* Vol. 44, No. 10, October 1996, pp. 1362–1366.

[24] Kolundzija, B. M., et al., *Electromagnetic Modeling of Composite Wire and Plate Structures—Software and User's Manual,* Norwood, MA: Artech House, 1995.

[25] Dragovic, M. B., and B. D. Popovic, "Combined Trigonometric and Polynomial Expansion for Current Along Cylindrical Antennas of Large Length," *Proc. 1st ICAP,* London, England, 1978, pp. 408–412.

[26] Kolundzija, B. M., "Effect of a Wire End in Thin-Wire Analysis," *Proc. IEEE AP-S Symposium,* Syracuse, NY, 1988, pp. 843–846.

[27] De Smedt, R., and J. G. Van Bladel, "Field Singularities at the Tip of a Metallic Cone of Arbitrary Cross Section," *IEEE Trans. Antennas Propagat.,* Vol. AP-34, 1986, pp. 865–871.

[28] Wu, T. T., and R. W. P. King, "Tapered Antenna and Its Application to the Junction Problem for Thin Wires," *IEEE Trans. Antennas Propagat.,* Vol. AP-24, No. 1, January 1975, pp. 42–45.

[29] Knepp, D. L., and J. Goldhirsh, "Numerical Analysis of Electromagnetic Radiation Properties of Smooth Conducting Bodies of Arbitrary Shape," *IEEE Trans. Antennas Propagat.,* Vol. AP-29, No. 3, May 1981, pp. 479–487.

[30] Glisson, A. W., and D. R. Wilton, "Simple and Efficient Numerical Methods for Problems of Electromagnetic Radiation and Scattering from Surfaces," *IEEE Trans. Antennas Propagat.,* Vol. AP-28, No. 5, September 1980, pp. 593–603.

[31] Wang, N. N., J. H. Richmond, and M. C. Gilreath, "Sinusoidal Reaction Formulation for Radiation and Scattering from Conducting Surfaces," *IEEE Trans. Antennas Propagat.,* Vol. AP-23, No. 3, May 1975, pp. 376–382.

[32] Newman, E. H., and D. M. Pozar, "Electromagnetic Modeling of Composite Wire and Surface Geometries," *IEEE Trans. Antennas Propagat.,* Vol. AP-26, No. 6, November 1978, pp. 784–789.

[33] Singh, J., and A. T. Adams, "A Nonrectangular Patch Model for Scattering from Surfaces," *IEEE Trans. Antennas Propagat.,* Vol. AP-27, No. 4, July 1979, pp. 531–535.

[34] Newman, E. H., and P. Tulyathan, "A Surface Patch Model for Polygonal Plates," *IEEE Trans. Antennas Propagat.,* Vol. AP-30, No. 4, July 1982, pp. 588–593.

[35] Kolundzija, B. M., and B. D. Popovic, "Entire-Domain Galerkin Method for Analysis of Metallic Antennas and Scatterers," *Proc. IEE,* Pt. H, Vol. 140, No. 1, February 1993, pp. 1–10.

[36] Kolundzija, B. M., "On the Locally Continuous Formulation of Surface Doublets," *IEEE Trans. Antennas Propagat.,* Vol. AP-46, No. 12, December 1998, pp. 1879–1883.

[37] Kolundzija, B. M., "Electromagnetic Modeling of Composite Metallic and Dielectric Structures," *IEEE Trans. Microwave Th. Tech.,* MTT-47, No. 7, July 1999, pp. 1021–1032.

[38] Kolundzija, B. M., "Accurate Solution of Square Scatterer as Benchmark for Validation of Electromagnetic Modeling of Plate Structures," *IEEE Trans. Antennas Propagat.,* Vol. AP-46, July 1998, pp. 1009–1014.

[39] Sarkar, T. K., E. Arvas, and S. Ponnapalli, "Electromagnetic Scattering from Dielectric Bodies," *IEEE Trans. Antennas Propagat.,* Vol. AP-37, May 1989, pp. 673–676.

[40] Van Bladel, J., *Electromagnetic Fields,* New York: McGraw-Hill, 1964.

[41] Kolundzija, B. M., "On the Inclusion of Edge Effect into Surface Vector Basis Functions," *Proc. of IEEE AP-S Int. Symp.,* Atlanta, GA, 1998, pp. 282–285.

[42] Rao, S. M., D. R. Wilton, and A. W. Glisson, "Electromagnetic Scattering by Surfaces of Arbitrary Shape," *IEEE Trans. Antennas Propagat.,* Vol. AP-30, No. 3, May 1982, pp. 409–418.

[43] Virga, K. L., and Y. Rahmat-Samii, "RCS Characterization of a Finite Ground Plane with Perforated Apertures: Simulations and Measurements," *IEEE Trans. Antennas Propagat.,* Vol. AP-42, No. 3, July 1994, pp. 1491–1501.

[44] Peterson, A. F., S.C. Ray, and R. Mittra, *Computational Methods for Electromagnetics,* New York: IEEE Press and Oxford, U.K.: Oxford University Press, 1998.

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

[46] Nedelec, J. C., "Mixed Finite Elements in R3," *Num. Math.,* Vol. 35, 1980, pp. 315–341.

[47] Schaubert, D. H., D. R. Wilton, and A. W. Glisson, "A Tetrahedral Modeling Method for Electromagnetic Scattering by Arbitrarily Shaped Inhomogeneous Dielectric Bodies," *IEEE Trans. Antennas Propagat.,* Vol. AP-32, January 1984, pp. 77–85.

[48] Jin, J., *The Finite Element Method in Electromagnetics,* New York: John Wiley and Sons, 1993.

[49] Notaros, B. M., "Numerical Analysis of Dielectric Bodies of Arbitrary Inhomogeneity in the Electromagnetic Field," Ph.D. Thesis, University of Belgrade, 1995 (in Serbian).

[50] Crowly, C. W., P. P. Silvester, and H. Hurwitz, "Covariant Projection Elements for 3-D Vector Field Problems," *IEEE Trans. Magnetics,* Vol. 24, 1988, pp. 397–400.

[51] Silvester, P. P., and R. L. Ferrari, *Finite Elements for Electrical Engineers,* 3rd ed., New York: Cambridge University Press, 1996.

[52] Webb, J. P., and B. Forghani, "Hierarchical Scalar and Vector Tetrahedra," *IEEE Trans. Magnetics,* Vol. 29, No. 2, 1993, pp. 1495–1498.

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

[54] Kolundzija, B. M., V. V. Petrovic, and A. R. Djordjevic, "Comparison of MoM/SIE, MoM/VIE and FEM Based on Topological Analysis of Two Canonical Problems," *Proc. of IEEE AP-S Int. Symp.,* Atlanta, GA, 1998, pp. 278–281.

7

Treatment of Excitations

In practice, we excite structures in many ways. From the theoretical point of view, any excitation can be represented by impressed currents or impressed fields, possibly combined with the modeling of the complete feeding structure. A precise model of the excitation in conjunction with a precise model of the whole structure leads to accurate modeling of the electromagnetic fields at all spatial points. However, we are often interested only in one part of this field. This part can usually be obtained using a simplified model of the excitation.

The goal in modeling excitation is to represent it in the simplest way, which still enables accurate results for the response that we are interested in. There are three general types of problems to be considered here: scattering of free-space waves from structures, excitation of structures by voltage and current generators, and excitation of structures by guided waves. The excitation by free-space waves is given in Section 7.1. Voltage and current generators are elaborated in Section 7.2, with special care devoted to delta-function generators. Modeling of guided waves is presented in Section 7.3 with emphasis on the TEM magnetic-current frill approximation of a coaxial-line feed over infinite and finite ground plane.

7.1 Free-Space Waves

The source of a free-space wave is at a large distance from the structure analyzed. Hence, there is no coupling between the wave source and the structure. We can describe the excitation simply by impressed electric and magnetic field vectors, $\mathbf{E}_i$ and $\mathbf{H}_i$.

Among free-space waves, a plane wave is most often used. We use it whenever the far field of a transmitting antenna excites a structure situated in a homogeneous linear and isotropic medium. Let the permittivity and the permeability of the medium be ϵ and μ. (Such a medium can be lossy, in which case these parameters are complex.) Further, assume that the wave of angular frequency ω is incident on the structure from the direction defined by spherical angles ϕ and θ (Figure 7.1). The unit vector of this direction is expressed in terms of the spherical angles, as

$$\mathbf{n} = -\sin\theta\cos\phi\,\mathbf{i}_x - \sin\theta\sin\phi\,\mathbf{i}_y - \cos\theta\,\mathbf{i}_z \tag{7.1}$$

In general, the wave is elliptically polarized; that is, the electric field vector has both the ϕ and θ components, which are not in phase. Let $E_{0\phi}$ and $E_{0\theta}$ be the ϕ and θ components of the electric field at the coordinate origin. The electric field vector at the origin is obtained in the form

$$\mathbf{E}_0 = E_{0\phi}\,\mathbf{i}_\phi + E_{0\theta}\,\mathbf{i}_\theta \tag{7.2}$$

where the unit vectors $\mathbf{i}_\phi$ and $\mathbf{i}_\theta$ are expressed in terms of the spherical angles as

$$\mathbf{i}_\phi = -\sin\phi\,\mathbf{i}_x + \cos\phi\,\mathbf{i}_y \qquad \mathbf{i}_\theta = \cos\theta\cos\phi\,\mathbf{i}_x + \cos\theta\sin\phi\,\mathbf{i}_y - \sin\theta\,\mathbf{i}_z \tag{7.3}$$

Finally, the electric and magnetic fields at an arbitrary field point have the form

$$\mathbf{E}_i(\mathbf{r}) = \mathbf{E}_0 e^{-j\beta\mathbf{r}\cdot\mathbf{n}} \qquad \mathbf{H}_i(\mathbf{r}) = \frac{1}{Z}\,\mathbf{n}\times\mathbf{E}_i(\mathbf{r}) \tag{7.4a,b}$$

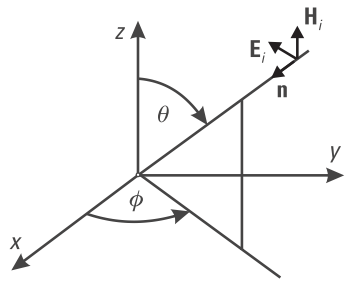

Figure 7.1 Plane wave excitation.

where $\mathbf{r} = x\mathbf{i}_x + y\mathbf{i}_y + z\mathbf{i}_z$ is the position vector of the field point, $\beta = \omega\sqrt{\epsilon\mu}$ is the phase coefficient of the medium, and $Z = \sqrt{\mu/\epsilon}$ is the intrinsic impedance of the medium.

In some cases, we use other types of free-space waves [1]. For example, a paraboloidal reflector is illuminated by the spherical wave of a primary feed antenna. Such a spherical wave can be expressed as

$$\mathbf{E}_i(\mathbf{r}) = C\frac{e^{-j\beta r}}{r}\mathbf{F}(\phi, \theta),\ r = |\mathbf{r}| \tag{7.5}$$

where C is a properly chosen constant, and $\mathbf{F}(\phi, \theta)$ is the characteristic function of the primary feed antenna. The corresponding magnetic field is expressed by (7.4b).

7.2 Voltage and Current Generators

A port of a structure has the form of two metallic terminals placed in a dielectric at an electrically small distance (Figure 7.2). The port can be excited in many ways, using various distributions of impressed currents or impressed fields localized in the vicinity of the terminals. The electrically small excitation region that contains the terminals and the impressed sources represents a lumped (concentrated) generator. For brevity, we shall refer to it as a generator.

Generators differ regarding the distribution of the impressed sources and the shape of the terminals. The terminals are usually modeled in the same way

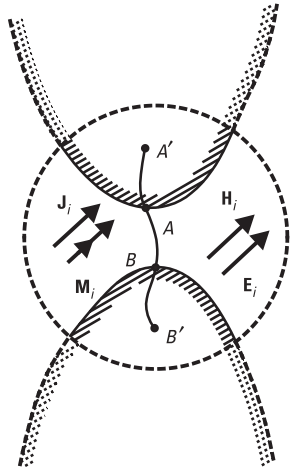

Figure 7.2 The definition of generator.

as the rest of the structure (e.g., by right truncated cones). Hence, from the implementation point of view, the generators differ regarding the distribution of the impressed sources.

To facilitate numerical modeling, the metallic terminals are usually considered to be made of a PEC, and the dielectric between the terminals is considered to be lossless. Electric currents and charges are induced over the terminal surfaces as a result of excitation by the impressed sources. The port represents an extreme discontinuity of the structure, which causes a very strong local electric field. This field is due predominantly to charges of the opposite sign induced on the terminals (i.e., the field due to the other parts of the structure can be neglected). Since the excitation region is electrically small, the field in the region is quasi-static. Hence, the total current going out of one terminal is equal to the total current going into the other terminal. Also, all points of a terminal surface are at the same potential.

Since the excitation region is electrically small, the quantities of interest, like the ratio of the voltage and the total current at the input terminals and the relative distribution of currents and fields outside the excitation regions, are practically independent of the distribution of the impressed sources. It is not necessary to include various types of localized impressed currents into the electromagnetic modeling. The question is which distribution of impressed sources is optimal from the implementation and postprocessing perspectives. Two classes of generators are convenient: the ideal current generator, which forces the desired total current at the terminals, and the ideal voltage generator, which forces the desired voltage between the terminals.

The ideal current generator is easily implemented in the MoM solution of the field integral equations by requiring that the approximation of the total current be equal to the desired value at the terminals. The ideal voltage generator is easily implemented if it is defined as a delta-function generator (see Section 7.2.1). The data postprocessing is easier for the ideal voltage generator. When the excitation is performed by the ideal voltage generator and the field integral equations are solved by the MoM, we know both the voltage and the current at input terminals. For the ideal current generator, only the currents are known and the voltage should be determined by postprocessing [2]. The results for any combination of ideal currents or voltage generators can be obtained using only one class of generators [3]. Hence, we recommend delta-function generators as the basic type of generators in electromagnetic modeling of composite metallic and dielectric structures.

7.2.1 Delta-Function Generator

The delta-function generator is a point-like ideal voltage generator. Originally, it was used in the analytical solution of the Halen equation [4]. There are

several ways in which it can be defined. The simplest is to require that the potential difference between two infinitely close points A and B of the generator terminals be equal to the desired generator EMF, whatever the shape of the terminals. The precise specification of impressed sources can be avoided by requiring that the circulation of the impressed electric field along the contour A′ABB′ shown in Figure 7.2 be equal to the desired EMF.

The ratio of the voltage and current at the input terminals (which equals the input impedance for one-port structures) partly depends on the actual shape of the input terminals. The delta function possesses serious theoretical and, in some instances, numerical problems if it is assumed to excite terminals that have the form of solid or hollow cylinders. Consider a wire-dipole antenna driven by such a generator [Figure 7.3(a)]. If the width of the gap δ tends to zero, the capacitance between the dipole arms becomes infinitely large. Therefore, the capacitive current component also becomes infinite. A low-order approximation of current along the dipole arms usually cannot follow this sharp rise in current intensity in the immediate vicinity of the generator. The imperfect current approximation therefore makes acceptable the otherwise meaningless generator model and is even known to yield relatively accurate impedance values. However, for a higher-order current approximation, the antenna susceptance becomes divergent.

These difficulties are avoided if the terminals are made in the form of two conical segments and the delta-function generator is connected to the cone apexes [Figure 7.3(b)]. If the distance between the apexes, δ, tends to zero, the capacitance between two conical segments tends to some finite and relatively small value. The current varies relatively slowly in the immediate vicinity of the generator. Hence, both the low-order and higher-order current approximations give accurate and stable results.

For example, consider a quarter-wavelength monopole antenna of height-to-radius ratio $h/a = 10$, above a perfectly conducting ground plane, base-

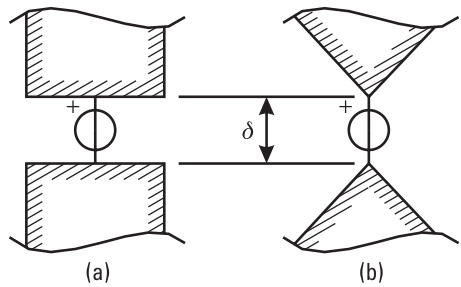

Figure 7.3 Delta-function generator with terminals in the form of (a) solid cylinders and (b) cones.

driven by a delta-function generator. The antenna feeding area was modeled in two ways, by a short cylinder and by a cone of height a, as shown by the antenna symmetrical equivalent in the inset of Figure 7.4. The end of the monopole was modeled by a flat disc. The currents along the feeding cylinder (cone) and ending flat disc are approximated by first-order polynomials. The current along the rest of the monopole was adopted as a polynomial of degree N. The solution of the EFIE was performed by the Galerkin method. Figure 7.4 shows the antenna admittance plotted against N. For the cylindrical feed, the susceptance dramatically increases with increasing N. For the conical feed, the results are practically independent of N. However, if a relatively thin antenna (e.g., $h/a = 30$) is modeled by a single cylinder, stable results for the admittance are obtained.

In practice, we usually tend to decrease the parasitic capacitance between input terminals. If the excitation region of an antenna is electrically small and small with respect to the overall antenna size, the influence of this capacitance on the input impedance of the antenna is negligible (as will be shown in Section 7.3). In all these cases, we can use a delta-function generator with terminals in the form of two conical segments or even two cylindrical segments. If the excitation region of an antenna is not electrically small, additional care should be devoted to the modeling of the terminals and impressed sources. Alternatively, the excitation should be treated as excitation by guided waves.

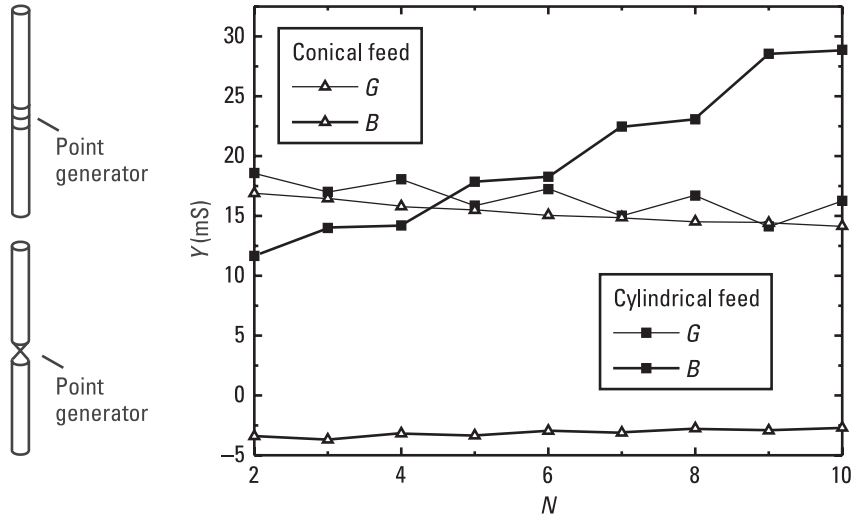

Figure 7.4 Admittance of a quarter-wavelength monopole antenna ($h/a = 10$) against the degree N of the polynomial approximation for currents along the main part of the monopole. The symmetrical equivalents of the monopole with two models of the excitation region are sketched in the inset.

7.3 Guided Waves

Very often, antennas and microwave devices are excited by guided waves. The device is connected to one or more guiding structures (transmission lines or waveguides). The basic assumptions for the guiding structures are that they are uniform (i.e., that the geometry and material properties are the same for all cross sections), lossless, and that only the dominant mode can propagate. The excitation of the device is performed using the dominant mode propagating along the guiding structure towards the device (called the direct, incident, or forward wave). When the dominant mode reaches the region in which the guiding structure is connected to the device, it is partly transmitted into the device and partly reflected (back to the guiding structure). The reflected part consists of the dominant mode that propagates backward (reflected or backward wave) and higher-order modes. Since the operating conditions are selected so that only the dominant mode can propagate, all higher-order modes are evanescent.

Consider the treatment of the excitation by guided waves when the MoM is applied to solve the SIEs. In general, such analysis includes not only determination of induced and equivalent currents over the boundary surfaces of the device, but also determination of the reflected wave and higher-order modes inside the guiding structure (see Section 7.3.1). In some cases, the influence of the higher-order modes on the field solution inside the structure and its vicinity can be neglected, thus enabling the excitation by the dominant guided wave to be modeled by equivalent (electric and magnetic) currents at the opening of the guiding structure. A typical case is the excitation by a TEM wave that propagates along a coaxial line (see Section 7.3.2).

7.3.1 Exact Modeling

Consider the junction of a guiding structure and a device such that the guiding structure is coupled with the device only through the opening, as is shown by the dashed line in Figure 7.5(a).[1] Let the z-axis be located along the guiding structure and the opening placed in the xOy-plane. The electric field inside the guiding structure can be approximately represented as a series of dominant modes (direct and reflected) and the first n higher-order modes in the form

$$\mathbf{E}(x, y, z) = a_d \mathbf{E}_0(x, y) e^{-\gamma_0 z} + a_r \mathbf{E}_0(x, y) e^{\gamma_0 z} + \sum_{i=1}^{n} a_i \mathbf{E}_i(x, y) e^{\gamma_i z} \tag{7.6}$$

1. Examples are shielded structures like coaxial lines and rectangular waveguides.

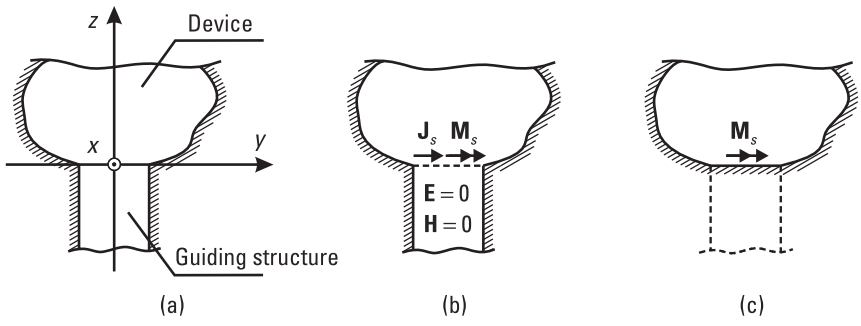

Figure 7.5 A junction of a guiding structure and a device: (a) original case, (b) the guiding structure is replaced by equivalent electric and magnetic currents, and (c) the opening is closed by a PEC.

where $\mathbf{E}_i(x, y)$ and γ_i are the electric field at the opening and the propagation coefficient of the ith mode, respectively; $i = 0$ corresponds to dominant mode; a_d is a given complex weighting coefficient; and a_r and a_i are unknown coefficients to be determined. In a similar way, we can represent the magnetic field inside the guiding structure.[2]

The influence of the electromagnetic field inside the guiding structure on the outside electromagnetic field can be modeled by replacing the field inside the guiding structure with equivalent electric and magnetic currents at the opening [Figure 7.5(b)]. The equivalent currents are determined as

$$\mathbf{J}_s(x, y) = \mathbf{i}_z \times \mathbf{H}(x, y, z = 0) \qquad \mathbf{M}_s(x, y) = -\mathbf{i}_z \times \mathbf{E}(x, y, z = 0) \tag{7.7a,b}$$

where $\mathbf{i}_z$ is the unit vector along the z-axis. These equivalent currents represent the excitation of the device. This excitation can be further simplified as follows. The zero field inside the guiding structure will not change if we put a PEC surface over the opening just inside the structure, leaving the equivalent currents outside. The equivalent electric currents are now shorted by the PEC [6]. These currents can be removed and replaced by identical electric currents induced in the PEC, leaving the equivalent magnetic currents. Thus, the excitation of the device by the guided structure is represented as the excitation by the equivalent magnetic currents given by (7.7b) over the PEC placed at the original opening of the guiding line [Figure 7.5(c)].

2. For a coaxial transmission line and a rectangular waveguide, the dominant modes are the TEM wave and the TE_{10} mode, respectively. If the device, together with the coaxial line, forms a rotationally symmetrical structure, only the TM_{0n} modes should be included into the expansion [5]. In general, we should include both the TE_{mn} and TM_{mn} modes.

The currents induced in the PEC and the equivalent magnetic currents can be determined by requiring that the tangential component of the total electric or magnetic field on the PEC just outside the device be equal to zero. For example, in the Galerkin method, the testing is performed for the tangential component of the electric field by functions used to approximate the induced electric currents and for the tangential component of magnetic field by functions used to represent the equivalent magnetic currents.

The main shortcoming of this approach is that for various types of guiding structures, we need to implement different forms of (7.6) in a general algorithm. Moreover, for some guiding structures, the determination of the dominant and higher-order modes is a relatively complicated stand-alone task. Finally, the approach is not easily applicable to guiding structures whose interaction with the device is not strictly limited to the opening of the guiding structure (e.g., for microstrip lines). All these difficulties can be overcome if guiding structures are modeled in the same way as any other structure (e.g., by metallic plates), and delta-function generators are applied to excite the guiding structures.

As the first example, let us consider a monopole antenna above a finite ground plane fed by a coaxial line [Figure 7.6(a)]. At one end of the coaxial line, its inner conductor is extended into the antenna. The outer conductor is connected to the ground plane. At the other end of the line, flat ends of the inner and outer conductor are connected by a short wire. A delta-function generator is placed at the junction of the wire and the flat end of the outer

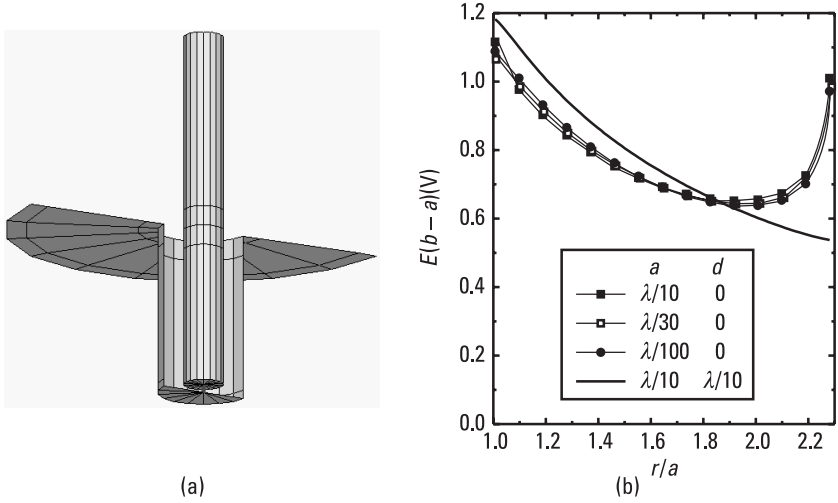

Figure 7.6 Exact modeling of coaxial-line excitation: (a) geometrical model, and (b) normalized magnitude of the electric field versus the radial coordinate at the opening and at distance of $\lambda/10$ from the opening.

conductor. Let the height of the monopole antenna, the radius of the ground plane, and the length of the coaxial line be $\lambda/4$ (where λ is the free-space wavelength), and let the ratio of the inner and outer radii of the coaxial line be $a/b = 2.3$.

Figure 7.6(b) shows the magnitude of the radial component of the electric field versus the radial coordinate, r, at the coaxial-line opening normalized by the voltage at the opening for three opening radii ($a = \lambda/100$, $\lambda/30$, and $\lambda/10$). In addition, normalized results at a distance $d = \lambda/10$ from the opening are given. At this distance, the magnitude of the electric field varies as $1/r$, which is characteristic for the dominant TEM mode. At the opening, the electric field shows singular behavior when the field point approaches the edge, due to higher-order modes added to the dominant direct and reflected TEM modes. The results indicate that the higher-order modes created in the vicinity of the opening vanish towards the generator, as expected.

As the second example, let us consider an open-ended waveguide antenna, shown in Figure 7.7. At the opposite end, the waveguide is closed by a flat PEC plate. The waveguide is excited by a wire probe placed in the vicinity of the closed waveguide end. The width, height, and length of the waveguide (positioned along x-, y- and z-axes) are $a = 0.6\lambda$, $b = 0.3\lambda$, and $l = 1.25\lambda$, respectively, where λ is the free-space wavelength.

Figure 7.7 shows the magnitude of the y-component of the electric field at the waveguide opening and at distance $d = \lambda/2$ from the opening. At $d = \lambda/2$ the magnitude of the electric field varies as $\sin(\pi x/a)$, where x is the Cartesian coordinate along the waveguide width, which is characteristic for the dominant TE_{10} mode. At the opening, the electric field becomes singular when the field point approaches the horizontal edges due to higher-order modes

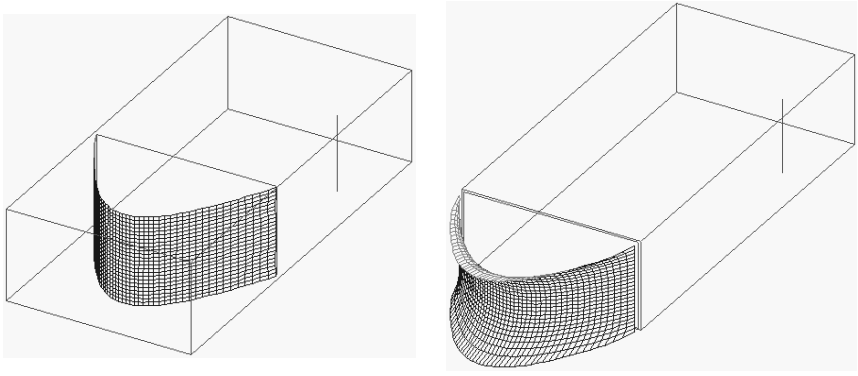

Figure 7.7 Magnitude of the transverse electric field in a rectangular waveguide: (a) at the opening, and (b) at distance $d = \lambda/2$ from the opening.

added to the dominant direct and reflected TE_{10} modes. The results indicate that the higher-order modes created in the vicinity of the opening vanish towards the generator.

Any guiding structure can be modeled as shown above for the coaxial line and the rectangular waveguide. In some cases, the field distribution at the opening does not depend much on the device that is excited. For example, the relative field distribution at the coaxial-line opening practically does not depend on the size of the ground plane and the lengths of the monopole and the line (except for electrically small ground planes and electrically short monopole and line). Moreover, this distribution is close to the TEM mode. Similarly, the relative field distribution at the waveguide opening is close to the TE_{10} mode. In all these cases, the excitation by guided structures can be treated approximately, neglecting higher-order modes. For example, the excitation by a coaxial line is modeled by the corresponding TEM magnetic-current frill (see Section 7.3.2), while the excitation by a rectangular waveguide is modeled by the corresponding TE_{10} magnetic-current sheet.

7.3.2 TEM Magnetic-Current Frill

Consider a vertical monopole antenna of a circular cross section, representing a simple protrusion of the inner coaxial-line conductor. The line drives the antenna through an infinite, perfectly conducting ground plane (Figure 7.8). The distribution of the electric field at the electrically small openings is close

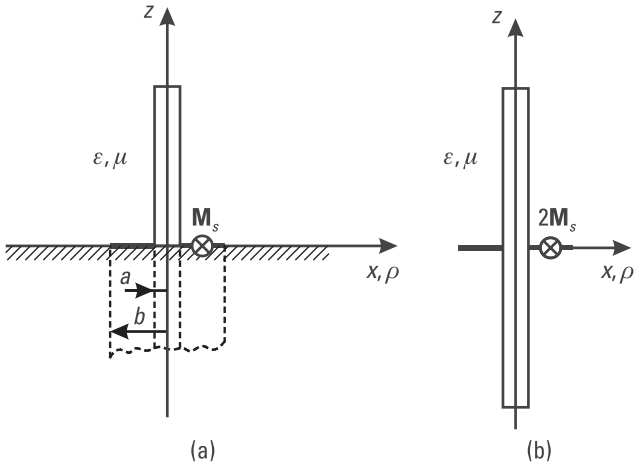

Figure 7.8 Monopole antenna above ground plane driven by coaxial line: (a) field inside the line is replaced by TEM magnetic-current frill at the opening; (b) dipole antenna is obtained after applying the image theory.

to the TEM mode, as presented in Figure 7.6. Hence, the electromagnetic field at the opening can be approximated by the TEM mode only [7, 8].

For the TEM mode, the field components at the opening are of the form

$$E_\rho(\rho) = \frac{V}{\rho \ln(b/a)} \qquad H_\phi(\rho) = \frac{I(0)}{2\pi\rho} \tag{7.8a,b}$$

where V is the voltage, $I(0)$ is the current intensity at the opening, and ρ is the distance of the point on the opening considered from the z-axis.

The perfectly conducting ground plane can be extended to cover the opening, provided that an annular layer of surface magnetic currents is placed at the former opening immediately above the plane as indicated in Figure 7.8(a). According to (7.7b) and (7.8a), these magnetic currents are circular. Their density, with respect to the reference direction shown in the figure, is given by

$$M_{s\phi}(\rho) = -E_\rho(\rho) \tag{7.9}$$

If the coaxial line voltage, V, is assumed to be known, $M_{s\phi}(\rho)$ is defined by (7.8a) and (7.9).

The image theory can be applied to obtain the equivalent system shown in Figure 7.8(b). In that system, the magnetic-current density of the frill is twice the frill of Figure 7.8(a). These magnetic currents are the source of the impressed electric field for the symmetrical dipole antenna. This procedure can be generalized to an arbitrary antenna with multiple coaxial feeds.

The impressed electric field due to the TEM magnetic-current frill shown in Figure 7.8(b) at a point $P(x, 0, z)$ lying in the xOz plane, has the following components [5]:

$$E_x(x, 0, z) = \frac{-4zV}{\ln(b/a)} \int_a^b \int_0^\pi \frac{\cos\phi}{R} \frac{dg(R)}{dR} d\phi\, d\rho \tag{7.10}$$

$$E_z(x, 0, z) = \frac{-4V}{\ln(b/a)} \int_0^\pi g(R)\Big|_{\rho=a}^{\rho=b} d\phi$$

where R is the distance between the source and the field points, ϕ is the spherical angle, $g(R) = e^{-j\beta R}/R$ is the Green's function, and β is the free-space phase coefficient. Similarly, the impressed magnetic field due to the frill at point P has only the y-component, which is computed from the formula

$$H_y(x,\ 0,\ z) = \frac{j4\omega\epsilon_0 V}{\ln(b/a)} \int_a^b \int_0^\pi \cos\phi g(R)\, d\phi\, d\rho \tag{7.11}$$

The integrals (7.10) and (7.11) can be evaluated numerically with ease, unless field point P is very close to the magnetic current frill.

Expressions (7.10) and (7.11) are important when analyzing a general structure excited by a coaxial line. For the vertical monopole shown in Figure 7.8, whose analysis is performed by solving the EFIE based on the EBC, we need only the electric field along the monopole axis. The x-component of the field due to the frill is equal to zero, and the z-component is evaluated as

$$E_z(0,\ 0,\ z) = \frac{-4\pi V}{\ln(b/a)} g(R)\Big|_{\rho=a}^{\rho=b} \qquad R = \sqrt{\rho^2 + z^2} \tag{7.12}$$

As a consequence of the magnetic-current frill excitation, the first derivative of the current along the monopole has a discontinuity at the frill location. The tangential component of the electric field vector, just above the frill (at $z = 0+$), has only the radial component given by (7.8a). The per-unit-length charge at the monopole base is related to this radial component as $Q' = 2\pi\epsilon_0 a E_\rho(\rho = a)$. Using continuity equation (6.2), the first derivative of the current at the frill location is obtained as

$$\left.\frac{dI(z)}{dz}\right|_{z=0+} = -j\frac{2\pi\omega\epsilon_0 V}{\ln(b/a)} = -j\beta Y_c V \tag{7.13}$$

where Y_c is the characteristic admittance of the coaxial line. The expression on the right hand side of (7.13) is imaginary. The active current component (in phase with the voltage) has zero derivative with respect to z at $z = 0+$, and the reactive current component must have the derivative at $z = 0+$ as in (7.13).

Condition (7.13) can be easily incorporated into any smooth approximation of the antenna current distribution (Chapter 6). If the point-matching method is applied to determine the unknown coefficients in the approximation for currents, the incorporation of condition (7.13) significantly improves the stability and accuracy of the results [5]. However, if the Galerkin method is used, accurate results for the current distribution are also obtained without taking this condition into account.

For example, consider a monopole antenna of height $h = 0.25$m and the ratio of the coaxial line radii $b/a = 2.3$ (Figure 7.8). The geometrical model consists of one cylindrical wire and one flat disk as the cap. Figure 7.9 shows

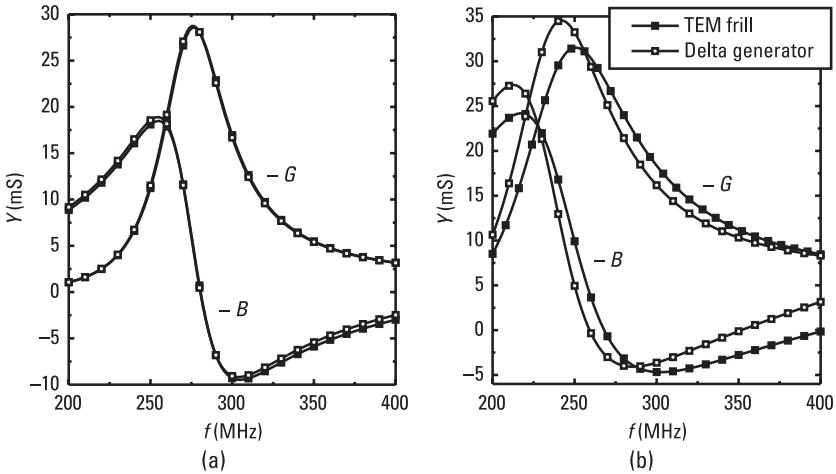

Figure 7.9 Input admittance of monopole antenna ($h = 0.25\lambda$) above infinite ground plane versus frequency for two ratios of the monopole height and radius: (a) $h/a = 100$ and (b) $h/a = 10$.

the antenna input admittance[3] for two ratios of the monopole height and radius: $h/a = 100$ and $h/a = 10$. The results are compared with those of a monopole antenna fed by a delta-function generator with a conical terminal, as described in Section 7.2.1. The currents along the feeding cone and the disc cap are approximated by first-order polynomials, while the current along the cylindrical part is approximated by a third-order polynomial. The condition for the first derivative (7.13) is not included in the approximation of the currents along the monopole. The unknowns are determined by the Galerkin method applied to the EFIE. By increasing the order of the approximation of currents along the cylinder, the results practically do not change, which means that sufficiently accurate results for the input admittance can be obtained with a relatively small number of unknowns without inclusion of (7.13).

For $h/a = 100$ ($a = \lambda/250$), the results obtained by the TEM frill and the delta-function generator almost coincide. We can use a delta-function generator instead of a TEM magnetic current frill whenever $a < \lambda/250$ ($b < 0.01\lambda$). By increasing the electrical size of the frill, the difference between the results increases, as is shown for $h/a = 10$ ($a = \lambda/25$). This difference occurs because the two kinds of generators have differently shaped input terminals and because, due to their electrical size, the influence of the terminal capacitance on the antenna input impedance is not negligible. For $a > \lambda/25$ ($b > 0.1\lambda$),

3. The input admittance is calculated as a ratio of the current at the monopole base, $I(0)$, and the specified voltage, V, of the TEM magnetic-current frill (i.e., as $Y = I(0)/V$).

the difference between the voltage calculated across the frill and the specified voltage V becomes significant. The input admittance cannot be obtained in the same way as for an ideal voltage generator [i.e., by using expression $Y = I(0)/V$], but by postprocessing near-field data. By further increasing the electrical size of the frill, the influence of higher order-modes on the overall solution should also be taken into account [5].

7.4 Transfer of Excitation

Consider a vertical monopole antenna driven by a coaxial line through a finite PEC plate. Following the same procedure as for the infinite plane in the previous section, we replace the fields inside the coaxial line by the TEM magnetic-current frill over the opening, below which the plate is extended [Figure 7.10(a)]. However, since the plate is finite, we cannot remove the plate by applying the image theory. The frill produces a discontinuous distribution of the excitation field over the plate. At points bellow the frill, the intensity of the electric-field-vector component tangential to the plate equals half the intensity of the surface magnetic currents. At other points of the plate, the field is zero. Consequently, the electric-charge distribution and the first derivative of the electric surface current are also discontinuous over the plate. The geometrical modeling of the plate tends to be quite complicated. First, we need to model the wire-to-plate junction [Figure 7.10(b)]. Second, to enable a discontinuous charge distribution at the frill edge, the meshing of the plate should approximately follow the outer frill bound [Figure 7.10(c)].

Such complicated geometrical modeling can be avoided if the size of the frill is substantially smaller than the size of the plate. The local currents and charges induced in the plate due to the frill are similar to those induced in an infinite plane, so that they can be approximately replaced by the mirror image of the frill. The frill and its image produce (on the excited monopole) an impressed field that is twice the field produced by the frill alone. The approach in which the original excitation by the frill is replaced by the impressed field, which is twice the field produced by the original frill alone, is called the transfer of excitation [9]. This case is shown in Figure 7.10(d). Since the frill is removed, the current and charge distributions are not discontinuous over the plate, and the simple geometrical model shown in Figure 7.10(e) suffices.

For example, let the height and radius of the vertical monopole be h = 0.25m and a = 0.0025m, let the ratio of the outer and inner frill radii be b/a = 2.3, and let the monopole be placed in the middle of a square plate of side d = 1m. Figure 7.11 shows the admittance of the antenna versus frequency for three antenna models shown in Figure 7.10(b, c, e), respectively. The results

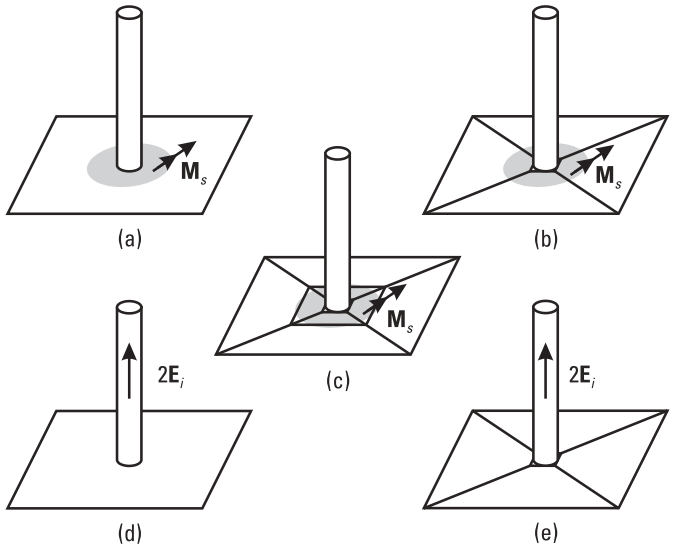

Figure 7.10 Monopole antenna above finite ground plane driven by coaxial line: (a) field inside the line is replaced by TEM magnetic-current frill at the opening covered by a PEC; (b) wire-to-plate junction is modeled; (c) meshing of the ground plane approximately follows the outer frill bound; (d) the influence of the frill is replaced by equivalent impressed field along the wire; and (e) corresponding geometrical model.

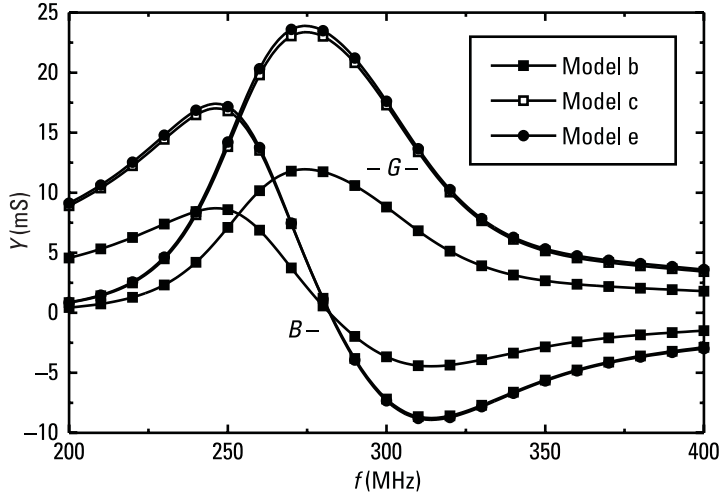

Figure 7.11 Input admittance of monopole antenna above square plate driven by TEM magnetic current frill (h = 0.25m, a = 0.0025m, b/a = 2.3, d = 1m) obtained for antenna models shown in Figure 7.10(b, c, e).

obtained by the first model differ significantly from those obtained by the second and third model, which almost coincide. By increasing the order of approximation, the results obtained by the second and third model are much more stable than those obtained by the first model. It is concluded that the transfer of excitation enables accurate analysis using simple geometrical models.

7.5 Summary

Using impressed sources (possibly combined with the modeling of the complete feeding structure), any practical excitation can be modeled in more or less detail. The goal in modeling the excitation is to represent it in the simplest way, which still yields accurate results for the part of the response that we are interested in. There are three general types of problems: scattering of free-space waves from structures, excitation of structures by voltage and current generators, and excitation of structures by guided waves. The plane-wave excitation and excitation by delta-function generators (with a conical model of terminals and, possibly, a more complete model of the feeding structure) can be used to simulate almost any excitation encountered in practice.

References

[1] Olver, D., et al., *Microwave Horns and Feeds,* London: IEE, and New York: IEEE Press, 1994.

[2] Miller, E. K., and F. J. Deadrick, "Some Computational Aspects of Thin-Wire Modeling," in *Computer Techniques for Electromagnetics,* R. Mittra, (ed.), Oxford, U.K.: Pergamon Press, 1973.

[3] Djordjevic, A. R., et al., *AWAS for Windows, Version 2.0: Analysis of Wire Antennas and Scatterers (Software and User's Manual),* Norwood, MA: Artech House, 2002.

[4] King, R. W. P., *The Theory of Linear Antennas,* Boston, MA: Harvard University Press, 1956.

[5] Popovic, B. D., M. D. Dragovic, and A. R. Djordjevic, *Analysis and Synthesis of Wire Antennas,* Chichester, U.K.: Research Study Press, 1982.

[6] Harrington, R. F., *Time-Harmonic Electromagnetic Field,* New York: McGraw-Hill, 1961.

[7] Tsai, L. L., "A Numerical Solution for Near and Far Fields of an Annular Ring of Magnetic Current," *IEEE Trans. Antennas Propagat.,* Vol. AP-20, September 1972, pp. 569–576.

[8] Thiele, G. A., "Wire Antennas," in *Computer Techniques for Electromagnetics,* R. Mittra, (ed.), Oxford, U.K.: Pergamon Press, 1973.

[9] Kolundzija, B. M., and B. D. Popovic, "Simplified Treatment of Wire-to-Plate Junctions with Magnetic Current Frill Excitation," *Proc. IEE,* Pt. H, Vol. 141, pp. 133–137.

8

Test Procedure

The application of the MoM to vector equations has some specific features when compared with its application to scalar equations, particularly if the testing is performed in nonorthogonal coordinate systems. These features are elaborated in Section 8.1.

For field equations, the inner product is usually adopted in the form of an integral. Hence, all test procedures can be realized as the weighted point-matching method. The problems that occur in the application of the method are the choice of matching points and weighting coefficients, as well as the evaluation of fields due to currents over generalized building elements. We deal with these problems in Section 8.2.

Application of the Galerkin method requires special care. First, the solution depends substantially on the form of the SIE that is tested. Second, the matrix elements can be evaluated much more efficiently if the starting inner product is properly transformed. Finally, in some cases the testing can be simplified by using the concept of the generalized scalar formulation of the SIEs. These topics are considered in Section 8.3.

The optimal choice of testing procedure is discussed in Section 8.4.

8.1 Testing of Vector Equations in Nonorthogonal Coordinate Systems

In Chapter 2 the MoM is elaborated using the example of real scalar linear operator equations. However, field equations in the frequency domain are of the vector complex type. Hence, let us briefly repeat the basic steps of the MoM, following an example of such equations. We start from

$$L\mathbf{F} = \mathbf{G} \tag{8.1}$$

where L is a linear operator, $\mathbf{F}$ is an unknown vector complex function (response), and $\mathbf{G}$ is a known vector complex function (excitation). As the first step, we approximate the unknown function as

$$\mathbf{F} = \sum_{i=1}^{N} a_i \mathbf{F}_i \tag{8.2}$$

where N is the order of approximation, $\mathbf{F}_i$ are known mutually independent basis functions, and a_i are unknown coefficients to be determined. After substituting (8.2) into (8.1), we obtain the approximate equation in the form

$$\sum_{i=1}^{N} a_i L\mathbf{F}_i \cong \mathbf{G} \tag{8.3}$$

Then we take the inner product of the left and right side of the approximate equation and weighting functions, $\mathbf{W}_j, j = 1, \ldots, N$, resulting in a determined system of linear equations:

$$\sum_{i=1}^{N} a_i \langle \mathbf{W}_j, L\mathbf{F}_i \rangle = \langle \mathbf{W}_j, \mathbf{G} \rangle \qquad j = 1, \ldots, N \tag{8.4}$$

Finally, the unknown coefficients are obtained by solving this system.

For brevity, let us denote the first vector in the inner products in (8.4) by **A**, and the second vector by **B**. In field problems, the inner product of two complex vectors is usually formulated as

$$\langle \mathbf{A}, \mathbf{B} \rangle = \int_{\Omega} \mathbf{A}^* \cdot \mathbf{B} d\Omega \tag{8.5}$$

where (*) denotes complex conjugate. Generally, the integration in (8.5) is performed numerically as

$$\langle \mathbf{A}, \mathbf{B} \rangle = \sum_{k=1}^{M} C_k \mathbf{A}^*(\mathbf{r}_k) \cdot \mathbf{B}(\mathbf{r}_k) \tag{8.6}$$

where $\mathbf{r}_k$ and C_k are position vectors of integration points and weighting coefficients of the numerical integration formula. Evaluation of various vectors at

integration points and their dot products is always performed using a coordinate system, which is generally curvilinear and nonorthogonal. In such systems the p- and s-components of a vector are not identical to its p- and s-projections (see Section 4.1.3). Consequently, each vector in the inner product can be represented in two ways, resulting in four different formulas for the dot product.

For example, let us consider 2-D coordinate systems defined over surface patches, which are widely used in solving the MoM/SIE.[1] Basically, evaluation of the tangential components of the vectors at the integration point means evaluation of the p- and s-components of the vector at this point: A_p, A_s, B_p, and B_s. The tangential components of **A** and **B** are written as

$$\mathbf{A}_{\tan} = A_p \mathbf{i}_p + A_s \mathbf{i}_s \qquad \mathbf{B}_{\tan} = B_p \mathbf{i}_p + B_s \mathbf{i}_s \tag{8.7a,b}$$

where $\mathbf{i}_p$ and $\mathbf{i}_s$ are unit vectors. Determination of the p- and s-components is performed in two steps. First, according to (4.16), we determine the p- and s-projections of **A** and **B**: A'_p, A'_s, B'_p, and B'_s. Then, according to (4.18), we express the p- and s-components as linear combinations of the p- and s-projections. Hence, (8.7) can be alternatively expressed in terms of A'_p, A'_s, B'_p, and B'_s as

$$\mathbf{A}_{\tan} = A'_p \mathbf{i}'_p + A'_s \mathbf{i}'_s \qquad \mathbf{B}_{\tan} = B'_p \mathbf{i}'_p + B'_s \mathbf{i}'_s \tag{8.8a,b}$$

$$\mathbf{i}'_p = \frac{\mathbf{i}_p - (\mathbf{i}_p \cdot \mathbf{i}_s)\mathbf{i}_s}{1 - (\mathbf{i}_p \cdot \mathbf{i}_s)^2} \qquad \mathbf{i}'_s = \frac{\mathbf{i}_s - (\mathbf{i}_p \cdot \mathbf{i}_s)\mathbf{i}_p}{1 - (\mathbf{i}_p \cdot \mathbf{i}_s)^2} \tag{8.9a,b}$$

Since $\mathbf{i}_p \cdot \mathbf{i}'_p = 1$, $\mathbf{i}_p \cdot \mathbf{i}'_s = 0$, $\mathbf{i}_s \cdot \mathbf{i}'_p = 0$, and $\mathbf{i}_s \cdot \mathbf{i}'_s = 1$, the vectors $\mathbf{i}'_p$ and $\mathbf{i}'_s$ are called inverse unit vectors.

Now we can establish the four formulas for evaluation of the dot product in a nonorthogonal coordinate system: (1) $\mathbf{A}_{\tan}$ and $\mathbf{B}_{\tan}$ are decomposed along the unit vectors; (2) $\mathbf{A}_{\tan}$ and $\mathbf{B}_{\tan}$ are decomposed along the inverse unit vectors; (3) $\mathbf{A}_{\tan}$ and $\mathbf{B}_{\tan}$ are decomposed along the unit and their inverse vectors, respectively; and (4) $\mathbf{A}_{\tan}$ and $\mathbf{B}_{\tan}$ are decomposed along the inverse unit vectors and the unit vectors, respectively. The formulas are summarized in Table 8.1.

Consisting of three operations only, the last two formulas listed in Table 8.1 have advantage with respect to the first two formulas. For the point-matching and Galerkin methods, the testing functions are usually formulated

1. The procedures can be easily generalized to 3-D coordinate systems defined over volume bricks, which are widely used in solving the MOM/VIE and the FEM.

Table 8.1
Summary of Formulas for Dot Product in Nonorthogonal Coordinate Systems

A in Terms of	B in Terms of	$\mathbf{A} \cdot \mathbf{B}$
Components	Components	$A_p B_p + A_s B_s + (A_p B_s + A_s B_p)\mathbf{i}_p \cdot \mathbf{i}_s$
Projections	Projections	$[1 - (\mathbf{i}_p \cdot \mathbf{i}_s)^2]^{-1}[A'_p B'_p + A'_s B'_s - (A'_p B'_s + A'_s B'_p)\mathbf{i}_p \cdot \mathbf{i}_s$
Components	Projections	$A_p B'_p + A'_s B_s$
Projections	Components	$A'_p B_p + A'_s B_s$

to have only the p- or s-components. Since the evaluation of field projections is easier than the the evaluation of field components, the fourth formula has advantage with respect to the third formula. In other words, the optimal procedure is to decompose vector field equations into projections and testing functions into components.

The last two formulas enable the MoM equations to be organized as the weighted point-matching method given by (2.65). For example, for the fourth formula, the matrices $\mathbf{W}$, $\mathbf{B}$, $\mathbf{F}_{\mathrm{PM}}$, and $\mathbf{G}_{\mathrm{PM}}$ are defined as

$$\mathbf{W} = [\{w_j(\mathbf{r}_k)\}_p \{w_j(\mathbf{r}_k)\}_s]_{N\times M} \qquad \mathbf{B} = \mathrm{diag}(b_1, b_1, \ldots, b_M, b_M) \tag{8.10a,b}$$

$$\mathbf{F}_{\mathrm{PM}} = \begin{bmatrix} \{Lf_i(\mathbf{r}_k)\}'_p \\ \{Lf_i(\mathbf{r}_k)\}'_s \end{bmatrix}_{M\times N} \qquad \mathbf{G}_{\mathrm{PM}} = \begin{bmatrix} \{g(\mathbf{r}_k)\}'_p \\ \{g(\mathbf{r}_k)\}'_s \end{bmatrix}_{M\times 1} \tag{8.10c,d}$$

8.2 Weighted Point-Matching Method

Common to all MoM/SIEs is that they are linear operator equations in terms of surface electric or magnetic currents, $\mathbf{J}_s$ and $\mathbf{M}_s$ (see Chapter 4). The geometry of boundary surfaces is modeled by three types of building elements: generalized quadrilaterals, triangles, and wires. Unknown currents are expanded and equations are imposed in local curvilinear nonorthogonal coordinate systems associated with the building elements. The main problems in the application of the weighted point-matching method are the choice of matching (integration) points and weighting coefficients and the imposition of point-matching equations for these points. In particular, the main problem of imposing these equations is the evaluation of field integrals due to known basis functions.

8.2.1 Choice of Matching (Integration) Points and Weighting Coefficients

For the ordinary point-matching method, there is no rule for the optimal choice of matching points. One possible goal may be to satisfy more or less uniformly an equation along wire axes and over surface patches. This objective can be achieved if the matching points are relatively uniformly distributed.

To begin, consider a straight wire. According to (6.13), the current is approximated by a polynomial expansion of order n, which automatically satisfies the condition that the current be zero at the wire ends. Such expansion consists of $n-1$ singlets multiplied by unknown coefficients. If the unknown coefficients are determined by the point-matching method, we need $n-1$ matching points. To obtain the uniform distribution of matching points, the axis is subdivided into $n-1$ subsegments of equal lengths, and the matching points are chosen in the middle of each segment [Figure 8.1(a)].

The same distribution can be applied for the weighted point-matching method, except that the number of matching (integration) points should be increased. In addition, we need to choose weighting coefficients. The simplest choice for the weighting coefficients associated with these matching points is that they equal the subsegment lengths. Such a choice of matching points and weighting coefficients corresponds to the application of the midpoint rule for numerical integration of (8.5).

The accuracy of integration depends on the number of integration points, as well as on the behavior of the basis functions, $\mathbf{F}_i$, and their fields, $L\mathbf{F}_i$. Usually, $\mathbf{F}_i$ are simple smooth functions, while $L\mathbf{F}_i$ are not well behaved. This causes the number of matching points needed for a required accuracy to be much greater than the number of unknown coefficients. Consequently, the

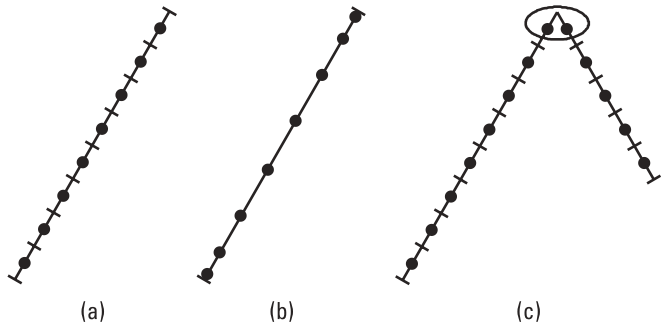

Figure 8.1 Matching points along wires: (a) uniform distribution along single segment, (b) Gauss-Legendre distribution along single segment, and (c) uniform distribution at junction of two wires.

integration of $\mathbf{F}_j \cdot L\mathbf{F}_i$ (the Galerkin method) requires fewer points than the integration of $L\mathbf{F}_j \cdot L\mathbf{F}_i$ (the least-squares method).

In general, the wire axis can be curvilinear, defined by the parametric equation $\mathbf{r} = \mathbf{r}(s)$, $-1 \le s \le 1$, where s is an arbitrary local coordinate associated with the wire generatrix (see Section 5.1.1). Instead of the midpoint rule, we can apply an arbitrary formula for numerical integration. Using $d\Omega = dl$ and $dl = e_s\,ds$ (e_s is the Lame coefficient), the inner product (8.5) is transformed as

$$\langle \mathbf{A}, \mathbf{B} \rangle = \int_C \mathbf{A}^* \cdot \mathbf{B}\,dl = \int_{-1}^{1} \mathbf{A}^* \cdot \mathbf{B} e_s\,ds \approx \sum_{k=1}^{M} c_k e_s(s_k) \mathbf{A}^*[\mathbf{r}(s_k)] \cdot \mathbf{B}[\mathbf{r}(s_k)] \tag{8.11}$$

where s_k are arguments and c_k are weighting coefficients of the integration formula along the interval $(-1, 1)$. According to the authors' experience, the greatest efficiency is achieved if the integration points (and the weighting coefficients) are chosen according to the arguments (and the weighting coefficients) of a Gauss-Legendre integration formula, as shown in Figure 8.1(b). The minimal number of integration points needed by the Galerkin method often approaches the number of unknowns. Obviously, the same distribution of matching points should be applied in the ordinary point-matching method if we want that the point-matching solution be close to the Galerkin solution.

The same distributions, uniform and Gauss-Legendre, can be successfully applied to arbitrary wire structures if weighted point-matching is applied. However, such choices of matching points are not always adequate for the ordinary point-matching method. For example, let us consider a polynomial expansion of currents along two interconnected wires consisting of $n_1 - 1$ singlets along the first wire, $n_2 - 1$ singlets along the second wire, and a doublet along both wires. It is not clear where to position the matching point that corresponds to the doublet. If we add it to one wire, we force better satisfaction of the equation along this wire, which is undesirable. If we place it at the interconnection of the two wires, where the electric field varies very quickly, relatively small changes in its position may cause relatively big changes in the final result [1–3]. To overcome these difficulties, we can add one matching point to each wire in the vicinity of the junction [Figure 8.1(c)] and sum the point-matching equations for these points [3]. We thereby obtain a relatively stable solution. However, this is not a pure point-matching method any more.

Based on the above reasoning, we can choose matching points and weighting coefficients for generalized quadrilaterals defined by the parametric equation $\mathbf{r} = \mathbf{r}(p, s)$, $-1 \le p \le 1$, $-1 \le s \le 1$, where p and s are arbitrary local

coordinates associated with the surface (see Section 5.2.1). Using $d\Omega = dS$ and $dS = |\mathbf{a}_p \times \mathbf{a}_s|\, dp\, ds$, the inner product (8.5) is transformed as

$$\langle \mathbf{A}, \mathbf{B} \rangle = \int_S \mathbf{A}^* \cdot \mathbf{B} dS = \int_{-1}^{1} \int_{-1}^{1} \mathbf{A}^* \cdot \mathbf{B} |\mathbf{a}_p \times \mathbf{a}_s|\, dp\, ds \tag{8.12}$$

$$\approx \sum_{k=1}^{M_k} \sum_{l=1}^{M_l} c_k c_l |\mathbf{a}_p(p_k, s_k) \times \mathbf{a}_s(p_k, s_k)| \mathbf{A}^*[\mathbf{r}(p_k, s_k)] \cdot \mathbf{B}[\mathbf{r}(p_k, s_k)]$$

where p_k and c_k, (s_l and c_l) are the arguments and weighting coefficients of the integration formula of the order M_p (M_s), applied to the interval (−1, 1) along the p-coordinate (s-coordinate). The integration points should be placed over the quadrilateral at (p_k, s_k), $k = 1, \ldots, M_p$, $l = 1, \ldots, M_s$. For example, Figure 8.2 shows two distributions of matching points over a bilinear surface: uniform and Gauss-Legendre. For each matching point, two equations are imposed, representing the projections of the approximate equation onto local unit vectors. Similarly, we can select matching points and weighting coefficients for generalized triangles.

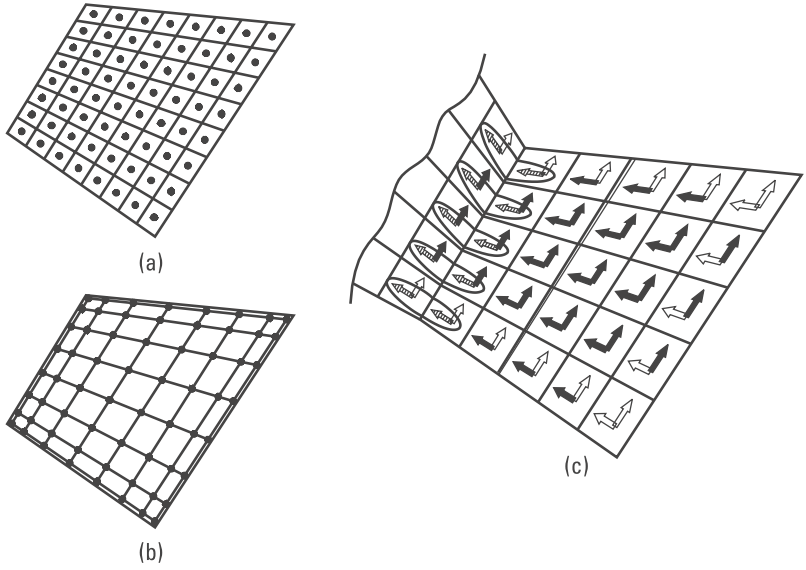

Figure 8.2 Matching points over quadrilaterals: (a) uniform distribution, (b) Gauss-Legendre distribution, and (c) distribution of equation components for the ordinary point-matching method.

This procedure should be slightly modified for the ordinary point-matching method [4, 5]. For example, consider a polynomial expansion of currents over one bilinear surface, as explained in Section 6.2.4. The s-component has the form (6.45), and the p-component has the same form, except that the indices p and s are interchanged. If the orders of approximation along the p- and s-coordinate are n_p and n_s, respectively, the number of basis functions is $(n_p + 1)(n_s + 1)$. We start from $n_p + 1$ by $n_s + 1$ uniformly distributed matching points and two equations imposed for each of them, representing projections of the approximate equation onto the local unit vectors. The condition that the component of the current normal to free edges equal zero is satisfied by omitting the corresponding edge basis functions from the expansion. Consequently, we omit the equations that correspond to the nontangential unit vectors at the matching points close to the free edges, as shown by the hollow arrows in Figure 8.2. The condition that the component of the current normal to an interconnected edge be continuous across the edge is satisfied by grouping into doublets the pairs of the corresponding edge basis functions on the plates that share this edge. Consequently, we sum the equations that correspond to the nontangential unit vectors at the neighboring matching points on the two plates, as shown by the hatched arrows in Figure 8.2. The grouping is not straightforward if the number of matching points along the edge on one plate is not equal to the number of matching points along the edge on the other plate. This difficulty represent a serious shortcoming of the point-matching method when compared with other test procedures.

8.2.2 Field Integrals of Currents over Generalized Quadrilaterals

The main problem in imposing the point-matching equations based on the MoM/SIEs is the evaluation of the electric and magnetic fields at matching points due to the approximation of the electric and magnetic surface currents. According to (3.60), the fields can be expressed as simple linear combinations of L and K operators acting on the currents. In particular, it is convenient that initial basis functions used to approximate both types of currents have the same form, $\mathbf{F}_{ij}(p, s)$, given by (6.41b), defined over a generalized quadrilateral. Thus, we reduce the problem to evaluation of $L(\mathbf{F}_{ij})$ and $K(\mathbf{F}_{ij})$.

In the evaluation of $L(\mathbf{F}_{ij})$ and $K(\mathbf{F}_{ij})$, we start from (3.44), in which $\mathbf{J}_s$ is replaced by $\mathbf{F}_{ij}$ according to (6.41b), $\mathbf{r}'$ by $\mathbf{r}(p, s)$ according to (5.9), and dS by $|\mathbf{a}_p \times \mathbf{a}_s| \, dp\, ds$ according to (5.16). $\nabla_s \cdot \mathbf{F}_{ij}$ is determined by applying (6.31), $\nabla g(R)$ is expressed as $dg(R)/dR\,\mathbf{i}_R$ using (3.31), and $\mathbf{i}_R$ is expressed as $[\mathbf{r} - \mathbf{r}(p, s)]/R$ according to (3.31c) and (3.27b,c). We obtain

$$L(\mathbf{F}_{ij}) = \gamma \int_{-1}^{1}\int_{-1}^{1} f_i(p) g_j(s) \mathbf{a}_s(p, s) g(R)\, dp\, ds \tag{8.13a}$$
$$- \frac{1}{\gamma} \int_{-1}^{1}\int_{-1}^{1} f_i(p) \frac{\partial g_j(s)}{\partial s} \frac{\mathbf{r} - \mathbf{r}(p, s)}{R} \frac{dg(R)}{dR}\, dp\, ds$$

$$K(\mathbf{F}_{ij}) = \int_{-1}^{1}\int_{-1}^{1} f_i(p) g_j(s) \mathbf{a}_s(p, s) \times \frac{\mathbf{r} - \mathbf{r}(p, s)}{R} \frac{dg(R)}{dR}\, dp\, ds \tag{8.13b}$$

If the matching points belong to the quadrilateral surface over which $\mathbf{F}_{ij}$ is defined, the second integral in the expression for $L(\mathbf{F}_{ij})$ and the integral in $K(\mathbf{F}_{ij})$ should be calculated in the principal-value sense.

The evaluation of these integrals is simplified if the parametric equation $\mathbf{r}(p, s)$ for the generalized quadrilateral is represented as a polynomial in terms of the parametric coordinates, p and s, and the functions f_i and g_j are formulated as simple power functions [i.e., $f_i(p) = p^i$ and $g_j(s) = s^j$]. In that case $L(\mathbf{F}_{ij})$ and $K(\mathbf{F}_{ij})$ are expressed in terms of integrals

$$P_{ij} = \int_{-1}^{1}\int_{-1}^{1} p^i s^j g(R)\, dp\, ds \qquad Q_{ij} = \int_{-1}^{1}\int_{-1}^{1} p^i s^j \frac{1}{R} \frac{dg(R)}{dR}\, dp\, ds \tag{8.14a,b}$$

Since the evaluation of potentials is based only on the integrals P_{ij}, they are called potential integrals. The integrals Q_{ij} appear only in the evaluation of fields, and they are called field integrals. In particular, if the generalized quadrilateral is formulated as a bilinear surface given by (5.23), $L(\mathbf{F}_{ij})$ and $K(\mathbf{F}_{ij})$ are obtained as

$$L(\mathbf{F}_{ij}) = \gamma(\mathbf{r}_s P_{ij} + \mathbf{r}_{ps} P_{i+1j}) \tag{8.15a}$$
$$+ \frac{j}{\gamma}(\mathbf{r}_{c0} Q_{ij-1} + \mathbf{r}_p Q_{i+1j-1} + \mathbf{r}_s Q_{ij} + \mathbf{r}_{ps} Q_{i+1j})$$

$$K(\mathbf{F}_{ij}) = (\mathbf{r}_{c0} \times \mathbf{r}_s) Q_{ij} + (\mathbf{r}_p \times \mathbf{r}_s + \mathbf{r}_{c0} \times \mathbf{r}_{ps}) Q_{i+1j} + (\mathbf{r}_p \times \mathbf{r}_{ps}) Q_{i+2j} \tag{8.15b}$$

where $\mathbf{r}_{c0} = \mathbf{r}_c - \mathbf{r}$. The above expressions are derived for the basis functions directed along the s-parametric coordinate. Similar expressions are obtained for

basis functions directed along the p-coordinate, except that the vectors $\mathbf{r}_p$ and $\mathbf{r}_s$, as well as the subscripts i and j, are interchanged. Thus, the main problem in imposing point-matching equations is the evaluation of the potential and field integrals (8.14a,b).

Potential and field integrals (8.14a,b) cannot be solved analytically, but only numerically. If the matching point is far away from a quadrilateral, a simple numerical integration suffices. However, if the point matching is placed inside the quadrilateral or close to it, the free-space Green's function shows rapid variations when the source point approaches the matching point. In that case, a simple numerical integration is usually inefficient. The efficiency can be improved if the fast-varying part is extracted and integrated analytically. For example, let the matching point of the position vector $\mathbf{r}$ be placed in the generalized quadrilateral of parametric equation $\mathbf{r}(p, s)$. The potential integral can be written in the form

$$P_{ij} = \int_{-1}^{1}\int_{-1}^{1}\left[p^i s^j g(R) - p_0^i s_0^j \frac{1}{4\pi R_0}\right] dp\, ds + \frac{p_0^i s_0^j}{4\pi}\int_{-1}^{1}\int_{-1}^{1}\frac{dp\, ds}{R_0} \tag{8.16a}$$

$$R = |\mathbf{r} - \mathbf{r}(p, s)| \qquad R_0 = |\mathbf{r} - \mathbf{r}_0(p, s)| \tag{8.16b,c}$$

$$\mathbf{r}_0(p, s) = \mathbf{r}(p_0, s_0) + \mathbf{a}_p(p_0, s_0)(p - p_0) + \mathbf{a}_s(p_0, s_0)(s - s_0) \tag{8.16d}$$

where p_0 and s_0 are local p- and s-coordinates of the matching point, and $\mathbf{a}_p$ and $\mathbf{a}_s$ are unitary vectors of the generalized quadrilateral, evaluated according to (5.13). Note that $\mathbf{r}_0(p, s)$ defines a parallelogram tangential to the generalized quadrilateral at the matching point (Figure 8.3). The integrand of the first integral on the right side of (8.16a) tends to zero when the source point approaches the field (matching) point. Hence, the numerical integration of this integral is much more efficient than the numerical evaluation of the complete potential integral. The second integral on the right side of (8.16a) can be evaluated analytically [6] once for all integrals P_{ij}. The same expressions are valid when the matching point is close to the generalized quadrilateral, except that p_0 and s_0 represent the local p- and s-coordinates of the source point nearest to the matching point. In a similar way, we can improve efficiency of numerical integration for the field integrals.

8.2.3 Reduced Kernel in Field Integrals of Currents Along Generalized Wires

Since the surface of a generalized wire can be represented as a generalized quadrilateral, (8.13) can be applied to generalized wires, too. By applying the thin-wire approximation, (8.13) can be significantly simplified.

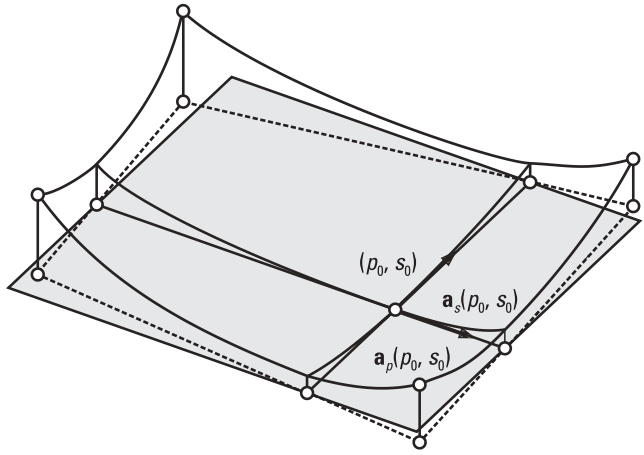

Figure 8.3 Evaluation of potential integrals for generalized quadrilaterals.

After substituting $f_i(p) = 1$, $|\mathbf{a}_p \times \mathbf{a}_s| = |\mathbf{a}_p|\,|\mathbf{a}_s|$, and $\mathbf{i}_s = \mathbf{a}_s / |\mathbf{a}_s|$, the basis function (6.41b) reduces to

$$\mathbf{F}_{ij}(p, s) = \frac{g_j(s)}{|\mathbf{a}_p|}\mathbf{i}_s = \mathbf{F}_j(s) \tag{8.17}$$

The basis functions for generalized wires are usually not expressed as surface-density vectors (8.17), but in the form of total currents along the wires. Let $g_{ja}(s)$ be the total current along the wire due to the basis function $\mathbf{F}_j(s)$. After multiplying the magnitude of $\mathbf{F}_j(s)$ by the circumference of the wire, $2|\mathbf{a}_p|$, we obtain $g_{ja}(s) = 2g_j(s)$. According to (5.10) and (5.13b), the s-unitary vector becomes

$$\mathbf{a}_s(p, s) = \mathbf{a}_{sa}(s) + \frac{\partial a(s)}{\partial s}\mathbf{i}_\rho(p) \qquad \mathbf{a}_{sa}(s) = \frac{\partial \mathbf{r}_a(s)}{\partial s} \tag{8.18}$$

where $\mathbf{a}_{sa}$ is the unitary vector of the wire axis. Expressions (8.13a, b) can be written as

$$L(\mathbf{F}_j) = \gamma \int_{-1}^{1} g_j(s) \left\{ \int_{-1}^{1} [\mathbf{a}_{sa}(s) + \frac{\partial a(s)}{\partial s}\mathbf{i}_\rho(p)] g(R)\, dp \right\} ds \tag{8.19a}$$

$$- \frac{1}{\gamma} \int_{-1}^{1} \frac{\partial g_j(s)}{\partial s} \left\{ \int_{-1}^{1} [\mathbf{r} - \mathbf{r}_a(s) - a(s)\mathbf{i}_\rho(p)] \frac{1}{R} \frac{dg(R)}{dR}\, dp \right\} ds$$

$$K(\mathbf{F}_j) = \int_{-1}^{1} g_j(s) \left\{ \int_{-1}^{1} \left[\mathbf{a}_{sa}(s) + \frac{\partial a(s)}{\partial s} \mathbf{i}_\rho(p) \right] \times [\mathbf{r} - \mathbf{r}_a(s) - a(s)\mathbf{i}_\rho(p)] \frac{1}{R} \frac{dg(R)}{dR} dp \right\} ds \tag{8.19b}$$

The point $p = 0$ can be adopted in an arbitrary way for each value of the coordinate s (i.e., for each cross section of the generalized wire). It is convenient to choose the point $p = 0$, so that this point, the axis point that belongs to the same cross section, and the field point define a plane perpendicular to the cross section. In that case, the functions in brackets in (8.19) are symmetrical with respect to this plane.

If the field point is not close to the p-integration contour, the integration can be performed accurately by the mid-point rule applied to two points, $p = -0.5$ and $p = +0.5$. R is replaced by

$$R_a = R(p = \pm 0.5) = \sqrt{[\mathbf{r} - \mathbf{r}_a(s)]^2 + a(s)^2} \tag{8.20}$$

Using $\mathbf{i}_\rho(-0.5) + \mathbf{i}_\rho(0.5) = 0$, the sums of brackets in (8.19a,b) for $p = -0.5$ and $p = +0.5$ are reduced as

$$\sum_{p=\pm 0.5} \left[\mathbf{a}_{sa}(s) + \frac{\partial a(s)}{\partial s} \mathbf{i}_\rho(p) \right] = 2\mathbf{a}_{sa}(s) \tag{8.21a}$$

$$\sum_{p=\pm 0.5} [\mathbf{r} - \mathbf{r}_a(s) - a(s)\mathbf{i}_\rho(p)] = 2[\mathbf{r} - \mathbf{r}_a(s)] \tag{8.21b}$$

$$\sum_{p=\pm 0.5} \left[\mathbf{a}_{sa}(s) + \frac{\partial a(s)}{\partial s} \mathbf{i}_\rho(p) \right] \times [\mathbf{r} - \mathbf{r}_a(s) - a(s)\mathbf{i}_\rho(p)] = 2\mathbf{a}_{sa}(s) \times [\mathbf{r} - \mathbf{r}_a(s)] \tag{8.21c}$$

Expressions (8.19a,b) can be written as

$$L(\mathbf{F}_j) = \gamma \int_{-1}^{1} g_{ja}(s) \mathbf{a}_{sa}(s) g(R_a)\, ds - \frac{1}{\gamma} \int_{-1}^{1} \frac{\partial g_{ja}(s)}{\partial s} \frac{\mathbf{r} - \mathbf{r}_a(s)}{R_a} \frac{dg(R_a)}{dR_a} ds \tag{8.22a}$$

$$K(\mathbf{F}_j) = \int_{-1}^{1} g_{ja}(s)\mathbf{a}_{sa}(s) \times \frac{\mathbf{r} - \mathbf{r}_a(s)}{R_a} \frac{dg(R_a)}{dR_a} ds \tag{8.22b}$$

If the parametric equation of the wire axis is represented as a polynomial of the parametric coordinate s, and the functions g_{ja} are formulated as simple power functions [i.e., $g_{ja}(s) = s^j$], $L(\mathbf{F}_j)$ and $K(\mathbf{F}_j)$ are expressed as linear combinations of the potential and field integrals

$$P_j = \int_{-1}^{1} s^j g(R_a)\, ds \qquad Q_j = \int_{-1}^{1} s^j \frac{1}{R_a} \frac{dg(R_a)}{dR_a} ds \tag{8.23a,b}$$

In particular, for a right truncated cone given by (5.5b), $L(\mathbf{F}_j)$ and $K(\mathbf{F}_j)$ are expressed as

$$L(\mathbf{F}_j) = \gamma \mathbf{r}_s P_j + \frac{j}{\gamma}(\mathbf{r}_{c0} Q_{j-1} + \mathbf{r}_s Q_j) \qquad K(\mathbf{F}_j) = (\mathbf{r}_{c0} \times \mathbf{r}_s) Q_j \tag{8.24a,b}$$

where $\mathbf{r}_{c0} = \mathbf{r}_c - \mathbf{r}$. Expressions (8.24a,b) represent a special case of (8.15) obtained for $\mathbf{r}_p, \mathbf{r}_{ps} = 0$. Thus, the main problem in imposing point-matching equations is the evaluation of the potential and field integrals (8.23a,b).

The potential and field integrals (8.23a,b) can be solved only numerically. If the field point is far away from the axis, a simple numerical integration suffices. Otherwise, the numerical integration should be combined with the analytical one. For example, the potential integral can be written as

$$P_j = \int_{-1}^{1} \left[s^j g(R_a) - s_0^j \frac{1}{4\pi R_{a0}} \right] ds + \frac{s_0^j}{4\pi} \int_{-1}^{1} \frac{ds}{R_{a0}} \tag{8.25a}$$

$$R_{a0} = \sqrt{[\mathbf{r} - \mathbf{r}_{a0}(s)]^2 + a_0(s)^2} \qquad \mathbf{r}_{a0}(s) = \mathbf{r}_a(s_0) + \mathbf{a}_s(s_0)(s - s_0) \tag{8.25b,c}$$

where s_0 is the s-coordinate of the point on the axis closest to the matching (field) point, and $\mathbf{a}_s$ is the unitary vector of the wire axis at $s = s_0$ evaluated according to (5.13). Note that $\mathbf{r}_{a0}(s)$ represents a straight segment tangential to the axis of the generalized wire at the matching point (Figure 8.4). The

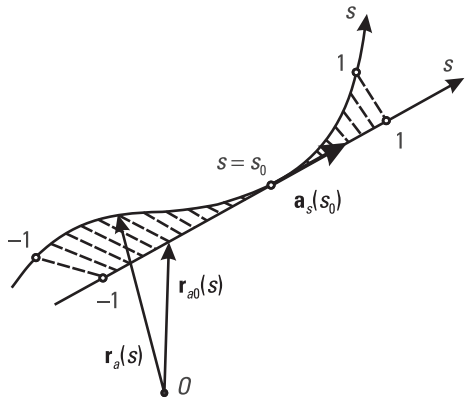

Figure 8.4 Evaluation of potential integrals for generalized wires.

integrand of the first integral on the right side of (8.25a) tends to zero when the source point approaches the field (matching) point. Hence, the numerical integration of this integral is much more efficient than the numerical evaluation of the complete potential integral. The second integral on the right side of (8.25a) is a closed-form integral, which can be evaluated analytically once for all integrals P_j. The efficiency can be further improved by extraction of quasisingular terms of higher order [7]. In a similar way we can improve the efficiency of numerical integration of the field integrals.

8.3 Galerkin Method

8.3.1 Choice of the SIE Form

As elaborated in Chapter 4, each SIE can be basically written in two forms. Each form can be obtained from the other by taking the cross-product with the unit vector normal to the boundary surface ($\mathbf{n}$). For the point-matching method or the least-squares method, it is not important which of these forms is tested. Both forms lead to the same results. However, for the Galerkin method, results depend on which of these two forms is tested.

Ideal testing yields a system of linear equations in which only diagonal elements are different from zero. (The solution of the system requires a small number of operations, one division per unknown.) However, such testing can be achieved in very rare cases. Usually, the system matrix is full, and system solution can fail if the system is not stable enough. Regarding the stability, it is desirable that the leading terms of equations be positioned on the main diagonal. In the opposite case, which is undesirable, the diagonal elements are

equal to zero or are small when compared with other matrix elements. Hence, for the Galerkin method applied to SIEs, we should adopt the form of SIE that results in relatively large (dominant) diagonal elements.

Consider an electric-current pulse basis function, $\mathbf{J}_s$, defined over a square patch [Figure 8.5(a)]. $\mathbf{J}_s$ is a constant vector over the patch and zero elsewhere. The direction of the electric field produced by the pulse, $\mathbf{E}_s(\mathbf{J}_s)$, coincides with the direction of this basis function. The corresponding magnetic field, $\mathbf{H}(\mathbf{J}_s)$, has only the component perpendicular to this direction. In that case $\mathbf{J}_s \cdot (\mathbf{n} \times \mathbf{J}_s)$, $\mathbf{J}_s \cdot [\mathbf{n} \times \mathbf{E}(\mathbf{J}_s)]$, and $\mathbf{J}_s \cdot \mathbf{H}(\mathbf{J}_s)$ are equal to zero. Consequently, if the SIE is written in the form which contains only the terms $\mathbf{n} \times \mathbf{J}_s$, $\mathbf{n} \times \mathbf{E}(\mathbf{J}_s)$, and $\mathbf{H}(\mathbf{J}_s)$, the testing performed by $\mathbf{J}_s$ results in zero diagonal elements.

Consider a magnetic-current pulse basis function, $\mathbf{M}_s$, defined over a square patch [Figure 8.5(b)]. The electric and magnetic fields due to this pulse are dual to the field produced by an electric-current pulse. Hence, $\mathbf{M}_s \cdot (\mathbf{n} \times \mathbf{M}_s)$, $\mathbf{M}_s \cdot [\mathbf{n} \times \mathbf{H}(\mathbf{M}_s)]$, and $\mathbf{M}_s \cdot \mathbf{E}(\mathbf{M}_s)$ are equal to zero. Consequently, if the SIE is written in the form which contains the terms $\mathbf{n} \times \mathbf{M}_s$, $\mathbf{n} \times \mathbf{H}(\mathbf{M}_s)$, and $\mathbf{E}(\mathbf{M}_s)$, the testing performed by $\mathbf{M}_s$ results in zero diagonal elements.

For basis functions defined over arbitrary quadrilateral patches, the diagonal elements are not zero, but relatively small when compared with some other matrix elements. If testing is performed by electric-current basis functions, the SIE should be written in a form which contains the terms $\mathbf{J}_s$, $\mathbf{E}_{\tan}(\mathbf{J}_s)$, and $\mathbf{H}_{\tan}(\mathbf{J}_s)$. If testing is performed by magnetic-current basis functions, the SIE should be written in a form that contains the terms $\mathbf{M}_s$, $\mathbf{H}_{\tan}(\mathbf{M}_s)$, and $\mathbf{n} \times \mathbf{E}_{\tan}(\mathbf{M}_s)$.

For each SIE we can determine the preferred form with respect to the testing by electric-current and magnetic-current basis functions. Summary of these forms for the most important SIEs is given in Table 8.2. All SIEs are written for boundary surfaces of dielectric materials. The first three SIEs can be also regarded as SIEs for PEC surfaces, when testing is performed only by electric-current basis functions.

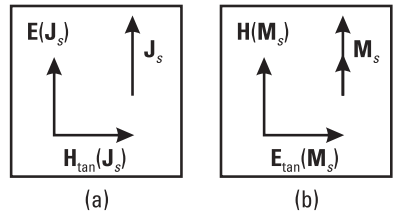

Figure 8.5 The proper choice of the SIE form regarding the testing by (a) electric currents and (b) magnetic currents.

Table 8.2
Summary of Preferred Forms of SIE for Testing by Electric- and Magnetic-Current Basis Functions

	$\mathbf{J}_s$	$\mathbf{M}_s$
EFIE	$\mathbf{E}_{\tan}^{(0-i)} = 0$	$\mathbf{n}_{ij} \times \mathbf{E}^{(0-i)} = 0$
MFIE	$\mathbf{n}_{ij} \times \mathbf{H}^{(0-i)} = 0$	$\mathbf{H}_{\tan}^{(0-i)} = 0$
CFIE	$\alpha \mathbf{E}_{\tan}^{(0-i)} + \mathbf{n}_{ij} \times \mathbf{H}^{(0-i)} = 0$	$\alpha \mathbf{n}_{ij} \times \mathbf{E}^{(0-i)} - \mathbf{H}_{\tan}^{(0-i)} = 0$
CRIE	$\mathbf{E}_{\tan}^{(0-i)} = \alpha \mathbf{E}_{\tan}^{(0-j)}$	$\mathbf{H}_{\tan}^{(0-i)} = \beta \mathbf{H}_{\tan}^{(0-j)}$
	$(\mathbf{n}_{ij} \times \mathbf{H}^{(0-i)} = \beta \mathbf{n}_{ij} \times \mathbf{H}^{(0-j)})$	$(\mathbf{n}_{ij} \times \mathbf{E}^{(0-i)} = \alpha \mathbf{n}_{ij} \times \mathbf{E}^{(0-j)})$
PMCHW	$\mathbf{E}_{\tan}^{(i)} = \mathbf{E}_{\tan}^{(j)}$	$\mathbf{H}_{\tan}^{(i)} = \mathbf{H}_{\tan}^{(j)}$
	$(\mathbf{n}_{ij} \times \mathbf{H}^{(i)} = \mathbf{n}_{ij} \times \mathbf{H}^{(j)})$	$(\mathbf{n}_{ij} \times \mathbf{E}^{(i)} = \mathbf{n}_{ij} \times \mathbf{E}^{(j)})$

For CRIEs, it is common to test the E-equation by $\mathbf{J}_s$ basis functions and the H-equation by $\mathbf{M}_s$ basis functions. According to the authors' experience, such testing provides stable and accurate results for some CRIEs, like the PMCHW. However, for some CRIEs, like the Muller method, such testing can lead to an unstable solution and even to an undetermined system of linear equations.

For example, consider a small cubical scatterer made of a dielectric whose parameters tend to those of the surrounding medium. In the limiting case, this cube is completely transparent (i.e., the scattered field vanishes). The E- and H-Muller equations reduce to (4.80). Let the equivalent currents be approximated by rooftop basis functions (12 electric-current and 12 magnetic-current basis functions). If the E-equation is tested by the electric-current basis functions and the H-equation is tested by the magnetic-current basis functions, we obtain an undetermined system of linear equations (i.e., 6 equations are dependent on the other 18 equations). If the parameters of the cube and the surrounding medium differ slightly, the system of linear equations becomes determined, but unstable. By increasing the difference in parameters of the cube and the surrounding medium, the results stabilize.

According to Table 8.2, it is possible to perform testing in an alternative way: the E-equation by $\mathbf{M}_s$ basis functions and the H-equation by $\mathbf{J}_s$ basis functions. The Muller equation gives stable results for an arbitrarily small difference between the electric parameters of the cube and the surrounding medium.

8.3.2 Impedance Integrals Due to Currents over Generalized Quadrilaterals

Once we adopt the form of the SIE and the basis functions, the main problem in imposition of the Galerkin method is the evaluation of matrix elements (impedances). According to (8.4), (8.5), and Table 8.2, impedances most often have the following general forms:

$$Z_{kl}^{L} = \langle \mathbf{F}_l,\, L\mathbf{F}_k \rangle = \int_{S_l} \mathbf{F}_l^{*} \cdot L\mathbf{F}_k \, dS \tag{8.26a}$$

$$Z_{kl}^{\mathbf{n}\times L} = \langle \mathbf{F}_l,\, \mathbf{n} \times L\mathbf{F}_k \rangle = \int_{S_l} \mathbf{F}_l^{*} \cdot (\mathbf{n} \times L\mathbf{F}_k)\, dS \tag{8.26b}$$

$$Z_{kl}^{K} = \langle \mathbf{F}_l,\, K\mathbf{F}_k \rangle = \int_{S_l} \mathbf{F}_l^{*} \cdot K\mathbf{F}_k \, dS \tag{8.27a}$$

$$Z_{kl}^{\mathbf{n}\times K} = \langle \mathbf{F}_l,\, \mathbf{n} \times K\mathbf{F}_k \rangle = \int_{S_l} \mathbf{F}_l^{*} \cdot (\mathbf{n} \times K\mathbf{F}_k)\, dS \tag{8.27b}$$

where L and K are linear operators given by (3.44), $\mathbf{F}_k$ is the kth basis function defined over the surface S_k, and $\mathbf{F}_l$ is the lth basis function defined over the surface S_l. We can evaluate impedances by using the weighted point-matching method, as explained in Section 8.2.1. However, the efficiency of evaluation can be increased if the above expressions are properly transformed. In this section we shall consider transformations of impedances Z_{kl}^{L} and Z_{kl}^{K} [8].

After applying the operators L and K, the impedances (8.26a) and (8.27a) are obtained in the form

$$Z_{kl}^{L} = \gamma \int_{S_l} \int_{S_k} \mathbf{F}_k(\mathbf{r}_k) \cdot \mathbf{F}_l(\mathbf{r}_l)\, g(R)\, dS_k\, dS_l \tag{8.28a}$$

$$- \frac{1}{\gamma} \int_{S_l} \int_{S_k} \nabla_k \cdot \mathbf{F}_k(\mathbf{r}_k)\, \nabla_l\, g(R) \cdot \mathbf{F}_l(\mathbf{r}_l)\, dS_k\, dS_l$$

$$Z_{kl}^{K} = \int_{S_l}\int_{S_k} [\mathbf{F}_k(\mathbf{r}_k) \times \nabla_l g(R)] \cdot \mathbf{F}_l(\mathbf{r}_l)\, dS_k\, dS_l \tag{8.28b}$$

Using the identity $\nabla \cdot (f\mathbf{A}) = \mathbf{A} \cdot \nabla f + f \nabla \cdot \mathbf{A}$ valid for an arbitrary scalar function f and arbitrary vector function $\mathbf{A}$, the second integral in the expression for Z_{kl}^{L} is transformed as

$$\int_{S_l} \mathbf{F}_l(\mathbf{r}_l) \cdot \nabla_l g(R)\, dS_l = \int_{S_l} \nabla_l \cdot [\mathbf{F}_l(\mathbf{r}_l) g(R)]\, dS_l - \int_{S_l} \nabla_l \cdot \mathbf{F}_l(\mathbf{r}_l) g(R)\, dS_l \tag{8.29}$$

Using the divergence theorem, the first integral on the right side of (8.29) is transformed into $\int_{C_l} [\mathbf{F}_l(\mathbf{r}_l) g(R)] \cdot d\mathbf{c}_l$, where $d\mathbf{c}_l = dc_i \mathbf{i}_l$, $\mathbf{i}_l$ is the unit vector placed in the surface S_l and perpendicular to the contour C_l. The scalar product $\mathbf{F}_l(\mathbf{r}_l) \cdot \mathbf{i}_l$ along the contour C_l represents the component of the lth basis function normal to the contour. For the basis functions used in the analysis (patch basis functions, doublets, and multiplets), such a component is always equal to zero. Hence, the first surface integral on the right side of (8.29) is equal to zero. The expression for Z_{kl}^{K} is transformed using

$$[\mathbf{F}_k(\mathbf{r}_k) \times \nabla_l g(R)] \cdot \mathbf{F}_l(\mathbf{r}_l) = [\mathbf{F}_l(\mathbf{r}_l) \times \mathbf{F}_k(\mathbf{r}_k)] \cdot \nabla_l g(R) \tag{8.30a}$$

$$\nabla_l g(R) = \frac{\mathbf{r}_l - \mathbf{r}_k}{R} \frac{dg(R)}{dR} \tag{8.30b}$$

Finally, after some transformations, the expressions (8.28a,b) are written in the form

$$Z_{kl}^{L} = \gamma \int_{S_l}\int_{S_k} [\mathbf{F}_k(\mathbf{r}_k) \cdot \mathbf{F}_l(\mathbf{r}_l) + \frac{1}{\gamma^2} \nabla_k \cdot \mathbf{F}_k(\mathbf{r}_k) \nabla_l \cdot \mathbf{F}_l(\mathbf{r}_l)] g(R)\, dS_k\, dS_l \tag{8.31a}$$

$$Z_{kl}^{K} = \int_{S_l}\int_{S_k} (\mathbf{r}_k - \mathbf{r}_l) \cdot [\mathbf{F}_k(\mathbf{r}_k) \times \mathbf{F}_l(\mathbf{r}_l)] \frac{1}{R} \frac{dg(R)}{dR}\, dS_k\, dS_l \tag{8.31b}$$

Since the patch basis functions, doublets, and multiplets represent linear combinations of initial basis functions given by (6.41b), the above impedances

can be expressed as linear combinations of the initial impedances (i.e., the impedances due to the initial basis functions). The initial impedances are evaluated according to (8.31) with the basis functions $\mathbf{F}_k$ and $\mathbf{F}_l$ replaced by the initial basis functions, $\mathbf{F}_{i_k j_k}$ and $\mathbf{F}_{i_l j_l}$,

$$\mathbf{F}_{i_k j_k} = \frac{f_{i_k}(p_k) g_{j_k}(s_k)}{|\mathbf{a}_{pk} \times \mathbf{a}_{sk}|} \mathbf{a}_{sk} \qquad \mathbf{F}_{i_l j_l} = \frac{f_{i_l}(p_l) g_{j_l}(s_l)}{|\mathbf{a}_{pl} \times \mathbf{a}_{sl}|} \mathbf{a}_{sl} \tag{8.32a,b}$$

where p_k and s_k (p_l and s_l) are the local p- and s-coordinates of the kth (lth) element, and $\mathbf{a}_{pk}$ and $\mathbf{a}_{sk}$ ($\mathbf{a}_{pl}$ and $\mathbf{a}_{sl}$) are the corresponding unitary vectors. $\nabla \cdot \mathbf{F}_{i_k j_k}$ and $\nabla \cdot \mathbf{F}_{i_l j_l}$ are determined according to (6.31). The surface elements, dS_k and dS_l, are expressed as $|\mathbf{a}_{pk} \times \mathbf{a}_{sk}| \, dp_k \, ds_k$ and $|\mathbf{a}_{pl} \times \mathbf{a}_{sl}| \, dp_l \, ds_l$, respectively, according to (5.16). Finally, the initial impedances are obtained as

$$Z^L_{i_k j_k i_l j_l} = \gamma \int_{S_l} \int_{S_k} \left[f_{i_k}(p_k) g_{j_k}(s_k) f_{i_l}(p_l) g_{j_l}(s_l) \mathbf{a}_{sk} \cdot \mathbf{a}_{sl} \right. \tag{8.33a}$$

$$\left. + \frac{1}{\gamma^2} f_{i_k}(p_k) \frac{dg_{j_k}(s_k)}{ds_k} f_{i_l}(p_l) \frac{dg_{j_l}(s_l)}{ds_l} \right] g(R) \, dp_k \, ds_k \, dp_l \, ds_l$$

$$Z^K_{i_k j_k i_l j_l} = \int_{S_l} \int_{S_k} f_{i_k}(p_k) g_{j_k}(s_k) f_{i_l}(p_l) g_{j_l}(s_l) (\mathbf{r}_k - \mathbf{r}_l) \tag{8.33b}$$

$$\cdot (\mathbf{a}_{sk} \times \mathbf{a}_{sl}) \frac{1}{R} \frac{dg(R)}{dR} \, dp_k \, ds_k \, dp_l \, ds_l$$

The evaluation of these integrals is simplified if the parametric equation for a generalized quadrilateral is represented as a polynomial of the parametric coordinates, p and s, and functions f_{i_k}, g_{j_k}, f_{i_l}, and g_{j_l} are expressed as simple power functions [i.e., $f_{i_k}(p_k) = p_k^{i_k}$, $g_{j_k}(s_k) = s_k^{j_k}$, $f_{i_l}(p_l) = p_l^{i_l}$, and $g_{j_l}(s_l) = s_l^{j_l}$]. In that case, the impedances in (8.33) are expressed in terms of the integrals

$$I^L_{i_k j_k i_l j_l} = \gamma \int_{-1}^{1} \int_{-1}^{1} p_l^{i_l} s_l^{j_l} P_{i_k j_k} \, dp_l \, ds_l \qquad I^K_{i_k j_k i_l j_l} = \int_{-1}^{1} \int_{-1}^{1} p_l^{i_l} s_l^{j_l} Q_{i_k j_k} \, dp_l \, ds_l \tag{8.34a,b}$$

where $P_{i_k j_k}$ and $Q_{i_k j_k}$ are the potential and field integrals given by (8.14). Since the integrals $I^L_{i_k j_k i_l j_l}$ ($I^K_{i_k j_k i_l j_l}$) occur only in the evaluation of initial impedances $Z^L_{i_k j_k i_l j_l}$ ($Z^K_{i_k j_k i_l j_l}$), they are called L-impedance (K-impedance) integrals. For the evaluation of these integrals, a simple numerical integration suffices. For the same accuracy, the K-impedance integrals require more integration points than the L-impedance integrals.

In particular, if the generalized quadrilateral is formulated as a bilinear surface, the impedances $Z^L_{i_k j_k i_l j_l}$ and $Z^K_{i_k j_k i_l j_l}$ are expressed in terms of the impedance integrals as

$$\begin{aligned} Z^L_{i_k j_k i_l j_l} = {} & (\mathbf{r}_{sk} \cdot \mathbf{r}_{sl}) I^L_{i_k, j_k, i_l, j_l} + (\mathbf{r}_{sk} \cdot \mathbf{r}_{psl}) I^L_{i_k, j_k, i_l+1, j_l} \\ & + (\mathbf{r}_{psk} \cdot \mathbf{r}_{sl}) I^L_{i_k+1, j_k, i_l, j_l} + (\mathbf{r}_{psk} \cdot \mathbf{r}_{psl}) I^L_{i_k+2, j_k, i_l+1, j_l} \\ & - \frac{j_k j_l}{\gamma^2} I^L_{i_k, j_k-1, i_l, j_l-1} \end{aligned} \tag{8.35a}$$

$$\begin{aligned} Z^K_{i_k j_k i_l j_l} = {} & (\mathbf{t}_c \cdot \mathbf{t}_s) I^K_{i_k, j_k, i_l, j_l} + (\mathbf{t}_c \cdot \mathbf{t}_{ps} + \mathbf{r}_{pl} \cdot \mathbf{t}_{sl} + \mathbf{r}_{pk} \cdot \mathbf{t}_{sk}) I^K_{i_k+1, j_k, i_l+1, j_l} \\ & + (\mathbf{t}_{sk} \cdot \mathbf{t}_c - \mathbf{r}_{pl} \cdot \mathbf{t}_s) I^K_{i_k, j_k, i_l+1, j_l} - (\mathbf{t}_{sl} \cdot \mathbf{t}_c - \mathbf{r}_{pk} \cdot \mathbf{t}_s) I^K_{i_k+1, j_k, i_l, j_l} \\ & - \mathbf{r}_{pl} \cdot \mathbf{t}_{sk} I^K_{i_k, j_k, i_l+2, j_l} - \mathbf{r}_{pk} \cdot \mathbf{t}_{sl} I^K_{i_k+2, j_k, i_l, j_l} \\ & - \mathbf{r}_{pl} \cdot \mathbf{t}_{ps} I^K_{i_k+1, j_k, i_l+2, j_l} + \mathbf{r}_{pk} \cdot \mathbf{t}_{ps} I^K_{i_k+2, j_k, i_l+1, j_l} \end{aligned} \tag{8.35b}$$

where $\mathbf{t}_c = \mathbf{r}_{ck} - \mathbf{r}_{cl}$, $\mathbf{t}_s = \mathbf{r}_{sk} \times \mathbf{r}_{sl}$, $\mathbf{t}_{sk} = \mathbf{r}_{sk} \times \mathbf{r}_{psl}$, $\mathbf{t}_{sl} = \mathbf{r}_{sl} \times \mathbf{r}_{psk}$, $\mathbf{t}_{ps} = \mathbf{r}_{psk} \times \mathbf{r}_{psl}$, and the vectors $\mathbf{r}_{ck}$, $\mathbf{r}_{pk}$, $\mathbf{r}_{sk}$ and $\mathbf{r}_{psk}$ ($\mathbf{r}_{cl}$, $\mathbf{r}_{pl}$, $\mathbf{r}_{sl}$ and $\mathbf{r}_{psl}$) define the kth (lth) bilinear surface according to (5.23). Recall that the above expressions are derived for the basis and testing functions directed along the s-parametric coordinate. If the basis functions or the testing functions are directed along the p-coordinate, the same expressions are obtained for each impedance type, except that the vectors $\mathbf{r}_{pk}$ and $\mathbf{r}_{sk}$, or the vectors $\mathbf{r}_{pl}$ and $\mathbf{r}_{sl}$, and subscripts i and j are interchanged. Evaluation of these eight types of impedances is reduced to only two classes of integrals given by (8.34). The time needed for combining the impedance integrals into the final matrix elements is negligible when compared with the time required to evaluate these integrals. Hence, the matrix fill-in time is practically equal to the time needed for the evaluation of the impedance integrals.

The system of linear equations can be constructed in two ways:

1. Compute the impedance integrals (8.34a,b) directly by two surface integrations and combine them into the matrix elements. The first integration is used for the evaluation of the potential and field integrals at the integration points of the second surface integration.
2. Implement the Galerkin method as the weighted point-matching method. In this case, the first integration remains the same as when the impedance integrals are directly evaluated. The second integration is performed through multiplication of a weighting matrix and the matrix of point-matching equations.

By inspection of the numbers of operations needed in these two cases, the second integration in the weighted point-matching method is found to be two to four times more time consuming than if the impedance integrals are directly evaluated.

Owing to the transformation (8.29), the initial impedances $Z^L_{i_k j_k i_l j_l}$ represent a linear combination of L-impedance integrals. If these impedances are evaluated through the weighted point-matching method, the transformation (8.29) cannot be applied and the initial impedances represent a linear combination of both the L- and K-impedance integrals. Hence, by using the same number of matching points, the initial impedances $Z^L_{i_k j_k i_l j_l}$ are evaluated more accurately in the direct way than through the weighted point-matching procedure.

For example, consider a small equilateral hexagonal scatterer of side a in a vacuum (Figure 8.6, inset). The scatterer is excited by a uniform plane

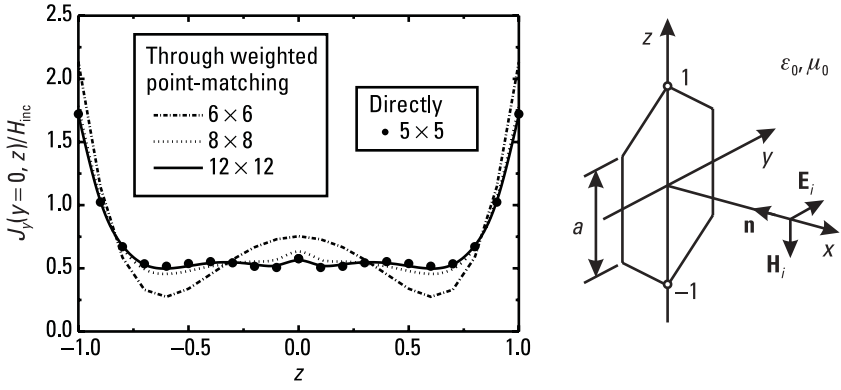

Figure 8.6 Normalized magnitude of the y-component of the surface current along the z-axis of hexagon scatterer ($a = 0.0536\lambda$) obtained when the impedances are evaluated directly (5×5 integration points per trapezoid) using the weighted point-matching method (6×6, 8×8, and 12×12 matching points per trapezoid).

electromagnetic wave normally incident on the scatterer surface. The incident wave is polarized in the direction of the y-axis. The scatterer is modeled by two trapezoids obtained by dividing the scatterer along the z-axis. The analysis is performed by solving the EFIE by the Galerkin method. The matrix elements (impedances) are evaluated in two ways: directly (with 5×5 integration points per trapezoid) and using the weighted point-matching method (with 6×6, 8×8, and 12×12 matching points per trapezoid). Each component of the surface currents over the two trapezoids is approximated by a double polynomial of the fourth order along both coordinates (i.e., $n_p = n_s = 4$).

Figure 8.6 shows the magnitude of the y-component of the surface current density vector along the z-axis of the hexagon ($a = 0.0536\lambda$) normalized with respect to the intensity of the incident magnetic field. The direct evaluation of the matrix elements, which includes (8.29), requires a much smaller number of integration points than the evaluation using the weighted point-matching method, which does not include (8.29).

8.3.3 Simplified Testing Based on Generalized Scalar Formulation of SIEs

The impedances $Z_{kl}^{\mathbf{n}\times L}$ and $Z_{kl}^{\mathbf{n}\times K}$ given by (8.26b) and (8.27b) can be written in terms of the basis functions in a similar manner to that shown in the previous section for the impedances Z_{kl}^{L} and Z_{kl}^{K} given by (8.26a) and (8.27a). If the parametric equation for a generalized quadrilateral is represented as a polynomial in the parametric coordinates, p and s, and the functions f_{i_k}, g_{j_k}, f_{i_l}, and g_{j_l} in (8.32) are expressed as simple power functions [i.e., $f_{i_k}(p_k) = p_k^{i_k}$, $g_{j_k}(s_k) = s_k^{j_k}$, $f_{i_l}(p_l) = p_l^{i_l}$, and $g_{j_l}(s_l) = s_l^{j_l}$], the impedances $Z_{kl}^{\mathbf{n}\times L}$ and $Z_{kl}^{\mathbf{n}\times K}$ can be expressed as linear combinations of the integrals:

$$I_{i_k j_k i_l j_l}^{\mathbf{n}\times L} = \gamma \int_{-1}^{1}\int_{-1}^{1} \frac{1}{|\mathbf{a}_p \times \mathbf{a}_s|}\, p_l^{i_l} s_l^{j_l} P_{i_k j_k}\, dp_l\, ds_l \tag{8.36a}$$

$$I_{i_k j_k i_l j_l}^{\mathbf{n}\times K} = \int_{-1}^{1}\int_{-1}^{1} \frac{1}{|\mathbf{a}_p \times \mathbf{a}_s|}\, p_l^{i_l} s_l^{j_l} Q_{i_k j_k}\, dp_l\, ds_l \tag{8.36b}$$

For parallelograms, $|\mathbf{a}_p \times \mathbf{a}_s|$ is constant, and the above integrals reduce to the L and K impedance integrals given by (8.34a,b), respectively. When the angle between two quadrilateral edges is equal to π, $|\mathbf{a}_p \times \mathbf{a}_s|$ is equal to zero at the common node of these edges [i.e., the integrands in (8.36) are singular func-

tions]. This also happens when the length of one edge tends to zero (when the quadrilateral degenerates into a triangle). However, if the angle is close to π or one edge is relatively short, these integrands are quasisingular functions. In all these cases, the evaluation of (8.36) requires many more integration points than the evaluation of the L and K impedance integrals.

Using the concept of the generalized scalar formulation of integral field equations (see Sections 4.1.2 and 4.1.7), the integrals (8.36a,b) can be replaced by the L and K impedance integrals, respectively [9]. Start with the impedances (8.26a,b). The basis functions $\mathbf{F}_k$ and $\mathbf{F}_l$ in (8.26a,b) can be represented as linear combinations of the initial basis functions $\mathbf{F}_{i_k j_k}$ and $\mathbf{F}_{i_l j_l}$ given by (8.32). The impedances (8.26a, b) can be expressed as linear combinations of the initial impedances (i.e., impedances due to the initial basis functions)

$$Z^{L}_{i_k j_k i_l j_l} = \int_{S_l} (\mathbf{F}_{i_l j_l} \cdot L\mathbf{F}_{i_k j_k})\, dS \qquad Z^{\mathbf{n}\times L}_{i_k j_k i_l j_l} = \int_{S_l} \mathbf{F}_{i_l j_l} \cdot (\mathbf{n}_l \times L\mathbf{F}_{i_k j_k})\, dS \tag{8.37a,b}$$

According to (8.32b), the basis functions $\mathbf{F}_{i_l j_l}$ are directed along the s-coordinate. The expressions for basis functions directed along the p-coordinate are obtained by interchanging the p- and s-coordinates and the indices in (8.32b). If the functions f and g are of the same type, the p and s variants of $\mathbf{F}_{i_l j_l}$ can be written as

$$(\mathbf{F}_{i_l j_l})_p = F_{i_l j_l}\, \mathbf{a}_{pl} \qquad (\mathbf{F}_{i_l j_l})_s = F_{i_l j_l}\, \mathbf{a}_{sl} \tag{8.38a,b}$$

Hence, we have the p and s variants of the initial impedances given by (8.37a,b). For (8.37a) these variants can be written as

$$(Z^{L}_{i_k j_k i_l j_l})_p = \int_{S_l} F_{i_l j_l}\, \mathbf{a}_{pl} \cdot L\mathbf{F}_{i_k j_k}\, dS \qquad (Z^{L}_{i_k j_k i_l j_l})_s = \int_{S_l} F_{i_l j_l}\, \mathbf{a}_{sl} \cdot L\mathbf{F}_{i_k j_k}\, dS \tag{8.39a,b}$$

The value of the mixed product is not changed by a cyclical permutation of its terms, and these variants for (8.37b) can be written as

$$(Z^{\mathbf{n}\times L}_{i_k j_k i_l j_l})_p = \int_{S_l} F_{i_l j_l}(\mathbf{a}_{pl} \times \mathbf{n}_l) \cdot L\mathbf{F}_{i_k j_k}\, dS \tag{8.40a}$$

$$(Z^{\mathbf{n}\times L}_{i_k j_k i_l j_l})_s = \int_{S_l} F_{i_l j_l}(\mathbf{a}_{sl} \times \mathbf{n}_l) \cdot L\mathbf{F}_{i_k j_k}\, dS \tag{8.40b}$$

After replacing the cross-products in (8.40) by the complementary unitary vectors

$$\mathbf{a}_{pl} \times \mathbf{n}_l \to -\mathbf{a}_{sl} \qquad \mathbf{a}_{sl} \times \mathbf{n}_l \to \mathbf{a}_{pl} \tag{8.41a,b}$$

the p and s variants of the initial impedances given by (8.40) convert to complementary initial impedances given by (8.39); that is,

$$(Z^{\mathbf{n}\times L}_{i_k j_k i_l j_l})_s \to (Z^{L}_{i_k j_k i_l j_l})_p \qquad (Z^{\mathbf{n}\times L}_{i_k j_k i_l j_l})_p \to -(Z^{L}_{i_k j_k i_l j_l})_s \tag{8.42a,b}$$

Thus, we avoid evaluation of integral (8.36a), since the initial impedances given by (8.39) can be expressed in terms of the L-impedance integrals (8.34a). Similarly, starting from the impedances (8.27a,b), evaluation of the integral (8.36b) can be replaced by evaluation of the K-impedance integrals (8.34b).

The consequence of (8.41) on the solution of the SIEs is as follows. By applying (8.41b), we replace the s-projections of $\mathbf{n} \times \mathbf{H}(\mathbf{n} \times \mathbf{E})$ in the SIEs by the p-projections of $\mathbf{H}(\mathbf{E})$ multiplied by e_p/e_s, where e_p and e_s are the Lame coefficients for the p- and s-coordinates. Similarly, by applying (8.41b) we replace the p-projections of $\mathbf{n} \times \mathbf{H}(\mathbf{n} \times \mathbf{E})$ by the s-projections of $\mathbf{H}(\mathbf{E})$ multiplied by e_s/e_p. Such replacements actually correspond to a particular generalized scalar formulation of the SIEs.

Consider the CFIE given by (4.30). After applying (8.41), the CFIE is obtained in the generalized scalar form (4.32), which is written for projections instead of components, where

$$\alpha_{sp} = -\frac{1}{\alpha}\frac{e_p}{e_s} \qquad \alpha_{ps} = \frac{1}{\alpha}\frac{e_s}{e_p} \tag{8.43a,b}$$

In order that the generalized scalar formulation has a unique solution, the coefficients α_{sp} and α_{ps} should satisfy the condition (4.37). Since e_p and e_s are always greater than zero, the condition (4.37) is satisfied whenever $\alpha > 0$. Solution of such a generalized scalar form of the CFIE by the Galerkin method requires evaluation of the L-impedance integrals only. In general, most SIEs

can be written in a generalized scalar form whose solution by the Galerkin method requires evaluation of the L- and K-impedance integrals only.

8.4 Choice of Optimal Test Procedure

The following can be seen in Chapter 2 and previous sections of the current chapter:

- The point-matching method is easiest to implement, although there are some difficulties in the automatic choice of matching points for complex problems.
- For the same number of unknowns, the least-squares method is much more time consuming than the point-matching and the Galerkin methods.
- The efficiency of the Galerkin method can be increased by a proper transformation of impedances when the analysis time becomes close to that of the point-matching method for the same number of unknowns; if, in addition, the matrix of the system of linear equations is symmetrical, this analysis time is halved.

However, the main criterion in the choice of the optimal test procedure is the number of unknowns required for a specified accuracy. It is shown that the Galerkin method usually requires a much smaller number of unknowns than the other test procedures [10–12].

As a typical example, let us consider a resonant vertical monopole antenna above a perfectly conducting ground plane, base-driven by a TEM magnetic-current frill, the equivalent of which is shown in the inset of Figure 8.7. The monopole is modeled by a cylinder and a flat disc. In addition to the current continuity, the approximations of currents over the cylinder and the flat disc satisfy the quasistatic relation at the frill and the center of the flat disc. The approximations for the cylinder and the disc are polynomials of degree N and 2, respectively. The analysis is based on the EFIE combined with five test procedures: the point-matching method with a uniform distribution of points; the point-matching method with a nonuniform distribution according to the Gauss-Legendre integration formula; least-squares testing; Galerkin testing; and using subdomain sine testing functions (which corresponds to solving the Hallen equation by means of the point-matching method [13]).

Figure 8.7 shows the relative error in the antenna admittance ($b/a = 100$, $b/a = 2.3$) versus the number of unknowns in the analysis. The relative error is defined as $E = |Y - Y_0|/|Y_0|$, where Y_0 is the "exact" value obtained by

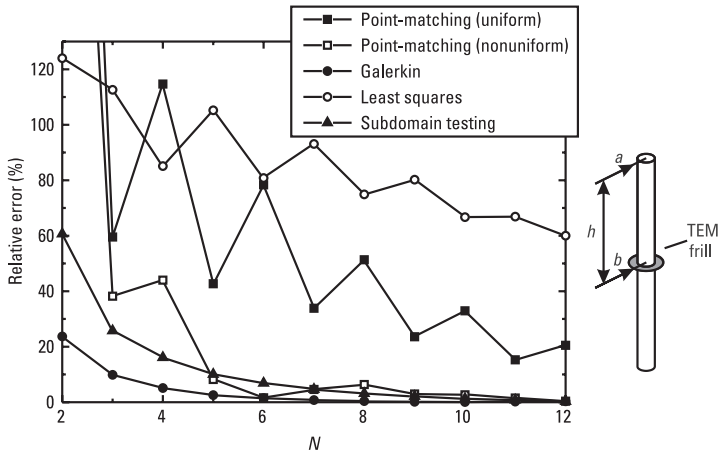

Figure 8.7 Relative error of the admittance of halfwave dipole (sketched in the inset) against the number of unknowns in the analysis for various test procedures. (After: [12].)

several methods with a sufficiently high degree of the current approximation. The following is seen:

- Galerkin testing is somewhat better than the subdomain testing and considerably better than other types of testing.
- Nonuniform point-matching is more accurate than uniform point-matching.
- Least-squares testing is the worst procedure.

The reason that the least-squares is the worst procedure can be explained by the fact that the polynomial approximation of currents cannot provide good satisfaction of the boundary conditions in the immediate vicinity of discontinuities. Attempting to satisfy the boundary condition with the least-squares error, we minimize the error locally (in the vicinity of discontinuities), which causes degradation of the overall solution. Obviously, by adopting more sophisticated basis functions, which enable easy satisfaction of the boundary conditions in the vicinity of the discontinuities, results obtained by all these testing procedures should converge. However, implementation of such basis functions is much more complicated, and the corresponding analysis would be much more time consuming. Having all this in mind, Galerkin testing seems to be optimal procedure.

8.5 Summary

For vector field equations, the inner product is usually formulated as the integral of the dot-product of two vector functions. In general, the testing is performed in nonorthogonal coordinate systems. For efficient evaluation of the inner products, the vector field equations should be decomposed into coordinate projections (along the inverse unit vectors), while testing functions should be decomposed into coordinate components (along the unit vectors). Such a decomposition also enables all test procedures to be realized as the weighted point-matching method. In particular, if building elements and currents are represented by polynomials in terms of parametric coordinates (as explained in Chapters 5 and 6), an application of the weighted point-matching method requires evaluation of only two classes of integrals (polynomial potential and field integrals).

The solution obtained by the Galerkin method depends significantly on the form of the SIE tested. Basically, there are two forms. Choosing the wrong form can even result in an undetermined system of linear equations. Hence, special care is devoted to the choice of the form that leads to stable and accurate solutions. To simplify testing, these forms can be further modified by using the concept of the generalized scalar formulation of the field equations (introduced in Chapter 4). In particular, if building elements and currents are represented by polynomials in terms of parametric coordinates, the evaluation of the matrix elements (impedances) can be reduced to evaluation of only two classes of integrals (L and K impedance integrals). Such direct evaluation of matrix elements is shown to be much more efficient than applying the weighted point-matching method.

Numerical experiments show that for the same accuracy, the Galerkin method requires fewer unknowns than the other test procedures. For the same number of unknowns, the analysis time of the Galerkin method is close to that of the point-matching method. Hence, regarding the efficiency of analysis, the Galerkin method is the optimal test procedure.

References

[1] Pearson, L. W., and C. M. Butler, "Inadequacies of Collocation Solutions to Pocklington-Type Models of Thin Wire Structures," *IEEE Trans. Antennas Propagat.,* Vol. AP-23, No. 2, March 1975, pp. 295–298.

[2] Butler, C. M., and D. R. Wilton, "Analysis of Various Numerical Techniques Applied to Thin-Wire Scatterer," *IEEE Trans. Antennas Propagat.,* Vol. AP-23, No. 4, July 1975, pp. 534–540.

[3] Popovic, B. D., M. B. Dragovic, and A. R. Djordjevic, *Analysis and Synthesis of Wire Antennas,* Chichester, U.K.: Research Studies Press, 1982.

[4] Djordjevic, A. R., and B. M. Kolundzija, "Analysis of Rectangular Scatterers," *Proc. XXIV Conf. ETAN,* Pt. 2, Pristina, Serbia, 1980, pp. 463–470 (in Serbian).

[5] Kolundzija, B. M., and A. R. Djordjevic, "Analysis of Dipole Antenna with Corner Reflector," *Proc. 7th Colloquium Microwave Commun.,* Budapest, Hungary, 1982, pp. 319–322.

[6] Wilton, D. R., et al., "Potential Integrals for Uniform and Linear Source Distributions on Polygonal and Polyhedral Domains," *IEEE Trans. Antennas Propagat.,* Vol. AP-32, No. 3, March 1984, pp. 276–281.

[7] Kolundzija, B. M., and B. D. Popovic, "A New, Rapid and Accurate Method for Evaluation of Potential Integrals Occurring in Thin Wire Antenna Problems," *Proc. 7th ICAP, IEE Conf. Publ. 274,* York, England, 1987, pp. 35–38.

[8] Kolundzija, B. M., "Electromagnetic Modeling of Composite Metallic and Dielectric Structures," *IEEE Trans. Microwave Theory Techniques,* Vol. AP-47, No. 7, July 1999, pp. 1021–1032.

[9] Kolundzija, B. M., "Generalized Combined Field Integral Equation," *Digest IEEE AP-S Int. Symp.,* Vol. 2, July 1996, pp. 852–855.

[10] Miller, E. K., and F. J. Deadrick, "Some Computational Aspects of Thin-Wire Modeling," in *Computer Techniques for Electromagnetics,* R. Mittra (ed.), Oxford, U.K.: Pergamon Press, 1973

[11] Kolundzija, B. M., "Electromagnetic Modeling of Wire to Plate Structures," D. Sc. Thesis, University of Belgrade, Department of Electrical Engineering, 1990 (in Serbian).

[12] Kolundzija, B. M., and B. D. Popovic, "Entire-Domain Galerkin Method for Analysis of Generalized Wire Antennas and Scatterers," *Proc. IEE H,* Vol. 139, No. 1, February 1992, pp. 17–24.

[13] Wilton, D. R., and C. M. Butler, "Efficient Numerical Technique for Solving Pocklington's Equation and Their Relationship to Other Methods," *IEEE Trans. Antennas Propagat.,* Vol. AP-24, 1976, pp. 83–86.

9

Practical Examples

Once we have well-developed algorithms (i.e., tools) for electromagnetic modeling of composite metallic and dielectric structures, we can use these tools to solve real-life problems. For example, starting from the known concept of the device (antenna, microwave circuit, and so forth), we can optimize their dimensions to fulfill given requirements. In addition, we can investigate how small changes in device dimensions affect the device properties to determine tolerances with which the fabrication must be performed. In some cases, it is convenient to systematically change the geometrical parameters of the device model and construct design curves. If we do not know which type of the device can satisfy given requirements, we can try different structures before deciding which one should be optimized. Besides design tasks, we sometimes use the modeling tools to better understand or verify a given device (either instead of measurements or in combination with them).

However, for efficient usage of tools for electromagnetic modeling, we must have some experience in both applied electromagnetics (e.g., antennas) and numerical modeling. Such experience gives us hints as to what can be modeled, how, what computer resources (in time and memory) are required, and so forth. In particular, it enables us to decide which details should be modeled and how precisely. Without such discrimination (i.e., taking unnecessary details into account), many problems would not fit into given resources. However, selection of the wrong details to consider (e.g., exclusion of crucial details from the structure model) leads to erroneous results.

In what follows it will be shown how some real-life problems can be efficiently model on a PC computer. The problems are solved using the WIPL-D software package, which includes some of the most powerful techniques elabo-

rated in the book [1]. All calculations are performed on Pentium at 733 MHz with 768-MB RAM.

9.1 TV-UHF Panel Antenna with Radome

For television broadcasting in the UHF range (470–860 MHz), we often use broadband panel antennas made of two or four dipoles. The antenna *voltage standing wave ratio* (VSWR) must be less than 1.2. It is difficult to achieve such a low VSWR in the whole frequency range. In the past, the design of such antennas was performed by the "cut and try method" based on the antenna measurements, which was a tedious and time-consuming job. Today, we can significantly speed up the design procedure using electromagnetic modeling techniques.

For example, we present the design procedure of a panel antenna consisting of four dipoles for the frequency range 470–800 MHz. The design was performed in two steps. In the first step, we optimized the panel antenna consisting of two dipoles. In the second step, two panel antennas made of two dipoles were merged into a single antenna and a matching circuit was added.

The complete starting model of the two-dipole panel antenna, together with a radome, is shown in Figure 9.1(a). The antenna consists of two flat

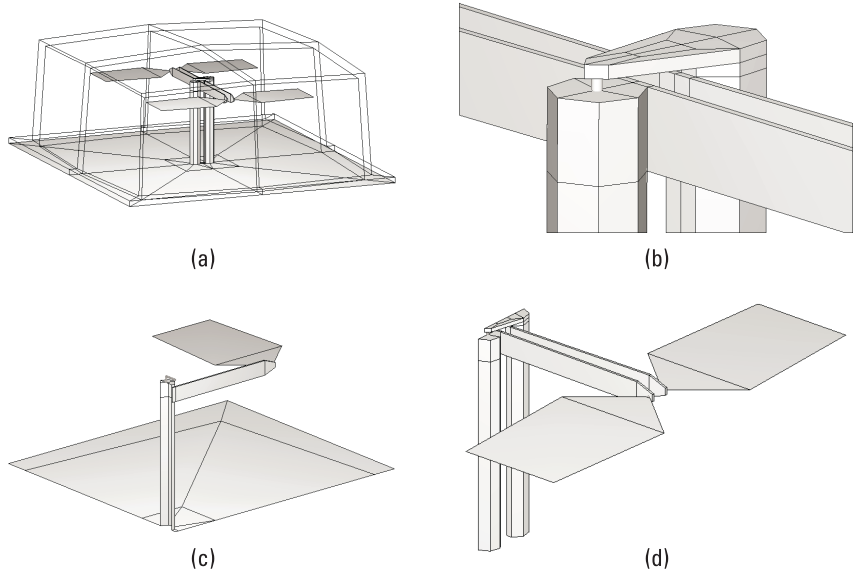

Figure 9.1 Two-dipole TV UHF panel antenna: (a) complete model, (b) feeding area, (c) model with simplified feed, and (d) simplified model with infinite ground plane.

(broadband) dipoles in front of a metallic reflector. The dipoles are fed by a power divider consisting of two striplines connected to the balun (balance-to-unbalance transformer). The balun is made in the form of a two-wire line (thick metallic cylinders) connected to the reflector on one side and to the striplines on the other side. The thick cylinder, shown left, represents the outer conductor of a coaxial feeder, whose inner conductor goes from the backside of the reflector through the cylinder. The other end of the inner conductor is connected by a thick flat plate to the thick cylinder, shown right [Figure 9.1(b)]. The flat dipoles and the reflector are modeled as infinitesimally thin plates, while the striplines are modeled taking into account their thickness. The coaxial feeder is modeled by a TEM magnetic current frill placed on the top of a fat cylinder, shown left. In order to speed up the analysis, the symmetry of the problem was utilized.

Before the starting model was optimized, we fabricated it (without radome) to check if the electromagnetic modeling is reliable enough. Figure 9.2 shows the input impedance of the antenna. Good agreement between theoretical (numerical) and measured results is observed. We also tried a number of simplified models. First, we replaced the thick flat plate (which connects inner conductor of the coaxial feeder to right thick cylinder) by a thin wire, thus introducing another symmetry plane into the model. In that case, only one quarter of the structure should be defined [Figure 9.1(c)]. As a result, the radiation resistance practically does not change, but the antenna reactance significantly increases (Figure 9.2). Hence, such a model cannot be used for impedance optimization, but is still good for far-field calculations. Then we

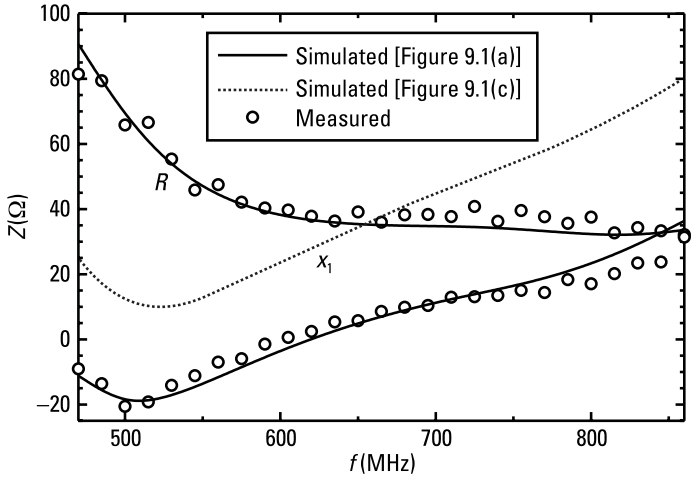

Figure 9.2 Input impedance of TV UHF-panel antenna without radome (shown in Figure 9.1).

tried to represent the striplines in the model without radome by infinitesimally thin plates. The results for the radiation resistance and antenna reactance became unacceptable. Finally, we tried to replace the reflector with a PEC plane [Figure 9.1(d)]. The results for the antenna impedance practically did not change. Hence, this model is convenient for impedance optimization, but cannot be used for far-field calculations.

Starting from the model shown in Figure 9.1(d), we optimized the size and shape of the dipoles and the striplines in order to minimize the VSWR of the antenna with respect to 50Ω. After 20 iterations, we obtained the model for which VSWR was less than 1.2 in the whole frequency range of interest. The fabricated model was modified, and the theoretical results were confirmed by measurements. Thereafter two such antennas were merged into a single unit. Measurements showed that the input impedance of one two-dipole antenna in the presence of another identical antenna was practically the same as when the antenna was alone. Hence, without further optimization of the antenna shape, we designed a matching circuit (T-junction, whose two arms are continued by impedance transformers made of coaxial-line sections). Figure 9.3 shows the final results for VSWR. Good agreement between calculated and measured results is observed. For more details see [2].

Consider the influence of a radome on radiation patterns. Figure 9.4 gives four groups of results for the gain in the main direction of the two-dipole TV UHF panel antenna. The results obtained with an air-filled radome agree very well with those obtained without a radome, which proves the reliability of the numerical model applied. The results for a radome without losses are shifted

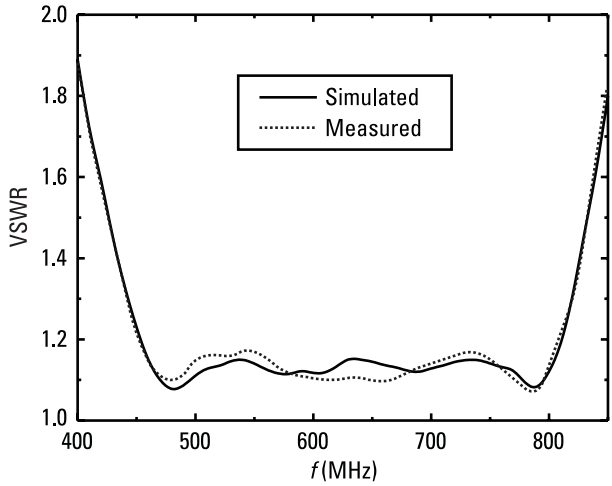

Figure 9.3 VSWR of optimized TV UHF panel antenna consisting of four dipoles.

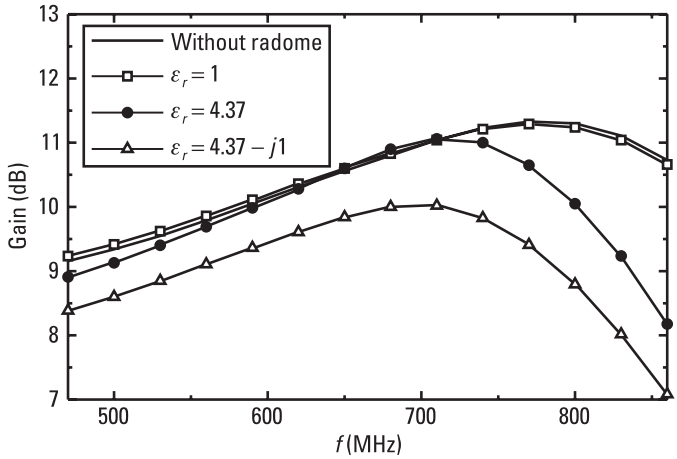

Figure 9.4 Influence of a radome on the gain of a TV UHF panel antenna (shown in Figure 9.1).

toward lower frequencies, as is expected. The same shift is observed for a lossy radome with additional attenuation of 1 dB to 2 dB.

9.2 Horn Antennas

Classical horn antennas are metallic sectoral, pyramidal, and conical horn antennas. Today, there are many varieties of corrugated, dielectric-loaded, and solid dielectric horns. [The *dielectric rod antenna* (DRA) is only a special case of solid dielectric horns.] All these antennas are widely used as feeds for large reflectors and lens antennas, high-gain elements in phased arrays, and standard-gain antennas in antenna measurements. Hence, there is continuous interest in developing methods for the analysis of such antennas.

The classic analysis of purely metallic horn antennas is based on the equivalence principle using the approximate aperture field distribution, giving satisfactory results for the main beam of the radiation pattern and the antenna gain [3]. The 2-D GTD model, which includes edge-diffracted fields, yields an improvement in the E-plane pattern [4]. A more accurate analysis is based on the solution of integral equations by the MoM (subdomain basis functions) or FDTD [5, 6]. However, these methods cannot be efficiently used for the analysis of electrically large horns. The efficiency of the MoM (subdomain basis functions) can be partly improved if it is hybridized with a full-wave stepped-waveguide method used for modeling the horn interior [7, 8]. However, once developed for horns of circular or rectangular geometries, this technique cannot

be applied for other geometries. Finally, this technique cannot evaluate the input impedance of a horn antenna if it is fed by a coaxial line.

The analysis of dielectric-loaded and solid dielectric horns is much more difficult than the analysis of purely metallic horns. Analytical methods give a good insight into how these antennas operate. However, good results for the radiation pattern are obtained for the main beam only [9]. For a precise analysis, we can use the MoM applied to the SIEs. However, methods based on subdomain approximation of currents require too many unknowns for arbitrarily shaped horns. For cylindrical and conical antennas, the efficiency can be increased utilizing the rotational symmetry [10]. In general, the efficiency can be improved by using the entire-domain approximation for surface currents, as will be shown by the following two examples.

As a first example, let us consider the pyramidal horn antenna shown in Figure 9.5(a). The horn is excited by a waveguide, which is fed by a coaxial line. Two geometrical models of such antenna are used. In both of them, it is supposed that the antenna is made of a perfect conductor. The finite-thickness metallic walls are modeled as infinitesimally thin plates, resulting in the surface currents that represent the sum of interior and exterior currents. When the antenna admittance is evaluated, the coaxial feed is modeled by a TEM magnetic current frill placed at the bottom of the waveguide. The frill excites a wire probe. A symmetry plane is used to speed up computations, as indicated in Figure 9.5(b). The same model can be used for evaluation of the radiation pattern. However, if only the far field is needed, the coaxial feed can be replaced by a dipole antenna situated in the waveguide in such a manner that two symmetry planes can be used [Figure 9.5(c)].

Figure 9.6(a) shows H-plane patterns for a 20-dB standard-gain horn (l_a = 4.87 in, l_b = 3.62 in, l_3 = 10.06 in, a = 0.9 in, b = 0.4 in, and $l_1 + l_2$ = 1.57 in) at 10 GHz [11]. The theoretical results are obtained using 800 unknowns, and the analysis time is less than one minute. The electrical length of the structure is greater than 10 wavelengths in the free space. Figure 9.6(b) shows the input admittance of the horn antenna (a = 1.37 in, b = 0.62 in, l_a = 6.3 in, l_b = 4.8 in, l_1 = 0.3 in, l_2 = 1.33 in, and l_3 = 4.72 in) versus frequency. [The dimensions not drawn in Figure 9.5(a) are the height and the diameter of the wire probe, h = 0.28 in and $2r$ = 0.067 inches.] Good agreement with measured results in both cases is observed.

As the second example, we consider a solid dielectric cone with a metallic sleeve [Figure 9.7(a)] [12]. The antenna is excited by a circular waveguide, which extends to a short metallic skirt that covers the beginning of the dielectric cone. A solid dielectric cone (without the skirt and sleeve) can support the HE_{11} mode, which has inherently very good aperture pattern symmetry. However, the radiation from the cone lateral surface interferes with and degrades the open-

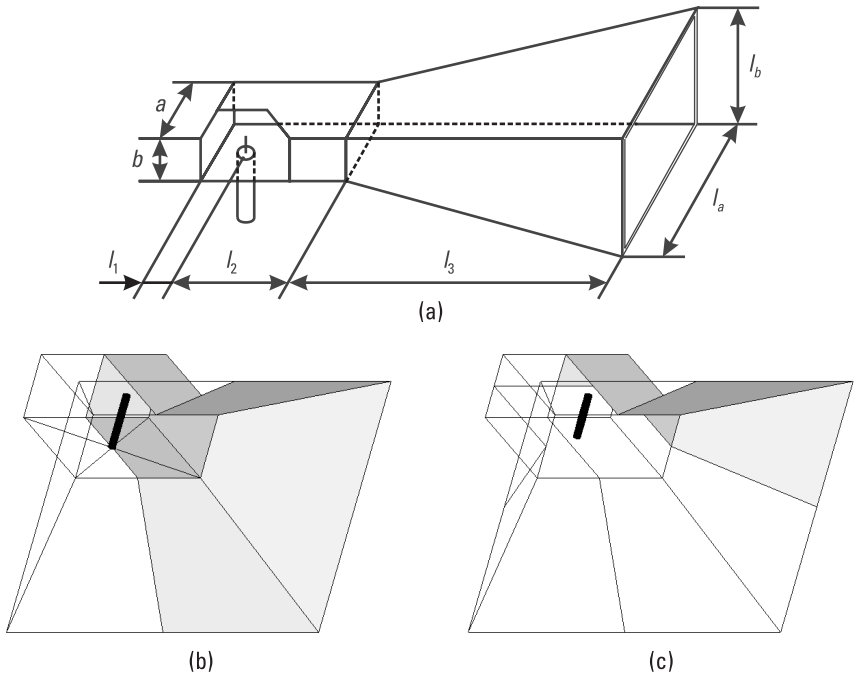

Figure 9.5 Pyramidal horn antenna: (a) sketch, (b) model with a symmetry plane, and (c) model with two symmetry planes.

end radiation. A metallic sleeve is used to shield the lateral surface radiation, but it must be far enough from the dielectric cone to prevent perturbation of the cone HE_{11} field structure. The role of the skirt is to provide that the sleeve is far enough from the dielectric cone at the throat of the feeding waveguide. The cylindrical waveguide, sleeve, and skirt are modeled as infinitesimally thin plates. The cylindrical waveguide is fed by a small dipole antenna, such that two symmetry planes are used [Figure 9.7(b)].

Figure 9.8(a) shows the gain normalized to the maximum value for the antenna without a skirt (a = 11.65 mm, b = 42.5 mm, c = 56.5 mm, L = 125 mm, and ϵ_r = 2.57). Theoretical results are obtained using 1,969 unknowns, and analysis time is about 8 minutes. The electrical length and the diameter of the structure are seven and four wavelengths, respectively. The computed results agree quite well with measurements. The differences in the H-plane near the –40-dB level may be related with the incomplete representation of the horn outer wall. The antenna without a skirt exhibits poor pattern symmetry between the E- and H-planes. Figure 9.8(b) shows the horn gain when the skirt is taken into account. The theoretical results reproduce the perfect measured pattern symmetry.

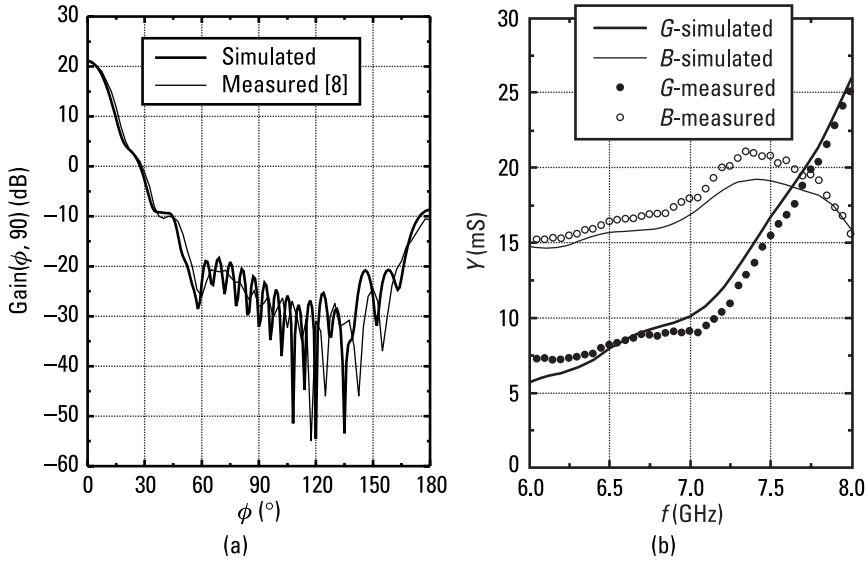

Figure 9.6 (a) Gain in *H*-plane of a 20-dB standard-gain horn (l_a = 4.87 in, l_b = 3.62 in, l_3 = 10.06 in, a = 0.9 in, b = 0.4 in, and $l_1 + l_2$ = 1.57 in) at 10 GHz; and (b) input admittance of a horn antenna (a = 1.37 in, b = 0.62 in, l_a = 6.3 in, l_b = 4.8 in, l_1 = 0.31 in, l_2 = 1.33 in, l_3 = 4.72 in, h = 0.28 in, and $2r$ = 0.067 in) versus frequency. (After: [11].)

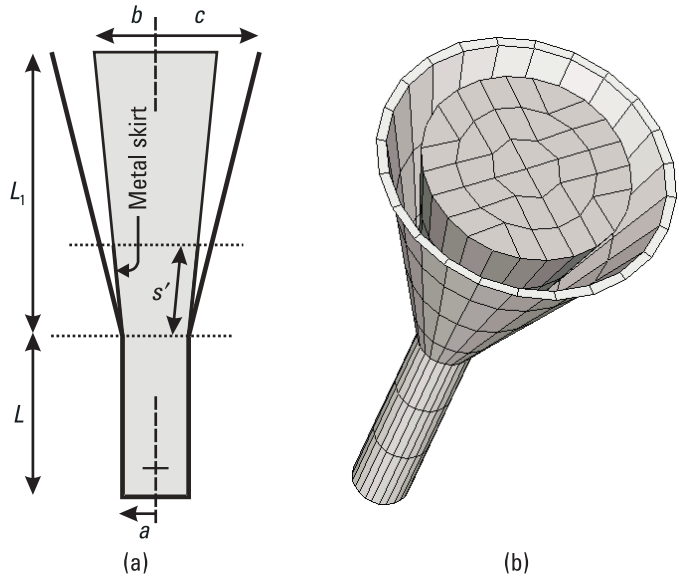

Figure 9.7 Shielded solid dielectric horn: (a) geometry and (b) model.

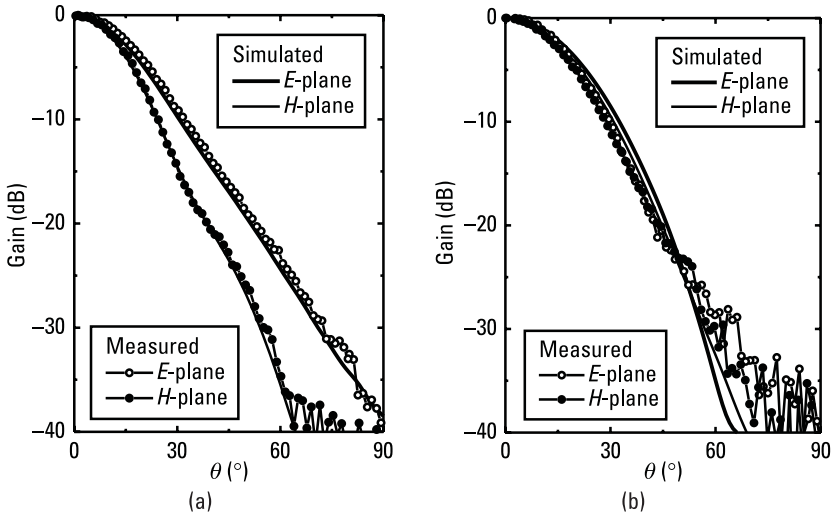

Figure 9.8 Gain in E- and H-planes of shielded solid dielectric horn (a = 11.65 mm, b = 42.5 mm, c = 56.5 mm, L = 125 mm, and ϵ_r = 2.57): (a) without skirt and (b) with skirt. (After: [12].)

9.3 Paraboloidal Reflector Antenna with Feed and Feed Struts

The scattered field from an electrically large reflector is usually evaluated by high-frequency techniques: the GO combined with the *aperture field* (AF) method, or the PO. Sophisticated methods include the GTD and the *physical theory of diffraction* (PTD) [13–15]. However, these techniques cannot be used for analysis of feeds for such antennas. A precise analysis of feeds can be performed by the MOM, as elaborated in the previous section. Hence, one possible way to determine the far field due to a given paraboloidal antenna illuminated by a given feed is to determine the near field of the feed using the MoM, and then to expose the reflector to this near field and use high-frequency techniques to determine the far field. However, in this way we do not take into account the blockage effect due to the feed and feed struts. The only way to take all these effects precisely into account is to apply the MoM to the whole structure. Another case when the MoM should be applied for the reflector surface instead of the high-frequency techniques is when the surface is strongly shaped to create a specific radiation pattern.

For example, let us consider a paraboloidal reflector antenna with a feed and feed struts [16], whose model is shown in Figure 9.9. The paraboloidal reflector is electrically very large. Therefore, its analysis requires a very large number of unknowns. However, the currents over the reflector vary relatively

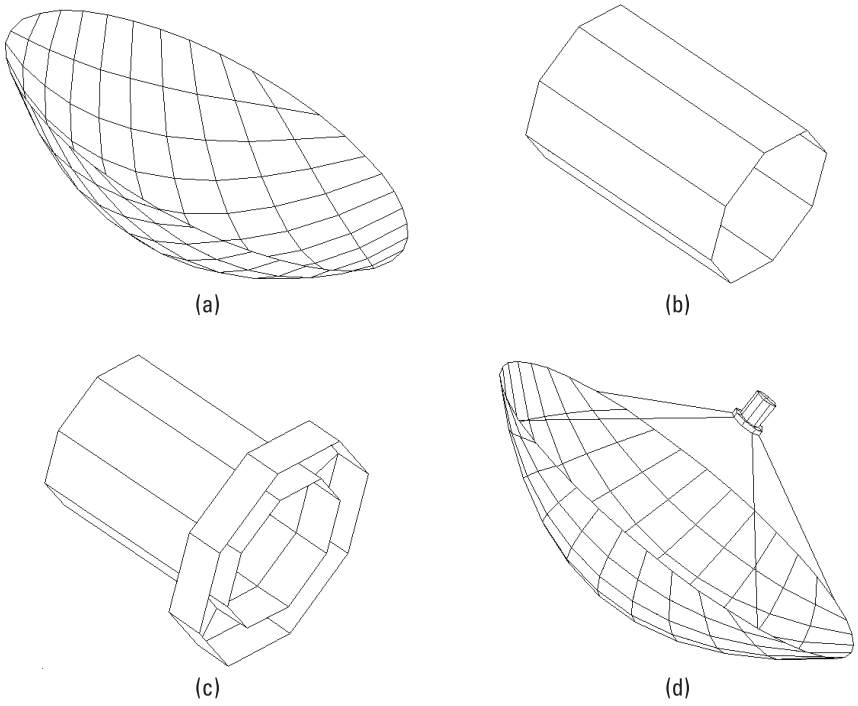

Figure 9.9 Paraboloidal reflector antenna with feed and feed struts: (a) paraboloidal reflector illuminated by Hertzian dipole, (b) circular waveguide feed, (c) circular waveguide feed with a choke, and (d) complete model.

slowly compared with the currents over the feed. Obviously, we need many more unknowns per wavelength squared for the feed than for the reflector. To determine the optimal size of patches and orders of approximation for currents, it is recommended to analyze scattering separately from the reflector and radiation of the feed before analyzing the complete model.

The scattering from the paraboloidal reflector is analyzed using a Hertzian dipole antenna for illumination [Figure 9.9(a)]. (The dipole is placed in the focal point of the reflector.) The reflector is modeled by bilinear surfaces as explained in Section 5.5. The maximum size of the bilinear surfaces that yields satisfactory results depends on the part of radiation pattern of interest. When the main lobe and few side lobes are considered, the maximum size of the bilinear surface is $S = 3\lambda^2$. The minimal order of approximation along each coordinate is two to three, resulting in three to five unknowns per λ^2. In particular, for the analysis of prime-focus symmetrical reflector antennas, two symmetry planes are used, needing minimally one unknown per λ^2. When the far field should be evaluated for all directions in the space, the maximum size

of the bilinear surface is $S = 1.5\lambda^2$. In that case, the minimal order of approximation along each coordinate is again 2 to 3, resulting in 6 to 10 unknowns per λ^2. For the analysis of prime-focus symmetrical reflector antennas, two symmetry planes are used, needing minimally two unknowns per λ^2. Figure 9.10 shows the relative power in the H-plane for an antenna whose diameter and focal distance are $D = 10.65\lambda$ and $f = 2.66\lambda$. The results are obtained using 144 patches. However, utilizing two symmetry planes, the number of unknowns is only 400. (Almost identical results are obtained using only 100 patches and 300 unknowns.) Good agreement with the results from [14] obtained by the PO + GTD method is observed.

In this example we tested two kinds of feeds: the circular waveguide feed with and without choke. The cylindrical surface of the waveguide is modeled by flat rectangular patches [Figure 9.9(b) and 9.9(c)]. The order of approximation along each coordinate is about four per wavelength. The best results are obtained if the n-sided polygon aperture of the model and the circular aperture of the original feed are of equal surfaces. (It can be shown than $n = 8$ suffices in most cases.) Figure 9.11 shows the relative power of cylindrical waveguide of radius 0.5λ and length 1.5λ. If two symmetry planes are used, the number of unknowns is about 100. Good agreement with results from [10] is observed.

Once we determine the optimal analysis parameters for the reflector and the feed, the complete antenna is easily analyzed. The feed-support struts are modeled by four wires that connect the choke of the circular waveguide with the reflector edge [Figure 9.9(d)]. Figure 9.12 shows the gain in the E-plane of such an antenna ($D = 20\lambda$, $f = 10\lambda$) with and without the feed support struts. The number of unknowns is $N = 653$, and the analysis time is $t = 5$

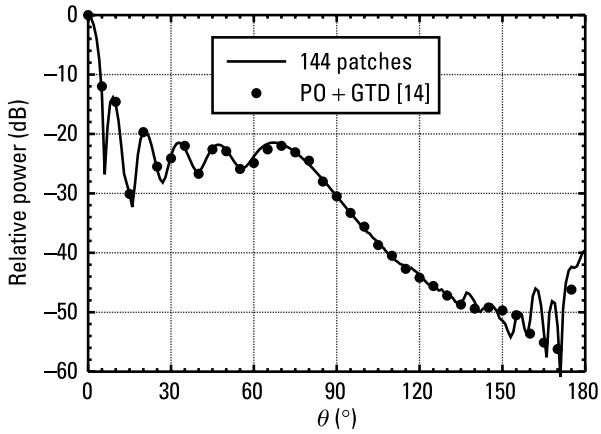

Figure 9.10 Relative power in the H-plane for paraboloidal antenna illuminated by Hertzian dipole antenna ($D = 10.65\lambda$ and $f = 2.66\lambda$).

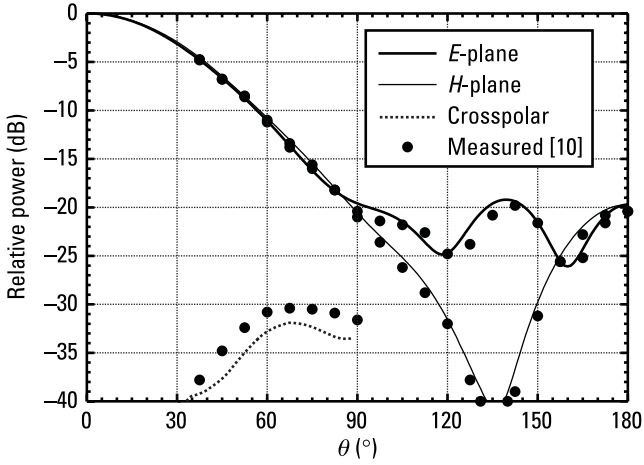

Figure 9.11 Relative power due to cylindrical waveguide feed ($r = 0.5\lambda$, $l = 1.5\lambda$).

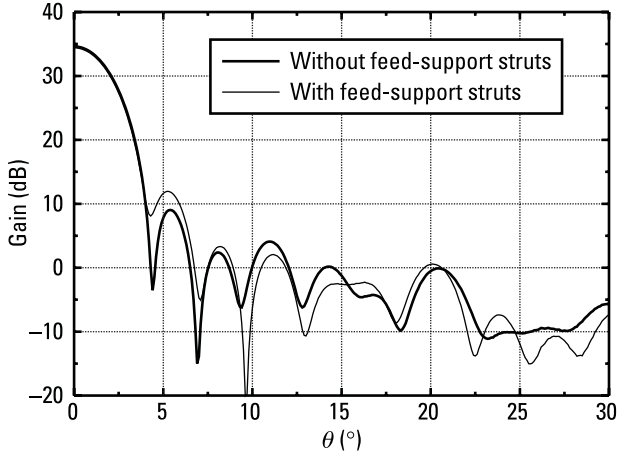

Figure 9.12 Gain of a paraboloidal reflector antenna fed by cylindrical waveguide with a choke ($D = 20\lambda$, $f = 10\lambda$). Results are obtained with and without feed struts taken into account. (After: [16].)

minutes. The struts do not influence the main lobe and increase the first sidelobe level for 2 dB.

9.4 Stacked Patch Antenna Mounted on an Airplane

Microstrip antenna technology has been one of the most rapidly developing topics in the antenna field in the last two decades. Their nonelectrical characteris-

tics usually make microstrip antennas preferrable to other types of radiators (i.e., low profile, lightweight, and conformal). However, when compared with traditional antenna elements, they suffer from a number of serious drawbacks (e.g., narrow bandwidth, high losses of the feeding network, or low power handling capacity). A lot of research remains to be done to overcome these drawbacks. Hence, there is a continuous interest in developing methods for the analysis of such antennas.

Most often the width of the substrate that surrounds the antenna patch is much greater than the substrate thickness. The antenna thus behaves in the same way as if the substrate is infinite. The most efficient way to include such substrates into analysis is to use Green's function for infinite material layers [17, 18]. However, if the width of the surrounding substrate is comparable to the substrate thickness, then the substrate must be modeled as a finite dielectric body. This happens often with multilayered structures, or when the microstrip patch antenna is built into the body of a ground vehicle or an airplane.

For example, let us consider a stacked patch antenna built in the fuselage of an airplane, shown in Figures 9.13 and 9.14. At the central operating frequency (f = 1.5 GHz), the electrical length of the airplane (l = 12m) is 60 wavelengths. The airplane body is electrically very large, and its analysis requires very large number of unknowns. Currents over the airplane body vary relatively slowly when compared with currents over the antenna and its vicinity. Hence, we need many more unknowns per wavelength squared for the antenna and its vicinity than for the rest of the airplane body. If we use the same density

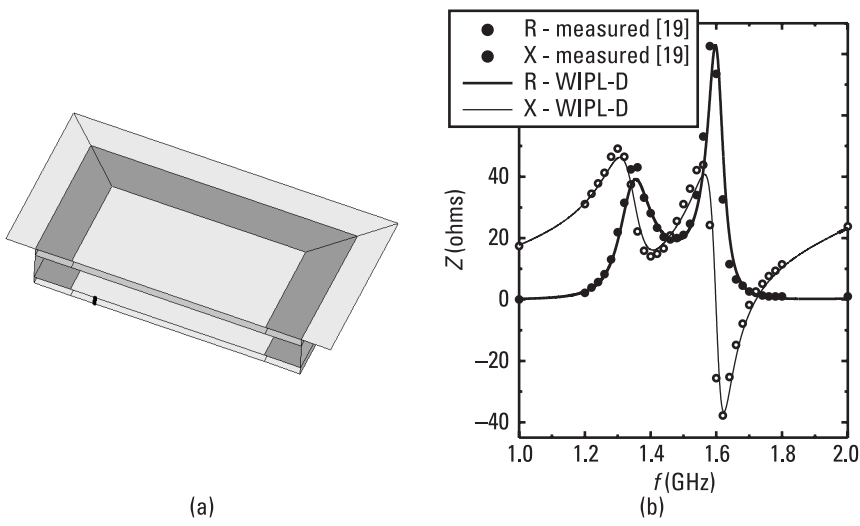

Figure 9.13 Stacked patch antenna: (a) model and (b) input impedance versus frequency.

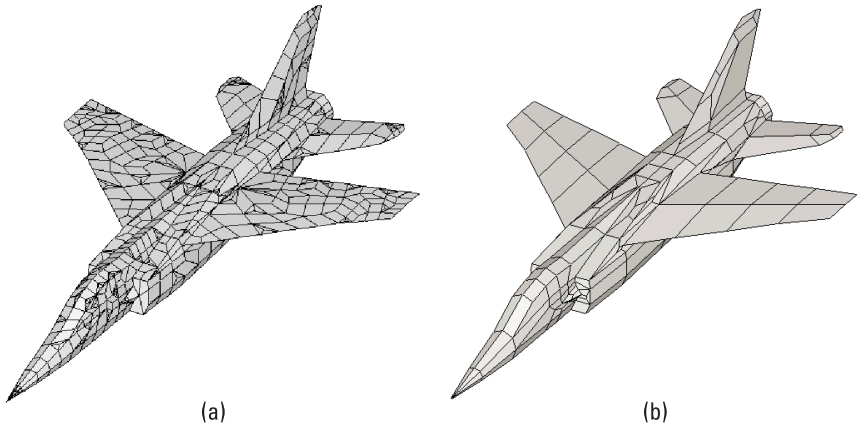

Figure 9.14 Airplane models made of (a) small patches and (b) large patches.

of unknowns for all parts of the structure, the total number of unknowns would be much greater that can be handled nowadays by PCs. Hence, it is very important to determine the optimal orders of approximation for currents before the analysis of the complete model. This can be done if we analyze scattering from the fuselage separately from radiation of the stacked patch antenna.

The stacked patch antenna consists of three square dielectric layers of side a = 91 mm. The first and the third layer have the same thickness ($d_1 = d_3$ = 3.175 mm) and the same relative permittivity ($\epsilon_{r1} = \epsilon_{r3} = 2.1$). The thickness of the second layer is d_2 = 11 mm, and its relative permittivity is $\epsilon_{r2} = 1.04$. The square active patch of side b = 66.2 mm is placed between the first and second layer, while the passive patch of the same size is placed at the top of the third layer. The bottom of the antenna and antenna sides are enclosed in a metallic cavity, which serves as a ground. The active patch is fed by a coaxial line positioned at a distance D = 24.75 mm from the center of the antenna, so that the antenna has one symmetry plane. The model of one half of the structure is shown in Figure 9.13(a). Figure 9.13(b) shows the input impedance of the antenna versus frequency. (The number of unknowns is 421 with the symmetry option applied.) Good agreement with measured data from [19] can be observed.

We started from the conventional airplane model consisting of 1,816 triangles. However, as elaborated in Chapters 5 and 6, modeling by bilinear surfaces is approximately two times more efficient than modeling by triangles. Hence, we converted the triangular model into a quadrilateral model in two steps. First, wherever it was possible, we merged pairs of neighboring triangles into convex quadrilaterals. Then, each orphan triangle (about 10% of all triangles) was subdivided into quadrilaterals, leaving a small triangular hole in the

middle of the original triangle. Thus we obtained the model shown in Figure 9.14(a). The patches are still relatively small at 1.5 GHz, so that low-order basis functions suffice. However, due to the large number of patches, the number of unknowns is too high to be handled by PCs. Efficiency can be further improved if we merge existing quadrilaterals into larger quadrilaterals, so that higher-order basis functions can be used. As shown in Chapter 6, these require fewer unknowns per wavelength squared than lower-order basis functions. This job was done manually, resulting in the model shown in Figure 9.14(b). Then, scattering from an airplane illuminated broadside (looking from the bottom) was analyzed. It can be shown that order two to three per wavelength for each coordinate suffices. In particular, the approximation order in the shadowed region can be decreased to two per wavelength for each coordinate.

Based on the above experience, we analyzed the stacked patch antenna mounted at the bottom of the airplane. Figure 9.15 shows the gain in azimuthal plane. The results are compared with the case when the antenna is not built into the airplane body. The difference in the results shows how much the radiation in the azimuthal plane is suppressed by the airplane body.

9.5 Base-Station Antenna with Cosecant Characteristic at 60 GHz

In the recent years, *wireless local area network* (WLAN) systems at 60 GHz are in the scope of research, development, and standardization. An important part of such systems is a base (remote) station antenna. Placed at a ceiling of height

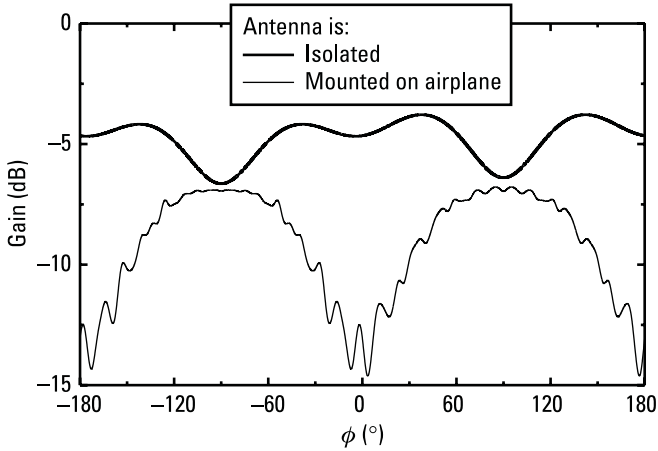

Figure 9.15 Gain in azimuthal plane due to stacked patch antenna mounted on the bottom of an airplane.

h_2, such an antenna should provide uniform power-level coverage of terminal antennas placed at height h_1 and maximal distance R_{max}. According to Figure 9.16, the corresponding ideal relative field pattern can be written in the form

$$F(\theta) = \begin{cases} F_{max}\dfrac{\cos\theta_{max}}{\cos\theta} & \theta < \theta_{max} \\ 0 & \theta > \theta_{max} \end{cases} \qquad \theta_{max} = \arctan\frac{R_{max}}{h_1 - h_2} \tag{9.1}$$

where F_{max} is the maximum of the relative field pattern, which occurs at the elevation angle θ_{max}. (In practice, θ_{max} takes values relatively close to 90°.) It is not possible to realize the pattern step and completely suppress the radiation for $\theta > \theta_{max}$. What is really important is to avoid interference with other base stations and the terminal antennas belonging to them. This condition can be fulfilled if the relative field pattern in the upper halfspace is not greater than some $F_2 << F_{max}$. In addition, we can allow that it linearly changes from F_{max} to F_2 for $\theta_{max} < \theta < 90°$.

Such a pattern can be approximately achieved by various types of antennas. One possible type is a quarter-wavelength monopole antenna mounted in the middle of a shaped circular reflector [20] (Figure 9.17). However, due to rotational symmetry, such an antenna cannot radiate along the axis (i.e., for $\theta = 0°$). Inability to cover one direction or even a small spatial angle around this direction does not discredit this antenna as a candidate for the base-station antenna. Namely, it rarely happens that the terminal antenna is placed in this direction. Even if it is placed in this direction, it can communicate with the

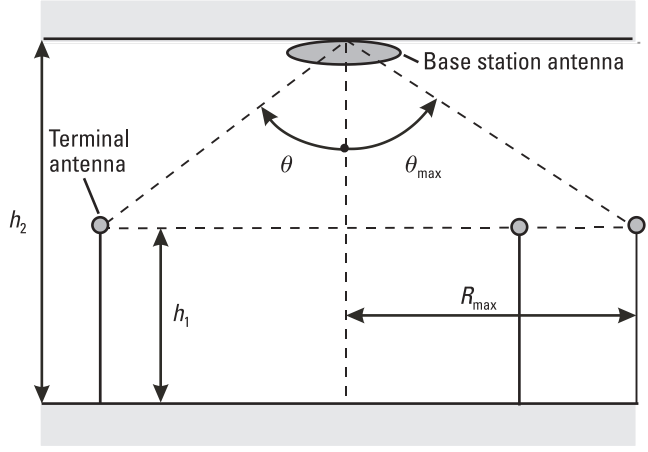

Figure 9.16 Working scenario of indoor base-station antenna.

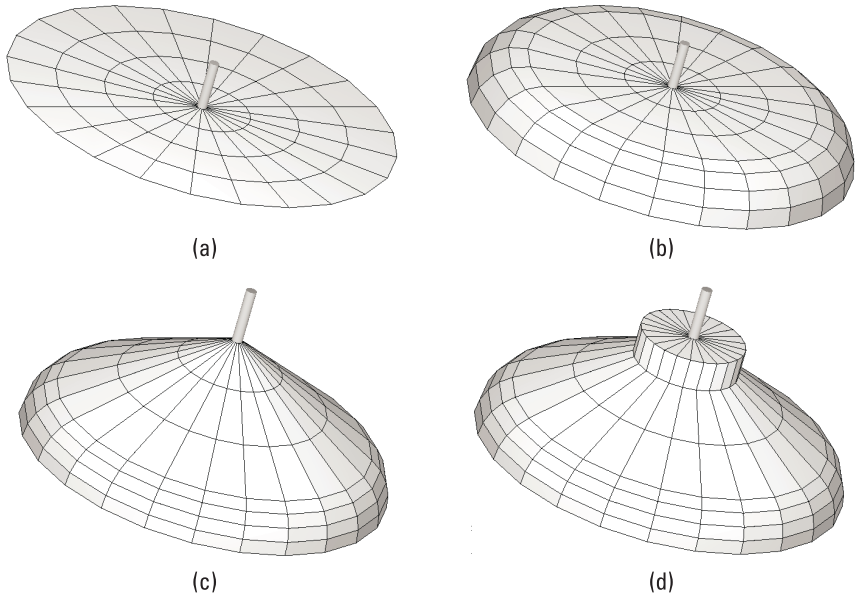

Figure 9.17 Sketch of monopole antenna mounted on a shaped conical reflector: (a) flat reflector with sharp edges, (b) flat reflector with round edges, (c) conical reflector with round edges, and (d) conical reflector with round edges and pedestal bump.

base station by locally reflected waves. Finally, it is always possible to slightly move the terminal antenna to ensure good communication. Hence, we allowed that the relative field pattern is a linear function when θ goes from 0° to some small $\theta_1 << \theta_{max}$.

The simplest form is a monopole antenna mounted on a flat circular reflector [Figure 9.17(a)]. It has often been investigated in the past. The first results were obtained by measurement [21]. A conical monopole above a flat disk is analyzed using spheroidal functions [22, 23]. For electrically large reflectors, the GTD can successfully be applied [24, 25]. This solution can be further improved by using the hybrid method, which combines the MoM for the monopole with the GTD for the disk [26, 27]. However, for reflectors whose diameter is either comparable with the wavelength or much smaller than the wavelength, the MoM is preferred [28]. Combining some of the above-mentioned methods, [29] presents an exhaustive investigation of the monopole antenna above flat disk. The elevation angle of peak gain varies from $\theta = 34°$ to $\theta = 61°$ for $0.5\lambda < a < 2.3\lambda$, where a is the flat disk radius.

To obtain greater elevation angles of peak gain, as required for base-station antennas, the reflector radius should be increased much above $a = 2.3\lambda$. For efficient analysis of such electrically large reflectors, we utilize rotational

symmetry. (The number of unknowns depends only on the total length of the generatrix, and three to four unknowns per wavelength are required.) After analyzing several examples, we realize that by increasing the radius, the diffraction lobes become very pronounced. To reduce these undesirable ripples in the radiation pattern, reflector edges are rounded [Figure 9.17(b)]. By increasing the curvature radius of the round edge up to $R_e = \lambda/2$, the diffraction lobes decrease very quickly. By increasing the radius above $R_e = \lambda$, the radiation pattern practically does not change. Hence, we always use $R_e = \lambda$ for rounded edges.

By increasing the radius of the flat circular reflector with round edges, the elevation angle of peak gain increases very slowly. For example, for $a = 8\lambda$ and $a = 16\lambda$, the elevation angle is $\theta_{max} = 73°$ and $\theta_{max} = 78°$, respectively. The elevation angle close to $\theta_{max} = 90°$ can be achieved by using only very large flat reflectors. To decrease the overall antenna size, conical reflectors can be used instead of flat ones, as in Figure 9.17(c). Figure 9.18 shows the elevation angle of peak gain, θ_{max}, versus the reflector radius normalized by wavelength, a/λ, for five cone angles, $\theta_0 = 70°$, 80°, 90°, 100°, and 110°, where the cone angle is measured between the monopole and the cone generatrix. By increasing θ_0 by 10°, θ_{max} also increases, on average, by 10°.

Once a flat circular reflector is successfully designed to give the peak gain in the prescribed direction, an additional technique is required to shape only the radiation in the low and middle angle directions (e.g., up to $\theta = 45°$). This control can be performed by creating bumps over the reflector. The simplest bump has the form of a pedestal. A pedestal whose radius is smaller than $b = \lambda$ and height is smaller than $h = 0.2\lambda$ can be used to control the radiation pattern

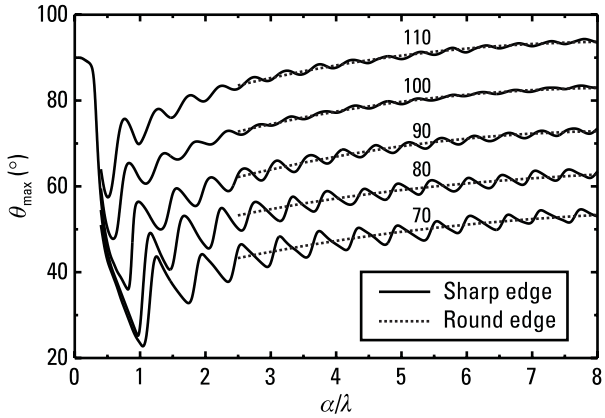

Figure 9.18 Peak gain versus reflector radius normalized by wavelength, a/λ, for five cone angles $\theta_0 = 70°$, 80°, 90°, 100°, or 110°.

in low- and middle-angle directions without affecting the main lobe. For example, let us consider a quarter-wavelength monopole antenna mounted on a pedestal of radius $b = 0.9\lambda$, which is placed on an infinite ground plane. Figure 9.19 shows the gain versus the elevation angle for four different heights: $h = 0.05\lambda$, 0.1λ, 0.15λ, and 0.2λ. The gain at low-angle directions increases by increasing the height. At the same time, the gain at middle-angle directions decreases.

If a conical reflector with round edges is combined with a pedestal bump [Figure 9.17(d)], a flexible antenna model is obtained. To design a base-station antenna with the cosecant characteristic, we used the following procedure:

- Starting from a fixed reflector radius, the peak of the main beam is positioned at θ_{max} by changing the cone angle.
- Sidelobes are fitted to the cosecant curve by optimizing the pedestal bump.
- The return loss is minimized by varying the monopole height and radius.

If F_2 is not small enough, the whole procedure is repeated for a greater reflector radius.

For example, let $a = 40$ mm, $f_{min} = 59$ GHz, $f_{max} = 64$ GHz, $\theta_{max} = 76°$, $\theta_1 \leq 10°$, $F_{max}/F_2 \leq -15$ dB, and $s_{11} \leq -10$ dB. According to these parameters, the antenna was designed, fabricated, and measured [20]. Figure 9.20 shows the gain of the optimal antenna versus the elevation angle at $f = 60$ GHz. Similar radiation patterns are obtained in the whole frequency range.

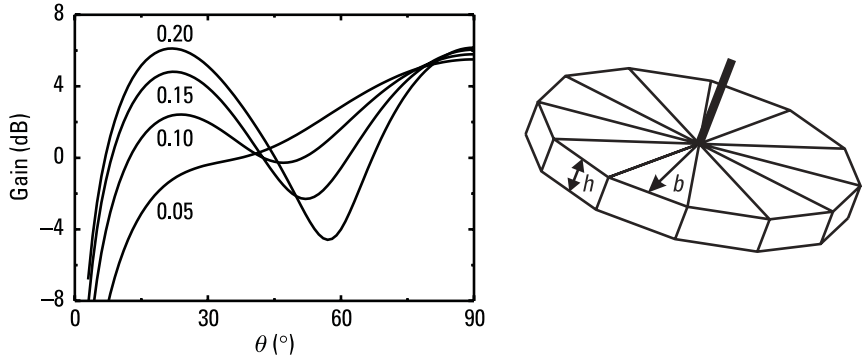

Figure 9.19 Gain versus elevation angle of quarter-wavelength monopole antenna placed on infinite ground plane with pedestal bump of radius $b = 0.9\lambda$, for different heights h.

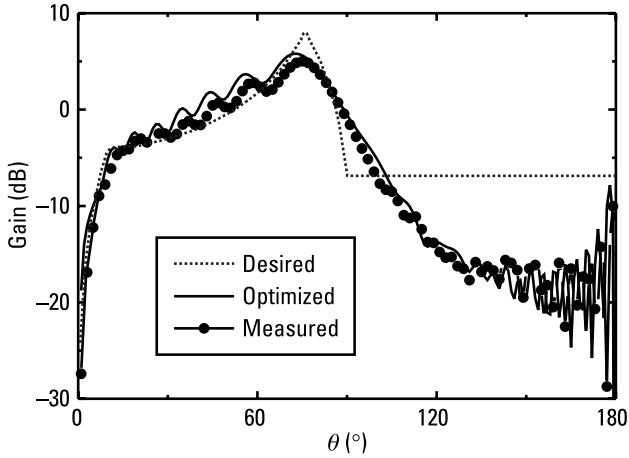

Figure 9.20 Gain versus elevation angle of indoor base-station antenna given in Figure 9.17(d): simulation and measurement are performed at f = 60 GHz. (After: [20].)

The simulated and measured results show good agreement. These results are very close to the desired pattern.

9.6 Summary

This chapter demonstrates how the tools for the electromagnetic modeling of composite metallic and dielectric structures can be used in real-life design processes. Examples are presented for a TV-UHF panel antenna with a radome, 20-dB gain horn antenna, shielded solid dielectric horn antenna, paraboloidal reflector antenna with feed and feed struts, stacked patch antenna mounted on an airplane, and base-station antenna with cosecant characteristic at 60 GHz.

References

[1] Kolundzija, B. M., J. S. Ognjanović, and T. K. Sarkar, *WIPL-D: Electromagnetic Modeling of Composite Metallic and Dielectric Structures, Software and User Manual,* Norwood, MA: Artech House, 2000.

[2] Kolundzija, B. M., B. Miletic, and J. Surutka, "New Approach in Design of TV UHF Panel Antenna," (invited paper—plenary session), *Proc. XLIII Conf. ETRAN,* Zlatibor, Serbia, 1999.

[3] Jasik, H., (ed.), *Antenna Engineering Handbook,* New York: McGraw-Hill, 1961.

[4] Yu, J. S., R. C. Rudduck, and L. J. Peters, "Comprehensive Analysis for E-Plane of Horn Antennas by Edge Diffraction Theory," *IEEE Trans. Antennas Propagat.,* Vol. AP-14, March 1996, pp. 138–149.

[5] Moheb, H., and L. Shafai, "Application of Integral Equations to Numerical Solution of Radiation from Horns," *Program and Abstracts of North American Radio Science Meeting,* London, Ontario, Canada, June 1991, p. 285.

[6] Katz, D. S., et al., "FTDT Analysis of Electromagnetic Wave Radiation from System Containing Horn Antennas," *IEEE Trans. Antennas Propagat.,* Vol. AP-39, August 1991, pp. 1203–1212.

[7] Kuhn, E., and V. Hombach, "Computer-Aided Analysis of Corrugated Horns with Axial or Ring-Loaded Radial Slots," *Proc. ICAP,* Pt. 1, 1983, pp. 127–131.

[8] Liu, K., et al., "Analysis of Pyramidal Horn Antennas Using Moment Methods," *IEEE Trans. Antennas Propagat.,* Vol. AP-41, October 1993, pp. 1379–1389.

[9] Salema, C., C. Fernandes, and R. K. Jha, *Solid Dielectric Horn Antennas,* Norwood, MA: Artech House, 1998.

[10] Olver, A. D., et al., *Microwave Horns and Feeds,* New York: IEEE Press, 1994.

[11] Kolundzija, B. M., et al., "Efficient Analysis of Horn Antennas by Using WIPL Code at Personal Computers," *Proc. IEEE AP-S Int. Symp.,* Baltimore, MD, June 1996, pp. 268–271.

[12] Fernandes, C., C. Salema, and M. Silveirinha, "WIPL-D Compared with Theory and Experiment," *Proc. 17th Applied Computational Electromagnetics Conf.,* Monterey, CA, March 2001, pp. 277–284.

[13] Narasimhan, M. S., and S. Christopher, "A New Method for Analysis of the Near and Far Fields of Paraboloidal Reflectors," *IEEE Trans. Antennas Propagat.,* Vol. AP-32, January 1984, pp. 13–19.

[14] Bach, H., "Comment on 'A New Method of Analysis of the Near and Far Fields of Paraboloidal Reflectors'," *IEEE Trans. Antennas Propagat.,* Vol. AP-33, January 1985, pp. 116–117.

[15] Duan, D. W., and Y. Rahmat-Samii, "A Generalized Diffraction Synthesis Technique for High Performance Reflector Antennas," *IEEE Trans. Antennas Propagat.,* Vol. AP-43, January 1995, pp. 27–40.

[16] Kolundzija, B. M., and Sarkar T., "Analysis of Prime-Focus Reflector Antennas by WIPL Code," *Proc. 26th EuMc,* Prague, Czech Republic, September 1996, pp. 781–785.

[17] Mosig, J. R., and F. E. Gardiol, "General Integral Equation Formulation for Microstrip Antennas and Scatterers," *Proc. IEE H,* Vol. 132, No. 7, December 1985, pp. 424–432.

[18] Michalski, K. A., and D. Zheng, "Electromagnetic Scattering and Radiation by Surfaces of Arbitrary Shape in Layered Media, Part I: Theory," *IEEE Trans. Antennas Propagat.,* Vol. AP-38, No. 3, March 1990, pp. 335–344.

[19] Gentili, G. G., et al., "Improved Green's Function Formulation for the Analysis of Rectangular Stacked Patch Antennas Enclosed in a Cavity," *Digest of IEEE AP-S Int. Symposium,* Vol. 2, Baltimore, MD, 1996, pp. 1070–1073.

[20] Kolundzija, B. M., V. Branković, and S. Zimmermann, "Design of Monopole Antenna Mounted on a Shaped Circular Reflector for mm-Wave Applications," *Proc. 10th ICAP,* Edinburgh, Scotland, 1997, pp. 1460–1463.

[21] Meier, A. S., and W. P. Summers, "Measured Impedance of Vertical Antennas over Finite Ground Planes," *Proc. IEEE,* Vol. 37, June 1949, pp. 609–616.

[22] Leitner, A., and R. D. Spence, "Effect of a Circular Ground Plane on Antenna Radiation," *J. Appl. Phys.,* Vol. 21, October 1950, pp. 1001–1006.

[23] Hann, R. F., and J. G. Fikioris, "Impedance and Radiation Pattern of Antennas Above Flat Discs," *IEEE Trans. Antennas Propagat.,* Vol. AP-21, January 1973, pp. 97–100.

[24] Green, H. E., "Impedance of a Monopole on the Base of a Large Cone," *IEEE Trans. Antennas Propagat.,* Vol. AP-17, November 1969, pp. 703–706.

[25] Green, H. E., "Radiation from a Monopole Excited Cone," *IEEE Trans. Antennas Propagat.,* Vol. AP-17, November 1969, pp. 707–715.

[26] Thiele, G. A., and T. H. Newhouse, "A Hybrid Technique for Combining Moment Methods with the Geometrical Theory of Diffraction," *IEEE Trans. Antennas Propagat.,* Vol. AP-23, January 1975, pp. 62–69.

[27] Awadalla, K. H., and T. S. M. Maclean, "Input Impedance of a Monopole Antenna at the Center of a Finite Ground Plane," *IEEE Trans. Antennas Propagat.,* Vol. AP-26, March 1978, pp. 244–248.

[28] Richmond, J. H., "Monopole Antenna on Circular Disk," *IEEE Trans. Antennas Propagat.,* Vol. AP-32, December 1984, pp. 1282–1287.

[29] Weiner, M. M., "Monopole Element at the Center of a Circular Ground Plane Whose Radius Is Small or Comparable to a Wavelength," *IEEE Trans. Antennas Propagat.,* Vol. AP-35, May 1987, pp. 488–495.

About the Authors

Branko M. Kolundzija is an associate professor at the School of Electrical Engineering at the University of Belgrade in Yugoslavia, where he received his B.Sc., M.Sc., and D.Sc. He is an author of several books and software packages, including *WIPL-D: Electromagnetic Modeling of Composite Metallic and Dielectric Structures, Software and User's Manual* (Artech House, 2000). His main areas of interest are numerical electromagnetics, antennas, microwaves, and EMC. His e-mail address is kol@kiklop.etf.bg.ac.yu.

Antonije R. Djordjević is a professor at the School of Electrical Engineering at the University of Belgrade in Yugoslavia, where he received his B.Sc., M.Sc., and D.Sc. He is a corresponding member of the Serbian Academy of Sciences and Arts. He is an author of several books and software packages, including *AWAS for Windows Version 2.0: Analysis of Wire Antennas and Scatterers, Software and User's Manual* (Artech House, 2002). His main areas of interest are numerical electromagnetics, antennas, microwaves, fast digital signal interconnects, and EMC. His e-mail address is edjordja@etf.bg.ac.yu.

Index