ESSENTIALS OF COMPUTATIONAL ELECTROMAGNETICS

ESSENTIALS OF COMPUTATIONAL ELECTROMAGNETICS

Xin-Qing Sheng
Beijing Institute of Technology, China

Wei Song
Beijing Institute of Technology, China

This edition first published 2012
© 2012 John Wiley & Sons Singapore Pte. Ltd.

Registered office
John Wiley & Sons Singapore Pte. Ltd., 1 Fusionopolis Walk, #07-01 Solaris South Tower, Singapore 138628

For details of our global editorial offices, for customer services and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com.

All Rights Reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, scanning, or otherwise, except as expressly permitted by law, without either the prior written permission of the Publisher, or authorization through payment of the appropriate photocopy fee to the Copyright Clearance Center. Requests for permission should be addressed to the Publisher, John Wiley & Sons Singapore Pte. Ltd., 1 Fusionopolis Walk, #07-01 Solaris South Tower, Singapore 138628, tel: 65-66438000, fax: 65-66438008, email: enquiry@wiley.com.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Designations used by companies to distinguish their products are often claimed as trademarks. All brand names and product names used in this book are trade names, service marks, trademarks or registered trademarks of their respective owners. The Publisher is not associated with any product or vendor mentioned in this book. This publication is designed to provide accurate and authoritative information in regard to the subject matter covered. It is sold on the understanding that the Publisher is not engaged in rendering professional services. If professional advice or other expert assistance is required, the services of a competent professional should be sought.

Library of Congress Cataloging-in-Publication Data

Sheng, Xin-Qing, 1968-
Essentials of computational electromagnetics / Xin-Qing Sheng, Wei Song.
p. cm.
Includes bibliographical references and index.
ISBN 978-0-470-82962-2 (hardback)
1. Electromagnetism–Data processing. 2. Electromagnetism–Mathematical models. I. Song, Wei. II. Title.
QC760.54.S54 2012
537.0285–dc23

2011045255

Set in 11/13 pt Times by Thomson Digital, Noida, India

Contents

Preface ix

1 Mathematical Formulations for Electromagnetic Fields 1
1.1 Deterministic Vector Partial Differential System of the Electromagnetic Fields 1
1.1.1 Maxwell's Equations 1
1.1.2 Constitutive Relations 3
1.1.3 Boundary Conditions 3
1.1.4 Maxwell's Equations in the Frequency Domain 5
1.1.5 Uniqueness Theorem 6
1.2 Vector Wave Equation of the Electromagnetic Fields 8
1.3 Vector Integral Equation of the Electromagnetic Fields 8
1.3.1 Equivalence Principle 9
1.3.2 Solution of Maxwell's Equation in Free Space 11
1.3.3 Integral Equations of Metallic Scattering Problems 14
1.3.4 Integral Equation of Homogeneous Dielectric Scattering Problems 16
1.3.5 Integral Equation of Inhomogeneous Dielectric Scattering Problems 19
1.3.6 Integral Equations of Scattering in Layered Medium 20
References 28

2 Method of Moments 29
2.1 Scattering from 3D PEC Objects 29
2.1.1 Formulation of the Problem 30
2.1.2 Discretization in MoM 30
2.1.3 Choice of Basis and Testing Functions 31
2.1.4 Discretized Integral Equation (DIE) and the Numerical Behavior Analysis 34
2.1.5 Handling of Singularity 36

2.1.6 Comparison of EFIE and MFIE 71
2.1.7 Interior Resonance Problem 73
2.1.8 Fast Multipole Method 74
2.1.9 Calculation of Scattered Fields 86
2.1.10 Writing Computer Program 89
2.1.11 Numerical Examples 94
2.1.12 Parallel Technology 100
2.1.13 Strong Scalability 106
2.1.14 Weak Scalability 107
2.2 Scattering from Three-Dimensional Homogeneous Dielectric Objects 109
2.2.1 Mathematic Formulation of the Problem 111
2.2.2 Discretized Forms and Their Numerical Performance 112
2.2.3 Numerical Examples 118
2.2.4 Implementation of Single Integral Equation and the Numerical Characteristics 122
2.3 Scattering from Three-Dimensional Inhomogeneous Dielectric Objects 128
2.3.1 Mathematic Formulation of the Problem 129
2.3.2 Rooftop Basis Functions 130
2.3.3 Discretization of the VIE 131
2.3.4 Singularity Processing 134
2.3.5 Fast Solution of the Discretized VIE 135
2.3.6 Numerical Examples 136
2.4 Essential Points in MoM for Solving Other Problems 136
2.4.1 Scattering from Two-Dimensional Objects 138
2.4.2 Scattering from Periodic Structures 141
2.4.3 Scattering from Two-and-Half-Dimensional Objects 144
2.4.4 Radiation Problems 146
References 150

3 Finite-Element Method 153
3.1 Eigenmodes Problems of Dielectric-Loaded Waveguides 153
3.1.1 Functional Formulation 154
3.1.2 Choice of Basis Functions 159
3.1.3 Discretization of the Functional 161
3.1.4 Imposition of the Boundary Condition 164
3.1.5 Solution of the Generalized Eigenvalue Equation 165
3.1.6 Computer Programming 166
3.1.7 Numerical Examples 170
3.2 Discontinuity Problem in Waveguides 170
3.2.1 Functional Formulation 171

3.2.2 Choice of the Basis Functions 174
3.2.3 Discretization of the Functional 176
3.2.4 Solution of the Linear Equations 178
3.2.5 Extraction of the Scattering Parameters 180
3.2.6 Numerical Examples 182
3.3 Scattering from Three-Dimensional Objects 184
3.3.1 Mathematic Formulation of the Problem 184
3.3.2 Writing Computer Program 187
3.3.3 Numerical Results 190
3.4 Node-Edge Element 192
3.4.1 Construction of Node-Edge Element 192
3.4.2 Implementation of Node-Edge Element 193
3.4.3 Numerical Examples 195
3.5 Higher-Order Element 196
3.6 Finite-Element Time-Domain Method 200
3.7 More Comments on FEM 203
References 205

4 Finite-Difference Time-Domain Method 207
4.1 Scattering from a Three-Dimensional Objects 207
4.1.1 FDTD Solution Scheme 208
4.1.2 Perfectly Matched Layers 209
4.1.3 Yee Discretizing Scheme 215
4.1.4 Discretization of the Scatterer Model 220
4.1.5 Treatment on the Curved Boundary 220
4.1.6 Determination of the Unit Size and the Time Step 222
4.1.7 Plane Waves in Time Domain 223
4.1.8 Calculation of Incident Plane Waves in Time Domain 225
4.1.9 Calculation of the Radar Cross Section 227
4.1.10 Computer Programing and Numerical Examples 229
4.2 Treatment for Special Problems 233
4.2.1 Treatments for Thin Metallic Wires 233
4.2.2 Treatments for Dispersive Media 235
4.2.3 Treatments for Lumped Elements 237
4.3 Comparison of the MoM, FEM and FDTD Methods 239
References 240

5 Hybrid Methods 243
5.1 Hybrid High-Frequency Asymptotic Methods and Full-Wave Numerical Methods 244
5.1.1 Hybird Physical Optics Method and FEM 244
5.1.2 Hybrid Physical Optics Method and Moment Method 248

5.2 Hybrid Full-Wave Numerical Methods 251
5.2.1 Hybrid FE-BI-MLFMA 252
5.2.2 Hybrid Method Combining EFIE and MFIE 266
5.2.3 Hybrid Method Combining FEM and Mode-Matching Method 271
References 276
Index 277

Preface

Computational electromagnetics (CEM) has evolved as an independent, vibrant discipline since the early 1960s. A large number of scholarly written works are already available which are exclusively dedicated to various aspects of numerical methods and their applications to electromagnetic (EM) problems. Despite the wide variety of these works, they all follow the same research paradigm, which essentially consists of discretizing the Maxwell's equations, obtaining the numerical solutions through writing computer programs, and validating the numerical solution by using theoretical values and/or experimental data. This research paradigm shows that discretization and computer realization are the essentials in CEM. Without discretization, the digital computer cannot be utilized; without computer realization, the numerical method cannot be validated and thus is not convincing. Although large amount of papers and books have been published in CEM, there are only three main full-wave numerical methods as far as discretization is concerned. These include the method of moment (MoM), the finite-element method (FEM), and the finite-difference time-domain (FDTD) method. These methods have distinct ways of discretization and possess fundamentally different numerical characteristics and performances. Generally, MoM discretizes the integral equation representations of EM problems, FEM discretizes the variational functional formulations of EM problems, while FDTD directly discretizes the Maxwell's equations in differential form. Numerous published works can be found which are exclusively devoted to either of the aforementioned methods. However, only few give a broad and general overview of the essential features innate in these methods, or the summary was made too early to include recent developments. As far as program realization is concerned, practical issues, such as numerical techniques and tricks in the conversion of discretized formulations into computer programs, play important roles. However, the majority of existing works primarily restrict themselves to presenting the final numerical results without specifying the aforementioned issues. Not addressing these practical aspects is very unfortunate in CEM, since it is not at all straightforward to convert discretized formulations into reliable computer programs as the designers would expect. In fact, significant effort is spent on the program realization in CEM. Obviously, writing a good computer program requires years of

experience – and it is in general rather difficult, if not impossible, to set rules for good programming practices. Nontheless, experience shows that studying typical programs can provide useful guidelines for writing efficient and professional computer programs. More to the point, it is hard to expect a monograph without discussions on numerical characteristics of the methods can serve the reader to acquire an in-depth understanding of the subject matter.

In this book, the authors will elaborate the above three methods in CEM in a paratactic manner, based on the authors' own research experiences as well as the information available in the current literatures, by thoroughly analyzing typical cases. If you look at the evolving history of CEM, several typical EM problems can be identified, which play a key role. The landmarks of progress in CEM are achieved by pursuing better solutions for these problems. This book focuses on these EM problems. Through solving these problems, the numerical methods are introduced and the essential principles, i.e., the techniques improving the numerical efficiency, and the skills in writing computer programs, are detailed. In addition, the numerical performances of the methods are discussed by analyzing the numerical results presented in the book. To help the reader obtain a fast and deep understanding of the major essentials about CEM, three codes are attached in this book, which are the computation of 3D scattering using MoM, FEM, and FDTD respectively. For issues related to broadening the scopes of the individual numerical techniques and their applications to other relevant problems, and aspects related to rigorous mathematical proofs, the reader is referred to the references available - they will not be elaborated further here.

Xin-Qing Sheng
January, 2012, Beijing

1

Mathematical Formulations for Electromagnetic Fields

There are various mathematical formulations for representing electromagnetic fields. They are equivalent to each other mathematically, but distinct in terms of numerical behavior. In this chapter, we will give a brief introduction of the three kinds of mathematical formulations employed in computational electromagnetics: vector partial differential equations, vector wave equations, and vector integral equations, which are the mathematical foundations of the finite-difference time-domain (FDTD) method, the finite-element method (FEM), and the method of moment (MoM) respectively.

1.1 Deterministic Vector Partial Differential System of the Electromagnetic Fields

A complete system of vector partial differential equations to determine the electromagnetic (EM) fields in an interested domain comprises three parts: the Maxwell's equations, the constitutive relations, and the boundary conditions. We will introduce them in the consecutive sections.

1.1.1 Maxwell's Equations

Based on fundamental studies carried out by Ampere and Faraday *et al.*, Maxwell introduced the concept of displacement current density (measured in A/m^2) and established the following system of equations that gives a complete mathematical description of the electromagnetic fields:

$$\nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0 \quad \text{(Faraday's law of induction)} \tag{1.1}$$

Essentials of Computational Electromagnetics, First Edition. Xin-Qing Sheng and Wei Song.
© 2012 John Wiley & Sons Singapore Pte. Ltd. Published 2012 by John Wiley & Sons Singapore Pte. Ltd.

$$\nabla \times \mathbf{H} - \frac{\partial \mathbf{D}}{\partial t} = \mathbf{J} \quad \text{(Ampere's law with Maxwell's invention)} \tag{1.2}$$

$$\nabla \cdot \mathbf{D} = \rho \quad \text{(Gauss' law for electric fields)} \tag{1.3}$$

$$\nabla \cdot \mathbf{B} = 0 \quad \text{(Gauss' law for magnetic fields)} \tag{1.4}$$

Here, the variables on the left-hand side (LHS) of the equations are the physical EM field quantities:

E: electric field intensity (vector), in $(\mathrm{V\,m^{-1}})$,
D: electric flux density (vector), in $(\mathrm{C\,m^{-2}})$,
H: magnetic field intensity (vector), in $(\mathrm{A\,m^{-1}})$,
B: magnetic flux density (vector), in $(\mathrm{Wb\,m^{-2}})$;

The variables on the right-hand side (RHS) of the equations are quantities depicting the sources that excite the EM fields:

J: volume current density (vector), in $(\mathrm{A\,m^{-2}})$,
ρ: free volume charge density (scalar), in $(\mathrm{C\,m^{-3}})$.

A further link to Maxwell's equations is established by the law of charge conservation

$$\nabla \cdot \mathbf{J} = -\frac{\partial \rho}{\partial t} \tag{1.5}$$

By applying the divergence operation to both sides of (1.1) we can obtain (1.4). Again, applying divergence operation to (1.2) and plugging (1.5) lead to (1.3). These indicate only three out of the above five equations are independent. Moreover, the three independent equations are not all required in solving a particular problem. For example, in solving electrostatic problems, we only use (1.1) and (1.3), because magnetic fields don't exist in this kind of problem and all other equations become irrelevant. Similarly, only (1.2) and (1.3) are used in magnetostatic problems. As far as EM waves excited by time-varying electric current, time-varying equivalent electric or magnetic current are concerned, generally, using (1.1) and (1.2) is enough. Obviously, these equations are not adequate in determining the unknown physical quantities. Take the EM wave problem as an example. Vector equations (1.1) and (1.2) provide us with 6 scalar equations and 12 unknowns contained in **E**, **H**, **D**, and **B**, if we assume the driving sources of the system **J** is known. Therefore, in order to render the above equations solvable, we need to invoke further relations, that is, the constitutive relations depicting the relationship between **D**, **B** and **E**, **H**, together with the relationship between **J** and **E**.

1.1.2 Constitutive Relations

The constitutive relations of a material are conventionally obtained by experiments or by establishing a microcosmic model of the material. Generally speaking, for many real-world materials, the constitutive relations can be written as

$$\mathbf{D} = \varepsilon \mathbf{E} \tag{1.6}$$

$$\mathbf{B} = \mu \mathbf{H} \tag{1.7}$$

$$\mathbf{J} = \sigma \mathbf{E} \tag{1.8}$$

Here ε, μ and σ stands for permittivity, in $(\mathrm{F\,m^{-1}})$, permeability, in $(\mathrm{H\,m^{-1}})$, and electrical conductivity, in $(\mathrm{S\,m^{-1}})$, respectively. If these constitutive parameters vary with the spatial position, the material is *inhomogeneous*. Otherwise, the material is *homogeneous*. If these parameters are frequency dependent, the material is *dispersive*. Plasma, water, biological tissues and wave-absorbing materials are examples of *dispersive* materials. Otherwise, the material is *nondispersive*. If the material has directional properties so that the constitutive parameters take a tensor form, the material is termed *anisotropic*. For example, permittivity in plasma and permeability in ferrite are tensors. Of course, there are also special materials whose constitutive relations are so complex that they cannot be expressed by (1.6)–(1.8). Take *chiral medium* for example. The electric displacement vector is related to not only the electric field intensity, but also the magnetic flux density. On the other hand, the magnetic flux density is related to both the magnetic field intensity and the electric field intensity. In this text, we will limit our discussion to the normal materials with constitutive relations expressed by (1.6)–(1.8). In particular for time-independent constitutive relations, (1.1) and (1.2) can be written as

$$\nabla \times \mathbf{E} = -\mu \frac{\partial \mathbf{H}}{\partial t} \tag{1.9}$$

$$\nabla \times \mathbf{H} = \varepsilon \frac{\partial \mathbf{E}}{\partial t} + \mathbf{J} \tag{1.10}$$

1.1.3 Boundary Conditions

Maxwell's equations, in concert with the constitutive relations, are still inadequate to uniquely determine the EM fields in a given domain. Further conditions, namely, the boundary conditions are necessary. The boundary conditions are different for various problems. We will give some examples in the following. In the electromagnetic area, many closed-domain problems, such as the eigenvalue problem of resonant cavity and

the scattering from metallic target, the bounding surface of the metal body involved defines the boundary. If we consider the metal to be perfect electric conductor (PEC), then the boundary condition of this kind of problem can be written as:

$$\hat{n} \times \mathbf{E} = 0 \tag{1.11}$$

or

$$\hat{n} \times \nabla \times \mathbf{H} = 0 \tag{1.12}$$

where $\hat{n}$ is the outward unit vector normal to the boundary. In mathematics, (1.11) is termed *the first-type (Dirichelet) boundary condition*. It specifies the values that the variables need to take on the boundary of the domain. Equation (1.12) is termed *the second-type (Neumann) boundary condition*. It specifies the value that the derivative of a solution needs to take on the boundary of the domain. In some problems, the boundary can be regarded as a perfect magnetic conductor (PMC). So we have

$$\hat{n} \times \mathbf{H} = 0 \tag{1.13}$$

or

$$\hat{n} \times \nabla \times \mathbf{E} = 0 \tag{1.14}$$

In contrast to the closed-domain problems, the boundary condition of open-domain problems (e.g., the radiation or scattering problem) normally cannot be written as either of the above conditions. Instead, the outer-boundary condition specifies the relation of the solution and the derivative of the solution on the outer boundary of the computational domain. For example, for the radiation or the scattering problem in free space, the boundary condition at infinity is:

$$\lim_{r \to \infty} r \left[\nabla \times \begin{pmatrix} \mathbf{E} \\ \mathbf{H} \end{pmatrix} + jk_0 \hat{r} \begin{pmatrix} \mathbf{E} \\ \mathbf{H} \end{pmatrix} \right] = 0 \tag{1.15}$$

where $r = \sqrt{x^2 + y^2 + z^2}$ and $\hat{r}$ stands for the unit vector in the radial direction. This is the so-called *Sommerfeld radiation condition*. Actually in electromagnetics, in order to simplify the problem, there are many other boundary conditions derived according to the physical characteristics of the problem. For instance, the surface of the ocean or the ground is approximated by the impedance boundary conditions. Local absorbing boundary conditions can be derived according to either the behavior of the scattered field in a scattering problem, or the transmission modes of waveguides in a propagation problem. Discussion of these conditions is beyond the scope of this text. The interested reader is referred to [1] for more in-depth discussion on impedance boundary conditions,

Sections 11.7–11.9 in [2] for more information on local absorbing boundary condition, and [3] for more details on absorbing boundary conditions based on waveguide modes.

1.1.4 Maxwell's Equations in the Frequency Domain

In the above three subsections we have introduced the Maxwell's equations, the constitutive relations, and the boundary conditions required to uniquely determine the EM fields in a given domain. Actually, very often we are only interested in the EM field or wave at a single frequency. Or to simplify the problem, we first study the time-harmonic (sinusoidal) EM field behavior, and then obtain the time-varying field through inverse Fourier transformation (IFT). For a time-harmonic EM field, the field can be expressed by the product of a time-independent complex vector and a time-dependent phasor $e^{j\omega t}$, where ω is the angular frequency. Thus, (1.9) and (1.10) can be written in the frequency (or phasor) domain as

$$\nabla \times \mathbf{E} = -j\omega\mu\mathbf{H} \tag{1.16}$$

$$\nabla \times \mathbf{H} = j\omega\varepsilon\mathbf{E} + \mathbf{J} \tag{1.17}$$

In the above we have omitted the phasor $e^{j\omega t}$. In order to clarify the physical meaning of the above equations, displacement electric and magnetic currents are included in the concept of generalized electric and magnetic currents. Under this concept, the induced electric and magnetic current densities are expressed as:

$$\mathbf{J}' = (\sigma + j\omega\varepsilon)\mathbf{E} = y\mathbf{E} \tag{1.18}$$

$$\mathbf{M} = j\omega\mu\mathbf{H} = z\mathbf{H} \tag{1.19}$$

where $y = \sigma + j\omega\varepsilon$ and $z = j\omega\mu$ stand for the media property and have the same units as admittance and impedance in unit length, respectively. Thus, the preceding equations can be understood uniformly as: an electric field is generated by time-varying magnetic current, and a magnetic field is generated by a time-varying electric current. To be specific, we separate the induced currents from the imposed driving source currents, and write the above equations as:

$$-\nabla \times \mathbf{E} = z\mathbf{H} + \mathbf{M}^i \tag{1.20}$$

$$\nabla \times \mathbf{H} = y\mathbf{E} + \mathbf{J}^i \tag{1.21}$$

where the electric and magnetic current densities, $\mathbf{J}^i$ and $\mathbf{M}^i$ are the driving sources. Equations (1.20) and (1.21) are the widely used equations to describe the time-

harmonic EM fields, which are often referred to as *the frequency-domain version of the electromagnetic equations.*

1.1.5 Uniqueness Theorem

As previously stated, after invoking the constitutive relations and the boundary conditions, we can solve the Maxwell's equations and determine the EM fields in a given domain. In this section, we will further illustrate under which circumstances the EM fields are uniquely determined. We will use the EM equations in the frequency domain as an example. The uniqueness theorem in the time domain can be obtained in a similar way, or the reader is referred to [4] for more details.

Consider the driving electric and magnetic sources $\mathbf{J}$ and $\mathbf{M}$ enclosed in an isotropic region S. The EM fields excited by these sources will satisfy (1.20) and (1.21). Suppose there are two sets of solutions $\mathbf{E}^a$, $\mathbf{H}^a$ and $\mathbf{E}^b$, $\mathbf{H}^b$, with their differences denoted as:

$$\delta\mathbf{E} = \mathbf{E}^a - \mathbf{E}^b \tag{1.22}$$

$$\delta\mathbf{H} = \mathbf{H}^a - \mathbf{H}^b \tag{1.23}$$

Subtracting the equations superscripted by 'a' from those superscripted by 'b' we obtain

$$\nabla \times \delta\mathbf{E} = -z\delta\mathbf{H} \tag{1.24}$$

$$\nabla \times \delta\mathbf{H} = y\delta\mathbf{E} \tag{1.25}$$

Then, multiplying (1.24) by $\delta\mathbf{H}^*$, multiplying the conjugate of (1.25) by $\delta\mathbf{E}$, and subtracting the resulting products yields

$$\delta\mathbf{E} \cdot \nabla \times \delta\mathbf{H}^* - \delta\mathbf{H}^* \cdot \nabla \times \delta\mathbf{E} = z|\delta\mathbf{H}|^2 + y^* \lfloor \delta\mathbf{E} \rfloor^2 \tag{1.26}$$

From vector identities we know that the LHS of (1.26) equals $-\nabla \cdot \left(\delta\mathbf{E} \times \delta\mathbf{H}^*\right)$. Thus

$$\nabla \cdot (\delta\mathbf{E} \times \delta\mathbf{H}^*) + z|\delta\mathbf{H}|^2 + y^* \lfloor \delta\mathbf{E} \rfloor^2 = 0 \tag{1.27}$$

Integrating both side of (1.27) over region V enclosed by S, and applying Gauss' divergence theorem give rise to

$$\oint_S (\delta\mathbf{E} \times \delta\mathbf{H}^*) \cdot dS + \int_V \left(z|\delta\mathbf{H}|^2 + y^* \lfloor \delta\mathbf{E} \rfloor^2\right) d\tau = 0 \tag{1.28}$$

If the result of the surface integration in (1.28) is zero, the volume integration must also be zero. Hence, we have

$$\int_V \left[\mathrm{Re}(z)|\delta\mathbf{H}|^2 + \mathrm{Re}(y)\lfloor\delta\mathbf{E}\rfloor^2\right] d\tau = 0 \tag{1.29}$$

$$\int_V \left[\mathrm{Im}(z)|\delta\mathbf{H}|^2 - \mathrm{Im}(y)\lfloor\delta\mathbf{E}\rfloor^2\right] d\tau = 0. \tag{1.30}$$

It is known from (1.18) that, for an isotropic lossy medium, $\mathrm{Re}(y)$ is a positive number. Therefore, as long as the medium is wholly lossy, (1.29) holds only when $\delta\mathbf{E} = 0$ holds everywhere in the medium regardless of the value of the lossy factor. Then, from (1.30) we can further know that $\delta\mathbf{H} = 0$ holds everywhere in the medium. On the other hand, for a lossless medium, (1.30) holds, as long as the stored electric energy equals the stored magnetic energy. This indicates resonance may exist and the solutions of the EM fields may no longer be unique. This multisolution brings challenges to computational electromagnetics. The interior resonance in MoM is one of the examples. We will show more details in Chapter 2.

Consequently, we can obtain the *uniqueness theorem*. This theorem states that, in a lossy region, the electromagnetic field solution will be uniquely specified in any one of the following three situations. The first situation is when the tangential components of electric field $\mathbf{E}$ over the boundary surface S are given. The second one is when the tangential components of magnetic field $\mathbf{H}$ over the boundary surface S are obtained. The last one is when the tangential components of electric field $\mathbf{E}$ over part of the boundary surface S and the tangential components of magnetic field $\mathbf{H}$ over the rest of it are presented. In a lossless region, there are unlimited resonant solutions, whose electric field energy stored in the region enclosed by S equals the stored magnetic field energy.

Problems

1.1 Different EM fields can be generated by different sources in the same region. Consider two sets of sources, $\mathbf{J}^{\mathrm{a}}$, $\mathbf{M}^{\mathrm{a}}$ and $\mathbf{J}^{\mathrm{b}}$, $\mathbf{M}^{\mathrm{b}}$, which generate the EM fields $\mathbf{E}^{\mathrm{a}}$, $\mathbf{H}^{\mathrm{a}}$ and $\mathbf{E}^{\mathrm{b}}$, $\mathbf{H}^{\mathrm{b}}$ in the same region, respectively. Find the relation between these EM fields.

1.2 Suppose a magnetic current $\hat{e}_\theta\delta(\mathbf{r} - \mathbf{r}_a)$ generates the magnetic field on the surface of the metallic target as

$$\mathbf{H} = -jk\frac{e^{-jkr}}{4\pi r}\mathbf{H}^{PS}$$

where $\mathbf{H}^{PS}$ is the scattered magnetic field by the target under the incident plane wave with unit amplitude of the electric field. Now we have a magnetic current $\mathbf{M}$ on the surface of the metallic target. Prove that the θ-component of the magnetic field

generated by the magnetic current $\mathbf{M}$ at the position of $\mathbf{r}_a$ is

$$H_\theta = -jk\frac{e^{-jkr}}{4\pi r}\int_s \mathbf{M}\cdot\mathbf{H}^{PS}dS.$$

1.2 Vector Wave Equation of the Electromagnetic Fields

The EM equations in the preceding discussion are systems of equations made up of two equations, with (1.9) and (1.10) expressed in the time domain, and (1.20) and (1.21) specified in the frequency domain. They all comprise two unknown vectors $\mathbf{E}$ and $\mathbf{H}$. In this section, we aim at eliminating one of them, and convert the system of equations into one equation. Take (1.20) and (1.21) in the frequency domain as an example. In order to eliminate $\mathbf{H}$ in (1.20), we proceed as follows.

First, divide both sides of (1.20) by z before applying the curl operation on them. Then, we substitute (1.21) into the resultant equation to produce:

$$\nabla\times\left(\frac{1}{z}\nabla\times\mathbf{E}\right)+y\mathbf{E}=-\mathbf{J}^i-\nabla\times\left(\frac{1}{z}\mathbf{M}^i\right) \tag{1.31}$$

Similarly, we can obtain

$$\nabla\times\left(\frac{1}{y}\nabla\times\mathbf{H}\right)+z\mathbf{H}=-\mathbf{M}^i+\nabla\times\left(\frac{1}{y}\mathbf{J}^i\right) \tag{1.32}$$

Equations (1.31) and (1.32) are the *vector EM wave equations*. In a homogeneous medium, (1.31) and (1.32) reduce to:

$$\nabla\times(\nabla\times\mathbf{E})-k^2\mathbf{E}=-z\mathbf{J}^i-\nabla\times\mathbf{M}^i \tag{1.33}$$

$$\nabla\times(\nabla\times\mathbf{H})-k^2\mathbf{H}=-y\mathbf{M}^i+\nabla\times\mathbf{J}^i \tag{1.34}$$

where $k=\sqrt{-zy}$ is the *wave number* in the medium.

Problems

1.3 Suppose the EM wave propagates in the chiral medium, whose constitutive relation is

$$\begin{aligned}\mathbf{D} &= \varepsilon\mathbf{E}+\xi\mathbf{H}\\ \mathbf{B} &= \varsigma\mathbf{E}+\mu\mathbf{H}\end{aligned}$$

Find the vector EM wave equation in this medium.

1.3 Vector Integral Equation of the Electromagnetic Fields

The EM integral equations (IEs) are diverse for different problems. However, they can be derived by a general approach. That is, by using the equivalence principle, we

separate the original problem into two parts, the regular one and the irregular one. The regular part can be solved analytically, and the solution can then be used in deriving the integral equation of the unknown equivalent source introduced in the irregular part. In the following we will take the integral equation of scattering problem in free space as an example to illustrate the approach to establish the EM IEs. First, we will introduce the equivalence principle, which is a theoretical basis to support the aforementioned separation. Second, we will discuss the EM fields generated by a source in free space, that is, the solution of the Maxwell's equation with sources in free space. Finally, we take three kinds of scatterers as examples and enumerate the respective approaches to derive their IEs. Specifically, the three kinds of scatterers considered are the metallic body, homogeneous dielectric body and inhomogeneous dielectric body.

1.3.1 Equivalence Principle

The *equivalence principle* is derived from the uniqueness theorem [5]. One of its important utilities is that it tells us how the problem can be transformed while keeping equivalence with the original problem. Various forms of the equivalence principle exist. Here, we concentrate on three widely used forms. Consider a situation shown in Figure 1.1a. S is a hypothetical closed surface with sources and inhomogeneous material inside. Outside S is the source-free homogeneous region. Obviously the sources and the heterogeneity enclosed in S make the field calculation challenging. However, if we only consider the field outside S, then this problem is equivalent to a regular problem with sources on S for which a solution may be written straightforwardly. Being equivalent means that these problems dictate the same EM fields in the domain of interest. The original problem is illustrated in Figure 1.1a, and the three equivalent schemes are shown in Figures 1.1b–d respectively:

In the surface equivalent problem illustrated in Figure 1.1b, the source inside S is removed and the field inside S is zero everywhere. Additionally, we postulate electric and magnetic surface currents on the surface of S, with current densities $\mathbf{J}_s$ and $\mathbf{M}_s$ given by

$$\mathbf{M}_s = \mathbf{E} \times \mathbf{n} \tag{1.35}$$

$$\mathbf{J}_s = \mathbf{n} \times \mathbf{H} \tag{1.36}$$

Then, from the continuity condition on the boundary, we know that the tangential EM fields on boundary S equals that of the original problem. According to the uniqueness theorem, the solution of this problem outside S is the same as that of the original problem. More to the point, since the fields inside S are null everywhere, we can further suppose the material inside S is homogeneous and is the same as the material outside S. In this way, the original problem is equivalent to a regular problem comprised of a homogeneous medium and a set of sources $\mathbf{J}_s$ and $\mathbf{M}_s$. We are able to provide an

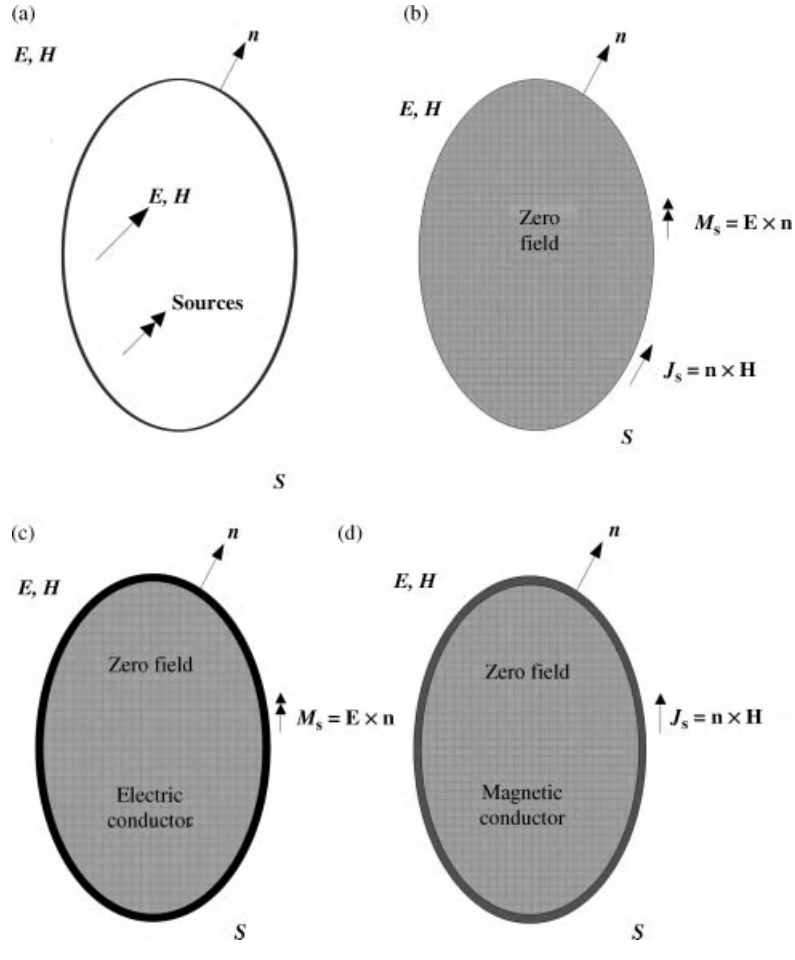


Figure 1.1 Illustration of the surface equivalence principle.

analytical solution to this problem, which will be presented in Section 1.3.2. This popular equivalence scheme is also known as the *Huygens' principle*. It is worth noting that there must be simultaneous equivalent electric and magnetic currents. Otherwise, we cannot assure the consistency of the fields outside S with the original problem while keeping the fields inside S to be zero.

In contrast to the first scheme, the second one assumes the medium inside S to be a perfect electric conductor (PEC). Thus, the electric field inside S is zero, as shown in Figure 1.1c. Then, we only need to define equivalent magnetic current density $\mathbf{M}_s$ on S to assure the fields equivalence outside S:

$$\mathbf{M}_s = \mathbf{E} \times \mathbf{n} \tag{1.37}$$

Again, as known from the continuity condition on the boundary, the tangential field on S equals that of the original problem. According to the uniqueness theorem, the resultant EM fields outside S are the same as that of the original problem. In this way, the original problem is equivalent to the EM fields generated by the magnetic current density $\mathbf{M}_s$ on the surface of a PEC body. If we choose S as a regular shape such as a sphere, then the analytical solution can be obtained.

The third scheme is a counterpart of the second one. As can be seen in Figure 1.1d, we only need to replace the PEC with a perfect magnetic conductor (PMC), and the equivalent magnetic current density $\mathbf{M}_s$ with the equivalent electric current density $\mathbf{J}_s$ satisfying

$$\mathbf{J}_s = \mathbf{n} \times \mathbf{H} \tag{1.38}$$

Again, the continuity condition on the boundary assures that the tangential field on S equals that of the original problem. According to the uniqueness theorem, the EM fields solution outside S equals that of the original problem. Therefore, the original problem is equivalent to the EM fields generated by the electric current density $\mathbf{J}_s$ on the surface of a PMC body. For S having a canonical form, such as a sphere, we can obtain an analytical solution.

Problems

1.4 Devise examples for demonstrating the importance of the equivalence principle.

1.3.2 Solution of Maxwell's Equation in Free Space

The solution to Maxwell's equation with sources in free space can be obtained analytically [6]. In the following, we use an electric current as an example to illustrate how to introduce a vector potential and a scalar potential to assist solving the problem. In free space, the following equation holds:

$$\nabla \cdot \mathbf{H} = 0 \tag{1.39}$$

According to the vector identity $\nabla \cdot \nabla \times \mathbf{A} = 0$, we can define a *vector potential* $\mathbf{A}$ satisfying

$$\mathbf{H} = \nabla \times \mathbf{A} \tag{1.40}$$

By substituting (1.40) into (1.16) we get

$$\nabla \times (\mathbf{E} + j\omega\mu\mathbf{A}) = 0 \tag{1.41}$$

We know that the curl of the gradient of any scalar field is always the zero vector. Mathematically, this is expressed as the vector identity $\nabla \times \nabla\phi = 0$. So we can further

introduce a *scalar potential* ϕ in order to express $\mathbf{E}$:

$$\mathbf{E} + j\omega\mu\mathbf{A} = \nabla\phi \tag{1.42}$$

Substituting (1.40) and (1.42) into (1.17) we have:

$$\nabla \times \nabla \times \mathbf{A} - k^2\mathbf{A} = \mathbf{J} + j\omega\varepsilon\nabla\phi \tag{1.43}$$

where $k = \omega\sqrt{\mu\varepsilon}$. By applying the vector identity, we can rewrite (1.43) as

$$\nabla(\nabla \cdot \mathbf{A}) - \nabla^2\mathbf{A} - k^2\mathbf{A} = \mathbf{J} + j\omega\varepsilon\nabla\phi \tag{1.44}$$

Apparently $\mathbf{A}$ cannot be uniquely determined by (1.40). We need to further specify $\nabla \cdot \mathbf{A}$. For simplicity, in view of (1.44) we choose

$$\nabla \cdot \mathbf{A} = j\omega\varepsilon\phi \tag{1.45}$$

Thus, (1.44) becomes a vector differential equation comprising only vector potential $\mathbf{A}$

$$\nabla^2\mathbf{A} + k^2\mathbf{A} = \mathbf{J} \tag{1.46}$$

This is the *vector Helmholtz equation*. By applying the divergence operator to both sides of (1.43), and substituting (1.45), we arrive at the scalar Helmholtz equation describing the scalar potential:

$$\nabla^2\phi + k^2\phi = -\frac{1}{j\omega\varepsilon}\nabla \cdot \mathbf{J} \tag{1.47}$$

Up to here we can see the convenience of introducing $\mathbf{A}$. In certain orthogonal coordinate system, the vector Helmholtz equation can be converted into scalar Helmholtz equations. For example, in the Cartesian coordinate system, we have:

$$\nabla^2 A_x + k^2 A_x = -J_x \tag{1.48}$$

$$\nabla^2 A_y + k^2 A_y = -J_y \tag{1.49}$$

$$\nabla^2 A_z + k^2 A_z = -J_z \tag{1.50}$$

where A_x, A_y, A_z and J_x, J_y, J_z are the components of $\mathbf{A}$ and $\mathbf{J}$ in the Cartesian coordinate system. Next consider the well-known three-dimensional (3D) Green's function for solving the scalar Helmholtz equation

$$G(\mathbf{r}|\mathbf{r}') = \frac{e^{-jk|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|} \tag{1.51}$$

where r' and r denote the positions of the source point and the field point, respectively. Then, solution to the above scalar Helmholtz equation satisfying the radiation condition can be written as:

$$\mathbf{A}(\mathbf{r}) = -\int \mathbf{J}(\mathbf{r}')G(\mathbf{r}|\mathbf{r}')d\tau' \tag{1.52}$$

The scalar potential ϕ can be obtained readily through (1.45)

$$\phi(\mathbf{r}) = \frac{1}{j\omega\varepsilon}\nabla \cdot \mathbf{A}(\mathbf{r}) \tag{1.53}$$

It can also be obtained by solving (1.47):

$$\phi(\mathbf{r}) = -\frac{1}{j\omega\varepsilon}\int \nabla' \cdot \mathbf{J}(\mathbf{r}')G(\mathbf{r}|\mathbf{r}')d\tau' \tag{1.54}$$

Therefore, the electric field $\mathbf{E}$ can be obtained in two forms. One is by substituting (1.52) and (1.53) into (1.42) and getting

$$\mathbf{E} = -j\omega\mu\int\left[\mathbf{I} + \frac{1}{k^2}\nabla\nabla\cdot\right]\mathbf{J}Gd\tau' \tag{1.55}$$

where $\mathbf{I}$ is the identity matrix. The other is by substituting (1.52) and (1.54) into (1.42) and obtain

$$\mathbf{E} = -j\omega\mu\int\left[\mathbf{J} + \frac{1}{k^2}\nabla(\nabla'\cdot\mathbf{J})\right]Gd\tau' \tag{1.56}$$

Note the difference between these two forms. In the former, both the divergence operator $\nabla\cdot$ and the gradient operator ∇ operate on the field point $\mathbf{r}$. In other words they are both imposed on the Green's function G, which results in a higher order of singularity in the integrand. However, because the equivalent source current is not involved, in some problems such as the far-field calculation, it leads to a concise expression. Therefore, this form of $\mathbf{E}$ is usually used in the far-field computation. In the latter form the divergence operator $\nabla\cdot$ operates on the field point $\mathbf{r}$ and the gradient operator ∇ operates on the source point $\mathbf{r}'$, respectively. In other words, the divergence operator is imposed onto the equivalent source current, while the gradient operator ∇ is imposed onto the Green's function G. As a consequence, the order of singularity in the integrand is lower than the former, and this form is often used in near-field calculation. Substitution of (1.52) into (1.40) yields

$$\mathbf{H} = -\int \mathbf{J}\times\nabla Gd\tau' \tag{1.57}$$

For a compact notation, we can introduce two integrodifferential operators **L** and **K**, defined as

$$\mathbf{L}(\mathbf{X}) = -jk\int\left[\mathbf{X} + \frac{1}{k^2}\nabla(\nabla'\cdot\mathbf{X})\right]Gd\tau' \tag{1.58}$$

$$\mathbf{K}(\mathbf{X}) = -\int\mathbf{X}\times\nabla Gd\tau' \tag{1.59}$$

Thus, the electric field **E** and the magnetic field **H** can be expressed as

$$\mathbf{E} = Z\mathbf{L}(\mathbf{J}) \tag{1.60}$$

$$\mathbf{H} = \mathbf{K}(\mathbf{J}) \tag{1.61}$$

where $Z = \sqrt{\mu/\varepsilon}$ is the wave impedance. By applying the duality of EM fields, the electric and magnetic fields generated by the equivalent current can be obtained:

$$\mathbf{E} = -\mathbf{K}(\mathbf{M}) \tag{1.62}$$

$$\mathbf{H} = \frac{1}{Z}\mathbf{L}(\mathbf{M}) \tag{1.63}$$

Utilizing the principle of linear superposition, the final EM fields caused by the electric and magnetic equivalent currents can be written as

$$\mathbf{E} = Z\mathbf{L}(\mathbf{J}) - \mathbf{K}(\mathbf{M}) \tag{1.64}$$

$$\mathbf{H} = \frac{1}{Z}\mathbf{L}(\mathbf{M}) + \mathbf{K}(\mathbf{J}) \tag{1.65}$$

Problems

1.5 Use vector identities to prove that

$$\int\left[\mathbf{I} + \frac{1}{k^2}\nabla\nabla\cdot\right]\mathbf{J}Gd\tau' = \int\left[\mathbf{J} + \frac{1}{k^2}\nabla(\nabla'\cdot\mathbf{J})\right]Gd\tau'.$$

1.3.3 Integral Equations of Metallic Scattering Problems

Consider an EM wave $\mathbf{E}^i, \mathbf{H}^i$ impinging on a metallic body with surface S. A scattered field will be generated by this metallic body. In this section, we will introduce the method to construct the integral equation describing the scattered fields. We can apply

the first scheme of the aforementioned surface equivalence principle on S. That is, the scattered fields outside S is equal to the fields in a homogeneous medium generated by the equivalent surface currents on S:

$$\mathbf{M} = \mathbf{E} \times \hat{n} \tag{1.66}$$

$$\mathbf{J} = \hat{n} \times \mathbf{H} \tag{1.67}$$

Since the tangential electric field on the metallic surface equals zero, the equivalent magnetic current is zero. Therefore, the scattered field is expressed solely by $\mathbf{J}$

$$\mathbf{E}^s = Z\mathbf{L}(\mathbf{J}) \tag{1.68}$$

$$\mathbf{H}^s = \mathbf{K}(\mathbf{J}) \tag{1.69}$$

The superposition of the incident and scattered fields yields the fields in the presence of the scatterer:

$$\mathbf{E} = \mathbf{E}^i + Z\mathbf{L}(\mathbf{J}) \tag{1.70}$$

$$\mathbf{H} = \mathbf{H}^i + \mathbf{K}(\mathbf{J}) \tag{1.71}$$

The tangential electric field on the metallic surface equals zero, yielding

$$\left[\mathbf{E}^i + Z\mathbf{L}(\mathbf{J})\right]\Big|_t = 0 \tag{1.72}$$

where $|_t$ in (1.72) denotes the tangential components. Taking the cross product of $\mathbf{n}$ with both sides of (1.71), and then by applying (1.67) we obtain

$$\mathbf{J} - \hat{n} \times \mathbf{K}(\mathbf{J}) = \hat{n} \times \mathbf{H}^i \tag{1.73}$$

Equation (1.72) is often termed the *electric field integral equation* (EFIE), since it is constructed according to the boundary condition of the electric field. Equation (1.73) is often termed the *magnetic field integral equation* (MFIE), since it is constructed according to the boundary condition of the magnetic field. Theoretically, the EFIE and the MFIE are equivalent to each other. However, they are significantly different from the computational point of view. In the EFIE (1.72), the unknown equivalent electrical current $\mathbf{J}$ appears only in the integral operator $\mathbf{L}$, while in the MFIE (1.73), $\mathbf{J}$ appears both inside and outside the integral operator $\mathbf{K}$. Consequently, mathematically, they are of different integration types. The EFIE belongs to the *Fredholm integral equation of the first kind*, while the MFIE belongs to the *Fredholm integral equation of the second kind.* We will discuss the difference on their numerical behavior in Chapter 2.

1.3.4 Integral Equation of Homogeneous Dielectric Scattering Problems

In this section we will discuss the approach to construct the integral equation describing the scattered field from a homogeneous dielectric scatterer. Consider the following situation: in a homogeneous background medium having constitutive parameters of ε_1 and μ_1, a scatterer with ε_2 and μ_2 is illuminated by the EM wave $\mathbf{E}^i, \mathbf{H}^i$. Denote the surface of the dielectric body by S and the EM fields outside S by $\mathbf{E}_1$ and $\mathbf{H}_1$. To apply the first scheme of the surface equivalent principle on S, we first obtain the equivalent currents on S:

$$\mathbf{M}_1 = \mathbf{E}_1 \times \hat{n}_1 \tag{1.74}$$

$$\mathbf{J}_1 = \hat{n}_1 \times \mathbf{H}_1 \tag{1.75}$$

where $\hat{n}_1$ is the outward normal unit vector. Thus, the scattered field outside S can be obtained by applying equivalent currents on S in the background ε_1, μ_1:

$$\mathbf{E}^s = Z_1 \mathbf{L}_1(\mathbf{J}_1) - \mathbf{K}_1(\mathbf{M}_1) \tag{1.76}$$

$$\mathbf{H}^s = \frac{1}{Z_1}\mathbf{L}_1(\mathbf{M}_1) + \mathbf{K}_1(\mathbf{J}_1) \tag{1.77}$$

where $Z_1 = \sqrt{\mu_1/\varepsilon_1}$. The total fields equal to the superposition of the incident and scattered fields dictates:

$$\mathbf{E}_1 - Z_1 \mathbf{L}_1(\mathbf{J}_1) + \mathbf{K}_1(\mathbf{M}_1) = \mathbf{E}^i \tag{1.78}$$

$$\mathbf{H}_1 - \frac{1}{Z_1}\mathbf{L}_1(\mathbf{M}_1) - \mathbf{K}_1(\mathbf{J}_1) = \mathbf{H}^i \tag{1.79}$$

Taking the cross product of $\hat{n}_1$ with both sides of (1.78), (1.79), and applying (1.74), (1.75) produces

$$-\mathbf{M}_1 - \hat{n}_1 \times [Z_1 \mathbf{L}_1(\mathbf{J}_1) - \mathbf{K}_1(\mathbf{M}_1)] = \hat{n}_1 \times \mathbf{E}^i \tag{1.80}$$

$$\mathbf{J}_1 - \hat{n}_1 \times \left[\frac{1}{Z_1}\mathbf{L}_1(\mathbf{M}_1) + \mathbf{K}_1(\mathbf{J}_1)\right] = \hat{n}_1 \times \mathbf{H}^i \tag{1.81}$$

Equation (1.80) is constructed according to the electric field outside the scatterer. Accordingly, it is termed the *exterior electric field integral equation*. Likewise (1.81) is termed the *exterior magnetic field integral equation*. These two equations are used to describe the EM fields outside the scatterer. Theoretically, they are equivalent instead of independent. Since both equations contain two unknowns $\mathbf{J}_1$ and $\mathbf{M}_1$, we still

need one more independent equation to uniquely determine $\mathbf{J}_1$ and $\mathbf{M}_1$. This equation can be obtained through the interior EM relations. Since the scatterer is homogeneous, the interior EM field can also be obtained through postulating the scenario of the homogeneous medium ε_2, μ_2 with a group of equivalent current sources on S satisfying

$$\mathbf{M}_2 = \mathbf{E}_2 \times \hat{n}_2 \tag{1.82}$$

$$\mathbf{J}_2 = \hat{n}_2 \times \mathbf{H}_2 \tag{1.83}$$

where $\mathbf{E}_2$ and $\mathbf{H}_2$ are the interior EM field on surface S, and $\hat{n}_2$ is the inward normal unit vector. Under this postulation, and due to the source-free condition inside the dielectric body, the interior EM fields can be expressed as

$$\mathbf{E}_2 = Z_2\mathbf{L}_2(\mathbf{J}_2) - \mathbf{K}_2(\mathbf{M}_2) \tag{1.84}$$

$$\mathbf{H}_2 = \frac{1}{Z_2}\mathbf{L}_2(\mathbf{M}_2) + \mathbf{K}_2(\mathbf{J}_2) \tag{1.85}$$

where $Z_2 = \sqrt{\mu_2/\varepsilon_2}$. Taking the cross product of $\hat{n}_2$ on both sides of (1.84), (1.85), and applying (1.82), (1.83) produce

$$-\mathbf{M}_2 = \hat{n}_2 \times [Z_2\mathbf{L}_2(\mathbf{J}_2) - \mathbf{K}_2(\mathbf{M}_2)] \tag{1.86}$$

$$\mathbf{J}_2 = \hat{n}_2 \times \left[\frac{1}{Z_2}\mathbf{L}_2(\mathbf{M}_2) + \mathbf{K}_2(\mathbf{J}_2)\right] \tag{1.87}$$

These two equations are termed the *interior electric field integral equation* and the *interior magnetic field integral equation*, respectively. These two equations are used to describe the EM fields inside the dielectric scatterer, and they are equivalent. According to the continuity condition on the boundary of the two media, $\mathbf{J}_1 = -\mathbf{J}_2$, $\mathbf{M}_1 = -\mathbf{M}_2$. For the convenience of notation, we denote $\mathbf{J}_1 = \mathbf{J}$, $\mathbf{M}_1 = \mathbf{M}$, $\mathbf{E}_1 = \mathbf{E}$, $\mathbf{H}_1 = \mathbf{H}$, and $\hat{n}_1 = \hat{n}$. Thus, we can write the exterior EFIE and MFIE as

$$\mathbf{M} + \hat{n} \times [Z_1\mathbf{L}_1(\mathbf{J}) - \mathbf{K}_1(\mathbf{M})] = -\hat{n} \times \mathbf{E}^i \tag{1.88}$$

$$\mathbf{J} - \hat{n} \times \left[\frac{1}{Z_1}\mathbf{L}_1(\mathbf{M}) + \mathbf{K}_1(\mathbf{J})\right] = \hat{n} \times \mathbf{H}^i \tag{1.89}$$

The interior EFIE and MFIE are written as

$$\mathbf{M} = -\hat{n} \times [Z_2\mathbf{L}_2(-\mathbf{J}) - \mathbf{K}_2(-\mathbf{M})] \tag{1.90}$$

$$\mathbf{J} = \hat{n} \times \left[\frac{1}{Z_2}\mathbf{L}_2(-\mathbf{M}) + \mathbf{K}_2(-\mathbf{J})\right] \tag{1.91}$$

Obviously, one from the exterior equations (1.88) and (1.89), in concert with another from the interior equations (1.90) and (1.91), can form the system of equations to determine the unknowns $\mathbf{J}$ and $\mathbf{M}$.

However, the above system of equations encounters problems in the practical computation. We will discuss that in Chapter 2. Here, we will give another system of equations widely used in computation. We can simply subtract (1.88) from (1.90), and subtract (1.89) from (1.91) to obtain

$$-(Z_1\mathbf{L}_1 + Z_2\mathbf{L}_2)(\mathbf{J}) + (\mathbf{K}_1 + \mathbf{K}_2)(\mathbf{M}) = \mathbf{E}^i \tag{1.92}$$

$$-(\mathbf{K}_1 + \mathbf{K}_2)(\mathbf{J}) - \left(\frac{1}{Z_1}\mathbf{L}_1 + \frac{1}{Z_2}\mathbf{L}_2\right)(\mathbf{M}) = \mathbf{H}^i \tag{1.93}$$

In contrast to the previous system of equations, each of the equations in this system is made up of both the interior and the exterior IEs. In Chapter 2, we will discuss in detail the reason why this system of equation bears better numerical characteristics. This system of equation is proposed by Poggio, Miller, Chang, Harrington and Wu together, and is often abbreviated as the *PMCHW equation*.

All the above integral equations are provided in the form of systems of equations made up of two independent equations. In the following we will introduce how to use a single integral equation to determine the EM fields scattered by a homogeneous dielectric scatterer. Let us introduce a different equivalent scheme. Choose the boundary of a scatterer as the surface in the equivalence principle. Postulate the following equivalent model: the EM fields in the interior region are the same as the original one, the electric fields in the exterior region are chosen such that the tangential electric field be continuous on both sides of the boundary, and the exterior medium is the same as the scatterer everywhere. The continuity of the tangential electric fields on the boundary makes the equivalent magnetic current disappeard. Thus, the interior EM fields can be expressed by the single unknown, that is, the equivalent electric current $\mathbf{J}_e$:

$$\mathbf{E}^s = Z_2\mathbf{L}_2(\mathbf{J}_e) \tag{1.94}$$

$$\mathbf{H}^s = \mathbf{K}_2(\mathbf{J}_e) \tag{1.95}$$

Due to the absence of the incident EM fields in the interior region, the total interior fields are given by

$$\mathbf{E}_2 = Z_2\mathbf{L}_2(\mathbf{J}_e) \tag{1.96}$$

$$\mathbf{H}_2 = \mathbf{K}_2(\mathbf{J}_e) \tag{1.97}$$

Through (1.96) and (1.97) the original E- and H- fields on the boundary are expressed by using the single unknown variable $\mathbf{J}_e$. Then, follow the idea of the first equivalence

scheme, we can arrive at (1.80) and (1.81), in which the equivalent currents can be expressed by incorporating (1.96) and (1.97) as

$$\mathbf{M}_1 = \mathbf{E}_1 \times \mathbf{n}_1 = \mathbf{E}_2 \times \mathbf{n}_1 = Z_2 \mathbf{L}_2(\mathbf{J}_e) \times \mathbf{n}_1 \tag{1.98}$$

$$\mathbf{J}_1 = \mathbf{n}_1 \times \mathbf{H}_1 = \mathbf{n}_1 \times \mathbf{H}_2 = \mathbf{n}_1 \times \mathbf{K}_2(\mathbf{J}_e) \tag{1.99}$$

Substituting (1.98) and (1.99) into (1.80) and (1.81) produces

$$Z_2 \mathbf{L}_2(\mathbf{J}_e) - Z_1 \mathbf{L}_1(\mathbf{n}_1 \times \mathbf{K}_2(\mathbf{J}_e)) + \mathbf{K}_1(Z_2 \mathbf{L}_2(\mathbf{J}_e) \times \mathbf{n}_1) = \mathbf{E}^i \tag{1.100}$$

$$\mathbf{K}_2(\mathbf{J}_e) - \frac{1}{Z_1} \mathbf{L}_1(Z_2 \mathbf{L}_2(\mathbf{J}_e) \times \mathbf{n}_1) - \mathbf{K}_1(\mathbf{n}_1 \times \mathbf{K}_2(\mathbf{J}_e)) = \mathbf{H}^i \tag{1.101}$$

Equation (1.100) is equivalent to (1.101). Choosing either one can lead to the solution of the only unknown $\mathbf{J}_e$. After $\mathbf{J}_e$ is determined, we can obtain the far-field by using (1.94) and (1.95). Let's take a look at this form of the integral equation. Although the number of equation reduces to one, we have to pay an extra price, that is, the complexity of the equation is increased. To be specific, the operators operate on the equivalent current $\mathbf{J}_e$ twice.

We can conveniently call (1.100) the *single electric field integral equation (SEFIE)* for a homogeneous medium, and (1.101) the *single magnetic field integral equation (SMFIE)* for a homogeneous medium. Their numerical properties are quite different. Because in SEFIE the operators imposed on $\mathbf{J}_e$ are different, the condition number of this kind of equations is worse than that from single operator equations. On the other hand, in SMFIE the operators imposed on $\mathbf{J}_e$ are the same, which leads to a better condition number than the single operator equations. We will show numerical examples of this comparison in Section 2.2.4.

1.3.5 Integral Equation of Inhomogeneous Dielectric Scattering Problems

In contrast to the surface equivalent currents introduced for the metallic or the homogeneous dielectric scatterers, the equivalent volume currents need to be defined for the inhomogeneous dielectric scattering problem. Consequently, the integral equations are *volume integration equations (VIEs)*. Given the position dependent dielectric constant $\varepsilon(r)$ and the conductivity $\sigma(r)$ of the inhomogeneous scatterer, the induced electric current density on the scatterer is:

$$\mathbf{J}_{eq}(\mathbf{r}) = [\sigma(\mathbf{r}) + j\omega(\varepsilon(\mathbf{r}) - \varepsilon_0)]\mathbf{E}(\mathbf{r}) = \tau(\mathbf{r})\mathbf{E}(\mathbf{r}) \tag{1.102}$$

According to Maxwell's equations, this current density can be written as

$$\mathbf{J}_{eq}(\mathbf{r}) = \frac{\tau}{j\omega\varepsilon + \sigma} \nabla \times \mathbf{H} \tag{1.103}$$

The fields generated by this induced electric current are

$$\mathbf{E}^s = Z\mathbf{L}(\mathbf{J}_{eq}) \tag{1.104}$$

$$\mathbf{H}^s = \mathbf{K}(\mathbf{J}_{eq}) \tag{1.105}$$

where the operators $\mathbf{L}$ and $\mathbf{K}$ are of similar forms to the previous ones, only the integration is operated on the volume. The fact that the total field equals the superposition of the incident and scattered fields dictates

$$\mathbf{E} = Z\mathbf{L}(\mathbf{J}_{eq}) + \mathbf{E}^i \tag{1.106}$$

and

$$\mathbf{H} = \mathbf{K}(\mathbf{J}_{eq}) + \mathbf{H}^i \tag{1.107}$$

Plugging (1.102) into (1.106) yields the *electric field volume integral equation (EFVIE)*:

$$\mathbf{E} = Z\mathbf{L}(\tau\mathbf{E}) + \mathbf{E}^i \tag{1.108}$$

Substituting (1.103) into (1.107) yields the *magnetic field volume integral equation (MFVIE)*:

$$\mathbf{H} = \mathbf{K}\left(\frac{\tau}{j\omega\varepsilon + \sigma} \nabla \times \mathbf{H}\right) + \mathbf{H}^i \tag{1.109}$$

Note that the above volume integral equations for inhomogeneous scatterer are different from the surface integral equations for metallic scatterer. Both EFVIE of (1.108) and MFVIE of (1.109) are Fredholm integral equations of the second kind.

1.3.6 Integral Equations of Scattering in Layered Medium

Many practical problems, such as the vehicles on the ground, ships on the sea, buried objects under the ground can all be categorized as scattering problems in a layered medium. Solving this kind of problems is fairly complex and challenging. Similar to the scattering problem in free space, the scatterer with regular shapes can be solved analytically. As for general cases, a numerical method is the only way to compute scattering by arbitrary targets in layered medium. The analytical solution can generally be applied in the following procedure: first, calculate the reflection and refraction on

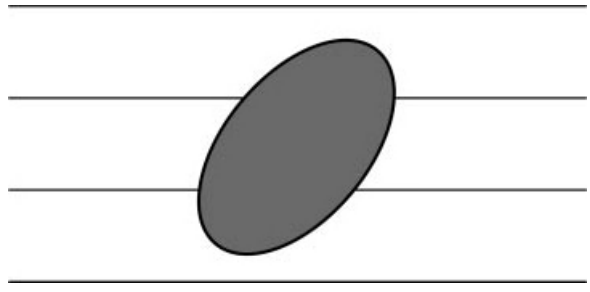


Figure 1.2 The PEC scatterer in a layered medium.

the interface of different layers; then calculate the scattering in the homogeneous medium; finally, consider their interaction. Due to the scope of this text, we are not going to present details for analytical solution. What we desire to address here is the formulation for the scattering problems in layered medium, which can be implemented by numerical methods. In contrast to the problem in free space, this problem involves the use of a more complicated and more intractable Green's function [7].

Consider the scattering by the metallic object in a layered medium as sketched in Figure 1.2. As in a homogeneous space, we first construct the integral equations involving the equivalent electric current. To this end, we need to obtain the Green's function for the layered medium. Represent the Maxwell's equations in the frequency domain as

$$\nabla \times \mathbf{E} = -\mathrm{j}\omega\mu\mathbf{H} \tag{1.110}$$

$$\nabla \times \mathbf{H} = \mathrm{j}\omega\varepsilon\mathbf{E} + \mathbf{J} \tag{1.111}$$

Since the background medium is homogeneous in both x and y directions, we introduce the following 2D Fourier transformation on (1.110) and (1.111), and obtain an equation with derivative only operating on z:

$$\tilde{f}\left(\mathbf{k}_\rho; z\right) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(\mathbf{r})\, e^{j\mathbf{k}_\rho \cdot \boldsymbol{\rho}} dxdy \tag{1.112}$$

The corresponding 2D inverse Fourier transformation is

$$f(\mathbf{r}) = \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \tilde{f}\left(\mathbf{k}_\rho; z\right) e^{-j\mathbf{k}_\rho \cdot \boldsymbol{\rho}} dk_x dk_y \tag{1.113}$$

By using the following plane wave integral representation of Bessel function:

$$J_n(\rho) = \frac{j^n}{2\pi}\int_0^{2\pi} e^{-j\rho\cos\phi} e^{-jn\phi} d\phi \tag{1.114}$$

(1.113) can be further transformed into a 1D integral form:

$$f(\mathbf{r}) = \frac{1}{2\pi}\int_0^\infty \tilde{f}(k_\rho) J_0(k_\rho \rho) k_\rho dk_\rho \tag{1.115}$$

Equation (1.115) is often referred to as the *Sommerfeld integration*, denoted by $\mathbb{S}_0$. Decompose the operator ∇ as $\nabla = \nabla_t + \nabla_z$, and substitute it into (1.110) and (1.111). Then upon using $\nabla_t = -j\mathbf{k}_\rho$, we obtain

$$\frac{d\tilde{\mathbf{E}}_t}{dz} = \frac{1}{j\omega\varepsilon}\left(k^2 - \mathbf{k}_\rho \mathbf{k}_\rho \cdot\right)\left(\tilde{\mathbf{H}}_t \times \hat{e}_z\right) + \frac{\mathbf{k}_\rho}{\omega\varepsilon}\tilde{J}_z \tag{1.116}$$

$$\frac{d\tilde{\mathbf{H}}_t}{dz} = \frac{1}{j\omega\mu}\left(k^2 - \mathbf{k}_\rho \mathbf{k}_\rho \cdot\right)\left(\hat{e}_z \times \tilde{\mathbf{E}}_t\right) - \hat{e}_z \times \tilde{\boldsymbol{J}}_t \tag{1.117}$$

$$\tilde{E}_z = -\frac{1}{j\omega\varepsilon}\left[j\mathbf{k}_\rho \cdot \left(\tilde{\mathbf{H}}_t \times \hat{e}_z\right) - \tilde{J}_z\right] \tag{1.118}$$

$$\tilde{H}_z = -\frac{1}{j\omega\mu}\left[j\mathbf{k}_\rho \cdot \left(\hat{e}_z \times \tilde{\mathbf{E}}_t\right)\right] \tag{1.119}$$

where $k^2 = \omega^2\mu\varepsilon$. Rotate the coordinate system from $(\hat{k}_x, \hat{k}_y)$ to $(\hat{\mathrm{k}}_\rho, \hat{e}_z \times \hat{\mathrm{k}}_\rho)$, denoted by $(\hat{e}_u, \hat{e}_v)$. Then, $\tilde{\mathbf{E}}_t$ and $\tilde{\mathbf{H}}_t$ can be decomposed in $(\hat{e}_u, \hat{e}_v)$ into

$$\tilde{\mathbf{E}}_t = \hat{u}\, V^e + \hat{v}\, V^h \tag{1.120}$$

$$\tilde{\mathbf{H}}_t \times \hat{e}_z = \hat{u}\, I^e + \hat{v}\, I^h \tag{1.121}$$

Substituting (1.120) and (1.121) into (1.116) and (1.117), we can find that (1.116) and (1.117) are equivalent to the following two independent decoupled transmission-line equations:

$$\frac{dV^p}{dz} = -jk_z Z^P I^p + v^P \tag{1.122}$$

$$\frac{dI^p}{dz} = -jk_z\, Y^P\, V^p + i^P \tag{1.123}$$

where the superscript p stands for e or h, and

$$k_z = \sqrt{k^2 - k_\rho^2} \tag{1.124}$$

$$Z^e = \frac{k_z}{\omega\varepsilon} \tag{1.125}$$

$$Z^h = \frac{\omega\mu}{k_z} \tag{1.126}$$

The source terms v^p, i^p are

$$\begin{array}{lll} v^e & = & \dfrac{k_\rho}{\omega\varepsilon}\tilde{J}_z, \qquad i^e = -\tilde{J}_u \\ v^h & = & 0, \qquad\qquad i^h = -\tilde{J}_v \end{array} \tag{1.127}$$

By using (1.118)–(1.121), the EM field in the frequency domain can be expressed as

$$\tilde{\mathbf{E}} = \hat{u}V^e + \hat{v}V^h - \hat{e}_z \frac{1}{j\omega\varepsilon}\left(jk_\rho I^e + \tilde{J}_z\right) \tag{1.128}$$

$$\tilde{\mathbf{H}} = -\hat{u}I^h + \hat{v}I^e + \hat{e}_z \frac{k_\rho}{\omega\mu} V^h \tag{1.129}$$

where (V^e, I^e) and $\left(V^h, I^h\right)$ denote the TM and TE polarized waves with respect to the z direction, respectively. The EM fields in the spatial domain can be obtained by applying the inverse transformation (1.115) to expressions (1.128) and (1.129). Thus, the solving of the Green's function for the layered medium is converted into solving the Green's function for the transmission-line equations (1.122) and (1.123). To be specific, this involves solving for two kinds of Green's functions as illustrated in Figure 1.3: $V_i^p(z|z'), I_i^p(z|z')$ generated by the shunt electric current source $\delta(z - z')$ at z'; and $V_v^p(z|z'), I_v^p(z|z')$ generated by the series electric voltage source $\delta(z - z')$ at z', satisfying, respectively, the following equations

$$\frac{dV_i^p}{dz} = -jk_z Z^P I_i^p \tag{1.130}$$

$$\frac{dI_i^p}{dz} = -jk_z Y^P V_i^p + \delta(z - z') \tag{1.131}$$

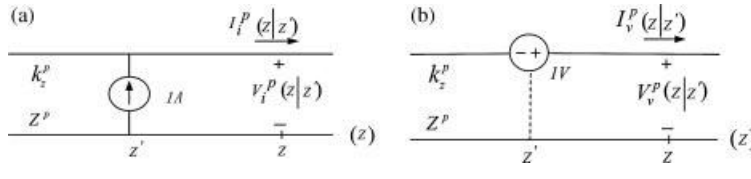


Figure 1.3 Green's functions for the transmission lines.

$$\frac{dV_v^p}{dz} = -jk_z Z^P I_v^p + \delta(z - z') \tag{1.132}$$

$$\frac{dI_v^p}{dz} = -jk_z Y^P V_v^p \tag{1.133}$$

where δ is the Dirac delta.

The reciprocity theorem allows us to write:

$$\begin{aligned} V_i^p(z|z') &= V_i^p(z'|z) \\ I_v^p(z|z') &= I_v^p(z'|z) \\ V_v^p(z|z') &= -I_i^p(z'|z) \\ I_i^p(z|z') &= -V_v^p(z'|z) \end{aligned} \tag{1.134}$$

From (1.122)–(1.123), and (1.128)–(1.133), one obtains the spectral domain expression of the EM fields in the layered medium generated by electric current source $\mathbf{J}$:

$$\tilde{\mathbf{E}} = \int_V \underline{\underline{\tilde{G}}}^e \mathbf{J}\, dV \tag{1.135}$$

$$\tilde{\mathbf{H}} = \int_V \underline{\underline{\tilde{G}}}^h \mathbf{J}\, dV \tag{1.136}$$

where $\underline{\underline{\tilde{G}}}^e$ and $\underline{\underline{\tilde{G}}}^h$ are the *dyadic Green's functions* in the layered medium:

$$\underline{\underline{\tilde{G}}}^e = \begin{bmatrix} -V_i^e & 0 & \dfrac{k_\rho}{\omega\varepsilon} V_v^e \\ 0 & -V_i^h & 0 \\ \dfrac{k_\rho}{\omega\varepsilon} I_i^e & 0 & \dfrac{1}{j\omega\varepsilon}\left[\dfrac{k_\rho^2}{j\omega\varepsilon} I_v^e - \delta(z - z')\right] \end{bmatrix} \tag{1.137}$$

$$\underline{\underline{\tilde{G}}}^h = \begin{bmatrix} 0 & I_i^h & 0 \\ -I_i^e & 0 & \dfrac{k_\rho}{\omega\varepsilon} I_v^e \\ 0 & -\dfrac{k_\rho}{\omega\mu} V_i^h & 0 \end{bmatrix} \tag{1.138}$$

Applying the inverse transformation (1.115) to the dyadic Green's functions in spectral domain (1.137) and (1.138), we get the dyadic Green's functions in spatial domain. However, (1.137) contains the k_ρ^2 term, whose inverse Fourier transformation produces difficulty in obtaining numerical solutions. To circumvent this problem, we need to find another type of EM fields expression equivalent to

(1.135) – *mixed-potentials expression*. As in free space, we consider only the electric currents and introduce the vector potential $\mathbf{A}$ and scalar potential ϕ satisfying

$$\mathbf{H} = \nabla \times \mathbf{A} \tag{1.139}$$

$$\mathbf{E} = -j\omega\mu\mathbf{A} - \nabla\phi \tag{1.140}$$

and

$$\mathbf{A} = \int_{V'} \underline{\underline{G}}^A \mathbf{J}\, dV' \tag{1.141}$$

Thus, according to (1.136), (1.139) and (1.141), we obtain

$$\underline{\underline{G}}^h = \nabla \times \underline{\underline{G}}^A \tag{1.142}$$

Note that $\underline{\underline{G}}^A$ satisfying (1.142) is not unique. Therefore, we can choose the following form of $\underline{\underline{G}}^A$ in the spectral domain in the coordinate system (u, v, z)

$$\tilde{\underline{\underline{G}}}^A = \begin{bmatrix} \tilde{G}^A_{vv} & 0 & 0 \\ 0 & \tilde{G}^A_{vv} & 0 \\ \tilde{G}^A_{zu} & 0 & \tilde{G}^A_{zz} \end{bmatrix} \tag{1.143}$$

Expressing (1.142) in the spectral domain and substituting (1.143) into it yields

$$\begin{aligned} \tilde{\nabla}\times \tilde{\underline{\underline{G}}}^A &= \left(-jk_\rho \hat{e}_u + \frac{d}{dz}\hat{e}_z\right) \times \begin{bmatrix} \tilde{G}^A_{vv} & 0 & 0 \\ 0 & \tilde{G}^A_{vv} & 0 \\ \tilde{G}^A_{zu} & 0 & \tilde{G}^A_{zz} \end{bmatrix} \\ &= \begin{bmatrix} 0 & -\dfrac{d\tilde{G}^A_{vv}}{dz} & 0 \\ jk_\rho \tilde{G}^A_{zu} + \dfrac{d\tilde{G}^A_{vv}}{dz} & 0 & jk_\rho \tilde{G}^A_{zz} \\ 0 & -jk_\rho \tilde{G}^A_{vv} & 0 \end{bmatrix} \end{aligned} \tag{1.144}$$

By comparing (1.138) with (1.144) we get

$$\tilde{G}^A_{vv} = \frac{1}{j\omega\mu} V^h_i \tag{1.145}$$

$$\tilde{G}^A_{zz} = \frac{1}{j\omega\varepsilon} I^e_v \tag{1.146}$$

$$\tilde{G}_{zu}^{A} = \frac{1}{jk_\rho}\left(I_i^h - I_i^e\right) \tag{1.147}$$

In order to uniquely determine the vector potential $\mathbf{A}$, we need to further fix $\nabla \cdot \mathbf{A}$. Following conventions, we choose

$$\nabla \cdot \mathbf{A} = j\omega\varepsilon\phi \tag{1.148}$$

In order to determine the scalar potential ϕ, we first express it as

$$\phi = \frac{1}{j\omega\varepsilon}\int_{V'} \boldsymbol{\psi} \cdot \mathbf{J} dV' \tag{1.149}$$

Substituting (1.141) and (1.149) into (1.148) leads to

$$\nabla \cdot \underline{\underline{G}}^{A} = \boldsymbol{\psi}. \tag{1.150}$$

Expressing (1.150) in the spectral domain, and substituting (1.143) into it yields

$$\begin{aligned} \nabla \cdot \tilde{\underline{\underline{G}}}^{A} &= \left(-jk_\rho \hat{e}_u + \frac{d}{dz}\hat{e}_z\right) \cdot \begin{bmatrix} \tilde{G}_{vv}^{A} & 0 & 0 \\ 0 & \tilde{G}_{vv}^{A} & 0 \\ \tilde{G}_{zu}^{A} & 0 & \tilde{G}_{zz}^{A} \end{bmatrix} \\ &= \begin{pmatrix} -jk_\rho \tilde{G}_{vv}^{A} + \dfrac{d\tilde{G}_{zu}^{A}}{dz} & 0 & \dfrac{d\tilde{G}_{zz}^{A}}{dz} \end{pmatrix} \end{aligned} \tag{1.151}$$

Postulate $\tilde{\boldsymbol{\psi}}$ in the following form:

$$\tilde{\boldsymbol{\psi}} = -\nabla' \tilde{K} + \tilde{C}\,\hat{e}_z \tag{1.152}$$

where $\tilde{K}$ is the scalar potential kernel and $\tilde{C}$ is the correction factor. In order to find $\tilde{K}$ and $\tilde{C}$, we substitute (1.151) and (1.152) into (1.150). Noting $\nabla' = j\mathbf{k}_\rho + \hat{z}d/dz'$, we get

$$\tilde{K} = \tilde{G}_{vv}^{A} - \frac{1}{jk_\rho}\frac{d\tilde{G}_{zu}^{A}}{dz} \tag{1.153}$$

$$\tilde{C} = \frac{d\tilde{G}_{zz}^{A}}{dz} + \frac{d\tilde{K}}{dz'} \tag{1.154}$$

By using (1.145)–(1.147), (1.152) can be written as

$$\tilde{K} = \frac{1}{j\omega\mu}V_i^h + \frac{1}{k_\rho^2}\frac{d}{dz}\left(I_i^h - I_i^e\right) \tag{1.155}$$

Then, substituting (1.131) into (1.155), and by using (1.124)–(1.126), we find

$$\tilde{K} = \frac{j\omega\varepsilon}{k_\rho^2}\left(V_i^e - V_i^h\right) \tag{1.156}$$

Substituting (1.146) and (1.156) into (1.154) leads to

$$\tilde{C} = \frac{1}{j\omega\varepsilon}\frac{d}{dz}I_v^e + \frac{j\omega\varepsilon}{k_\rho^2}\frac{d}{dz'}\left(V_i^e - V_i^h\right) \tag{1.157}$$

Finally, by using (1.130)–(1.134) we arrive at

$$\tilde{C} = \frac{k^2}{k_\rho^2}\left(V_v^h - V_v^e\right) \tag{1.158}$$

Thus far the vector- and scalar potentials $\tilde{\mathbf{A}}$ and $\tilde{\phi}$ in the spectral domain are obtained. The corresponding potentials $\mathbf{A}$ and ϕ in the spatial domain can be conveniently obtained by applying Sommerfeld integration to $\tilde{\mathbf{A}}$ and $\tilde{\phi}$. In the following, we will discuss more details of every entry in the Sommerfeld integration of $\tilde{\mathbf{A}}$ and $\tilde{\phi}$. Since Cartesian coordinate is usually used in the spatial domain, we transform the frequency-domain expression of $\tilde{\mathbf{A}}$ with coordinates (u, v, z) into the Cartesian coordinate system:

$$\underline{\underline{\tilde{G}}}^A = \begin{bmatrix} \tilde{G}_{vv}^A & 0 & 0 \\ 0 & \tilde{G}_{vv}^A & 0 \\ \cos\varphi\,\tilde{G}_{zu}^A & \sin\varphi\,\tilde{G}_{zu}^A & \tilde{G}_{zz}^A \end{bmatrix} \tag{1.159}$$

The diagonal entries in (1.159) are independent of φ. Therefore, their Sommerfeld integrations are all of zero order. On the other hand, the off-diagonal entries are function of φ. As a result, the Sommerfeld integrations of these entries can be converted into the first order form by using the following identity:

$$\mathbb{S}_0\left\{\begin{pmatrix}\sin n\varphi \\ \cos n\varphi\end{pmatrix}\tilde{f}\left(k_\rho\right)\right\} = (-j)^n\begin{pmatrix}\sin n\varphi \\ \cos n\varphi\end{pmatrix}\mathbb{S}_n\left\{\tilde{f}\left(k_\rho\right)\right\} \tag{1.160}$$

where $\mathbb{S}_n$ denotes the Sommerfeld integration of order n, expressed as

$$\mathbb{S}_n\left\{\tilde{f}\left(k_\rho\right)\right\} = \frac{1}{2\pi}\int_0^\infty \tilde{f}\left(k_\rho\right)J_n\left(k_\rho\rho\right)k_\rho^{n+1}dk_\rho \tag{1.161}$$

As for the entries in $\tilde{\phi}$, since they are all independent of φ, their Sommerfeld integrations are all of zero order.

Substituting (1.141) and the spatial domain counterpart of (1.152) into (1.140) we obtain the expression for electric field. Then, after performing some mathematical manipulations involving the Gauss' theorem, the desired mixed-potential representation

of E fields in the layered medium generated by electric current **J** is obtained:

$$\mathbf{E} = -j\omega\mu \int_V \underline{\underline{G}}^A \mathbf{J}\, dV + \frac{1}{j\omega\varepsilon} \nabla \left[\int_V K\, \nabla' \cdot \mathbf{J}\, dV + \int_V C\, J_z dV \right] \tag{1.162}$$

With the above expression, we can construct the integration equation about the equivalent electric current **J** on the surface of the PEC in exactly the same way as we do in free space. Thereby, we can use MoM to solve for the integral equations in a procedure analogous to the moment method for a free-space problem. The difference stands in the Sommerfeld integration involved in the matrix elements in the layered problem. If we calculate this integration directly without special treatment, this integration will be very time consuming. Concerning the special treatment of the Sommerfeld integration, the reader is referred to [8].

Problems

1.6 Prove the relation $\nabla_t = -j\mathbf{k}_\rho$ used in deriving (1.116)–(1.119).
1.7 Prove the reciprocity relations (1.134) for the transmission-line equations.

References

1. Senior, T.B.A. (1960) "Impedance boundary conditions for imperfectly conducting surface." *Applied Scientific Research, Section B*, **8**, 418–436.
2. Peterson, A.F., Ray, S.L., and Mittra, R. (1998) *Computational Methods for Electromagnetics*, IEEE Press, New York.
3. Ise, K., Inoue, K., and Koshiba, M. (1990) "Three-dimensional finite-element solution of dielectric scattering obstacles in a rectangular waveguide." *IEEE Transactions on Antennas and Propagation*, **38**(9), 1352–1359.
4. Stratton, J.A. (1941) *Electromagnetic Theory*, McGraw-Hill, New York.
5. Harrington, R.F. (1961) *Time-Harmonic Electromagnetic Fields*, McGraw-Hill, New York.
6. Tai, C.T. (1971) *Dyadic Green's Functions in Electromagnetic Theory*, International Textbook Company, Scratton, PA.
7. Michalski, K.A. and Mosig, J.R. (1997) "Multilayered media Green's functions in integral equation formulations." *IEEE Transactions on Antennas and Propagation*, **45**(3), 508–519.
8. Aksun, M.I. (1996) "A Robust Approach for the Derivation of Closed-form Green's Functions." *IEEE Transactions on Microwave Theory and Techniques*, **44**(5), 651–658.

2

Method of Moments

The *method of moments* (MoM) is a full-wave numerical method that discretizes the integral equations of electromagnetic fields. Since the integral equations satisfy the Sommerfeld radiation condition automatically , MoM is especially suitable for solving open problems, such as scattering and radiation problems. The development history of MoM in electromagnetics shows that MoM not only arose from solving open problems, but also becomes sophisticated in pursuing better solutions for this kind of problems. Therefore, to teach MoM by the case study of solving scattering problem is not only concise but also grasp the essentials of this method. As expected, plenty of content has been developed through research on MoM in the last 40 years. Even if only the scattering problem is considered, it is not possible to list all the relevant issues in one chapter. However, the essentials in MoM for solving any problems usually lie in the following four aspects: (1) choice of basis and testing functions; (2) handling of singularities; (3) identifying the relationship between the discretized forms of integral equations and their numerical performances; (4) accelerating the solution of the discretized matrix equations in MoM. The important progress of MoM in these four aspects is well embodied in solving scattering from a three-dimensional (3D) object. This chapter will discuss these four issues explicitly by studying EM scattering from a 3D object. The key points in solving other problems, such as scattering from two-dimensional (2D) objects, periodically arranged objects or axially symmetric objects, and radiation problems, will be provided at the end of this chapter. The readers can complete the solution in a similar manner or refer to [1,2].

2.1 Scattering from 3D PEC Objects

A sophisticated MoM solution of the EM scattering from a 3D perfect electric conductor (PEC) is rich in tricks for handling numerical difficulties as well as implementing fast algorithms. Let us introduce these difficulties and techniques by

Essentials of Computational Electromagnetics, First Edition. Xin-Qing Sheng and Wei Song.
© 2012 John Wiley & Sons Singapore Pte. Ltd. Published 2012 by John Wiley & Sons Singapore Pte. Ltd.

presenting an advanced MoM solution for this EM scattering problem in a complete and subtle way.

2.1.1 Formulation of the Problem

Radar is a target detection system that operates through transmitting electromagnetic waves and receiving the scattering waves from objects to determine the range and speed of the targets, for example, aircraft, ships, tanks and motor vehicles. Since many of these targets are mainly metallic, the associated scattering problems in the real world can be regarded as scattering from PEC objects. Ordinarily, the radar operates at a distance from the objects. As a consequence, usually the incident wave can be approximated by a plane wave, that is,

$$\mathbf{E}^i(\mathbf{r}) = (\cos\alpha\,\hat{\theta} + \sin\alpha\,\hat{\phi})e^{-j\mathbf{k}^i\cdot\mathbf{r}} \tag{2.1}$$

$$\mathbf{H}^i(\mathbf{r}) = \frac{1}{Z}\hat{k}^i \times \mathbf{E}^i(\mathbf{r}) \tag{2.2}$$

where $\hat{\theta}$, $\hat{\phi}$ are the unit vectors in the spherical coordinate system, α is the polarization angle, Z is the wave impedance and $\mathbf{k}^i$ is the wave vector:

$$\mathbf{k}^i = k_0(\sin\theta^i\cos\phi^i\,\hat{x} + \sin\theta^i\sin\phi^i\,\hat{y} + \cos\theta^i\hat{z}) \tag{2.3}$$

where θ^i, ϕ^i are the coordinates of the incident angle in the spherical coordinate system. From Section 1.3.3, we can establish the electric field integral equation (EFIE) on the surface of the PEC object

$$\left[\mathbf{E}^i + Z\,\mathbf{L}(\mathbf{J})\right]\Big|_t = 0 \tag{2.4}$$

and the magnetic field integral equation (MFIE),

$$\mathbf{J} - \hat{n}\times\mathbf{K}(\mathbf{J}) = \hat{n}\times\mathbf{H}^i \tag{2.5}$$

In (2.4), the symbol $|_t$ denotes the tangential components. Operators $\mathbf{L}$, and $\mathbf{K}$ are as defined in Section 1.3.2. Mathematically, the above two equations are equivalent. One can solve either and obtain the unknown equivalent electric current density $\mathbf{J}$. Once $\mathbf{J}$ is calculated, the scattering field in the far-field can be calculated in a straightforward way. Thus, it is critical to solve (2.4) or (2.5). Evidently, it is impossible to solve the integral equations for metallic objects with arbitrary shape without employing computers. To this end, the integral equations need to be discretized first.

2.1.2 Discretization in MoM

The procedure of discretization in MoM will be outlined in this section. The steps involved in the procedure will be discussed in greater detail in the subsequent sections.

Generally, an integral equation can be written as

$$Lf = g \tag{2.6}$$

where L is the linear operator, g is the known "source" function, and f is the unknown "solution" function. Express f in terms of the linear combination of a set of basis functions (BFs) $\{f_1, f_2, f_3, \ldots\}$

$$f = \sum_j \alpha_j f_j \tag{2.7}$$

where the a priori unknown scalar expansion coefficients α_j are to be determined. Substituting (2.7) into (2.6), and taking the linearity property of L into account gives rise to the following equation

$$\sum_j \alpha_j L f_j = g \tag{2.8}$$

Then, choose a set of testing functions (TFs) $\{\omega_1, \omega_2, \omega_3, \ldots\}$. Take the inner-product between the testing functions and the functions on both sides of (2.8), to obtain

$$\sum_j \alpha_j \langle \omega_i, L f_j \rangle = \langle \omega_i, g \rangle \tag{2.9}$$

where $i = 1, 2, 3, \ldots$. This system of equations can be written in matrix form as

$$[A]\{\alpha\} = \{g\} \tag{2.10}$$

where

$$\{g\} = \{\langle \omega_i, g \rangle\} \tag{2.11}$$

$$[A] = \left[\langle \omega_i, L f_j \rangle\right] \tag{2.12}$$

If $[A]$ is nonsingular, then its inverse exists. In this case, $\{\alpha\}$ is given by

$$\{\alpha\} = [A]^{-1}\{g\} \tag{2.13}$$

The solution for f is given by (2.7). The above steps constitute the discretization procedure of MoM. The name *method of moments* stems from the convention referring to $\int x^n f(x)dx$ as the n$^{\text{th}}$ moment of f. When x^n is replaced by an arbitrary ω_n, we still continue calling the integral "a moment of f."

2.1.3 Choice of Basis and Testing Functions

It is clear from the outline of the MoM described in the above section that the first step in discretizing the integral equations is an appropriate choice of basis and testing functions. Actually, the choice can be quite flexible. Therefore, in the following, we

will compare convenience, accuracy and efficiency of implementation when choosing different basis and testing functions.

The basis functions can be categorized into two classes, that is, the entire-domain basis functions and the subdomain basis functions. The entire-domain basis functions are fixed and continuous in the entire definition domain of the operator L. This kind of basis functions was used by Ritz, Galerkin and other scholars to solve particular problems in earlier times. However, as far as many general problems, especially the 2D and 3D problems, it is very hard to predict the approximate solutions. Therefore, it is challenging to construct this type of basis functions, and that is why they have very few applications at present. On the other hand, employing basis functions defined on a finite support has gained wide popularity. The subdomain basis functions are only fixed and continuous in the special subdomain of the definition domain of the operator, while they take zero value outside this subdomain. The subdomain bases are normally constructed in the following manner: First, the entire simulation domain is partitioned into an adequate number of subdomains. Then, the function on each subdomain is constructed by using the discrete function values at certain locations in this subdomain through the polynomial interpolation method. Finally, the undetermined function on the entire domain is expressed as the composition of the functions on each of the subdomains. It is evident that two critical issues are involved here: the geometry shape of the subdomain, and the choice of the interpolation parameters. There are a many shapes of subdomains. Generally speaking, an arbitrary surface can be approximated fairly well with triangles, while tetrahedra are suitable if the region of interest is a volume. This is because a triangle is a simplex for a 2D manifold, and a tetrahedron is a simplex for a volume. Commercial software such as *Hypermesh* and *Ansys* can provide the mesh solution to regions of arbitrary geometry. Although smaller size of subdomains results in a more accurate model, the large number of subsections also leads to substantial degree of freedom that will cause severe computational burden. In practice, if we use triangular facets or tetrahedra as meshing elements, mesh size of $\lambda/8$ is adequate (λ is the operating wavelength). The choice of interpolation parameters depends on the specific integral equations. Convergence consideration require the solution to be in a *Sobolev space* (S^n) specified by the structure of integral equations of interest (inhabitants of S^n are by definition n-times continuously differentiable).

Next, we will construct the basis functions for the surface integral equation (2.4) and use this example to illustrate the principle in constructing the basis functions. Here, we choose triangular facets as the basic form of the subdomain. Since L includes the divergence operator, the expression of the solution function needs to have bounded divergence. When we apply this condition to the common edges of adjacent triangles, the normal component of the electric current density is required to be continuous. To this end, we choose a pair of adjacent triangular facets that share one common edge (referred to as the positive triangle T_i^+ and the negative one T_i^-) as

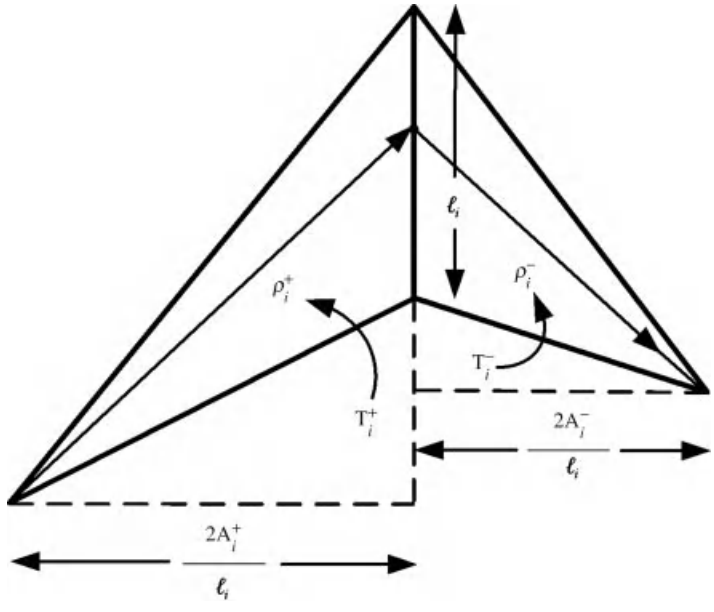


Figure 2.1 RWG basis function.

the support of the subdomain basis function $\boldsymbol{g}_i$, and choose the normal component of electric current density at the edges on the triangle $J_i (i = 1, 2, 3)$ as the interpolation parameters, as shown in Figure 2.1. It is worth noting here that the basis functions in T_i^+ and T_i^- are defined in a slightly different way (check ρ_i^+ and ρ_i^- in Figure 2.1). Thus, the equivalent electric current can be interpolated as

$$\mathbf{J} = \sum_{i=1}^{3} J_i \mathbf{g}_i \tag{2.14}$$

where

$$\mathbf{g}_i = \begin{cases} \dfrac{l_i}{2A_i^+} \rho_i{}^+ & \text{in } T_i^+ \\ \dfrac{l_i}{2A_i^-} \rho_i{}^- & \text{in } T_i^- \\ 0 & \text{elsewhere} \end{cases} \tag{2.15}$$

l_i denotes the length of the ith edge, $A_i^{\pm}$ is area of $T_i^{\pm}$, and $\rho_i^{\pm}$ corresponds to the ith position vector in $T_i^{\pm}$.

The *Rao–Wilton–Glisson (RWG) vector basis function* is constructed in this way. It was first proposed in [3] by Rao, Wilton and Glisson. It can be shown that the normal component of g_i on the ith edge equals unity, while that on the jth $(j \neq i)$ edge equals zero. This assures that the normal component of the interpolation function $\mathbf{J}$ on the ith edge of the triangle corresponds to J_i. Since the adjacent triangles share a common edge, the continuity of the RWG basis function on the interface of different triangles is assured.

In practical applications, there are numerous choices for testing functions, for example, Dirac delta functions, piecewise linear functions connecting the centers of adjacent triangles, or basis functions. The thus-obtained schemes are called *point-matching, line-matching, and Galerkin methods*, respectively. From the implementation point of view, point matching is the easiest, and Galerkin's method is the most complicated scheme. However, Galerkin's method is the most stable one as far as the computational performance is concerned. Besides, the choice of the testing function is closely related to the particular operator in the integral formulation for a particular problem. For example, if the gradient operator in L is imposed on the Green's function, the order of the singularity will increase, resulting in difficulty in the computation and a decrease of accuracy of the results. To circumvent this problem, the vector identity is used to transpose the gradient operator to the testing functions. Under this circumstance, we cannot use the Dirac delta function as the testing function.

2.1.4 Discretized Integral Equation (DIE) and the Numerical Behavior Analysis

When the basis functions are selected, the equivalent electric surface current $\mathbf{J}$ can be expressed as

$$\mathbf{J} = \sum_{i=1}^{N_S} J_i \mathbf{g}_i \tag{2.16}$$

where N_S denotes the number of the total edges of the triangular facets discretizing S. The RWG basis function $\boldsymbol{g}_i$ has been introduced in the preceding section. Substituting (2.16) into (2.4), and choosing $\boldsymbol{g}_i$ as the testing function yields the following discretized form of the electric field integral equation:

$$\left[P^{TE}\right]\{J\} = \left\{b^{TE}\right\} \tag{2.17}$$

where

$$P_{ij}^{TE} = -Z \int_S \mathbf{g}_i \cdot \mathbf{L}(\mathbf{g}_j)\, dS \tag{2.18}$$

$$b_i^{TE} = \int_S \mathbf{g}_i \cdot \mathbf{E}^i dS \tag{2.19}$$

The above discretized integral equation is called the *tangential electric-field integral equation (T-EFIE)*, since the testing function is tangential to the basis function in the EFIE. We can also use $\hat{\boldsymbol{n}} \times \boldsymbol{g}_i$, which is normal to the basis function, as the testing function and obtain the following discretized *normal electric-field integral equation (N-EFIE)*

$$\left[P^{NE}\right]\{J\} = \left\{b^{NE}\right\} \tag{2.20}$$

where

$$P_{ij}^{NE} = -Z\int_S \hat{n} \times \mathbf{g}_i \cdot \mathbf{L}(\mathbf{g}_j)\, dS \tag{2.21}$$

$$b_i^{NE} = \int_S \hat{n} \times \mathbf{g}_i \cdot \mathbf{E}^i dS \tag{2.22}$$

Similarly, substituting (2.16) into (2.5), and choosing $\boldsymbol{g}_i$ as the testing function yields the following discretized *tangential magnetic-field integral equation (T-MFIE)*:

$$\left[P^{TH}\right]\{J\} = \left\{b^{TH}\right\} \tag{2.23}$$

where

$$P_{ij}^{TH} = \int_S \mathbf{g}_i \cdot \left[\mathbf{g}_j - \hat{n} \times \mathbf{K}(\mathbf{g}_j)\right] dS \tag{2.24}$$

$$b_i^{TH} = \int_S \mathbf{g}_i \cdot \hat{n} \times \mathbf{H}^i dS \tag{2.25}$$

Using $\hat{\boldsymbol{n}} \times \boldsymbol{g}_i$ as the testing function yields *normal magnetic-field integral equation (N-MFIE)*

$$\left[P^{NH}\right]\{J\} = \left\{b^{NH}\right\} \tag{2.26}$$

where

$$P_{ij}^{NH} = \int_S \hat{n} \times \mathbf{g}_i \cdot \left[\mathbf{g}_j - \hat{n} \times \mathbf{K}(\mathbf{g}_j)\right] dS \tag{2.27}$$

$$b_i^{NH} = \int_S \hat{n} \times \mathbf{g}_i \cdot \hat{n} \times \mathbf{H}^i dS \tag{2.28}$$

Mathematically, the above four integral equations are equivalent. However, they possess distinctive numerical properties. Let us analyze the diagonal entries in the coefficient matrices. It can be found that the integrands for $i = j$ in (2.21) and (2.27) are

Table 2.1 The condition numbers of the coefficient matrices of different schemes when in the simulation of a PEC sphere scattering

Number of Unknowns	$[P^{TE}]$	$[P^{NE}]$	$[P^{TH}]$	$[P^{NH}]$
3456	1.920×10^2	1.486×10^9	1.52×10^1	2.182×10^9

equal to zero, while those in (2.18) and (2.24) are not. Moreover, the integrands in (2.18) and (2.24) yield greater values on entries $i = j$ than those on entries $i \neq j$. Therefore, $[P^{TE}]$ and $[P^{TH}]$ are quasi-*diagonally dominant matrix*, resulting in better matrix behavior than $[P^{NE}]$ and $[P^{NH}]$, which are diagonally weak. Furthermore, $[P^{TH}]$ should bear a better condition number than that of $[P^{TE}]$, because the former originates from MFIE, which is the Fredholm integral equation of the second kind, and the latter originates from EFIE, which is the Fredholm integral equation of the first kind. In order to verify these statements, a PEC sphere with its diameter being equal to one wavelength is modeled, and the condition numbers of the aforementioned matrices are listed and compared in Table 2.1. It can be seen that $[P^{TH}]$ bears the best condition number, which is followed by $[P^{TE}]$, while $[P^{NE}]$ and $[P^{NH}]$ are already ill-conditioned.

2.1.5 Handling of Singularity

Next, we will introduce details of the procedures involved in calculating the elements in $[P^{TH}]$ and $[P^{TE}]$. When the field point is far from the source point, the calculation is straightforward. That is because the integrands of the elements are smooth. Consequently, the integrals can be calculated using the *Gauss–Legendre formulae* on a triangle, that is,

$$\frac{1}{\Delta}\int_T f(\mathbf{r})dS = \sum_{i=1}^{n} \omega_i f(\mathbf{r}_i) \tag{2.29}$$

where Δ is the area of the triangle, and T denotes the triangular integral area. The number n is chosen according to the accuracy desired. If the accuracy requirement is not very high, $n = 1$ is sufficient. This means only the center of the triangle is taken into accounted in the integral calculation. Hence, the weight ω_1 equals 1. The choice of $n = 4$ leads to more accurate numerical results. In this case the center of the triangle as well as all the three vertices are involved in the integral calculation, and their corresponding weights ω_i ($i = 1,2,3,4$) are 3/4, 1/12, 1/12, and 1/12, respectively. If the requirement on the accuracy is even higher, we can choose $n = 7$. In that case not only the center and the three vertices of the triangle, but also the center points of the edges are included in the right-hand side (RHS) of (2.29). The corresponding weights ω_i ($i = 1, 2, \ldots, 7$) are 9/20, 1/20, 1/20, 1/20, 2/15, 2/15, and 2/15, respectively. While a greater n value can in general lead to higher accuracy, it also decreases the computational efficiency. In the MoM calculations, the calculation of the matrix entries is very

time consuming, and critically affects the computational efficiency of the method. As a consequence, the value of n is usually chosen by balancing accuracy and efficiency. Normally, sufficient accuracy can be provided by $n = 4$. When the field point is very close to the source point, the calculation of the matrix elements can become a tricky undertaking. The computational complication is due to the dramatic change in the integrand when $k|\mathbf{r} - \mathbf{r}'| \ll 1$. In particular, when the field and the source subdomains coincide, the integrand becomes singular. On the other hand, the accuracy in computing these diagonal elements greatly affects the accuracy of the final MoM results. Therefore, special techniques need to be developed to take special care of these cases. The following discussion is devoted to this end.

The evaluation of singular integrals can be performed by both numerical and analytical techniques. By applying the singularity-extraction technique [4], the singular terms are extracted from the kernel and integrated analytically. The remaining bounded difference integrand can be integrated numerically. This method is well established for constant and linear basis functions, like the RWG basis functions. However, when higher-order basis functions are applied, this analytical technique become challenging and singularity-cancellation techniques such as Duffy's method [5–10] are more popularly chosen in favor of the purely numerical quadrature schemes. We will introduce the singularity-extraction method in this section, and present Duffy's method in the following section.

2.1.5.1 Singularity-Extraction Technique

Consider the integral $\int_S \vec{g}_i \cdot L(\vec{g}_j)\, dS$. According to the vector identity

$$\mathbf{b} \cdot \nabla a = \nabla \cdot (a\mathbf{b}) - a\nabla \cdot \mathbf{b} \tag{2.30}$$

the second term in $\int_S \vec{g}_i \cdot L(\vec{g}_j)\, dS$ can be expressed as

$$\int_S \int_{S'} \mathbf{g}_i \cdot \nabla(\nabla' \cdot \mathbf{g}_j G)\, dS'\, dS = \int_S \int_{S'} \left[\nabla \cdot (\mathbf{g}_i \nabla' \cdot \mathbf{g}_j G) - \nabla \cdot \mathbf{g}_i \nabla' \cdot \mathbf{g}_j G\right] dS' dS \tag{2.31}$$

Then, by applying Gauss's law, the first entry at the RHS leads to the integral contribution

$$\int_S \int_{S'} \nabla \cdot (\mathbf{g}_i \nabla' \cdot \mathbf{g}_j G)\, dS'\, dS = \oint_l \mathbf{g}_i \int_{S'} \nabla' \cdot \mathbf{g}_j\, G dS' \cdot \hat{m} dl \tag{2.32}$$

where l denotes the outer boundary of the triangle, and $\hat{m}$ denotes the unit outward vector normal to the boundary. Since the normal component of $\mathbf{g}_i$ equals unity on the corresponding ith edge and zero on the remaining edges, the above integral over the outer boundary of the triangle reduces to the integral over the ith edge. On the other hand, the integrals on the adjacent triangles cancel each other due to the identical

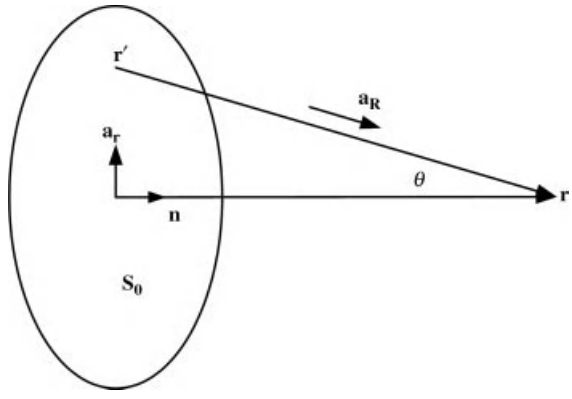


Figure 2.2 A very small region S_0 enclosing the singular point.

interpolating parameters and the opposite directions. Thus, we obtain

$$\int_S \mathbf{g}_i \cdot \mathbf{L}(\mathbf{g}_j) dS = -jk \int_S \int_{S'} \left[\mathbf{g}_i \cdot \mathbf{g}_j - \frac{1}{k^2} \nabla \cdot \mathbf{g}_i \nabla' \cdot \mathbf{g}_j \right] G dS' dS = \int_S MG dS \qquad (2.33)$$

Conventionally the above integral is rewritten into two parts:

$$\int_S MG dS = \int_{S-S_0} MG dS + \int_{S_0} MG dS, \qquad (2.34)$$

where S_0 is an infinitesimally small region enclosing the singular point, as indicated in Figure 2.2. The first term on the RHS of the equation (2.34) is usually named the principle value integration (with the notation $PV \int_S$), and the second entry is called the singularity residual term. Since the integrand in (2.34) is singular of order one, the singularity residual integral equals zero. In order to calculate the principle value integration in (2.34) numerically, the Green's function G is rewritten as

$$G = \frac{1}{4\pi R}(e^{-jkR} - 1) + \frac{1}{4\pi R} \qquad (2.35)$$

where $\mathbf{R} = \mathbf{r} - \mathbf{r}'$ and R is the magnitude of $\mathbf{R}$. Using Taylor series expansion of the first term at the RHS and keeping only the first two dominant terms we have:

$$\frac{1}{4\pi R}(e^{-jkR} - 1) \approx -\frac{k^2}{8\pi} R + j \frac{k}{24\pi}(k^2 R^2 - 6) \qquad kR \ll 1 \qquad (2.36)$$

Without the presence of any singularity, this entry can be calculated in a standard way. The analytical expression of the second entry in (2.35) will be given later.

Unlike $\int_S \mathbf{g}_i \cdot \mathbf{L}(\mathbf{g}_j) dS$, the singularity residual in $\int_S \mathbf{g}_i \cdot \hat{n} \times \mathbf{K}(\mathbf{g}_j) dS$ is no longer zero, since the derivatives in ∇G increase the order of the singularity by one. Next we

calculate the singularity residual in $K(\mathbf{g}_j)$. Since S_0 is an infinitesimally small region, the equivalent electric current on it can be approximated using a constant value. As a result, we obtain

$$\int_{S_0} \mathbf{g}_j \times \nabla G dS \approx \mathbf{g}_j \times \int_{S_0} \nabla G dS \tag{2.37}$$

Approximating ∇G to the first order, that is,

$$\nabla G = \nabla \frac{e^{-jkR}}{4\pi R} \approx \frac{1}{4\pi} \nabla \frac{1}{R} \tag{2.38}$$

Then, (2.37) reads

$$\int_{S_0} \mathbf{g}_j \times \nabla G dS \approx \frac{1}{4\pi} \mathbf{g}_j \times \int_{S_0} \nabla \frac{1}{R} dS \tag{2.39}$$

As shown in Figure 2.2, we can express $\nabla(1/R)$ as

$$\nabla \frac{1}{R} = -\frac{\mathbf{a}_R}{R^2} = -\frac{\hat{n}}{R^2} \cos\theta + \frac{\hat{a}_r}{R^2} \sin\theta \tag{2.40}$$

where $\mathbf{a}_R = (\mathbf{r} - \mathbf{r}')/R$. By applying the symmetry of the integral region, the component in the $\hat{a}_r$ direction equals zero, which yields

$$\int_{S_0} \nabla \frac{1}{R} dS = -\hat{n} \int_{S_0} \frac{\cos\theta}{R^2} dS = -\hat{n} \int_{S_0} d\Omega = -\hat{n}\Omega \tag{2.41}$$

Here, Ω is the solid angle of S_0. For a smooth surface, $\Omega = 2\pi$. Thus, the singularity residual in the integral $K(\mathbf{g}_j)$ is

$$\int_{S_0 \to 0} \mathbf{g}_j \times \nabla G dS' \approx -\frac{\Omega}{4\pi} \mathbf{g}_j \times \hat{n} \tag{2.42}$$

which leads to

$$\int_S \mathbf{g}_i \cdot \hat{n} \times \mathbf{K}(\mathbf{g}_j) dS = \frac{\Omega}{4\pi} \int_S \hat{n} \times \mathbf{g}_i \cdot (\hat{n} \times \mathbf{g}_j) dS + \int_S \hat{n} \times \mathbf{g}_i \cdot \left(PV \int_{S'} \mathbf{g}_j \times \nabla G dS' \right) dS \tag{2.43}$$

where the notation $PV\int_{S'}$ denotes the principle value integral, which is calculated in the following. Express ∇G as

$$\nabla G = -\frac{1}{4\pi} \frac{\mathbf{R}}{R^3} (jkR + 1) e^{-jkR} \tag{2.44}$$

Similar to (2.35), ∇G can be rewritten into two parts:

$$\nabla G = \frac{1}{4\pi} \frac{\mathbf{R}}{R^3} \left[-(jkR + 1) e^{-jkR} + 1 + \frac{1}{2} k^2 R^2 \right] - \frac{1}{4\pi} \frac{\mathbf{R}}{R^3} \left(1 + \frac{1}{2} k^2 R^2 \right) \tag{2.45}$$

The first term in (2.45) can be expended using the Taylor series expansion:

$$\frac{1}{4\pi}\frac{\mathbf{R}}{R^3}\left[-(jkR+1)e^{-jkR}+1+\frac{1}{2}k^2R^2\right] \approx \frac{k^3}{4\pi}\mathbf{R}\left[\frac{kR}{8}\left(1-\frac{1}{18}k^2R^2\right)+j\frac{1}{3}\left(1-\frac{1}{10}k^2R^2\right)\right] \tag{2.46}$$

where $kR \ll 1$. It can be seen from (2.45) that there is no singularity involved in the first term, so it can be calculated directly. The next problem is how to calculate the second term at the RHS of (2.35) and (2.45). For clarity we shall denote the triangular integral region where the source point resides by T and its boundary by ∂T. Denote the three vertices of the triangle by $r_i (i = 1, 2, 3)$, and the edges opposite to them by $E_i (i = 1, 2, 3)$. Given r_0 is the projection point of the field point r on T. Construct a polar coordinate system (P, ϕ) with r_0 as origin on T, as shown in Figure 2.3. In this way, $\mathbf{R}$ can be expressed as

$$\mathbf{R} = \mathbf{r} - \mathbf{r}_0 - (\mathbf{r}' - \mathbf{r}_0) = \mathbf{d} - \mathbf{P} \tag{2.47}$$

Then, the calculation of the second term at the right-hand sides of (2.35) and (2.45) becomes the following principle value integration

$$I_n(\mathbf{r}) = \int_{T-T_\varepsilon} \frac{1}{R^n} dT \qquad n = 1,3 \tag{2.48}$$

$$I_{nP}(\mathbf{r}) = \int_{T-T_\varepsilon} \frac{\mathbf{P}}{R^n} dT \qquad n = 1,3 \tag{2.49}$$

where T_ε is an infinitesimally small circular region with radius ε centered at r_0. It is worth noting that when we calculate (2.34) and (2.43), there are other variables besides

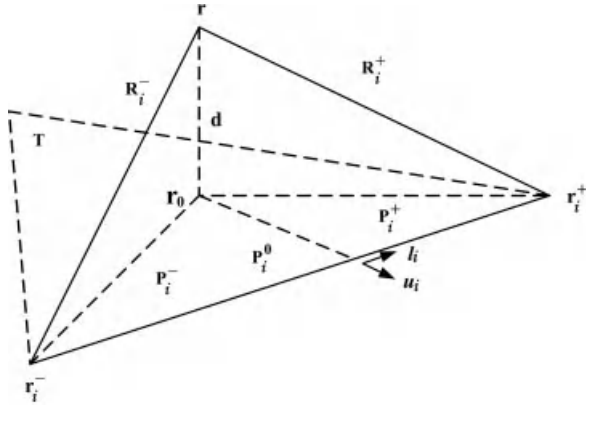


Figure 2.3 The projection of the field point in the triangle of the source plane.

G and ∇G, namely, M and $\mathbf{g}_j$. However, within the integral region, these variables change very slowly compared with the singularity introduced in G and ∇G. As a consequence, these variables are taken outside the integral as constants in the calculation. According to the surface divergence formula in polar coordinate system, we obtain

$$\frac{1}{R} = \nabla_S \cdot \left(\frac{R}{P}\hat{P}\right) \tag{2.50}$$

where P and $\hat{P}$ are the magnitude and directional unit vector of $\mathbf{P}$, respectively. By applying Gauss' divergence theorem, (2.50) can be further written as

$$\begin{aligned} I_1 &= \int_{T-T_\varepsilon} \frac{1}{R} dT \\ &= \int_{T-T_\varepsilon} \nabla_S \cdot \left(\frac{R}{P}\hat{P}\right) dT \\ &= \int_{\partial T} \left(\frac{R}{P^2}\mathbf{P}\right) \cdot \hat{u} dl + \int\limits_{\partial T_\varepsilon} \frac{R}{\varepsilon}(\hat{\varepsilon}\cdot\hat{u})\varepsilon d\phi \\ &= \int_{\partial T} \left(\frac{R}{P^2}\mathbf{P}\right) \cdot \hat{u} dl - d\int\limits_{\partial T_\varepsilon} d\phi \end{aligned} \tag{2.51}$$

where $\hat{u}$ denotes the outward unit vector normal to the integral path.

On edge i, $\hat{l}_i \cdot \hat{u}_i = 0$, and $\mathbf{P}_i \cdot \hat{u}_i = \mathbf{P}_i^0$ is a constant. Additionally,

$$\frac{R}{P^2} = \frac{1}{R} + \frac{d^2}{P^2 R} \tag{2.52}$$

holds, so

$$I_1 = \sum_{i=1}^{3} P_i^0 \int\limits_{E_i} \left(\frac{1}{R} + \frac{d^2}{P^2 R}\right) dl - d\int\limits_{\partial T_\varepsilon} d\phi \tag{2.53}$$

Next, we have

$$\int\limits_{E_i} \frac{dl}{R} = \int_{l_i^-}^{l_i^+} \frac{dl}{\sqrt{l^2 + (P_i^0)^2 + d^2}} = \ln \frac{R_i^+ + l_i^+}{R_i^- + l_i^-} \tag{2.54}$$

$$\begin{aligned} \int_{E_i} \frac{1}{P^2 R} dl &= \int_{l_i^-}^{l_i^+} \frac{dl}{\left[l^2 + (P_i^0)^2\right]\sqrt{l^2 + (P_i^0)^2 + d^2}} \\ &= \frac{1}{|P_i^0 d|}\left[tg^{-1}\frac{|d|l_i^+}{|P_i^0|R_i^+} - tg^{-1}\frac{|d|l_i^-}{|P_i^0|R_i^-}\right] \end{aligned} \tag{2.55}$$

and

$$\int_{\partial T_\varepsilon} d\phi = \sum_{i=1}^{3} \left(tg^{-1} \frac{l_i^+}{P_i^0} - tg^{-1} \frac{l_i^-}{P_i^0} \right) \tag{2.56}$$

where the plus or the minus sign in the superscript of the variables in (2.54), (2.55), and (2.56) denotes the corresponding end on that edge. Applying the following trigonometric identity

$$tg^{-1}\alpha - tg^{-1}\beta = tg^{-1} \frac{\alpha - \beta}{1 + \alpha\beta} \tag{2.57}$$

we can obtain:

$$tg^{-1} \frac{l_i^+}{P_i^0} - tg^{-1} \frac{|d| l_i^+}{\left|P_i^0\right| R_i^+} = tg^{-1} \frac{P_i^0 l_i^+}{\left(R_i^0\right)^2 + |d| R_i^+} \tag{2.58}$$

where $\left(R_i^0\right)^2 = \left(P_i^0\right)^2 + d^2$. Let

$$f_i = \ln \frac{R_i^+ + l_i^+}{R_i^- + l_i^-} \tag{2.59}$$

and

$$\beta_i = tg^{-1} \frac{P_i^0 l_i^+}{\left(R_i^0\right)^2 + |d| R_i^+} - tg^{-1} \frac{P_i^0 l_i^-}{\left(R_i^0\right)^2 + |d| R_i^-} \tag{2.60}$$

then we obtain the following compact expression

$$I_1 = \int_{T - T_\varepsilon} \frac{1}{R} dT = \sum_{i=1}^{3} (P_i^0 f_i - |d| \beta_i) \tag{2.61}$$

Next, we will focus on the calculation of I_{1P}. Applying the surface gradient operator in the polar coordinates we can obtain

$$\nabla_s R = \frac{\mathbf{P}}{R} \tag{2.62}$$

Thus, I_{1P} can be expressed as

$$\begin{aligned} I_{1P} &= \int_{T - T_\varepsilon} \nabla_s R \, dT \\ &= \int_{\partial T} R \, \hat{u} dl + \int_{\partial T_\varepsilon} R \, \hat{u} dl \\ &= \sum_{i=1}^{3} \hat{u}_i \int_{l_i^-}^{l_i^+} \sqrt{l^2 + \left(P_i^0\right)^2 + d^2} dl \\ &= \frac{1}{2} \sum_{i=1}^{3} \hat{u}_i \left[\left(R_i^0\right)^2 f_i + R_i^+ l_i^+ - R_i^+ l_i^+ \right] \end{aligned} \tag{2.63}$$

In a similar approach, $1/R^3$ can be rewritten as

$$\frac{1}{R^3} = -\nabla \cdot \left(\frac{1}{PR}\hat{P}\right) \tag{2.64}$$

which leads to

$$I_3 = \beta_i / d \tag{2.65}$$

By applying

$$\frac{\mathbf{P}}{R^3} = -\nabla_S \left(\frac{1}{R}\right) \tag{2.66}$$

I_{3P} can be obtained:

$$I_{3P} = -\sum_{i=1}^{3} \hat{u}_i f_i \tag{2.67}$$

It is worth noting that when $R_i^{\pm} + l_i^{\pm} = 0$, a singularity still exists in the expression of f_i. However, because R_i^0 and P_i^0 are concurrently zero at the same time, we can still calculate the scalar product of f_i with R_i^0 or P_i^0 in I_1 or I_{1P}, which turns out to be zero. The term of $\hat{u}_i f_i$ in I_{3P} cancels with its counterpart in the adjacent triangle. Another remark is that in order to avoid the singularity in I_3 when $d = 0$, we normally calculate $d \cdot I_3$.

2.1.5.2 Duffy's Method

Despite the widespread usage, the singularity-extraction method has a number of disadvantages. First, the singular term extracted is often determined by utilizing Taylor series expansion, which experiences increased complexity with the complexity of the basis functions, geometry and Green's functions involved. Second, the extraction method leads to a problem-specific program because the source subdomain geometry, basis function, and asymptotic form of the Green's function are inextricably linked in the analytically evaluated self-term integral [11]. In this section, we will introduce a general-purpose scheme, which is flexible in dealing with higher-order bases and complex geometry and Green's functions. This scheme employs substitution of variables such that the Jacobian of the transformation cancels the singularity. Different transformations lead to various schemes, such as the polar-angular coordinate method [5,12], Duffy's method [6,7], and so on. Duffy's method is a general one in treating the kind of problem where singularity appears at vertices of the integration domain. It has wide applications in 2D and higher-dimensional problems, especially with the frequently used triangular or pyramid-shaped integration domain. In this section, we will introduce the principle of this method.

I. *Duffy coordinate transformation*

Duffy's method is based on Duffy transformation. Consider a triangular integration domain S as shown in Figure 2.4a. The distance R of a point (x, y) to the origin can be expressed as

$$R = \sqrt{x^2 + y^2} \tag{2.68}$$

Suppose $f(x, y)$ is a function without any singularity in the integration domain. However, $f(x, y)/R$ contains a singularity at the origin. Consider the numerical integration of $f(x, y)/R$ on S

$$I = \int_S \frac{f(x,y)}{R} dS = \int_{x=0}^{1} \int_{y=0}^{x} \frac{f(x,y)}{\sqrt{x^2 + y^2}} dydx \tag{2.69}$$

Let us introduce the Duffy transformation:

$$y = xu \tag{2.70}$$

Thus, the triangular domain S in Figure 2.4a is transformed into the square domain as shown in Figure 2.4b. By substituting (2.68) and (2.70) into (2.69) we obtain:

$$R = \sqrt{x^2 + x^2u^2} = x\sqrt{1 + u^2} \tag{2.71}$$

The Jacobian of this Duffy transformation is

$$J = \begin{vmatrix} \dfrac{\partial x}{\partial x} & \dfrac{\partial y}{\partial x} \\ \dfrac{\partial x}{\partial u} & \dfrac{\partial y}{\partial u} \end{vmatrix} = \begin{vmatrix} \dfrac{\partial x}{\partial x} & \dfrac{\partial (xu)}{\partial x} \\ \dfrac{\partial x}{\partial u} & \dfrac{\partial (xu)}{\partial u} \end{vmatrix} = x \tag{2.72}$$

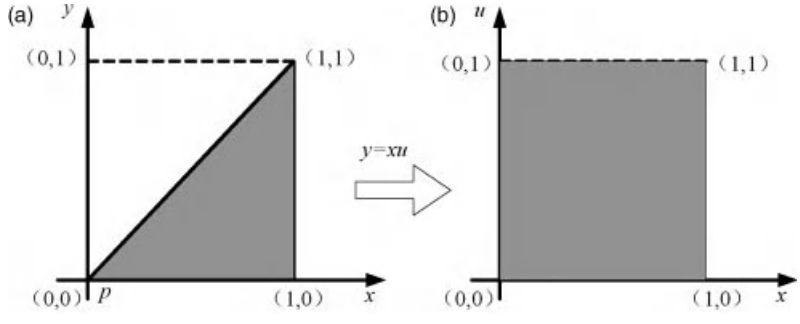


Figure 2.4 Duffy coordinate transformation. (a) Original integration domain; (b) Transformed integration domain.

Substituting (2.71) and (2.72) into (2.69) yields

$$\begin{aligned}\int_S \frac{f(x,y)}{R}dS &= \int_{x=0}^{1}\int_{y=0}^{x}\frac{f(x,y)}{\sqrt{x^2+y^2}}dydx \\ &= \int_{x=0}^{1}\int_{u=0}^{1}\frac{f(x,ux)}{x\sqrt{1+u^2}}Jdudx \\ &= \int_{x=0}^{1}\int_{u=0}^{1}\frac{f(x,ux)}{\sqrt{1+u^2}}dudx\end{aligned} \tag{2.73}$$

Obviously, after this transformation, the integration kernel $f(x,ux)/\sqrt{1+u^2}$ doesn't have any singularity in the transformed integration domain. In this way, the integration with a singularity in (2.69) in the (x,y) coordinate system is transformed into a bounded integration in the new coordinate system (x,u). From the above analysis, it can be seen that the essential technique in a transformation-based singularity-cancellation method is to apply proper coordinate transformation, whose Jacobian cancels the singularity in the integration kernel.

Before we go further into the details of Duffy's method, let us introduce another way of calculating the Jacobian determinant in the coordinate transformation. Consider the transformation between a curvilinear triangular patch in a Cartesian coordinate system and a flat isosceles right triangular patch in the parameter coordinate system (u,v), as illustrated in Figure 2.5. Then, an arbitrary vector defined by point (x,y,z) in Cartesian coordinate system can be expressed as a function of u and v, that is,

$$r(u,v) = x(u,v)\hat{x} + y(u,v)\hat{y} + z(u,v)\hat{z} \tag{2.74}$$

whose covariant and contravariant base vectors $\mathbf{l}_1$, $\mathbf{l}_2$, $\mathbf{l}^1$ and $\mathbf{l}^2$ are given by

$$\mathbf{l}_1 = \frac{\partial \mathbf{r}(u,v)}{\partial u} = \frac{\partial x}{\partial u}\hat{x} + \frac{\partial y}{\partial u}\hat{y} + \frac{\partial z}{\partial u}\hat{z} \tag{2.75}$$

$$\mathbf{l}_2 = \frac{\partial \mathbf{r}(u,v)}{\partial v} = \frac{\partial x}{\partial v}\hat{x} + \frac{\partial y}{\partial v}\hat{y} + \frac{\partial z}{\partial v}\hat{z} \tag{2.76}$$

$$\mathbf{l}^1 = \nabla u = \frac{\partial u}{\partial x}\hat{x} + \frac{\partial u}{\partial y}\hat{y} + \frac{\partial u}{\partial z}\hat{z} \tag{2.77}$$

$$\mathbf{l}^2 = \nabla v = \frac{\partial v}{\partial x}\hat{x} + \frac{\partial v}{\partial y}\hat{y} + \frac{\partial v}{\partial z}\hat{z} \tag{2.78}$$

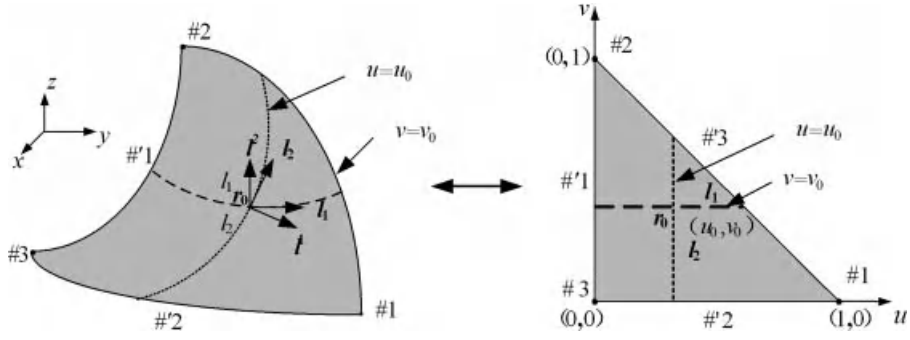


Figure 2.5 The covariant and contravariant basis vectors on a curvilinear triangular patch in the Cartesian and parameter coordinate system.

These base vectors at point $\mathbf{r}_0$ in the Cartesian and parameter coordinate systems, are illustrated in Figure 2.5. Figure 2.6 corresponds to the case when $\mathbf{r}_0$ falls on an edge of the triangular patch. Detailed discussion on the properties of these basis vectors can be found in [10]. Here, we only present a useful equation to calculate the Jacobian of this mapping from the Cartesian coordinate system to the parameter coordinate system:

$$J = |\mathbf{l}_1 \times \mathbf{l}_2| \tag{2.79}$$

II. *Duffy's method in MoM*

Because of the convenience and the high accuracy, Duffy's method has drawn considerable attention in MoM. The conventional Duffy's method can effectively handle the singularity when the field and source triangles coincide. Here, we use a curvilinear triangular element as an example to illustrate the method.

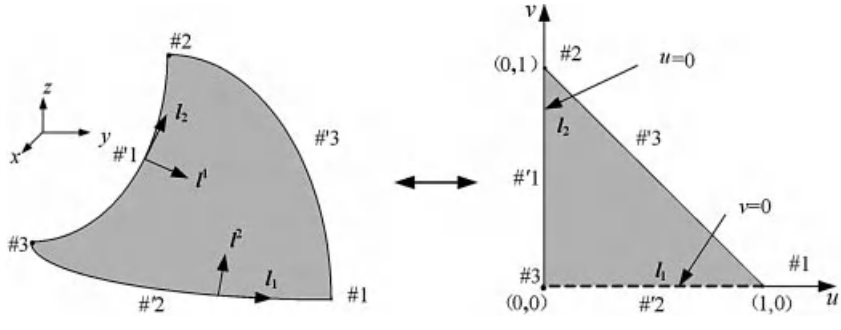


Figure 2.6 The covariant and contravariant basis vectors on a curvilinear triangular patch in the Cartesian and parameter coordinate system.

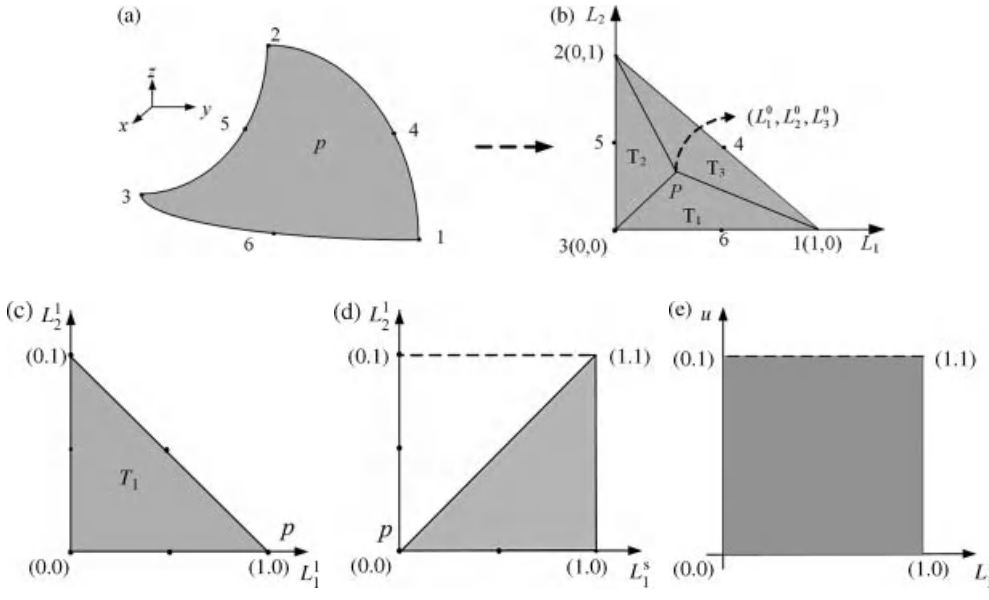


Figure 2.7 Illustration of the Duffy transformation technique. (a) The curvilinear triangle in the Cartesian coordinate system. (b) The same triangle in the parameter coordinate system being partitioned into three subtriangles. (c) The parameter coordinate system of each subtriangle $\left(L_1^e, L_2^e\right)$. (d) The subtriangle in the transformed coordinate system $\left(L_1^s, L_2^e\right)$. (e) The final integration domain in the parameter coordinate system $\left(L_1^s, u\right)$.

Let us still use the integral kernel $f(\mathbf{r}, \mathbf{r}')/R$ and write the integration as

$$I = \int_{S'} \frac{f(\mathbf{r}, \mathbf{r}')}{\mathbf{R}} dS' \tag{2.80}$$

In the following, Duffy's method is applied to evaluate I. The Duffy transformation is illustrated in Figure 2.7. The sample points of the testing function on the original triangle (Figure 2.7a) can be chosen in the same way as in a normal integration element, for example, using the Gauss–Legendre rules. However, due to the use of the specific expression of the distance between two points on two triangle integration domains, numerical integration usually deals only with the first-order elements. Consequently, we need to transform the curvilinear triangular element into a triangular one. As shown in Figure 2.7, the curvilinear triangle in Cartesian coordinate (a) is transformed into an isosceles right triangle, denoted by T_0, in the parameter coordinate system (b). Then, the integration in (2.80) on the curvilinear triangular element is transformed into integration on T_0, written as

$$I = \int_{L_1=0}^{1} \int_{L_2=0}^{1-L_1} \frac{f(\mathbf{r}, \mathbf{r}')}{\mathbf{R}} J_1 dL_2 dL_1 \tag{2.81}$$

where J_1 is the Jacobian of the mapping, given by (2.79), and (L_1, L_2) are the area coordinates (ref. Chapter 3 for details.). Although the integration is done in the parameter coordinate system, the function $f(\mathbf{r}, \mathbf{r}')$ still involves the information about the curvilinear triangle in the Cartesian coordinate system. By using the mapping function, we can find the Cartesian coordinates of an arbitrary point p with parameter coordinates $\left(L_1^0, L_2^0, L_3^0\right)$, which further leads to the geometry information of this point on the curvilinear triangle. If p falls on the triangle that is both the source triangle and the observation triangle, singularity appears in the numerical integration involving p.

Consider the curvilinear triangular patch that corresponds to a triangle T_0 in the parameter coordinate system shown in Figure 2.7. Connecting p with the vertices of T_0 partitions T_0 into three subtriangles (T_1, T_2, T_3), as plotted in Figure 2.7b. Obviously, the integration on T_0 can be obtained as the summation of the integrations on the three subtriangles T_1, T_2 and T_3, that is,

$$I = \int_{T_0} \frac{f(\mathbf{r}, \mathbf{r}')}{R} dS = \sum_{e=1}^{3} \int_{T_e} \frac{f(\mathbf{r}, \mathbf{r}')}{R} dS \tag{2.82}$$

Apparently, the point p still causes a singularity in the integration on the subtriangles. Also, because the shapes of these subtriangles are different from that in Figure 2.4a, further transformation is necessary before Duffy transformation shall be applied. In order to transform the integration domain into the standard form for the Duffy transformation, let us introduce another parameter coordinate system $\left(L_1^e, L_2^e\right)$ $(e = 1, 2, 3)$ on subtriangle T_e. In Figure 2.7c, the parameter coordinate system $\left(L_1^1, L_2^1\right)$ on T_1 is illustrated for example. Then, under the parameter coordinate system $\left(L_1^e, L_2^e\right)$, I can be expressed by

$$I = \int_{L_1=0}^{1} \int_{L_2=0}^{1-L_1} \frac{f(\mathbf{r}, \mathbf{r}')}{R} J_1 dL_2 dL_1 = \sum_{e=1}^{3} \int_{L_1^e=0}^{1} \int_{L_2^e=0}^{1-L_1^e} \frac{f(\mathbf{r}, \mathbf{r}')}{R} J_1 J_2 dL_2^e dL_1^e \tag{2.83}$$

The relationship of $\left(L_1^e, L_2^e\right)$ and (L_1, L_2) can be expressed by:

$$\begin{bmatrix} L_1 \\ L_2 \\ L_3 \end{bmatrix} = [M_1] \begin{bmatrix} L_1^1 \\ L_2^1 \\ L_3^1 \end{bmatrix} = \begin{bmatrix} L_1^0 & 0 & 0 \\ L_2^0 & 1 & 0 \\ L_3^0 & 0 & 1 \end{bmatrix} \begin{bmatrix} L_1^1 \\ L_2^1 \\ L_3^1 \end{bmatrix} \tag{2.84}$$

$$\begin{bmatrix} L_1 \\ L_2 \\ L_3 \end{bmatrix} = [M_2] \begin{bmatrix} L_1^2 \\ L_2^2 \\ L_3^2 \end{bmatrix} = \begin{bmatrix} 1 & L_1^0 & 0 \\ 0 & L_2^0 & 0 \\ 0 & L_3^0 & 1 \end{bmatrix} \begin{bmatrix} L_1^2 \\ L_2^2 \\ L_3^2 \end{bmatrix} \tag{2.85}$$

$$\begin{bmatrix} L_1 \\ L_2 \\ L_3 \end{bmatrix} = [M_3] \begin{bmatrix} L_1^3 \\ L_2^3 \\ L_3^3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & L_1^0 \\ 0 & 1 & L_2^0 \\ 0 & 0 & L_3^0 \end{bmatrix} \begin{bmatrix} L_1^3 \\ L_2^3 \\ L_3^3 \end{bmatrix} \tag{2.86}$$

Correspondingly, the Jacobian J_2 for mapping from $\left(L_1^e, L_2^e\right)$ to (L_1, L_2) is

$$J_2 = |M_e| = L_e^0,\ e = 1, 2, 3 \tag{2.87}$$

We still cannot directly apply the Duffy transformation to the integration domain as shown in Figure 2.7c. Therefore, we need to further apply variable substitution. Introduce L_1^s

$$L_1^s = \left(1 - L_1^e\right) \tag{2.88}$$

The corresponding Jacobian is

$$J_3 = \left|\frac{\partial L_1^s}{\partial L_1^e}\right| = 1 \tag{2.89}$$

Substituting (2.88) into (2.83) leads to

$$I = \sum_{e=1}^{3} \int_{L_1^s=0}^{1} \int_{L_2^e=0}^{L_1^s} \frac{f(\mathbf{r}, \mathbf{r}')}{R} J_1 J_2 J_3 dL_2^e dL_1^s \tag{2.90}$$

The integration domain is shown in Figure 2.7d. By comparing with Figure 2.4a, we can directly apply the Duffy transformation to the integration domain in (2.90). Referring to the treatment of (2.69), let us introduce another variable

$$L_2^e = uL_1^s \tag{2.91}$$

whose Jacobian is

$$J_4 = \left|\frac{\partial L_2^e}{\partial u}\right| = L_1^s \tag{2.92}$$

Substituting (2.91) into (2.83) yields

$$I = \sum_{e=1}^{3} \int_{u=0}^{1} \int_{L_1^s=0}^{1} \frac{f(\mathbf{r}, \mathbf{r}')}{R} J_1 J_2 J_3 J_4 du dL_1^s \tag{2.93}$$

By comparing (2.80) with (2.93), it can be seen that the integration on the curvilinear triangle has been transferred to the parameter coordinate system

$\left(L_1^s, u\right)$, as shown in Figure 2.7e, with the corresponding Jacobian J being

$$J = J_1 J_2 J_3 J_4 = J_1 L_1^s L_e^0 \tag{2.94}$$

From the above discussion it can be seen that the essential of the Duffy transformation is using the Jacobian introduced by coordinate transformation to cancel the first-order singularity in the integration kernel. The expression for calculating the Jacobian is provided by (2.94). Next, we will investigate the expression of the integration kernel $f(\mathbf{r}, \mathbf{r}')/\mathbf{R}$ under the parameter coordinate system. Since $f(\mathbf{r}, \mathbf{r}')$ is nonsingular in the integration domain, here we examine the expression for R. In the subtriangle T_e, the distance $\mathbf{R}$ between an arbitrary point $\mathbf{r}$ with parametric coordinates $\left(L_1^e, L_2^e, L_3^e\right)$ with point p is

$$\mathbf{R} = \left| \sum_{i=1}^{3} L_i^e \mathbf{r}_i^e - \mathbf{r}_p \right| = \left| \sum_{i=1}^{3} L_i^e \left(\mathbf{r}_i^e - \mathbf{r}_p\right) \right| \tag{2.95}$$

where $\mathbf{r}_i^e$ is the position vector in T_e corresponding to the vertex $L_i^e = 1$. Then, from (2.88) and (2.91), we can obtain

$$L_1^e = 1 - L_1^s \tag{2.96}$$

Since $L_1^e + L_2^e + L_3^e = 1$, we have

$$L_3^e = 1 - L_1^e - L_2^e \tag{2.97}$$

Take $e = 1$ as an example. So, $\mathbf{r}_i^1 = \mathbf{r}_p$ Substituting (2.91), (2.96) and (2.97) into (2.95) yields

$$\begin{aligned} \mathbf{R} &= \left| \sum_{i=1}^{3} L_i^1 \left(\mathbf{r}_i^1 - \mathbf{r}_p\right) \right| \\ &= \left| (1 - L_1^s)\left(\mathbf{r}_1^1 - \mathbf{r}_p\right) + u L_1^s \left(\mathbf{r}_2^1 - \mathbf{r}_p\right) + L_1^s (1 - u)\left(\mathbf{r}_3^1 - \mathbf{r}_p\right) \right| \\ &= L_1^s \left| \left(\mathbf{r}_3^1 - \mathbf{r}_1^1\right) + u\left(\mathbf{r}_2^1 - \mathbf{r}_3^1\right) \right| \end{aligned} \tag{2.98}$$

Then, substituting (2.94) and (2.98) into (2.93) we obtain

$$\begin{aligned} I &= \sum_{e=1}^{3} \int_{u=0}^{1} \int_{L_1^s=0}^{1} \frac{f(\mathbf{r}, \mathbf{r}')}{L_1^s \left| \left(\mathbf{r}_3^1 - \mathbf{r}_1^1\right) + u\left(\mathbf{r}_2^1 - \mathbf{r}_3^1\right) \right|} J_1 L_1^s L_e^0 du dL_1^s \\ &= \sum_{e=1}^{3} \int_{u=0}^{1} \int_{L_1^s=0}^{1} \frac{f(\mathbf{r}, \mathbf{r}')}{\left| \left(\mathbf{r}_3^1 - \mathbf{r}_1^1\right) + u\left(\mathbf{r}_2^1 - \mathbf{r}_3^1\right) \right|} J_1 L_e^0 du dL_1^s \end{aligned} \tag{2.99}$$

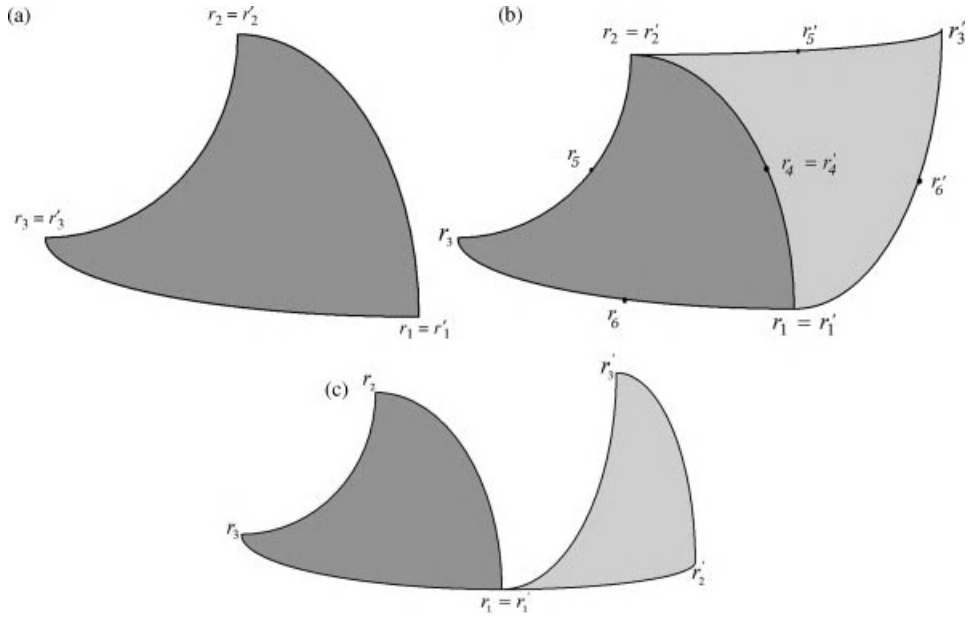


Figure 2.8 Singularity of different cases in the improved Duffy transformation method, when the field- and the source elements share (a) a common facet, (b) a common edge, or (c) a common vertex.

Equation 2.99 clearly shows that the Jacobian introduced by the Duffy transformation canceled the singularity in the integration kernel. As a result, we obtain a normal integration that can be dealt with standard numerical integration technique such as the Gaussian integration method.

III. *Improved Duffy transformation method*

In the above section, the implementation of Duffy's method is briefly introduced. The flaw of this method is that this technique is only applicable in treating the case when the source element and the field element coincide. For the singularity when the field point falls close to the source element, an improved Duffy transformation method is employed in order to achieve better computational accuracy [8–10].

As illustrated in Figure 2.8, in the improved Duffy transformation method, various singularity problems are handled differently according to the singularity case, namely, when the field element shares a common facet, a common edge or a common vertex with the source element. The basic idea is, to introduce proper coordinate transformation according to the singularity case, transform the singularity point of the original coordinate system to the origin of the new coordinate system, and then apply the Duffy transformation to cancel the singularity.

The MoM matrix element of a curvilinear triangle element in a Cartesian coordinate system is written as:

$$Z_{ij} = j\omega\mu \int_{T_i}\int_{T_j} \left[\mathbf{B}_i(\mathbf{r}) \cdot \mathbf{B}_j(\mathbf{r}') - \frac{1}{k^2} \nabla \cdot \mathbf{B}_i(\mathbf{r}) \nabla' \cdot \mathbf{B}_j(\mathbf{r}') \right] G(\mathbf{r}, \mathbf{r}') dS' dS \quad (2.100)$$

As we do in the conventional Duffy transformation, we first express the integration domain in the parameter coordinate system:

$$Z_{ij} = \int_{L_1=0}^{1} \int_{L_2=0}^{1-L_1} \int_{L_1'=0}^{1} \int_{L_2'=0}^{1-L_1'} \left[f\left(L_1, L_2; L_1', L_2'\right)/R \right] dL_2' dL_1' dL_2 dL_1 \quad (2.101)$$

where (L_1, L_2) and $\left(L_1', L_2'\right)$ correspond to the area (parametric) coordinates of the field and source triangles. $f\left(L_1, L_2; L_1', L_2'\right)/R$ is the integration kernel, and $f\left(L_1, L_2; L_1', L_2'\right)$ is nonsingular in the integration domain (Figure 2.9). For the singularity due to the common facet or common edge, the improved Duffy's method follows the procedure below:

1. Introduce the parameter coordinates for the source and the field integration domains.
2. Introduce relative coordinates that map the singular point to the coordinates origin. Express the inner integration domain (for the source element) in terms of the relative coordinates.
3. Change the integration sequence so that the integration about the relative coordinates becomes the outer integration.
4. Apply the Duffy transformation to the relative coordinates, so that the Jacobian cancels the singularity in the integration kernel and (2.101) becomes a normal integration.

First, for convenience, introduce another set of parameter coordinates $\left(\eta_1, \eta_2; \eta_1', \eta_2'\right)$ for both the field and the source integration domains (triangles):

$$\begin{cases} \eta_1 = 1 - L_1 \\ \eta_2 = 1 - L_1 - L_2 \\ \eta_1' = 1 - L_1' \\ \eta_2' = 1 - L_1' - L_2' \end{cases} \quad (2.102)$$

In the following, we will use I defined as

$$I = \int_{\eta_1=0}^{1} \int_{\eta_2=0}^{\eta_1} \int_{\eta_1'=0}^{1} \int_{\eta_2'=0}^{\eta_1'} \left[f\left(\eta_1, \eta_2; \eta_1', \eta_2'\right)/R \right] d\eta_2' d\eta_1' d\eta_2 d\eta_1 \quad (2.103)$$

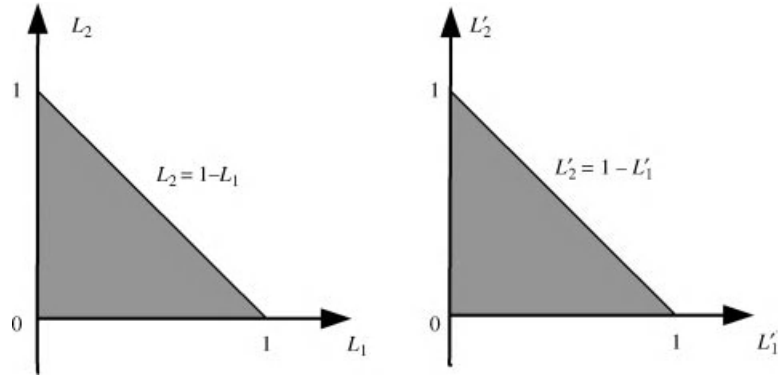


Figure 2.9 Curvilinear triangle elements in parameter coordinate system.

in explaining the singularity treatment. Apparently, (2.103) involves a fourfold integration, whose original integration domain is illustrated in Figure 2.10a, with corresponding coordinate systems $(\eta_1, \eta_2; \eta_1', \eta_2')$. We first exchange the integration sequence of η_2 and η_1':

$$I = \int_{\eta_1=0}^{1} \int_{\eta_1'=0}^{1} \int_{\eta_2=0}^{\eta_1} \int_{\eta_2'=0}^{\eta_1'} \left[f\left(\eta_1, \eta_2; \eta_1', \eta_2'\right)/R \right] d\eta_2' d\eta_2 d\eta_1' d\eta_1 \tag{2.104}$$

The integration domain is illustrated in Figure 2.10b. Obviously, a singularity exists when the field triangle coincides with the source one, that is,

$$\begin{cases} \eta_1' = \eta_1 \\ \eta_2' = \eta_2 \end{cases} \tag{2.105}$$

Define (u_1, u_2) as

$$\begin{cases} u_1 = \eta_1' - \eta_1 \\ u_2 = \eta_2' - \eta_2 \end{cases} \tag{2.106}$$

Then, the singular point is shifted to the origin of the coordinate system (u_1, u_2). Substitute (2.106) into (2.104) and the integration becomes

$$I = \int_{\eta_1=0}^{1} \int_{u_1=-\eta_1}^{1-\eta_1} \int_{\eta_2=0}^{\eta_1} \int_{u_2=-\eta_2}^{u_1+\eta_1-\eta_2} [f(\eta_1, \eta_2; u_1, u_2)/R] du_2 d\eta_2 du_1 d\eta_1 \tag{2.107}$$

In the following, we change the integration sequence from $(\eta_1, u_1; \eta_2, u_2)$ to $(u_1, u_2; \eta_1, \eta_2)$, which naturally causes the change of the integration domains:

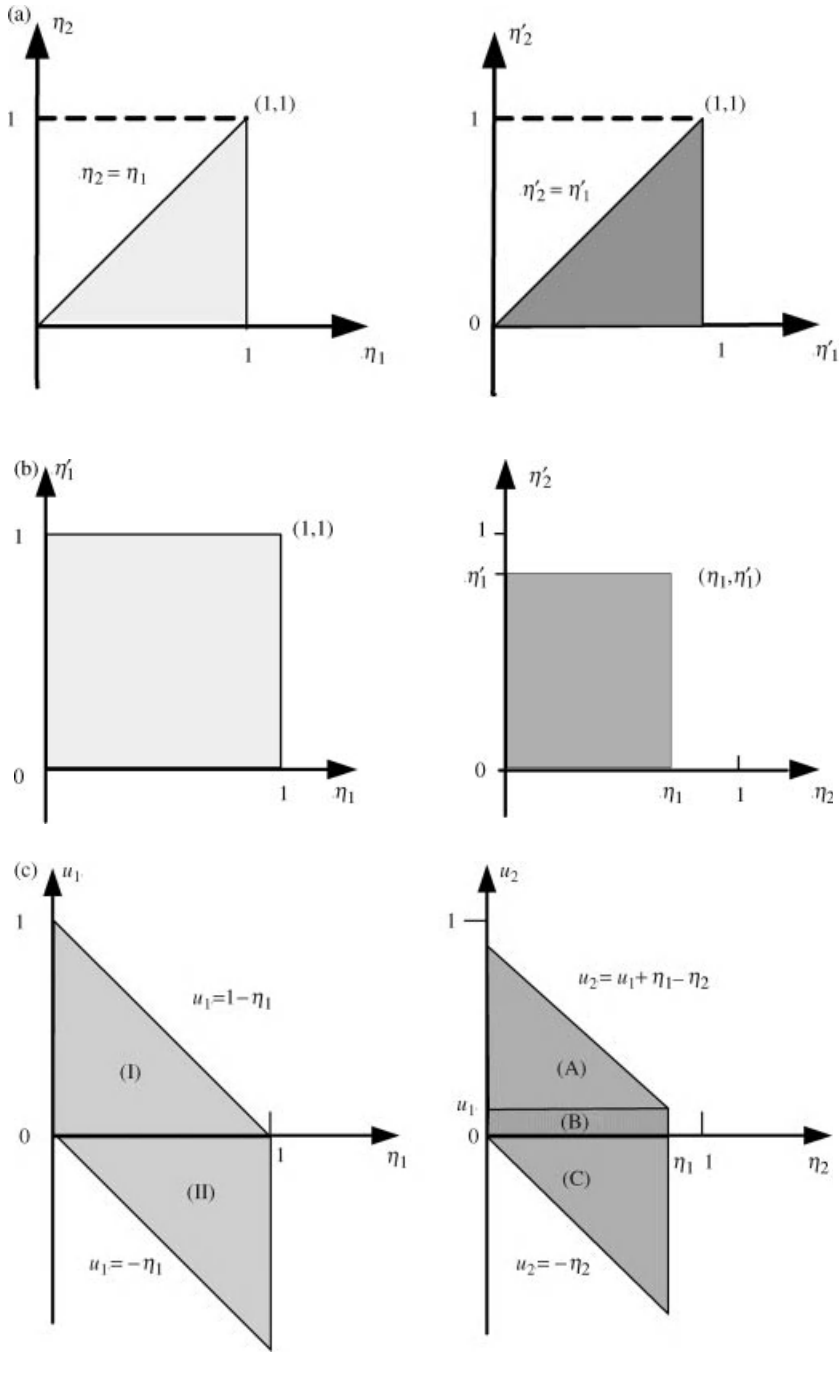


Figure 2.10 The improved Duffy's method.

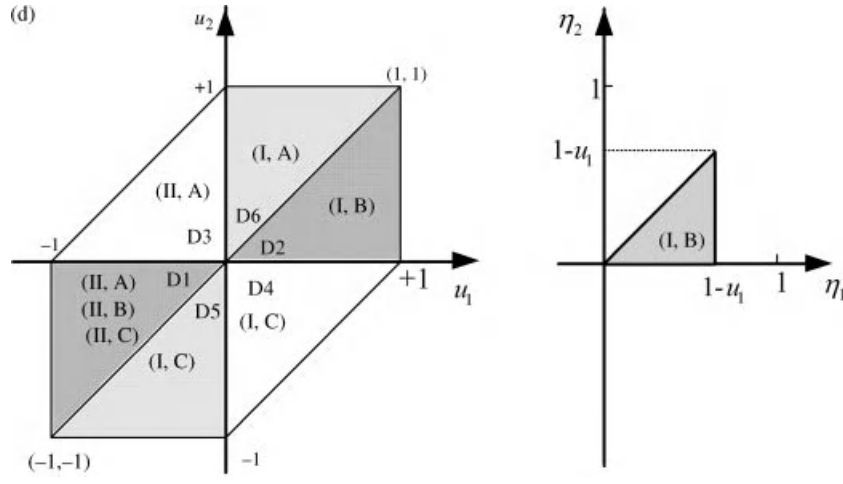


Figure 2.10 (*Continued*)

1. Partition the integration domain in the coordinate system (η_1, u_1) into two triangles, and denote them by I and II. Exchange the integration sequence of u_1 and η_1 on each of the triangle, resulting in:

$$\begin{aligned}\int_{\eta_1=0}^{1}\int_{u_1=-\eta_1}^{1-\eta_1} \ldots du_1 d\eta_1 &= \int_{\eta_1=0}^{1}\int_{u_1=-\eta_1}^{0} \ldots du_1 d\eta_1 + \int_{\eta_1=0}^{1}\int_{u_1=0}^{1-\eta_1} \ldots du_1 d\eta_1 \\ &= \int_{u_1=-1}^{0}\int_{\eta_1=-u_1}^{1} \ldots d\eta_1 du_1 + \int_{u_1=0}^{1}\int_{\eta_1=0}^{1-u_1} \ldots d\eta_1 du_1\end{aligned} \tag{2.108}$$

Analogously, partition the integration domain under coordinate system (η_2, u_2) into three subdomains, denoted by A, B, and C. Exchange integration sequence of u_2 and η_2 on each of the subdomain, leading to:

$$\begin{aligned}&\int_{\eta_2=0}^{\eta_1}\int_{u_2=-\eta_2}^{u_2+\eta_1-\eta_2} \ldots du_2 d\eta_2 \\ &\quad= \int_{\eta_2=0}^{\eta_1}\int_{u_2=-\eta_2}^{0} \ldots du_2 d\eta_2 + \int_{\eta_2=0}^{\eta_1}\int_{u_2=0}^{\eta_1} \ldots du_2 d\eta_2 \\ &\qquad+ \int_{\eta_2=0}^{\eta_1}\int_{u_2=\eta_1}^{u+\eta_1-\eta_2} \ldots du_2 d\eta_2 \\ &\quad= \int_{u_2=-\eta_1}^{0}\int_{\eta_2=-u_2}^{\eta_1} \ldots d\eta_2 du_2 + \int_{u_2=0}^{u_1}\int_{\eta_2=0}^{\eta_1} \ldots d\eta_2 du_2 \\ &\qquad+ \int_{u_2=u_1}^{u_1+\eta_1}\int_{\eta_2=0}^{u_1+\eta_1-u_2} \ldots d\eta_2 du_2\end{aligned} \tag{2.109}$$

After the above treatment, the original fourfold integration becomes a combination of two new double integrations composed of 2 and 3 subdomains under coordinate systems $(u_1, \eta_1; u_2, \eta_2)$, respectively, as shown in Figure 2.10c. Based on this, further exchange the integration sequence of η_1 and u_2, then we get the resultant integration on 6 (2×3) subdomains ((I,A), (I,B), (I, C), (II,A), (II,B) and (II,C)), as plotted in Figure 2.10d. The integration on each of the subdomain is denoted as $D_n (n = 1, 2, \cdots, 6)$:

$$D_n = \iiiint [f(u_1, u_2; \eta_1, \eta_2)/R] d\eta_2 d\eta_1 du_2 du_1 \tag{2.110}$$

Obviously, all D_n are triangular regions, with singular point corresponding to the origin of coordinate system (u_1, u_2). It is worth noting that in $(u_1, u_2; \eta_1, \eta_2)$, the integration domain corresponding to (η_1, η_2) are composed of six parts as well. In Figure 2.10d we only show the one corresponding to (I,B) for example.

Through the above operation, we have moved the singularity to the origin of the relative coordinate system, and have transformed the outer double integration to the relative coordinate system. By doing this, the integration is ready to be handled by the Duffy-transformation technique. In the following, we will show the details of the consequence techniques in terms of the singularity cases: common facet singularity, common edge and common point singularities.

A. *Treatment for common facet singularity*

An arbitrary position vector on a second-order curvilinear triangle can be expressed as

$$\mathbf{r}(L_1, L_2) = \sum_{i=1}^{6} \varphi_i(L_1, L_2) \mathbf{r}_i \tag{2.111}$$

$$\mathbf{r}'(L_1', L_2') = \sum_{i=1}^{6} \varphi_i(L_1', L_2') \mathbf{r}_i' \tag{2.112}$$

where $\mathbf{r}$ and $\mathbf{r}'$ denotes the field point and the source point, and φ_i is given in Table 2.2. As shown in Figure 2.8a, since the field and the source facets coincide,

$$\mathbf{r}_i = \mathbf{r}_i' \qquad i = 1, 2, \ldots, 6 \tag{2.113}$$

Table 2.2 Coefficients $\varphi_i(x_1, x_2)$(Common facet)

φ_1	$x_1(2x_1 - 1)$	φ_4	$4x_1x_2$
φ_2	$x_2(2x_2 - 1)$	φ_5	$4x_2(1 - x_1 - x_2)$
φ_3	$(1 - x_1 - x_2)(1 - 2x_1 - 2x_2)$	φ_6	$4x_1(1 - x_1 - x_2)$

Then, the distance between the field point and the source point is

$$\begin{aligned} R &= \left|\mathbf{r}(L_1, L_2)\text{-}\mathbf{r}'\left(L_1', L_2'\right)\right| \\ &= \left|\sum_{k=1}^{6}\left[\varphi_k(L_1, L_2) - \varphi_k\left(L_1', L_2'\right)\right]\mathbf{r}_k\right| \end{aligned} \tag{2.114}$$

From the definition of the parameter coordinate system, on the singular point (when the field point and the source point coincide), relations

$$\begin{cases} L_1' = L_1 \\ L_2' = L_2 \end{cases} \tag{2.115}$$

hold. Substituting (2.115) into (2.102) yields

$$\begin{cases} \eta_1' = \eta_1 \\ \eta_2' = \eta_2 \end{cases} \tag{2.116}$$

Then, by substituting (2.116) into (2.106) we obtain

$$\begin{cases} u_1 = 0 \\ u_2 = 0 \end{cases} \tag{2.117}$$

In other words, the singular point corresponds to the origin of coordinate system (u_1, u_2).

According to the symmetry of the integration domain, the integration on the six parts of the integration domain in Figure 2.10d can be combined into integration on three parts. For example, D_1 and D_2 can be combined as

$$\begin{aligned} I_1 &= D_1 + D_2 \\ &= \int_{u_1=0}^{1}\int_{u_2=0}^{u_1}\int_{\eta_1=0}^{1-u_1}\int_{\eta_2=0}^{\eta_1}[f(u_1, u_2; \eta_1, \eta_2) \\ &\quad + f(-u_1, -u_2; \eta_1 + u_1, \eta_2 + u_2)]/R d\eta_2 d\eta_1 du_2 du_1 \end{aligned} \tag{2.118}$$

Similarly,

$$\begin{aligned} I_2 &= D_3 + D_4 \\ &= \int_{u_1=0}^{1}\int_{u_2=u_1-1}^{0}\int_{\eta_1=-u_2}^{1-u_1}\int_{\eta_2=-u_2}^{\eta_1}[f(u_1, u_2; \eta_1, \eta_2) \\ &\quad + f(-u_1, -u_2; \eta_1 + u_1, \eta_2 + u_2)]/R d\eta_2 d\eta_1 du_2 du_1 \end{aligned} \tag{2.119}$$

$$\begin{aligned} I_3 &= D_5 + D_6 \\ &= \int_{u_1=0}^{1}\int_{u_2=u_1}^{1}\int_{\eta_1=u_2-u_1}^{1-u_1}\int_{\eta_2=0}^{\eta_1-(u_2-u_1)}[f(u_1, u_2; \eta_1, \eta_2) \\ &\quad + f(-u_1, -u_2; \eta_1 + u_1, \eta_2 + u_2)]1/R d\eta_2 d\eta_1 du_2 du_1 \end{aligned} \tag{2.120}$$

Thus, (2.104) can be rewritten as

$$I=\int_{\eta_1=0}^{1}\int_{\eta_1'=0}^{1}\int_{\eta_2=0}^{\eta_1}\int_{\eta_2'=0}^{\eta_1'} f\left(\eta_1,\eta_2;\eta_1',\eta_2'\right)/R d\eta_2' d\eta_2 d\eta_1' d\eta_1=\sum_{i=1}^{3} I_i \tag{2.121}$$

When the singularity is transferred to the vertex of a triangular domain, the Duffy transformation can be applied straightforwardly. Obviously, the integration domain corresponding to I_1 is a standard Duffy-transformation domain. We can directly introduce parameters (w,x) and let

$$\begin{cases} u_1=w \\ u_2=wx \end{cases} \tag{2.122}$$

The Jacobian is

$$J=\begin{vmatrix} \dfrac{\partial u_1}{\partial w} & \dfrac{\partial u_2}{\partial w} \\ \dfrac{\partial u_1}{\partial x} & \dfrac{\partial u_2}{\partial x} \end{vmatrix}=\begin{vmatrix} \dfrac{\partial w}{\partial w} & \dfrac{\partial (wx)}{\partial w} \\ \dfrac{\partial w}{\partial x} & \dfrac{\partial (wx)}{\partial x} \end{vmatrix}=w \tag{2.123}$$

and the integration becomes

$$\begin{aligned} I_1&=\int_{w=0}^{1}\int_{x=0}^{1}\int_{\eta_1=0}^{1-w}\int_{\eta_2=0}^{\eta_1}\begin{bmatrix} f(w,wx;\eta_1,\eta_2) \\ +f(-w,-wx;\eta_1+w,\eta_2+wx) \end{bmatrix}\frac{1}{R}w d\eta_2 d\eta_1 dx dw \\ &=\int_{w=0}^{1}\int_{x=0}^{1}\int_{L_1=U_1}^{V_1}\int_{L_2=U_2}^{V_2}\left[f\left(L_1,L_2;L_1',L_2'\right)+f\left(L_1',L_2';L_1,L_2\right)\right]\frac{w}{R}dL_2 dL_1 dx dw \end{aligned} \tag{2.124}$$

where U_1, V_1, U_2 and V_2 are listed in Table 2.3. Analogously, in subdomains D_4 and $D_{6,}$ corresponding to I_2 and I_3, introduce, respectively,

$$\begin{cases} u_1=wx \\ u_2=w(x-1) \end{cases} \tag{2.125}$$

and

$$\begin{cases} u_1=wx \\ u_2=w \end{cases} \tag{2.126}$$

(The reader is referred to [8] for more details.). It can be shown that the Jacobians of these two transformations also equal w. Substituting (2.125) and (2.126) into (2.119) and (2.120) respectively, yields:

$$I_2 = \int_{w=0}^{1}\int_{x=0}^{1}\int_{\eta_1=-u_2}^{1-u_1}\int_{\eta_2=-u_2}^{\eta_1} \times \begin{bmatrix} f(wx, w(x-1); \eta_1, \eta_2) \\ +f(-wx, -w(x-1); \eta_1 + wx, \eta_2 + w(x-1)) \end{bmatrix} \frac{w}{R} d\eta_2 d\eta_1 dx dw$$

$$= \int_{w=0}^{1}\int_{x=0}^{1}\int_{L_1=U_1}^{V_1}\int_{L_2=U_2}^{V_2} \left[f\left(L_1, L_2; L_1', L_2'\right) + f\left(L_1', L_2'; L_1, L_2\right) \right] \frac{w}{R} dL_2 dL_1 dx dw \tag{2.127}$$

$$I_3 = \int_{w=0}^{1}\int_{x=0}^{1}\int_{\eta_1=0}^{1-u_1}\int_{\eta_2=0}^{\eta_1} \begin{bmatrix} f(wx, w; \eta_1, \eta_2) \\ +f(-wx, -w; \eta_1 + wx, \eta_2 + w) \end{bmatrix} \frac{w}{R} d\eta_2 d\eta_1 dx dw$$

$$= \int_{w=0}^{1}\int_{x=0}^{1}\int_{L_1=U_1}^{V_1}\int_{L_2=U_2}^{V_2} \left[f\left(L_1, L_2; L_1', L_2'\right) + f\left(L_1', L_2'; L_1, L_2\right) \right] \frac{w}{R} dL_2 dL_1 dx dw \tag{2.128}$$

Because $f\left(L_1, L_2; L_1', L_2'\right)$ does not possess any singularities in the integration domain, we only need consider the term related to R. Substitute the above coordinate transformation into (2.114), on D_1 and D_2, the expression for R is

$$R = \left| \mathbf{r}(L_1, L_2) - \mathbf{r}'\left(L_1', L_2'\right) \right| = \left| \sum_{k=1}^{6} c_k \mathbf{r}_k \right| \tag{2.129}$$

where the interpolation coefficients c_k are listed in Table 2.3. It can be seen that c_k have a common factor w, which can be canceled by the Jacobian of the Duffy transformation. In the singular case (2.115), the expression R/w gives a positive value. As a result, the integration kernels in the transformed expressions (2.124), (2.127) and (2.128) are free of singularity.

Problem:

2.1 Explain why the remaining part of expression R except for w, is non-zero in the case of a singularity.

Table 2.3 Coefficients c_k of R (Common facet)

c_k\Area	D_1 and D_2	D_3 and D_4	D_5 and D_6
(L_1', L_2')	$L_1' = L_1 - w,$ $L_2' = L_2 + w(1-x)$	$L_1' = L_1 - wx, \quad L_2' = L_2 + w$	$L_1' = L_1 - wx,$ $L_2' = L_2 - w(1-x)$
(U_1, V_1)	$U_1 = w, V_1 = 1$	$U_1 = wx, V_1 = 1 + w(x-1)$	$U_1 = wx, V_1 = 1 - w(1-x)$
(U_2, V_2)	$U_2 = 0, V_2 = 1 - L_1$	$U_2 = 0, V_2 = 1 - L_1 - w(x-1)$	$U_2 = w(1-x), V_2 = 1 - L_1$
c_1	$w(-2w + 4L_1 - 1)$	$wx(-2wx + 4L_1 - 1)$	$wx(-2wx + 4L_1 - 1)$
c_2	$w(-1+x)(-2wx + 2w + 4L_2 - 1)$	$-w(4L_2 + 2w - 1)$	$-w(-1+x)(2wx - 2w + 4L_2 - 1)$
c_3	$wx(-2wx - 3 + 4L_1 + 4L_2)$	$w(-1+x)(-2wx + 2w + 4L_2 - 3 + 4L_1)$	$w(-2w - 3 + 4L_1 + 4L_2)$
c_4	$4w(w - wx - L_1 + L_1x + L_2)$	$-4w(-wx + L_1 - xL_2)$	$-4w(-wx^2 + wx - L_1 + L_1x - L_2x)$
c_5	$-4w(wx - wx^2 + 2L_2x + 1 - L_1 - L_2 - x + L_1x)$	$4w(-wx + w - xL_2 + 2L_2 - 1 + L_1)$	$4w(w - wx - 2L_2 + 1 - L_1 - x + L_1x + L_2x)$
c_6	$-4w(-wx + L_1x - 1 + L_1 + L_2)$	$-4w(wx - wx^2 + 2L_1x - L_1 - x + xL_2)$	$-4w(-wx + L_1 - x + L_1x + L_1x)$

B. *Treatment for common edge singularity*

In order to illustrate the treatment of singularity due to the fact that the field and source triangles share a common edge, let us consider the case as illustrated in Figure 2.8b.

$$\begin{cases} \mathbf{r}_1 = \mathbf{r}'_1 \\ \mathbf{r}_2 = \mathbf{r}'_2 \\ \mathbf{r}_4 = \mathbf{r}'_4 \end{cases} \tag{2.130}$$

The distance between the field and source points is

$$R = \left|\mathbf{r}(L_1, L_2)\text{-}\mathbf{r}'\left(L'_1, L'_2\right)\right| = \left|\sum_{k=1}^{6} c_k \mathbf{r}_k + c_7 \mathbf{r}'_3 + c_8 \mathbf{r}'_5 + c_9 \mathbf{r}'_6\right| \tag{2.131}$$

A singularity appears on the common edge. Therefore, the relationships of the area coordinates corresponding to singularity are

$$\begin{cases} L_1 = L'_1 \\ L_2 = L'_2 \\ L_3 = L'_3 = 0 \end{cases} \tag{2.132}$$

Substituting (2.132) into (2.102) and (2.106) yields $u_1 = u_2 = \eta_2 = 0$. Due to this relation, we need to introduce a three-dimensional (3D) Duffy transformation in (u_1, u_2, η_2) in order to eliminate the singularity. We will show the details of the 3D Duffy transformation by using the example of D_1.

$$D_1 = \int_{u_1=-1}^{0} \int_{u_2=u_1}^{0} \int_{\eta_1=-u_1}^{1} \int_{\eta_2=-u_2}^{\eta_1-(u_2-u_1)} f(u_1, u_2; \eta_1, \eta_2)/R d\eta_2 d\eta_1 du_2 du_1 \tag{2.133}$$

The integration domains of D_1 under coordinate systems (u_1, u_2), (η_1, η_2) are shown in Figures 2.11a and c, respectively. First, introduce

$$\begin{cases} z_1 = -u_1 \\ z_2 = -u_2 \end{cases} \tag{2.134}$$

Thus, the corresponding integration domain corresponding to (u_1, u_2) is changed into (z_1, z_2), as illustrated in Figure 2.11b. Exchange the integration sequence of η_1 and η_2, and the integration domain becomes that shown in Figure 2.11d. Introduce a new parameter $S = \eta_2 + (z_1 - z_2)$, then substitute the expressions for z_1, z_2 and S into (2.133) yields

$$D_1 = \int_{z_1=0}^{1} \int_{z_2=0}^{z_1} \int_{S=z_1}^{1} \int_{\eta_1=S}^{1} f(-z_1, -z_2; \eta_1, S - (z_1 - z_2))/R d\eta_1 dS dz_2 dz_1 \tag{2.135}$$

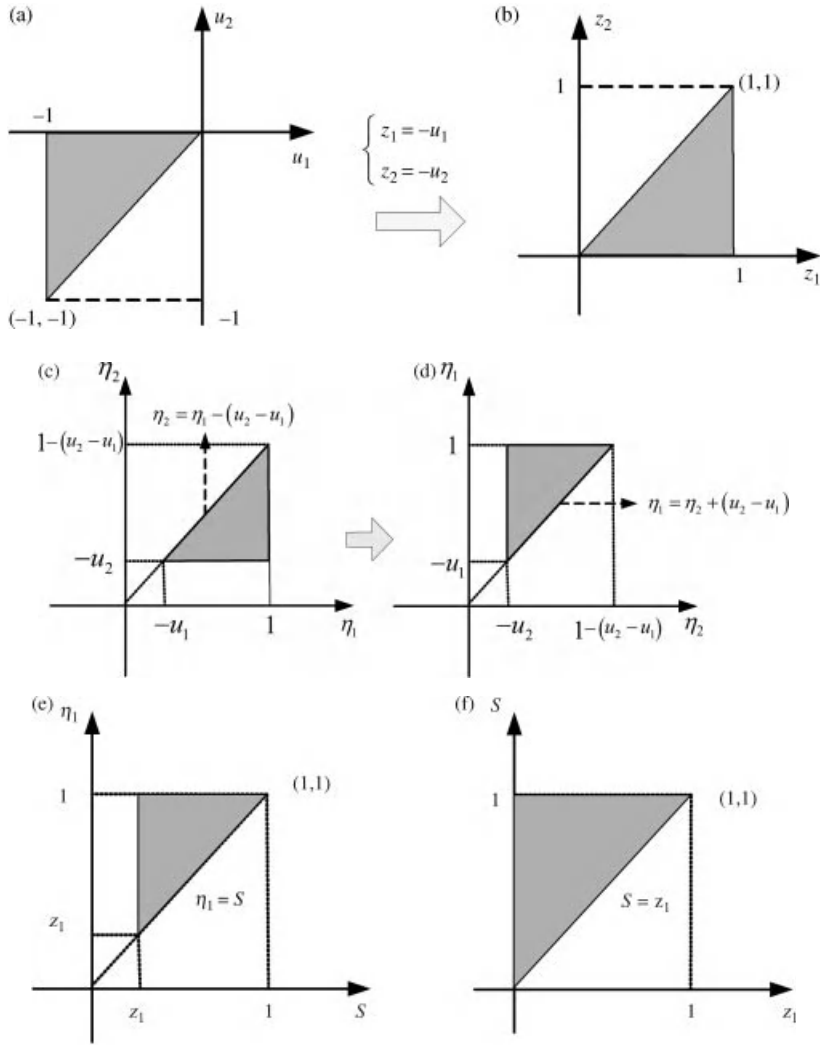


Figure 2.11 Coordinate transformation applied on D_1.

The corresponding integration domain is shown in Figure 2.11e. It can be seen that $z_1 \leq S \leq 1$, as plotted in Figure 2.11f.As shown in Figure 2.12, after the above transformation, the outer 3D integration domain of D_1 is a tetrahedron ABCD with a vertex on the coordinate origin. Due to the symmetry of the integration domain, we can change the integration sequence of D_1 to be

$$D_1 = \int_{S=0}^{1}\int_{z_1=0}^{S}\int_{z_2=0}^{z_1}\int_{\eta_1=S}^{1} f(-z_1,-z_2;\eta_1,S-(z_1-z_2))/R d\eta_1 dz_2 dz_1 dS \tag{2.136}$$

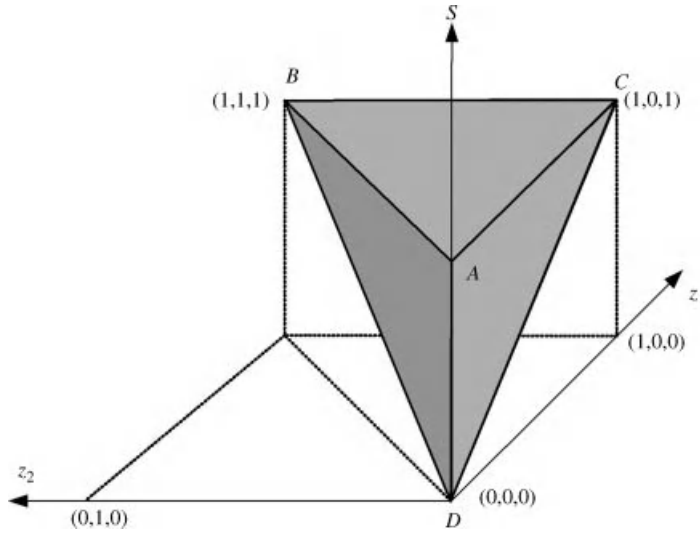


Figure 2.12 The outer 3D integration domain of D_1.

Then, introduce the following 3D Duffy transformation,

$$\begin{cases} S = w \\ z_1 = wx_1 \\ z_2 = wx_1x_2 \end{cases} \tag{2.137}$$

whose Jacobian J is expressed as

$$J = \left|\frac{\partial(S, z_1, z_2)}{\partial(w, x_1, x_2)}\right| = \begin{vmatrix} \dfrac{\partial S}{\partial w} & \dfrac{\partial z_1}{\partial w} & \dfrac{\partial z_2}{\partial w} \\ \dfrac{\partial S}{\partial x_1} & \dfrac{\partial z_1}{\partial x_1} & \dfrac{\partial z_2}{\partial x_1} \\ \dfrac{\partial S}{\partial x_2} & \dfrac{\partial z_1}{\partial x_2} & \dfrac{\partial z_2}{\partial x_2} \end{vmatrix} = w^2 x_1 \tag{2.138}$$

From the above process, we can find that the key point is to introduce new parameters (z_1, z_2, S) to express the relationship between (u_1, u_2, η_2) and obtain the integration domain as shown in Figure 2.12. In this way we can apply the Duffy transformation under parameter coordinate system (z_1, z_2, S). In Figures 2.13–2.17, the Duffy transformation on the remaining five integration domains are illustrated. Since on every one of these domains in coordinate system (z_1, z_2, S), the coordinates can be expressed in the same way as (2.137), the corresponding Jacobians are all $w^2 x_1$. The coefficients of R in (2.131), c_k are listed in Table 2.4. It can be seen that c_k share a common

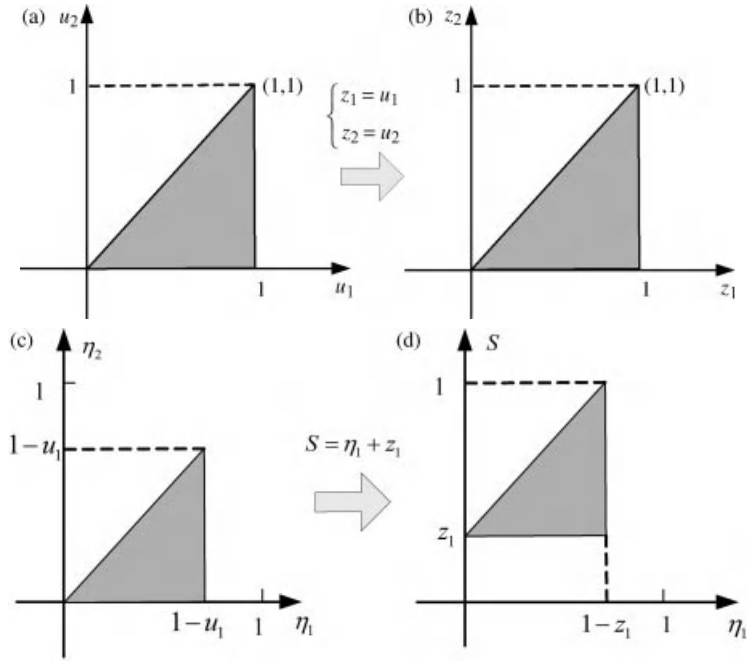


Figure 2.13 Transformation on D_2.

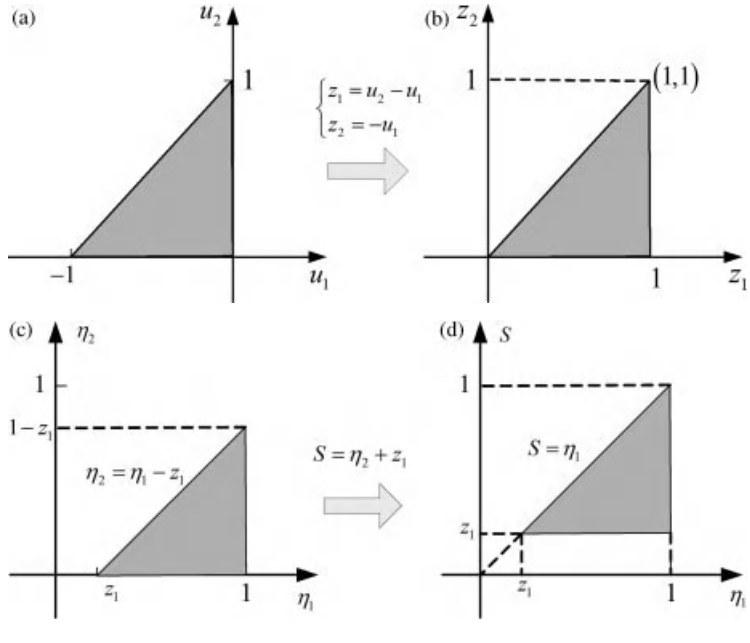


Figure 2.14 Transformation on D_3.

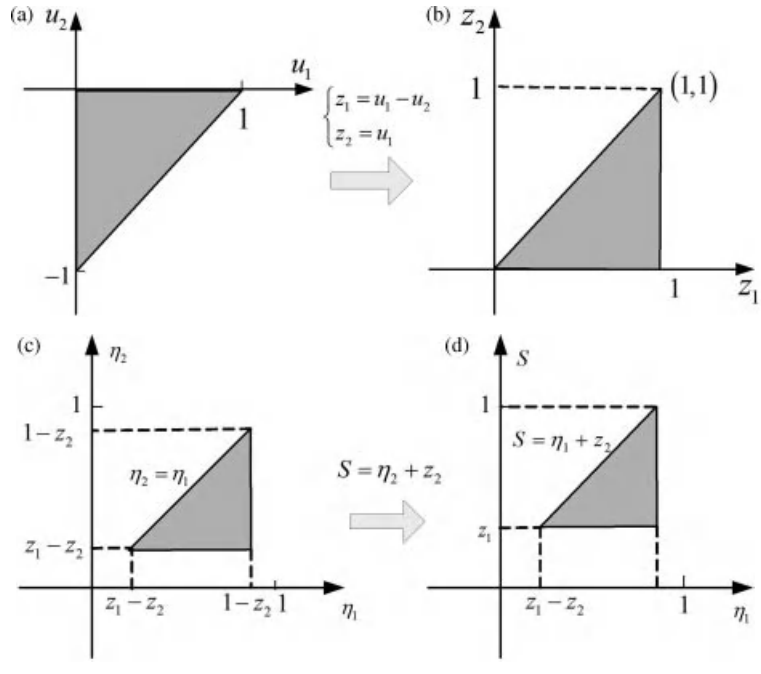


Figure 2.15 Transformation on D_4.

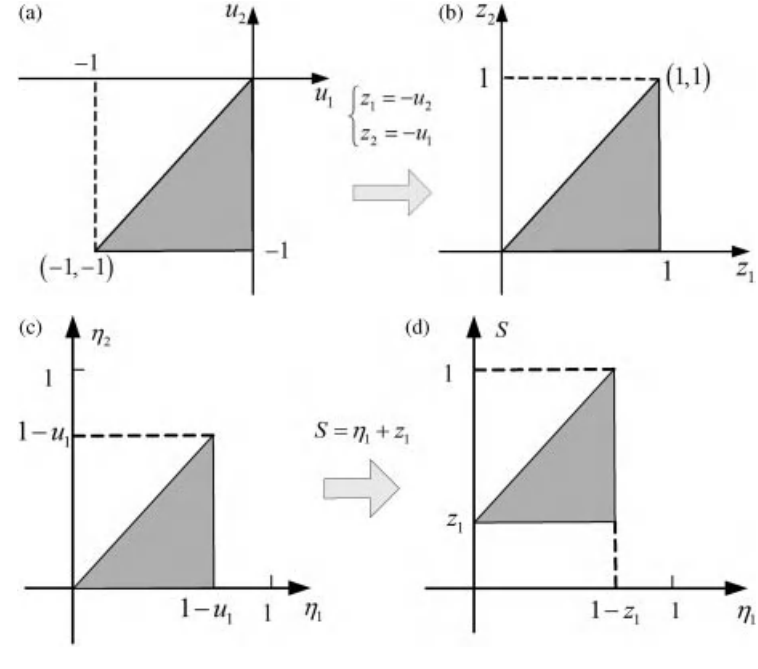


Figure 2.16 Transformation on D_5.

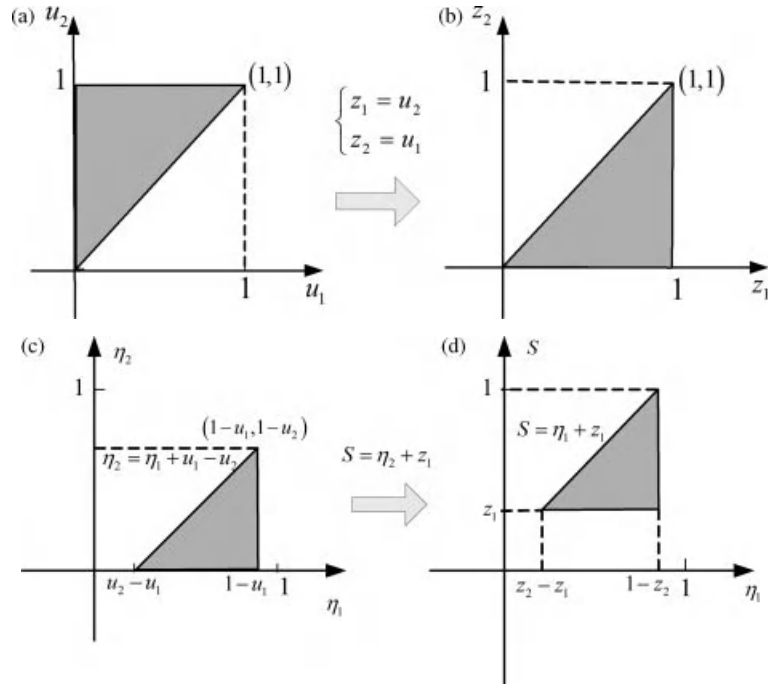


Figure 2.17 Transformation on D_6.

factor w, and the remaining parts are positive in the integration domain. Because the w in c_k is canceled by that in the Jacobian of the Duffy transformation, the resultant integration kernel is free of a singularity. Then, the final integration I can be expressed as:

$$I = \sum_{i=1}^{6} I_i = \sum_{i=1}^{6} \int_{x_1=0}^{1} \int_{x_2=0}^{1} \int_{w=0}^{1} \int_{L_1=V_i}^{U_i} \frac{w^2 x_1}{R} f\left(L_1, L_2, L_1', L_2'\right) dL_1 dw dx_2 dx_1 \tag{2.139}$$

where V_i and U_i can be found in Table 2.4.

C. *Treatment for common vertex singularity*

As for the singularity caused by the common vertex shared by the observation and the source facets, we can consider the case when

$$\mathbf{r}_1 = \mathbf{r}_1' \tag{2.140}$$

for example. The distance R between the field and source points on the second-order curvilinear element can be expressed as

$$R = \left|\mathbf{r}(L_1, L_2) - \mathbf{r}\left(L_1', L_2'\right)\right| = \left|\sum_{k=1}^{6} \left[c_k \mathbf{r}_k + c_{(k+6)} \mathbf{r}_k'\right]_6\right| \tag{2.141}$$

Table 2.4 Coefficients c_k of R (Common edge)

c_k\Area	D_1	D_2	D_3
(L_2, L_1', L_2')	$L_2 = 1 - L_1 - w(1 - x_1 + x_1x_2)$ $L_1' = L_1 + wx_1,$ $L_2' = 1 - L_1 - w$	$L_2 = 1 - L_1 - w(1 - x_1),$ $L_1' = L_1 - wx_1$ $L_2' = 1 - L_1 + w(2x_1 - x_1x_2 - 1)$	$L_2 = 1 - L_1 - w(1 - x_1)$ $L_1' = L_1 + wx_1x_2,$ $L_2' = 1 - L_1 - w$
(V_i, U_i)	$V_1 = 0, \quad U_1 = 1 - w$	$V_2 = wx_1, \quad U_2 = 1 - w(1 - x_1)$	$V_3 = 0, U_3 = 1 - w$
c_1	$-wx_1(2wx_1 + 4L_1 - 1)$	$-wx_1(2wx_1 - 4L_1 + 1)$	$-wx_2x_1(2wx_1x_2 + 4L_1 - 1)$
c_2	$wx_1(x_2 - 1)(2wx_1x_2 - 2wx_1 + 4w + 4L_1 - 3)$	$-wx_1(-1 + x_2)(2wx_1x_2 - 6wx_1 + 4L_1 - 3 + 4w)$	$-wx_1(-2wx_1 + 4w - 3 + 4L_1)$
c_3	$w(1 - x_1 + x_1x_2)(-1 + 2w - 2wx_1 + 2wx_1x_2)$	$w(-1 + x_1)(1 - 2w + 2wx_1)$	$w(-1 + x_1)(1 - 2w + 2wx_1)$
c_4	$-4wx_1(-w - 2L_1 + L_1x_2 + 1)$	$-4wx_1(-2wx_1 + wx_1x_2 + 2L_1 - L_1x_2 - 1 + w)$	$4wx_1(x_2w + L_1 - x_2 + L_1x_2)$
c_5	$-4w(-1 + L_1 + w - wx_1 + wx_1x_2) \times (1 - x_1 + x_1x_2)$	$-4w(1 - L_1 - w + wx_1)(-1 + x_1)$	$4w(-1 + L_1 + w - wx_1)(-1 + x_1)$
c_6	$4wL_1(1 - x_1 + x_1x_2)$	$-4wL_1(-1 + x_1)$	$-4wL_1(-1 + x_1)$
c_7	$-w(-1 + x_1)(1 + 2wx_1 - 2w)$	$-w(-x_1 + x_1x_2 + 1)(-1 - 2wx_1 + 2wx_1x_2 + 2w)$	$-w(x_2x_1 - 1)(1 + 2wx_1x_2 - 2w)$
c_8	$-4w(-1 + L_1 + w)(-1 + x_1)$	$4w(-1 + L_1 - 2wx_1 + wx_1x_2 + w) \times (-x_1 + x_1x_2 + 1)$	$-4w(-1 + L_1 + w)(x_2x_1 - 1)$
c_9	$4w(L_1 + wx_1)(-1 + x_1)$	$4w(-L_1 + wx_1)(-x_1 + x_1x_2 + 1)$	$4w(L_1 + wx_1x_2)(x_2x_1 - 1)$

Table 2.4 (*Continued*)

c_k\Area	D_4	D_5	D_6
(L_2, L_1', L_2')	$L_2 = 1 - L_1 - w(1 - x_1x_2)$ $L_1' = L_1 - wx_1x_2,$ $L_2' = 1 - L_1 + w(x_1x_2 + x_1 - 1)$	$L_2 = 1 - L_1 - w$ $L_1' = L_1 + wx_1x_2,$ $L_2' = 1 - L_1 - w(x_1x_2 - x_1 + 1)$	$L_2 = 1 - L_1 - w$ $L_1' = L_1 + wx_1x_2,$ $L_2' = 1 - L_1 - w(x_1x_2 - x_1 + 1)$
(V_i, U_i)	$V_4 = wx_1x_2, \quad U_4 = 1 - w(1 - x_1x_2)$	$V_5 = 0, \quad U_5 = 1 - w$	$V_6 = 0, \quad U_6 = 1 - w$
c_1	$-wx_1x_2(2wx_1x_2 - 4L_1 + 1)$	$-wx_1x_2(2wx_1x_2 + 4L_1 - 1)$	$-wx_1x_2(2wx_1x_2 - 4L_1 + 1)$
c_2	$-wx_1(4wx_1x_2 + 2wx_1 - 4L_1 + 3 - 4w)$	$-wx_1(x_2 - 1)(2wx_1x_2 - 2wx_1 - 3 + 4L_1 + 4w)$	$-wx_1(-1 + x_2)(2wx_1x_2 + 2wx_1 - 4w + 3 - 4L_1)$
c_3	$w(-1 + x_1x_2)(1 - 2w + 2wx_1x_2)$	$w(-1 + 2w)$	$w(-1 + x_1)(1 - 2w + 2wx_1)$
c_4	$4wx_1(wx_1x_2^2 + wx_1x_2 - L_1x_2 - L_1 + x_2 - wx_2)$	$4wx_1(-wx_1x_2 + x_1wx_2^2 - L_1 + 2L_1x_2 - x_2 + wx_2)$	$4wx_1(wx_2^2x_1 + L_1 - 2L_1x_2 + x_2 - wx_2)$
c_5	$-4w(1 - L_1 - w + wx_1x_2)(-1 + x_1x_2)$	$-4w(-1 + L_1 + w)$	$-4w(1 - L_1 - w + wx_1)(-1 + x_1)$
c_6	$-4wL_1(-1 + x_1x_2)$	$4wL_1$	$-4L_1w(-1 + x_1)$
c_7	$-w(x_1 - 1)(1 + 2wx_1 - 2w)$	$-w(-1 + x_1)(1 - 2w + 2wx_1)$	$-w(-1 + 2w)$
c_8	$4w(1 - L_1 + wx_1x_2 + wx_1 - w)(x_1 - 1)$	$-4w(-1 + L_1 + w - wx_1 + wx_1x_2) \times (-1 + x_1)$	$-4w(1 - L_1 + wx_1x_2 - w)$
c_9	$-4w(-L_1 + wx_1x_2)(x_1 - 1)$	$4w(L_1 + wx_1x_2)(-1 + x_1)$	$4w(-L_1 + wx_1x_2)$

In this case, a singularity only appears on the vertex expressed mathematically as

$$\begin{cases} L_1 = L_1' = 1 \\ L_2 = L_2' = 0 \\ L_3 = L_3' = 0 \end{cases} \tag{2.142}$$

Substituting (2.142) into (2.102) yields

$$\eta_1 = \eta_2 = \eta_1' = \eta_2' = 0 \tag{2.143}$$

Apparently, the singular point (2.143) already corresponds to the origin of the four-dimensional (4D) coordinate system. Therefore, we don't need to introduce relative parameters, and the next task is to transform the integration domain in order to apply 4D Duffy transformation. To do this, we first exchange the integration sequence of η_2 and η_1', which gives

$$\begin{aligned} I &= \int_{\eta_1=0}^{1}\int_{\eta_2=0}^{\eta_1}\int_{\eta_1'=0}^{1}\int_{\eta_2'=0}^{\eta_1'}\left[f\left(\eta_1,\eta_2;\eta_1',\eta_2'\right)/R\right]d\eta_2'd\eta_1'd\eta_2d\eta_1 \\ &= \int_{\eta_1=0}^{1}\int_{\eta_1'=0}^{1}\int_{\eta_2=0}^{\eta_1}\int_{\eta_2'=0}^{\eta_1'}\left[f\left(\eta_1,\eta_2;\eta_1',\eta_2'\right)/R\right]d\eta_2'd\eta_2d\eta_1'd\eta_1 \end{aligned} \tag{2.144}$$

The integration domain after this transformation is illustrated in Figure 2.18a. Partitioning this domain by using the tridiagonal line into two parts as shown in Figure 2.18b, then (2.144) can be further rewritten as:

$$\begin{aligned} I &= \int_{\eta_1=0}^{1}\int_{\eta_1'=0}^{1}\int_{\eta_2=0}^{\eta_1}\int_{\eta_2'=0}^{\eta_1'}\left[\frac{f\left(\eta_1,\eta_2;\eta_1',\eta_2'\right)}{R\left(\eta_1,\eta_2;\eta_1',\eta_2'\right)}\right]d\eta_2'd\eta_2d\eta_1'd\eta_1 \\ &= \int_{\eta_1=0}^{1}\int_{\eta_1'=0}^{\eta_1}\int_{\eta_2=0}^{\eta_1}\int_{\eta_2'=0}^{\eta_1'}\left[\frac{f\left(\eta_1,\eta_2;\eta_1',\eta_2'\right)}{R\left(\eta_1,\eta_2;\eta_1',\eta_2'\right)}\right]d\eta_2'd\eta_2d\eta_1'd\eta_1 \\ &\quad + \int_{\eta_1'=0}^{1}\int_{\eta_1=0}^{\eta_1'}\int_{\eta_2=0}^{\eta_1}\int_{\eta_2'=0}^{\eta_1'}\left[\frac{f\left(\eta_1,\eta_2;\eta_1',\eta_2'\right)}{R\left(\eta_1,\eta_2;\eta_1',\eta_2'\right)}\right]d\eta_2'd\eta_2d\eta_1d\eta_1' \\ &= \int_{\eta_1=0}^{1}\int_{\eta_1'=0}^{\eta_1}\int_{\eta_2=0}^{\eta_1}\int_{\eta_2'=0}^{\eta_1'}\left[\frac{f\left(\eta_1,\eta_2;\eta_1',\eta_2'\right)}{R\left(\eta_1,\eta_2;\eta_1',\eta_2'\right)}+\frac{f\left(\eta_1',\eta_2';\eta_1,\eta_2\right)}{R\left(\eta_1',\eta_2';\eta_1,\eta_2\right)}\right]d\eta_2'd\eta_2d\eta_1'd\eta_1 \end{aligned} \tag{2.145}$$

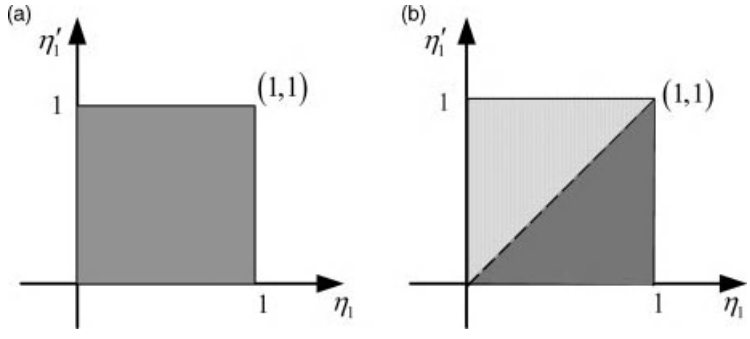


Figure 2.18 Partition of the integration domain.

Note that in (2.145), we have made use of the symmetry of the parameters and the integration domain. In the final integration domain of (2.145), we can apply the following 4D Duffy transformation:

$$\begin{cases} \eta_1 = w \\ \eta_2 = wx_1 \\ \eta_1' = wx_2 \\ \eta_2' = wx_2x_3 \end{cases} \tag{2.146}$$

whose Jacobian J is expressed as

$$J = \left|\frac{\partial(\eta_1, \eta_2, \eta_1', \eta_2')}{\partial(w, x_1, x_2, x_3)}\right| = \det\left(\begin{vmatrix} \frac{\partial\eta_1}{\partial w} & \frac{\partial\eta_1}{\partial x_1} & \frac{\partial\eta_1}{\partial x_2} & \frac{\partial\eta_1}{\partial x_3} \\ \frac{\partial\eta_2}{\partial w} & \frac{\partial\eta_2}{\partial x_1} & \frac{\partial\eta_2}{\partial x_2} & \frac{\partial\eta_2}{\partial x_3} \\ \frac{\partial\eta_1'}{\partial w} & \frac{\partial\eta_1'}{\partial x_1} & \frac{\partial\eta_1'}{\partial x_2} & \frac{\partial\eta_1'}{\partial x_3} \\ \frac{\partial\eta_2'}{\partial w} & \frac{\partial\eta_2'}{\partial x_1} & \frac{\partial\eta_2'}{\partial x_2} & \frac{\partial\eta_2'}{\partial x_3} \end{vmatrix}\right) = w^3x_2 \tag{2.147}$$

Substituting (2.146) and (2.147) into (2.144) yields

$$I = \int_{w=0}^{1}\int_{x_1=0}^{1}\int_{x_2=0}^{1}\int_{x_3=0}^{1}\left[\frac{f(L_1, L_2; L_1', L_2')}{R(L_1, L_2; L_1', L_2')} + \frac{f(L_1', L_2'; L_1, L_2)}{R(L_1', L_2'; L_1, L_2)}\right] w^3x_2 dx_3 dx_2 dx_1 dw \tag{2.148}$$

Substituting (2.146) into (2.141) we can obtain c_k:

$$\begin{cases} c_1 = -w(-1+x_2)(2wx_2-3+2w) \\ c_2 = w(-1+x_1)(-2w+2wx_1+1) \\ c_3 = wx_1(-1+2wx_1) \\ c_4 = 4w(-1+w)(-1+x_1) \\ c_5 = -4w^2(-1+x_1)x_1 \\ c_6 = -4w(-1+w)x_1 \end{cases} \begin{cases} c_7 = 0.0 \\ c_8 = -wx_2(-1+x_3)(-2wx_2+2wx_2x_3+1) \\ c_9 = -wx_2x_3(-1+2wx_2x_3) \\ c_{10} = -4w(-1+wx_2)x_2(-1+x_3) \\ c_{11} = 4w^2{x_2}^2(-1+x_3)x_3 \\ c_{12} = 4w(-1+wx_2)x_2x_3 \end{cases} \tag{2.149}$$

From (2.149), we can see that c_k share a common factor w, and R/w is positive in the integration domain. Since this factor w in R can be canceled by the Jacobian of the Duffy transformation, the resultant integration kernel is free of a singularity. In this way, the original singular integration is transformed into a normal integration under the coordinate system (w, x_1, x_2, x_3).

IV. *Numerical results*

From the above two sections, we can see that the conventional Duffy's method only deals with the singularity caused by the common facets, while the improved Duffy's method also deals with the singularity caused by the common edge and common vertex. Naturally, we expect an improved accuracy from the improved Duffy's method. To verify this, the RCS of a PEC sphere with diameter of 10λ under homogeneous plane wave incidence is calculated by using these two Duffy's methods. In both numerical experiments, hierarchical basis functions of 1.5 order are adopted and GMRES is used as the solver, with the convergence threshold being 10^{-3}. The root-mean-square (RMS) errors of the bistatic RCS are compared in Figure 2.19. It can be clearly seen that the improved Duffy's method provides better numerical accuracy in the results.

2.1.6 Comparison of EFIE and MFIE

As have been pointed out in Section 2.1.4, that EFIE is a Fredholm integral equation of the first kind, while MFIE is a Fredholm integral equation of the second kind. Therefore, the discretized matrix derived from MFIE bears a better condition number than that from EFIE. If we solve the final discretized equation using the iteration method, the MFIE solution converges much faster than the EFIE one. This is the major difference between the EFIE and the MFIE in terms of numerical behavior. Besides, they are also different in other senses, as presented in the following.

If the object is so thin that its thickness can be ignored, then the discretized function of this object for the unknown equivalent electric currents on the upper surface equals its counterpart on the lower surface. As a result, the coefficient matrix of the final

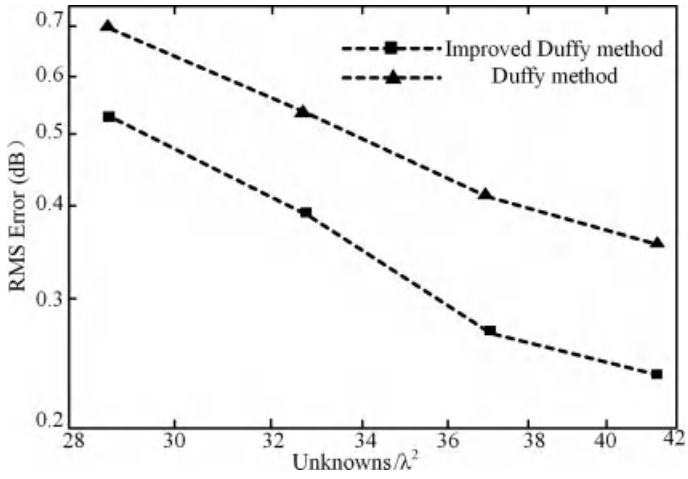


Figure 2.19 Numerical accuracy of the conventional and the improved Duffy's methods.

equation is always singular and we can hardly get any reasonable solution. However, it can be shown after careful analysis that the coefficients for the upper surface are identical to those for the lower surface. So we can introduce a new variable defined as the sum of the current on the upper surface and its counterpart on the lower surface. In this way, the number of the variables reduces from n to $n/2$. Because the $n/2$ equations corresponding to the currents on the upper surface are linearly independent, we can solve the newly introduced variables. Then, the scattered field can be calculated by using these variables. On the other hand, the above is not valid for the discretized MFIE. This is because there are two terms in the discretized MFIE. In one term, the coefficients for the upper surface are identical to those for the lower surface, while in the other, the aforementioned coefficients have the same absolute values but opposite signs. Consequently, we cannot introduce a variable to reduce the number of the unknowns. To conclude, only EFIE can be applied to the modeling of a very thin object.

Next, let us consider the low-frequency case. As is well known, there are two terms in operator L, namely, a vector one and a scalar one. When the frequency is very low, the vector term is close to zero, and the scalar one dominates. As can be seen from the mathematical expression of L, the value of the scalar term depends only on the divergence of the electric current source. It is known that the divergence of a curl is zero. This means when the frequency is very low, the electric current source may no longer be unique in EFIE. The ill-conditioned EFIE matrix is a reflection of this nonuniqueness. This problem does not exist in MFIE. Therefore, MFIE is more suitable for the low-frequency problems. Of course EFIE can still be used, subject to some special treatment. The reader is referred to Section 10.6 in [2] for more details. It should be noted that the low frequency referred to in this book is with respect to the electrical size of the segments in the meshes. More specifically, the "low-frequency problem" corresponds to

the case in which the smallest discretized dimension is much smaller than 0.1λ. It cannot be determined solely by the value of the operating frequency.

Another remark worth mentioning here is that usually the MoM based on EFIE shows better accuracy than that based on MFIE, especially when the mesh size is not sufficiently fine. The reason is as follows. The singularity residual term in MFIE is not zero. Accurate calculation of this term involves the calculating of the solid angle, which is highly related to the boundary geometry of the object. However, in most practical computations this angle is assumed to be 2π.

2.1.7 *Interior Resonance Problem*

From the uniqueness theorem introduced in Chapter 1, we know that the EM field cannot be uniquely determined from only the tangential components on the boundary for a lossless region. However, the above EFIE or MFIE is developed by using only the tangential electric or magnetic field on the boundary. This fact leads to the interior resonance problem in MoM. To be specific, the discretized EFIE and MFIE are singular at the resonant frequencies of the cavity formed by the boundary of an object. A good review about various methods in diminishing the interior resonance is given in Chapter 6 in [2]. In the following, we only introduce the most convenient and widely used scheme. We can simply combine EFIE and MFIE in the following way

$$\left(\alpha\left[P^{TE}\right] + Z(1-\alpha)\left[P^{TH}\right]\right)\{J\} = \left\{\alpha b^{TE} + Z(1-\alpha)b^{TH}\right\} \tag{2.150}$$

where α is a weight between 0 and 1. Multiplying MFIE with the characteristic impedance makes the addends in the two matrices comparable. Normally (2.150) is called the *combined field integral equation (CFIE)*. Next, we will show why the CFIE can diminish the interior resonance. For convenience, we choose $\alpha = 0.5$. The above function is actually constructed based on the following boundary condition

$$\mathbf{E}|_t + Z\hat{n}^+ \times \mathbf{H} = 0 \tag{2.151}$$

where $\hat{n}^+$ denotes an outwards unit vector, while $\hat{n}^-$ in the following denotes an inwards one. According to $\mathbf{J} = \sigma\mathbf{E}$ and $\mathbf{J} = \hat{n}^- \times \mathbf{H}$, $\sigma = 1/Z$ can be deduced from (2.151). This indicates a lossy boundary, resulting in complex resonant frequencies in the enclosed region. Since the computed frequencies are all real numbers, the interior resonance problem is circumvented. This is an explanation. The rigorous proof is presented in the following. Calculating the integral of the dot product of equation (2.151) with its conjugate over the problem region, yields,

$$\int_S |\mathbf{E}_t|^2 + Z^2|\mathbf{H}_t|^2 dS + 2Z\left\{\text{Real}\int_S [\mathbf{E}\times\mathbf{H}^*]\cdot(-\hat{n}^+)dS\right\} = 0 \tag{2.152}$$

The second integrand is the Poynting vector denoting the energy flowing into the interested region, which yields a non-negative value. So, in the vicinity of the boundary inside the region, $\mathbf{E}_t$ and $\mathbf{H}_t$ are simultaneously zeros, which indicates the interior resonance problem does not exist.

As a result, the CFIE (2.150) is popularly used in MoM. Numerical experiments show that when α equals 0.2, (2.150) gives a fairly good condition number and relatively fast convergence. As α increases, the convergence becomes slower but the accuracy increases. This is easy to explain from the analysis in Section 2.1.6: EFIE leads to higher accuracy and a larger condition number in the discretized equation, while MFIE leads to lower accuracy and a smaller condition number in the discretized equation.

2.1.8 Fast Multipole Method

In previous sections, the procedure is discussed on discretizing the integral equations to obtain a well-behaved system of linear equations (or linear system). Next, we will show how to solve the linear system rapidly. There are two types of algorithms for the numerical solution of a linear system, namely, the direct methods such as the Gaussian elimination method, and the iterative methods such as the conjugate gradient (CG) method. Denote the number of degrees of freedom (unknowns) by N. Because the coefficient matrix of the linear system is a full matrix, the direct method requires computer memory of $O(N^2)$ and operations of $O(N^3)$. Alternatively, the iteration method requires computer memory of $O(N^2)$ and operations of $O(N^2)$ for each iteration. Such large amounts of computer resource requirement has restricted the application of MoM so greatly, that before the 1990s, MoM was only applied to electrically small objects. After the 1990s, things changed and now MoM is applicable to electrically large problems as well. This change is mainly attributed to the *fast multipole method* (FMM) proposed in [13] by Rokhlin. This technique on the application of electromagnetic problems was discussed in great detail in [14] and was successfully applied to CEM in [15]. In the following, we will introduce this technique, which greatly reduces the computer memory requirement and accelerates the matrix-vector multiplication.

2.1.8.1 Basic Idea of Fast Multipole Method

As we know, the amount of calculation in the iteration method depends mainly on the calculation of matrix-vector multiplication. From the physical point of view, the aforementioned multiplication is the action of the source on the field. As in MoM, when the field points coincide with the field points, this action can also be considered as the interactions of equivalent electric currents. The first step of the fast multipole method is to divide the unknown equivalent currents into subgroups. This can be done in the following way: First, define a cube that can just enclose the object. Second, partition this cube into smaller boxes (the size of the box is to be discussed in detail in the following).

Figure 2.20 The conventional implementation of the far interaction.

We index and record only those boxes cut by the object boundary, because currents do not exist in the remaining boxes. It is obvious that the computational cost depends only on the nonempty boxes. For any nonempty box, all other boxes are interacted with it in either one of the following two manners: near interaction and far interaction. Generally speaking, when the centers of two boxes are no more than half wavelength away, the two boxes are near interacted. Otherwise they are far interacted.

Let us now analyze the far interaction of boxes A and B. Suppose there are 100 unknowns in A and B, respectively, as illustrated in Figure 2.20. If we use the conventional way to evaluate the interaction between them, 100×100 operations are required. FMM provides a new way to perform the far interaction. The basic idea is to separate the interaction into three steps: aggregation, translation, and disaggregation. First, the equivalent currents corresponding to the 100 unknowns are aggregated into the center of A. By doing this, we obtain a new set of functions with the following important property: the far interaction of all the equivalent currents in A to a field point in the far zone can be performed by "translation" and "disaggregation" of these new sets of functions. This is called aggregation. In the translation step, the functions in the center of A are "transferred" to the center of B. Then, in the disaggregation step, they are "distributed" to the electric currents corresponding to the 100 unknowns, which complete the fast multipole operation evaluating the far interaction between A and B. This fast multipole scheme is illustrated in Figure 2.21. In the next section, we will show that the plane wave function has the required translation property. Moreover, fewer plane waves are usually needed than the original unknowns. In other words, the

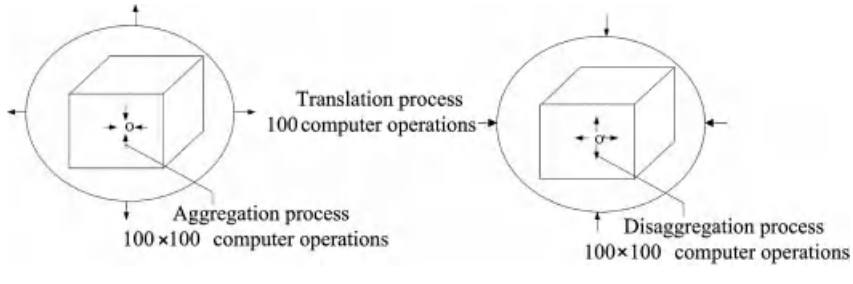


Figure 2.21 The FMM implementation of the far interaction.

translation of the new functions between the center of A and B requires less than 100 operations. This is the mechanism in FMM to accelerate the far interaction between A and B. The separation of the interaction origins from the multipole expansion of the Green's function, thus comes the name. Because it is very hard to achieve a satisfactory accuracy in the multipole expansion of the Green's function in performing the near interaction, this scheme works only for the far-interaction case. That is the reason we group the interaction into the near and the far ones.

Next, let us consider the remaining question: what is the proper size of the box. From the previous analysis we know that the number of calculations is linear with the number of unknowns in the process of translation. However, it has almost a quadratic dependence of the number of unknowns in the process of aggregation or disaggregation. Actually, to aggregate 100 unknowns into 100 plane waves requires 100×100 operations. On the other hand, to disaggregate 100 plane waves to 100 unknowns also requires 100×100 operations. Therefore, if we only consider the far interactions of only two boxes, the computation cost in FMM will exceed that in the conventional method. However, when we consider the far interactions of 100 boxes, it will change. Since there are $100 \times 100 = 10000$ unknowns, 10000×10000 operations are involved in the conventional cost method. On the other hand, by applying FMM, each box requires 100×100 operations in aggregation. Since there are 100 boxes, $100 \times 100 \times 100$ operations are required in the process of aggregation. Obviously, the process of disaggregation also involves $100 \times 100 \times 100$ operations. As to the translation, another $100 \times 100 \times 100$ operations are necessary since there are 100 boxes and 100×100 translations each requiring 100 operations will be performed. As a result, the FMM performs the total interactions using three $100 \times 100 \times 100$ operations, which is a significant saving in the operations compared to the conventional way. As can be seen, the box cannot be too large, since too large size of the box will cause excessive operations in the aggregation and the disaggregation processes. Meanwhile, it cannot be too small; otherwise the operations in the translation will be overwhelming. Strictly speaking, if we have N unknowns in total, and partition them into M groups, then each group supports approximately N/M unknowns. According to the above analysis, the aggregation and the disaggregation require computer operations of $O(N^2/M)$ each, while the translation step involves $O(MN)$ operations. So the total interaction needs operations of $O(N^2/M + MN)$. It can be shown that when $M = N^{1/2}$, the amount of operations is minimized as $O(N^{3/2})$. With this choice of M, the computer memory requirement is also of $O(N^{3/2})$.

2.1.8.2 Mathematic Formulation of FMM

As introduced in Section 2.1.4, the entries in the coefficient matrix of the discretized integral equation (DIE) can be expressed as:

$$Z_{ij} = \int_S \int_{S'} \mathbf{g}_i \cdot P(G) \mathbf{g}_j dS' dS \tag{2.153}$$

where $P(G)$ is imposed on the Green's function G. Suppose $\{x\}$ and $\{y\}$ denote the unknowns in two boxes (A and B) that are far from each other. Then, the action of $\{x\}$ on $\{y\}$ can be expressed as

$$\{y\} = [Z]\{x\} \tag{2.154}$$

From the previous section we know that the FMM divide the matrix-vector multiplication into three concrete steps, viz., the aggregation, translation and disaggregation. In this section, more details will be presented. The aforementioned division rests on two elementary identities. The first, an expansion of the kernel in the integral in (2.153), is a form of addition theorem of Green's function (Gegenbauer's addition theorem) [16],

$$\frac{e^{-jk|\mathbf{r}+\mathbf{d}|}}{|\mathbf{r}+\mathbf{d}|} = -jk\sum_{l=0}^{\infty}(-1)^l(2l+1)j_l(kd)h_l^{(2)}(kr)P_l(\hat{d}\cdot\hat{r}) \tag{2.155}$$

where j_l is the spherical Bessel function of the first kind, $h_l^{(2)}$ is a spherical Hankel function of the second kind, P_l is the Legendre polynomial, and $d < r$. It is worth noting that when $l < z$, the amplitudes of functions $j_l(z)$ and $h_l^{(2)}(z)$ remain approximately as constants. However, when $l < z$, $j_l(z)$ experiences very rapid attenuation, while $h_l^{(2)}(z)$ experiences very rapid increase. So when $d \ll r$, (2.155) can be truncated with high precision:

$$\frac{e^{-jk|\mathbf{r}+\mathbf{d}|}}{|\mathbf{r}+\mathbf{d}|} \approx -jk\sum_{l=0}^{L}(-1)^l(2l+1)j_l(kd)h_l^{(2)}(kr)P_l(\hat{d}\cdot\hat{r}) \tag{2.156}$$

Normally, adequate accuracy can be provided by choosing $L = kd + 2\ln(kd+\pi)$. The second identity is the plane-wave expansion of j_lP_l in (2.156):

$$4\pi(-j)^l j_l(k_0d)P_l(\hat{d}\cdot\hat{r}) = \oint e^{-j\mathbf{k}\cdot\mathbf{d}}P_l(\hat{k}\cdot\hat{r})d^2\hat{k} \tag{2.157}$$

where the notation $d^2\hat{k}$ denotes the integration being performed over the whole unit spherical surface. This integration can be achieved by using the Gaussian numerical integral method. To be specific, we choose L points in interval $[0,\pi]$, so that $\cos\theta$ in $[-1,1]$ satisfies the Gauss–Lengendre L point integral equations. The θ value and the weights of the L points in the integral can be directly obtained by using the subroutine "gauleg" in reference [17]. As to the choice of ϕ, $2L$ points can be evenly chosen in $[0,2\pi]$. Thus, the RHS of (2.157) can be written as

$$4\pi(-j)^l j_l(k_0d)P_l(\hat{d}\cdot\hat{r}) = \sum_{p=1}^{K}\omega_p e^{-j\mathbf{k}_p\cdot\mathbf{d}}P_l(\hat{k}_p\cdot\hat{r}) \tag{2.158}$$

where $K = 2L^2$, ω_p denotes the weight, $\hat{k}_p = (\sin\theta_p \cos\phi_p, \sin\theta_p \sin\phi_p, \cos\theta_p)$, $\mathbf{k}_p = k\hat{k}_p$, and (θ_p, ϕ_p) denote the spherical coordinates of the sample point on the unit sphere. Substituting (2.158) into (2.156), and then changing the order of the addends gives rise to

$$\frac{e^{-jk|\mathbf{r}+\mathbf{d}|}}{|\boldsymbol{r}+\boldsymbol{d}|} = -j\frac{k}{4\pi}\sum_{p=1}^{K}\omega_p V(\mathbf{k}_p\cdot\mathbf{d})T_p(kr,\hat{k}_p\cdot\hat{r}) \tag{2.159}$$

where

$$V(\mathbf{k}_p\cdot\mathbf{d}) = e^{-j\mathbf{k}_p\cdot\mathbf{d}} \tag{2.160}$$

$$T_p(kr,\hat{k}_p\cdot\hat{r}) = \sum_{l=0}^{L}(-j)^l(2l+1)h_l^{(2)}(kr)P_l(\hat{k}_p\cdot\hat{r}) \tag{2.161}$$

Notice that the $V(\mathbf{k}_p\cdot\mathbf{d})$ in the RHS of (2.159) is independent of the choice of $\mathbf{r}$, and $T_p(kr,\hat{k}_p\cdot\hat{r})$ is independent of the choice of $\mathbf{d}$. This indicates that (2.159) has separated the direct interaction in Green's function into the far-field translation and the near-field aggregation or disaggregation. For a concise statement, but without losing generality, we use an example of $P(G) = G$ for illustration. As illustrated in Figure 2.22, we define $\mathbf{r} = \mathbf{r}_o - \mathbf{r}_{o'}$ and $\mathbf{d} = \mathbf{r}_{om} - \mathbf{r}_{o'm'}$. From (2.159), the FMM expression of impedance-matrix elements in (2.153) can be obtained:

$$Z_{ij} = \sum_{p=1}^{K}\mathbf{D}_{ip}T_p\mathbf{A}_{pj} \tag{2.162}$$

where

$$\mathbf{D}_{ip} = \int_{S'}\omega_p\boldsymbol{g}_i e^{-j\mathbf{k}_p\cdot\mathbf{r}'_{mo'}}dS' \tag{2.163}$$

$$\mathbf{A}_{pj} = \int_{S} e^{-j\mathbf{k}_p\cdot\mathbf{r}_{mo}}\boldsymbol{g}_j dS \tag{2.164}$$

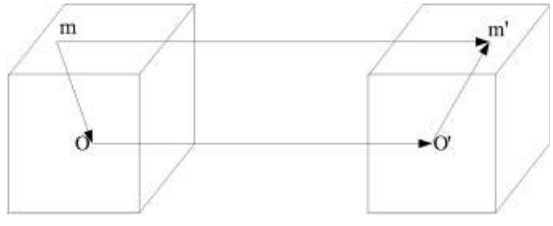


Figure 2.22 The implementation principle of FMM.

Note that $\mathbf{D}_{ip}$ and $\mathbf{A}_{pj}$ are vectors, while T_p is a scalar. Thus, the matrix-vector multiplication in (2.154) can be expressed as

$$\{y\} = [\mathbf{D}][T][\mathbf{A}]\{x\} \tag{2.165}$$

where the matrix entries of $[\mathbf{D}]$ and $[\mathbf{A}]$ are given by (2.163) and (2.164). The matrix element of $[T]$ is $T_{pq} = T_p\delta_{pq}$, where δ_{pq} is the Kronecker delta. Apparently $[T]$ is a diagonal matrix. Equation 2.165 shows the FMM procedure of separating the direct interaction into aggregation, translation and disaggregation in a mathematical way. Figure 2.22 is used for illustration. Points O and O' are the centers of "boxes" A and B, respectively. $[A]\{x\}$ denotes the aggregation step, which aggregates the basis functions in A into K plane-wave functions, denoted by $\{x_1\}$. $[T]\{x_1\}$ denotes the translation step. In this step, the plane waves $\{x_1\}$ are translated from the center O to the center O', yielding K plane waves centered on O', denoted by $\{y_1\}$. Finally, $[D]\{y_1\}$ denotes the disaggregation step. In this step, the plane waves centered on O' are disaggregated into basis functions in B before the final result $\{y\}$ is obtained. An important fact is that matrices $[A]$, $[T]$ and $[D]$ are independent of each other. This means that for a different box B, we can use the same $[A]$, while for different box A, we do not need change matrix $[D]$. In other words, when we calculate the interaction of A and a different B, we use the same $\{x_1\}$. Similarly, the result $\{y_1\}$ from a different box A can always form a total $\{y_1\}$ to be disaggregated in B. The only step that needs to be recalculated is the step of translation, on any change of A or B. Fortunately, the amount of calculation in that step is linear with the number of plane waves, because $[T]$ is a diagonal matrix.

Problems

2.2 Prove the identity (2.155), namely

$$\frac{e^{-jk|\mathbf{r}+\mathbf{d}|}}{|\mathbf{r}+\mathbf{d}|} = -jk\sum_{l=0}^{\infty}(-1)^l(2l+1)j_l(kd)h_l^{(2)}(kr)P_l(\hat{d}\cdot\hat{r}).$$

2.3 Write a program to numerically study the truncation of (2.155).
2.4 Prove the identity (2.157), namely

$$4\pi(-j)^l j_l(k_0 d)P_l(\hat{d}\cdot\hat{r}) = \oint e^{-j\mathbf{k}\cdot\mathbf{d}}P_l(\hat{k}\cdot\hat{r})d^2\hat{k}.$$

2.5 Write a program to numerically verify (2.157).

2.1.8.3 Multilevel Fast-Multipole Algorithm (MLFMA)

It has been stated in Section 2.1.8.1, that the "box" in the FMM can be neither too large nor too small. If the boxes are too large, despite the fact that the translation can be calculated efficiently, the aggregation and disaggregation cannot be done efficiently.

On the other hand, if the boxes are too small, the calculation for the translation is not efficient. Therefore, the size of the boxes needs to be carefully chosen in order to obtain the best efficiency in the FMM. In this section, a new way of implementing the FMM is introduced. The basic idea is to partition the unknowns into boxes of different levels. As one progresses from the finest level to the coarsest level, the box becomes larger. Then, the matrix-vector multiplication is decomposed into two procedures: constructing multipole expansions for each nonempty box at all levels, and constructing local multipole expansions contributed from well-separated boxes at all levels, which includes the translation and the disaggregation. In the first procedure, the multipole expansions are computed at the finest level and then the expansions for larger boxes (coarser level) are obtained using interpolation and shifting. We can still call this the aggregation process. Before we describe the translation process, we need to define the "second-nearby boxes." The second-nearby boxes are well-separated boxes at the present level but they are neighbor boxes at the parent level. At the coarsest level, the local multipole expansions contributed from well-separated boxes are calculated using translation. At the other levels, the local expansions include the contribution from parent boxes by using shifting and anterpolation [18] and the contribution from the second-nearby boxes by using translation. The anterpolation matrix is the transpose of the interpolation matrix. Disaggregation process takes place only in the finest level. This method is termed the *multilevel fast-multipole algorithm (MLFMA)*. Next, we use the step of the disaggregation to illustrate this algorithm. As shown in Figure 2.23, suppose there is a large group made up of four subgroups with totally $4m$ unknowns, which conventionally require $16m^2$ computer operations in the aggregation step. Alternatively, if the aggregation is done first in the subgroup which contains m unknowns each, $4m^2$ operations are needed. Then, translate the four groups of plane waves from the center of each subgroup to the center of the large group, and combine them into m plane waves, $4m$ more operations are needed. Next, construct $4m$ plane waves centered at the large group by interpolating the m plane waves obtained from the previous operation. This is the aggregation step in the large group. In the following section it is shown that this interpolation takes $64m$ operations. In this way, this implementation requires altogether $4m^2 + 68m$ operations. Obviously when m is a big number, $4m^2 + 68m$ is much smaller than $16m^2$.

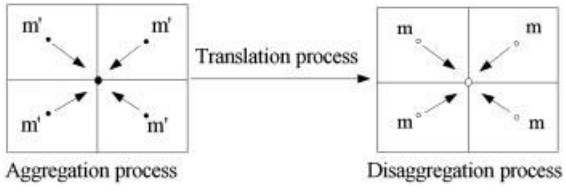


Figure 2.23 The implementation principle of MLFMA.

In order to present a complete MLFMA implementation procedure and without losing generality, we use an example of a square computational domain to illustrate the procedure. As is shown in Figure 2.24, we partition the square domain into four small squares, and label these groups as level-1, the centers of which are labeled using "*." Then, further divide the squares of level-1 into four squares, and repeat this action until the edge of the smallest square is approximately half the wavelength. For the convenience of explanation, we present here only three levels of MLFMA. Figure 2.24b and c give the grouping in level-2 and level-3, respectively, in which "○" and "·" denote the center of each level of group.

Analogous to FMM, MLFMA decomposes the matrix-vector multiplication into three processes, viz., aggregation, translation and disaggregation. The aggregation process takes place from the finest level, before the plane waves are translated to the center of a coarser level and interpolated. The disaggregation process is opposite to it. As to the translation process, MLFMA is almost the same as the FMM. The difference is that the translation in MLFMA only takes place in between the second-nearby boxes.

To begin with, let us consider the aggregation at the level-2 for example. The aggregation is performed through the interpolation of the translated plane waves from the finer lever namely the level-3. Suppose there are N_3 unknowns in each small square in level-3. Thus, there are $M = N/N_3$ small squares altogether. Since the number of plane waves aggregated in the center of the small squares "·" is normally not greater than the number of the unknowns N_3, the operation involved by the translation from centers "·" to centers "○" is normally less than $N_3 \times M = N$. It can be shown by analogous analysis that the operation involved by that the translation in other levels will not exceed N either. In the following section we will demonstrate that the interpolation in each level involves approximately $O(N)$ operations. In this way the operation in the process of aggregation in each level is $O(N)$. Likewise, the operation in the process of disaggregation in each level is also $O(N)$. Now consider the process of translation. As illustrated in Figure 2.24c, the small squares in shallow shading in level-3 are all "second-nearby boxes" of center "a," while the remaining cells are not. It is not hard to imagine, that the "second-nearby boxes" of any given center would not exceed a certain number, denoted as constant C. As for this very example, this number is not going to exceed 27. As was mentioned previously, the process of the translation is only taking place at centers of "second-nearby boxes," because a farther interaction can be obtained by the disaggregation of a coarser level. For example, only the plane waves centered at the small squares in shallow shading are translated to center "a." The dark shading squares are interacted with "a" through the disaggregation of plane waves on center "A" of level-2. If there are N_3 unknowns in level-3, then there are $M = N/N_3$ small squares. Since the number of plane waves aggregated in each small square is less than the number of unknowns N_3, and the translation only takes places between the "second-nearby boxes," the translation in this level need approximately $C \times N_3 \times M = C \times N$ operations.

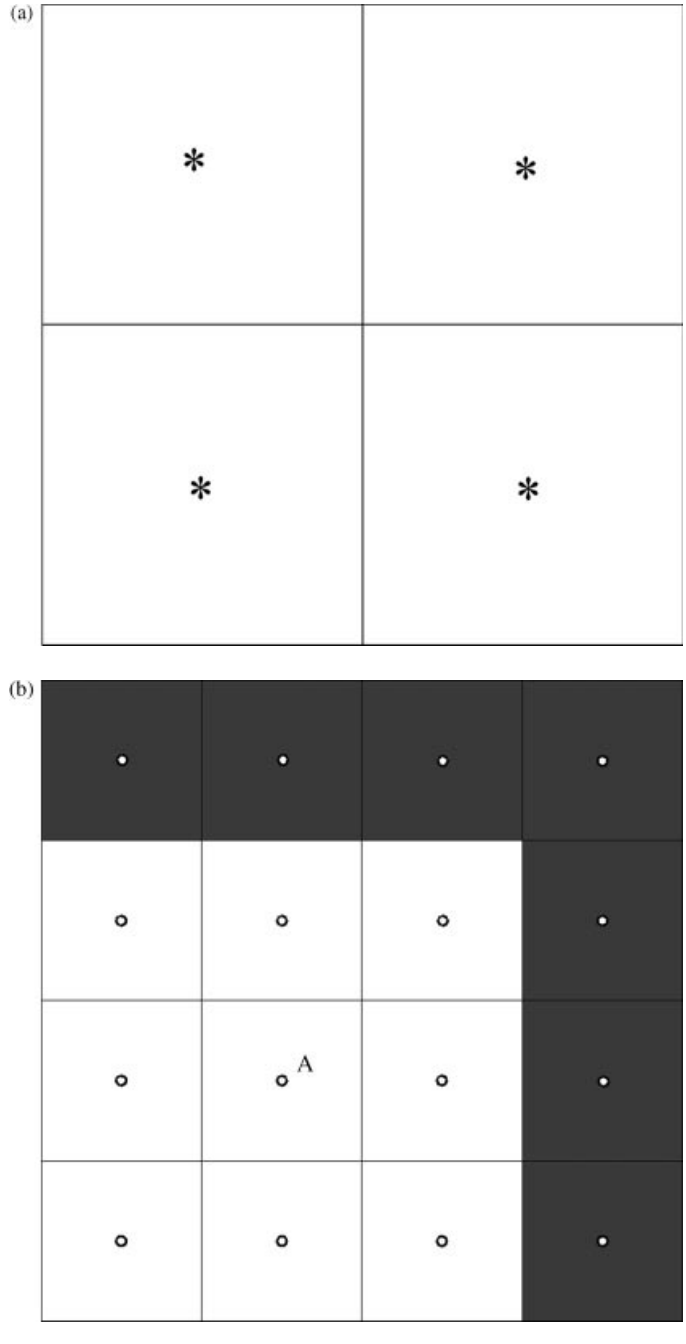


Figure 2.24 A three-level MLFMA. (a) Level-1, (b) level-2, (c) level-3.

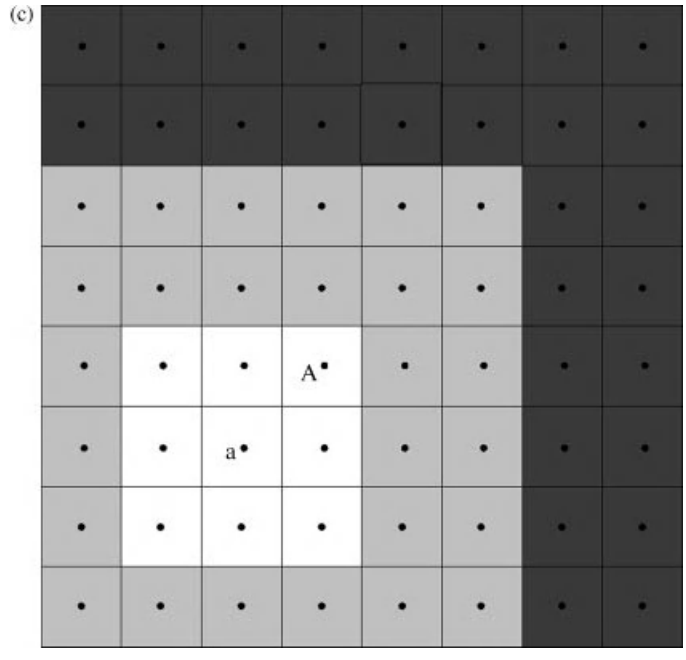


Figure 2.24 (*Continued*)

Thus, the aggregation, translation and disaggregation in MLFMA in each level involves operation of $O(N)$. Usually, $\log N$ levels are used for problems with N unknowns, so the total number of operations will be $O(N \log N)$. It is easy to determine that the memory requirement will not exceed $O(N \log N)$ in MLFMA.

2.1.8.4 Mathematic Formulation in MLFMA

Compared with FMM, there are four additional mathematic operations in MLFMA. Two of them are in the aggregation procedure: one is to translate the plane waves from the center of a finer-level group to that of a coarser-level one; the other is to achieve a large number of plane waves by interpolating a small number of plane waves. The other two are in the disaggregation procedure: one is to translate the plane waves from the center of a coarser-level group to that of a finer-level one; the other is to obtain a small number of plane waves by anterpolating a large number of plane waves. Apparently, the two operations in the disaggregation are the inverse operations of those in the aggregation. So we only concentrate on the mathematic formulation in the process of aggregation. The translation of the center of plane waves is relatively simple. In the first step, we need to multiply a phase factor by the expression of the plane waves according to the center distance. Then, summing up the plane waves with the same $\hat{k}_p$ yields a series of plane waves centered in the coarser box. Note that the plane waves so obtained

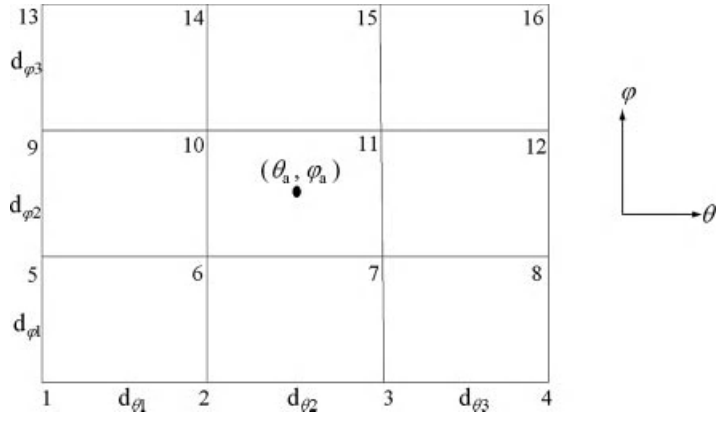


Figure 2.25 The interpolation in the MLFMA.

are less than those required in a coarser level. Since the kernel in EFIE or MFIE has a very smooth impact in the far-field, additional plane waves can be obtained through interpolation. There are many interpolation schemes. Here, we only introduce a popular bicubic interpolation briefly. Suppose $\{y_c\}$ are the amplitudes of a series of plane waves corresponding to the propagation direction of (θ_i, ϕ_j) on a coarse level grid, as shown in Figure 2.25. We aim to calculate the amplitudes of a series of plane waves $\{y_p\}$ corresponding to the propagation direction (θ_i, ϕ_j) on a coarse-level grid. For example, in order to calculate y_a corresponding to (θ_a, ϕ_a) in Figure 2.25, we can express y_a according to [17] as

$$y_a = \sum_{i=1}^{4} \sum_{j=1}^{4} c_{ij} t^{i-1} u^{j-1} \tag{2.166}$$

where

$$t = (\theta_a - \theta_6)/d_{\theta_2} \tag{2.167}$$

$$u = (\phi_a - \phi_6)/d_{\phi_2} \tag{2.168}$$

and $c_{ij} = cl(k)(k = 4 \times (i-1) + j)$ can be obtained by using

$$cl(k) = \sum_{i=1}^{16} \omega_{ki} \times z(i) \tag{2.169}$$

where $z(i)(i = 1, 2, \ldots, 4)$ denote the amplitudes of the plane waves corresponding to points 6,7,10,11; $z(i)(i = 5, 6, \ldots, 8)$, $z(i)(i = 9, 10, \ldots, 12)$, $z(i)(i = 13, 14, \ldots, 16)$ denote the derivatives of θ, ϕ, and their cross derivatives, respectively.

Moreover,

$$[\omega] = \begin{bmatrix}
1 & 0 & -3 & 2 & 0 & 0 & 0 & 0 & -3 & 0 & 9 & -6 & 2 & 0 & -6 & 4 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 & -9 & 6 & -2 & 0 & 6 & -4 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 9 & -6 & 0 & 0 & -6 & 4 \\
0 & 0 & 3 & -2 & 0 & 0 & 0 & 0 & 0 & 0 & -9 & 6 & 0 & 0 & 6 & -4 \\
0 & 0 & 0 & 0 & 1 & 0 & -3 & 2 & -2 & 0 & 6 & -4 & 1 & 0 & -3 & 2 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 3 & -2 & 1 & 0 & -3 & 2 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -3 & 2 & 0 & 0 & 3 & -2 \\
0 & 0 & 0 & 0 & 0 & 0 & 3 & -2 & 0 & 0 & -6 & 4 & 0 & 0 & 3 & -2 \\
0 & 1 & -2 & 1 & 0 & 0 & 0 & 0 & 0 & -3 & 6 & -3 & 0 & 0 & -4 & 2 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & -6 & 3 & 0 & -2 & 4 & -2 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -3 & 3 & 0 & 0 & 2 & -2 \\
0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & -3 & 0 & 0 & -2 & 2 \\
0 & 0 & 0 & 0 & 0 & 1 & -2 & 1 & 0 & -2 & 4 & -2 & 0 & 1 & -2 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 2 & -1 & 0 & 1 & -2 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & -1 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 2 & -2 & 0 & 0 & -1 & 1
\end{bmatrix} \tag{2.170}$$

By using the difference method to calculate the derivatives, we can express {z} as

$$\{z(i)\} = [R][C]\{y\} \tag{2.171}$$

where

$$[\mathrm{C}] = \begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 1
\end{bmatrix} \tag{2.172}$$

and $\{y\}$ denote the amplitudes of the plane waves on a coarse grid. We may notice that $[R]$ is a diagonal matrix, with its diagonal elements r_i given by:

$$\begin{aligned}
r_i &= 1 && i = 1,2,3,4 \\
r_i &= \frac{1}{d_{\theta 1} + d_{\theta 2}} && i = 5,8 \\
r_i &= \frac{1}{d_{\theta 2} + d_{\theta 3}} && i = 6,7 \\
r_i &= \frac{1}{d_{\phi 1} + d_{\phi 2}} && i = 9,10 \\
r_i &= \frac{1}{d_{\phi 2} + d_{\phi 3}} && i = 11,12 \\
r_{13} &= \frac{1}{(d_{\theta 1} + d_{\theta 2})(d_{\phi 1} + d_{\phi 2})} \\
r_{14} &= \frac{1}{(d_{\theta 2} + d_{\theta 3})(d_{\phi 1} + d_{\phi 2})} \\
r_{15} &= \frac{1}{(d_{\theta 2} + d_{\theta 3})(d_{\phi 2} + d_{\phi 3})} \\
r_{16} &= \frac{1}{(d_{\theta 1} + d_{\theta 2})(d_{\phi 2} + d_{\phi 3})}
\end{aligned} \tag{2.173}$$

Thus, $\{cl\}$ can be obtained through the following equation:

$$\{cl\} = [W][R][C]\{y\} \tag{2.174}$$

From (2.166) it can be seen that $64m$ computation operations are required when m plane waves are interpolated into $4m$ plane waves, as stated previously in Section 2.1.8.3.

2.1.9 Calculation of Scattered Fields

Once the equivalent electric currents $\mathbf{J}$ on the surface of a PEC body are obtained, from Section 1.3.2, we know that the scattered field at an arbitrary point in the space can be calculated using

$$\mathbf{E}^S = -j\omega\mu \int \left[1 + \frac{1}{k^2} \nabla\nabla \cdot \right] \mathbf{J}(\mathbf{r}') G dS' \tag{2.175}$$

$$\mathbf{H}^S = -\int \mathbf{J}(\mathbf{r}') \times \nabla G dS' \tag{2.176}$$

In the following, we will give a more concise approximation expression. The far-field is defined as the region for which $r > 2D^2/\lambda$, where D is the largest distance between two points on the surface of the object. Generally, a radar receives only the far-field

backscattered signals. Under the far-field condition, the above Green's function G can be approximately expressed as

$$G(\mathbf{r}|\mathbf{r}') \approx \frac{e^{-jkr}}{4\pi r} e^{jk\mathbf{r}'\cdot\hat{r}} = g_1(r)g_2(\theta,\phi) \tag{2.177}$$

where

$$g_1(r) = \frac{e^{-jkr}}{4\pi r} \tag{2.178}$$

and

$$g_2(\theta,\phi) = e^{jk\mathbf{r}'\cdot\hat{r}} \tag{2.179}$$

Apparently, in the spherical coordinate system, g_1 is only a function of r but not a function of θ, ϕ, while g_2 depends only on θ, ϕ, but not on r. Thus, in the spherical coordinate system,

$$\begin{aligned} \nabla G &= g_1\nabla g_2 + g_2\nabla g_1 \\ &= g_1 O\left(\frac{1}{r}\right) + g_2\left[-jk\hat{r}g_1 + O\left(\frac{1}{r^2}\right)\right] \\ &= -jkG\hat{r} + O\left(\frac{1}{r^2}\right) \end{aligned} \tag{2.180}$$

where $O(\cdot)$ denotes the order of magnitude. Ignoring higher-order terms yields

$$\nabla G \approx -jkG\hat{r} \tag{2.181}$$

So we obtain

$$\begin{aligned} \nabla\nabla\cdot[\mathbf{J}(\mathbf{r}')G] &= \nabla[\mathbf{J}(\mathbf{r}')\cdot\nabla G] \\ &\approx \nabla[-jkG\hat{r}\cdot\mathbf{J}(\mathbf{r}')] \\ &\approx -jk[\hat{r}\cdot\mathbf{J}(\mathbf{r}')]\cdot\nabla G \\ &\approx -k^2[\hat{r}\cdot\mathbf{J}(\mathbf{r}')]G\hat{r} \end{aligned} \tag{2.182}$$

Substituting (2.182) into (2.175), we will find that under the spherical coordinate, E^S only has θ, ϕ components, written as

$$\mathbf{E}^S = -j\omega\mu\mathbf{N}_t = -jkZ\mathbf{N}_t \tag{2.183}$$

where Z denotes the wave impedance in a homogeneous space, and N_t denotes the tangential components of the following $\mathbf{N}$ with respect to $\hat{r}$:

$$\mathbf{N} = \frac{e^{-jkr}}{4\pi r}\int \mathbf{J}(\mathbf{r}')e^{jk\mathbf{r}'\cdot\hat{r}}dS' \tag{2.184}$$

Equation 2.183 is the *scattered field* expression in the far-field. By an analogous approach we can obtain

$$\mathbf{H}^S = -jk\hat{r} \times \mathbf{N}_t = \frac{1}{Z}\hat{r} \times \mathbf{E}^S \tag{2.185}$$

In order to highlight the scattering property of the object, we use the evenly scattered power density to normalize the actual scattered power density from the object. In this way a widely used physical quantity – the *radar cross section* (RCS) σ is defined as:

$$\sigma(\theta, \phi, \theta^i, \phi^i) = \lim_{r \to \infty} 4\pi r^2 \frac{\left|\mathbf{E}^S(\theta, \phi)\right|^2}{\left|\mathbf{E}^i(\theta^i, \phi^i)\right|^2} \tag{2.186}$$

The RCS can usually be further categorized as the bistatic RCS and the monostatic RCS. The bistatic RCS is calculated by fixing the incident angle and observing the scattering in arbitrary directions from the object. It is widely used in the inverse problem. In the monostatic RCS calculation, the observation angle is always opposite to the incident angle. So it is also termed the backscattering RCS. By sweeping the incident angle, the monostatic RCS for a desired range of angle can be obtained. When the observation angle is the same as the incident angle, the RCS is also termed the forward scattering RCS. Furthermore, antenna engineers often split the linear polarization (mostly used in the radar system) of a wave into components along $\hat{\theta}$ and $\hat{\phi}$ directions of the spherical coordinate system. While these components may vary with the angle θ and ϕ, each is confined to a plane containing the direction of propagation, nonetheless, and they are orthogonal. If we want to express the polarization characteristics, we need to use the following polarization scattering matrix

$$\begin{bmatrix} \sigma_{\theta\theta} & \sigma_{\theta\phi} \\ \sigma_{\phi\theta} & \sigma_{\phi\phi} \end{bmatrix} \tag{2.187}$$

where $\sigma_{\theta\theta}$ denotes the *vertically polarized RCS*, corresponding to the case with vertically polarized incident field ($\hat{\theta}$) and vertically polarized received field ($\hat{\theta}$); $\sigma_{\phi\phi}$ denotes the *horizontally polarized RCS*, corresponding to the case with horizontally polarized incidence ($\hat{\phi}$) and horizontally polarized received field ($\hat{\phi}$); $\sigma_{\theta\phi}$ or $\sigma_{\phi\theta}$ denote the *crosspolarized RCS*, corresponding to the cases with vertically polarized incidence ($\hat{\theta}$) and horizontally polarized received field ($\hat{\phi}$), or horizontally polarized incidence ($\hat{\phi}$) and vertically polarized incidence received field ($\hat{\theta}$).

Problem

2.6 Prove the following approximation used in (2.182)

$$\nabla[-jkG\hat{r} \cdot \mathbf{J}(\mathbf{r}')] \approx -jk[\hat{r} \cdot \mathbf{J}(\mathbf{r}')] \cdot \nabla G$$

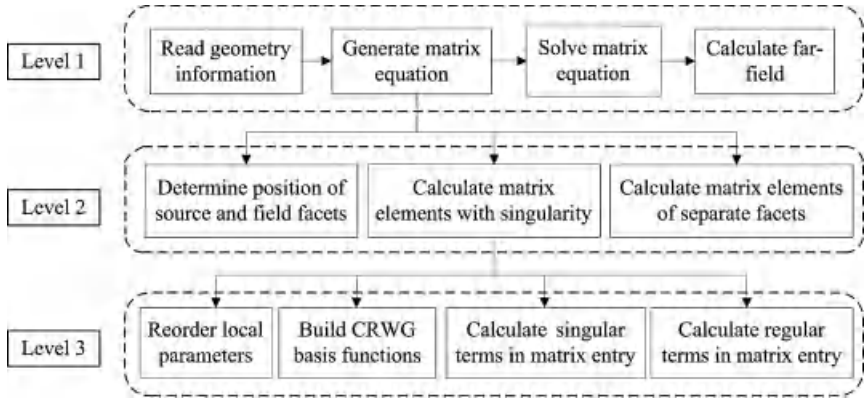
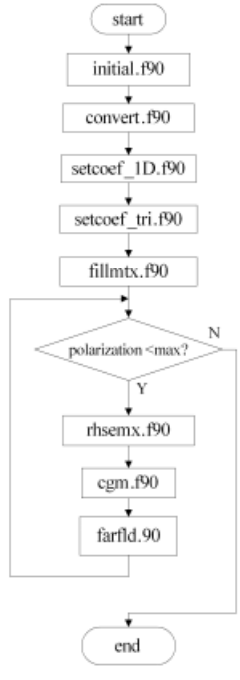


Figure 2.26 The hierarchical chart for the MoM computer program.

2.1.10 Writing Computer Program

By now we have presented the complete discretized formulation in calculating the scattering by an arbitrary PEC body. The next step is to translate this formulation into computer language, that is, to write the computer program. To this end, we usually first sketch out the hierarchical chart and the float chart for the computer program. In the following, we will use the calculation of the scattering by a PEC object as an example to demonstrate how to write the computer program. Figures 2.26 and 2.27 demonstrate the hierarchical and float chart for computing scattering by a PEC object, respectively. In the hierarchical chart, Level-1 lists four major parts of the program for four functions, namely, obtaining the geometry information of the object, forming the matrix equation; solving the matrix equation, and calculating the far-field pattern. The subroutines for obtaining the geometry information, solving the matrix equation and calculating the far-field are straightforward, and their structures are simple and can be easily written. However, the one forming the matrix equation is complex. To understand and write this subroutine, more explanation is given in Level-2. As shown in Figure 2.26, in Level-2 the three major parts of this subroutine for three functions are listed: determining the relative position of the source and field facets; calculating matrix entries with a singularity; and calculating matrix entries without a singularity. Again, the subroutines for implementing all functions are simple and can be easily written except for the one treating the singularity. Therefore, Level-3 provides a detailed explanation about the singularity treatment procedure, including reordering the local parameters, building the curvilinear-RWG (CRWG) basis functions, calculating the singularity terms in matrix entries, and calculating the regular terms in matrix entries. As a result, the structure of the whole program is clearly illustrated in this chart by using a hierarchical style, which makes this hierarchical chart a useful tool in helping us grasp the logical relationship of the subroutines involved. The float charts as shown in Figure 2.27 are



Notes for the figure

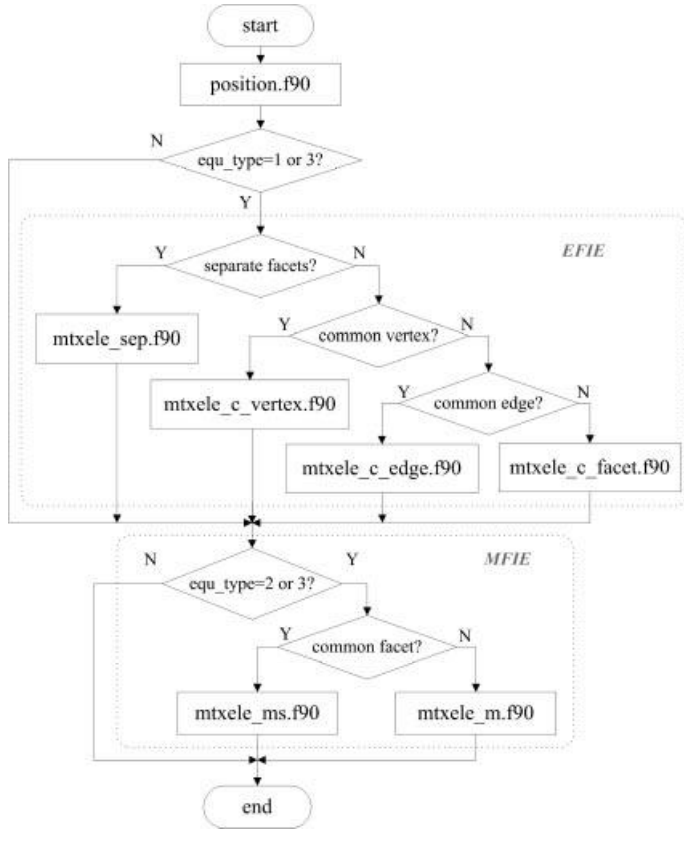
initial.f90	Read control parameters & set the RCS receiver parameters.
convert.f90	Read and convert geometry data.
setcoef_1D.f90	Set the Gauss-Legendre points and the weights
setcoef_tri.f90	Set the Gauss points and the weights based on the triangle.
fillmtx.f90	Fill the matrix.
rhsemx.f90	Set up the right-hand-side of the matrix equation.
cgm.f90	Solve the matrix equation.
farfld.f90	Calculate the EM field in the far zone.
index: polarization	*1: theta(V), 2: phi(H), 3: Both*

(a)

Figure 2.27 The float charts of the computer program. (a) MOM.f90; (b) cfiex.f90.

also useful tools in understanding and writing computer programs. There are two float charts for the main program and for one major subroutine respectively. The functions of the subroutines involved and the key parameters are detailed in the accompanying tables. From these charts, the reader can gain a clear view about how the program is executed.

As to each subroutine, it is essential to clearly define the input and the output arguments, as well as clarify which variable is to be stored and which is not. The subroutine retrieving the geometry information of the object defines the following



Notes for the figure

position.f90	Determine relationship according to the number of COMMON nodes, and group the position relationship of the field- and source- triangles into different types: separate, sharing a node, an edge, or a facet.
mtxele_sep.f90	Calculate the EFIE matrix elements for separated facets.
mtxele_c_vertex.f90	Calculate the EFIE matrix elements for the common vertex singularity.
mtxele_c_edge.f90	Calculate the EFIE matrix elements for the common edge singularity.
mtxele_c_facet.f90	Calculate the EFIE matrix elements for the common facet singularity.
mtxele_ms.f90	Calculate the MFIE matrix elements for the common facet singularity.
mtxele_m.f90	Calculate the MFIE matrix elements for other cases.
Index: equ_type	*1:EFIE, 2:MFIE, 3:CFIE*

(b)

Figure 2.27 (*Continued*)

output arguments: the number of triangular patches that discretize the PEC surface (stored using integer variable *maxpatch*), the number of edges (stored using integer variable *maxedge*), the number of vertices (stored using integer variable *maxnode*), the indices of three vertices on a given triangular patch (stored using integer array *ipatch*(*3,maxpatch*)), the coordinate information of all the vertices (stored using integer array *xyznode*(*3,maxnode*)), the indices of the two vertices on a given edge and the indices of the two triangles sharing this common edge (stored using integer array *iedge*(*4,maxedge*)). These data are read from data files.

The subroutine for constructing the matrix equation takes the above output arguments as input, and outputs a coefficient matrix $[A]$ (stored using the complex array *cmatrix*(*maxedge,maxedge*)) and a vector $\{J\}$ (stored using the complex array *crhs* (*maxedge*)). The size of *cmatrix* (*maxedge,maxedge*) determines the memory requirement of the whole program. The coefficient matrix $[A]$ is normally constructed by using a double loop. Since each loop index changes from 1 to *maxedge* to calculate one matrix element, altogether $maxedge \times maxedge$ operations are involved. Obviously, this is a very time-consuming process. It is worth noting that the basis function defined on every edge corresponds to two triangles, therefore the integration needs to be done over the two triangles.

The subroutine solving the matrix equation, and the one calculating the far-field patterns have obvious input and output arguments, so we do not discuss them in detail here. This programming structure does not include the implemention of MLFMA. If we implement MLFMA, the above structure needs to be adjusted. After the geometry information collection subroutine, we need to add a subroutine to perform a sorting function, in order to obtain: the total group number, the subjection of the triangles with the group, and the direct neighbor, the second-nearby group, and the well-separated group of each group for the calculations of near interaction, translation and far interaction, respectively. Correspondingly, the subroutine constructing the matrix equation can be split further into two subroutines. One calculates the near interaction. It can be extended readily from the original subroutine. The major difference is the use of a one-dimensional array instead of the *cmatrix (maxedge,maxedge)*, to store the matrix elements, because the matrix for the near interaction is fairly sparse. The other one calculates the aggregation, translation and disaggregation matrices for far interaction. The subroutine for solving the DIEs is also different. Only an iteration method, such as the CG method can be applied, with the matrix-vector multiplication calculated by using MLFMA.

It is acknowledged that to write a computer program is not a challenging task by itself. The difficulty lies in obtaining the correctness, efficiency, and robustness of the code. In other words, the code needs to work correctly and efficiently in any circumstance. Generally, one can translate the discretized equations into a computer program in a short time, but one may also find there are bugs here and there and the program is not working at all. These bugs can be categorized into two types: one is the computer language error, such as the violation of the computer language rule, wrong data type definition, and array misoperation, and so on. The other is related to the

discretized equation itself. Bugs of the former type are easy to locate, especially with the aid of the "Debug" tool provided by various high-level programming platforms. However, bugs belonging to the latter type are hard to find. They include mathematical derivation mistakes, bias understanding of knowledge, lacking of consideration, and fault in choosing of the discretized method. Mathematical derivation mistakes can be corrected after careful examination, while finding the remaining errors requires long term exploration, which is also a study process towards a deep understanding of the method. How this process proceeds varies from person to person. As far as the authors' experience is concerned, designing various simple but typical models and verifying by applying laws of physics is an effective approach.

Despite the fact that it is straightforward to convert an explicit formulation system into computer programs, particular care needs to be devoted to the program readability, portability, efficiency and the requirement on the computer memory. A few suggestions are enumerated below. In order to assure the accuracy and efficiency, we need to pay special attention to the following:

- In a summation operation, we can change the order of the addends, so that smaller numbers are added before the big ones. This can prevent the small numbers being "eaten" by the big ones. For example, if $A > B > C$, then $(C + B) + A$ is better than $(A + B) + C$.
- Try to avoid the subtraction between two big but comparable numbers. Otherwise, the number of significant numbers will be lost. For example, if $X \ll R$, then $\sqrt{2.0 \cdot R \cdot X + X^2}$ is better than $\sqrt{(R + X)^2 - R^2}$.
- Try to avoid a small number to be the divisor and a big number to be the multiplier.
- Always simplify the expression to reduce the computer operations. For example, $X - Y$ is better than $(X^2 - Y^2)/(X + Y)$.
- Try to avoid the computation between data with different data types.

The last two are also helpful in improving the efficiency of the program. Besides, to improve the efficiency, we also need to take care about:

- Make full use of the recursions. For example, if we compute x^{255} directly, 254 operations are needed. Instead, if we use the equality $x^{255} = x \cdot x^2 \cdot x^4 \cdot x^8 \cdot x^{16} \cdot x^{32} \cdot x^{64} \cdot x^{128}$, then only 7 operations are involved.
- Avoid using unnecessary arrays in the program.
- Determine the sequence of loops according to the computer language. Consider the following Fortran program:

```
   DO 10 I=1,50
   DO 10 J=1,50
10 A(I,J)=0.0
```

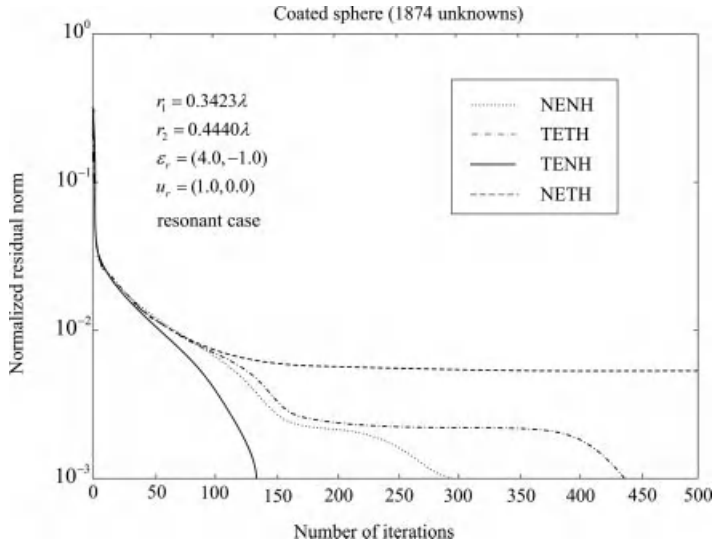


Figure 2.28 The iteration convergence history of the three DIEs by using the CG method.

This code runs much slower than:

```
   DO 10 J=1,50
   DO 10 I=1,50
10 A(I,J)=0.0
```

That is because in Fortran, the elements in matrix A are arranged in the memory in the following order

```
A(1,1),A(2,1),...,A(50,1),A(1,2),A(2,2),...
```

After debugging the program, we need to comment in detail on the input, output arguments as well as the method used and the reference literatures. Experienced code writers often view this as an important and necessary step, because with the aid of these comments, the code writer himself or other people can read, use and modify the code fairly conveniently in the future. Meanwhile, it is also recommended to record typical input and output data for further modification and debugging in the future.

2.1.11 Numerical Examples

Going through debugging, the program is ready to be run on computers and provide numerical results. In the following, we use a few examples to illustrate the numerical characteristics and accuracy of the aforementioned MoM schemes. As is known, there

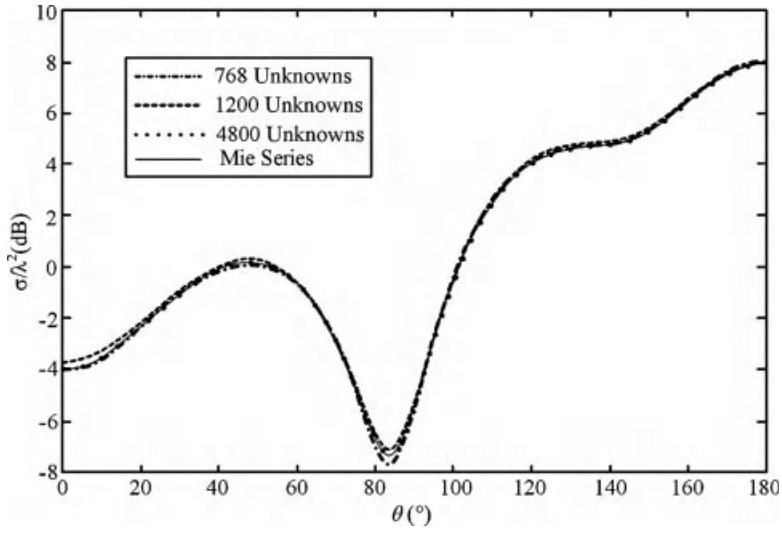


Figure 2.29 The bistatic RCS of a PEC sphere with radius $r = 0.45\lambda$.

is a rigorous analytical result (*Mie series*) for RCS of a metallic sphere. So, next we use this example to examine the accuracy of MoM. Unless we note specifically, all the programs in this book are written using the high-level computer language Fortran. The numerical experiments in this and the next sections are performed on the *Tsinghua Tong-fang Extra* 1000L server (P4 2G CPU, 1 Gb memory). The operation system is Red Hat Linux 8.0.

First, let us calculate the bistatic RCS of a PEC sphere with radius $r = 0.45\lambda$ (where λ is the free-space wavelength). The surface of the sphere is discretized with 512 triangular patches. There are 768 edges with an average length of 0.109λ. Figure 2.28 plots the iteration convergence history calculated by using different DIEs with reference to the theoretical results given by the Mie series. It can be seen that the solution of the CFIE converges the fastest. It is followed by the MFIE, and the EFIE shows the slowest convergence. Figure 2.29 plots the bistatic RCS curves calculated by using different DIEs and by using the Mie series. Good agreement in the figure has indicated fairly satisfactory accuracy in all the numerical results. It can also be seen that the EFIE and CFIE results are comparable and are slightly better than the MFIE one. However, it is shown by following numerical experiment that the three DIE results converge to the analytical result with the increase of the spatial discretization resolution. We increase the mesh density of the sphere step by step, and obtain three meshes made up of 1200, 2700 and 4800 edges, with average edge sizes of 0.087λ, 0.058λ and 0.044λ, respectively. Simulations are performed based on these three meshes.

Figure 2.30 shows the number of iterations to reach convergence of the three DIEs by using CG method as a function of the number of unknowns. It can be seen that the

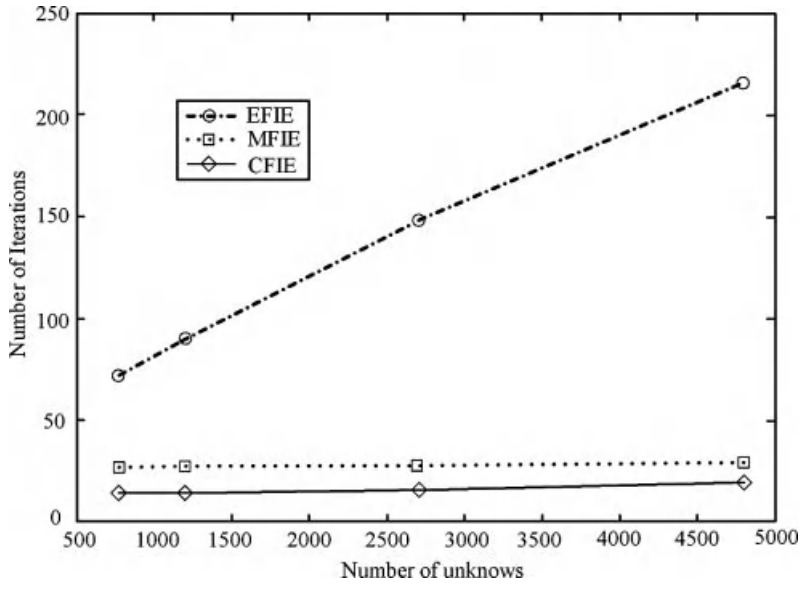


Figure 2.30 The number of iterations to reach convergence of the three DIEs by using a CG method as a function of number of unknowns.

iteration numbers in the MFIE and CFIE are almost constant, while that of the EFIE increases rapidly with the increase of the spatial resolution. Thus, the analytical statement in Section 2.1.6 is verified: with the decrease of the frequency, the condition number of the EFIE becomes worse, while those of the MFIE and CFIE do not. Figure 2.31 verifies that as the spatial resolution increases, the three DIEs results are convergent to the Mie series result.

Next, we will examine the saving on computer resources by incorporating MLFMA in MoM. Using the above model, the saving in the computer memory and the CPU time by employing MLFMA is shown in Table 2.5. The degradation on accuracy brought by MLFMA is small, as illustrated in Figure 2.32.

Due to the limited electrical size of the above PEC sphere, only two-level MLFMA can be applied. Therefore, the saving on the computer resources is not so evident. In order to further show the capability of MLFMA, we apply the 2-, 3-, 4-, 5- and 6-level MLFMA to calculate the RCS of PEC spheres with radii of 0.75λ, 1.5λ, 3λ, 6λ, and 12λ, respectively. The computer resources in the calculation are presented in Table 2.6. It can be seen from this table that the computer memory and the CPU time increase roughly linearly with the number of the unknowns in MLFMA, which verifies the analytical prediction in Section 2.1.8.3. It also indicates the overwhelming advantage of MLFMA in calculating the electrically very large problems. For example, to calculate the RCS of PEC spheres with diameter of 12λ, only 474 Mb computer memory and 21 minutes are taken in the MLFMA program. Without using MLFMA, it

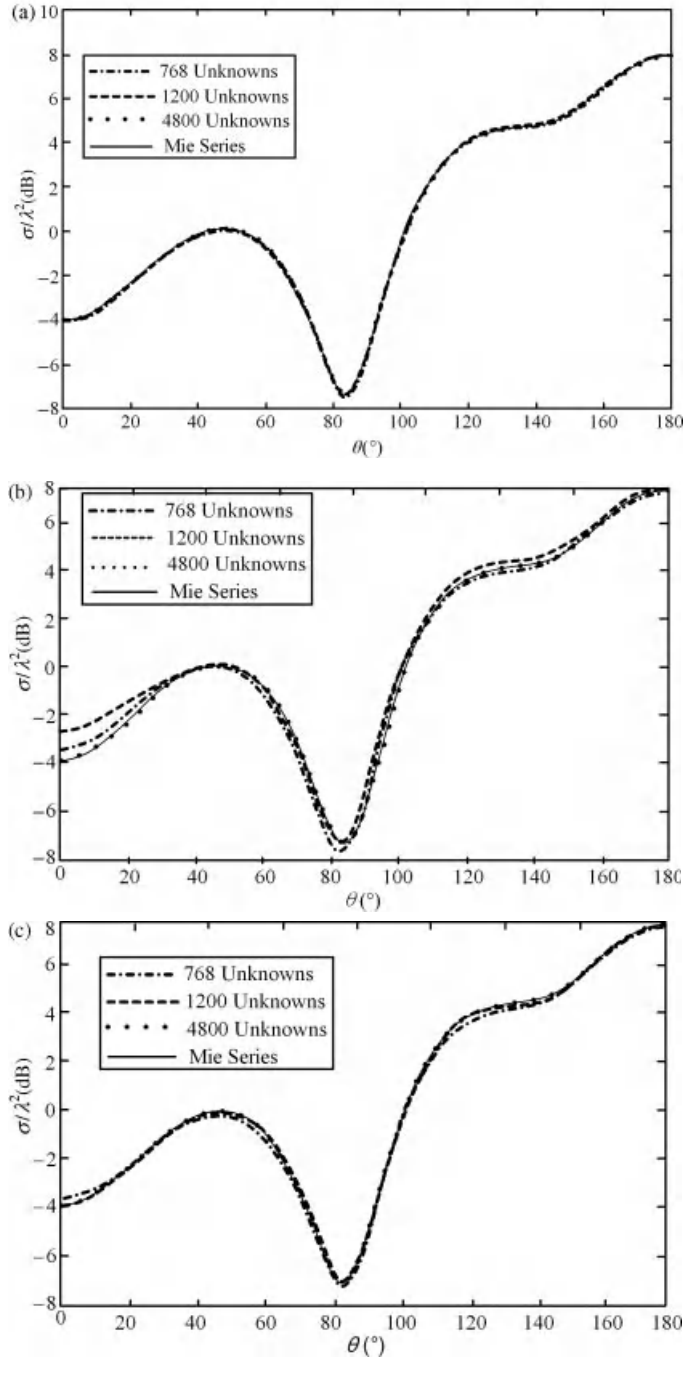


Figure 2.31 The accuracy convergence with the spatial resolution of the three DIEs in calculating the RCS of a metallic sphere. (a) EFIE (b) MFIE (C) CFIE.

Table 2.5 The saving on the computer memory and computational time by employing two-level MLFMA

Number of unknowns	768	1200	2700	4800
memory (without using MLFMA)/Mb	4.9	11.8	58.8	185.2
memory (using MLFMA)/Mb	2.8	5.2	18.7	52.6
CPU (without using MLFMA)/s	8.3	17.0	77.7	240.6
CPU (using MLFMA)/s	6	11.0	45.4	137.2

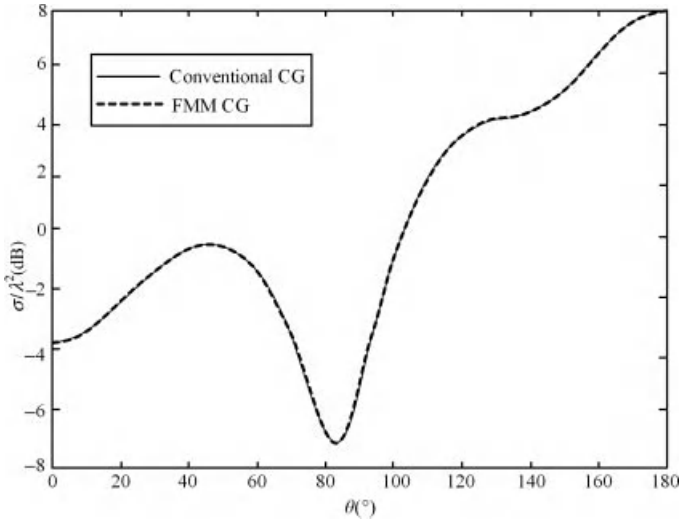


Figure 2.32 The bistatic RCS of a PEC sphere of radius $r = 0.45\lambda$ calculated by using a two-level MLFMA.

will take approximately 2.4×10^6 Mb computer memory and about one month to complete the calculation. Although approximations are introduced in the MLFMA, adequate accuracy has been verified by Figure 2.33.

Table 2.6 The saving on the computer memory and computational time by employing MLFMA

Diameter of sphere (λ)	0.75	1.5	3.0	6.0	12.0
Number of unknowns	768	2700	10 800	43 200	172 800
memory (without MLFMA)/Mb	4.9	58.8	935.1	1.5E5	2.4E6
memory (with MLFMA)/Mb	2.3	7.4	29.5	118.0	474.0
CPU (without MLFMA)/s	9	57	912	1.5E5 (Est.)	2.4E6 (Est.)
CPU (with MLFMA)/s	6	17	69	281	1246

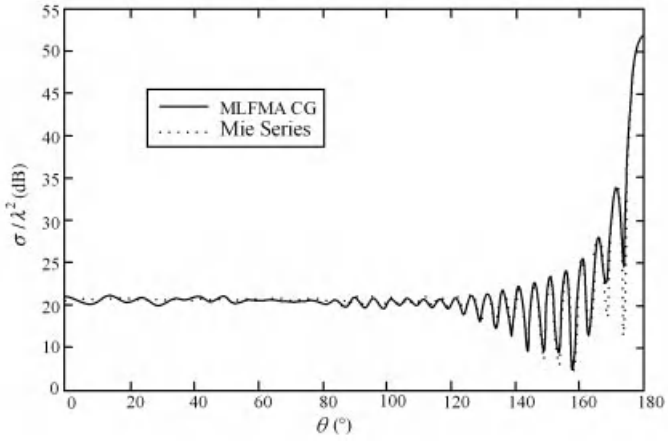


Figure 2.33 The bistatic RCS of a PEC sphere of radius of 12.0λ calculated by using a 6-level MLFMA.

The above numerical experiments have shown that the MoM incorporating MLFMA can accurately calculate the amplitude of the scattering (or the RCS) from an electrically large target. In many present radar systems, such as the range radar, the synthetic aperture radar, and the three-dimensional interferometric imaging radar, the phase information is as important as the amplitudes. Therefore, we need to examine the accuracy of the above MoM algorithm on phase. Let us still use the PEC sphere as an example. The phase and amplitude of the scattering are calculated by using the above MoM algorithm and the physical optics (PO) method. These numerical results are compared with the analytical solution. It can be seen from Figure 2.34 that the above MoM is able to calculate both the amplitude and the phase of the scattering precisely.

Now we can come to a conclusion: the MoM incorporating MLFMA is a powerful simulation tool in studying the scattering property of metallic targets and microwave imaging. To make the best use of this tool, an essential point is to provide a series of measures to assure the reliability of the simulation resulting from the absence of the experiment or other kinds of results. To this end, we provide a double examine criterion according to the characteristics of the MoM. These criterions are: (1) examining the convergence of the results on meshes with different spatial resolutions; (2) examining the convergence of the results under setting different residual errors for iteration convergence. With these criterions, we can not only examine the reliability of the numerical method, but can also select the best computation parameters regarding the tradeoff of computation efficiency and accuracy. In the following, these criteria are applied to the simulation of scatterings from an aircraft-shaped target. Figure 2.35a plots the simulated phase curves of the backscattering from the target under setting different residual errors, while Figure 2.35b plots this under different spatial resolutions. From these numerical experiments we gain the

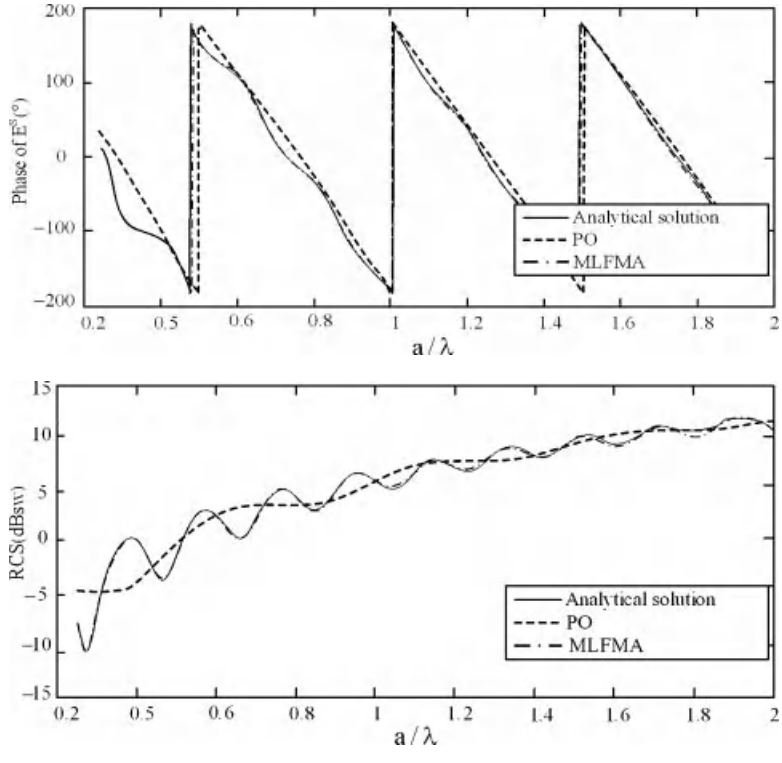


Figure 2.34 Comparison of the phase and amplitude of the scattering provided by different methods.

experience that, when the average edge size is approximately 0.1λ, and the residual error is 0.01, the simulation has adequate accuracy as well as satisfactory computational efficiency. We also come to the conclusion that the meshes of the model affect the computational efficiency and accuracy greatly.

2.1.12 Parallel Technology

2.1.12.1 Parallel MLFMA

A great number of numerical experiments show that in modeling scattering from PEC targets, MoM can provide satisfactory accuracy for the requirement in the practical engineering projects. It is shown that the computational efficiency and memory requirement are the bottleneck in the application of MoM in the engineering. Although MLFMA has greatly improved the computational efficiency and decreased the memory requirement of MoM intrinsically, the CPU time and memory requirement in the complex, electrically very large body in the real world is still far too expensive

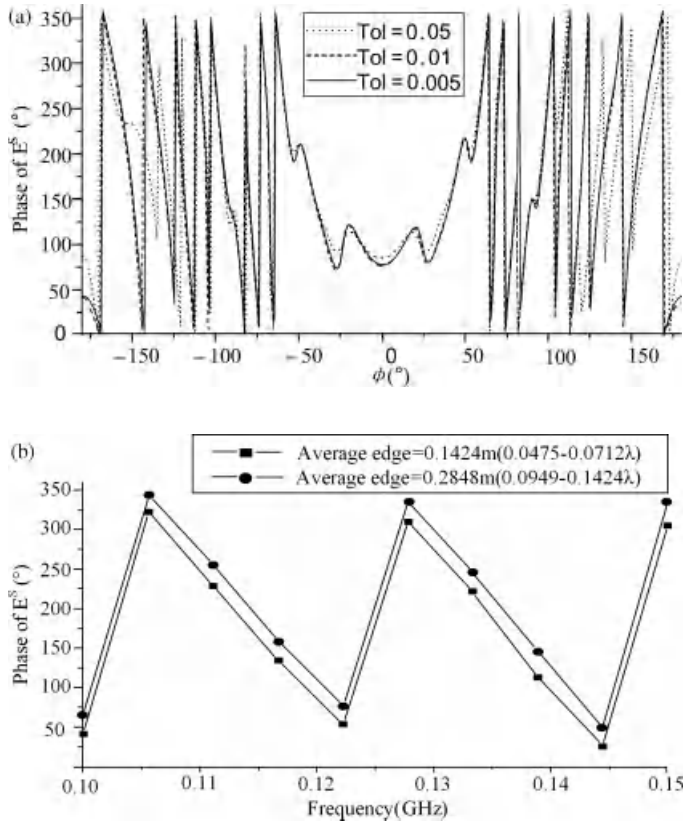


Figure 2.35 Simulated phase curves of the backscattering from an aircraft-shaped target under (a) different residual errors; (b) different spatial resolutions.

to model. There are two solutions to this problem. One is to analyze the scattering mechanism and simplify the target; the other is to further improve the computational efficiency. The former actually depends on the problem itself. So we only discuss the latter here. From the development of the present technology, computational efficiency can be improved in the following two manners. One is to combine MoM with high-frequency methods. As will be discussed in detail in Chapter 5, accuracy is sacrificed to some extent in the hybrid method. The other is to make full use of the *parallel computer technology*, to design and develop the parallel algorithms. In the following, a parallel algorithm based on MLFMA is briefly introduced.

Generally speaking, highly efficient parallel algorithms are designed and developed against specific parallel computer architecture. There are different kinds of architectures in parallel computer technology, such as *the parallel vector processor, the symmetric multiprocessor, the massively parallel processor and the Beowulf cluster*. Among these parallel architectures, the *Beowulf cluster* is the cheapest, most extendable, and easy to

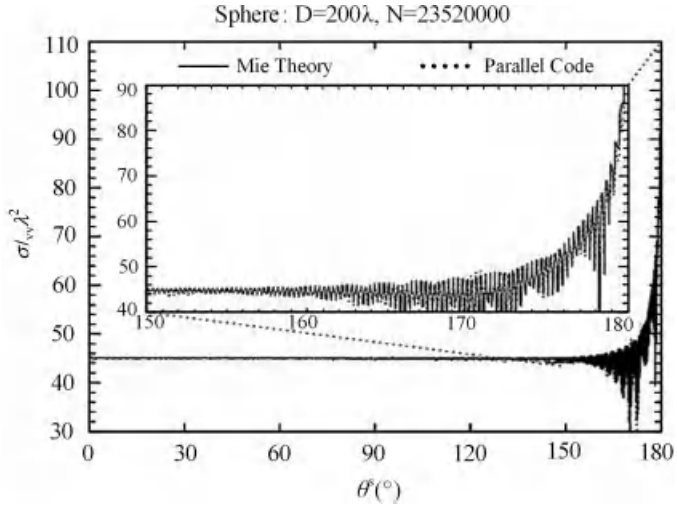


Figure 2.36 The bistatic RCS of a metallic sphere with diameter 200λ.

use one, and is consequently one of the most popular computer architectures. The following parallel MLFMA is developed based on the Beowulf cluster.

Before the *parallel MLFMA* is discussed, let us briefly introduce the *Beowulf cluster*. The hardware in a Beowulf cluster includes the computer nodes, the network connecting the nodes, the switches controlling the data transfer in the network, and the KVM for the switching of the keyboards and the monitors. In terms of the software, an operation system such as Linux controls the system operations. Besides, there are parallel program developing languages, such as MPI and the parallel program debugging tools, such as TotalView.

Various techniques have to be employed in the parallelization of MLFMA on a Beowulf cluster platform, for example, the distributed storage of the geometry information, the transformation of the one-dimensional arrays containing the "box" and "level" information, into Morton Key, and the special treating of the translation matrix. We are not going to enumerate all the designs and techniques here. The interested reader is referred to [19] for more details. What we discuss here is the parallelization of the translation process, which is a difficult but critical issue in the parallel MLFMA. There are numerous schemes. One scheme is to distribute the "boxes" to different computational nodes and store the translation matrix in the corresponding node. However, this scheme is not suitable for the boxes in the coarse levels. Since there are a few boxes in the coarse levels, it is hard to evenly distribute the workload. More to the point, due to the large number of plane waves in the coarse-level box, the considerable storage required by the translation matrix becomes the bottleneck of the computation. To circumvent this problem, a method is suggested [19–21]. It is that, the translation matrix is not stored, instead is real-time computed in a way similar to the one-dimensional FMM. However, further study shows

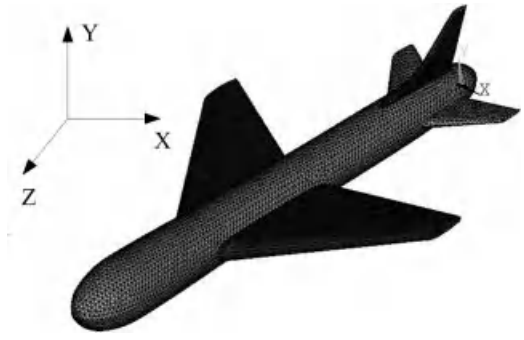


Figure 2.37 The model of the aircraft. (Reproduced with permission from X.M. Pan and X.Q. Sheng, "A sophisticated parallel MLFMA for scattering by extremely large targets," *IEEE Antennas and Propagation Magazine*, **50**, 3, 129–138, 2008. © 2008 IEEE.)

that, despite the saving on the memory, this method greatly decreases the computational efficiency. It is verified by numerical experiments that the following scheme is a good one, if the computer memory is not too restrictive. In this scheme, the translation matrices are still stored directly. However, the matrices in finer levels are distributed according to the boxes, while those in coarser levels are distributed to different nodes according to the propagation directions of the plane waves. This scheme is named the *hybrid-parallel MLFMA (hybrid-PMLFMA)*. The following examples show the numerical performance of the hybrid-PMLFMA. The first model is a PEC sphere with diameter of 200λ. It is meshed using more than 20 million unknowns. Figure 2.36 shows the comparison of the numerically calculated bistatic RCS of the sphere with the analytical Mie series solution. In the second example a simplified aircraft with length of 480λ is modeled by near 20 million unknowns. Figures 2.37 and 2.38 present the model

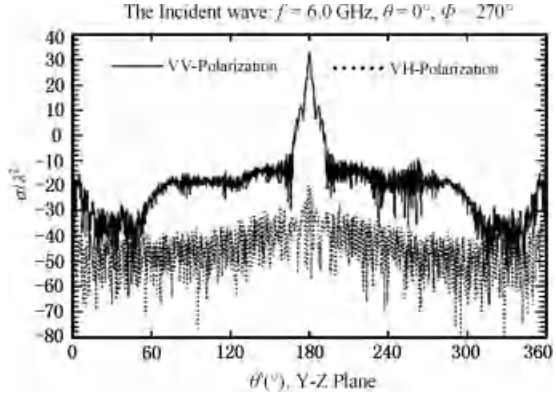


Figure 2.38 The bistatic RCS of the aircraft model. (Reproduced with permission from X.M. Pan and X.Q. Sheng, "A sophisticated parallel MLFMA for scattering by extremely large targets," *IEEE Antennas and Propagation Magazine*, **50**, 3, 129–138, 2008. © 2008 IEEE.)

Table 2.7 The computer resource used to calculate the scattering from a metallic sphere with a diameter of 200λ

Number of unknowns (million)	Total memory (GB)	Average memory of each node (GB)	Total time (hour)	Calculating the distance between boxes		Geometry information	
				Time (s)	Memory (MB)	Time (s)	Memory (MB)
23.52	56.96	1.78	3.5	7.5	198.0	71.0	45.1

and the bistatic RCS of this aircraft, respectively. Tables 2.7 and 2.8 show the computer resources used in these two examples, respectively. Both examples are performed on the parallel platform "*Liuhui*" of the Center of Electromagnetic Simulation (CEMS) in Beijing Institute of Technology (BIT). "*Liuhui*" is a Linux memory-distributed parallel cluster with 32 IBM blades. Each blade is equipped with P4 3.0 GHz processors and 4.0 GB memory. They are connected with Myrinet.

Parallel algorithms are usually needed to be redesigned and updated with the progress of computers. For example, computers with multicore CPU, which were not so mature several years ago, are now becoming a standard configuration. The current popular system of high-performance computing (HPC) is hierarchical architecture. It includes many computer nodes, which are memory distributed and connected with networks such as *infiniband*. Each computer node has many CPUs and each CPU has multicores, which share the memory. The parallel efficiency of the above presented pure MPI-based parallel MLFMA can be further improved by redesigning the parallel scheme on this kind of computer platform.

In the hybrid-parallel MLFMA, a transition level is introduced in the MLFMA-tree. For levels finer than the transition level, the MLFMA-tree is partitioned in such a way that each process gets an almost equal number of boxes. For levels coarser than the transition level, the plane waves are partitioned equally among all processes. Numerical experiments show that the choice of the transition level significantly

Table 2.8 The computer resource used to calculate the scattering from an aircraft with a length of 480λ

Number of unknowns (million)	Total memory (GB)	Average memory of each node (GB)	Total time (hour)	Calculating the distance between boxes		Geometry information	
				Time (s)	Memory (MB)	Time (s)	Memory (MB)
17.266	67.2	2.1	2.0	12.4	252.9	10.8	40.8

impacts the efficiency of the hybrid MPI parallel schemes [20]. A rule to suitably determine the transition level for the Hybrid-PMLFMA has been proposed in [20], which is the truncation number of L in (2.156) in the transition level equal to the number of processes. According to this rule, the transition level would vary with the number of processes to maximize efficiency even for the same target. In general, the transition level will be moved to the coarser level when more processes are employed. Hence, the efficiency will decrease because of the load imbalance problem arising from the levels near to the transition level in the MLFMA-tree. This explains why the scalability of the Hybrid-PMLFMA is not good for extremely large targets when too many computer nodes must be used to meet with the memory requirement.

2.1.12.2 MPI-OpenMP-PMLFMA

The multicore computers, which are of shared-memory architecture, actually boosts applications of OpenMP [22,23]. In the OpenMP parallelization, the master "thread" (each process has only one master thread) "forks" a specified number of slave "threads" and a task is divided among them. The threads share the same memory and run concurrently. The potential of the OpenMP parallelization has been partly revealed by [22], where the pure OpenMP parallel MLFMA (OpenMP-MLFMA) performs almost as efficiently as its MPI counterpart even for large targets with over 40 million unknowns. A major drawback of the OpenMP-MLFMA is that it can only work efficiently on shared-memory computers. Accordingly, its computational capability is limited by the hardware resource on a single computational node.

Based on the current hierarchical architecture of HPC computer platform, a parallel scheme of combining MPI with OpenMP can further improve parallel efficiency. In the MPI-OpenMP-PMLFMA, the MLFMA-tree is partitioned among processes to perform the MPI parallelization. Simultaneously, tasks on each process are further accelerated by employing the OpenMP multithreading parallelization. Compared with pure MPI PMLFMA, this MPI-OpenMP-PMLMFA enables us to employ the OpenMP parallelization among threads to avoid a large number of processors in the MPI parallelization. Thus, the unavoidable load imbalance in the pure MPI parallelization can be significantly depressed by combining the OpenMP parallelization.

The following will show some numerical experiments for the MPI-OpenMP-PMLFMA. These experiments are carried out on the Deep-Comp 7000 at the Chinese Academy of Sciences. Two computer platforms are used for the experiments.

- The blade cluster. It consists of IBM HS21SM servers connected by an InfiniBand network. Each server has two 4-core Xeon processors and 32 GB memory. Consequently, no more than 8 threads are forked in each process for computations on this platform. All of the experiments are carried out on it except for those about the strong scalability study.

- The IBM X3950M2 cluster. It consists of IBM X3950M2 servers connected by InfiniBand network. Each server has sixteen 4-core Xeon processors and 512GB memory. Two nodes are employed to test the strong scalability of the MPI-OpenMP-PMLFMA.

Because of the high parallel efficiency of near-field matrix filling, in the following discussion, we focus on the parallelization of far-field interaction (FFI). Let t_1 be the time for evaluating a single FFI by a single thread. Let t_C be the time for computing FFI by using $C = P \times T$ computing units, where P is the number of processes, and T is the number of threads forked in each process. Then, the parallel efficiency is defined as

$$\eta = \frac{t_1}{C \times t_C} \times 100\% \tag{2.188}$$

The efficiency of the MPI-OpenMP-PMLFMA is presented by computing scattering from two PEC spheres with different sizes, namely, *Sph-120* and *Sph-220*. *Sph-120* is a sphere with a diameter of 120λ, modeled by 11 267 532 unknowns. *Sph-220* is of 220λ in diameter, modeled by 36 792 012 unknowns.

2.1.13 Strong Scalability

The following numerical experiments are performed on the IBM X3950M2 cluster to show the strong scalability by increasing the number of computing units for a fixed problem. Four cases, where the number of threads varies from 1 to 8, are tested in the computations. These four cases are, respectively, case (1): 1 thread; case (2): 2 threads; case (3): 4 threads; case (4): 8 threads. The parallel efficiency for *Sph-220* is presented in Figure 2.39. As indicated in this figure, the parallel efficiency increases as more threads are forked. In particular, when the number of computing units reaches 128, the efficiency for the *Sph-220* in the pure MPI case (or 1-thread case) drops to be less than 65%, while that in the 8-thread case can still reach over 80%.

In fact, combining with the OpenMP, parallelization in the MPI not only benefits us in choosing a suitable transition level but also in cutting down the MPI overhead, as shown in Table 2.9. Compared with the pure MPI result, the MPI overhead is almost reduced by a factor of 2.0 when 2 threads are forked. The total MPI overhead consists of the overhead arising from communication operations and that from synchronizations. The former generally contributes little to the total overhead in FFI because most communication is carried out simultaneously in computations. The latter, arising from load imbalance, cannot be alleviated in the pure MPI parallel scheme. On the contrary, variations on workloads among processes can be deflated by using the OpenMP multithreading acceleration. This is another benefit offered by the MPI-OpenMP-PMLFMA.

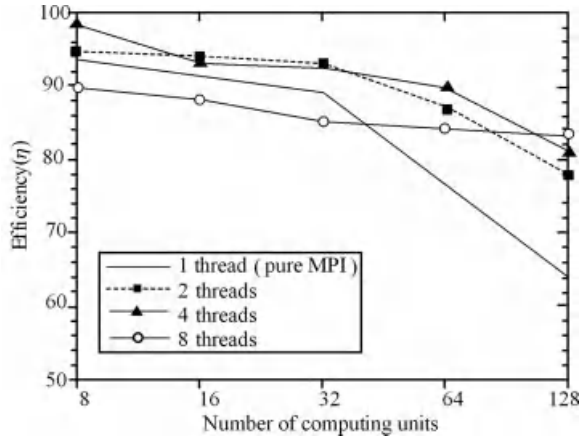


Figure 2.39 Strong scaling of the MPI-OpenMP-PMLFMA.

2.1.14 Weak Scalability

The following numerical experiments are performed to show the weak scalability of the MPI-OpenMP-PMLFMA by increasing the number of computing units while the number of unknowns on each computing unit remains approximately the same. The detailed information for the set of spheres is listed in Table 2.10. The CPU time for one FFI increases with the growth of the computing units, as shown in Figure 2.40. The explanation lies in the $O(N \log N)$ complexity of the MLFMA. Even though the number of unknowns on each computing unit approximately remains constant, the corresponding computational task increases at a speed of $O(N \log N)$. As a result, solution time will keep growing as the size of a target increases. The solution time cannot be greatly reduced by employing more processes because the parallel efficiency would decrease rapidly as revealed in the experiments on the strong scalability. On the contrary, to increase the number of threads is a good choice to relieve this bottleneck due to the benefits offered by the OpenMP parallelization.

Table 2.9 CPU time (s) used by different parts in one FFI, including the MPI overhead, when different number of threads are employed for *Sph-220.*

No.	Time for aggregation	Time for translation	Time for disaggregation	MPI overhead
1	5.51	12.75	6.37	3.57
2	2.87	6.57	3.39	1.78
4	1.53	3.37	1.81	0.99
8	0.88	1.82	0.99	0.61

Table 2.10 Information on spheres of different sizes

Diameter (wavelength)	Unknowns	No. of computing units
20	419 628	2
30	842 700	4
44	1 687 500	8
64	3 383 532	16
90	6 482 700	32
126	12 411 468	64
184	25 895 532	128
260	51 170 700	256

2.1.14.1 Capability of the MPI-OpenMP-PMLFMA

A PEC sphere with a diameter of 1200λ is employed to demonstrate the capability of the MPI-OpenMP-PMLFMA. This sphere, denoted by *Sph-1200*, involves more than 1 billion unknowns with an average mesh size of 0.12λ. In the computation, a 13-level MLFMA, where the 0th level is the coarsest one, is employed with the memory requirement of about 3.9 TB. In the computation, the 2nd–4th levels are parallelized by distributing FFPs and the other levels are parallelized by distributing boxes evenly among 200 processes. Each process spawns 8 threads to conduct the OpenMP acceleration. Table 2.11 details computational resources used in this calculation. Figure 2.41 plots the calculated bistatic RCS in comparison with the Mie series results. Since the forward scattering varies sharply, only this part of RCS is presented. Statistics show that the root-mean-square error of the RCS is about 0.8 dBsm.

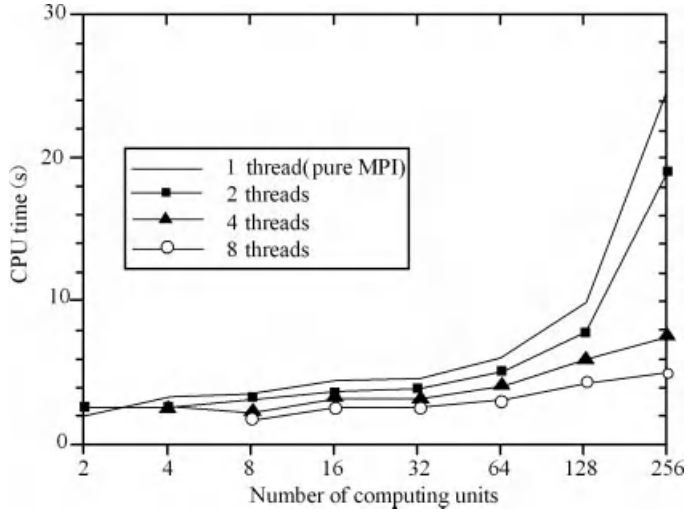


Figure 2.40 Weak scaling of the MPI-OpenMP-PMLFMA.

Table 2.11 Computational resources used by the MPI-OpenMP-PMLFMA

General information		
	Sph-1200	Airplane
Geometry size (wavelength)	1200	1600
Number of processes × threads	200 × 8	128 × 8
Number of levels	13	13
Size of the finest box (wavelength)	0.29	0.39
Total memory (TB)	3.9	1.3
Residual error	0.001	0.001
Number of iterations	106	83
Total time (Min)	131	63
Memory requirement (GB)		
	Sph-1200	Airplane
MLFMA tree and geometry data	298.0	27.8
Near-field interaction matrix	1740.1	740.4
Aggregation/disaggregation matrix	552.2	150.3
Interpolation matrix	87.4	59.9
Translation matrix	150.0	41.8
Far-field pattern	369.8	82.7
GMRES solution	553.2	150.2
CPU time (s)		
	Sph-1200	Airplane
Geometry information	46.1	6.3
Determination of truncation number	72.3	36.8
Setup of near-field interaction matrix	409.2	1456.1
Setup of far-field interaction matrix	370.8	112.2
One matrix-vector multiplication	41.9	15.1

To further demonstrate the capability of the MPI-OpenMP-PMLFMA, RCS from an airplane model is calculated. This is a target with complex geometrical shapes and has a largest dimension of 1600λ, involving 288 151 344 unknowns with an average mesh size of 0.12λ. A 13-level MLFMA is employed with the memory requirement of over 1.3 TB. In the computation, the 2nd–5th levels are parallelized by distributing FFPs and the other levels are parallelized by distributing boxes evenly among 128 processes. Each process spawns 8 threads to conduct the OpenMP acceleration. Table 2.11 details computational resources used in this calculation. Figure 2.42 plots the obtained bistatic RCS with an incident from the direction of $(0^\circ, 0^\circ)$.

2.2 Scattering from Three-Dimensional Homogeneous Dielectric Objects

From a physical point of view, the scattering from homogeneous dielectrics is totally different from that from a PEC body. However, from a numerical point of view, they are

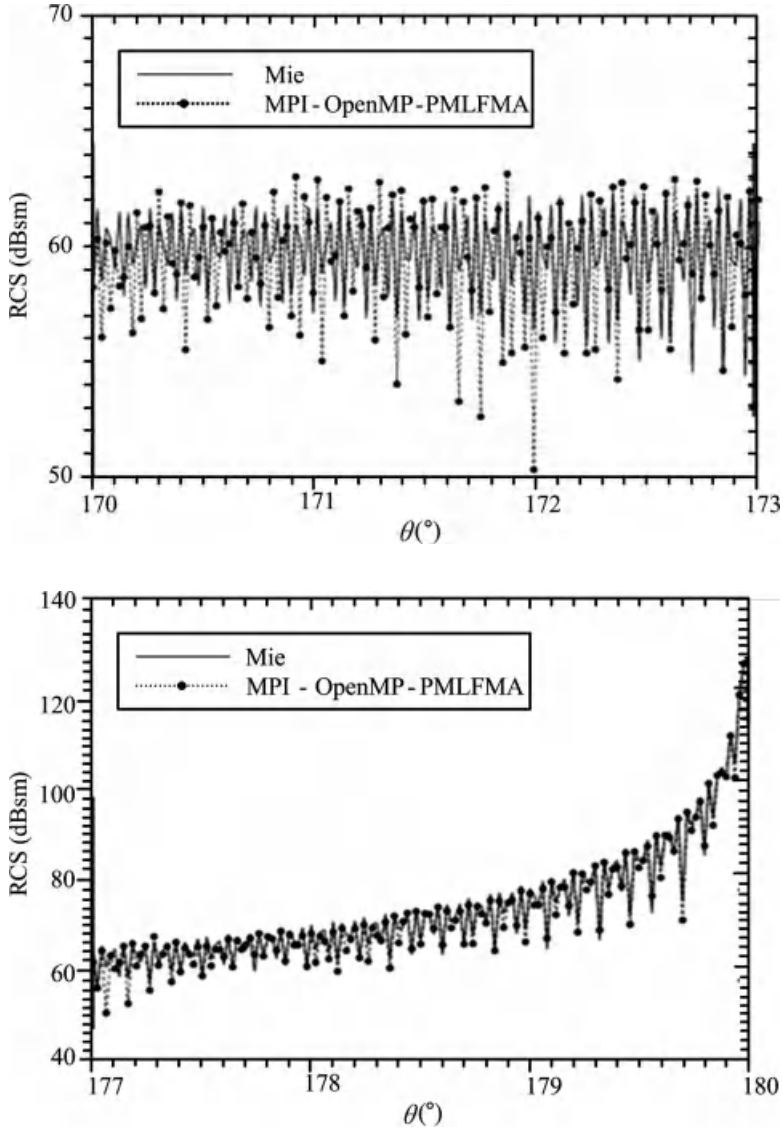


Figure 2.41 Bistatic RCS from *Sph-1200*.

similar to some extent. For example, the RWG basis functions, the singularity process techniques, and MLFMA that are used for the PEC problems can be imported readily to the MoM for solving dielectric problems. Of course, there are differences between them. Since the integral equations for these two kinds of problems are different, the discretized linear systems of equations are numerically distinct, which in turn leads to a large difference in the convergence speed when applying iteration methods. The difference, but not the analogy will be explained with great emphasis in this section.

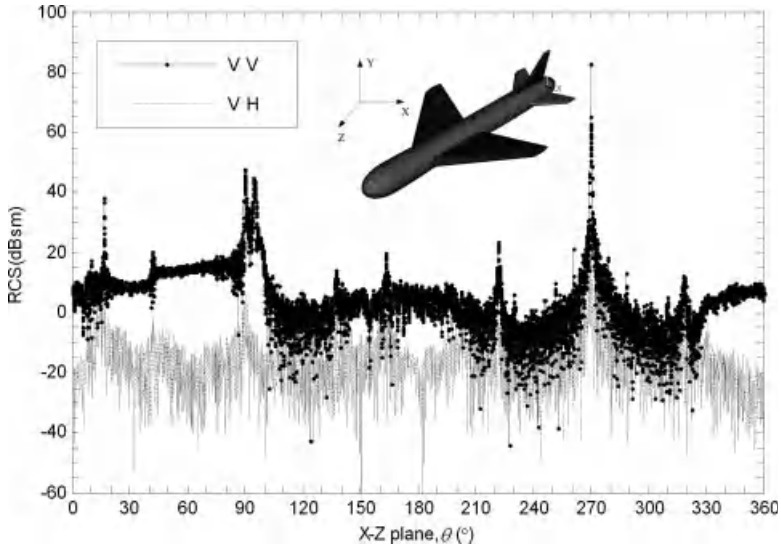


Figure 2.42 RCS from the airplane model.

2.2.1 *Mathematic Formulation of the Problem*

Similar to the scattering from a PEC body, here we consider a plane wave as the incident wave, with the mathematic expression given by (2.1) and (2.2). The integral equation for the scattering from homogeneous dielectrics has been constructed in Section 1.3.4. There are many types of the integral equations. They can be categorized into three types. The first type is a system of equations made up of two equations coming from the interior and the exterior regions, respectively. For example, the EFIE system made up of the interior EFIE and the exterior EFIE is written as:

$$\mathbf{M} + \mathbf{n} \times [Z_1 \mathbf{L}_1(\mathbf{J}) - \mathbf{K}_1(\mathbf{M})] = -\mathbf{n} \times \mathbf{E}^i \tag{2.189}$$

$$\mathbf{M} + \mathbf{n} \times [Z_2 \mathbf{L}_2(-\mathbf{J}) - \mathbf{K}_2(-\mathbf{M})] = 0 \tag{2.190}$$

The MFIE system made up of the interior MFIE and the exterior MFIE is written as:

$$\mathbf{J} - \mathbf{n} \times \left[\frac{1}{Z_1}\mathbf{L}_1(\mathbf{M}) + \mathbf{K}_1(\mathbf{J})\right] = \mathbf{n} \times \mathbf{H}^i \tag{2.191}$$

$$\mathbf{J} - \mathbf{n} \times \left[\frac{1}{Z_2}\mathbf{L}_2(-\mathbf{M}) + \mathbf{K}_2(-\mathbf{J})\right] = 0 \tag{2.192}$$

The second type is also a system of equations made up of two equations. But both equations consist of operators of the interior and the exterior regions. These are the famous PMCHW equations, written as:

$$-(Z_1\mathbf{L}_1 + Z_2\mathbf{L}_2)(\mathbf{J}) + (\mathbf{K}_1 + \mathbf{K}_2)(\mathbf{M}) = \mathbf{E}^i \tag{2.193}$$

$$-(\mathbf{K}_1 + \mathbf{K}_2)(\mathbf{J}) - \left(\frac{1}{Z_1}\mathbf{L}_1 + \frac{1}{Z_2}\mathbf{L}_2\right)(\mathbf{M}) = \mathbf{H}^i \tag{2.194}$$

The third type is a single equation with only one unknown. Although the number of equations is reduced, the equation gets much more complex. In particular, because the order of the operator increased, the selection of the basis functions and the singularity processing becomes more difficult. To overcome these difficulties, a weak discretized technique in [24] is usually used, which leads to the computation load of the single-equation form actually equivalent to that of the two-equation form. The following will first discuss the implementation and the numerical property of the first two types in Sections 2.2.2 and 2.2.3, respectively, and then the third type in Section 2.2.4.

2.2.2 *Discretized Forms and Their Numerical Performance*

As is done with a PEC object, we discretize the surface of the dielectrics using triangular patches, with the equivalent electric current $\mathbf{J}$ and the equivalent magnetic current $\mathbf{M}$ being expressed with the RWG basis functions, written as,

$$\mathbf{J} = \sum_{i=1}^{N_S} J_i \mathbf{g}_i \tag{2.195}$$

$$\mathbf{M} = \sum_{i=1}^{N_S} M_i \mathbf{g}_i \tag{2.196}$$

where N_S is the total edge number of the triangles in the computational domain S. $\mathbf{g}_i$ denotes the RWG vector basis function. Substituting (2.195)–(2.196) into (2.189)–(2.190), and using $\mathbf{g}_i$ as the test function, the following discretized system of equations can be obtained:

$$\left[P_1^{TE}\right]\{J\} + \left[Q_1^{TE}\right]\{M\} = \left\{b^{TE}\right\} \tag{2.197}$$

$$\left[P_2^{TE}\right]\{J\} + \left[Q_2^{TE}\right]\{M\} = \{0\} \tag{2.198}$$

where

$$P_{1ij}^{TE} = Z_1 \int_S \mathbf{g}_i \cdot \left[\hat{n} \times \mathbf{L}_1(\mathbf{g}_j)\right] dS \tag{2.199}$$

$$Q_{1ij}^{TE} = \int_S \mathbf{g}_i \cdot \left[\mathbf{g}_j - \hat{n} \times \mathbf{K}_1(\mathbf{g}_j)\right] dS \tag{2.200}$$

$$P_{2ij}^{TE} = -Z_2 \int_S \mathbf{g}_i \cdot \left[\hat{n} \times \mathbf{L}_2(\mathbf{g}_j)\right] dS \tag{2.201}$$

$$Q_{2ij}^{TE} = -\int_S \mathbf{g}_i \cdot \left[\mathbf{g}_j - \hat{n} \times \mathbf{K}_2(\mathbf{g}_j)\right] dS \tag{2.202}$$

$$b_i^{TE} = -\int_S \mathbf{g}_i \cdot (\hat{n} \times \mathbf{E}^i) dS \tag{2.203}$$

As introduced in Section 2.1.4, the above discretized equations are a *T-EFIE* system (abbreviated as *TE* in the equation notation), since the testing functions are tangential to the basis functions. We can also use the testing function of $\hat{n} \times \mathbf{g}_i$. Then, by applying the vector identity $(\hat{n} \times \mathbf{a}) \cdot (\hat{n} \times \mathbf{b}) = \mathbf{a} \cdot \mathbf{b}$ (when $\hat{n}$ is orthogonal to $\mathbf{a}$ or $\mathbf{b}$), we arrive at the *N-EFIE (NE)* system:

$$\left[P_1^{NE}\right]\{J\} + \left[Q_1^{NE}\right]\{M\} = \left\{b^{NE}\right\} \tag{2.204}$$

$$\left[P_2^{NE}\right]\{J\} + \left[Q_2^{NE}\right]\{M\} = \{0\} \tag{2.205}$$

where

$$P_{1ij}^{NE} = Z_1 \int_S \mathbf{g}_i \cdot \mathbf{L}_1(\mathbf{g}_j) dS \tag{2.206}$$

$$Q_{1ij}^{NE} = \int_S \mathbf{g}_i \cdot \left[\hat{n} \times \mathbf{g}_j - \mathbf{K}_1(\mathbf{g}_j)\right] dS \tag{2.207}$$

$$P_{2ij}^{NE} = -Z_2 \int_S \mathbf{g}_i \cdot \mathbf{L}_2(\mathbf{g}_j) dS \tag{2.208}$$

$$Q_{2ij}^{NE} = \int_S \mathbf{g}_i \cdot \left[\hat{n} \times \mathbf{g}_j - \mathbf{K}_2(\mathbf{g}_j)\right] dS \tag{2.209}$$

$$b_i^{NE} = -\int_S \mathbf{g}_i \cdot \mathbf{E}^i dS \tag{2.210}$$

Analogously, we can discretized the MFIE to obtain the discretized MFIE system. In contrast to the single operator involved in solving for PEC scatterer problem, the EFIE and MFIE systems for solving dielectric problems bear considerate numerical behavior due to the concurrent existence of operators **L** and **K** in each equation. Also, we can substitute (2.195)–(2.196) into (2.193)–(2.194), and use $\mathbf{g}_i$ as the testing function to obtain the *tangential PMCHW system* (*TPMCHW*):

$$\left[P_1^{TP}\right]\{J\} + \left[Q_1^{TP}\right]\{M\} = \left\{b_1^{TP}\right\} \tag{2.211}$$

$$\left[P_2^{TP}\right]\{J\} + \left[Q_2^{TP}\right]\{M\} = \left\{b_2^{TP}\right\} \tag{2.212}$$

where

$$P_{1ij}^{TP} = -\int_S \mathbf{g}_i \cdot \left[Z_1 \mathbf{L}_1(\mathbf{g}_j) + Z_2 \mathbf{L}_2(\mathbf{g}_j)\right] dS \tag{2.213}$$

$$Q_{1ij}^{TP} = \int_S \mathbf{g}_i \cdot \left[\mathbf{K}_1(\mathbf{g}_j) + \mathbf{K}_2(\mathbf{g}_j)\right] dS \tag{2.214}$$

$$P_{2ij}^{TP} = -\int_S \mathbf{g}_i \cdot \left[\mathbf{K}_1(\mathbf{g}_j) + \mathbf{K}_2(\mathbf{g}_j)\right] dS \tag{2.215}$$

$$Q_{2ij}^{TP} = -\int_S \mathbf{g}_i \cdot \left[\frac{1}{Z_1}\mathbf{L}_1(\mathbf{g}_j) + \frac{1}{Z_2}\mathbf{L}_2(\mathbf{g}_j)\right] dS \tag{2.216}$$

$$b_{1i}^{TP} = \int_S \mathbf{g}_i \cdot \mathbf{E}^i dS \tag{2.217}$$

$$b_{2i}^{TP} = \int_S \mathbf{g}_i \cdot \mathbf{H}^i dS \tag{2.218}$$

Substituting (2.195)–(2.196) into (2.193)–(2.194), then using $\hat{n} \times \mathbf{g}_i$ as the testing function, we can obtain the *normal PMCHW System* (*NPMCHW*):

$$\left[P_1^{NP}\right]\{J\} + \left[Q_1^{NP}\right]\{M\} = \left\{b_1^{NP}\right\} \tag{2.219}$$

$$\left[P_2^{NP}\right]\{J\} + \left[Q_2^{NP}\right]\{M\} = \left\{b_2^{NP}\right\} \tag{2.220}$$

where

$$P_{1ij}^{NP} = -\int_S \hat{n} \times \mathbf{g}_i \cdot \left[Z_1 \mathbf{L}_1(\mathbf{g}_j) + Z_2 \mathbf{L}_2(\mathbf{g}_j)\right] dS \tag{2.221}$$

$$Q_{1ij}^{NP} = \int_S \hat{n} \times \mathbf{g}_i \cdot \left[\mathbf{K}_1(\mathbf{g}_j) + \mathbf{K}_2(\mathbf{g}_j)\right] dS \tag{2.222}$$

$$P_{2ij}^{NP} = -\int_S \hat{n} \times \mathbf{g}_i \cdot \left[\mathbf{K}_1(\mathbf{g}_j) + \mathbf{K}_2(\mathbf{g}_j)\right] dS \tag{2.223}$$

$$Q_{2ij}^{NP} = -\int_S \hat{n} \times \mathbf{g}_i \cdot \left[\frac{1}{Z_1}\mathbf{L}_1(\mathbf{g}_j) + \frac{1}{Z_2}\mathbf{L}_2(\mathbf{g}_j)\right] dS \tag{2.224}$$

$$b_{1i}^{NP} = -\int_S \hat{n} \times \mathbf{g}_i \cdot \mathbf{E}^i dS \tag{2.225}$$

$$b_{2i}^{NP} = -\int_S \hat{n} \times \mathbf{g}_i \cdot \mathbf{H}^i dS \tag{2.226}$$

From the analysis of the matrix entries in Section 2.1.4, we know that the matrices $\left[P_i^{NE}\right]$, $\left[Q_i^{TE}\right]$, $\left[P_1^{TP}\right]$, $\left[Q_1^{NP}\right]$, $\left[P_2^{NP}\right]$ and $\left[Q_2^{TP}\right](i = 1, 2)$ are quasidiagonally dominant matrices, while $\left[P_i^{TE}\right]$, $\left[Q_i^{NE}\right]$, $\left[P_1^{NP}\right]$, $\left[Q_1^{TP}\right]$, $\left[P_2^{TP}\right]$ and $\left[Q_2^{NP}\right]$ are off-diagonally dominant matrices. Besides, since $\left[Q_i^{TE}\right]$, $\left[Q_1^{NP}\right]$ and $\left[P_2^{NP}\right]$ come from the Fredholm integral equation of the second kind, they should bear a better condition number than $\left[P_i^{NE}\right]$, $\left[P_1^{TP}\right]$ and $\left[Q_2^{TP}\right]$, which come from the Fredholm integral equation of the first kind. These facts can be illustrated using mathematic notations as

$$\left[Q_i^{TE}\right] \sim \left[Q_1^{NP}\right] \sim \left[P_2^{NP}\right] \sim \begin{bmatrix} \ddots & & \\ & 2 & \\ & & \ddots \end{bmatrix}, \quad \left[P_i^{NE}\right] \sim \left[P_1^{TP}\right] \sim \left[Q_2^{TP}\right] \sim \begin{bmatrix} \ddots & & \\ & 1 & \\ & & \ddots \end{bmatrix}$$

$$\left[P_i^{TE}\right] \sim \left[Q_i^{NE}\right] \sim \begin{bmatrix} \ddots & & \\ & 0 & \\ & & \ddots \end{bmatrix}, \quad \left[P_1^{NP}\right] \sim \left[Q_1^{TP}\right] \sim \left[P_2^{TP}\right] \sim \left[Q_2^{NP}\right] \sim \begin{bmatrix} \ddots & & \\ & 0 & \\ & & \ddots \end{bmatrix} \tag{2.227}$$

where the notations "2," "1," and "0" indicate the strength of dominance instead of the exact values of the numbers. Thus, the T-EFIE for homogeneous dielectric bodies can be roughly expressed as:

$$\begin{bmatrix} P_1 & Q_1 \\ P_2 & Q_2 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} \ddots & & \\ & 0 & \\ & & \ddots \end{bmatrix} & \begin{bmatrix} \ddots & & \\ & 2 & \\ & & \ddots \end{bmatrix} \\ \begin{bmatrix} \ddots & & \\ & 0 & \\ & & \ddots \end{bmatrix} & \begin{bmatrix} \ddots & & \\ & 2 & \\ & & \ddots \end{bmatrix} \end{bmatrix} \tag{2.228}$$

The N-EFIE is roughly expressed as:

$$\begin{bmatrix} P_1 & Q_1 \\ P_2 & Q_2 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} \ddots & & \\ & 1 & \\ & & \ddots \end{bmatrix} & \begin{bmatrix} \ddots & & \\ & 0 & \\ & & \ddots \end{bmatrix} \\ \begin{bmatrix} \ddots & & \\ & 1 & \\ & & \ddots \end{bmatrix} & \begin{bmatrix} \ddots & & \\ & 0 & \\ & & \ddots \end{bmatrix} \end{bmatrix} \tag{2.229}$$

The structures of the matrices in (2.228) and (2.229) reveal that these matrices are not block-diagonal dominant, and likely ill-conditioned. In fact, it is verified by numerical experiments in the following section that, both T-EFIE and N-EFIE are ill-conditioned equation systems. In order to solve this problem, we add the T-EFIE and N-EFIE systems and obtain the *combined EFIE system* (*TENE*). The matrix in TENE is written as

$$\begin{bmatrix} P_1 & Q_1 \\ P_2 & Q_2 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} \ddots & & \\ & 1 & \\ & & \ddots \end{bmatrix} & \begin{bmatrix} \ddots & & \\ & 2 & \\ & & \ddots \end{bmatrix} \\ \begin{bmatrix} \ddots & & \\ & 1 & \\ & & \ddots \end{bmatrix} & \begin{bmatrix} \ddots & & \\ & 2 & \\ & & \ddots \end{bmatrix} \end{bmatrix} \tag{2.230}$$

Although this matrix is no longer ill-conditioned, it suffers from interior resonance. This is because both the source equation systems are EFIEs. Of course we can combine TENE with the *combined MFIE system* (*THNH*). The resulting equation system is neither ill-conditioned nor suffers from interior resonance. But due to the complexity we are not going to present it here. The interested reader is referred to [25] for details.

What we will explain next is the popularly used PMCHW system of equations. It can be easily found out that the matrix in the tangential discretized system of equations can be expressed as:

$$\begin{bmatrix} P_1 & Q_1 \\ P_2 & Q_2 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} \ddots & & \\ & 1 & \\ & & \ddots \end{bmatrix} & \begin{bmatrix} \ddots & & \\ & 0 & \\ & & \ddots \end{bmatrix} \\ \begin{bmatrix} \ddots & & \\ & 0 & \\ & & \ddots \end{bmatrix} & \begin{bmatrix} \ddots & & \\ & 1 & \\ & & \ddots \end{bmatrix} \end{bmatrix} \tag{2.231}$$

The matrix in the normal discretized PMCHW system of equations is written as:

$$\begin{bmatrix} P_1 & Q_1 \\ P_2 & Q_2 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} \ddots & & \\ & 0 & \\ & & \ddots \end{bmatrix} & \begin{bmatrix} \ddots & & \\ & 2 & \\ & & \ddots \end{bmatrix} \\ \begin{bmatrix} \ddots & & \\ & 2 & \\ & & \ddots \end{bmatrix} & \begin{bmatrix} \ddots & & \\ & 0 & \\ & & \ddots \end{bmatrix} \end{bmatrix} \tag{2.232}$$

Both matrices are quasidiagonally dominant matrices. They should bear low condition numbers, and the normal PMCHW one is even lower. Moreover, neither of these two systems of equations suffers interior resonance. This is because each equation in the PMCHW system is made up from the interior operators and the exterior operators, and there is no resonant frequency in the exterior unlimited space.

Problems

2.7 The PMCHW formulation is the combination of the interior EFIE with exterior EFIE, and the interior MFIE with the exterior MFIE. Following the same idea, but with different combination coefficients, we can obtain

$$(\mathbf{L}_1 + \mathbf{L}_2)(\mathbf{J}) - \left(\frac{1}{Z_2}\mathbf{K}_1 + \frac{1}{Z_2}\mathbf{K}_2\right)(\mathbf{M}) = -\frac{1}{Z_1}\mathbf{E}^i$$

$$(Z_1\mathbf{K}_1 + Z_2\mathbf{K}_2)(\mathbf{J}) + (\mathbf{L}_1 + \mathbf{L}_2)(\mathbf{M}) = -Z_1\mathbf{H}^i$$

which is usually named the CTF formulation. Compare the numerical performance of CTF with PMCHW.

2.8 Following a similar idea, we can obtain the following CNF formulation

$$\hat{n} \times (\mathbf{K}_1 - \mathbf{K}_2)(\mathbf{M}) + \hat{n} \times (-Z_1\mathbf{L}_1 + Z_2\mathbf{L}_2)(\mathbf{J}) = -\hat{n} \times \mathbf{E}^i(\mathbf{r})$$

$$\hat{n} \times (\mathbf{K}_1 - \mathbf{K}_2)(\mathbf{J}) + \hat{n} \times \left(\frac{1}{Z_1}\mathbf{L}_1 + \frac{1}{Z_2}\mathbf{L}_2\right)(\mathbf{M}) = -\hat{n} \times \mathbf{H}^i(\mathbf{r})$$

Combining CNF with CTP yields the JMCFIE formulation. Analyze the numerical performance of JMCFIE.

2.2.3 *Numerical Examples*

In the previous section, we analyzed the numerical property of various discretized integral equations. In this section, we will use numerical experiments to verify our analysis. First, let us calculate the RCS of a dielectric sphere with radius $r = 0.5\lambda_0$ (where λ_0 is the free-space wavelength) and relative permittivity $\varepsilon_r = 4.0$. The sphere is discretized by 1152 triangular patches, with 1728 edges of an average length of $0.08\lambda_0$. Figures 2.43a and b present the RCS curves calculated by using different DIEs with reference of the theoretical results given by the Mie series. It can be seen that the discretized TENE, TPMCHW and NPMCHW results agree well with the theoretical result, while the TE and NE routines give totally wrong results. Since the condition numbers in the discretized TE and NE equations are greater than 1×10^6, the accumulated error introduced by using single precision *LU* algorithm ruins the results. If we use the CG method instead of the *LU* algorithm, we cannot get a convergent result at all. The iterative convergence histories for the discretized TENE, TPMCHW and NPMCHW are plotted in Figure 2.44. These numerical experiments verify that the discretized TE and NE equations are ill-conditioned ones that hardly work, while the discretized TENE, TPMCHW and NPMCHW equations are normal ones. Furthermore, the NPMCHW has the best condition number, which is followed by TPMCHW, and TENE has the worst condition number.

Next, let us calculate the resonant case. The model is a dielectric sphere with radius $r = 0.444\lambda_0$ and $\varepsilon_r = 4.0$. Figure 2.45 shows the bistatic RCS calculated by using the discretized TENE and TPMCHW equations. They are compared with the theoretical results. As can be seen, under the resonant frequency, the discretized TENE is ill-conditioned and gives a totally wrong result, while the discretized TPMCHW equation still provides very accurate results. This verified that the interior resonant problem resides in TENE but not in TPMCHW. A dielectric sphere with $r = 0.45$ m and $\varepsilon_r = 4.0$ is computed using the CG method. The iteration numbers under different working frequencies are compared in Figure 2.46. It is clearly shown that the discretized TENE equation suffers the resonant problem, while the TPMCHW and NPMCHW equations do not.

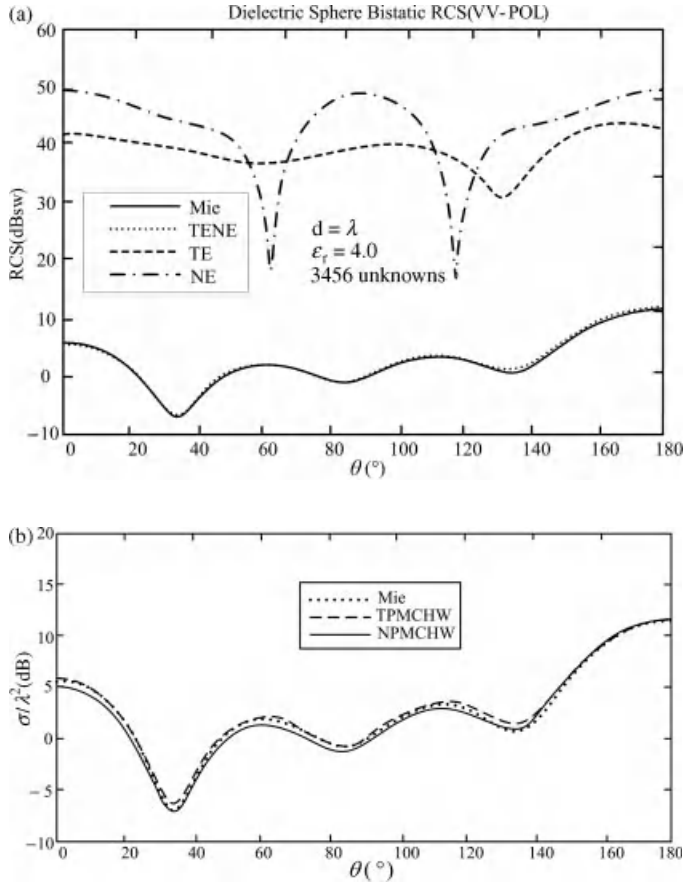


Figure 2.43 RCS curves calculated by using the five different DIEs with reference of the theoretical results given by the Mie series. (Reproduced with permission from X.Q. Sheng, J. M. Jin, J.M. Song *et al.*, "Solution of combined-field integral equation using multilevel fast multipole algorithm for scattering by homogeneous bodies," *IEEE Transactions on Antennas and Propagation*, **46**, 11, 1718–1726, 1998. © 1998 IEEE.)

The following numerical experiments aim at comparing the discretized TPMCHW and NPMCHW equations. The dielectric sphere with radius $r = 0.5\lambda_0$ and relative permittivity $\varepsilon_r = 4.0$ is discretized by four kinds of meshes. The total edge numbers in these meshes are 768, 1200, 2700 and 4800, respectively. We use the CG method to solve the discretized TPMCHW and NPMCHW equations under these meshes. The iteration numbers to achieve convergence are plotted in Figure 2.47. It can be seen, with decreasing frequency, that the discretized TPMCHW equation tends to be ill-conditioned, while the discretized NPMCHW equation behaves in a relatively steady way.

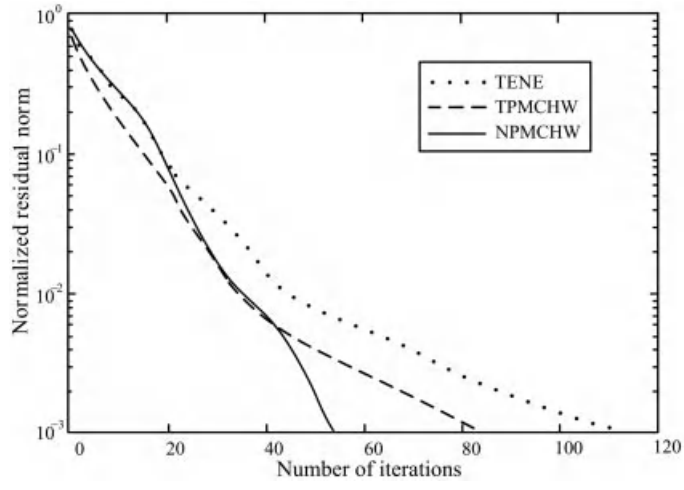


Figure 2.44 The iterative convergence curves for the discretized TENE, TPMCHW and NPMCHW using the CG method.

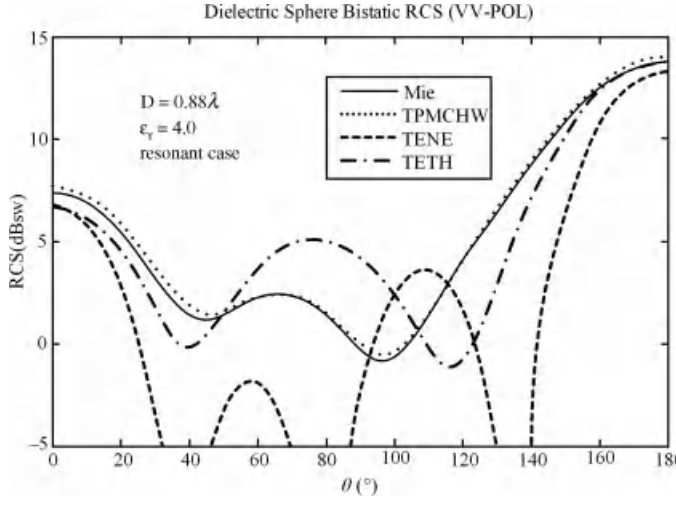


Figure 2.45 The bistatic RCS of the dielectric sphere calculated by the TENE and TPMCHW schemes. (Reproduced with permission from X.Q. Sheng, J.M. Jin, J.M. Song *et al.*, "Solution of combined-field integral equation using multilevel fast multipole algorithm for scattering by homogeneous bodies," *IEEE Transactions on Antennas and Propagation*, **46**, 11, 1718–1726, 1998. © 1998 IEEE.)

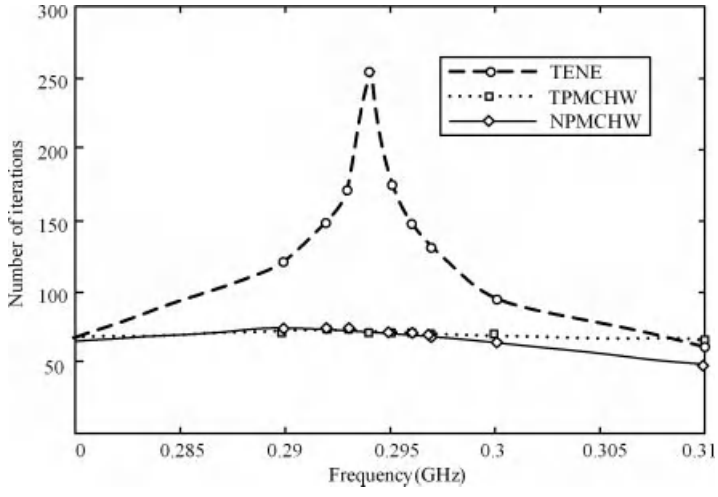


Figure 2.46 The iteration numbers as functions of frequency used in the CG method to solve different equation systems.

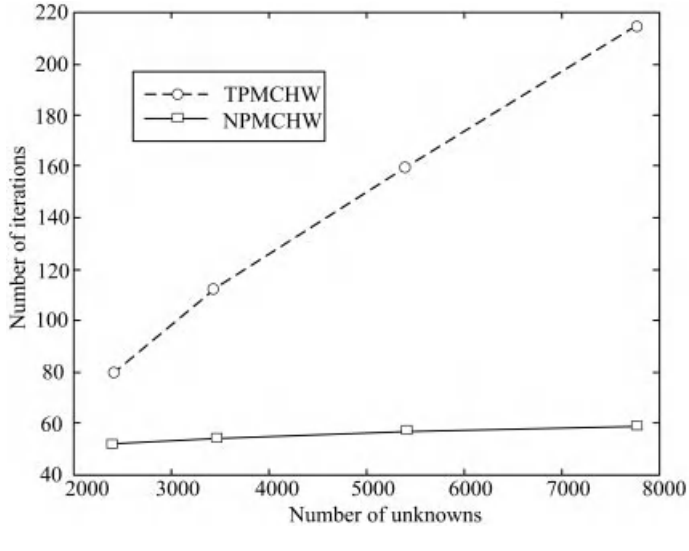


Figure 2.47 The iteration numbers to achieve convergence in the CG method as a function of the spatial resolution.

2.2.4 *Implementation of Single Integral Equation and the Numerical Characteristics*

From Section 1.3.4, we know that beside the system of the aforementioned two integral equations PMCHW, there are other systems of single integral equation for the homogeneous dielectric scatterer problem. They are the single EFIE

$$Z_2\mathbf{L}_2(\mathbf{J}_e) - Z_1\mathbf{L}_1(\mathbf{n}_1 \times \mathbf{K}_2(\mathbf{J}_e)) + \mathbf{K}_1(Z_2\mathbf{L}_2(\mathbf{J}_e) \times \mathbf{n}_1) = \mathbf{E}^i \tag{2.233}$$

and the single MFIE

$$\mathbf{K}_2(\mathbf{J}_e) - \frac{1}{Z_1}\mathbf{L}_1(Z_2\mathbf{L}_2(\mathbf{J}_e) \times \mathbf{n}_1) - \mathbf{K}_1(\mathbf{n}_1 \times \mathbf{K}_2(\mathbf{J}_e)) = \mathbf{H}^i \tag{2.234}$$

In contrast to the system of two integral equations, the integro-differential operators $\mathbf{L}$ and $\mathbf{K}$ in the single integral equation (2.233) or (2.234) are imposed twice, which are equivalent to higher-order integro-differential operators. As a result, it is quite challenging to discretize them directly, so weak solution discretization is widely used instead. In the following, we will illustrate the discretization of the single integral equation by using the example of (2.234). Let us introduce intermediate variables $\mathbf{E}, \mathbf{H}$ satisfying

$$\mathbf{E} = \mathbf{L}_2(\mathbf{J}_e) \tag{2.235}$$

$$\mathbf{H} = \mathbf{K}_2(\mathbf{J}_e) \tag{2.236}$$

Substituting (2.235) and (2.236) into (2.234) gives rise to

$$\mathbf{H} - \frac{1}{Z_1}\mathbf{L}_1(Z_2\mathbf{E} \times \mathbf{n}_1) - \mathbf{K}_1(\mathbf{n}_1 \times \mathbf{H}) = \mathbf{H}^i \tag{2.237}$$

First, let us discretize (2.235) and (2.236), that is, use triangular patches to model the surface of the homogeneous dielectrics, and use the RWG basis function to express the equivalent electric current $\mathbf{J}_e$. In order to explicitly express the tangential electric and magnetic components on every edge, we choose the tangential unit vector of each edge as the testing functions. Thus, (2.235) and (2.236) become:

$$\{E\} = \left[W^E\right]\{J_e\} \tag{2.238}$$

$$\{H\} = \left[W^M\right]\{J_e\} \tag{2.239}$$

where

$$W_{ij}^E = -\frac{1}{l_i}\int_{l_i} \hat{l}_i \cdot K_2(\vec{\mathbf{g}}_j)\, dl \tag{2.240}$$

$$W_{ij}^{M} = \frac{1}{l_i} Z_2 \int_{l_i} \hat{l}_i \cdot L_2(\mathbf{g}_j)\, dl \tag{2.241}$$

and $\hat{l}_i$ is the tangential unit vector corresponding to the ith edge. In the discretization of (2.237), we express $\hat{n} \times \mathbf{E}$ and $\hat{n} \times \mathbf{H}$ using the RWG basis functions. In order to examine the matrix equations under different testing functions, we use $\hat{l}_i$ and $\hat{n} \times \mathbf{g}_i$ as testing functions to discretize (2.237). They correspond to the line match and the Galerkin match, respectively.

$$\left(\left[P^{LM}\right]\right)\{E\} + \left[Q^{LM}\right]\{H\} = \left\{b^{LM}\right\} \tag{2.242}$$

$$\left(\left[P^{GM}\right]\right)\{E\} + \left[Q^{GM}\right]\{H\} = \left\{b^{GM}\right\} \tag{2.243}$$

Substituting (2.238) and (2.239) into (2.242) and (2.243) respectively, yields:

$$\left\{\left[P^{LM}\right]\left[W^{E}\right] + \left[Q^{LM}\right]\left[W^{M}\right]\right\}\{J_e\} = \left\{b^{LM}\right\} \tag{2.244}$$

$$\left\{\left[P^{GM}\right]\left[W^{E}\right] + \left[Q^{GM}\right]\left[W^{M}\right]\right\}\{J_e\} = \left\{b^{GM}\right\} \tag{2.245}$$

Analogously, using testing functions $\hat{n} \times \hat{l}_i$ and $\mathbf{g}_i$ to discretize (2.233) yields:

$$\left\{\left[P^{LE}\right]\left[W^{E}\right] + \left[Q^{LE}\right]\left[W^{M}\right]\right\}\{J_e\} = \left\{b^{LE}\right\} \tag{2.246}$$

$$\left\{\left[P^{GE}\right]\left[W^{E}\right] + \left[Q^{GE}\right]\left[W^{M}\right]\right\}\{J_e\} = \left\{b^{GE}\right\} \tag{2.247}$$

For a neat expression, we denote the equations of (2.244), (2.245), (2.246), and (2.247) as MFIE-G, MFIE-L, EFIE-G and EFIE-L, respectively. These four equations are derived either from EFIE or MFIE. Solely using any one of them leads to the interior resonance problem. In order to eliminate the interior resonance, we combine the EFIE (2.233) and MFIE (2.234) and obtain the combined-field integral equation (CFIE). In this book, we study the following two combinations, denoted as CFIE-GG and CFIE-GL, respectively.

$$\text{CFIE-GG}: \quad \alpha \times \text{EFIE-G} + (1-\alpha) \times \text{MFIE-G} \tag{2.248}$$

$$\text{CFIE-GL}: \quad \alpha\sqrt{\varepsilon_1/\mu_1} \times \text{EFIE-G} + (1-\alpha)\bar{l} \times \text{MFIE-L} \tag{2.249}$$

where α is the weight, with value range being, $(0, 1)$, $\bar{l}$ is the average edge size of the triangular patch. The choice of α will affect the speed of convergence. In the following numerical experiments, α equals 0.2 in the CFIE-GG experiment, and α equals 0.94 in the CFIE-GL experiment.

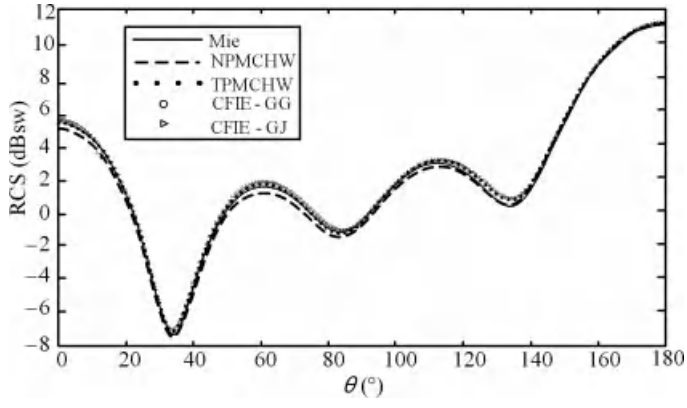


Figure 2.48 The bistatic RCS of a lossless dielectric sphere with radius $r = 0.5\,\text{m}$, and $\varepsilon_r = 4.0$.

Next, we compare the numerical behavior among the above single discretized equation forms and the PMCHW equations by numerical experiments. The frequency of the incident waves in the following is 300 MHz. The computation platform is Xeon 3.66G CPU, and 16G Memory, on an IBM X366 server.

Figure 2.48 shows the bistatic RCS of a lossless dielectric sphere with radius $r = 0.5\,\text{m}$, and $\varepsilon_r = 4.0$, by using MFIE-L,CFIE-GG,CFIE-GL,TPMCHW, NPMCHW schemes, and an analytical method. The total edge number of the triangles is 3072 in MoM simulations. It can be seen that numerical results agree well with the analytical solution. Figure 2.49 shows the bistatic RCS of a lossy dielectric sphere with

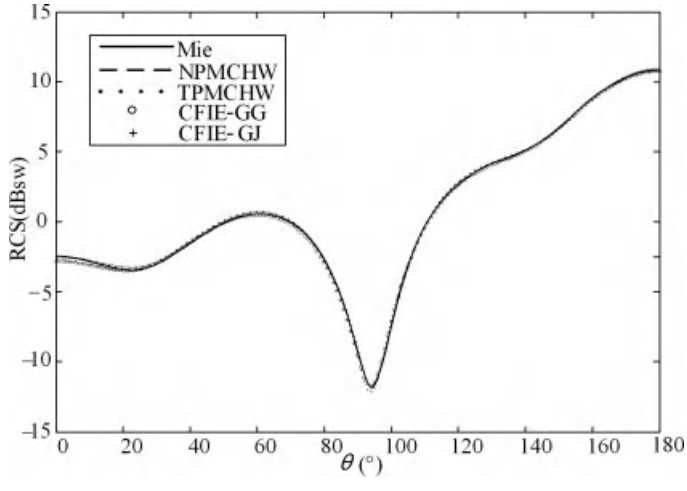


Figure 2.49 The bistatic RCS of a lossless dielectric sphere with radius $r = 0.5\,\text{m}$, and $\varepsilon_r = 4 - 100j$.

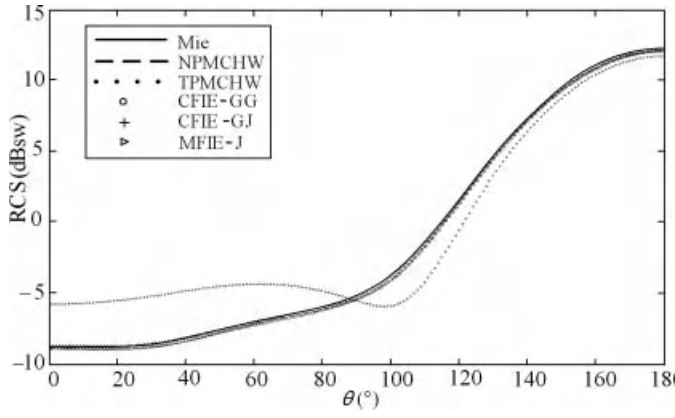


Figure 2.50 The bistatic RCS of a lossless dielectric sphere on resonant frequency ($\varepsilon_r = 2.25$, $k_1 r = 2.7466$).

radius $r = 0.5$ m, and $\varepsilon_r = 4 - 100j$. The methods and discretization are the same as the above modeling. Again, numerical results agree well with the analytical solution.

Next, we will study the robustness of these MoM algorithms on resonance. Figures 2.50 and 2.51 show the bistatic RCS of a lossless dielectric sphere on resonant frequencies. For a dielectric sphere with relative permittivity $\varepsilon_r = 2.25$, $k_1 r = 2.7466$ and $k_1 r = 4.4934$ correspond to the resonance in MFIE. From Figures 2.50 and 2.51 it can be seen that, only MFIE-L fails to convergent to the analytical solution on resonance, indicating the absence of the interior resonance in the remaining four MoM algorithms.

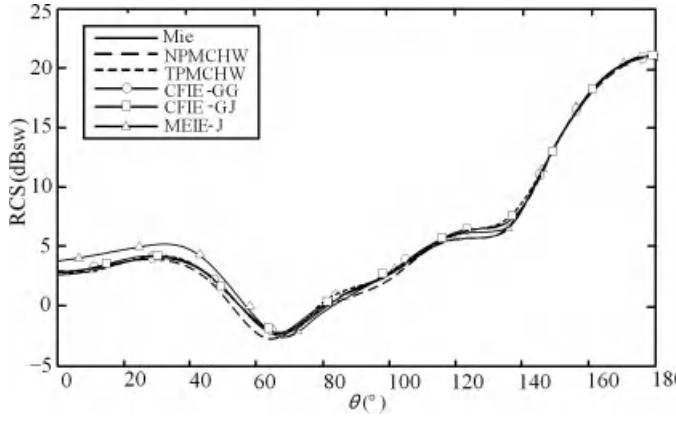


Figure 2.51 The bistatic RCS of a lossless dielectric sphere on resonant frequency ($\varepsilon_r = 2.25$, $k_1 r = 4.4934$).

Table 2.12 Time used in RCS calculation for sphere of radius r, dielectric constants ε_r and meshed by N edges: $r = 0.5\lambda_0, \varepsilon_r = 4.0, N = 3072$

Discretized equations	Iteration numbers	Iteration time (s)	Time for constructing matrices (s)
MFIE-L	20	21.57	480.48
CFIE-GG	20	20.92	1443.10
CFIE-GL	14	15.29	944.34
NPMCHW	58	30.70	655.21
TPMCHW	404	203.48	658.67

We can come to the conclusion that, as with PMCHW, the single integral equations CFIE-GG and CFIE-GL bears the merits of high accuracy and no interior resonance; whereas MFIE-L suffers interior resonance.

In the following, we will compare the computer run time in the aforementioned five MoM schemes. The total computation time in a moment method includes mainly the time to construct the matrices and the iteration time. In constructing the matrices, NPMCHW and TPMCHW use almost the same time, because they are only different in the testing functions selected. Consider an EM scattering problem, with total edge number in meshes N. From (2.211)–(2.212), and (2.219)–(2.220) we know that, TPMCHW and NPMCHW involve computation of four $N \times N$ full matrices. From (2.244)–(2.249), we know that, MFIE-L involves computation of four $N \times N$ full matrices, while CFIE-GG and CFIE-GL need to construct six $N \times N$ full matrices. Then, we come to the conclusion that the time to construct matrices in the MFIE-L, PMCHW, CFIE-GL and CFIE-GG are in ascending order. Tables 2.12–2.14 list the examples of RCS calculation for spheres of different radii and dielectric constants.

From Tables 2.12–2.14 we can see that the time to construct matrices in the MFIE-L, PMCHW, CFIE-GL and CFIE-GG are indeed in ascending order, which verifies the above analysis.

Table 2.13 Time used in RCS calculation for sphere of radius r, dielectric constants ε_r and meshed by N edges: $r = 0.125\lambda_0, \varepsilon_r = 64, N = 3072$

Discretized equations	Iteration numbers	Iteration time (s)	Time for constructing matrices (s)
MFIE-L	24	25.24	945.40
CFIE-GG	63	64.93	3211.51
CFIE-GL	23	24.41	2133.62
NPMCHW	627	324.63	1746.11
TPMCHW	840	365.35	1818.95

Table 2.14 Time used in RCS calculation for sphere of radius r, dielectric constants ε_r and meshed by N edges: $r = 0.5\lambda_0, \varepsilon_r = 4 - 100j, N = 3072$

Discretized equations	Iteration numbers	Iteration time (s)	Time for constructing matrices (s)
MFIE-L	9	10.22	522.79
CFIE-GG	18	18.98	1486.37
CFIE-GL	12	12.36	962.08
NPMCHW	114	59.53	699.64
TPMCHW	155	76.46	702.11

Next, we will compare the iteration time when using the CG method in the aforementioned five MoM schemes. It can be seen from Tables 2.12–2.14 that the convergence in MoM schemes using a single integral equation is faster than those based on PMCHW. Table 2.15 lists the iteration numbers for spheres with different dielectric constants. The spatial resolutions in the meshes are the same, with the average edge length $0.12\lambda_g$ (λ_g is the wavelength of the incident wave in the medium). It can be seen from Table 2.15 that the iteration number increases rapidly with the increase of the dielectric constants in the NPMCHW schemes, while those in the single-equation schemes remain almost constant. To be specific, to model media with increasing dielectric constants, the iteration numbers in MFIE-L, CFIE-GL, CFIE-GG, NPMCHW and TPMCHW are in ascending order. Moreover, the single-integral equation schemes converge much faster than the NPMCHW schemes, especially when the dielectric constant is large.

It is reasonable that the single-equation scheme MFIE-L converges faster than the PMCHW schemes. This is because the second term in the LHS of (2.234) is similar to an $\mathbf{L} \cdot \mathbf{L}$ operator. From [23] we know that, twice imposing of the integro-differential operator in EFIE is equivalent to doubling the integro-differential operation in MFIE: $(\hat{n} \times \mathbf{L})^2 = -\frac{1}{4}\mathbf{I} + \mathbf{K}^2$.

Table 2.15 Comparison of the convergence speed (by using the CG method)

Radii of the sphere (m)	Dielectric constant ε_r	Iteration number				
		MFIE-L	CFIE-GG	CFIE-GL	NPMCHW	TPMCHW
0.8165	1.5	18	17	11	26	648
0.5000	4.0	20	20	14	58	404
0.1768	32.0	21	42	20	273	469
0.1250	64.0	24	63	23	627	840

Therefore, the RHS of (2.234) shows the property brought by the $\mathbf{K}$ operator. As we acknowledge that the matrix equation derived from discretizing the $\mathbf{K}$ operator bears a good condition number, we naturally expect a faster convergence as a result.

There is another thing worth noting: the discretization of operators $\mathbf{L}_1$, $\mathbf{K}_1$, $\mathbf{L}_2$, $\mathbf{K}_2$ in CFIE-GL and CFIE-GG involves four matrices in conventional implementation. However, in order to assure good numerical behavior in both the EFIE and the MFIE matrices, in practice we use two different testing function to discretize $\mathbf{L}_1$, $\mathbf{K}_1$. As a consequence, $\mathbf{L}_1$ and $\mathbf{K}_1$ correspond to four matrices each. Together with the two full matrices brought by discretizing $\mathbf{L}_2$ and $\mathbf{K}_2$, there are six matrices. It is verified if only one testing function is used to discretize $\mathbf{L}_1$ and $\mathbf{K}_1$, both the accuracy and the speed of convergence degrade greatly in the combined integration equation system.

2.3 Scattering from Three-Dimensional Inhomogeneous Dielectric Objects

From a physical point of view, the scattering from an inhomogeneous dielectric body is more complex than that from a homogeneous dielectrics or a metallic body. Numerically they are quite different as well. This is because the integral equations in this kind of problems are volume integral equations (VIEs) instead of the surface integral equations. Naturally, we can no longer apply triangular discretization for a volume. Instead, we need volume elements such as tetrahedra. So a usual procedure is, first construct basis functions analogous to the RWG basis functions, then discretize the VIE using the Galerkin's method, and finally solve the equations by employing a fast technique such as the FMM. This procedure is a direct extension of the surface integral MoM introduced in Sections 2.1 and 2.2. The reader is referred to [27] for more details. The accuracy in this scheme is satisfactory, but the efficiency and convenience of applying is not as good as the scheme using cuboids as simplex. Because the system of linear equations obtained by using cuboids as discretizing elements can be solved by a technique that is faster than the FMM. This technique is the fast Fourier transform (FFT). Although the computational complexity in FFT ($O(N \lg N)$) is the same as that of the FMM, numerical experiments show that the coefficient of the computational complexity in the former is much smaller than the latter. Of course, on the other hand, the cuboids cannot model the boundary of the inhomogeneous dielectric scatterer as accurately as the tetrahedra. However, the modeling error in an inhomogeneous dielectric problem plays a less crucial role in the final RCS result than in the PEC or homogeneous dielectric cases. Therefore, to calculate the scattering from inhomogeneous dielectrics, when the requirement on accuracy is not very high, the use of cuboids is a good choice with combined consideration of accuracy and efficiency. In this section, we will introduce a volume MoM of this kind that was proposed by Zwamborn and van der Berg [28].

2.3.1 Mathematic Formulation of the Problem

The VIE depicting the scattering from an inhomogeneous dielectric body has already been developed in Section 1.3.5. If only the permittivity is inhomogeneous in the material, we get

$$\mathbf{E} + j\omega\mu\left(1 + \frac{1}{k_0^2}\nabla\nabla\cdot\right)\mathbf{A} = \mathbf{E}^i \tag{2.250}$$

where

$$\mathbf{A} = \int_V [\sigma + j\omega\varepsilon_0(\varepsilon_r - 1)]\mathbf{E}GdV \tag{2.251}$$

Obviously (2.250) is an equation of unknown electric field $\mathbf{E}$. Since the electric field is only tangentially but not normally continuous on the interface of different media, we can model the tangential electric field using the basis functions taking tangential electric field components as interpolating parameters. From Section 2.1.5, we gain the experience that a good way to calculate the matrix entries is to transpose one of the ∇ operators in (2.250) to the testing function. On the other hand, the choice of the tangential testing functions makes this transposition cumbersome. To address this problem, we change the unknown vector variable in (2.250) from electric field $\mathbf{E}$ to electric displacement $\mathbf{D}$. Introducing $\varepsilon = \varepsilon_0\varepsilon_r - j\sigma/\omega$, then (2.250) is rewritten as

$$\frac{\mathbf{D}}{\varepsilon} - (k_0^2 + \nabla\nabla\cdot)\mathbf{A} = \mathbf{E}^i \tag{2.252}$$

where $k_0 = \omega\sqrt{\varepsilon_0\mu_0}$, and

$$\mathbf{A} = \frac{1}{\varepsilon_0}\int_V G(\mathbf{r} - \mathbf{r}')\chi(\mathbf{r}')\mathbf{D}(\mathbf{r}')d\mathbf{r}' \tag{2.253}$$

where

$$\chi(\mathbf{r}) = \frac{\varepsilon(\mathbf{r}) - \varepsilon_0}{\varepsilon(\mathbf{r})} \tag{2.254}$$

$$G(\mathbf{r}|\mathbf{r}') = \frac{e^{-jk_0|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r} - \mathbf{r}'|} \tag{2.255}$$

Equation 2.252 is the VIE introduced in this section. It is also a widely used VIE for solving inhomogeneous problems. Similarly, we can get the following VIE based on

the MFIE:

$$\mathbf{H} + \frac{1}{j\omega}\int_V (\chi \nabla \times \mathbf{H}) \times \nabla G dV = \mathbf{H}^i \tag{2.256}$$

Details about this integral equation will not be elaborated in this book.

2.3.2 *Rooftop Basis Functions*

A kind of basis function depicting the vector variables on the cuboidal volume elements is introduced in this section. The interpolating parameter is chosen as the normal unit vector on the center of each face of the cuboid, as shown in Figure 2.52. The coordinates of the start point and the end point of the cuboid is (x_1, y_1, z_1) and (x_2, y_2, z_2), respectively. For a neat representation, we construct a new local coordinate (ξ, ς, η), which is related to the original coordinate (x, y, z) by

$$\xi = \frac{x - x_1}{x_2 - x_1} \tag{2.257}$$

$$\varsigma = \frac{y - y_1}{y_2 - y_1} \tag{2.258}$$

$$\eta = \frac{z - z_1}{z_2 - z_1} \tag{2.259}$$

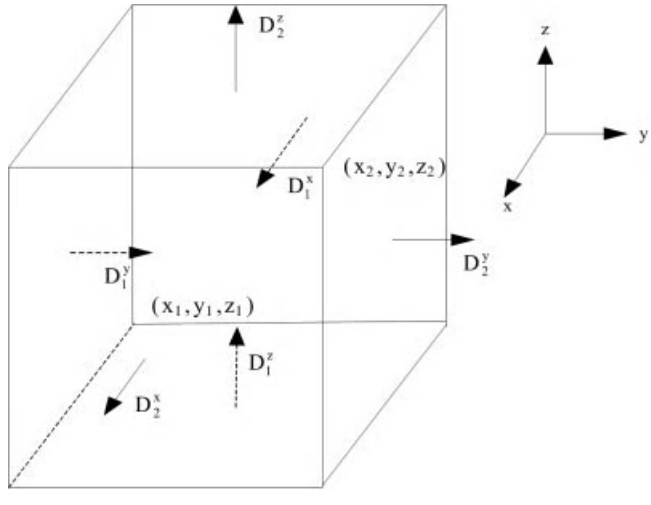


Figure 2.52 Illustration of the rooftop basis functions.

Thus, the components of an arbitrary vector $\vec{\mathbf{u}}$ on this cuboid are written as

$$u^x = \sum_{i=1}^{2} u_i^x L_i^x \tag{2.260}$$

$$u^y = \sum_{i=1}^{2} u_i^y L_i^y \tag{2.261}$$

$$u^z = \sum_{i=1}^{2} u_i^z L_i^z \tag{2.262}$$

where

$$L_1^x = 1 - \xi, \quad L_2^x = \xi \tag{2.263}$$

$$L_1^y = 1 - \varsigma, \quad L_2^y = \varsigma \tag{2.264}$$

$$L_1^z = 1 - \eta, \quad L_2^z = \eta \tag{2.265}$$

We will study the variation in the basis function corresponding to each of the interpolation parameter in (2.260)–(2.262). Because every interpolation parameter is shared by the adjacent cuboids, the corresponding basis function is defined on the adjacent cuboids. Its value increases in one cuboid linearly from 0 to 1, and decreases in the other linearly from 1 to 0, like the rooftop. So this kind of basis function is termed a *rooftop basis function (RBF)*.

2.3.3 Discretization of the VIE

The space is discretized uniformly using $\Delta x, \Delta y, \Delta z$ in the x, y, z directions and a cuboid discretization is obtained. Then, we denote the center of each cuboid as $\mathbf{r}_{i,j,k} = ((i-1/2)\Delta x, (j-1/2)\Delta y, (k-1/2)\Delta z)$, and consider the permittivity ε in every cuboid as a constant, denoted by $\varepsilon_{i,j,k} = \varepsilon\big(\mathbf{r}_{i,j,k}\big)$. In contrast to the discretization introduced in Sections 2.1 and 2.2, we express not only the unknown $\mathbf{D}$, but also the intermediate variable $\mathbf{A}$ with basis functions. That is, in the cuboid (i, j, k) we have:

$$D_{i,j,k}^x(\mathbf{r}) = \sum_{p=1}^{2} D_{p(i,j,k)}^x L_p^x(\mathbf{r}) \tag{2.266}$$

$$D^y_{i,j,k}(\mathbf{r}) = \sum_{p=1}^{2} D^y_{p(i,j,k)} L^y_p(\mathbf{r}) \tag{2.267}$$

$$D^z_{i,j,k}(\mathbf{r}) = \sum_{p=1}^{2} D^z_{p(i,j,k)} L^z_p(\mathbf{r}) \tag{2.268}$$

$$A^x_{i,j,k}(\mathbf{r}) = \sum_{p=1}^{2} A^x_{p(i,j,k)} L^x_p(\mathbf{r}) \tag{2.269}$$

$$A^y_{i,j,k}(\mathbf{r}) = \sum_{p=1}^{2} A^y_{p(i,j,k)} L^y_p(\mathbf{r}) \tag{2.270}$$

$$A^z_{i,j,k}(\mathbf{r}) = \sum_{p=1}^{2} A^z_{p(i,j,k)} L^z_p(\mathbf{r}) \tag{2.271}$$

It will be shown later that by doing this, the order of singularity is decreased, and the singularity processing in the matrix element computation is considerably simplified. Nevertheless, this process obviously introduces additional error, and the condition number of the matrix will also be higher. Adding the expression for the above **D** and **A**, and grouping the functions with the same interpolation parameters together, we obtain the following **D** and **A** expressions in the entire domain:

$$\mathbf{D}(\mathbf{r}) = \sum_{i,j,k} \sum_{q=1}^{2} D^{(q)}_{i,j,k} \phi^{(q)}_{i,j,k}(\mathbf{r}) \tag{2.272}$$

$$\mathbf{A}(\mathbf{r}) = \sum_{i,j,k} \sum_{q=1}^{2} A^{(q)}_{i,j,k} \phi^{(q)}_{i,j,k}(\mathbf{r}) \tag{2.273}$$

where $\phi^{(1)}_{i,j,k}(\mathbf{r})$, $\phi^{(2)}_{i,j,k}(\mathbf{r})$ and $\phi^{(3)}_{i,j,k}(\mathbf{r})$ are the RBFs defined on the adjacent cuboids in the x, y, z directions respectively, given by:

$$\phi^{(1)}_{i,j,k}(\mathbf{r}) = (L^x_{1(i,j,k)} + L^x_{2(i+1,j,k)})\hat{e}_x \tag{2.274}$$

$$\phi^{(2)}_{i,j,k}(\mathbf{r}) = (L^{y}_{1(i,j,k)} + L^{y}_{2(i,j+1,k)})\hat{e}_y \tag{2.275}$$

$$\phi^{(3)}_{i,j,k}(\mathbf{r}) = (L^{z}_{1(i,j,k)} + L^{z}_{2(i,j,k+1)})\hat{e}_z \tag{2.276}$$

Because the interpolation parameters defined on the boundary of the scatterer is only used by one cuboid, special care is needed in the programming when defining the corresponding basis functions. Using the Galerkin method, that is, using $\phi^{(q)}_{m,n,p}$ as the testing function and calculate its inner-product with (2.252)

$$\begin{aligned}\left\langle \phi^{(q)}_{m,n,p}, E^i(\mathbf{r}) \right\rangle = \left\langle \phi^{(q)}_{m,n,p}(\mathbf{r}), \frac{\mathbf{D}(\mathbf{r})}{\varepsilon(\mathbf{r})} \right\rangle - k_0^2 \left\langle \phi^{(q)}_{m,n,p}(\mathbf{r}), \mathbf{A}(\mathbf{r}) \right\rangle \\ - \left\langle \phi^{(q)}_{m,n,p}(\mathbf{r}), \nabla\nabla\cdot\mathbf{A}(\mathbf{r}) \right\rangle, \qquad q = 1,2,3\end{aligned} \tag{2.277}$$

where $\langle , \rangle$ denotes the inner-product of two vector functions. Applying the same techniques as introduced in Section 2.1.5, transpose one of the double ∇ operators imposed on $\mathbf{A}$ in the last term of (2.277) to the testing function, (2.277) can be further simplified as:

$$\begin{aligned}\left\langle \phi^{(q)}_{m,n,p}, E^i(\mathbf{r}) \right\rangle = \left\langle \phi^{(q)}_{m,n,p}(\mathbf{r}), \frac{\mathbf{D}(\mathbf{r})}{\varepsilon(\mathbf{r})} \right\rangle - k_0^2 \left\langle \phi^{(q)}_{m,n,p}(\mathbf{r}), \mathbf{A}(\mathbf{r}) \right\rangle \\ + \left\langle \nabla\cdot\phi^{(q)}_{m,n,p}(\mathbf{r}), \nabla\cdot\mathbf{A}(\mathbf{r}) \right\rangle, \qquad q = 1,2,3\end{aligned} \tag{2.278}$$

Substituting (2.266)–(2.271) into (2.278) yields

$$\left\{ e^{i,p}_{m,n,p} \right\} = \left[u^{(p,q)}_{m,n,p;i,j,k} \right] \left\{ D^{(q)}_{i,j,k} \right\} - \left[k_0^2 v^{(p,q)}_{m,n,p;i,j,k} - w^{(p,q)}_{m,n,p;i,j,k} \right] \left[A^{(q)}_{i,j,k} \right] \tag{2.279}$$

where

$$e^{i,q}_{m,n,p} = \left\langle \phi^{(p)}_{m,n,p}, E^i(\mathbf{r}) \right\rangle \tag{2.280}$$

$$u^{(p,q)}_{m,n,p;i,j,k} = \left\langle \phi^{(p)}_{m,n,p}, \frac{\varepsilon_0}{\varepsilon} \phi^{(q)}_{i,j,k} \right\rangle \tag{2.281}$$

$$v_{m,n,p;i,j,k}^{(p,q)} = \left\langle \phi_{m,n,p}^{(p)}, \phi_{i,j,k}^{(q)} \right\rangle \tag{2.282}$$

$$w_{m,n,p;i,j,k}^{(p,q)} = \left\langle \nabla \cdot \phi_{m,n,p}^{(p)}, \nabla \cdot \phi_{i,j,k}^{(q)} \right\rangle \tag{2.283}$$

In order to solve (2.279), we need to further identify the relationship of $D_{m,n,p}^{(q)}$ and $A_{m,n,p}^{(q)}$. This can be done by analyzing (2.253). We will discuss how to do this in a rapid way in the following two sections.

2.3.4 Singularity Processing

Noticeably, without involving singularities, the matrix entries expressed by (2.280)–(2.283) are easy to calculate. Singularity only exists in (2.253) where the relationship of $D_{m,n,p}^{(q)}$ and $A_{m,n,p}^{(q)}$ is depicted. We can use tricks analogous to Section 2.1.5 to process precisely, but the procedure is complex. Here, we introduce a more convenient but less accurate skill. In short, we use the averaged **A** in one sphere to replace **A**. This variable is defined as:

$$[\mathbf{A}(\mathbf{r})] = \frac{\int_{|\mathbf{r}''|<(1/2)\Delta r} \mathbf{A}(\mathbf{r}+\mathbf{r}'')d\mathbf{r}''}{\int_{|\mathbf{r}''|<(1/2)\Delta r} d\mathbf{r}''} \tag{2.284}$$

Substituting (2.284) into (2.253), and swapping the order of the integral gives rise to

$$[\mathbf{A}(\mathbf{r})] = \frac{1}{\varepsilon_0}\int_S [G](\mathbf{r}-\mathbf{r}')\chi(\mathbf{r}')\mathbf{D}(\mathbf{r}')d\mathbf{r}' \tag{2.285}$$

where

$$[G](\mathbf{r}) = \frac{\int_{|\mathbf{r}''|<(1/2)\Delta r} \mathbf{G}(\mathbf{r}+\mathbf{r}'')d\mathbf{r}''}{\int_{|\mathbf{r}''|<(1/2)\Delta r} d\mathbf{r}''} \tag{2.286}$$

By applying the addition theorem for Green's function, that is, (2.155) in Section 2.1.8.2, and the plane-wave expansion of the product of Bessel function and Legendre polynomial $j_l P_l$ that is, (2.156), the above integral (2.286) can be

written as:

$$[G](\mathbf{r}) = \begin{cases} \dfrac{\left(1+\dfrac{1}{2}jk_0\Delta r\right)\exp\left(-\dfrac{1}{2}jk_0\Delta r\right)-1}{\dfrac{1}{6}\pi k_0^2(\Delta r)^3} & \text{if} \quad |\mathbf{r}|=0 \\ \dfrac{\exp(-jk_0|\mathbf{r}|)\left[\dfrac{\sin\left(\frac{1}{2}k_0\Delta r\right)}{\frac{1}{2}k_0\Delta r}-\cos\left(\frac{1}{2}k_0\Delta r\right)\right]}{\dfrac{1}{3}\pi(k_0\Delta r)^2|\mathbf{r}|} & \text{if} \quad |\mathbf{r}|>\dfrac{1}{2}\Delta r \end{cases} \tag{2.287}$$

Thus, (2.285) can be discretized as the following without a singularity

$$A_{m,n,p}^{(q)} = \frac{\Delta V}{\varepsilon_0}\sum_{m',n',p'} G_{m-m',n-n',p-p'}\chi_{m',n',p'}^{(q)} D_{m',n',p'}^{(q)} \tag{2.288}$$

where $\Delta V = \Delta x\Delta y\Delta z$. Substituting (2.288) into (2.279) yields a system of linear equations with a single unknown $\{D\}$. For convenience, we represent this in the following form

$$[P]\{D\} = \{e^i\} \tag{2.289}$$

2.3.5 *Fast Solution of the Discretized VIE*

Since the computational domain of the VIE is three-dimensional, the number of unknowns in (2.289) increases rapidly with the increase of the electric size of the scatterer. More to the point, because the matrix $[P]$ is a full matrix, it will be very computationally expensive if we use a direct method (such as LU decomposition) to solve (2.289). Therefore, an iteration method (such as the CG method) is normally used in solving (2.289). The major overhead in an iteration method is the matrix-vector multiplication. We divide the multiplication of matrix $[P]$ and vector $\{D\}$ into two steps: first, $A_{m,n,p}^{(q)}$ is calculated from (2.288); then the product of $[P]$ and $\{D\}$ is calculated by using (2.279). Denote the number of the unknowns by N. Because matrices $[u], [v], [w]$ are all sparse matrix, and there are $O(N)$ nonzero elements, the computational load and the memory requirement in the second step are both $O(N)$. If the first step is performed directly, the computational load and the memory requirement are both $O(N^2)$. If we observe (2.288) carefully, we may find that the RHS

of (2.288) is the convolution of G and $\chi^q D^q$. Therefore, fast discretized Fourier transformation (DFT) can be used in evaluating (2.288):

$$A^{(q)}_{m,n,p} = \frac{\Delta V}{\varepsilon_0} DFT^{-1}\left\{ DFT\left\{ G_{m,n,p} \right\} \cdot DFT\left\{ \chi^{(q)}_{m,n,p} D^{(q)}_{m,n,p} \right\} \right\} \tag{2.290}$$

Through using DFT, we can calculate $A^{(q)}_{m,n,p}$ by using $O(N)$ memory, and the workload of $O(N \lg N)$, which is a significant saving on the computational resources.

Problem

2.9 Prove (2.287).

2.3.6 Numerical Examples

In the above sections we have discussed the discretization, singularity process, and fast computation of the electric VIE. We will verify the above techniques using numerical experiments. Here, we consider a cladded lossy dielectric sphere, with the center coordinate of (a, a, a), where a is the outer radius of the cladding. The inner radius a_1 and the outer radius a_2 satisfy $k_0 a_1 = 0.163$ and $k_0 a_2 = 0.314$. The relative permittivity and conductivity of the core and the cladding are $\varepsilon_{r1} = 72$, $\sigma_1 = 0.9\ \mathrm{S\ m^{-1}}$ and $\varepsilon_{r2} = 7.5$, $\sigma_2 = 0.05\ \mathrm{S\ m^{-1}}$, respectively. The incident wave is x_1 polarized plane wave with amplitude of 1, incident from negative x_3 direction, with a frequency of 100 MHz. We discretize the sphere with two meshes of different spatial resolutions: in each direction, the coarse mesh contains 15 grids (denoted as $15 \times 15 \times 15$), and the fine mesh contains 29 grids (denoted as $29 \times 29 \times 29$). Figure 2.53 presents the comparison of the field distribution calculated by simulations using these two kinds of meshes and by Mie series [29]. It can be seen that the numerical results converge to the analytical results.

We need to clarify here that for targets showing strong inhomogeneity, that is, those with large difference in permittivity and conductivity in different portions, or for electrically large lossless inhomogeneous dielectrics, the convergence speed in (2.289) is very slow. In some problems we cannot even achieve convergence at all. For these kinds of problems, the result obtained by using the method introduced in this section is often with large error.

2.4 Essential Points in MoM for Solving Other Problems

In the above sections we use the example of a three-dimensional scattering problem to illustrate the solution procedure, the essential points, the difficulties, the acceleration techniques and the numerical performance of MoM. From the case study, we understand the essential points in MoM, such as the importance of the basis functions

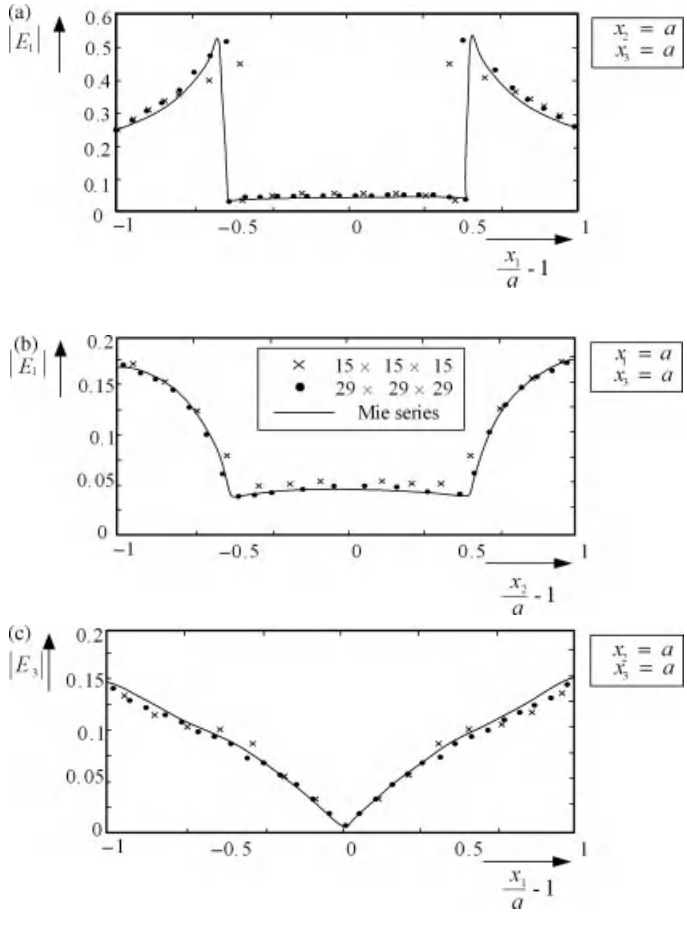


Figure 2.53 The comparison of the field distribution given by simulations using these two kind of meshes and by the Mie series. (Reproduced with permission from P. Zwamborn and P.M. van der Berg, "The three-dimensional weak form of the conjugate gradient FFT method for solving scattering problems," *IEEE Microwave Theory and Techniques*, **40**, 9, 1757–1766, 1992. © 1992 IEEE.)

choice, the difficulties for a singularity, the significance of the acceleration techniques, have general sense in MoM for any other applications. In other words, when we use MoM to solve other problems, we will follow the same procedure and pay attention to these essential points. To be more specific, we need to consider the choice of the basis functions, the numerical performance of the discretized matrix equation, the singularity handling, and the acceleration in solution of the matrix equation, and so on. Of course, the developments of MoM for different applications are different in the detailed and concrete level. We will choose different basis functions, different techniques of singularity handling and acceleration. In this section, we will briefly

discuss the MoM solutions of four other problems. We emphasize the unique treatment according to the unique characteristic of each problem, in order to give the readers some flavor of both the general principle in MoM and the applications of the method. These four kinds of problems are: the scattering from two-dimensional (2D) objects, that from periodic structures, that from two-and-half dimensional objects, and the radiation problems. As to the more complex scattering and radiation from layered media, the reader is referred to [30,31].

2.4.1 *Scattering from Two-Dimensional Objects*

When an object is infinitely long in one dimension and the constitutive parameters are evenly distributed in that dimension, we can simplify this 3D object as a 2D object. Strictly speaking, this kind of object does not exist in the real world. Despite this, when the size of the object in one dimension is much larger than that in any other dimension, and the constitutive parameters are quasievenly distributed over that dimension, then in practice, we can conveniently regard this object as a 2D object. As is indicated in Section 3.12 in [15], any EM wave can be expressed by the sum of a pair of transverse-magnetic (TM) wave and transverse-electric (TM) wave, with regard to one direction (such as the z direction). In the TM_z wave, the magnetic field is in the plane perpendicular (transverse) to z, while in the TE_z wave, the electric field is in the plane perpendicular to z. For a 2D object, due to the even distribution of the constitutive parameters in one dimension, the transverse magnetic incident wave causes only transverse magnetic scattering, but not transverse electric scattering, and vice versa. Thus, the scattering from a 2D object can be decomposed into two independent cases namely the TM and TE cases. We only consider the scattering from a 2D metallic target with an incident TM_z wave normal to the z direction to demonstrate the major difference of a 2D scattering problem with the 3D one. The reader is referred to the second chapter in [2] for the TE case, scattering from a 2D homogeneous and inhomogeneous dielectrics, and the eighth chapter in [2] for the oblique incident problem.

As we know, there are only three components in the incident TM_z wave, namely, E_z^i, H_x^i, H_y^i. Here, E_z^i can be expressed as

$$E_z^i = e^{-jk(x\cos\phi^i + y\sin\phi^i)} \tag{2.291}$$

where ϕ^i is the angle between the incident direction and the x-axis. According to the uniformity of both the incident field and the target, we can know from the Maxwell's equations that there are only three components in the scattered waves as well, namely E_z^S, H_x^S, H_y^S. Thus, the equivalent surface currents on the 2D metallic body contains only J_z component, which is homogeneous in the z direction. So $\nabla \cdot \mathbf{J}_z = 0$ holds, and consequently the EFIE (2.4) in Section 2.1.1 can be simplified as

$$E_z^i = jkZA_z \tag{2.292}$$

where

$$A_z = \int_S J_z \frac{e^{-jk|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|} dS \tag{2.293}$$

Due to the homogeneity of J_z in the z direction, we can construct a rectangular local coordinate system defined by two axes, that is, z and t. t is tangential to the metallic surface and normal to z. Rewrite (2.293) as

$$\begin{aligned} A_z &= \int_t J_z dt \int_{-\infty}^{\infty} \frac{e^{-jk\sqrt{|\rho-\rho'|^2+(z-z')^2}}}{4\pi\sqrt{|\boldsymbol{\rho}-\boldsymbol{\rho}'|^2+(z-z')^2}} dz \\ &= \int_t J_z \frac{1}{4j} H_0^2(k|\boldsymbol{\rho}-\boldsymbol{\rho}'|) dt \end{aligned} \tag{2.294}$$

where $\frac{1}{4j}H_0^{(2)}(k|\vec{\rho}-\vec{\rho}'|)$ is the Green's function in a homogeneous 2D plane. Segment the boundary of the body into the N fragments ΔC_n. Since there is no integral operator in (2.292), the equivalent electric current in every fragment can be approximated by a constant. In other words, the basis functions can be chosen as:

$$\phi(\boldsymbol{\rho}) = \begin{cases} 1 & \rho \text{ on } \Delta C_n \\ 0 & \text{otherwise} \end{cases} \tag{2.295}$$

Substituting (2.295) into (2.292), and by using point matching, that is, using the Dirac delta defined in the center of each fragment as a testing function, we obtain the following discretized integral equation:

$$[P]\{J\} = \{b\} \tag{2.296}$$

where

$$P_{mn} = \frac{kZ}{4} \int_{\Delta C_n} H_0^{(2)}\left(k\sqrt{(x-x_m)^2+(y-y_m)^2}\right) dt \tag{2.297}$$

$$b_m = e^{-jk(x_m \cos\phi^i + y_m \sin\phi^i)} \tag{2.298}$$

When $m \neq n$, there is no singularity in the integrand in (2.297). So we can either calculate the matrix element P_{mn} directly using a numerical integral method, or using the following approximation equation:

$$P_{mn} = \frac{kZ}{4} \Delta C_n H_0^{(2)}(k\sqrt{(x-x_m)^2+(y-y_m)^2}) dt \tag{2.299}$$

When $m = n$, a singularity exists in the Hankel function in (2.297). So we need to calculate it analytically. By applying the following approximated expansion of Hankel function under small argument:

$$H_0^{(2)}(z) \approx 1 - j\frac{2}{\pi}\ln\left(\frac{\gamma z}{2}\right) \tag{2.300}$$

where $\gamma = 1.781\ldots$ is Euler's constant, we obtain

$$P_{mn} = \frac{kZ}{4}\Delta C_n\left[1 - j\frac{2}{\pi}\ln\left(\frac{\gamma k\Delta C_n}{4e}\right)\right] \tag{2.301}$$

where $e = 2.718\ldots$. The solution of (2.296) can be accelerated by the 2D FMM, and the reader is referred to [2] for more details. However, the number of unknowns increases slowly with the electrical size of the object in the 2D problems. For example, the unknown in a metallic cylinder with a diameter of 20 wavelengths is less than 1000. Therefore, even by using the explicit LU decomposition method we can get the equivalent electric current solution quickly. As in the 3D scattering problem, (2.297) suffers interior resonant. Again, this problem can be avoided by using the CFIE combining (2.297) with the MFIE.

Once we obtain the equivalent electric current, we can calculate the *echo width* of a target, which is analog to the RCS in 3D problems:

$$\sigma(\phi, \phi^i) = \lim_{\rho\to\infty} 2\pi\rho\frac{\left|\mathbf{E}^S(\phi)\right|^2}{\left|\mathbf{E}^i(\phi^i)\right|^2} \tag{2.302}$$

By applying the following approximated expansion of Hankel function under large argument:

$$H_0^{(2)}(z) \approx \sqrt{\frac{2j}{\pi z}}e^{-jz} \tag{2.303}$$

we can obtain the scattering in the far-field,

$$E^s(\phi^s) = kZF\int_C J_z(x,y)e^{jk(x\cos\phi + y\sin\phi)}dl \tag{2.304}$$

where

$$F(\rho) = \frac{1}{\sqrt{8\pi k\rho}}e^{-j(k\rho+3\pi/4)} \tag{2.305}$$

Substituting (2.304) into (2.302) gives rise to the concise approximation of the echo width:

$$\sigma(\phi) = \frac{kZ^2}{4}\left|\int_C J_z(x,y)e^{jk(x\cos\phi + y\sin\phi)}dl\right|^2 \tag{2.306}$$

2.4.2 *Scattering from Periodic Structures*

Periodic structures are of great importance due to the vast applications in the real world. For example, the frequency-selective surface (FSS) is a typical periodic structure. A variety of periodic structures exists, as can be found in the literature, such as Chapter 7 in [2]. In this section, we only calculate the TM scattering from a one-dimensional periodic structure made by metal strips to illustrate the features when MoM is applied in periodic-structure modeling.

Consider the scattering from a periodic structure (as shown in Figure 2.54) illuminated by a TM wave expressed by (2.290). The equivalent current on the metal strip contains only a z component, denoted as $J_z(x)$. According to the Floquet theorem, $J_z(x)$ is expressed by the Floquet exponents:

$$J_z(x) = \sum_{-\infty}^{\infty} c_n \phi_n(x) \tag{2.307}$$

where c_n is the coefficient, and $\phi_n(x)$ is the Floquet exponent, written as

$$\phi_n(x) = e^{-jk_{xn}x} \tag{2.308}$$

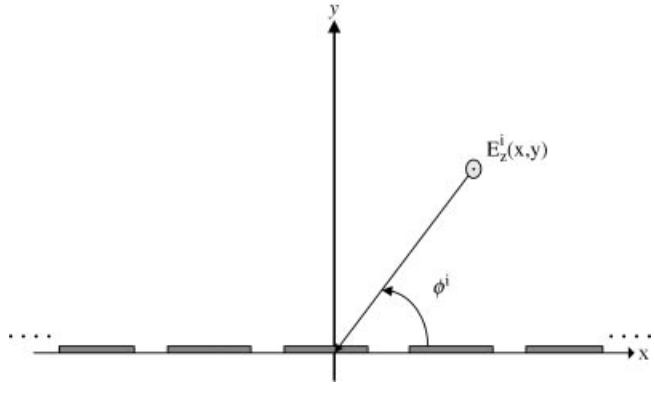


Figure 2.54 The one-dimensional periodic structure.

with

$$k_{xn} = k_x - \frac{2\pi}{a}n \tag{2.309}$$

where a is the dimension of the period, k_x is the wave number in the x direction.

It can be deduced from (2.308) that

$$J_z(x+a) = J_z(x)e^{-jk_x a} \tag{2.310}$$

As illustrated in the last section, we can construct the following EFIE on the strip:

$$jkZ\int_{-\infty}^{\infty} J_z(x')\frac{1}{4j}H_0^{(2)}(k|x-x'|dx' = E_z^i(x) \tag{2.311}$$

Substituting (2.310) into (2.311) yields

$$jkZ\int_{\text{Single Strip}} J_z(x')G_p(x-x')dx' = E_z^i(x) \tag{2.312}$$

where

$$G_p(x) = \frac{1}{4j}\sum_{l=-\infty}^{\infty} H_0^{(2)}(k|x-la|)e^{-jlk_x a} \tag{2.313}$$

Compared with (2.292), (2.312) is only different in the Green's function G_P, which is for the periodic structure and is the summation of infinite series. So we apply exactly the same segmentation manner, that is, to divide the single strip into N segments and model the equivalent current in every segment (ΔC_n) using a constant. Then, the adoption of point matching leads to the following discretized equation:

$$[P]\{J\} = \{b\} \tag{2.314}$$

where

$$P_{mn} = jkZ\int_{\Delta C_n} G_p(x_m - x')dx' \tag{2.315}$$

$$b_m = e^{-jk_x x_m} \tag{2.316}$$

The next key point is to calculate P_{mn}. According to the approximated expansion of Hankel function under large argument (2.304), the convergence rate in the series in (2.313) is low, as expressed by $O\left(1/\sqrt{l}\right)$. In order to accelerate the computation, we

can apply the following tricks. First, express (2.315) in a convolutional form:

$$P_{mn} = jkZB(x) * G_p(x)|_{x=x_m-x_n} \tag{2.317}$$

where

$$B(x) = \begin{cases} 1 & x \text{ on } \Delta C_n \\ 0 & \text{otherwise} \end{cases} \tag{2.318}$$

By applying the convolution theorem:

$$B(x) * G_p(x) = F^{-1}\left\{\bar{B}(f)\bar{G}_p(f)\right\} \tag{2.319}$$

where $\bar{B}(f)$ and $\bar{G}_P(f)$ are the spectrum functions of $B(x)$ and $G_P(x)$, respectively, we obtain

$$\bar{B}(f) = \frac{\sin \pi fb}{\pi f} \tag{2.320}$$

$$\bar{G}_p(f) = \frac{1}{2ja}\sum_{-\infty}^{\infty}\delta\left(f - \frac{l}{a} + \frac{k_x}{2\pi}\right)\frac{1}{\beta_y} \tag{2.321}$$

where b is the length of the segment ΔC_n, and

$$\beta_y = \begin{cases} \sqrt{k^2 - (2\pi f)^2} & k^2 \geq (2\pi f)^2 \\ -\sqrt{(2\pi f)^2 - k^2} & k^2 < (2\pi f)^2 \end{cases} \tag{2.322}$$

Thus, P_{mn} can be calculated by

$$P_{mn} = \frac{kZ}{2a}\sum_{-\infty}^{\infty}\left[\frac{\sin \pi\, fb}{\pi f}\frac{e^{j2\pi f(x_m-x_n)}}{\beta_y}\right]\Bigg|_{f=\frac{l}{a}-\frac{k_x}{2\pi}} \tag{2.323}$$

It is easy to determine the convergence rate in the series in (2.323) as $O(1/l^2)$, which is much higher than that in (2.313). The skills presented here are only one of the convenient acceleration skills. Even faster techniques are introduced in Section 7.5 in [2].

Problem

2.10 Write a program to compute P_{mn} using (2.315) and (2.323) respectively.

2.4.3 *Scattering from Two-and-Half-Dimensional Objects*

The so-called two-and-half-dimensional objects refer to the objects showing axial symmetry in both the geometry shape and constitutive parameters. Many flying targets such as the guided missile and the rocket can be regarded as two-and-half dimensional objects. The scattering from these kinds of targets can be calculated directly by using the 3D moment method introduced in Sections 2.1–2.3. However, we can make use of the structural features and obtain a simpler and more efficient scheme. Next, we use the scattering from a cylindrical metallic shell with finite length (as shown in Figure 2.55) to briefly demonstrate the simplifying procedure. The reader can work out the detailed processes or refer to [2].

Similar to the 3D scattering problem, we can construct the following EFIE on the surface of the cylinder:

$$\left[\mathbf{E}^i + Z\mathbf{L}(\mathbf{J})\right]\big|_t = 0 \tag{2.324}$$

where the operator L on the surface of the PEC shell can be expressed as

$$\mathbf{L}(\mathbf{J}) = -jk\int_{z_0}^{z_n}\int_0^{2\pi}\left[\mathbf{J} + \frac{1}{k^2}\nabla(\nabla'\cdot\mathbf{J})\right]\frac{e^{-jkR}}{4\pi R}ad\phi' dz' \tag{2.325}$$

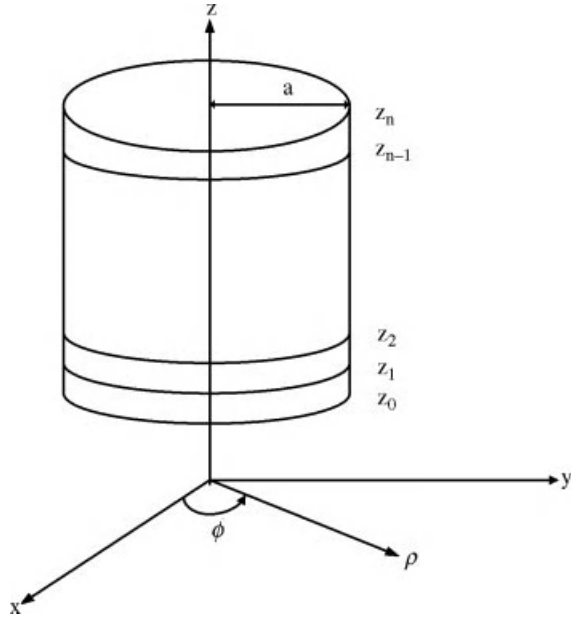


Figure 2.55 A cylindrical metallic shell with finite length.

where

$$R = \sqrt{(z-z')^2 + 4a^2 \sin^2\left[\frac{1}{2}(\phi - \phi')\right]} \tag{2.326}$$

According to the axial symmetry of the structure, the unknown equivalent current, the incident field and the corresponding Green's function can be expressed as:

$$J_\phi(\phi', z') = \sum_{m=-\infty}^{\infty} J_{\phi m}(z') e^{jm\phi'} \tag{2.327}$$

$$J_z(\phi', z') = \sum_{m=-\infty}^{\infty} J_{zm}(z') e^{jm\phi'} \tag{2.328}$$

$$E^i_\phi(\phi, z) = \sum_{m=-\infty}^{\infty} E^i_{\phi m}(z) e^{jm\phi} \tag{2.329}$$

$$E^i_z(\phi, z) = \sum_{m=-\infty}^{\infty} E^i_{zm}(z) e^{jm\phi} \tag{2.330}$$

and

$$\frac{e^{-jkR}}{4\pi R} = \sum_{m=-\infty}^{\infty} G_m(z - z') e^{jm\alpha} \tag{2.331}$$

where

$$E^i_{\phi m}(z) = \frac{1}{2\pi} \int_{-\pi}^{\pi} E^i_\phi(a, \alpha, z) e^{-jm\alpha} d\alpha \tag{2.332}$$

$$E^i_{zm}(z) = \frac{1}{2\pi} \int_{-\pi}^{\pi} E^i_z(a, \alpha, z) e^{-jm\alpha} d\alpha \tag{2.333}$$

and

$$\begin{aligned} G_m(z - z') &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{e^{-jkR}}{4\pi R} e^{-jm\alpha} d\alpha \\ &= \frac{1}{4\pi^2} \int_0^{\pi} \frac{e^{-jkR}}{R} e^{-jm\alpha} d\alpha \end{aligned} \tag{2.334}$$

where $\alpha = \phi - \phi'$. Substituting (2.327)–(2.331) into (2.325), and by equating the coefficients of the second harmonics in both sides of the equation, we get

$$j2\pi akZ\int_{z_0}^{z_n}\left[\left(1-\frac{m^2}{k^2a^2}\right)J_{\phi m}+j\frac{m}{k^2a}\frac{\partial J_{zm}}{\partial z'}\right]G_m(z-z')dz' = E^i_{\phi m} \tag{2.335}$$

$$j2\pi akZ\left[\int_{z_0}^{z_n}J_{zm}G_m(z-z')dz' + \frac{1}{k^2}\frac{\partial}{\partial z}\int_{z_0}^{z_n}\left(j\frac{m}{a}J_{\phi m}+\frac{\partial J_{zm}}{\partial z'}\right)G_m(z-z')dz'\right] = E^i_{zm} \tag{2.336}$$

These two scalar equations indicate that the mth harmonic incident wave in the ϕ direction causes only the mth harmonic scattering from the cylindrical metallic shell. Thus, the vector SIE (2.325) is transformed into two scalar line integral equations, which simplifies the problem significantly. The reader can try to discretize (2.335) and (2.336) or refer to Section 8.7 in [2].

2.4.4 Radiation Problems

There are two types of radiation, that is, passive radiation and the active one. We know that any object with a certain temperature can emit EM energy. The emission rate of diverse objects is different. The determination of the emission rate is a kind of passive radiation problem. According to the dynamic equilibrium on the temperature of an object, the emission rate of an object equals its absorption rate. Further, according to conservation of energy, the sum of the emission rate and the absorption rate is unity. Consequently, if we know the scattering from an object, we can calculate the scattering rate, in order to know the emission rate. Therefore, the key issue in the passive radiation problem is still a scattering problem, as details provided in [32]. In the active radiation problem, we mainly study the radiation of a certain structure under known excitation. Antenna problems are typical radiation problems. The difference between the active radiation problem and the scattering problem is they have different kinds of excitation: in the active radiation problem the source is close to the structure and there are various types of excitation, for example, the aperture pulse excitation in the wire antennas, the marginal excitation of coaxial cables, and the mode excitation in the waveguides; in the scattering problem the source is placed far away from the structure and the excitation can be deemed to be plane waves. From the numerical computational point of view, the methodology and numerical techniques are roughly the same. The major differences are: the ways in mimicking the excitation, and the convergence speed in the iterative method due to different excitations. Generally speaking, the convergence speed in the radiation problem is lower than that in the scattering problem. Moreover, the final physical quantity is different. For the active radiation problem, we are not only interested in the far-field physical quantities such as the radiation pattern and the gain,

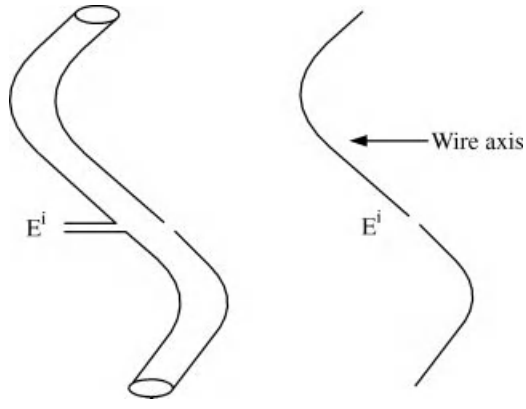


Figure 2.56 A PEC wire with limited length.

but we also care about the near-field quantities such as the input impedance. In this section we will illustrate the difference in terms of computation in the active radiation and the scattering problems using the example of a wire antenna.

Figure 2.56 shows a wire antenna made of metal wire, excited by an aperture pulse. As is known from Section 1.3.5, on the surface of the metallic line, we can construct the following EFIE:

$$[j\omega\mathbf{A} + \nabla\Phi]|_t = \mathbf{E}^i|_t \tag{2.337}$$

where

$$\mathbf{A} = \mu\int_S \mathbf{J}GdS \tag{2.338}$$

$$\Phi = \frac{1}{\varepsilon}\int_S \rho GdS \tag{2.339}$$

$$\rho = -\frac{1}{j\omega}\nabla\cdot\mathbf{J} \tag{2.340}$$

$$\mathbf{E}^i = \delta(l_i)\hat{l} \tag{2.341}$$

and $\hat{l}$ is the unit vector along the wire axis. There are two differences in (2.337) compared with (2.4) in Section 2.1.1. First, $\mathbf{E}^i$ is expressed as a vector pulse function. Secondly, instead of using $\mathbf{L}$, a vector $\mathbf{A}$ and a scalar Φ are used in this form. We choose these because we want to use the same techniques as in Section 2.2.3. That is, we can use basis functions to express the intermediate variables $\mathbf{A}$ and Φ, in order to decrease

the order of singularity. Before discretizing (2.337), we can further simplify (2.337)–(2.340) according to the structural characteristics of the wire antennas. Because the wire is very thin, the current is approximately regarded as a line current **I** flowing along the wire axis of the thread. Thus, (2.337)–(2.340) can be rewritten as

$$j\omega \mathbf{A}_l - \frac{\partial \Phi}{\partial l} = \mathbf{E}_l^i \tag{2.342}$$

where

$$\mathbf{A} = \mu \int_l \mathbf{I} G dl \tag{2.343}$$

$$\Phi = \frac{1}{\varepsilon} \int_l l \rho G dl \tag{2.344}$$

$$\rho = -\frac{1}{j\omega} \frac{\partial \mathbf{I}}{\partial l} \tag{2.345}$$

Discretize the wire axis of the wire into N segments. Denote the starting point of the nth fragment as n^-, the middle point as n, the ending point as n^+, and the length as Δl_n. Denote the fragment between the middle points of the $(n-1)$th and the nth fragments as fragment n^-, and that between the middle points of the $(n+1)$th and the nth fragments as fragment n^+, with their lengths being denoted as Δl_n^- and Δl_n^+, respectively. Using the constant basis function introduced in Section 2.4.1 to express **I**, **A**, $\Phi(n^+)$, and $\rho(n^+)$, specifically, suppose **I** and **A** are constants $\mathbf{I}(n)$ and $\mathbf{A}(n)$ in the nth segment, and Φ and ρ are constants $\Phi(n^+)$ and $\rho(n^+)$ in the n^+th segment. Thus, (2.342)–(2.346) can be discretized as

$$j\omega \mathbf{A}_l(m) + \frac{\Phi(m^-) - \Phi(m^+)}{\Delta l_m} = \mathbf{E}_l^i \tag{2.346}$$

where

$$\mathbf{A}(m) = \mu \sum_n \mathbf{I}(n) \int_{\Delta l_n} G dl \tag{2.347}$$

$$\Phi(m^+) = \frac{1}{\varepsilon} \sum_n \rho(n^+) \int_{\Delta l_{n^+}} G dl \tag{2.348}$$

$$\rho(n^+) = -\frac{1}{j\omega} \frac{I(n+1) - I(n)}{\Delta l_{n^+}} \tag{2.349}$$

where m and n denote the location of the field point and the source point, respectively. Choosing the same testing functions with the basis functions and calculate the inner-product with (2.346), after substituting (2.347)–(2.349) we obtain

$$[Z]\{I\} = \{V\} \tag{2.350}$$

where

$$\begin{aligned} Z_{mn} &= j\omega\mu\Delta l_n \cdot \Delta l_m \psi(n,m) \\ &\quad + \frac{1}{j\omega\varepsilon}[\psi(n^+,m^+) - \psi(n^-,m^+) - \psi(n^+,m^-) + \psi(n^-,m^-)] \end{aligned} \tag{2.351}$$

$$V = \left\{ \begin{array}{c} 0 \\ \vdots \\ V_i = 1 \\ \vdots \\ 0 \end{array} \right\} \tag{2.352}$$

$$\psi(m,n) = \frac{1}{\Delta l_n}\int_{\Delta l_n} G dl \tag{2.353}$$

The index i in V_i in (2.352) denotes the location of the feeding of the antenna. Solve (2.351) and we get the equivalent electric current $\{I\}$, from which we can calculate all the interested physical quantities. As far as antennas are concerned, people are interested in the input impedance, the radiation pattern, the directivity, the gain and the radiation efficiency. As we know, the gain is the product of the directivity and the radiation efficiency, while the latter is dependent on the specific feeding network and the input impedance of the antenna. With the feeding network being uncertain, the analysis in antennas is concentrated on the input impedance, the radiation pattern, and the directivity. The input impedance reflects the near-field behavior of the antennas. It is simply the rate of the voltage at the feeding point with the calculated equivalent current. The radiation pattern and the directivity reflect the far-field behavior of the antennas. From (2.183), the far-field radiated energy can be expressed as:

$$U = F(\theta,\phi) = \frac{1}{2Z}\left[|E_\theta(\theta,\phi)|^2 + |E_\phi(\theta,\phi)|^2\right] \tag{2.354}$$

where E_θ and E_ϕ are the two components of the antenna radiation in the far-field. Thus, the total radiation energy in the antennas can be written as:

$$P_r = \int_0^{2\pi}\int_0^{\pi} F(\theta,\phi)\sin\theta d\,\theta d\phi \tag{2.355}$$

Then, the radiation pattern of the antennas can be defined as

$$D(\theta,\phi)=4\pi\frac{F(\theta,\phi)}{P_r}=4\pi\frac{F(\theta,\phi)}{\int_0^{2\pi}\int_0^{\pi}F(\theta,\phi)\sin\theta d\theta d\phi} \tag{2.356}$$

The directivity of the antennas is defined as the maximum value of $D(\theta,\phi)$.

References

1. Harrington, R.F. (1993) *Field Computation by Moment Methods*, 2nd edn,IEEE Press, New York.
2. Peterson, A.F., Ray, S.L., and Mittra, R. (1998) *Computational Methods for Electromagnetics*, IEEE Press, New York.
3. Rao, S.M., Wilton, D.R., and Glisson, A.W. (1982) "Electromagnetic scattering by surfaces of arbitrary shape." *IEEE Transactions on Antennas and Propagation*, **30**(3), 409–418.
4. Graglia, R.D. (1988) "The use of parametric elements in the moment method solution of static and dynamic volume integral equations." *IEEE Transactions on Antennas and Propagation*, **36**(5), 636–646.
5. Khayat, M.A. and Wilton, D.R. (2008) "An improved transformation and optimized sampling scheme for the numerical evaluation of singular and near-singular potentials." *IEEE Antennas and Wireless Propagation Letters*, **7**, 377–380.
6. Duffy, M.G. (1982) "Quadrature over a pyramid of cube of integrands with a singularity at a vertex." *SIAM Journal on Numerical Analysis*, **19**(6), 1260–1262.
7. Donepudi, K.C., Jin, J.M., Velamparambil, S. *et al.* (2001) "A higher order parallelized multilevel fast multipole algorithm for 3D scattering." *IEEE Transactions on Antennas and Propagation*, **49**(7), 1069–1078.
8. Taylor, D.J. (2003) "Accurate and efficient numerical integration of weakly singular integrals in Galerkin EFIE solutions." *IEEE Transactions on Antennas and Propagation*, **51**(7), 1630–1637.
9. Hu, F.G., Wang, C.F., and Gan, Y.B. (2007) "Efficient calculation of interior scattering from large three-dimensional PEC cavities." *IEEE Transactions on Antennas and Propagation*, **55**(1), 167–177.
10. Deng, C.Q. (2011) Study of Advanced Techniques for Method of Moments, PhD Thesis, Beijing Institute of Technology, Beijing, China, (in Chinese).
11. Khayat, M.A. and Wilton, D.R. (2005) "Numerical evaluation of singular and near-singular potential integrals." *IEEE Transactions on Antennas and Propagation*, **53**(10), 3180–3190.
12. Rossi, L. and Cullen, P.J. (1999) "On the fully numerical evaluation of the linear-shape function times the 3D Green's function on a plane triangle." *IEEE Transactions on Microwave Theory and Techniques*, **47**(4), 398–402.
13. Rokhlin, V. (1990) "Rapid solution of integral equations of scattering theory in two dimensions." *Journal of Computational Physics*, **86**(2), 414–439.
14. Coifman, R., Rokhlin, V., and Wandzura, S. (1993) "The fast multipole method for the wave equation: a pedestrian prescription." *IEEE Antennas and Propagation Magazine*, **35**(3), 7–12.
15. Harrington, R.F. (1961) *Time-Harmonic Electromagnetic Fields*, McGraw-Hill, New York.
16. Song, J.M. and Chew, W.C. (1995) "Multilevel fast-multipole algorithm for solving combined field integral equations of electromagnetic scattering." *Microwave and Optical Technology Letters*, **10**(1), 14–19.
17. Press, W.H., Teukolsky, S.A., Vetterling, W.T., and Flannery, B.P. (1992) *Numerical Recipes in Fortran*, Cambridge University Press, New York.

18. Brandt, A. (1991) "Multilevel computations of integral transforms and partial interactions with oscillatory kernels." *Computer Physics Communications*, **65**, 24–38.
19. Velamparambil, S., Chew, W.C., and Song, J. (2003) "10 million unknowns: Is it that big?" *IEEE Antennas and Propagation Magazine*, **45**, 43–58.
20. Pan, X.M. and Sheng, X.Q. (2006) "A highly efficient parallel approach of multilevel fast multipole algorithm." *Journal of Electromagnetic Waves and Applications*, **20**, 1081–1092.
21. Pan, X.M. and Sheng, X.Q. (2008) "A sophisticated parallel MLFMA for scattering by extremely large targets." *IEEE Antennas and Propagation Magazine*, **50**(3), 129–138.
22. Pan, X.M., Pi, W.C., and Sheng, X.Q. (2011) "On OpenMP parallelization of the multilevel fast multipole algorithm." *Progress in Electromagnetics Research*, **112**, 199–213.
23. Qinn, M.J. (2004) *Parallel programming in C with MPI and OpenMP*, McGraw Hill, New York.
24. Yueng, M.S. (1999) "Single integral equation for electromagnetic scattering by three-dimensional homogeneous dielectric objects." *IEEE Transactions on Antennas and Propagation*, **47**(10), 1615–1622.
25. Sheng, X.Q., Jin, J.M., Song, J.M. *et al.* (1998) "Solution of combined-field integral equation using multilevel fast multipole algorithm for scattering by homogeneous bodies." *IEEE Transactions on Antennas and Propagation*, **46**(11), 1718–1726.
26. Hsiao, G.C. and Kleinman, R.E. (1997) "Mathematical foundations for error estimation in numerical solutions of integral equations in electromagnetic." *IEEE Transactions on Antennas and Propagation*, **45**(9), 316–328.
27. Lu, C.C. (2001) *Volume-surface integral equation in Fast and Efficient Algorithms in Computational Electromagnetics*, Artech House, Norwood MA.
28. Zwamborn, P. and van der Berg, P.M. (1992) "The three-dimensional weak form of the conjugate gradient FFT method for solving scattering problems." *IEEE Transactions on Microwave Theory and Techniques*, **40**(9), 1757–1766.
29. Mie, G. (1908) "Beiträge zur Optik trüber Medien, speziell kolloidaler Metallösungen, Leipzig." *Annals of Physics*, **330**, 377–445.
30. Michalski, K.A. and Mosig, J.R. (1997) "Multilayered media Green's functions in integral equation formulations." *IEEE Transactions on Antennas and Propagation*, **45**(3), 508–519.
31. He, J., Yu, T., Geny, N., and Carin, L. (2000) "Method-of-moments analysis of electromagnetic scattering from a general three-dimensional dielectric target embedded in a multi-layered medium." *Radio Science*, **35**(2), 305–313.
32. Tsang, L., Kong, J.A., and Ding, K.H. (2000) *Scattering of Electromagnetic Waves: Theories and Applications*, John Wiley Sons, New York.

3

Finite-Element Method

The *finite-element method* (FEM) is a full-wave numerical method that discretizes the variational formulation of a functional. The evolution of this method within the scope of electromagnetics traces back to the solving of two types of problems, namely, the eigenmode problems, and the boundary-value problems. The eigenmode problems usually have types of waveguide and cavity. Both waveguides and cavities can be further classified into three types: hollow, dielectric-loaded, or dielectric. The boundary-value problem can be categorized into closed-domain EM problems and open-domain scattering/radiating problem. Following the style in Chapter 2 of using examples to illustrate methods, we select three examples to illustrate FEM. The first is the dielectric-loaded waveguide, which is a classic example of the eigenmode problems. The second is the three-dimensional (3D) discontinuous waveguide problem, which is a representative problem of the closed-domain EM problems. The third concerns the 3D scattering problems, which are typical open-domain scattering/radiating problem. Their solutions unfold the essentials of FEM. In this chapter, we will provide complete and detailed illustration about the theoretical basis, the critical solving techniques and the typical tricks involved in FEM by solving the above three specific problems. At the end of this chapter, we will also briefly review the FEM solution for some other problems, and the reader is referred to references [1,2] for more details.

3.1 Eigenmodes Problems of Dielectric-Loaded Waveguides

A waveguide is a structure that guides the EM waves. Here, we are interested in the *regular waveguides* that are infinite along their axes, while having invariant cross sections in shape and material. From theoretical analysis, it can be deduced that a series of modes can exist independently in a regular waveguide [3]. The fields of these modes differ not only in the patterns on the waveguide cross section but also in the propagation constants along the axis. To determine these parameters is a kind of eigenmode

Essentials of Computational Electromagnetics, First Edition. Xin-Qing Sheng and Wei Song.
© 2012 John Wiley & Sons Singapore Pte. Ltd. Published 2012 by John Wiley & Sons Singapore Pte. Ltd.

problem. In FEM, this problem can be mathematically converted into a standard matrix eigenvalue problem or a generalized eigenvalue problem. Due to this fact, FEM is a better candidate in solving this kind of problem compared with the MoM, which is discussed in the previous chapter, or the finite-difference time-domain (FDTD) method to be introduced in the following chapter. Since the 1960s, when FEM was first applied to EM problems, FEM has made great progress. The development of FEM in solving the eigenvalue problem essentially comprises three aspects: (1) the study of various mathematical formulations of the functional; (2) the construction of proper basis functions; (3) the development of fast techniques for solving the generalized eigenvalue equations. These progresses are fully unfolded during the solving the eigenmodes of a dielectric-loaded waveguide. In the following, we will present the detailed procedure for solving this particular problem.

3.1.1 Functional Formulation

The *variational formulation of a functional* is generally derived from the partial differential equation (PDE) that describes the problem. Since various PDEs exist, there are numerous expressions for the functionals. These forms are different in the scope of application and numerical efficiency. For example, if we are interested in solving the eigenmodes of a dielectric-loaded waveguide, a functional can be constructed by using the longitudinal electric and magnetic fields [4], by using transverse electric and magnetic fields [5], or by using vector potentials and scalar potentials [6], and so on. For a combined consideration of the scope of application and numerical efficiency, the functional of the full electric/magnetic field derived from the vector wave equations introduced in Chapter 1 is a good choice [7]. This functional formulation is suitable for arbitrary material constitution, and the unknowns are solely the electric fields or the magnetic fields. Next, we will show a general procedure to derive the functional.

From Section 1.2 we know that the partial differential equation describing the fields in dielectric-loaded waveguide can be written as

$$\nabla \times \frac{1}{\mu_r} \nabla \times \mathbf{E} - k_0^2 \varepsilon_r \mathbf{E} = 0, \qquad \text{in } S \tag{3.1}$$

$$\hat{n} \times \mathbf{E} = 0 \qquad \text{on } \Gamma_1 \tag{3.2}$$

$$\hat{n} \times \nabla \times \mathbf{E} = 0 \qquad \text{on } \Gamma_2 \tag{3.3}$$

where S is the waveguide cross section, whose boundary might comprise a perfectly electric conducting boundary (PEC) Γ_1 and a perfectly magnetic conducting boundary (PMC) Γ_2. Actually, the physical boundary of this waveguide is wholly PEC. However, for some modes, according to the symmetry of the field distribution, we can regard the symmetry plane as PMC. So if only these modes are concerned, the computational

domain can be reduced by employing PMC. Applying the dot product on both sides of (3.1) with an infinitesimal and arbitrary variation of $\mathbf{E}$, and then integrating over the cross section S yields

$$\int_S \left[\nabla \times \left(\frac{1}{\mu_r}\nabla \times \mathbf{E}\right) - k_0^2 \varepsilon_r \mathbf{E}\right] \cdot \delta\mathbf{E} dS = 0 \tag{3.4}$$

Applying the dot product on both sides of (3.3) with an infinitesimal and arbitrary variation of $\mathbf{E}$, and integrating on Γ_2, we obtain

$$-\frac{1}{\mu_r}\int_{\Gamma_2} \hat{n} \times (\nabla \times \mathbf{E}) \cdot \delta\mathbf{E} d\Gamma = 0 \tag{3.5}$$

where the coefficient $-1/\mu_r$ is added to simplify the following expression. Adding (3.4) and (3.5) results in

$$\int_S \left[\nabla \times \left(\frac{1}{\mu_r}\nabla \times \mathbf{E}\right) - \varepsilon_r k_0^2 \mathbf{E}\right] \cdot \delta\mathbf{E} dS - \frac{1}{\mu_r}\int_{\Gamma_2} \hat{n} \times (\nabla \times \mathbf{E}) \cdot \delta\mathbf{E} d\tau = 0 \tag{3.6}$$

If $\delta\mathbf{E}$ in the first term and in the second term are independent and arbitrary, (3.6) is equivalent with the system of equations (3.1) and (3.3). In other words, the solution to (3.6) can simultaneously satisfy (3.1) and (3.3). Note that (3.2) is not included in (3.6). This is because once $\mathbf{E}$ is fixed to some extent by the first-type (Dirichelet) of boundary condition, $\delta\mathbf{E}$ should be zero and can no longer be arbitrary as defined. By assuming $\delta\mathbf{E}$ in the first and the second terms as identical and applying the following Green's theorem,

$$\int_S [u(\nabla \times \mathbf{a}) \cdot (\nabla \times \mathbf{b}) - \mathbf{a} \cdot (\nabla \times u\nabla \times \mathbf{b})] dS = \int_\Gamma u(\mathbf{a} \times \nabla \times \mathbf{b}) \cdot \hat{n} d\Gamma \tag{3.7}$$

(3.6) can be simplified as

$$\int_S \left[(\nabla \times \delta\mathbf{E}) \cdot \left(\frac{1}{\mu_r}\nabla \times \mathbf{E}\right) - \varepsilon_r k_0^2 \delta\mathbf{E} \cdot \mathbf{E}\right] dS = 0 \tag{3.8}$$

Note the assumption of $\delta\mathbf{E}$ in the first and the second terms in (3.6) as identical may lead to non-uniqueness of the solution and cause spurious solutions. The above equation is equivalent to the variation of the following functional

$$F(\mathbf{E}) = \frac{1}{2}\int_S \left[(\nabla \times \mathbf{E}) \cdot \left(\frac{1}{\mu_r}\nabla \times \mathbf{E}\right) - \varepsilon_r k_0^2 \mathbf{E} \cdot \mathbf{E}\right] dS \tag{3.9}$$

In this way, the original waveguide eigenvalue problem is equivalent to the following variational problem on imposing the electric boundary condition as:

$$\begin{cases} \delta F(\mathbf{E}) = 0 \\ \hat{n} \times \mathbf{E} = 0 \qquad \text{on } \Gamma_1 \end{cases} \tag{3.10}$$

According to the structural properties of the waveguide, the electric field of an arbitrary mode in a regular waveguide can be written as:

$$\mathbf{E}(x, y, z) = [\mathbf{E}_t(x, y) + \hat{\mathrm{z}} E_z(x, y)] e^{-j\beta z} \tag{3.11}$$

which leads to

$$\frac{\partial \mathbf{E}}{\partial z} = -j\beta \mathbf{E} \tag{3.12}$$

Substituting $\nabla = \nabla_t - j\beta\hat{z}$ into (3.9) yields

$$F(\mathbf{E}) = \int_S \left\{ \frac{1}{\mu_r} [(\nabla_t \times \mathbf{E}_t) \cdot (\nabla_t \times \mathbf{E}_t) + (\nabla_t E_z + j\beta \mathbf{E}_t) \cdot (\nabla_t E_z + j\beta \mathbf{E}_t)] - \varepsilon_r k_0^2 \mathbf{E} \cdot \mathbf{E} \right\} dS \tag{3.13}$$

By directly discretizing the functional (3.13), we obtain the generalized eigenvalue expression of k_0 determined by the propagation constant β. However, in the practical application, we usually need to determine β by using k_0. To this end, we apply the following transformation:

$$\begin{aligned} e_x &= \beta E_x \\ e_y &= \beta E_y \\ e_z &= -jE_z \end{aligned} \tag{3.14}$$

Thus, (3.13) becomes

$$\begin{aligned} F(\boldsymbol{e}) = \int_S \Bigg\{ \beta^2 \left[\frac{1}{\mu_r} (\boldsymbol{e}_t + \nabla_t e_z) \cdot (\boldsymbol{e}_t + \nabla_t e_z) - k_0^2 \varepsilon_r e_z^2 \right] \\ + \frac{1}{\mu_r} (\nabla_t \times \boldsymbol{e}_t \cdot \nabla_t \times \boldsymbol{e}_t) - k_0^2 \varepsilon_r \boldsymbol{e}_t \cdot \boldsymbol{e}_t \Bigg\} dS \end{aligned} \tag{3.15}$$

Apparently, the functional (3.15) involves only terms of k_0^2 and β^2, but no k_0 and β. Therefore, through discretization of the functional (3.15), the generalized eigenvalue expression of k_0 can be determined by the propagation constant β, or vice versa.

Problems

3.1 If (3.10) holds, it seems that we can only obtain

$$\int_S \left[\nabla \times \frac{1}{\mu_r}\nabla \times \mathbf{E} - \varepsilon_r k_0^2 \mathbf{E}\right] dS - \frac{1}{\mu_r}\int_{\Gamma_2} \hat{n} \times (\nabla \times \mathbf{E}) d\tau = 0$$

Please give the further explanation that (3.10) is equivalent to (3.1) and (3.3).

3.2 Consider the problem defined by the second-order differential equation

$$-\frac{\partial}{\partial x}\left(\alpha_x \frac{\partial \phi}{\partial x}\right) - \frac{\partial}{\partial y}\left(\alpha_y \frac{\partial \phi}{\partial y}\right) - \frac{\partial}{\partial z}\left(\alpha_z \frac{\partial \phi}{\partial z}\right) + \beta\phi = f$$

and the homogeneous boundary conditions

$$\phi = 0 \qquad \text{on } S_1$$
$$\left(\alpha_x \frac{\partial \phi}{\partial x} + \alpha_y \frac{\partial \phi}{\partial y} + \alpha_z \frac{\partial \phi}{\partial z}\right) \cdot \hat{n} + \gamma\phi = 0 \quad \text{on } S_2$$

Show that its associated operator is self-adjoint provided that α_x, α_y, α_z, β, and γ are real numbers or functions. Further show that the functional for this problem is given by

$$F(\phi) = \frac{1}{2}\int_V \left[\alpha_x \left|\frac{\partial \phi}{\partial x}\right|^2 + \alpha_y \left|\frac{\partial \phi}{\partial y}\right|^2 + \alpha_z \left|\frac{\partial \phi}{\partial z}\right|^2 + \beta|\phi|^2\right] dV$$
$$+ \frac{1}{2}\int_{S_2} \gamma|\phi|^2 dS - \frac{1}{2}\int_V \left[\phi f^* + f\phi^*\right] dV$$

3.3 Consider the vector wave equation

$$\nabla \times \left(\frac{1}{\varepsilon_r}\nabla \times \mathbf{H}\right) - k_0^2 \mu_r \mathbf{H} = \nabla \times \left(\frac{1}{\varepsilon_r}\mathbf{J}\right)$$

and the homogeneous boundary conditions,

$$\hat{n} \times \mathbf{H} = 0 \qquad \text{on } S_1$$
$$\frac{1}{\varepsilon_r}\hat{n} \times (\nabla \times \mathbf{H}) + \gamma_h \hat{n} \times (\hat{n} \times \mathbf{H}) = 0 \qquad \text{on } S_2$$

Show that the associated operator is self-adjoint provided that ε_r, μ_r, and γ_h are real-valued numbers or functions. Furthermore, show that the functional is given by:

$$F(\mathbf{H}) = \frac{1}{2}\int_V \left[\frac{1}{\varepsilon_r}(\nabla\times\mathbf{H})\cdot(\nabla\times\mathbf{H})^* - k_0^2\mu_r\mathbf{H}\cdot\mathbf{H}^*\right]dV + \frac{1}{2}\int_{S_2}\left[\gamma_h(\hat{n}\times\mathbf{H})\cdot(\hat{n}\times\mathbf{H})^*\right]dS - \frac{1}{2}\int_V\left[\mathbf{H}^*\cdot\left(\nabla\times\frac{1}{\varepsilon_r}\mathbf{J}\right) + \mathbf{H}\cdot\left(\nabla\times\frac{1}{\varepsilon_r}\mathbf{J}\right)^*\right]dV$$

3.4 Reconsider the problem defined in 3.2, but with the inhomogeneous boundary conditions

$$\phi = p \quad \text{on } S_1$$
$$\left(\alpha_x\frac{\partial\phi}{\partial x} + \alpha_y\frac{\partial\phi}{\partial y} + \alpha_z\frac{\partial\phi}{\partial z}\right)\cdot\hat{n} + \gamma\phi = q \quad \text{on } S_2$$

Apply the modified variational principle to show that the functional for this problem is given by

$$F(\phi) = \frac{1}{2}\int_V\left[\alpha_x\left|\frac{\partial\phi}{\partial x}\right|^2 + \alpha_y\left|\frac{\partial\phi}{\partial y}\right|^2 + \alpha_z\left|\frac{\partial\phi}{\partial z}\right|^2 + \beta|\phi|^2\right]dV + \frac{1}{2}\int_{S_2}\left(\gamma|\phi|^2 - q\phi^* - q^*\phi\right)dS - \frac{1}{2}\int_V\left(f^*\phi + f\phi^*\right)dV$$

3.5 Reconsider the problem defined in 3.3, but with inhomogeneous boundary conditions,

$$\hat{n}\times\mathbf{H} = \mathbf{Q} \quad \text{on } S_1$$
$$\frac{1}{\varepsilon_r}\hat{n}\times(\nabla\times\mathbf{H}) + \gamma_h\hat{n}\times(\hat{n}\times\mathbf{H}) = \mathbf{V} \quad \text{on } S_2$$

Show that the functional for this problem is

$$F(\mathbf{H}) = \frac{1}{2}\int_V\left[\frac{1}{\varepsilon_r}(\nabla\times\mathbf{H})\cdot(\nabla\times\mathbf{H})^* - k_0^2\mu_r\mathbf{H}\cdot\mathbf{H}^*\right]dV + \frac{1}{2}\int_{S_2}\left[\gamma_h(\hat{n}\times\mathbf{H})\cdot(\hat{n}\times\mathbf{H})^* + \mathbf{H}^*\cdot\mathbf{V} + \mathbf{H}\cdot\mathbf{V}^*\right]dS - \frac{1}{2}\int_V\left[\mathbf{H}^*\cdot\left(\nabla\times\frac{1}{\varepsilon_r}\mathbf{J}\right) + \mathbf{H}\cdot\left(\nabla\times\frac{1}{\varepsilon_r}\mathbf{J}\right)^*\right]dV$$

3.1.2 Choice of Basis Functions

As in MoM, the first step in discretizing the functional (3.15) is to choose proper basis functions, and express the unknown variable using superposition of the basis functions. Two factors play a pivotal role in the choice of basis functions: the discretization element, and the interpolation function. For the particular problem discussed in this section, the computational domain is a 2D domain, so obviously triangular patch is the most flexible simplex for constructing the basis functions. As for the proper choice of the interpolation function, it is not as obvious. In fact, people have spent years researching on this point. For a long time, the vector electric fields on the three nodes are taken as the interpolation parameters. But this choice leads to spurious solutions, which means unphysical solution may exist among the FEM results. In order to suppress these spurious solutions, people have been trying different treatments such as the penalty function method. However, effective treatment was not found until people realized that the improper choice of the interpolation parameters is the source of the problem in the late 1980s. Because by using the vector electric fields on the three nodes as the interpolation parameters, in the adjacent elements, not only the continuity of the tangential electric field, but also that of the normal electric field are enforced, which is not physically true. The following adjustment is applied in order to address this problem. As shown in Figure 3.1, E_z on the three nodes are still used as interpolation functions for expressing the longitudinal E_z. However, the tangential component E_{ti} on the midpoints of the triangle edges are used as the interpolation parameters for expressing the transverse E-field $\mathbf{E}_t$. The basis functions constructed in this way are termed *edge-element basis functions*. Before we present the mathematic expression for edge-element basis functions, let us define the area coordinates on the triangular patch. As shown in Figure 3.1, the three nodes are denoted anticlockwise as 1, 2 and 3. The edge is labeled after the opposite node. For example, edge "1" is the edge opposite node "1" and connecting nodes "2" and "3." The unit vector $\hat{e}_i (i = 1, 2, 3)$ on edge i

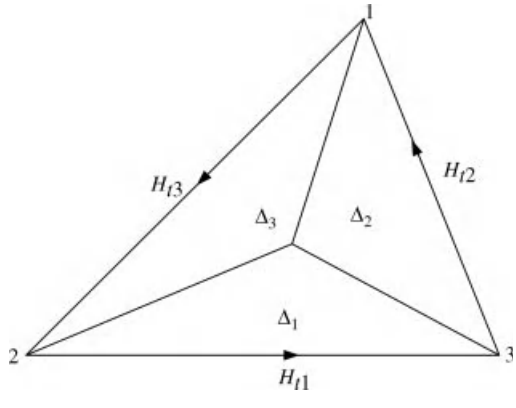


Figure 3.1 Illustration of the edge-elements in a triangle.

is labeled after the edge. For instance, $\hat{e}_1$ is the unit vector pointing form node "2" to "3." The area of the triangle is denoted as Δ. $P(x, y)$ is a point inside the triangle. Connecting $P(x, y)$ with the three nodes partitions the original triangle into three subtriangles, whose areas are denoted as Δ_1, Δ_2, and Δ_3, respectively. The areas can be explicitly expressed as

$$\Delta = \frac{1}{2}\begin{vmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{vmatrix} \tag{3.16}$$

and

$$\Delta_1 = \frac{1}{2}\begin{vmatrix} 1 & x & y \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{vmatrix} \tag{3.17}$$

$$\Delta_2 = \frac{1}{2}\begin{vmatrix} 1 & x_1 & y_1 \\ 1 & x & y \\ 1 & x_3 & y_3 \end{vmatrix} \tag{3.18}$$

Δ_3 can be expressed by using an analog. The area coordinates (L_1, L_2, L_3) of P are defined as

$$L_i = \frac{\Delta_i}{\Delta} \qquad \text{i} = 1, 2, 3 \tag{3.19}$$

In the literature they are also known as the natural coordinates. Obviously, they are not independent, with summation of 1. Consider the vector function constructed by using the area coordinates:

$$\mathbf{N}_1 = (L_2 \nabla L_3 - L_3 \nabla L_2) l_1 \tag{3.20}$$

where l_1 is the length of edge "1." According to the definition of the area coordinates, L_2 is a linear function, with values changing from 1 (corresponding to the node "2") to 0 (corresponding to the node "3"). Likewise, L_3 is a linear function, with values changing from 0 (corresponding to the node "2") to 1 (corresponding to the node "3"). Thus, we have $\hat{e}_1 \cdot \nabla L_2 = -1/l_1$, and $\hat{e}_1 \cdot \nabla L_3 = 1/l_1$. We can then verify the following relationship:

$$\hat{e}_1 \cdot \mathbf{N}_1 = 1 \qquad \text{on edge '1'} \tag{3.21}$$

Again, according to the definition of the area coordinates, the line parallel to edge "2" is an isoline for L_2. So ∇L_2 is perpendicular to edge "2." Additionally, on edge "2,"

$L_2 = 0$. Consequently,

$$\hat{e}_2 \cdot \mathbf{N}_1 = 0 \qquad \text{on edge '2'} \tag{3.22}$$

Analogously, it holds

$$\hat{e}_3 \cdot \mathbf{N}_1 = 0 \qquad \text{on edge '3'} \tag{3.23}$$

This shows on edge "1", the tangential component of $\mathbf{N}_1 E_{t1}$ equals the tangential field component at its "midpoint." Along the remaining edges the tangential components of $\mathbf{N}_1 E_{ti} (i = 2, 3)$ vanish. In a similar approach, we can define

$$\mathbf{N}_2 = (L_3 \nabla L_1 - L_1 \nabla L_3) l_2 \tag{3.24}$$

$$\mathbf{N}_3 = (L_1 \nabla L_2 - L_2 \nabla L_1) l_3 \tag{3.25}$$

In this way, we have constructed the basis functions corresponding to E_{t1}, E_{t2} and E_{t3}, known variously as "Whitney forms" or "edge bases." So the transverse electrical field can be expressed as

$$\mathbf{E}_t = \sum_{i=1}^{3} \mathbf{N}_i E_{ti} \tag{3.26}$$

The longitudinal electrical field can be expressed according to the definition of the area coordinate as

$$E_z = \sum_{i=1}^{3} L_i E_{zi} \tag{3.27}$$

Problem

3.6 Prove the following relation

$$\mathbf{N}_i \times \hat{n} = \mathbf{g}_i$$

where $\mathbf{N}_i$ is the basis function given by (3.20), (3.24), or (3.25), $\mathbf{g}_i$ is the RWG basis function given by (2.1.3), $\hat{n}$ is the normal unit vector of the triangular element.

3.1.3 Discretization of the Functional

By dividing the cross section of the waveguide into a number of small triangular elements, the unknowns in the functional (3.15) in each element can be expressed by using edge-element basis functions as introduced in the preceding section, as

$$\mathbf{e}_t^e = \sum_{i=1}^{3} \mathbf{N}_i^e e_{ti}^e = \{\mathbf{N}^e\}^T \{e_t^e\} = \{e_t^e\}^T \{\mathbf{N}^e\} \tag{3.28}$$

$$e_z^e = \sum_{i=1}^{3} L_i^e e_{zi}^e = \{L^e\}^T \{e_z^e\} = \{e_z^e\}^T \{L^e\} \tag{3.29}$$

where "T" denotes the transpose operation. Inserting (3.28) and (3.29) into (3.15) yields

$$F = \frac{1}{2}\sum_{e=1}^{M}\left(\left\{e_t^e\right\}^T\left[A_{tt}^e\right]\left\{e_t^e\right\} + \beta^2\begin{Bmatrix} e_t^e \\ e_z^e \end{Bmatrix}^T\begin{bmatrix} B_{tt}^e & B_{tz}^e \\ B_{zt}^e & B_{zz}^e \end{bmatrix}\begin{Bmatrix} e_t^e \\ e_z^e \end{Bmatrix}\right) \tag{3.30}$$

where

$$\left[A_{tt}^e\right] = \int_{S^e}\left[\frac{1}{\mu_r}\{\nabla_t \times \mathbf{N}^e\}\cdot\{\nabla_t \times \mathbf{N}^e\}^T - k_0^2\varepsilon_r^e\{\mathbf{N}^e\}\cdot\{\mathbf{N}^e\}^T\right]dS \tag{3.31}$$

$$\left[B_{tt}^e\right] = \int_{S^e}\frac{1}{\mu_r^e}\{\mathbf{N}^e\}\cdot\{\mathbf{N}^e\}^T dS \tag{3.32}$$

$$\left[B_{tz}^e\right] = \int_{s_e}\frac{1}{\mu_r^e}\{\mathbf{N}^e\}\cdot\{\nabla L^e\}^T dS \tag{3.33}$$

$$\left[B_{zt}^e\right] = \left[B_{tz}^e\right]^T \tag{3.34}$$

$$\left[B_{zz}^e\right] = \int_{s^e}\left[\frac{1}{\mu_r^e}\{\nabla_t L^e\}\cdot\{\nabla_t L^e\}^T - k_0^2\varepsilon_r^e\{L^e\}\cdot\{L^e\}^T\right]dS \tag{3.35}$$

where M is the total number of the triangular elements on the cross section, $\left(\varepsilon_r^e, \mu_r^e\right)$ denote the relative permittivity and permeability in the eth element. If $\left(\varepsilon_r^e, \mu_r^e\right)$ are constants inside one element (generally speaking this assumption is valid if the element is small enough), the integrals in (3.31)–(3.35) can be expressed in closed form. By applying

$$\int_{\Omega^e}(L_1)^k(L_2)^l(L_3)^m d\Omega = \frac{k!l!m!}{(k+l+m+2)!}2\Delta \tag{3.36}$$

where Δ is the area of the triangle, (3.31)–(3.35) can be written explicitly as

$$\left[A_{tt}^e\right] = \frac{1}{\mu_r^e}[T]_{3\times 3} - k_0^2\varepsilon_r^e[R]_{3\times 3} \tag{3.37}$$

$$\left[B_{tt}^e\right] = \frac{1}{\mu_r^e}[R]_{3\times 3} \tag{3.38}$$

$$\left[B_{tz}^e\right] = \frac{1}{\mu_r^e}[U]_{3\times 3} \tag{3.39}$$

$$\left[B_{zz}^e\right] = \frac{1}{\mu_r^e}[P]_{3\times 3} - k_0^2\varepsilon_r^e[Q]_{3\times 3} \tag{3.40}$$

where

$$T_{ij} = 4\Delta l_i l_j r_{i_1 i_2} \cdot r_{j_1 j_2} \tag{3.41}$$

$$R_{ij} = \frac{l_i l_j \Delta}{12}\left[\left(1+\delta_{i_1 j_1}\right)g_{i_1 j_1} - \left(1+\delta_{i_1 j_2}\right)g_{i_2 j_1} - \left(1+\delta_{i_2 j_1}\right)g_{i_1 j_2} + \left(1+\delta_{i_2 j_2}\right)g_{i_2 j_2}\right] \tag{3.42}$$

$$U_{ij} = \frac{l_i \Delta}{3}\left(g_{i_2 j} - g_{i_1 j}\right) \tag{3.43}$$

$$P_{ij} = \Delta g_{ij} \tag{3.44}$$

$$Q_{ij} = \frac{\left(1+\delta_{ij}\right)\Delta}{12} \tag{3.45}$$

In the above equations, the subscripts i_1 and i_2 are determined by the subscript i given by Table 3.1. In a similar way, the subscripts j_1 and j_2 are determined by the subscript j. r_{ij} and g_{ij} are given by

$$r_{ij} = \nabla L_i \times \nabla L_j \tag{3.46}$$

$$g_{ij} = \nabla L_i \cdot \nabla L_j \tag{3.47}$$

It is worth noting that when the integrand is complex, for example, in the case utilizing higher-order basis functions, the derivation of the formulae for the matrix entries is usually very complicated. In that case we can apply the numerical integral method as introduced in Section 2.1.5. Although error will be introduced by the numerical integral method, this error and the discretization error in FEM are verified to cancel each other [8]. Moreover, despite the minor additional time spent on the numerical integration, the calculation of the matrix entries takes a small portion of time in the whole FEM solution.

Labeling the unknown interpolation parameters globally (3.30) can be rewritten as

$$F = \frac{1}{2}\{e_t\}^T[A_{tt}]\{e_t\} + \frac{1}{2}\beta^2\begin{Bmatrix} e_t \\ e_z \end{Bmatrix}^T\begin{bmatrix} B_{tt} & B_{tz} \\ B_{zt} & B_{zz} \end{bmatrix}\begin{Bmatrix} e_t \\ e_z \end{Bmatrix} \tag{3.48}$$

Table 3.1 The label system of the edges and the nodes in the triangular element

i	i_1	i_2
1	2	3
2	3	1
3	1	2

Applying the variational principle to (3.48), we obtain the following mathematical standard generalized eigenvalue equation

$$\begin{bmatrix} A_{tt} & 0 \\ 0 & 0 \end{bmatrix} \begin{Bmatrix} e_t \\ e_z \end{Bmatrix} = -\beta^2 \begin{bmatrix} B_{tt} & B_{tz} \\ B_{zt} & B_{zz} \end{bmatrix} \begin{Bmatrix} e_t \\ e_z \end{Bmatrix} \tag{3.49}$$

Apparently, $[A_{tt}]$, $[B_{tt}]$, and $[B_{zz}]$ are all sparse symmetric matrices.

Problem

3.7 Write a program for generating $[A_{tt}]$, $[B_{tt}]$, $[B_{zz}]$ by the analytical method based on (3.42)–(3.52) and the numerical method by the Gauss–Legendre quadrature for triangles.

3.1.4 Imposition of the Boundary Condition

Before solving (3.49), we need to further impose boundary condition (3.2), that is, to force the tangential electric field on the boundary to be 0. In order to do that, we first transform the generalized eigenvalue equation (3.49) into a linear system of equation with the unknown β enclosed in the coefficient matrix. Second, find the equation describing the tangential electric field on the boundary. Third, set the coefficient before this variable to be 1, and all other coefficients to be 0. At the same time, set the RHS value to be the imposed value (0, in this case). Obviously, this modified function results in the solution of the tangential electric field on the boundary to be exactly the imposed value. It is also apparent that the operation is only applied to the row elements and the symmetry of the original matrix will be broken. To circumvent this problem, let us study the effect of other equations by imposing this condition. In this case, since the imposed value is 0, this term contributes null in other equations, regardless of the coefficient. Due to this fact, we can set the coefficients of this variable in other equations to be 0. More generally, if the imposed value is not 0, we can set the coefficients of this variable in other equations to be 0 and transfer the resultant variation to the RHS of the equations. This operation is applied to the column entries and thus the symmetry of the original matrix is maintained. Further observation shows that we can eliminate the imposed unknowns from the linear system of equation. This way of eliminating the imposed unknowns reduces the number of the unknowns of the system and leads to improvement of numerical efficiency, To implement this way, it is better to relabel the unknowns and let the label of the imposed unknowns be the last before we calculate the matrix entries. Otherwise, the complexity of programming will be increased due to the relabeling of the unknowns. Once the boundary condition is imposed, the linear system of equations changes back to the form of a generalized eigenvalue equation and the eigenvalues and the eigenvectors can be solved.

3.1.5 Solution of the Generalized Eigenvalue Equation

To solve equation in form of (3.49) is a classical mathematical problem. Mathematicians have explored this problem for a long time and various algorithms and skills have been developed. They can be grouped according to the essential ideas employed. The widely followed idea in early days was to obtain the eigenvalues and eigenvectors using the bisection algorithm after firstly tridiagonalizing the symmetric matrix by using Householder similarity transformation [9]. Recently, the Krylov subspace method has become popular. The basic idea is to decompose the matrix into the Hessenberg (or tridiagonal) form using Arnoldi or Lanczos algorithms, and keep using iterations similar to the QR algorithm to obtain all the eigenvalues and eigenvectors [10]. Methods of this kind can make full use of the sparsity of the matrix, resulting in low memory requirement and high computational efficiency. In the process of converting this idea to computer programs, mathematicians have developed various numerical techniques according to the characteristics of the computer operations to further reduce the memory requirement and increase the computational efficiency. These techniques evolve with the development of the computer technology. As a computational electromagnetics researcher, it is not an easy task, though possible, to grasp all these techniques and build computer routines accordingly. Fortunately, there are always foreseeing mathematicians who convert the state-of-the-art research into sophisticated black-box program packages with easy interface for a convenient utilization for researchers from other areas. On solving the eigenvalue problem, there were the *EISPACK* program package in the 1970s [11], the *IMSL* software package in the 1980s [12], and the *ARPACK* software package in the 1990s [13]. Moreover, the ARPACK package has been included internally in the commercial software MATLAB®, which is a powerful tool in computation and graphic-processing. They can be readily used. For this reason, this text will not cover the details of the algorithms and techniques in solving the eigenvalue problems. What we wish to discuss here, though, is the possibility to further simplify or transform the equation for E-field, so that it is more suitable to be solved by existing software.

Apparently, due to the singularity of the coefficient matrix in the LHS of (3.49), there exist a series of degenerate eigenvector with eigenvalue of 0. These eigenvectors do not physically exist. They originate from the applying of transformation (3.14). Hence, we further simply (3.49). From (3.49) we obtain

$$[B_{zt}]\{e_t\} + [B_{zz}]\{e_z\} = 0 \tag{3.50}$$

Eliminating $\{e_z\}$ in (3.49) yields

$$[A_{tt}]\{e_t\} = \beta^2[B'_{tt}]\{e_t\} \tag{3.51}$$

where

$$[B'_{tt}] = [B_{tz}][B_{zz}]^{-1}[B_{zt}] - [B_{tt}] \tag{3.52}$$

It can be shown that $[B'_{tt}]$ is still a symmetric matrix. However, compared with the original matrix, this matrix is no longer sparse. Next, we introduce another transformation without breaking the sparsity of the original matrix:

$$\begin{bmatrix} B_{tt} & B_{tz} \\ B_{zt} & B_{zz} \end{bmatrix} \begin{Bmatrix} e_t \\ e_z \end{Bmatrix} = \frac{\theta^2}{\theta^2 - \beta^2} \begin{bmatrix} B_{tt} + \dfrac{1}{\theta^2}[A_{tt}] & B_{tz} \\ B_{zt} & B_{zz} \end{bmatrix} \begin{Bmatrix} e_t \\ e_z \end{Bmatrix} \tag{3.53}$$

where $\theta^2 = k_0^2 \varepsilon_{\max} \mu_{\max}$, $\varepsilon_{\max}$ and $\mu_{\max}$ are the maximum permittivity and maximum permeability of the waveguide loading material, respectively. Obviously, θ is the propagation constant β of a TEM wave in a homogeneous ($\varepsilon_{\max}$, $\mu_{\max}$) material. Generally speaking, the propagation constant in a lossless dielectrically loaded waveguide follows that $\beta^2 < \theta^2$, and a mode closer to the dominant mode has a propagation constant β closer to θ, and hence a bigger $\theta^2/(\theta^2 - \beta^2)$. Consequently, the eigenvalues of (3.53) in ascending order correspond to the dominant to higher-order modes of a dielectrically loaded waveguide. This result is very useful since in reality we usually are only interested in the first few modes of the waveguide. For convenience, we write (3.51) and (3.53) uniformly as the following standard generalized eigenvalue equation

$$[A]\{e\} = \lambda [B]\{e\} \tag{3.54}$$

3.1.6 Computer Programming

From the above analysis, it can be seen that the computer program for the eigenvalue problem of a dielectric-loaded waveguide should be made up of four modules, that is, obtaining the waveguide meshing information, forming the discretized matrices, imposing the boundary condition and solving the eigenvalue equations. Before we go into details, let us clarify the storage method of the final generalized eigenvalue equation matrices, because it greatly affects the programming in every part.

The most straightforward way is to use two 2D arrays to store $[A]$ and $[B]$, respectively. Despite the convenience in the application, this manner does not make use of the sparsity of the matrix. Therefore, a large amount of memory is wasted and the computational efficiency is low. There are many storage methods making use of the sparsity of the matrix. Here, we introduce a *one-dimensional(1D) row-indexed storage method*. Take the sparse matrix $[A]$ with dimension $N \times N$ as an example. This method uses two 1D arrays *sa* and *ija* to store it. Array *sa* is used to store the nonzero entries in $[A]$, which are complex numbers in general case and real number if the problem is lossless. Array *ija* is an integer type array, and is used to store the column indices of the entries in *sa*. The detailed steps are as follows:

- The first N elements of *sa* store the diagonal entries of $[A]$, regardless of whether the value is zero or not. It is worth noting that usually the diagonal entries of the FEM matrices are nonzero values.

- The first N elements of *ija* store the indices of the first nonzero and nondiagonal entries in each row of $[A]$ in *sa*. If there is no such element in a row, then the element is given by the index of the last nonzero element in the upper row of $[A]$ in *sa* increased by one. The first element of *ija* is always $N+2$.
- The $(N+1)$th element of *sa* is reserved. We can fill it with an arbitrary value.
- The $(N+1)$th element of *ija* stores the index of the last nonzero and nondiagonal entries of $[A]$ in array *sa* increased by one.
- From the $(N+2)$th element, the nonzero and nondiagonal entries in every row of $[A]$ are stored in *sa*. The storage is performed in a succeeding manner from the first row. In each row, the entries are stored according to ascending column index order.
- From the $(N+2)$th element, the corresponding column indices of the *sa* elements in the original matrix $[A]$ are stored in *ija*.

Denote the sum of the total number of the nonzero and nondiagonal entries and the number of the diagonal entries of $[A]$ by *nnz*. Then, the sizes of the arrays *sa* and *ija* both equal $nnz+1$. The key point in effectively reassembling matrix $[A]$ is to determine the number of nonzero and nondiagonal entries in each row of $[A]$ and their corresponding column indices. Now, we are ready to further analyze the programming of the four modules.

In the first module, the waveguide meshing information is obtained. This module involves the following functions. First, retrieve the basic information from input files. Second, convert the information to the form required by the following routines. Now there are many commercial software such as ANSYS or Hypermesh that can automatically mesh the waveguide into triangular elements, and they usually provide: the total number of the nodes and triangle elements (stored using integer variables *maxnodes* and *maxeles*, respectively), the node coordinates (stored using real array $xynodes(2, maxnodes)$), the relation between the local and the global labels of the nodes in a triangular element, and the media type (stored using integer array $\ln(4, maxeles)$, with the first three columns storing the global indices of the three nodes and the last column indicating the type of the material). These are basic data and are readily provided. However, this information form cannot meet the requirement of the following programming, since the interpolation parameters used in the edge-element method are the tangential electric fields on the edges of the triangle elements. Therefore, we need to further label all the edges and provide the relations of these local and global labels. In order to conveniently impose boundary condition, we normally label the interior edges first and then label the edges on the boundary. For the same reason, we'd better relabel all the nodes, that is, label the interior nodes before the nodes on the boundary. Third, we need to obtain the number of the nonzero entries in each row of $[A]$. Now, the functions in this module is clear: first, read data from an input file and get: *maxnodes, maxeles, xynodes*(*2, maxnodes*, and ln(4, *maxeles*); second, separate the interior edges with the boundary ones, label them, and provide the relations of the local and global indices *le(3,maxeles)*; thirdly, relabel the nodes,

provide the new sets of *xynodes(2,maxnodes)* and *ln(4,maxeles)*; finally obtain the total number of the nonzero elements in each row of $[A]$. The programming of the first step is quite straightforward, and the processes in the second and the third steps are similar, so we only provide more details on the programming of the second and the last step in the following.

As we know that an edge in the interior of a 2D domain is always shared by two triangular elements, while that on the boundary is only owned by one element. This fact can be used in separating the interior and the boundary edges. The most direct way is to compare every edge in every triangle with all the edges from all other triangles. If there exist identical pairs, the edge is an interior one, otherwise it is on the boundary. This process requires a double loop from 1 to *maxeles*, and the computational complexity is $O(maxeles^2)$. This is too heavy to be accepted. Next, we will introduce another scheme. First, define $y = n_1\sqrt{n_2}$ for every edge, where n_1 and n_2 are the global indices for the nodes of the edge. Second, sort y by using a fast sorting algorithm, such as the *Heapsort* [14]. Then, compare the sorted sequence. If the entry in the sequence of y is unique, then the corresponding edge is on the boundary. If there are duplicated elements, the boundary edge can be separated by further comparing the corresponding nodes. It can be seen that most computations are spent on the *Heapsort*, which is $O(maxeles \lg maxeles)$ and is a significant improvement compared to the direct method. Moreover, based on this sorting, we can index y as desired, and obtain the relations between the triangular elements' local and global indices, that is, $le(3, maxeles)$, according to the relationship of y and the edges. Thus, the second step is accomplished.

For convenience, here we use $[B_{zz}]$, as an example to illustrate the process to determine the number of the nonzero elements in each row of the matrix. By examining the step in assembling the descretized matrix, we know that the total number of the nonzero elements in row i equals the total number of the nodes included in all the elements sharing the same node globally indexed with i. Therefore, we need to first identify the total number of the elements sharing the same node (stored using an integer array $ip(maxnodes)$) and the global indices of these elements (stored using a 2D integer array $ipele(maxadj, maxnodes)$). This is not hard to do. We only need to set a double loop for every element and for every node in the element. More specifically, the Fortran codes can be written as

```
do i = 1,maxeles
do j = 1,3
   n = ln(j,i)
   ip(n) = ip(n) + 1
   ipele(ip(n),n) = i
end do
end do
```

Then, by using *ip* and *ipele*, we retrieve the global indices of all other nodes that are in the elements sharing the same node globally indexed as *i*. By sorting *ipele* for each node, we can easily find the total number of the nonzero entries in row *i* of $[B_{zz}]$ (stored using integer array $iq(maxnodes)$).

Once we obtain *iq*, we can go into more details in the program of the assembling matrix module. The entries of the discretized matrices can be obtained directly according to (3.37)–(3.45). Here, we would like to talk more about the assembling of the matrices. Let us still use the example of $[B_{zz}]$. In the assembling procedure, we need to consider how to build the global matrix $[B_{zz}]$ by using the element matrix $[B_{zz}]^e_{3\times 3}$ generated by the element globally indexed as *e*. As we know, to store a sparse matrix $[B_{zz}]$ using "1D array indexed-by-row" scheme, we need to use two 1D arrays *sa* and *ija*. With *iq* obtained before, it is easy to get *ija*. Once we obtain *ija*, we can write the following Fortran code to assemble the matrix

```
do 10 i=1,3
do 10 j=1,3
    ii=ln(i,p)
    jj=ln(j,p)
    if(ii.eq.jj)then
       sa(ii)=sa(ii)+bzze(i,j)
else
       do k=ija(ii),jia(ii+1)-1
if(jj.eq.ija(k))then
      sa(k)=sa(k)+bzze(i,j)
      goto 10
else
 endif
enddo
  10  continue
  10  continue
```

where $bzze(i,j)$ denotes the elemental matrix $[B_{zz}]^e_{3\times 3}$. The module imposing the boundary condition can be written separately or be included in the matrix assembling module. Denote the total number of the nodes as maxinnodes. Since we have reindexed the nodes following the interior to boundary sequence, the boundary processing can be achieved by adding one line beneath the fourth line in the above codes:

```
if(ii.gt.maxinnodes) goto 10
```

In this way, array *sa* stores the *maxinnodes*-dimensional $[B_{zz}]$ with boundary condition readily imposed. As for the module solving for the eigenvalue and eigenfunction (3.54)

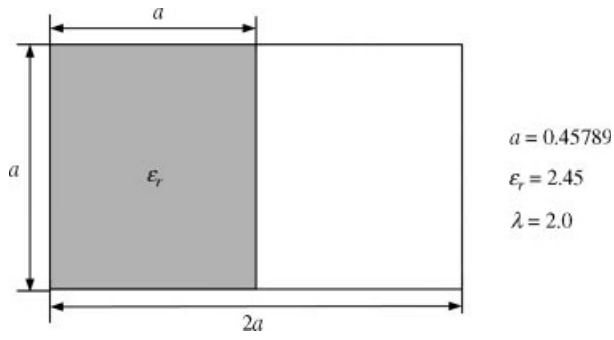


Figure 3.2 The cross section of the rectangular waveguide partially filled with dielectrics.

Table 3.2 Propagation constants of the half-filled waveguide

Eigenmodes	Analytical results	Numerical results
Dominant mode	2.51335	2.53346
The first-order mode	−5.41743	−5.51886
The second-order mode	−5.55289	−5.65088

(3.54), we can use software packages such as the ARPACK, and do not need to write codes ourselves.

3.1.7 Numerical Examples

Here, we only use one numerical example as in [15] to illustrate the effectiveness and the accuracy of this method. The geometry and dimension of the model is provided in Figure 3.2. In Table 3.2, the first three propagation constants of the partially dielectric-loaded waveguide calculated by using this method is compared with the analytical results. From the table it can be seen that the FEM results agree well with the mode-matching results. No spurious solution is found in the numerical results. In this numerical experiment, the whole cross section of the waveguide is divided into 100 triangular elements, which illustrates the accuracy of FEM. More numerical examples can be found in reference [15].

3.2 Discontinuity Problem in Waveguides

Considerable analysis about microwave devices such as filters, the directional coupler and the microwave circulator can be regarded as waveguide discontinuity problems. Such kinds of problem can be further grouped into open waveguide discontinuity problem and closed waveguide discontinuity problems. The discontinuity problem in

optical fibers belongs to the former and that in metallic waveguide belongs to the latter. These two groups of problems are essentially different. Since the scattered field in an open waveguide radiates into unlimited space, the problem domain is infinite. On the other hand, the scattered field in a closed waveguide is constrained inside the waveguide. Consequently, the problem domain is within the waveguide. If the radiation in the discontinuity problem in an open waveguide is weak and can be ignored, the analysis is the same as the problem in a closed waveguide. Otherwise, the FEM analysis becomes difficult. In that case, we usually avoided using FEM in practical applications but use MoM or the Hybrid FE-BI-MLFMA, to be introduced in Chapter 5 instead. From the above analysis it can be seen that as far as the 3D waveguide discontinuity problem is concerned, the closed waveguide one is more typical. Next, we will illustrate the theory and techniques in FEM solution by solving a discontinuity problem in a closed waveguide.

3.2.1 Functional Formulation

As shown in Figure 3.3, there is a dielectric block in a rectangular waveguide. When the incident TE_{10} wave propagates and encounters this block, a portion of the power carried by the incident wave will be reflected and another portion will pass and

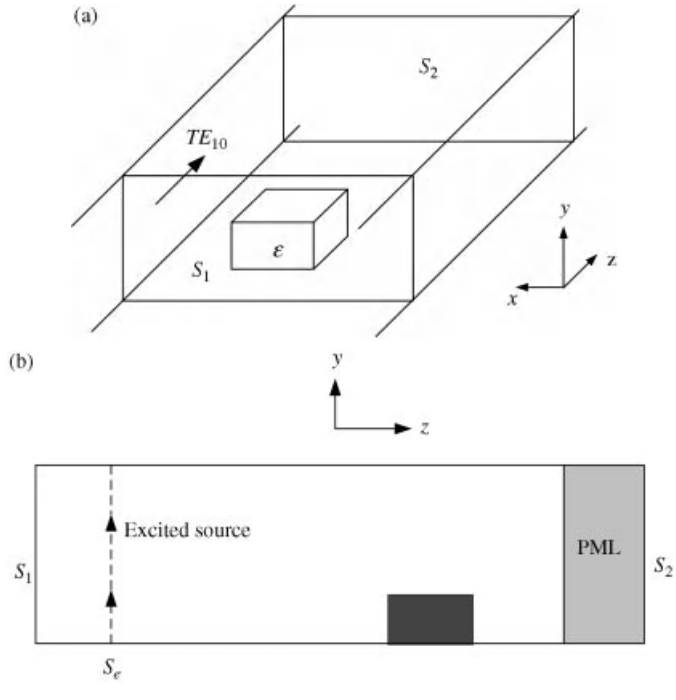


Figure 3.3 Model of a 3D discontinuous waveguide problem.

continue to propagate along the waveguide. Our goal is to calculate the amplitudes and phases of the scattered wave and the transmitted wave. In other words, we want to obtain the scattering matrix of the discontinuous two-port network, as defined in reference [3]. The key problem in this scattering matrix calculation is to determine the field distribution in the waveguide. To this end, we need to solve the Maxwell's equations with incidence of the TE_{10} wave in the problem domain defined by the waveguide. However, since the space inside an infinitely long waveguide is also infinite, it is neither possible nor necessary to solve Maxwell's equation in such a domain. Therefore, we need to truncate the infinite long waveguide. There are many forms of truncation. As shown in Figure 3.3a, we define two fictitious boundary surfaces S_1 and S_2 that are not far from the dielectric block and use the volume enclosed by S_1, S_2 and the boundary of the waveguide as the computational domain. Obviously, the key issue in a manner of truncation is to determine the boundary condition on S_1, S_2 (the third-type boundary condition). From the waveguide theory, the field on sections S_1 and S_2 can be obtained by the linear superposition of the eigenmodes of the waveguide. Thus, the boundary condition can be derived, as presented in reference [16]. The efficiency of this truncation is high, because S_1 and S_2 can be put very close to the dielectric block, and the computational domain is very small. However, the implementation is neither easy nor general. Since we have to precalculate the eigenmodes of the waveguide and write different codes for different waveguide. For this reason, we will introduce another scheme in the following. Despite the fact that it is not as efficient, the programming can be made very general.

As shown in Figure 3.3b, we use metallic boundaries S_1 and S_2 to terminate the problem domain in proper distances away from the dielectric scatterer. We still use the volume enclosed by S_1, S_2 and the boundary of the waveguide as the computational domain. Then, we add a source $\mathbf{J}^{\mathrm{e}} = \delta(x - a/2, z - z_e)\hat{e}_y$ on the center line of a fictitious surface S_e not far from S_1. The boundary condition is easy to implement. The problem here is whether it is equivalent to the original problem. To be specific, we are concerned as to whether the same scattering matrix as in the original problem can be achieved. First, let us compare the incidences in these two problems. The incidence on S_e in the postulated problem will excite multiple modes. If the waveguide is operating at the dominant mode, all higher-order modes will attenuate rapidly. This is equivalent to the original problem. If the waveguide is operating on multimode transmission, then multiple modes including the dominant mode are supported in the waveguide, and the postulated problem slightly differs from the original problem. However, from the retrieval of the scattering matrix method to be introduced in Section 3.2.3, we know that even for multimode transmission, we can still obtain the scattering matrix for the dominant mode through the analysis of the field distribution in the postulated problem. Let us move onto the transmitted waves in this problem. The transmitted wave in the postulated problem will be completely reflected by the metallic boundary S_2. The reflected wave will again impinge on the dielectric body and be partially reflected and partially transmitted. This will repeat again and again. As a result, the final field

distribution deviates from the original problem and we cannot obtain the required information. Therefore, we need to put in front of S_2 an absorbing material, in which the transmitted wave will be attenuated without reflection. Hence, when the wave reaches S_2, it is so weak that can be ignored. In this way, we can retrieve the scattering matrix for the dominant mode of the original problem from the field distribution in the postulated problem. One may wonder whether the presence of the metallic boundary S_1 will affect the field distribution. In fact, the theoretical analysis of the active transmission line [17] tells us that the metallic boundary S_1 on the left side of the source will only affect the power distribution but not the field distribution pattern on both sides. Now, we are ready to present the mathematic expression of the postulated problem in the following.

From Section 1.2 we know that the deterministic partial differential equation for this postulated problem is

$$\nabla \times [\mu_r]^{-1} \nabla \times \mathbf{E} - k_0^2 [\varepsilon_r] \mathbf{E} = -jk_0 Z_0 \mathbf{J}^{\mathrm{e}} \tag{3.55}$$

$$\hat{n} \times \mathbf{E} = 0 \qquad \text{on the boundary} \tag{3.56}$$

where $k_0 = \omega\sqrt{\mu_0 \varepsilon_0}$ and $Z_0 = \sqrt{\mu_0/\varepsilon_0}$ are the wave number and the wave impedance in free space respectively. The relative permittivity $[\varepsilon_r]$ and the relative permeability $[\mu_r]$ are determined according to the material of the scatterer. From reference [18] or Chapter 4 in this text we know that, the relative permittivity $[\varepsilon_r]$ and permeability $[\mu_r]$ of the *perfectly matched layer(PML)* takes an identical form, written as:

$$[\varepsilon_r] = [\mu_r] = \begin{bmatrix} \Lambda_x & 0 & 0 \\ 0 & \Lambda_y & 0 \\ 0 & 0 & \Lambda_z \end{bmatrix} \tag{3.57}$$

where

$$\Lambda_x = s_z, \Lambda_y = s_z, \Lambda_z = \frac{1}{s_z} \tag{3.58}$$

where $s_z = \beta - j\alpha$ is the complex coordinate stretching factor in the z direction, in which β determines the wavelength in the anisotropic absorber and α determines the rate of decay of the transmitted wave. The choice of larger α and β leads to faster damping of the transmitted wave. Consequently, higher computational efficiency can be achieved since the thinner layer of the PML results in a smaller computational domain. However, larger α and β also brings higher reflection between the discretized FEM elements and hence reduces accuracy in the final solution. So we need to consider the trade-off in choosing α and β. Fortunately, the final results are not sensitive to the choice of α and β, so they are not too hard to choose. Based on the authors' experience, the choice of $\alpha = \beta = 1.0$ and the PML thickness of 0.3 waveguide wavelength is usually adequate.

By comparing (3.55) with (3.1) in the preceding section, we find they are almost identical except for the excitation. By an analogous procedure, we can arrive at the variation of the functional for (3.55):

$$\begin{cases} \delta F(\mathbf{E}) = 0 \\ \hat{n} \times \mathbf{E} = 0 \qquad \text{on the boundary} \end{cases} \tag{3.59}$$

where

$$F(\mathbf{E}) = \frac{1}{2}\int_V \left[(\nabla \times \mathbf{E}) \cdot ([\mu_r]^{-1} \cdot \nabla \times \mathbf{E}) - k_0^2 \mathbf{E} \cdot [\varepsilon_r] \cdot \mathbf{E} \right] dV + jk_0 Z_0 \int_V \mathbf{J}^e \cdot \mathbf{E} dV \tag{3.60}$$

3.2.2 *Choice of the Basis Functions*

In contrast to the functional for the waveguide eigenmode problem (3.15), the functional for the discontinuous waveguide problem (3.60) is a volume integration. Correspondingly, the discretize elements are volume elements instead of surface ones. The tetrahedron is the most flexible simplex for a volume. The interpolation variables can be chosen as the tangential electric fields on the central points of every edge, as shown in Figure 3.4. It can be seen that the interpolation parameter here is simply

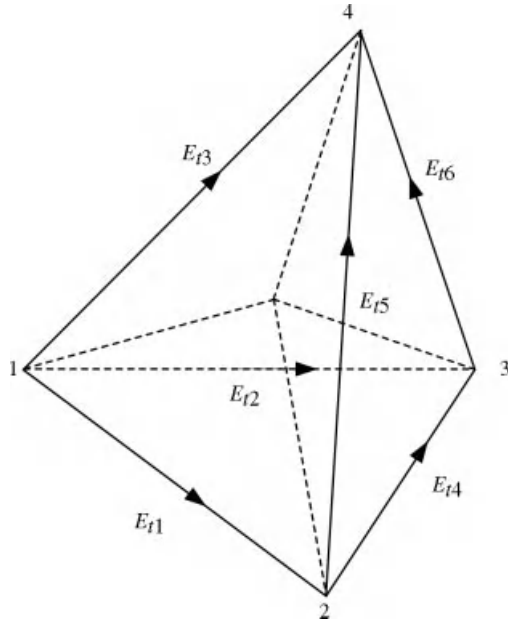


Figure 3.4 The edge-elements of FEM defined on a tetrahedron.

an analog of the edge-element interpolation parameter introduced in Section 3.1.2. Therefore, the edge-element basis function in Figure 3.4 can be constructed in a similar way to that of the triangular elements. First, let us define the volume coordinate on the tetrahedron. As shown in Figure 3.4, we label the four nodes of the tetrahedron as 1, 2, 3, and 4, and denote the volume of the tetrahedron by V. $P(x, y, z)$ is an arbitrary point in the tetrahedron. By connecting $P(x, y, z)$ and the four nodes, the original tetrahedron is partitioned into four smaller tetrahedra, with the volumes denoted as V_1, V_2, V_3, and V_4, respectively. They can be calculated straightforwardly as:

$$V = \frac{1}{3!}\begin{vmatrix} 1 & x_1 & y_1 & z_1 \\ 1 & x_2 & y_2 & z_2 \\ 1 & x_3 & y_3 & z_3 \\ 1 & x_4 & y_4 & z_4 \end{vmatrix} \tag{3.61}$$

and

$$V_1 = \frac{1}{3!}\begin{vmatrix} 1 & x & y & z \\ 1 & x_2 & y_2 & z_2 \\ 1 & x_3 & y_3 & z_3 \\ 1 & x_4 & y_4 & z_4 \end{vmatrix} \tag{3.62}$$

$$V_2 = \frac{1}{3!}\begin{vmatrix} 1 & x_1 & y_1 & z_1 \\ 1 & x & y & z \\ 1 & x_3 & y_3 & z_3 \\ 1 & x_4 & y_4 & z_4 \end{vmatrix} \tag{3.63}$$

V_3 and V_4 can be obtained analogously. The volume coordinates of $P(L_1, L_2, L_3, L_4)$ are defined as

$$L_i = \frac{V_i}{V}, \quad i = 1, 2, 3, 4 \tag{3.64}$$

Obviously these coordinates are not independent, with the following relationship holding

$$\sum_{i=1}^{4} L_i = 1 \tag{3.65}$$

After labeling the edges and the nodes of the tetrahedron as in Table 3.3, the corresponding basis function of the interpolation variable E_i can be expressed as

Table 3.3 Local indices of the edge with the corresponding nodes in the tetrahedron

Local index of the edge i	Local index of the vertex i_1	Local index of the vertex i_2
1	1	2
2	1	3
3	1	4
4	2	3
5	2	4
6	3	4

$$\mathbf{N}_i = (L_{i1}\nabla L_{i2} - L_{i2}\nabla L_{i1})l_i \tag{3.66}$$

where l_i is the length of edge i. Thus, the electric field in the tetrahedron can be expressed as

$$\mathbf{E} = \sum_{i=1}^{6} \mathbf{N}_i E_{ti} \tag{3.67}$$

3.2.3 Discretization of the Functional

If we divide the computational domain into a number of tetrahedra with neither gaps nor overlap in between, then, in each element, the unknowns in functional (3.60) can be expressed by using the basis function given in the preceding section as

$$\mathbf{E}^e = \sum_{i=1}^{6} \mathbf{N}_i^e E_i^e = \{\mathbf{N}^e\}^T\{E^e\} = \{E^e\}^T\{\mathbf{N}^e\} \tag{3.68}$$

Inserting (3.68) into (3.60) produces

$$F = \sum_{e=1}^{M} \left(\frac{1}{2}\{E^e\}^T[A^e]\{E^e\} - \{E^e\}^T\{b^e\}^e\right) \tag{3.69}$$

where

$$[A^e] = \int_{V^e} \left[\{\nabla \times \mathbf{N}^e\}\left[v_r^e\right]\{\nabla_t \times \mathbf{N}^e\}^T - k_0^2\{\mathbf{N}^e\}[\varepsilon_r]^e\{\mathbf{N}^e\}^T\right] dV \tag{3.70}$$

$$\{b^e\} = -jk_0Z_0\int_{V^e} \mathbf{J}\cdot\{\mathbf{N}^e\}^T dV \tag{3.71}$$

where $\left[\nu_r^e\right] = \left[\mu_r^e\right]^{-1}$, and $\left[\varepsilon_r^e\right]$, $\left[\mu_r^e\right]$ denote the relative permittivity and permeability tensors in element e. If they are constants inside one element (generally speaking this postulate is valid if the element is small enough), the integrals in (3.70)–(3.71) can be expressed in closed form. By applying

$$\iiint_{V^e} (L_1)^k (L_2)^l (L_3)^m (L_4)^n dV = \frac{k!l!m!n!}{(k+l+m+n+3)!} 6V^e \tag{3.72}$$

where V^e is the volume of the tetrahedron, (3.70) can be written explicitly as

$$[A^e] = [T]_{6\times 6} - k_0^2 [R]_{6\times 6} \tag{3.73}$$

where

$$T_{ij} = 4Vl_i l_j r_{i_1 i_2} \cdot \left[\nu_r^e\right] \cdot r_{j_1 j_2} \tag{3.74}$$

$$R_{ij} = \frac{l_i l_j V}{20} \left[\left(1+\delta_{i_1 j_1}\right) g_{i_1 j_1} - \left(1+\delta_{i_2 j_1}\right) g_{i_2 j_1} - \left(1+\delta_{i_1 j_2}\right) g_{i_1 j_2} + \left(1+\delta_{i_2 j_2}\right) g_{i_2 j_2}\right] \tag{3.75}$$

In the above equations, the subscripts i_1 and i_2 are determined by the subscript i given by Table 3.3. In a similar way, the subscripts j_1 and j_2 are determined by the subscript j. Furthermore, r_{ij} and g_{ij} are given by

$$r_{ij} = \nabla L_i \times \nabla L_j \tag{3.76}$$

$$g_{ij} = \nabla L_i \cdot \left[\varepsilon_r^e\right] \cdot \nabla L_j \tag{3.77}$$

Besides, the above integration can also be obtained by applying the following numerical integration equation

$$\iiint_{V^e} F(L_1, L_2, L_3, L_4) dL_1 dL_2 dL_3 = \sum_{i=1}^{m} W_i F(L_{1i}, L_{2i}, L_{3i}, L_{4i}) \tag{3.78}$$

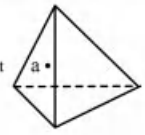
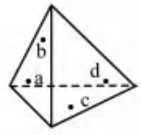
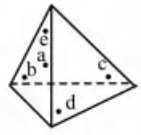
The m in (3.78) denotes the number of samples, where the sample coordinates $(L_{1i}, L_{2i}, L_{3i}, L_{4i})$ and the corresponding weights W_i (i = 1, 2, 3, 4) are given in Table 3.4. By indexing the interpolation parameters globally, we can rewrite (3.69) as

$$F = \frac{1}{2}\{E\}^T [A]\{E\} - \{E\}^T \{b\} \tag{3.79}$$

By calculating the variation of (3.79), we obtain the following system of equation

$$[A]\{E\} = \{b\} \tag{3.80}$$

Table 3.4 Numerical integration schemes of the tetrahedron

Type of element	Graphics	Error	Integration points	Volume coordinates	Weights
linear element		$R=O(h^2)$	a	$\frac{1}{4},\frac{1}{4},\frac{1}{4},\frac{1}{4}$	1
second-order element		$R=O(h^3)$	a	α,β,β,β	$\frac{1}{4}$
			b	β,α,β,β	$\frac{1}{4}$
			c	β,β,α,β	$\frac{1}{4}$
			d	β,β,β,α	$\frac{1}{4}$
				$\alpha=0.58541020$ $\beta=0.13819660$	
third-order element		$R=O(h^4)$	a	$\frac{1}{4},\frac{1}{4},\frac{1}{4},\frac{1}{4}$	$-\frac{4}{5}$
			b	$\frac{1}{3},\frac{1}{6},\frac{1}{6},\frac{1}{6}$	$\frac{9}{20}$
			c	$\frac{1}{6},\frac{1}{3},\frac{1}{6},\frac{1}{6}$	$\frac{9}{20}$
			d	$\frac{1}{6},\frac{1}{6},\frac{1}{3},\frac{1}{6}$	$\frac{9}{20}$
			e	$\frac{1}{6},\frac{1}{6},\frac{1}{6},\frac{1}{3}$	$\frac{9}{20}$

Obviously, $[A]$ is a sparse symmetric matrix. By imposing the boundary condition (3.56) to (3.80) as introduced in Section 3.1.5, we can solve for the unknown interpolation parameters $\{E\}$, and then obtain the field distribution in the whole domain by applying (3.68).

Problem

3.8 Write a program for generating $[A]$, $\{b\}$ by the analytical method based on (3.78) and the numerical method by the Gauss–Legendre quadrature for tetrahedra.

3.2.4 Solution of the Linear Equations

Generally speaking, methods used to solve the linear algebraic equations or matrix equation (3.80) can be categorized into two groups. One is the direct method such as the LU decomposition method. The other is the iteration method, such as the conjugate gradient method. These two kinds of methods originated a long time ago, and they are still evolving towards more stable and more efficient methods with the development of the computer technology. In the direct method, the sparseness usually has to be broken and cannot be fully taken advantage of in the solving procedure. How to take advantage of sparseness, in other words, how to minimize the number of nonzero fill-ins, is a key

factor in successfully developing an efficient direct solver. A great effort has been made in the recent years and various techniques have been designed to efficiently solve large sparse matrix equations. For example, by reordering the unknowns, the distribution of the nonzero entries will be improved, and the nonzero fill-ins required from the LU decomposition will be dramatically reduced. Figure 3.5a shows the distribution of the nonzero entries generated in the FEM solution of a discontinuous

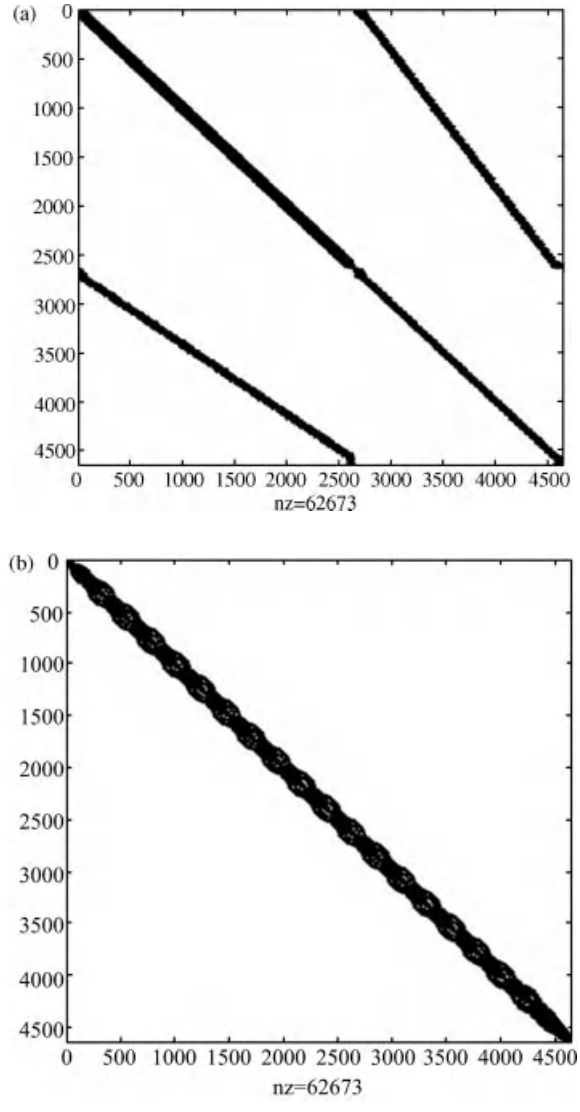


Figure 3.5 The distribution of the nonzero entries of the FEM matrices. (a) The original matrix; (b) after RCM sorting.

waveguide problem. After this matrix is processed by a reordering scheme termed reverse Cuthill–Mckee (RCM), the nonzero entries locate mainly in the vicinity of the diagonal line, as can be seen from Figure 3.5b. This will effectively suppress the increase of the nonzero elements introduced by LU decomposition. As to the iteration methods, a number of schemes based on the Krylov subdomain projection theory and the preprocess techniques, speeding up the iteration, have sprung up in recent years, such as the generalized minimal residual (GMRES), the quasiminimal residual (QMR), and the Incomplete LU decomposition preconditioning. A number of sophisticated techniques have been integrated in MATLAB®. For example, the direct method for solving the linear system of equation in MATLAB® is the *UMFPACK* package developed by T.A. Davis and Duff [19], which is based on the unsymmetric-pattern multifrontal method [19]. S*uperLU* is another example based on the direct method. It was developed by Demmel, Glibert, and Li *et al.*, and can be downloaded from Web site http://www.netlib.org/scalapack/prototype. The packages based on the iteration method, such as GMRES, QMR, and BICGSTAB, are all integrated in MATLAB®. For a computational electromagnetic researcher, it is important to select the most suitable one from these packages. Take the equation (3.80) as an example. The direct method and the iteration method are suitable for different cases. Normally, the direct method demands more computer memory, but it is numerically more stable and more efficient than the iteration method. Since the condition number of the finite-element matrix $[A]$ is generally not good, when the matrix dimension is large, convergence is often slow, and sometimes is even impossible to achieve. Figure 3.6 shows the iteration results while using the conjugate gradient method (CG) and the biconjugate gradient method (BCG) to solve the matrix equations in the above discontinuity problem. It can be seen that in solving this linear system of equations with 21 408 unknowns, after 10 000 iterations, CG results in the normalized residual norm of 0.03, and the BCG results do not converge at all. Therefore, we need to choose the method according to the specific problem and the corresponding matrices. For a special problem, we can normally try a combination of the two methods, that is, using the direct method such as the incomplete LU decomposition method to precondition the matrix, and then use the iteration method to solve the equations.

3.2.5 *Extraction of the Scattering Parameters*

Once the unknown parameter $\{E\}$ is solved for, from (3.80), the field in the whole computational domain can be obtained by (3.68), from which we can further extract the scattering parameters. The electric field along the longitudinal axis of the waveguide can be expressed as

$$E_y(x = a/2, y = b/2, z) = \sum_{i=1}^{d} c_i e^{js_i z} + n(z) \tag{3.81}$$

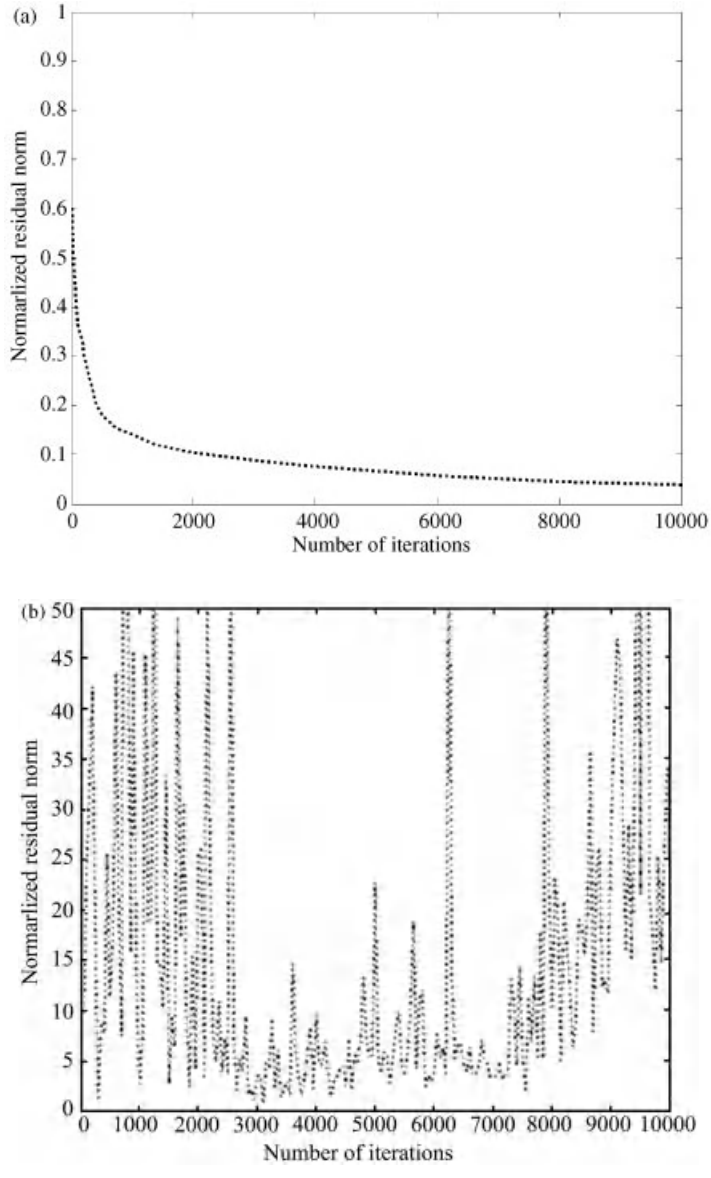


Figure 3.6 Iteration results of the FEM equation. (a) Conjugate gradient method (CG); (b) biconjugate gradient method (BCG).

where s_i is the propagation constant of the ith propagation mode. If s_i is positive, the wave propagates along the $-z$ direction. Otherwise, the wave propagates along the $+z$ direction. c_i is the corresponding complex coefficient, and $n(z)$ is the noise introduced by numerical computation. Obviously, the extraction of the scattering parameters requires obtaining s_i and c_i through the computation result

$E_y(x = a/2, y = b/2, z)$. There are many ways to do this. We can simply determine s_i and c_i by using three samples of $E_y(x = a/2, y = b/2, z)$ [20]. Of course, this way is only suited only for the dominant mode transmission case, and the numerical result is not stable. Alternatively, the Prony method can be used [21]. However, this method is sensitive to the noise $n(z)$. Here, we recommend an *estimation of signal parameters via rotational invariance technique, ESPRIT*. The user needs only to approximate d according to the physical phenomena and select $N > 2d + 1$ samples, in order to obtain s_i and c_i. Since it is beyond the scope of this text, the reader is referred to reference [22] for more details.

3.2.6 Numerical Examples

Since we have obtained the discretized formula to calculate the discontinuous waveguide, we can write the computer code accordingly. We do not give more details here since similar key issues have been discussed in Section 3.1.7. Here, we only demonstrate the effectiveness of the method by presenting a result from [23]. The geometry and dimension of the model is plotted in Figure 3.7. The relative permeability tensor of the ferrite is

$$[\mu_r] = \begin{bmatrix} 1 & 0 & 0.5j \\ 0 & 1 & 0 \\ -0.5j & 0 & 1 \end{bmatrix}$$

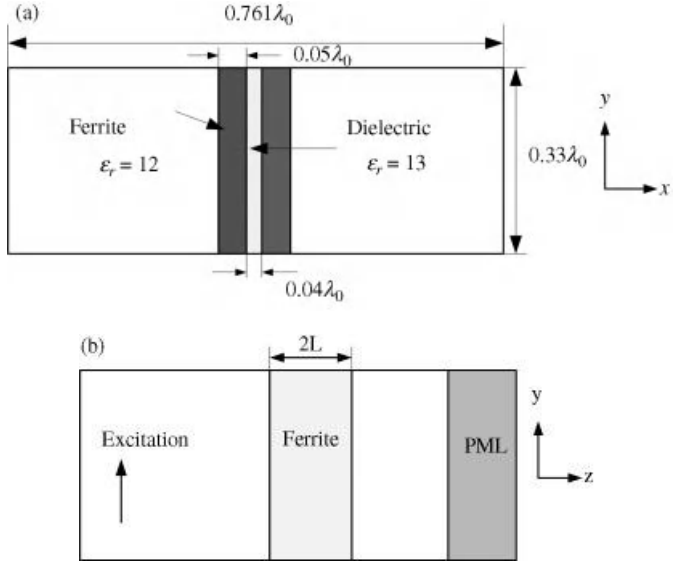


Figure 3.7 The waveguide model of the numerical experiment. (a) The x-y cross section; (b) the y-z cut-plane of the waveguide.

Figure 3.8a shows the calculated y component of the electric field along the z direction in a central y-z cut-plane of the waveguide with the length of the ferrite to be $L = 0.125\lambda_0$. It can be seen from the figure that when the incident wave encounters the ferrite, part of it is reflected, forming the standing wave due to the superposition with the incidence; and the other part gets transmitted and is absorbed by the PML. This numerical result agrees well with the analysis of the physical phenomena. The absolute value of the reflection coefficient obtained from the distribution of the electric field is shown in Figure 3.8b. Again, the numerical result agrees well with the mode-matching result, which verifies the accuracy of FEM. The interested reader is referred to references [16,23] for more numerical results.

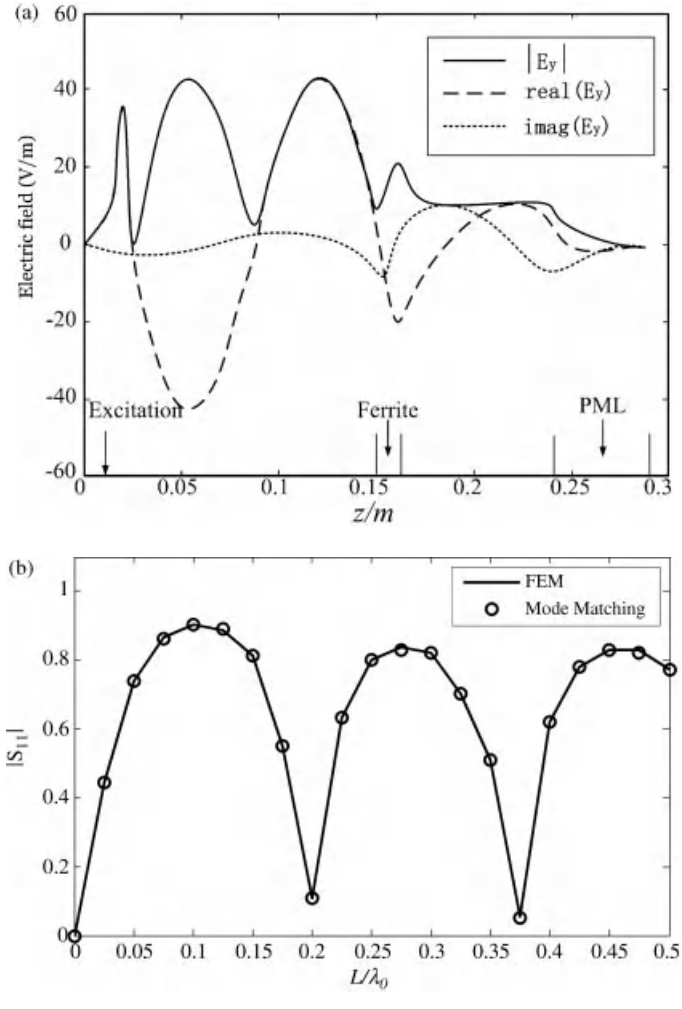


Figure 3.8 Numerical results. (a) The calculated E_y along the z direction in a central y-z cut-plane of the waveguide. (b) The absolute value of the reflection coefficients.

3.3 Scattering from Three-Dimensional Objects

3.3.1 Mathematic Formulation of the Problem

The scattering from a 3D object has been discussed with great emphasis under the topic of MoM in Chapter 2. The reason we come back to this problem here, is to envisage the method instead of being interested in the problem itself. Through the study of the FEM procedure in solving this problem, we aim to give the reader a flavor of the characteristics of FEM solution for this problem, and a comparison of FEM with MoM.

As shown in Figure 3.9, the complex scatterer is made up of three parts: part I is a perfect electric conductor (PEC), part II is a perfect magnetic conductor (PMC), and part III is an inhomogeneous dielectric. Next, we will show how to use FEM to calculate the EM scattering from this complex object. Since an EM wave cannot penetrate the perfect conductor boundary S_1 and S_2 and enters regions I and II, the computational domain is exterior to regions I and II. From Section 1.2 we know that in such a domain, the EM fields satisfy:

$$\nabla \times \left(\frac{1}{\mu_r}\nabla \times \mathbf{E}\right) - \varepsilon_r k_0^2 \mathbf{E} = 0 \tag{3.82}$$

$$\hat{n} \times \mathbf{E} = 0, \qquad \text{on } S_1 \tag{3.83}$$

$$\hat{n} \times (\nabla \times \mathbf{E}) = 0, \qquad \text{on } S_2 \tag{3.84}$$

Obviously, the above domain is infinite. This leads to an infinite number of interpolation functions after discretizaion and hence it is impossible to solve this

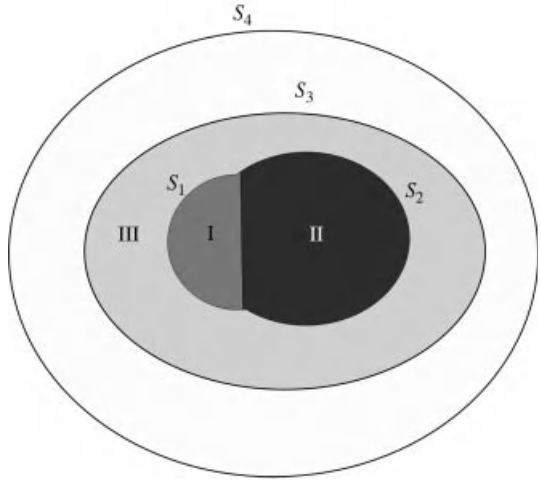


Figure 3.9 The FEM model of a complex scatterer.

system of equations. Therefore, the computational domain must be truncated. There are many means of truncation, and they can be approximately grouped into two classes. The first one is the whole-domain truncation techniques, that is, we use rigorous equations, such as the integral equations to truncate the unbounded domain. This kind of truncation bears the merit of being accurate. However, the sparseness of the FEM matrix is broken, which is to be further discussed in Chapter 5. The second one is the local-domain truncation techniques. The *perfectly matched layer* (PML) is one candidate. The common point of this type of truncation is that the sparseness of the FEM matrix is preserved, though we can only get an approximate solution. Here, we will introduce another simple local-domain truncation technique.

Let us use a fictitious surface of a sphere S_4 to truncate the unlimited space. In other words, we replace the unlimited space with the limited domain enclosed by surfaces S_1, S_2 and S_4. The next problem is to define the corresponding third-type boundary condition. A simple solution is to transfer the radiation condition in the infinity to S_4, that is,

$$\hat{r} \times \left(\nabla \times \mathbf{E}^S\right) \approx -jk_0 \hat{r} \times (\hat{r} \times \mathbf{E}^S) \tag{3.85}$$

Obviously, (3.85) holds only in an approximate sense. In order to increase the accuracy, the following is proposed in reference [24]:

$$\begin{aligned} \hat{r} \times \left(\nabla \times \mathbf{E}^S\right) \approx & -jk_0 \hat{r} \times (\hat{r} \times \mathbf{E}^S) + \beta(r)\nabla \times [\hat{r}(\nabla \times \mathbf{E}^S)_r] \\ & + \beta(r)\nabla_t(\nabla \cdot \mathbf{E}_t^S) \end{aligned} \tag{3.86}$$

where $\mathbf{E}^s$ denotes the scattered electric field, $\beta(r) = r/2(1 + jk_0 r)$ (r is the radius of the sphere), the subscript t labels the component transverse to $\hat{r}$, and the subscript r labels the tangential component. More to the point, it holds that $\mathbf{E}_t^s = -\hat{r} \times (\hat{r} \times \mathbf{E}^s)$ and $\mathbf{E}_r^s = \hat{r} \cdot \mathbf{E}^s$. As is pointed out in reference [24], (3.85) is derived after applying the power-series expansion to the scattering field, and taking the first-order approximation of the expansion. The second-order approximation leads to (3.86). Therefore, (3.85) is termed the first-order absorbing boundary condition (ABC), and (3.86) is the second-order absorbing boundary condition. Apparently, the greater r is, the more accurate (3.85) and (3.86) will be, and the more unknowns and more computation will be involved. In the practical application, S_4 is normally put 0.3λ away from the scatterer. For convenience, we write (3.85) and (3.86) uniformly as

$$\hat{r} \times \left(\nabla \times \mathbf{E}^S\right) + P(\mathbf{E}^S) = 0 \tag{3.87}$$

where P stands for the vector operator. To be specific, for the first-order ABC (3.85), P is defined as

$$P(\mathbf{E}^S) = jk_0 \hat{r} \times (\hat{r} \times \mathbf{E}^S) \tag{3.88}$$

For the second-order ABC (3.86), P is defined as

$$P(\mathbf{E}^S) = jk_0\hat{r} \times (\hat{r} \times \mathbf{E}^S) - \beta(r)\nabla \times [\hat{r}(\nabla \times \mathbf{E}^S)_r] - \beta(r)\nabla_t(\nabla \cdot \mathbf{E}_t^S) \tag{3.89}$$

Apparently, these conditions apply only to the scattered field. In order to obtain the boundary condition for the total field, we substitute $\mathbf{E}^s = \mathbf{E} - \mathbf{E}^i$ into (3.87) and obtain

$$\hat{r} \times (\nabla \times \mathbf{E}) + P(\mathbf{E}) = \hat{r} \times \left(\nabla \times \mathbf{E}^i\right) + P(\mathbf{E}^i) \tag{3.90}$$

Equation 3.90 can be further written as

$$\hat{r} \times (\nabla \times \mathbf{E}) + P(\mathbf{E}) = \mathbf{U}^i \tag{3.91}$$

where

$$\mathbf{U}^i = \hat{r} \times \left(\nabla \times \mathbf{E}^i\right) + P(\mathbf{E}^i) \tag{3.92}$$

where $\mathbf{E}^i$ denotes the incident electric field. For a plane wave, the expression is given by (2.1). Now, the partial differential equation describing the field in the 3D scattering problem is obtained. Compared with the PDE for the discontinuous waveguide problem, we can find the only difference lies in the boundary condition on boundary S_4. Using an analogous approach as in Section 3.1.1, we arrive at the variation of the functional (3.82):

$$\begin{cases} \delta F(\mathbf{E}) = 0 \\ \hat{n} \times \mathbf{E} = 0, \qquad \text{on } S_1 \end{cases} \tag{3.93}$$

Employing the first-order ABC, the functional is:

$$\begin{aligned} F(\mathbf{E}) = &\frac{1}{2}\iiint_V \left[\frac{1}{\mu_r}(\nabla \times \mathbf{E}) \cdot (\nabla \times \mathbf{E}) - k_0^2\varepsilon_r\mathbf{E} \cdot \mathbf{E}\right] dV \\ &+ \frac{1}{2}jk_0\iint_{S_4} \mathbf{E}_t \cdot \mathbf{E}_t dS + \iint_{S_4} \mathbf{U}^i \cdot \mathbf{E} dS \end{aligned} \tag{3.94}$$

Employing the second-order ABC, the functional is:

$$\begin{aligned} F(\mathbf{E}) = &\frac{1}{2}\iiint_V \left[\frac{1}{\mu_r}(\nabla \times \mathbf{E}) \cdot (\nabla \times \mathbf{E}) - k_0^2\varepsilon_r\mathbf{E} \times \mathbf{E}\right] dV \\ &+ \frac{1}{2}\iint_{S_4} [jk_0\mathbf{E}_t \cdot \mathbf{E}_t + \beta(\nabla \times \mathbf{E})_r^2 - \beta(\nabla \cdot \mathbf{E}_t)^2 dS \\ &+ \iint \mathbf{E} \times \mathbf{U}^i dS \end{aligned} \tag{3.95}$$

Once we obtain the functional, we can perform discretization as we do in the discontinuous waveguide problem and solve for the solution. Here, we would like to address more on the computation of the scattered field in the far zone. Since the functional (3.94) or (3.95) comprises only an unknown electrical field, we can obtain only the electrical-field distribution in the final result. From the Huygens' principle introduced in Section 1.3.1, we need to calculate the magnetic field in order to obtain the scattered field in the far zone. However, here we can express the magnetic field by directly applying Faraday's law of induction $\nabla \times \mathbf{E} = -j\omega\mu\mathbf{H}$. Then, applying the approximation of the far-field in Section 2.1.9 yields

$$\mathbf{E}^S = -\frac{e^{-jk_0 r}}{4\pi r}\iint_S \{\hat{r}\times[\hat{r}\times(\hat{n}\times\nabla\times\mathbf{E})] - jk_0(\hat{n}\times\mathbf{E}\times\hat{r})\}e^{jk_0\hat{r}\cdot\mathbf{r}'}dS' \qquad (3.96)$$

3.3.2 *Writing Computer Program*

In the attached code, we present a FEM program for computing the EM scattering problems to illustrate the computer programming. In this program, the total-field formulation and the first-order absorbing boundary conditions (ABCs) are applied. After reading the geometric information of the mesh files, the coefficient matrix can be generated by assembling the individual element matrices. The right-hand side vector of this equation is calculated by involving the incident wave. Then, the unknowns can be solved by using a matrix equations solver such as MUMPS in the attached codes. In the final step, the electric field on a selected closed surface is then used to compute the radar cross section (RCS).

Figure 3.10 demonstrates the hierarchical chart for computing scattering by a PEC object. Level-1 contains four steps to solve this problem: reading the geometric

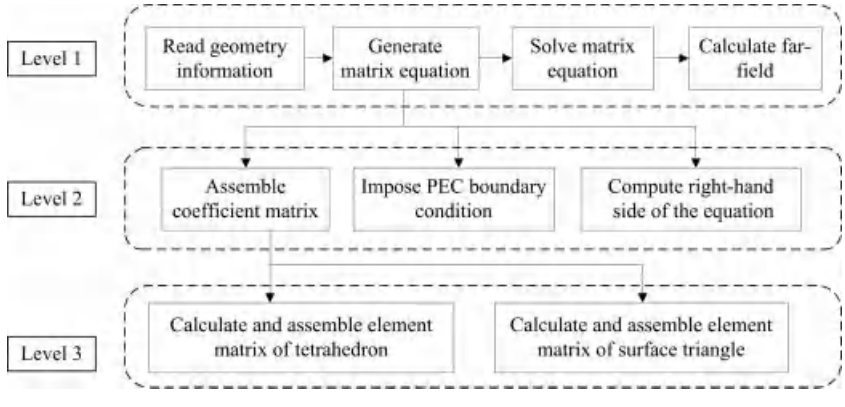
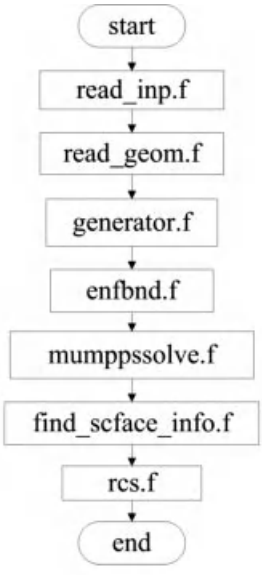


Figure 3.10 The hierarchical chart for the FEM computer program.

information of the mesh; generation of the coefficient matrix and the right-hand side vector, solving the matrix equation, and the far-field calculation that is a post process. The first and last step is straightforward. The step for solving matrix equation is also simple since a lot of commercial or noncommercial software is available. However, the process to generate the matrix equation is complex, and it is also the essential part of FEM. Therefore, we use a Level-2 to illustrate the detailed procedures in this step. These procedures include: Assembling the coefficient matrix, PEC boundary condition imposition, and computing the right-hand side vector. More specifically, in order to assemble the coefficient matrix, we need to calculate the element matrices for all tetrahedra and assemble them, and calculate the element matrices for all surface triangles and assemble them, as illustrated in Level-3. Two float charts for the main program and for one major subroutine, respectively, are presented in Figure 3.NaN. The function of the subroutines involved and the key parameters are detailed in the accompanying tables.

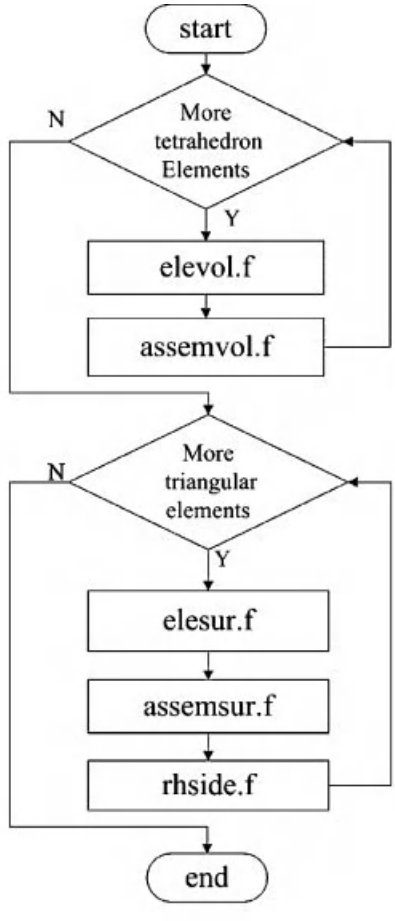


Notes for the figure

read_inp.f	Read control parameters & set the RCS receiver parameters. Read mesh index information.
read_geom.f	Read and convert geometry data.
generator.f	Generate coefficient matrix & right hand side.
enfbnd.f	Impose metal boundary condition.
mumppsolver.f	Solve the matrix equation.
find_scface_info.f	Find the scattering face geometric information.

(a)

Figure 3.11 The float charts of the computer program. (a) fem.f; (b) generator.f.



Notes for the figure

elevol.f	Compute element matrix (6 by 6) for tetrahedron.
assemvol.f	Assemble volume element (tetrahedron) matrix.
elesur.f	Compute element matrix (3 by 3) for triangle.
assemsur.f	Assemble surface element (triangle) matrix.
rhside.f	Compute right-hand side vector.

(b)

Figure 3.11 (*Continued*)

3.3.3 Numerical Results

Up to now, we have presented the FEM procedure in calculating the scattering from a 3D object. In terms of the procedure, the FEM solution in the 3D scattering is mainly the same as that in the discontinuous closed waveguide problem. The only difference stands in the truncation technique for the unbounded problem domain. If we apply PML in both cases, then even this difference disappears. However, in terms of the numerical result, the FEM in a closed waveguide is much better than the 3D scattering in open space. This is because the physical phenomena in the discontinuous closed waveguide are mainly the reflection and transmission caused by the discontinuity in the waveguide. Since the truncation surface, that is, the cut-plane of the waveguide, is relatively small, the computational domain is consequently of small size. On the other hand, the physical phenomenon in the 3D scattering problem is majorly scattering in the whole space. The truncation surface of the infinite space needs to enclose the 3D scatterer and be a certain distance away from it. Consequently, generally the computational domain is comparatively large. Compared with the MoM, the reader may find the FEM more general and convenient to use. Take the 3D scattering problem as an example. The MoM solution needs to treat the metallic, homogeneous dielectric and inhomogeneous dielectric targets differently. The skills applied, the difficulties encountered and the formulation used are different in each particular case. On the other hand, the FEM solves these problems in a similar approach with easier formulation. The advantage brought by the complexity in implementing MoM is better efficiency and accuracy

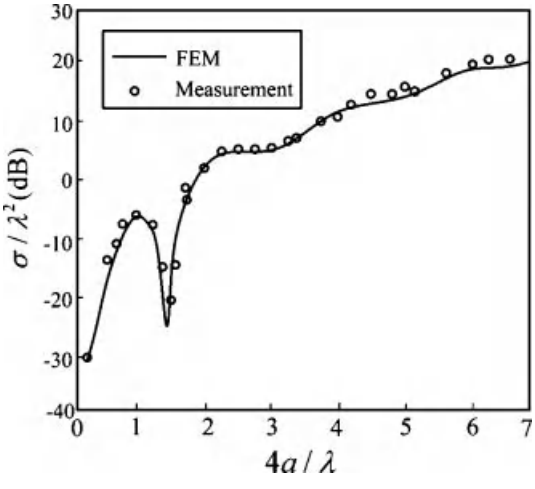


Figure 3.12 Backscatter RCS at normal incidence for a metallic cube of size a. (Reproduced with permission from A. Chatterjee, J.M. Jin and J.L. Volakis, "Edge-based finite elements and vector ABCs applied to 3D scattering," *IEEE Transactions on Antennas and Propagation*, **41**, 2, 221–226, 1993. © 1993 IEEE.)

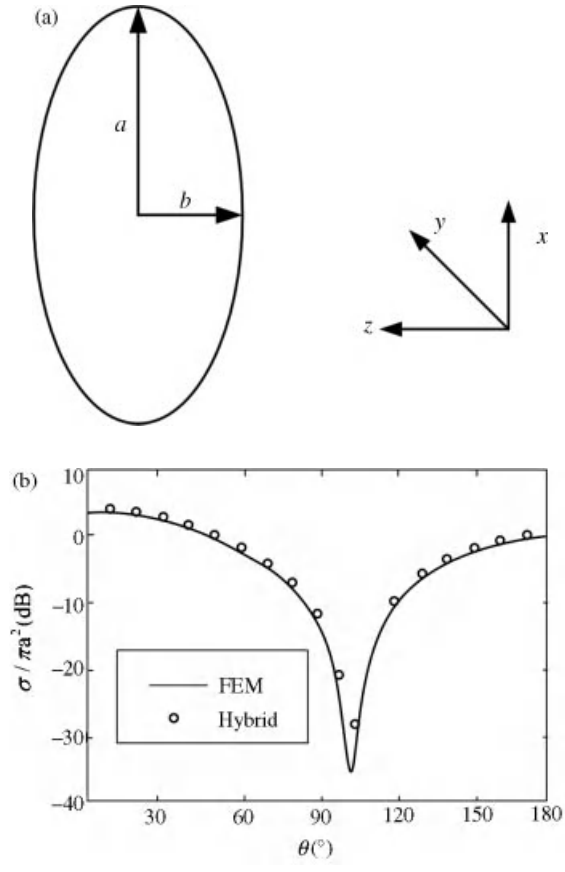


Figure 3.13 (a) The model of the homogeneous dielectric ellipsoid. (b) The bistatic RCS pattern of a lossy prolate spheroid ($k_0a = \pi/2$, $a/b = 2$, $\varepsilon_r = 4 - j$), where a and b are the major and minor axes of the spheroid respectively. (Reproduced with permission from A. Chatterjee, J.M. Jin and J.L. Volakis, "Edge-based finite elements and vector ABCs applied to 3D scattering," *IEEE Transactions on Antennas and Propagation*, **41**, 2, 221–226, 1993. © 1993 IEEE.)

in the result, especially in treating the metallic body. More comparison of these two methods, together with the comparison with the FDTD method will be fully discussed in Chapter 4. Next, we will illustrate the effectiveness of the FEM by two numerical examples. Figure 3.12 plots the comparison of the monostatic RCS calculated by FEM with the measured result of a metallic cube under incidences of different frequencies. Figure 3.13 plots the comparison of the FEM and the MoM solution of the bistatic RCS for a homogeneous lossy dielectric prolate spheroid ($k_0a = \pi/2$, $a/b = 2$, $\varepsilon_r = 4 - j$). In both simulations, the second-order absorbing boundary condition is used and the distance between the boundary and the scatterer is 0.15λ [25]. It can be seen that the FEM is validated by these results.

3.4 Node-Edge Element

As is stated in Section 3.1.2, the node element does not suit for discretizing the whole electric (or magnetic) vector functional. CEM researchers found the edge-element to be a proper element to replace the node element. However, with the development of FEM, people realized that the matrix generated by the edge-element sometimes suffers from a poor condition number. In particular, in self-adaptive FEM, due to the dense meshes in certain areas, the condition number is so bad that the accuracy and the efficiency of FEM are greatly affected.

An early solution addressing this problem is the tree-cotree splitting technique proposed in reference [26]. Based on [26], combining the node element and edge element, the *node-edge element* is presented to provide a simple and easy-to-utilize form of element that is free of spurious solutions and shows good matrix property [27].

3.4.1 Construction of Node-Edge Element

It has been shown that the condition number of the edge-element matrix becomes poor under the following two circumstances: (a) the working frequency is low and the electrical size of the target is small; (b) the target is complex in structure, which requires highly dense meshing to accurately model the distribution of the field. These two cases can be regarded equivalently as the same situation, that is, $k_0^2 = (2\pi/\lambda)^2 \to 0$. Under this condition, a solution of (3.9) combined with an arbitrary gradient field still satisfies (3.9). In other words, the solution to (3.9) in this case is not unique. Generally speaking, there are two types of eigenmodes in the functional (3.9): one corresponds to the nonphysical gradient field, and the other corresponds to the physical rotational field. Because we cannot separate the gradient field with the rotational one by using the edge-element, these two kinds of fields affect each other, resulting in poor matrix behavior. To address this problem, a practically efficient element, namely *the node-edge element* is proposed. According to the Helmholtz decomposition theory, we decompose the electric field into

$$\mathbf{E}(\mathbf{r}) = \mathbf{E}_g(\mathbf{r}) + \mathbf{E}_r(\mathbf{r}) \tag{3.97}$$

where $\mathbf{E}_g$ denotes the gradient field, and $\mathbf{E}_r$ denotes the rotational field, satisfying $\nabla \times \mathbf{E}_g = 0$ and $\nabla \cdot \mathbf{E}_r = 0$, respectively. The construction of the basis function for the node-edge element is based on this decomposition: the node element is used to model the gradient field $\mathbf{E}_g$ and the edge-element is applied to model the rotational field $\mathbf{E}_r$. To be specific, for the tetrahedral element, the node-edge element can be expressed as

$$\mathbf{E} = \sum_{i=1}^{4} \nabla L_i E_{gi} + \sum_{i=1}^{6} \mathbf{N}_i E_{ri} \tag{3.98}$$

where L_i denotes the volume coordinate on node i, and $\mathbf{N}_j$ denotes the basis function of the edge-element j. It is worth noting that the node element in (3.98) is significantly in contrast to the conventional node elements. First, the basis function in the node element in (3.98) is the gradient of the volume coordinate, but not the volume coordinate itself. Second, the interpolation function in the node element in (3.98) normally does not stand for the gradient of the node. Because ∇L_i is the linear combination of the edge-element basis function, the interpolation function E_{ri} still denotes the tangential component on the edge-element. Therefore, the continuity of the tangential field between adjacent elements is preserved. Moreover, the node-edge element constructed in this way is not independent. Consequently we need to eliminate the redundant interpolation functions to assure the independency of this type of basis function. There are a number of ways to do this. We will give one example of the node-edge element implementation.

Problem

3.9 Find the relation between E_{gi} and E_{ri} in (3.98) and E_{ti} in (3.67).

3.4.2 Implementation of Node-Edge Element

From the above analysis, the key issue in utilizing the node-edge element is to eliminate the redundant interpolation functions. Let us consider a 2D example. Generally speaking, the dimension of the null-spaces of the curl operator in (3.94) or (3.95) is normally the number of the nodes in the entire computation domain, denoted as N_v. So we choose all the nodes as independent interpolation functions. Usually, the number of interpolation variables should be the same as that in the edge-element method. Hence, we need only to keep $N–N_v$ interpolation variables on edges and eliminate the rest. There are various schemes to perform the elimination. Here, we introduce a scheme based on a traversal scheme in the graph theory. Construct a minimum traversal pass connecting all nodes in the meshes, and remove this pass from the original meshes. Then, the remaining edges and nodes form the elements on which the basis function of the node-edge element is defined. The traversal pass should satisfy: (1) this minimum pass is simply connected (or 1-connected); (2) this pass connects every two nodes in the meshes. Figure 3.14 shows an example of a 2D FEM meshes. From the traversal algorithm in the graph theory, the minimum traversal pass is found as shown in Figure 3.15. By removing the edges on this pass, we obtained the node-edge elements formed by all the nodes and the residual edges, as plotted in Figure 3.16.

We would like to mention two more issues here about the implementation of the node-edge element method. First, if the target comprises a conductive part, then all the nodes on the conductor are regarded as one node in the processing. For example, in Figure 3.15, if the boundary of the cross section of the waveguide is not PEC, then the dashed lines should also be included in the minimum traversal pass. Second, the

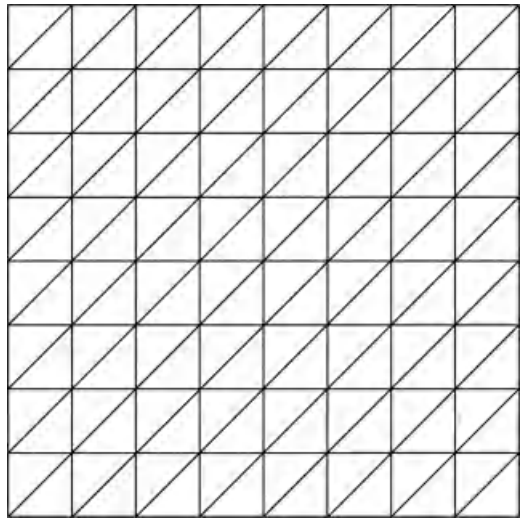


Figure 3.14 An example of 2D FEM meshes.

number of the edges on the minimum traversal pass equals the number of the nodes subtracted by one. Therefore, the increased number of the nodes is compensated by the reduced number of edges. Consequently, the dimension of the node-edge element matrix is equivalent to the edge-element matrix.

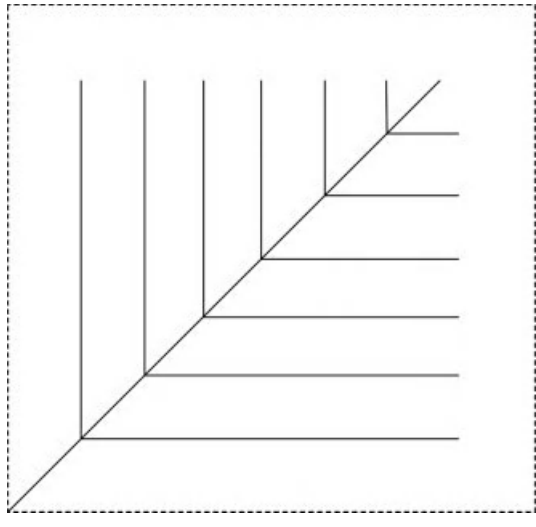


Figure 3.15 The minimum traversal pass connecting every two nodes in the meshes.

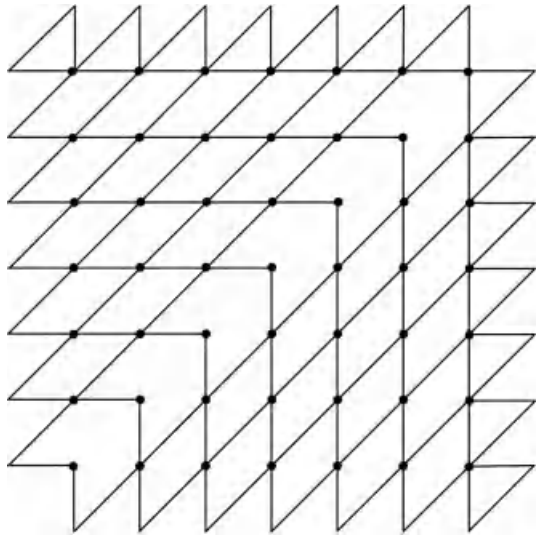


Figure 3.16 The node-edge elements.

3.4.3 Numerical Examples

In order to demonstrate the improved matrix behavior of the node-edge element, the field distribution in a 3D cavity excited by a point source is calculated by using FEM based on the node-edge element and the edge element methods, respectively. A CG iteration method is used to solve the matrix equation in both cases. The number of iterations required by these two schemes as functions of the mesh density are plotted in Figure 3.17. It can be seen from the figure that the iteration number required by the

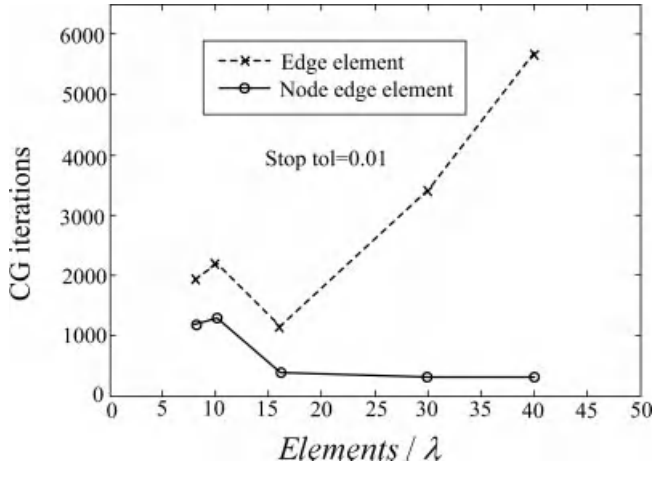


Figure 3.17 The numbers of iteration required as functions of the mesh density.

node-edge element scheme is lower than that in the edge-element scheme. Especially with the increase of the mesh density, the iteration number in the node-edge element method does not increase linearly as it does in the edge-element case.

3.5 Higher-Order Element

The error introduced by discretizing the continuous equation is termed the *truncation error*. This is unavoidable in any numerical method. We can only manage to reduce this error. Except for this, FEM also suffers *dispersion error*. To be specific, after discretizing the functional, the numerical phase velocity given by the discretized equation is different from the physical one when the EM wave propagates in the medium. This error is also termed *numerical dispersion*. This difference affects the accuracy of the final field solution in an accumulative way with the increase of the computational domain. An efficient solution to reduce these two kinds of error is the adoption of higher-order elements.

Beside the above errors, in using FEM we need to pay attention to another error, namely the *round-off error*. This error is not caused by discretization. Instead, it is introduced while using a computer to solve the discretized equation. This error is strongly related to the condition number of the matrix as well as the choice of the solver. As a result, when we construct higher-order element to reduce the truncation error and the dispersion error, we need to keep an eye on the condition number of the resultant matrix.

The next issue is how to construct higher-order elements suitable for the functional (3.9). There are two ways of constructing higher-order vector elements: *interpolatory basis function* scheme and *hierarchical basis functions* scheme. As is indicated by the name, the field in the interpolation higher-order element is expressed by interpolating the fields on a set of points on the element. In the hierarchical higher-order element, the field is expressed by adding new functions (and the corresponding unknown variables) to the lower-order basis functions. Therefore, the coefficients in the interpolation scheme correspond to the components of the physical field at the interpolation points, while those in the hierarchical scheme don't have an obvious physical interpretation. Generally speaking, the condition number of the matrix from the interpolation scheme is better than that from the hierarchical scheme. But the disadvantage of the interpolation scheme is that the higher-order element is not compatible with the lower-order one due to their completely different forms in the interpolatory basis functions. Fortunately, this incompatibility can be dealt with by employing a simple boundary-matching treatment. On the other hand, the poor condition number problem in the latter scheme can be solved by orthogonalizing the basis functions. In the following, we will introduce the construction of the higher-order vector element for (3.9) [28]. The reader is referred to references [29,30] for more on hierarchical higher-order elements and the *basis function orthogonalization process*.

In order to make the construction of the higher-order vector element easier to understand, we first introduce two kinds of interpolation polynomials. The first one is conveniently written using the auxiliary interpolatory polynomials of Silvester:

$$R_i(p,L) = \begin{cases} \dfrac{1}{i!}\displaystyle\prod_{k=0}^{i-1}(pL-k), & 1 \le i \le p \\ 1, & i = 0 \end{cases} \tag{3.99}$$

These functions are polynomials of degree i in L, where L is the interpolation point in the interval $[0,1]$. The parameter p indicates the number of uniform subintervals into which the interval is divided, and $R_i(p,L)$ equals unity at $L = i/p$ and equals zero at $L = 0, 1/p, 2/p, \cdots, (i-1)/p$. By applying Silvester polynomials, in the nature coordinate system, we can easily construct the scalar Lagrangian interpolation on the canonical elements. Take a 2D problem as an example, the following polynomial

$$\alpha_{ijk}(L_1,L_2,L_3) = R_i(p,L_1)R_j(p,L_2)R_k(p,L_3) \tag{3.100}$$

with $i+j+k = p$, is the pth-order Lagrange polynomial interpolating points within a triangle whose normalized area coordinates (L_1,L_2,L_3) are $(i/p, j/p, k/p)$.

The second one is the interpolation polynomial convenient for constructing the scalar higher-order basis functions. It is referred to in the following as the modified or shifted Silvester polynomial, defined as:

$$\begin{aligned} \hat{R}_i(p,L) &= R_{i-1}(p,L-1/p) \\ &= \begin{cases} \dfrac{1}{(i-1)!}\displaystyle\prod_{k=0}^{i-1}(pL-k), & 2 \le i \le p \\ 1, & i = 1 \end{cases} \end{aligned} \tag{3.101}$$

From the definition, it can be seen that the ith-order shifted Silvester polynomial $\hat{R}_i(p,L)$ is the (i–1)th-order Silvester polynomial when the interpolation points are left shifted by one.

With the above two polynomials, we can construct the higher-order vector element. Before we do that, let us observe the following identities:

$$\begin{aligned} (p+2)L_1^{\alpha}L_2^{\beta}\nabla L_1 &= \nabla\left(L_1^{\alpha+1}L_2^{\beta}\right) + \beta L_1^{\alpha}L_2^{\beta-1}\mathbf{N}_3 \\ (p+2)L_1^{\alpha}L_2^{\beta}\nabla L_2 &= \nabla\left(L_1^{\alpha}L_2^{\beta+1}\right) - \alpha L_1^{\alpha-1}L_2^{\beta}\mathbf{N}_3 \\ p &= \alpha+\beta-1 \ge 0, \qquad \alpha,\beta \ge 0 \end{aligned} \tag{3.102}$$

It can be seen that an arbitrary $(p+1)$th-order vector polynomial can be divided into two parts: the gradient of the a $(p+2)$th-order vector polynomial, and the product of the pth-order vector polynomial and the basis function of the edge-element. As we know, there are two parts in functional (3.9). One of them is the curl of the fields. The accuracy in calculating this part is significantly lower than the other entry. Moreover, the curl of any gradient fields is zero. Consequently, the gradient field in a complete $(p+1)$th-order vector polynomial contributes little to the accuracy of the computation results. In other words, in order to achieve pth-order accuracy after discretization, we don't need to construct a complete $(p+1)$th-order vector polynomial. Instead, we only need to use the product of a complete pth-order scalar polynomial and the basis functions.

For a triangular element, the basis functions of the pth-order vector basis functions are

$$\begin{aligned} \mathbf{N}^1_{ijk} &= C^1_{ijk} R_i(p+2, L_1)\hat{R}_j(p+2, L_2)\hat{R}_k(p+2, L_3)\mathbf{N}_1 \\ i &= 0, 1, \cdots p \qquad j,k = 1, 2, \cdots, p+1 \end{aligned} \tag{3.103}$$

$$\begin{aligned} \mathbf{N}^2_{ijk} &= C^2_{ijk} \hat{R}_i(p+2, L_1)R_j(p+2, L_2)\hat{R}_k(p+2, L_3)\mathbf{N}_2 \\ j &= 0, 1, \cdots p \qquad i,k = 1, 2, \cdots, p+1 \end{aligned} \tag{3.104}$$

$$\begin{aligned} \mathbf{N}^3_{ijk} &= C^3_{ijk} \hat{R}_i(p+2, L_1)\hat{R}_j(p+2, L_2)R_k(p+2, L_3)\mathbf{N}_3 \\ k &= 0, 1, \cdots p \qquad i,j = 1, 2, \cdots, p+1 \end{aligned} \tag{3.105}$$

where $i+j+k = p+2$, $\mathbf{N}_1$, $\mathbf{N}_2$ and $\mathbf{N}_3$ are the basis functions before normalization

$$\begin{aligned} \mathbf{N}_1 &= L_2 \nabla L_3 - L_3 \nabla L_2 \\ \mathbf{N}_2 &= L_3 \nabla L_1 - L_1 \nabla L_3 \\ \mathbf{N}_3 &= L_1 \nabla L_2 - L_2 \nabla L_1 \end{aligned} \tag{3.106}$$

Figure 3.18 illustrates the higher-order vector element constructed on triangular elements. The multiplying polynomials ahead of $\mathbf{N}_1$ in (3.103) corresponds to the pth-order complete interpolation polynomial constructed by using Pascal's triangle interpolation points under the Silvester interpolation and the shifted Silvester interpolation schemes. The reason to shift the nodes along edges "2" and "3" to the triangle's interior is as follows: $\mathbf{N}_1$ is normal to edges "2" and "3." If the shift is not applied and the interpolation points fall on the edges, the components normal to edges "2" and "3" are taken as the interpolation functions. Then, the continuity on the normal components is implied, which is a violation of the physical boundary condition on the interface. Equations 3.104 and 3.105 can be interpreted analogously. The normalization factor C^{β}_{ijk} in (3.103)–(3.105) is introduced in order to assure that the component of

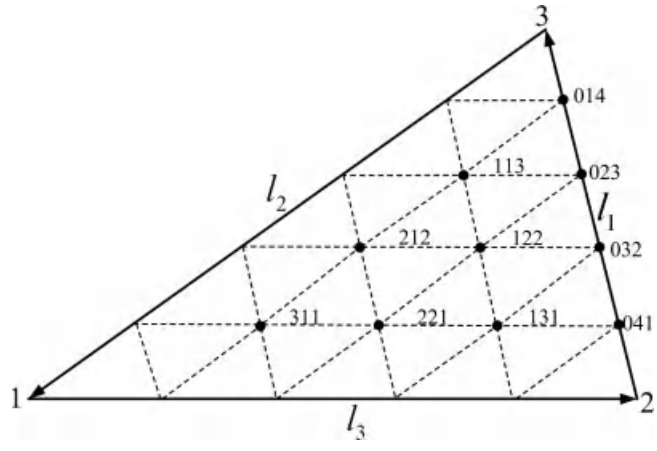


Figure 3.18 The higher-order vector element constructed on triangular elements.

$\mathbf{N}_{ijk}^{\beta}$ along $\mathbf{l}_\beta$ on the interpolation point takes the value of unity.

$$C_{ijk}^{\beta} = \frac{p+2}{p+2-i_\beta} l_\beta^{ijk} \tag{3.107}$$

with i_β taken to be i, j, k, for $\beta = 1, 2, 3$, respectively; l_β^{ijk} is the value of $l_\beta = |\mathbf{l}_\beta|$ at the interpolation point $\left(\frac{i}{p+2}, \frac{j}{p+2}, \frac{k}{p+2}\right)$, and $\mathbf{l}_\beta$ is defined as

$$\mathbf{l}_\beta = \frac{\partial \mathbf{r}}{\partial \mathbf{L}_\beta} \tag{3.108}$$

In many cases, when we use higher-order FEM to model the distribution of the fields, we need to employ isoparametric conformal elements to efficiently model the computational domain. The coordinate transformation for the isoparametric conformal elements is:

$$\mathbf{r} = \sum_{i+j+k=p} \alpha_{ijk}(L_1, L_2, L_3)\mathbf{r}_i \tag{3.109}$$

Thus, the scheme employing general higher-order isoparametric conformal vector elements is, to first determine C_{ijk}^{β} from (3.107)–(3.109), and then obtain the higher-order vector basis function $\mathbf{N}_{ijk}^{\beta}$ through (3.103)–(3.105).

Based on the above analysis, the number of interpolation parameters of the pth-order vector triangular elements discretizing (3.9) can be calculated in the following way: the number of interpolation parameters of the tangential components on each edge is $p + 1$, which results in $3(p + 1)$ parameters on the three edges. Inside the triangle there are $p(p + 1)/2$ interpolation points, and each corresponds to two parameters according to

the two directions. So there are $p(p+1)$ interior parameters. Therefore, the total number of the interpolation parameters of the pth-order vector parameters is $(p+1)(p+3)$.

In the same way we can calculate the number of interpolation functions of the pth-order vector tetrahedral elements that discretizes (3.9): the number of interpolation parameters of the tangential components on each edge is $p+1$, resulting in $6(p+1)$ parameters on the six edges. There are $4p(p+1)$ interpolation points on the four facets and each corresponds to two parameters according to the two directions. Inside the tetrahedron, there are $p(p^2-1)/6$ points and each point corresponds to three parameters according to the three directions, leading to $p(p^2-1)/2$ interior parameters. So there are altogether $(p+1)(p+3)(p+4)/2$ interpolation parameters. In a similar approach, the basis function in the pth-order vector tetrahedral elements is expressed as:

$$\begin{aligned} \mathbf{N}^1_{ijkl} &= C^1_{ijkl}\hat{R}_i(p+2,L_1)\hat{R}_j(p+2,L_2)R_k(p+2,L_3)R_l(p+2,L_4)\mathbf{N}_1 \\ i,j &= 1,2,\cdots p+1 \qquad k,l = 0,1,\cdots,p \end{aligned} \tag{3.110}$$

where

$$\mathbf{N}_1 = L_2\nabla L_3 - L_3\nabla L_2 \tag{3.111}$$

$$C^1_{ijkl} = \frac{p+2}{p+2-i-j} l_1^{(ijkl)} \tag{3.112}$$

and $l_1^{(ijkl)}$ is the value of $l_1 = |\mathbf{l}_1|$ at the interpolation point $\left(\frac{i}{p+2}, \frac{j}{p+2}, \frac{k}{p+2}, \frac{l}{p+2}\right)$. The remaining $\mathbf{N}^\beta_{ijkl}(\beta = 2,3,\cdots 6)$ can be obtained analogously.

Problems

3.10 Read references [29–31], and derive the hierarchical higher-order basis functions.

3.11 Write a program to generate the element matrix of (3.80) using interpolatory and hierarchical higher-order basis functions respectively, and compare their condition numbers.

3.6 Finite-Element Time-Domain Method

In the history of FEM, the *finite-element time-domain method* (FETD) is far less popular than FEM in the frequency domain [32]. This is mainly caused by the following two facts. First, the conventional FETD is quite expensive because of the inverse of large mass matrices in every step of the time-marching scheme. Second, among time-domain (TD) numerical methods, the FDTD method, which will be introduced in the

next chapter, is dominant due to its simplicity, generality and high efficiency. However, the FDTD method, which normally employs a grid mesh, is obviously less flexible than FETD for complex geometric modeling. In particular, FDTD is not compatible with commercial meshing software such as Hypermesh, Ansys. On the other hand, time-domain numerical methods become more and more important since many practical EM problems are going to involve wide band. Furthermore, from the point of view of exploring EM phenomenon, time-domain numerical methods are also indispensable due to the time-domain nature of EM waves. Therefore, FETD is always an important topic in CEM. Many novel formulations of FETD have been developed in recent years [33,34]. One of efficient formulations is the so-called *discontinuous Galerkin finite-element time-domain method* (DG-FETD) [34], which will be introduced in the following. The coupled first-order Maxwell's equations in a lossless region read:

$$\varepsilon \frac{\partial \mathbf{E}}{\partial t} - \nabla \times \mathbf{H} = -\mathbf{J} \tag{3.113}$$

$$\mu \frac{\partial \mathbf{H}}{\partial t} + \nabla \times \mathbf{E} = -\mathbf{M}_0 \tag{3.114}$$

where $\mathbf{E}$ and $\mathbf{H}$ are the electric and magnetic fields; $\mathbf{J}$ and $\mathbf{M}_0$ is applied electric current densities, ε, and μ denote the permittivity and permeability of the material, respectively.

Denote $\mathbf{N}_e$ and $\mathbf{N}_h$ as basis functions for $\mathbf{E}$ and $\mathbf{H}$, respectively. The Galerkin's weighted residual method is applied to minimize the residual of (3.113) and (3.114) in a subdomain:

$$\int_V \mathbf{N}_e \cdot \left(\varepsilon \frac{\partial \mathbf{E}}{\partial t} - \nabla \times \mathbf{H} + \mathbf{J} \right) dV = 0 \tag{3.115}$$

$$\int_V \mathbf{N}_h \cdot \left(\mu \frac{\partial \mathbf{H}}{\partial t} + \nabla \times \mathbf{E} + \mathbf{M} \right) dV = 0 \tag{3.116}$$

where V denotes the volume of the subdomain, which contains one or more vector elements. Performing integration by parts for (3.115) and (3.116) leads to

$$\int_V \mathbf{N}_e \cdot \left(\varepsilon \frac{\partial \mathbf{E}}{\partial t} + \mathbf{J} \right) dV = \int_V \nabla \times \mathbf{N}_e \cdot \mathbf{H}\, dV + \int_S \mathbf{N}_e \cdot (\hat{n} \times \mathbf{H}) dS \tag{3.117}$$

$$\int_V \mathbf{N}_h \cdot \left(\mu \frac{\partial \mathbf{H}}{\partial t} + \mathbf{M} \right) dV = -\int_V \nabla \times \mathbf{N}_h \cdot \mathbf{E}\, dV - \int_S \mathbf{N}_h \cdot (\hat{n} \times \mathbf{E}) dS \tag{3.118}$$

where S denotes the subdomain surface and $\hat{n}$ is the unit outward normal vector. Since in the discontinuous Galerkin method, the basis functions between different subdomains can be discontinuous, values of $(\hat{n} \times \mathbf{E})$ and $(\hat{n} \times \mathbf{H})$ on an interface should be determined by fields from both sides of S. Take the interface of the ith and the

*j*th subdomains as an example. It is proved in [35,36] that the correct fields on the interface will be

$$\left(Y^{(i)}+Y^{(j)}\right)(\hat{n}\times\mathbf{E}) = \hat{n}\times\left(Y^{(i)}\mathbf{E}^{(i)}+Y^{(j)}\mathbf{E}^{(j)}\right)-\hat{n}\times\hat{n}\times\left(\mathbf{H}^{(i)}-\mathbf{H}^{(j)}\right) \quad (3.119)$$

$$\left(Z^{(i)}+Z^{(j)}\right)(\hat{n}\times\mathbf{H}) = \hat{n}\times\left(Z^{(i)}\mathbf{H}^{(i)}+Z^{(j)}\mathbf{H}^{(j)}\right)+\hat{n}\times\hat{n}\times\left(\mathbf{E}^{(i)}-\mathbf{E}^{(j)}\right) \quad (3.120)$$

where $Z^{(m)}$ and $Y^{(m)}\left(Z^{(m)}=1/Y^{(m)}=\sqrt{\mu^{(m)}/\varepsilon^{(m)}}, m=i,j\right)$ are wave impedances and admittances for the *i*th and the *j*th subdomains, respectively. Then, the discretized system of equations by the DG-FETD for the *i*th subdomain is

$$\mathbf{M}_{ee}^{(i)}\frac{\partial\mathbf{e}^{(i)}}{\partial t} = \mathbf{K}_{ee}^{(i)}\mathbf{e}^{(i)}+\mathbf{K}_{eh}^{(i)}\mathbf{h}^{(i)}+\sum_j\left(\mathbf{L}_{ee}^{(ij)}\mathbf{e}^{(j)}+\mathbf{L}_{eh}^{(ij)}\mathbf{h}^{(j)}\right)+\mathbf{j}^{(i)} \quad (3.121)$$

$$\mathbf{M}_{hh}^{(i)}\frac{\partial\mathbf{h}^{(i)}}{\partial t} = \mathbf{K}_{he}^{(i)}\mathbf{e}^{(i)}+\mathbf{K}_{hh}^{(i)}\mathbf{h}^{(i)}+\sum_j\left(\mathbf{L}_{he}^{(ij)}\mathbf{e}^{(j)}+\mathbf{L}_{hh}^{(ij)}\mathbf{h}^{(j)}\right) \quad (3.122)$$

where $\mathbf{e}^{(i)}$ and $\mathbf{h}^{(i)}$ are vectors of the discretized electric and magnetic fields, $\mathbf{M}_{ee}^{(i)}$ and $\mathbf{M}_{hh}^{(i)}$ are the mass matrices, $\mathbf{K}_{ee}^{(i)}$, $\mathbf{K}_{hh}^{(i)}$, $\mathbf{K}_{eh}^{(i)}$ and $\mathbf{K}_{he}^{(i)}$ are the stiffness matrices, $\mathbf{j}^{(i)}$ is the discretized excitation vector for the *i*th subdomain. $\mathbf{L}_{ee}^{(ij)}$, $\mathbf{L}_{eh}^{(ij)}$, $\mathbf{L}_{hh}^{(ij)}$, and $\mathbf{L}_{he}^{(ij)}$ can be viewed as the coupling between the fields of the local and the neighboring subdomains. Expressions for these matrices and vectors are:

$$\begin{aligned}
\left[\mathbf{M}_{ee}^{(i)}\right]_{mn} &= \int_{S_i}\varepsilon\mathbf{N}_{em}^{(i)}\cdot\mathbf{N}_{en}^{(i)}ds, & \left[\mathbf{M}_{hh}^{(i)}\right]_{mn} &= \int_{S_i}\mu\mathbf{N}_{hm}^{(i)}\cdot\mathbf{N}_{hn}^{(i)}ds,\\
\left[\mathbf{K}_{ee}^{(i)}\right]_{mn} &= \int_{L_i}\mathbf{N}_{em}^{(i)}\cdot\frac{\hat{n}\times\hat{n}\times\mathbf{N}_{en}^{(i)}}{Z^{(i)}+Z^{(j)}}dl, & \left[\mathbf{K}_{hh}^{(i)}\right]_{mn} &= \int_{L_i}\mathbf{N}_{hm}^{(i)}\cdot\frac{\hat{n}\times\hat{n}\times\mathbf{N}_{hn}^{(i)}}{Y^{(i)}+Y^{(j)}}dl,\\
\left[\mathbf{K}_{eh}^{(i)}\right]_{mn} &= \int_{S_i}\nabla\times\mathbf{N}_{em}^{(i)}\cdot\mathbf{N}_{hn}^{(i)}ds+\int_{L_i}\mathbf{N}_{em}^{(i)}\cdot\frac{\left\{\hat{n}\times\mathbf{N}_{hn}^{(i)}\right\}Z^{(i)}}{Z^{(i)}+Z^{(j)}}dl,\\
\left[\mathbf{K}_{he}^{(i)}\right]_{mn} &= -\int_{S_i}\nabla\times\mathbf{N}_{hm}^{(i)}\cdot\mathbf{N}_{en}^{(i)}ds-\int_{L_i}\mathbf{N}_{hm}^{(i)}\cdot\frac{\left\{\hat{n}\times\mathbf{N}_{en}^{(i)}\right\}Y^{(i)}}{Y^{(i)}+Y^{(j)}}dl,\\
\left[\mathbf{L}_{ee}^{(ij)}\right]_{mn} &= -\int_{L_i}\mathbf{N}_{em}^{(i)}\cdot\frac{\hat{n}\times\hat{n}\times\mathbf{N}_{en}^{(j)}}{Z^{(i)}+Z^{(j)}}dl, & \left[\mathbf{L}_{hh}^{(ij)}\right]_{mn} &= -\int_{L_i}\mathbf{N}_{hm}^{(i)}\cdot\frac{\hat{n}\times\hat{n}\times\mathbf{N}_{hn}^{(j)}}{Y^{(i)}+Y^{(j)}}dl,\\
\left[\mathbf{L}_{eh}^{(ij)}\right]_{mn} &= \int_{L_i}\mathbf{N}_{em}^{(i)}\cdot\frac{\left\{\hat{n}\times\mathbf{N}_{hn}^{(j)}\right\}Z^{(j)}}{Z^{(i)}+Z^{(j)}}dl, & \left[\mathbf{L}_{he}^{(ij)}\right]_{mn} &= -\int_{L_i}\mathbf{N}_{hm}^{(i)}\cdot\frac{\left\{\hat{n}\times\mathbf{N}_{en}^{(j)}\right\}Y^{(j)}}{Y^{(i)}+Y^{(j)}}dl,\\
\left[\mathbf{J}^{(i)}\right]_m &= -\int_{S_i}\mathbf{N}_{em}^{(i)}\cdot\mathbf{J}ds
\end{aligned} \quad (3.123)$$

It is worth noting that the derivative of fields with time in the discretized system (3.121) and (3.122) is independent for the *i*th subdomain, thus the updating fields with time can be completed independently in each subdomain. Therefore, the inverse of large mass matrix in the conventional FETD can be avoided in this DG-FETD. The reason for avoiding the inverse of a large matrix is that the boundary conditions (3.119) and (3.120) used in DG-FETD do not involve the derivative of fields with time, which is essentially in contrast to the conventional FETD.

The time-marching scheme corresponding to (3.121) and (3.122) can be obtained by using the fourth-order Runge–Kutta (RK4) method [34]. To be more specific, the system in (3.121) and (3.122) can be written as

$$\frac{d\mathbf{y}}{dt} = \mathbf{A}\mathbf{y} + \mathbf{s} \tag{3.124}$$

where $\mathbf{A}$ is a matrix and $\mathbf{s}$ is an excitation. Then, by the RK4 scheme,

$$\mathbf{y}^{n+1} = \mathbf{y}^n + \frac{\Delta t}{6}(\mathbf{k}_1 + 2\mathbf{k}_2 + 2\mathbf{k}_3 + \mathbf{k}_4) \tag{3.125}$$

where

$$\mathbf{k}_1 = \mathbf{A}\mathbf{y}^n + \mathbf{s}^n, \qquad \mathbf{k}_2 = \mathbf{A}\left(\mathbf{y}^n + \frac{\Delta t}{2}\mathbf{k}_1\right) + \mathbf{s}^{n+\frac{1}{2}},$$

$$\mathbf{k}_3 = \mathbf{A}\left(\mathbf{y}^n + \frac{\Delta t}{2}\mathbf{k}_2\right) + \mathbf{s}^{n+\frac{1}{2}}, \qquad \mathbf{k}_4 = \mathbf{A}(\mathbf{y}^n + \Delta t\mathbf{k}_3) + \mathbf{s}^{n+1} \tag{3.126}$$

It is suggested in [34] that the stability criterion for the DG-FETD method with the RK4 time stepping scheme is

$$\Delta t \leq \frac{\sqrt{3}}{4}\frac{h_{\min}}{c} \tag{3.127}$$

where $h_{\min}$ denotes the shortest edge among the DG-FETD meshes of all subdomains and c denotes the speed of light.

It is worth noting here that although the edge elements, which can suppress the spurious modes when used in solving the second-order vector wave equations, cannot guarantee the nonspurious solution in the first-order Maxwell's equations. Based on numerical experiments [37], it is found that the choice of different order basis functions for $\mathbf{E}$ and $\mathbf{H}$ is necessary to guarantee a nonspurious vector FETD scheme.

3.7 More Comments on FEM

Similar to Section 2.4, in this section we mean to provide a more general and complete view about the utilization of FEM, in order to help the readers to deeply

understand the general essentials and the key points in FEM utilization introduced in the preceding sections. In contrast to the concrete case study in Section 2.4, in this section, issues including both the problem and the methodology are discussed in a more concise and general manner. These issues come from the authors' computing experience. We here only give a concise and general discussion because analogs of details can be found in Section 2.4, such as the conversion of a 3D problem to a 2D one, and the difference between radiation and scattering, and so on.

As far as the eigenmode problem is concerned, in Section 3.1 we only analyzed in detail the FEM solution of the dielectrically loaded waveguide. Here, we would like to supplement the FEM solution to other eigenmode problems briefly. As is mentioned in the beginning of this chapter, the eigenmode problem can be grouped into the waveguide eigenmode problem and the resonant cavity eigenmode problem. Basically, the FEM solutions to these two kinds of problems are the same. The difference stems from the problem domain. Since the problem domain in the cavity resonator is volume, the discretization elements should be edge elements defined on tetrahedra as introduced in Section 3.2. Although in the cavity resonator problem, more unknowns are involved that leads to heavy computation, the discretized formulation is simpler than that in the waveguide problem. Consequently, the derivation of the discretized formulation and the programming is even simpler than that of the waveguide problem. As far as the waveguide eigenmode problem is concerned, besides the dielectric-loaded waveguide, there are hollow-waveguide and open-waveguide problems. The hollow waveguide can be regarded as a special case of the dielectric-loaded waveguide. Therefore, the routines for the dielectric-loaded waveguide can readily be used for the hollow-waveguide problem. However, this is not the most efficient way. A better way is to use the scalar partial differential equation to describe the transverse magnetic (TM) mode and the transverse electric (TE) mode, and construct the corresponding scalar variation of functional and solve, as is introduced in reference [2]. As for the open waveguide, the problem is much harder due to the special process required by the infinite problem domain. The most widely used treatment is the infinite elements, that is, to use damping exponential function at some distance away from the waveguide boundary to mimic the field at infinity. The damping factor and the propagation constant of the waveguide are usually obtained by iteration. The reader is referred to reference [38,39].

The discontinuity problem in Section 3.2 and the scattering problem in Section 3.3 are all 3D FEM problems. For 2D problems, in terms of the formulation derivation and the computational resources demand, they are easier than the 3D ones. Analogous to the procedure in Section 2.4.1, the essential parts in 2D FEM are to simplify the Maxwell's equation into scalar partial differential equations according to the continuity of the EM fields in a certain direction, construct the scalar variational formulation of functional. The reader is referred to [2,40] for more details.

References

1. Silvester, P.P. and Ferrari, R.L. (1983) *Finite Elements for Electrical Engineering*, Cambridge University Press, Cambridge.
2. Jin, J.M. (1993) *The Finite-element method in Electromagnetics*, Wiley, New York.
3. Huang, H.J. (1963) *Microwave Principles*, Science Press, Beijing (in Chinese).
4. Cendes, Z.J. and Silvester, P. (1970) "Numerical solution of dielectric loaded waveguides: I. finite-element analysis." *IEEE Transactions on Microwave Theory and Techniques*, **18**(12), 1124–1131.
5. Angkaew, T., Matsuhara, M., and Kunagai, N. (1987) "Finite-element analysis of waveguide modes: A novel approach that eliminates spurious modes." *IEEE Transactions on Microwave Theory and Techniques*, **35**(2), 117–123.
6. Bardi, I. and Biro, O. (1991) "An efficient finite-element formulation without spurious modes for anisotropic waveguides." *IEEE Transactions on Microwave Theory and Techniques*, **39**(7), 1133–1139.
7. Lee, J.F., Sun, D.K., and Cendes, Z.J. (1991) "Full-waves analysis of dielectric waveguides using tangential vector finite elements." *IEEE Transactions on Microwave Theory and Techniques*, **39**(7), 1262–1271.
8. Zienkiewicz, C. and Taylor, R.L. (1989) Basic formulation and linear problems, in *The Finite-element method*, vol. **1**, 4th edn, McGraw-Hill, New York.
9. Wilkinson, J.H. and Reinsch, C. (1971) *Linear Algebra, Handbook for Automatic Computation*, vol. II, Springer-Verlag, New York.
10. Golub, G.H. and van Loan, C.F. (1996) *Matrix Computations*, 3nd edn, Johns Hopkins University Press, Baltimore.
11. Smith, B.T. *et al.* (1976) Lectures notes in computer science, in: *Matrix Eigensystem Routines: EISPACK Guide*, vol. **6**, 2nd edn, Springer-Verlag, New York.
12. Rice, J.R. (1983) *Numerical Methods, Software, and Analysis*, IMSL Reference Edition, McGraw-Hill, New York.
13. Lehoucq, R.B., Sorensen, D.C., and Yang, C. (1998) *ARPACK Users' Guide*, SIAM, Philadelphia.
14. Press, W.H., Teukolsky, S.A., Vetterling, W.T., and Flannery, B.P. (1992) *Numerical Recipes in Fortran: The Art of Scientific Computing*, 2nd edn, Cambridge University Press, New York.
15. Sheng, X.Q. (1996) Edge element and its applications, PhD Dissertation, (in Chinese).
16. Ise, K., Inoue, K., and Koshiba, M. (1991) "Three-dimensional finite-element method with edge elements for electromagnetic waveguide discontinuities." *IEEE Transactions on Microwave Theory and Techniques*, **39**(8), 1289–1295.
17. Liao, C.N. (1984) *Fundamentals of Microwave Technology*, National Defence Industrial Press, Changsha (in Chinese), pp. 68–70.
18. Sacks, Z.S., Kingsland, D.M., Lee, R., and Lee, J.F. (1995) "A perfectly matched anisotropic absorber for use as an absorbing boundary condition." *IEEE Transactions on Antennas and Propagation*, **43**(12), 1460–1463.
19. Davis, T.A. and Duff, I.S. (1997) "An unsymmetric-pattern multifrontal method for sparse LU factorization." *SIAM Journal of Matrix Analysis and Applications*, **19**(1), 140–158.
20. Chang, D.C. and Zheng, J.X. (1992) "Electromagnetic modeling of passive circuit elements in MMIC." *IEEE Transactions on Microwave Theory and Techniques*, **40**(9), 1741–1747.
21. Naishadham, K. and Lin, X.P. (1994) "Application of spectral domain Prony's method to the FDTD analysis of planar microstrip circuits." *IEEE Transactions on Microwave Theory and Techniques*, **42** (12), 2391–2398.
22. Roy, R., Paulraj, A., and Kailath, T. (1986) "ESPRIT-A subspace rotation approach to estimation of parameters of cisoids in noise." *IEEE Transactions on Acoustics Speech and Signal Processing, Signal Processing.*, **34**(10), 1340–1342.

23. Sheng, X.Q. and Yung, E.K.N. (2003) "Analysis of 3D ferrite-loaded waveguide discontinuities problems." *International Journal of RF and Microwave Computer-Aided Engineering*, **13**(5), 341–347.
24. Webb, J.P. and Kanellopoulos, V.N. (2007) "Absorbing boundary conditions for the finite element solution of the vector wave equation." *Microwave and Optical Technology Letters*, **2**(10), 370–372.
25. Chatterjee, A., Jin, J.M., and Volakis, J.L. (1993) "Edge-based finite elements and vector ABCs applied to 3D scattering." *IEEE Transactions on Antennas and Propagation*, **41**(2), 221–226.
26. Albanese, R. and Rubinacci, G. (1988) "Solution of three dimensional eddy current problems by integral and differential methods." *IEEE Transactions on Magnetics*, **24**(1), 98–101.
27. Sheng, X.Q. and Peng, Z. (2008) "A novel high-performance element in the electromagnetic finite-element method—node-edge element," *Journal of Systems Engineering and Electronics*, Vol. **15**, No. 5, pp. 878–881, Oct. 2008
28. Graglia, R.D., Wilton, D.R., and Peterson, A.F. (1997) "Higher order interpolatory vector bases for computational electromagnetics." *IEEE Transactions on Antennas and Propagation*, **45**(3), 329–342.
29. Webb, J.P. (1999) "Hierarchal vector basis functions of arbitrary order for triangular and tetrahedral finite elements." *IEEE Transactions on Antennas and Propagation*, **47**(8), 1244–1253.
30. Sun, D.K., Lee, J.F., and Cendes, Z. (2001) "Construction of nearly orthogonal Nedelec bases for rapid convergence with multilevel preconditioned solvers." *SIAM Journal of Scientific Computing*, **23**(4), 1053–1076.
31. Peterson, A.F. and Graglia, R.D. (2010) "Scale factors and matrix conditioning associated with triangular-cell hierarchical vector basis functions." *IEEE Antennas and Wireless Propagation Letters*, **9**, 40–43.
32. Lee, J.F. and Sacks, Z. (1995) "Whitney elements time domain (WETD) methods." *IEEE Transactions on Magnetics*, **31**(3), 1325–1329.
33. Riley, D.J., Jin, J.M., Lou, Z., and Petersson, L.E. (2006) "Total-and scattered-field decomposition technique for the finite-element time-domain method." *IEEE Transactions on Antennas and Propagation*, **54**(1), 35–41.
34. Zhu, B., Chen, J.F., Zhong, W.X., and Liu, Q.H. (2011) "A hybrid FETD-FDTD method with nonconforming meshes." *Computer Physics Communications*, **9**(3), 828–842.
35. Mohammadian, A.H., Shankar, V., and Hall, W.F. (1991) "Computation of electromagnetic scattering and radiation using a time-domain finite volume discretization procedure." *Computer Physics Communications*, **68**(1), 175–196.
36. Lu, T., Zhang, P., and Cai, W. (2004) "Discontinuous Galerkin methods for dispersive and lossy Maxwell equations and PML boundary conditions." *Journal of Computational Physics*, **200**, 549–580.
37. Chen, J.F., Liu, Q.H., Chai, M., and Mix, J.A. (2010) "A nonspurious 3D vector discontinuous Galerkin Finite-Element Time-Domain method." *Microwave and Wireless Components Letters*, **20**, 1–3.
38. Xu, S.J., Sheng, X.Q., Greiner, P. *et al.* (1993) "High-order finite-element analysis of scattering properties of II-VI semiconductor materials." *Chinese Journal Infrared and Millimeter Waves*, **12**(3), 201–209.
39. Rahman, B.M.A. and Davies, J.B. (1984) "Finite-element analysis of optical and microwave waveguide problems." *IEEE Transactions on Microwave Theory and Techniques*, **32**(6), 20–28.
40. McDougall, M.J. and Webb, J.P. (1989) "Infinite elements for the analysis of open dielectric waveguides." *IEEE Transactions on Microwave Theory and Techniques*, **37**(11), 1724–1731.

4

Finite-Difference Time-Domain Method

The *finite-difference time-domain method (FDTD)* is a full-wave numerical method that directly discretizes the time-domain partial differential form of Maxwell's equations. It was noted earlier that the partial differential is discretized by the finite difference. However, this technique was not found in a suitable form in the computational electromagnetics until the Chinese American K.S. Yee developed the Yee discretizing scheme in the 1960s. In the Yee discretizing scheme, the electric and magnetic field components are defined in an interleaving way both in a three-dimensional (3D) space and in time. This provides a beautifully simple picture of 3D space being filled by an interlinked array of Faraday's Law and Ampere's Law contours, representing the physical EM wave propagation. This Yee scheme is versatile and easy to use, and does not require much preknowledge. Due to these merits, this method is popular and gains vast applications in the computational electromagnetics. In this chapter, we still use a 3D scattering problem to illustrate the general idea and key issues of the Yee scheme. Then, some special treatments will be overviewed, followed by the comparison of the three computational methods, namely, MoM, FEM and FDTD, as a summary.

4.1 Scattering from a Three-Dimensional Objects

We have discussed the scattering problem for a 3D target in the previous two chapters. One of the reasons we still use this problem as an example is that the progress of the FDTD method largely happens during the pursuing of the accurate and fast solution to this problem. Moreover, through solving the same problem, the reader can make a fair comparison about the three numerical methods introduced in this book.

Essentials of Computational Electromagnetics, First Edition. Xin-Qing Sheng and Wei Song.
© 2012 John Wiley & Sons Singapore Pte. Ltd. Published 2012 by John Wiley & Sons Singapore Pte. Ltd.

4.1.1 FDTD Solution Scheme

The problem of scattering from a 3D target has been explained in the previous chapters. The only difference here is the form of the incident wave. The time-domain plane wave is used in accordance with the time-domain Maxwell's equations. The mathematical expression of the incident wave can be found in Section 4.1.7, and the time-domain Maxwell's equations have already been provided in Section 1.1.2. In this section, we will address the following two issues: the truncation of the unbounded problem domain, and the choice of the unknown field vector.

As introduced in Chapter 3, in numerical modeling, many geometries of interest are defined in an open region where the spatial domain of the computed field is unbounded in one or more coordinate directions. Computers can only store and calculate a limited amount of data. Therefore, an absorbing boundary condition (ABC) is often used to truncate the computational domain, which suppresses spurious reflections of the outgoing numerical wave analogs to an acceptable level. Generally, there are two groups of local-domain truncation ABCs: some derived from differential equations and others employing absorbing materials. Among the first group of ABCs, we've discussed the radiation boundary condition in Section 3.3.1. For FDTD, however, the most popular ABC is the one derived by Engquist and Majda with the discretization given by Mur [1,2]. This ABC is based on approximating the outgoing wave equation by linear expressions using a Taylor approximation. Material-based ABCs are realized by surrounding the computational domain with a lossy material that makes the outgoing fields attenuate. In this group, the perfectly matched layer (PML) technique [3] which was put forward by Berenger in 1994, and further developed in [4,5] demonstrated significantly better accuracy than other ABCs. Hence, it is widely used in the FDTD simulations. In this chapter, we will present a detailed treatment in realizing PML.

Note that the FDTD operations of the Yee algorithm with Mur's ABC or PML can be applied with equal validity to the scattered field as to the total field. Since in a scattering problem, the scattered field describes the scattering characteristics of the object, it is convenient to use the scattered field as the unknown field vector. However, in certain regions, such as the shadow area, the scattered field has equivalent amplitude and opposite phase to the incident field, resulting in the total field being close to zero. In this case, if we use the scattered field as the unknown field, a minor error in the scattered field will cause a relatively large error in terms of the total field. In this sense, in this kind of region, it's proper to use the total field as the unknown variable. To this end, the so-called *total-field/scattered-field technique* (TF/SF) [2,6] was proposed. As shown in Figure 4.1, the Yee space lattice is zoned into two distinct regions: the center region comprising the scatterers, where the total field is used as the unknowns, and the surrounding region, where only the scattered field is taken as the unknown variable. The virtual surface separating these two regions is referred to as the connection boundary. The connecting condition is applied on this boundary to

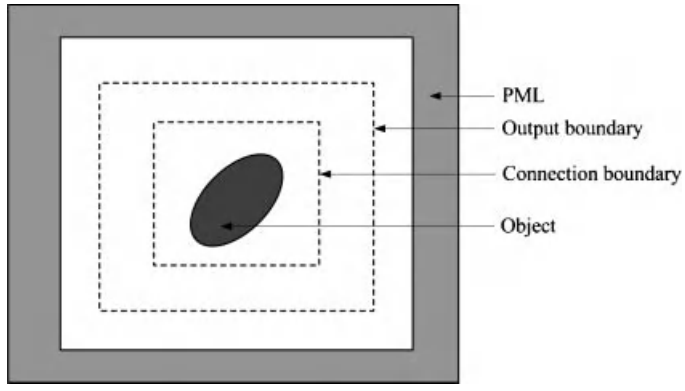


Figure 4.1 The FDTD model.

generate an arbitrary user-specified incident plane wave in the center region that simultaneously ensures consistency of the numerical space derivative operations across the interface. By adopting such a scheme, the incident field can be easily and uniformly enforced through the connection boundary. More details about the generation of the plane wave are given in Section 4.1.8. Outside the connection boundary, the scattered field will be absorbed by the ABC. In order to calculate the scattered field in the far zone, one closed surface is defined as the output boundary, on which Huygens' principle is applied. In principle, this output boundary can be in arbitrary shape enclosing the scatterer between the ABC and the connection boundary. In practice, we usually choose a box for convenience.

4.1.2 Perfectly Matched Layers

Since it was created by the French scholar Berenger [2], the *perfectly matched layers* (*PML*) has experienced continuous improvement and extension and has become an effective truncation scheme with neat expression. The PML has been utilized in Section 3.2.1. In short, when the scattered field enters from free space or one material into the PML, there is no reflection, and the wave is attenuated in the PMLs. How is this property obtained? Let us unfurl the design of this PML material [5]. Apply the local Wilcox series expansion [7] to the scattered field on point P from the inner boundary of the PML:

$$\boldsymbol{E}^S(u,v,w) = \frac{e^{-jkw}}{4\pi(r_1 r_2)^{1/2}} \sum_{n=0}^{\infty} \frac{\boldsymbol{E}_n^S(u,v)}{(r_1 r_2)^{n/2}} \tag{4.1}$$

where (u, v, w) are the local coordinate system with coordinate stretching coefficients $(h_1, h_2, 1)$ based on the normal Cartesian coordinate. $k = \omega/c, r_1 = r_{01}(u,v) +$

$w, r_2 = r_{02}(u,v) + w$, c is the phase velocity of the wave in the medium, and r_{01}, r_{02} are the principle radii with boundary on point P. These parameters satisfy: $h_1 = r_1/r_{01}$ and $h_2 = r_2/r_{02}$. First, let us apply the following complex transformation to the coordinate w in the absorbing material region:

$$\tilde{w} = s_w w, \quad (w > 0) \tag{4.2}$$

where $s_w = \kappa + \frac{\sigma}{j\omega\varepsilon}$ is a complex number. By inserting it into (4.1), we obtain:

$$\tilde{\boldsymbol{E}}^S(u,v,w) = \frac{e^{-k\frac{\sigma}{\omega\varepsilon}(w)} e^{-jk_0\kappa(w)}}{4\pi(\tilde{r}_1\,\tilde{r}_2)^{1/2}} \sum_{n=0}^{\infty} \frac{\boldsymbol{E}_n^S(u,v)}{(\tilde{r}_1\,\tilde{r}_2)^{n/2}} \tag{4.3}$$

where $\tilde{r}_1 = r_{01} + \tilde{w}$ and $\tilde{r}_2 = r_{02} + \tilde{w}$. The $\alpha \geq 0$ in (4.3) is the exponential attenuation factor introduced in the w direction. Note that (4.3) is an analytic function of $\tilde{w}$. Therefore, the field at the interface in the normal FDTD region ($w \to 0^-$) equals the field at the interface in the absorbing material region ($w \to 0^+$). In other words, this transformation assures the continuity of the field at the interface. As a result, the reflectionless property of the absorbing material is assured. Under this transformation, the operator ∇ can be expressed using the Cartesian coordinate system as

$$\nabla \to \bar{\nabla} = \hat{x}\frac{\partial}{\partial\bar{x}} + \hat{y}\frac{\partial}{\partial\bar{y}} + \hat{z}\frac{\partial}{\partial\bar{z}} = \hat{x}\frac{\partial}{\partial x} + \hat{y}\frac{\partial}{\partial y} + \hat{z}\frac{1}{s_Z}\frac{\partial}{\partial z} \tag{4.4}$$

The above transformation can be generalized into the other directions and gives:

$$\nabla \to \bar{\nabla} = \hat{x}\frac{\partial}{\partial\bar{x}} + \hat{y}\frac{\partial}{\partial\bar{y}} + \hat{z}\frac{\partial}{\partial\bar{z}} = \hat{x}\frac{1}{s_x}\frac{\partial}{\partial x} + \hat{y}\frac{1}{s_y}\frac{\partial}{\partial y} + \hat{z}\frac{1}{s_Z}\frac{\partial}{\partial z} \tag{4.5}$$

Obviously, under condition $s_x = s_y = 1$, (4.5) reduces to (4.4). Equation 4.5 can be expressed more concisely as

$$\bar{\nabla} = \bar{\bar{S}} \cdot \nabla \tag{4.6}$$

where

$$\bar{\bar{S}} = \hat{x}\hat{x}\left(\frac{1}{s_x}\right) + \hat{y}\hat{y}\left(\frac{1}{s_y}\right) + \hat{z}\hat{z}\left(\frac{1}{s_z}\right) \tag{4.7}$$

In this way, the modified Maxwell's equations become:

$$\bar{\nabla} \times \mathbf{E}^c = -j\omega\mu\mathbf{H}^c \tag{4.8}$$

$$\bar{\nabla} \times \mathbf{H}^c = j\omega\varepsilon\mathbf{E}^c \tag{4.9}$$

Note there is a difference between the fields in the above equations and the physical fields, which is denoted by the superscript "c," since the operator is different from the

conventional one in the Maxwell's equations. So, in the following, we will convert (4.8) and (4.9) to Maxwell's form through mathematical operations. Making use of the following identity in the Cartesian coordinate:

$$\nabla \times \left(\bar{\bar{S}}^{-1} \cdot \boldsymbol{a}\right) = \left(\det \bar{\bar{S}}\right)^{-1} \bar{\bar{S}} \cdot \left(\bar{\bar{S}} \cdot \nabla\right) \times \boldsymbol{a} \tag{4.10}$$

where $\boldsymbol{a}(\boldsymbol{r})$ is an arbitrary vector function, and $\det \bar{\bar{S}} = \left(s_x s_y s_z\right)^{-1}$. Then, (4.8) and (4.9) can be rewritten as

$$\nabla \times \left(\bar{\bar{S}}^{-1} \cdot \mathbf{E}^c\right) = -j\omega\mu \left(\det \bar{\bar{S}}\right)^{-1} \bar{\bar{S}} \cdot \mathbf{H}^c \tag{4.11}$$

$$\nabla \times \left(\bar{\bar{S}}^{-1} \cdot \mathbf{H}^c\right) = j\omega\varepsilon \left(\det \bar{\bar{S}}\right)^{-1} \bar{\bar{S}} \cdot \mathbf{E}^c \tag{4.12}$$

Let

$$\mathbf{E} = \bar{\bar{S}}^{-1} \cdot \mathbf{E}^c \tag{4.13}$$

$$\mathbf{H} = \bar{\bar{S}}^{-1} \cdot \mathbf{H}^c \tag{4.14}$$

Thus, (4.8) and (4.9) become

$$\nabla \times \mathbf{E} = -j\omega\mu \left(\det \bar{\bar{S}}\right)^{-1} \bar{\bar{S}} \cdot \bar{\bar{S}}\mathbf{H} \tag{4.15}$$

$$\nabla \times \mathbf{H} = j\omega\varepsilon \left(\det \bar{\bar{S}}\right)^{-1} \bar{\bar{S}} \cdot \bar{\bar{S}}\mathbf{E} \tag{4.16}$$

The above equations are the Maxwell's equations in the perfect matched layers. From these equations, we can derive the relative permittivity and permeability of the media in PML:

$$\bar{\bar{\varepsilon}}_r = \bar{\bar{\mu}}_r = \left(\det \bar{\bar{S}}\right)^{-1} \bar{\bar{S}} \cdot \bar{\bar{S}} = \bar{\bar{\Lambda}} \tag{4.17}$$

where

$$\bar{\bar{\Lambda}} = \hat{x}\hat{x}\Lambda_x + \hat{y}\hat{y}\Lambda_y + \hat{z}\hat{z}\Lambda_z \tag{4.18}$$

$$\Lambda_x = \frac{s_y s_z}{s_x}; \Lambda_y = \frac{s_z s_x}{s_y}; \Lambda_z = \frac{s_x s_y}{s_z} \tag{4.19}$$

Table 4.1 Special cases involved in the entire FDTD lattice implementing PML.

Special zones	Special ranges that require $s_p = 1$ (i.e. $\kappa_p = 1$, and $\sigma_p = 0$)
Lossless, isotropic interior zone	$p \in (x, y, z)$
PML planes at i_{UPML} $i \in (x, y, z)$	$p \in (x, y, z)$ and $p \neq i$
PML dihedral corners at (i_{UPML} and j_{UPML}) ($i, j \in (x, y, z)$ and $i \neq j$)	$p \in (x, y, z)$ and $p \neq i, p \neq j$
PML trihedral corners at (x_{UPML}, y_{UPML} and z_{UPML})	None

In the frequency domain, the complex coordinate stretching factors generally take the following forms:

$$s_x = \kappa_x + \gamma_x = \kappa_x + \frac{\sigma_x}{j\omega\varepsilon}$$

$$s_y = \kappa_y + \gamma_y = \kappa_y + \frac{\sigma_y}{j\omega\varepsilon}$$

$$s_z = \kappa_z + \gamma_z = \kappa_z + \frac{\sigma_z}{j\omega\varepsilon} \tag{4.20}$$

where the imaginary part of the stretching factor, that is, the attenuation coefficient of the EM wave, is frequency dependent. Equation (4.20) are the general equation describing the EM waves in PML. In applying this equation, we need to pay special attention to the parameters chosen according to the specific spatial region as listed in Table 4.1. For the example shown in Figure 4.2, we choose $\sigma_x = \sigma_y = \sigma_z = 0$ in the PML-free block. In the face block of PML$_x$, we need to choose $\sigma_x \neq 0, \sigma_y = \sigma_z = 0$.

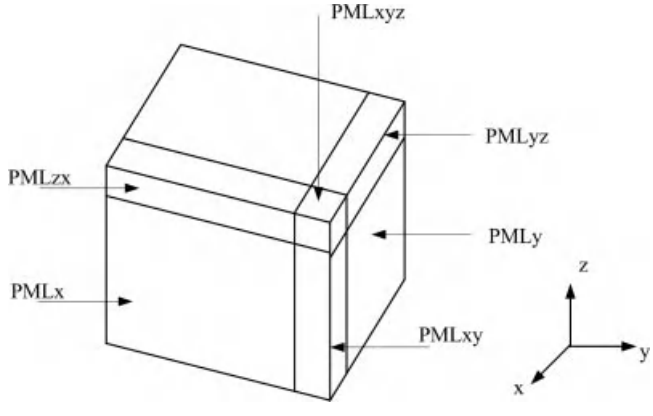


Figure 4.2 The PML blocks with different parameters.

In the following table, all the special cases involved in the entire FDTD lattice implementing PML are listed:

Since the thickness of PML is usually fixed in the computation, the electrical thickness corresponding to different frequency is different. So another advantage in choosing such attenuation coefficients (4.20) is that this PML can provide equivalent absorbing effect in case of various frequencies. However, this leads to extra effort in converting (4.15) and (4.16) from the frequency domain to the time domain. Next, we will discuss the implementation of PML. From (4.15)–(4.19), Ampere's law (4.16) in a matched PML is expressed as

$$\begin{bmatrix} \dfrac{\partial H_z}{\partial y} - \dfrac{\partial H_y}{\partial z} \\ \dfrac{\partial H_x}{\partial z} - \dfrac{\partial H_z}{\partial x} \\ \dfrac{\partial H_y}{\partial x} - \dfrac{\partial H_x}{\partial y} \end{bmatrix} = j\omega\varepsilon \begin{bmatrix} \dfrac{s_y s_z}{s_x} & 0 & 0 \\ 0 & \dfrac{s_z s_x}{s_y} & 0 \\ 0 & 0 & \dfrac{s_x s_y}{s_z} \end{bmatrix} \begin{bmatrix} E_x \\ E_y \\ E_z \end{bmatrix} \tag{4.21}$$

Directly substituting (4.20) into (4.21) and then transforming into the time domain would result in a computationally intensive convolution between the tensor coefficients and the E-field. Therefore, in the following we introduce an auxiliary differential equation (ADE) method [8,9] widely used in dealing with a general dispersive medium. By introducing auxiliary variables, this technique yields a composite self-consistent system. Let

$$D_x = \varepsilon \frac{s_z}{s_x} E_x, \tag{4.22}$$

$$D_y = \varepsilon \frac{s_x}{s_y} E_y, \tag{4.23}$$

$$D_z = \varepsilon \frac{s_y}{s_z} E_z. \tag{4.24}$$

Then, (4.21) is rewritten as

$$\begin{bmatrix} \dfrac{\partial H_z}{\partial y} - \dfrac{\partial H_y}{\partial z} \\ \dfrac{\partial H_x}{\partial z} - \dfrac{\partial H_z}{\partial x} \\ \dfrac{\partial H_y}{\partial x} - \dfrac{\partial H_x}{\partial y} \end{bmatrix} = j\omega \begin{bmatrix} s_y & 0 & 0 \\ 0 & s_z & 0 \\ 0 & 0 & s_x \end{bmatrix} \begin{bmatrix} D_x \\ D_y \\ D_z \end{bmatrix} \tag{4.25}$$

By substituting (4.20) into (4.25), and applying the inverse Fourier transform using the identity $j\omega f(\omega) \rightarrow (\partial/\partial t) f(t)$, we obtain a system of time-domain differential

equations for (4.25):

$$\begin{bmatrix} \dfrac{\partial H_z}{\partial y} - \dfrac{\partial H_y}{\partial z} \\ \dfrac{\partial H_x}{\partial z} - \dfrac{\partial H_z}{\partial x} \\ \dfrac{\partial H_y}{\partial x} - \dfrac{\partial H_x}{\partial y} \end{bmatrix} = \frac{\partial}{\partial t} \begin{bmatrix} \kappa_y & 0 & 0 \\ 0 & \kappa_z & 0 \\ 0 & 0 & \kappa_x \end{bmatrix} \begin{bmatrix} D_x \\ D_y \\ D_z \end{bmatrix} + \frac{1}{\varepsilon} \begin{bmatrix} \sigma_y & 0 & 0 \\ 0 & \sigma_z & 0 \\ 0 & 0 & \sigma_x \end{bmatrix} \begin{bmatrix} D_x \\ D_y \\ D_z \end{bmatrix} \tag{4.26}$$

By discretizing this system of equations in a standard Yee's leapfrogging manner as will be introduced in the following section, D_x, D_y, and D_z can be expressed in an explicit time-marching way. Further, let us examine (4.22). Inserting the expressions of s_i, it can be written as

$$\left(\kappa_x + \frac{\sigma_x}{j\omega\varepsilon}\right) D_x = \varepsilon \left(\kappa_z + \frac{\sigma_z}{j\omega\varepsilon}\right) E_x \tag{4.27}$$

Multiplying both sides of (4.27) by $j\omega$ and transforming the result into the time-domain yields

$$\frac{\partial}{\partial t}(\kappa_x D_x) + \frac{\sigma_x}{\varepsilon} D_x = \varepsilon \left[\frac{\partial}{\partial t}(\kappa_z E_x) + \frac{\sigma_z}{\varepsilon} E_x\right] \tag{4.28}$$

In a similar way, we can obtain the relations of D_y,E_y, and D_z, E_z in the time domain as

$$\frac{\partial}{\partial t}(\kappa_y D_y) + \frac{\sigma_y}{\varepsilon} D_y = \varepsilon \left[\frac{\partial}{\partial t}(\kappa_x E_y) + \frac{\sigma_x}{\varepsilon} E_y\right] \tag{4.29}$$

$$\frac{\partial}{\partial t}(\kappa_z D_z) + \frac{\sigma_z}{\varepsilon} D_z = \varepsilon \left[\frac{\partial}{\partial t}(\kappa_y E_z) + \frac{\sigma_y}{\varepsilon} E_z\right] \tag{4.30}$$

Again, the Yee's leapfrogging scheme can be used to discretize the above equations to obtain an explicit time-marching scheme with updating of D_i readily accomplished.

Updating of the components of **H** can be done in an analogous way. That is, starting with Faraday's Law (4.15), define the constitutive relation dual to (4.22)–(4.24) (e.g., $B_x = \mu \frac{s_z}{s_x} H_x$), obtain the new values of the components of **B** in the first step, and in the second step, obtain the new value of **H** using the newly updated **B**.

Finally, let us talk about the values of the PML parameters. In an FDTD program, we expect to use a moderate thickness of PML to attenuate the outgoing wave, which enables us to use a PEC boundary to terminate the whole computational domain. As mentioned before, we have freedom in changing the values of $s_i (i \in (x, y, z))$ along the corresponding i direction without ruining the reflectionless property of PML. But the discretization will cause numerical reflection. The more rapidly parameters in PML

change, the larger the reflection will be. Therefore, gradually varying parameters are usually adopted in PML, in order to reduce the numerical reflection at the interface of PML due to discretization. Take the z direction for example. Choose $s_z = 1 + \sigma'_z/j\omega\varepsilon_0$, where the conductivity is normalized by the local ε_r to assure the transverse invariance. Then, σ'_z is chosen to be spatially variant along the z-axis in Gedney's experiment [4]

$$\sigma'_z(z) = \frac{\sigma'_{\max}|z-z_0|^m}{d^m} \tag{4.31}$$

where d is the thickness of the PML, z_0 denotes the position of the PML interface, and m is an integer. Numerical experiments [4] showed that $m = 4$, and

$$\sigma'_{\max} = \frac{m+1}{\sqrt{\varepsilon_r}150\pi\Delta} \tag{4.32}$$

provide the minimized reflection from PML. Here, Δ is the FDTD spatial discretization.

Problem

4.1 Consider a PML used to terminate a conductive medium. Instead of (4.21), Ampere's law in this PML region can be expressed as

$$\begin{bmatrix} \frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z} \\ \frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x} \\ \frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} \end{bmatrix} = (j\omega\varepsilon + \sigma) \begin{bmatrix} \frac{s_y s_z}{s_x} & 0 & 0 \\ 0 & \frac{s_z s_x}{s_y} & 0 \\ 0 & 0 & \frac{s_x s_y}{s_z} \end{bmatrix} \begin{bmatrix} E_x \\ E_y \\ E_z \end{bmatrix}$$

Design an explicit updating approach for the E-field components.

4.1.3 Yee Discretizing Scheme

In this section, we will introduce how to discretize the time-domain Maxwell's equations in the *Yee scheme*. First, we define the discretized electric and magnetic field components in an interleaving way in a three-dimensional space, as shown in Figure 4.3. From this figure, it can be seen that each electric field component is surrounded by four magnetic field components, and each magnetic field component is surrounded by four electric field components. In this way, the partial derivative of an arbitrary E-component with respect to time, which equals the curl of the magnetic

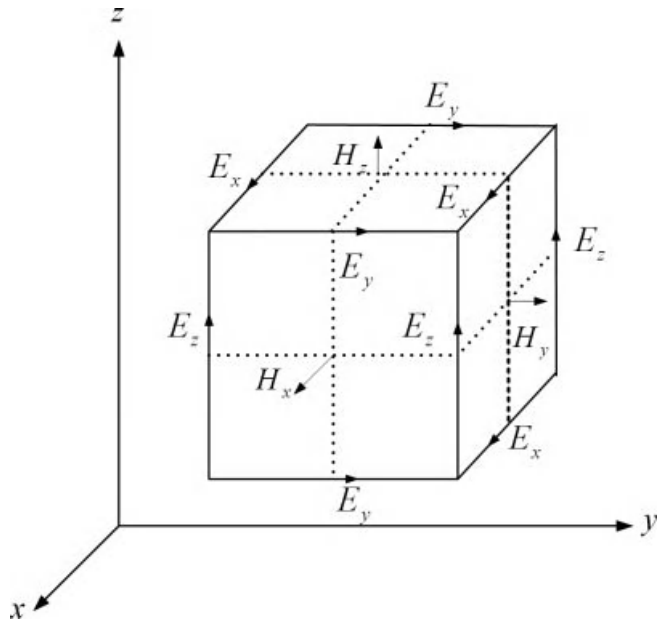


Figure 4.3 Yee's spatial grids.

field on that exact point, can be approximated by the central difference of the surrounding H-components. Analogously, the partial derivative of an arbitrary H-component with respect to time can be approximately expressed by the surrounding E-components. Second, in the time domain, the discretized electric and magnetic field components are also defined in an interleaving manner, which is often referred to as Yee's leapfrogging scheme. Then, the partial derivative of an arbitrary E (or H) component with respect to time is approximately expressed by the central difference of the neighboring discretized components in time. It is worth noting that some discretized E and H components are defined at half a time step, and half a grid position, the new variables **D** and **B** introduced in the previous section should be aligned with the corresponding **E** and **H**. More specifically, if we denote a certain spatial grid as

$$(i,j,k) = (i\Delta x, j\Delta y, k\Delta z) \tag{4.33}$$

where Δx, Δy, and Δz denote the grid dimensions in the x, y, and z directions, respectively. Denote the value of an arbitrary function u in this grid and on time step n as

$$u(i\Delta x, j\Delta y, k\Delta z, n\Delta t) = u|^n_{i,j,k} \tag{4.34}$$

where Δt is the temporal increment, then the scalar equation about the x component in (4.22) can be discretized as

$$\begin{aligned}&\frac{D_x|_{i+1/2,j,k}^{n+1}-D_x|_{i+1/2,j,k}^{n}}{\Delta t}\kappa_y+\frac{D_x|_{i+1/2,j,k}^{n+1}+D_x|_{i+1/2,j,k}^{n}}{2\varepsilon}\sigma_y\\&\quad=\frac{H_z|_{i+1/2,j+1/2,k}^{n+1/2}-H_z|_{i+1/2,j-1/2,k}^{n+1/2}}{\Delta y}-\frac{H_y|_{i+1/2,j,k+1/2}^{n+1/2}-H_y|_{i+1/2,j,k-1/2}^{n+1/2}}{\Delta z}\end{aligned}\tag{4.35}$$

Note in the above equation, $D_x|_{i+1/2,j,k}^{n+1/2}$is expressed in a semi-implicit approximation. Rewriting (4.35) we can obtain the explicit marching-on-time (MOT) expression of $D_x|_{i+1/2,j,k}^{n+1}$:

$$\begin{aligned}D_x|_{i+1/2,j,k}^{n+1}&=\left(\frac{2\varepsilon\kappa_y-\sigma_y\Delta t}{2\varepsilon\kappa_y+\sigma_y\Delta t}\right)D_x\Big|_{i+1/2,j,k}^{n}+\left(\frac{2\varepsilon\Delta t}{2\varepsilon\kappa_y+\sigma_y\Delta t}\right)\\&\cdot\left(\frac{H_z|_{i+1/2,j+1/2,k}^{n+1/2}-H_z|_{i+1/2,j-1/2,k}^{n+1/2}}{\Delta y}-\frac{H_y|_{i+1/2,j,k+1/2}^{n+1/2}-H_y|_{i+1/2,j,k-1/2}^{n+1/2}}{\Delta z}\right)\end{aligned}\tag{4.36}$$

Discretizing (4.24) and by introducing a semi-implicit approximation for $E_x|_{i+1/2,j,k}^{n+1/2}$, we arrive at

$$\begin{aligned}&\kappa_x\frac{D_x|_{i+1/2,j,k}^{n+1}-D_x|_{i+1/2,j,k}^{n}}{\Delta t}+\sigma_x\frac{D_x|_{i+1/2,j,k}^{n+1}+D_x|_{i+1/2,j,k}^{n}}{2\varepsilon}\\&\quad=\varepsilon\left(\kappa_z\frac{E_x|_{i+1/2,j,k}^{n+1}-E_x|_{i+1/2,j,k}^{n}}{\Delta t}+\sigma_z\frac{E_x|_{i+1/2,j,k}^{n+1}+E_x|_{i+1/2,j,k}^{n}}{2\varepsilon}\right)\end{aligned}\tag{4.37}$$

which leads to the explicit expression of $E_x|_{i+1/2,j,k}^{n+1}$

$$\begin{aligned}E_x|_{i+1/2,j,k}^{n+1}&=\left(\frac{2\varepsilon\kappa_z-\sigma_z\Delta t}{2\varepsilon\kappa_z+\sigma_z\Delta t}\right)E_x\Big|_{i+1/2,j,k}^{n}+\frac{1}{(2\varepsilon\kappa_z+\sigma_z\Delta t)\varepsilon}\\&\cdot\left[(2\varepsilon\kappa_x+\sigma_x\Delta t)D_x|_{i+1/2,j,k}^{n+1}-(2\varepsilon\kappa_x-\sigma_x\Delta t)D_x|_{i+1/2,j,k}^{n}\right]\end{aligned}\tag{4.38}$$

In a similar approach, the discretized E and H field components can all be expressed explicitly. For example, B_x and H_x can be updated by:

$$B_x\Big|_{i,j+1/2,k+1/2}^{n+3/2} = \left(\frac{2\varepsilon\kappa_y - \sigma_y\Delta t}{2\varepsilon\kappa_y + \sigma_y\Delta t}\right) B_x\Big|_{i,j+1/2,k+1/2}^{n+1/2} - \frac{2\varepsilon\Delta t}{2\varepsilon\kappa_y + \sigma_y\Delta t} \cdot \left(\frac{E_z\big|_{i,j+1,k+1/2}^{n+1} - E_z\big|_{i,j,k+1/2}^{n+1}}{\Delta y} - \frac{E_y\big|_{i,j+1/2,k+1}^{n+1} - E_y\big|_{i,j+1/2,k}^{n+1}}{\Delta z}\right) \tag{4.39}$$

$$H_x\Big|_{i,j+1/2,k+1/2}^{n+3/2} = \left(\frac{2\varepsilon\kappa_z - \sigma_z\Delta t}{2\varepsilon\kappa_z + \sigma_z\Delta t}\right) H_x\Big|_{i,j+1/2,k+1/2}^{n+1/2} + \frac{1}{(2\varepsilon\kappa_z + \sigma_z\Delta t)\mu} \cdot \left[(2\varepsilon\kappa_x + \sigma_x\Delta t)B_x\Big|_{i,j+1/2,k+1/2}^{n+3/2} - (2\varepsilon\kappa_x - \sigma_x\Delta t)B_x\Big|_{i,j+1/2,k+1/2}^{n+1/2}\right] \tag{4.40}$$

It is worth mentioning that, at the connection boundary of the total field and the scattered field, the above iterative equations need to be adjusted before being applied. The reason for this is as follows. The above equations suit for either pure total fields or pure scattered-fields. However, in the vicinity of the connection boundary, the total fields exist inside the boundary and the scattered fields exist outside it. Therefore, in the equations, we need to add the incident fields to the scattered fields to match the total fields inside the connection boundary, or to subtract the incident fields from the total-fields to match the scattered fields outside the connection boundary. This is the *connecting condition*. By doing this, the excitation is naturally added in the discretization equations. In computer programming, we need to pay extra attention to the different spatial position due to the interleaving scheme in Yee's spatial grid as well. Besides, the points on the corner and on the face of the connection boundary need to be treated separately, since the distributions of the total fields and the scattered fields are different. More details can be found in reference [8] (pages 84–85).

Problems

4.2 Derive the explicit discretized time-marching expressions for D_y, E_y, B_z and H_z.

4.3 Write the adjusted equations for the connection boundary.

4.4 Consider a 2D FDTD computational domain in the x-y plane with free-space truncated by PEC. The size of the whole domain is $500\Delta \times 300\Delta$, where Δ is the spatial increment in x and y directions. Choose $\Delta t = \Delta/(2c)$ as the temporal increment where c is the speed of light in the free space. A TE_z polarized wave involves the following EM components: E_x, E_y and H_z. Write a program updating

these components based on (4.36), (4.38) and the results of Problem 4.2. A hard line source of sine functions at ($x = 100\Delta$ to $200\Delta, y = 100\Delta$) operating in a short period ($t_s \leq = 50\Delta t$) can be added by forcing the H_z component at the source to be the defined value in the time-marching process. The program writer can define Δ, the starting phase and amplitude of the source and the total time step number. Choose the frequency to be $1/(10\Delta t)$. Test the program to see whether it provides reasonable results. Apply FFT to the temporal fields at point $(200\Delta, 200\Delta)$. Further investigate with sources of higher and lower frequency and explain the results.

4.5 In Problem 4.4, the outermost 25 layers of free space are replaced by PMLs. Test the program again to see the performance of the PML.

4.6 Based on Problem 4.4, we define the total-field region in the center with size $300\Delta \times 200\Delta$. Remove the hard source. Suppose source $S(x, y, t)$ is given on the connection boundary. Add the connection boundary based on the adjusted equations derived in Problem 4.3. See how the source is added through the connection boundary and how the total field is confined by the connection boundary.

4.7 A cylindrical PEC cavity resonator is assumed to be infinite long. The radius of the cavity is 0.15 m. The computational region is set to be (0.5 m × 0.5 m) meshed by (60×60) cells. The cavity is filled with vacuum and the outer material is PEC. Since the PEC layer has sufficient thickness, penetration of the EM field through the PEC can be neglected. A sine-modulated cosine pulse is excited at selected source positions inside the cavity to provide a wide-band excitation to excite all the possible TE modes: $\text{Source}(n) = \sin(2\pi f \cdot n \cdot dt) \cdot (1 - \cos(2\pi f \cdot n \cdot dt))$ where $f = 3$ GHz is the frequency parameter, and $dt = 1$ ps is the time increment. Plot the time-domain signal and the frequency spectrum within 0–3 GHz of the source. Find a proper total time step number. Obtain the lowest four spectra of the cavity resonator and compare them with the analytical results (0.586 GHz, 0.9721 GHz, 1.2198 GHz, and 1.3372 GHz). Change the spatial resolution to be (80×80) cells, and (120×120) cells and observe the change of numerical efficiency and accuracy under different spatial resolution.

4.8 Consider an air-filled rectangular PEC cavity resonator. The dimension of the air space between the inner walls is $a \times b \times l = 0.4\,\text{m} \times 0.2\,\text{m} \times 0.3\,\text{m}$, and the thickness of the PEC wall is 0.1 m. This air-filled cavity has resonant frequency $f_0 = \frac{1}{2\sqrt{\mu\varepsilon}} \sqrt{\left(\frac{m}{a}\right)^2 + \left(\frac{n}{b}\right)^2 + \left(\frac{p}{l}\right)^2}$ where the indices m, n, and p can take on all positive integer values, including zeros, as long as only one index is zero at a time. Obtain the lowest five spectra of the cavity resonator by using the above equation. Develop the FDTD program to calculate the lowest five spectra. In the program, choose $\Delta x = \Delta y = \Delta z = \Delta = 0.01$ m, and $\Delta t = \Delta/(2c)$, where c is the speed of light in the free space. Compare the numerical results with the analytical results.

4.1.4 Discretization of the Scatterer Model

The discretization in the FDTD method is simple. We can first define a box containing the scatterer as the total computational domain. Then, in the x, y, and z directions, we use Δx, Δy and Δz, respectively, to equally divide the domain into uniform orthogonal meshes. Note that in contrast to the FEM scheme, the cell in the FDTD grid does not correspond to the constitutive parameters of the material in a one-to-one way. Instead, it is the spatial point of the unknown variable that one-to-one corresponds to the constitutive parameter of the material. For example, in an anisotropic material, on a certain point where E_x is defined, the relative permittivity ε_r is the average dielectric constant in all the FDTD grids sharing this point. Therefore, in the same cell in a FDTD grid, the relative permittivity ε_r may be different on different points where E_y components are defined.

4.1.5 Treatment on the Curved Boundary

The above Yee scheme is based on rectangular grids. An oblique (or curved) surface has to be approximated by staircased mesh. It usually requires very fine meshes, which results in very small time steps and consequent extensive computation. In order to achieve a satisfactory accuracy without dramatically increasing the computational burden, a modified FDTD formulation based on the integral form of Maxwell's equations was proposed and is presented below. The curved surface can be either metallic or dielectric, and the treatment is different according to the types of material. For a metallic boundary, we use the integral form of the Faraday's law to obtain the modified iteration equations for the magnetic fields near the boundary. The effect of the curved boundary is fully accounted for in the modified equations for updating the magnetic fields. As a result, the electric fields can be updated in a conventional way. This is the *contour path FDTD (CP-FDTD or CFDTD)* scheme [11]. For the example shown in Figure 4.4, one cell of the FDTD grid is divided into two parts by a boundary, that is, the metallic part (the dark region) and the free space (the white region). We know that the iterative equation for the magnetic field component H_x in the free space is

$$H_x|^{n+1/2}_{i,j+1/2,k+1/2} = H_x|^{n-1/2}_{i,j+1/2,k+1/2} - \frac{\Delta t}{\mu_0}\left(\frac{E_z|^n_{i,j+1,k+1/2} - E_z|^n_{i,j,k+1/2}}{\Delta y} - \frac{E_y|^n_{i,j+1/2,k+1} - E_y|^n_{i,j+1/2,k}}{\Delta z}\right) \tag{4.41}$$

Next, let us explain the modification for the FDTD grid as illustrated in Figure 4.4. By applying the integral form of the Faraday's law to this grid

$$\oint_l E \cdot dl = -\mu_0 \frac{\partial}{\partial t}\int_S H \cdot dS \tag{4.42}$$

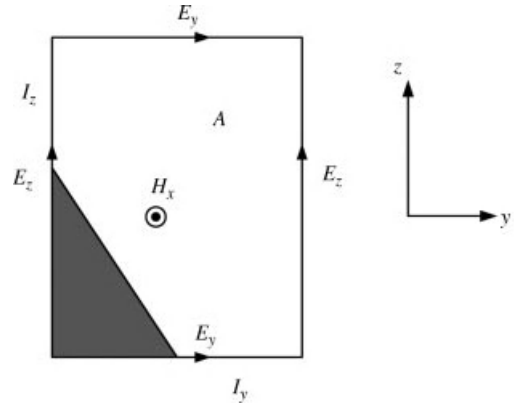


Figure 4.4 The treatment on a curved metallic boundary.

and by making use of the condition that the tangential electric field on a metallic surface equals zero, the modified time-marching equation for H_x can be expressed as

$$\begin{aligned} H_x\big|_{i,j+1/2,k+1/2}^{n+1/2} = H_x\big|_{i,j+1/2,k+1/2}^{n-1/2} - \frac{\Delta t}{\mu_0 A}\Big(&E_z\big|_{i,j+1,k+1/2}^{n} l_z\big|_{i,j+1,k+1/2} \\ &- E_z\big|_{i,j,k+1/2}^{n} l_z\big|_{i,j,k+1/2} - E_y\big|_{i,j+1/2,k+1}^{n} l_y\big|_{i,j+1/2,k+1} \\ &+ E_y\big|_{i,j+1/2,k}^{n} l_y\big|_{i,j+1/2,k}\Big) \end{aligned} \tag{4.43}$$

where A denotes the area of the free space part in this grid, and l_y, l_z denote the lengths of the free space segment on the bottom and left edges. The equations for the other two magnetic components can be obtained in the same way.

As for the dielectric curved boundary, the treatment is to define different effective constitutive parameter corresponding to different EM components. For example, if two materials are present in one cell, as shown in Figure 4.5, the effective permittivity $\varepsilon_y^e|_{i,j+1/2,k}$ for $E_y|_{i,j+1/2,k}$, and $\varepsilon_y^e|_{i,j,k+1/2}$ for $E_z|_{i,j,k+1/2}$ can be calculated as

$$\varepsilon_y^e\Big|_{i,j+1/2,k} = [\Delta y_2 \varepsilon_{y2} + (\Delta y - \Delta y_2)\varepsilon_{y1}]/\Delta y \tag{4.44}$$

$$\varepsilon_z^e\Big|_{i,j,k+1/2} = [\Delta z_2 \varepsilon_{z2} + (\Delta z - \Delta z_2)\varepsilon_{z1}]/\Delta z \tag{4.45}$$

In this example, the electric field components $E_y|_{i,j+1/2,k+1}$ and $E_z|_{i,j+1,k+1/2}$ are defined inside one material, so they do not require special treatment.

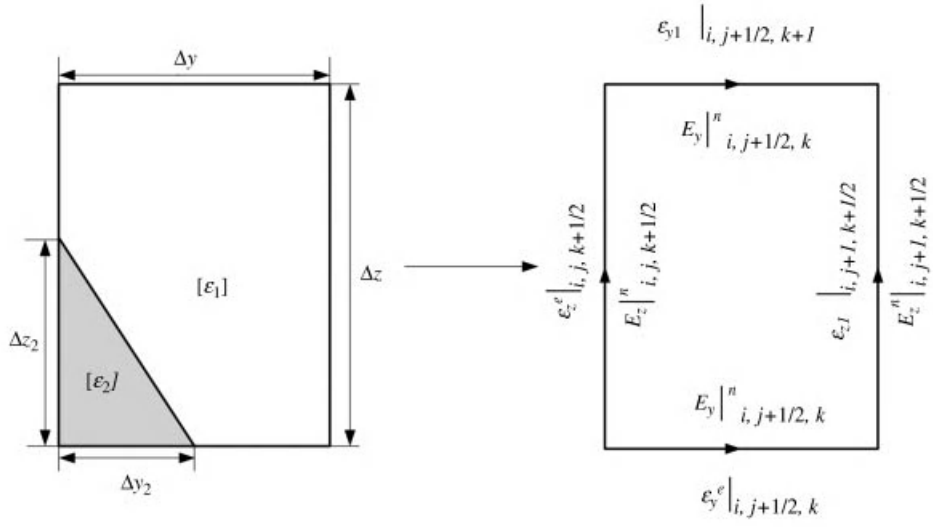


Figure 4.5 The FDTD grid comprising a dielectric boundary.

Problem

4.9 Reconsider the cylindrical PEC cavity resonator in Problem 4.7. Modify the program to employ the CP-FDTD to treat the curved boundary. Use the same parameters as in Problem 4.7 and mesh the whole computational domain by using (60×60) cells. Compare the program with the one in Problem 4.7.

4.1.6 Determination of the Unit Size and the Time Step

In FDTD programming, we need to choose proper spatial and temporal increments, that is, the grid size and the time step. Similar to the meshes in the FEM, the spatial increment in FDTD meshes need to be smaller than 0.1λ, where λ is the wavelength. Since a pulse, which has sufficient bandwidth, is a popular excitation in FDTD, λ should be the minimum wavelength corresponding to the highest frequency we are interested in. More to the point, a larger scatter object normally requires finer FDTD meshes due to the *numerical dispersion* that FDTD suffers. In order to assure the accuracy in the scattering result of a large object, we have to increase the spatial resolution to reduce the accumulative error caused by the numerical dispersion. Once the spatial increments are determined, the time step can be determined accordingly. It has been shown that, in order to assure the stability of the scheme, the time step needs to satisfy the following Courant–Friedrich–Levy (CFL) condition

$$v(\Delta t) \leq 1/\sqrt{\frac{1}{(\Delta x)^2} + \frac{1}{(\Delta y)^2} + \frac{1}{(\Delta z)^2}} \tag{4.46}$$

where v is the speed of the EM wave in the medium. This condition is also called the Courant stability condition. A detailed derivation of the CFL condition can be found in [1] or [10].

4.1.7 Plane Waves in Time Domain

In spherical coordinate system, the time-domain plane wave with incidence direction $\hat{r}(\theta, \phi)$ in a medium can be expressed as

$$\mathbf{E}^i = [E_\theta \hat{\theta} + E_\phi \hat{\phi}] f(t + \mathbf{r}' \cdot \hat{r}/v) \tag{4.47}$$

$$\mathbf{H}^i = \frac{1}{Z}\left[E_\phi \hat{\theta} - E_\theta \hat{\phi}\right] f(t + \mathbf{r}' \cdot \hat{r}/v) \tag{4.48}$$

where $\hat{\theta}$ and $\hat{\phi}$ are the unit vectors in the spherical coordinate; Z is the characteristic impedance of the wave in the space; and v is the phase velocity of the wave in the medium. $\mathbf{r}'$ denotes the coordinate of the discretized point. $f(t)$ can take various forms, such as the sine function, the Gaussian pulse, or the modified Gaussian pulse. If $f(t)$ is a sine function, we will obtain the response on a single frequency after calculation. On the other hand, if $f(t)$ is a Gaussian pulse with the frequency spectra shown in Figure 4.6b, then after calculation we will obtain the response for the frequency band from 0 to f_0. If $f(t)$ is a modified Gaussian pulse with the frequency spectra shown in Figure 4.7b, then after calculation, we will obtain the response for the frequency band from f_1 to f_2. For the convenience of computation, we usually compute in the Cartesian coordinate system, in which the amplitudes of the components of the above incident wave are:

$$E_x = E_\theta \cos\theta \cos\phi - E_\phi \sin\phi \tag{4.49}$$

$$E_y = E_\theta \cos\theta \sin\phi - E_\phi \cos\phi \tag{4.50}$$

$$E_z = -E_\theta \sin\theta \tag{4.51}$$

$$H_x = (E_\theta \sin\phi - E_\phi \cos\theta \cos\phi)/Z \tag{4.52}$$

$$H_y = (-E_\theta \cos\phi + E_\phi \cos\theta \sin\phi)/Z \tag{4.53}$$

$$H_z = -E_\theta \sin\phi/Z \tag{4.54}$$

More specifically, if $f(t)$ is a Gaussian pulse, the x component of the E field can be expressed as

$$E_x^i\big|_{i+1/2,j,k}^n = E_x \exp[-4\pi(t-t_0)^2/\tau^2] \tag{4.55}$$

where

$$t = n\Delta t + \mathbf{r}' \cdot \hat{r}/v \tag{4.56}$$

$$t_0 = \beta \Delta t \tag{4.57}$$

where Δt is the time step, τ determines the width of the pulse in the time domain, and t_0 denotes the time when the pulse is at peak. We will explain the meaning of β later. $\mathbf{r}' \cdot \hat{r}$ can be written in closed form as

$$\mathbf{r}' \cdot \hat{r} = ((i-1)+1/2)\Delta x \cos\phi \sin\theta + (j-1)\Delta y \sin\phi \sin\theta + (k-1)\Delta z \cos\theta \tag{4.58}$$

Note the $1/2$ displacement in (4.58) is the result of considering the specific spatial position where E_x component is defined in the grid (i,j,k). The expression for other components can be written analogously. Another point worth mentioning here is the choice of the parameters in the Gaussian pulse, that is, τ and β. Strictly speaking, the Gaussian pulse is an infinite temporal function. However, with the increase of the time, the signal becomes weaker and weaker so that the signal after the time t_p can be omitted. In terms of computational efficiency, a smaller t_p is preferred. However, a small t_p leads to strong high-frequency components in the numerical result. In this scenario, the spatial discretization becomes inadequate for the high-frequency components, which in turn will decrease the accuracy of the computational result. Therefore, there is a tradeoff in determining the lasting time of the Gaussian pulse. Moreover, this lasting time relates to τ as well. Consequently, the tradeoff also exists in the choice of τ. In practice, $\tau = t_0 = \beta\Delta t$ and duration of $2t_0$ are usually used. Apply Fourier transformation to (4.55) yields

$$\begin{aligned} E_x^i(i,j,k,f) &= E_x \frac{\tau}{2} \exp\left[-j2\pi f t_0 - \pi f^2 \tau^2/4\right] \\ &= E_x \frac{\beta\Delta t}{2} \exp\left[-j2\beta\pi\Delta t f - \pi f^2 (\beta\Delta t)^2/4\right] \end{aligned} \tag{4.59}$$

It is clear that the bandwidth of the Gaussian pulse f_0 is determined by β. If we wish to restrict the energy of the higher frequency to the extent that the maximum amplitude of the frequency component outside the frequency band is no larger than 5% of that within this band, then $f_0 \approx 2/(\beta\Delta t)$. Alternatively, if we target at the maximum amplitude of the frequency component outside the frequency band is no larger than 10% of that within the frequency band, then $f_0 \approx 1.7/(\beta\Delta t)$.

Next, we will calculate the backscattering from a metallic board within frequency band of 0–3 GHz as an example to illustrate the determination of τ and β. The board is of size 29 cm $\times$ 29 cm. For the working frequency of 3GHz, the free-space wavelength is 10 cm. According to the maximum spatial discretization step of one-tenth of a wavelength, the meshing grid needs to be no larger than 1 cm $\times$ 1 cm $\times$ 1 cm. Thus, according to the

Courant stability condition, the time increment cannot exceed $\Delta t_{\max} = 0.0192$ ns. Usually, we can choose $\Delta t = 0.8\Delta t_{\max} = 0.0154$ ns. From $f_0 \approx 2/(\beta\Delta t)$ and $f_0 = 3$ GHz, we can get $\beta = 43.3$. For convenience, we can use $\beta = 40$ instead.

Problem

4.10 Obtain the Gaussian pulse and the modified Gaussian pulse as shown in Figures 4.6 and 4.7. Change the parameters in the mathematical expressions to see the change of the curves.

4.1.8 Calculation of Incident Plane Waves in Time Domain

From the last paragraph of Section 4.1.3, the incident plane wave is added by the modification of the discrete iteration equation on the connection boundary, that is, the connecting condition. This procedure requires the calculation of all the incident fields on all the discretization points on all the time steps. There are various ways to do so. For

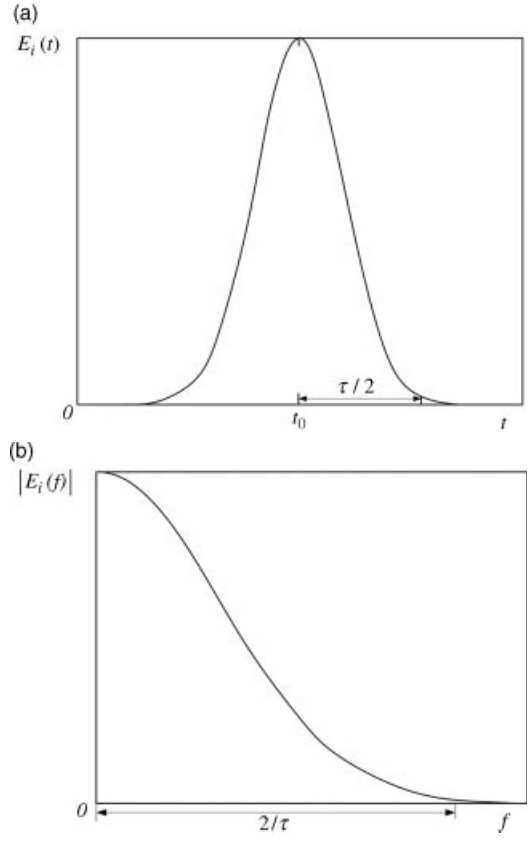


Figure 4.6 Gaussian pulse in (a) the time domain; (b) the frequency domain.

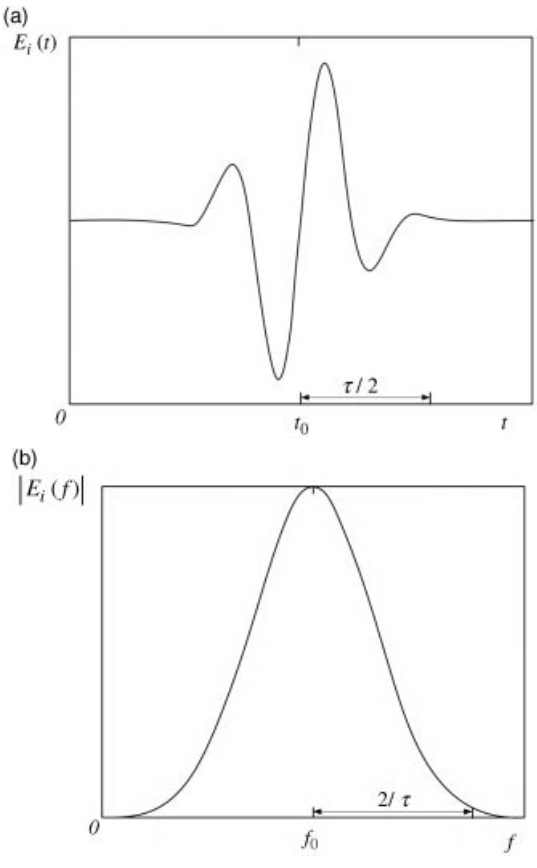


Figure 4.7 Modified Gaussian pulse in (a) the time domain; (b) the frequency domain.

example, we can input all the coordinates and time steps into (4.55) and calculate directly. This way is usually not the most efficient way. Here, we introduce another approach. First, we obtain the 1D FDTD solution of $E_i|_p^n$, $H_i|_{p+1/2}^{n+1/2}$ on the discretization points in a path in the direction of the incidence. We require Δt to be the same as that used in the original 3D FDTD problem. Denote the spatial discretization by Δr, the 1D FDTD iteration equations are

$$E_i|_p^{n+1} = E_i|_p^n - \frac{\Delta t}{\varepsilon \Delta r}\left(H_i|_{p+1/2}^{n+1/2} - H_i|_{p-1/2}^{n+1/2}\right) \tag{4.60}$$

$$H_i|_{p+1/2}^{n+1/2} = H_i|_{p+1/2}^{n-1/2} - \frac{\Delta t}{\mu \Delta r}\left(E_i|_{p+1}^n - E_i|_p^n\right) \tag{4.61}$$

Then, we use projection interpolation method to obtain the incident field on the connection boundary. In the following, we use $E_z|_{i,j,k+1/2}^n$ as an example. First, project

this point on a line passing through the origin and in the direction of the incident wave. The distance of the projection point from the origin is:

$$r = r' \cdot \hat{r} = i\Delta x \sin\theta_i \cos\phi_i + j\Delta y \sin\theta_i \sin\phi_i + (k+1/2)\Delta z \cos\theta_i \tag{4.62}$$

This projection point may fall between two 1D FDTD nodes. For example, $r = (p+w)\Delta r$, $0 < w < 1$, means it is between nodes p and $p+1$. By using a linear interpolation equation, $E_z|^n_{i,j,k+1/2}$ on the boundary is obtained:

$$E_z^n(i,j,k+1/2) = (1-w)E_i|^n_p + wE_i|^n_{p+1} \tag{4.63}$$

Problems

4.11 Based on Problem 4.6, obtain the source function $S(x,y,t)$ with the same frequency as the hard source in Problem 4.4. Run the simulation with this source and compare it with the simulation in Problem 4.4. See how the source is added through the connection boundary and how the total field is confined by the connection boundary.

4.12 Based on Problem 4.10, incorporate PMLs with thickness of 25 layers. Add a PEC cylinder with radius of 50Δ. Test the code again to observe the performance of the PML.

4.1.9 Calculation of the Radar Cross Section

From Section 2.1.9, we know that the far-zone fields can be obtained through the equivalent electric and magnetic current sources $\mathbf{J_s}$ and $\mathbf{M_s}$ on a closed boundary S' (the output boundary) as

$$E_\theta = -ZW_\theta - U_\phi \tag{4.64}$$

$$E_\phi = -ZW_\phi + U_\theta \tag{4.65}$$

where

$$\mathbf{W}(\omega) = j\omega \frac{e^{-j\omega r/c}}{4\pi rc} \int_{S'} \mathbf{J_s}(\omega) \exp(jkr' \cdot \hat{r}) dS' \tag{4.66}$$

$$\mathbf{U}(\omega) = j\omega \frac{e^{-j\omega r/c}}{4\pi rc} \int_{S'} \mathbf{M_s}(\omega) \exp(jkr' \cdot \hat{r}) dS' \tag{4.67}$$

where ω is the angular frequency. Based on this, we are able to use the near-field data obtained in a single FDTD simulation to efficiently and accurately calculate the complete far-field scattering response of an target, or radiation pattern of an antenna [6,12]. In FDTD, this technique is often termed the *near-to-far-field (N/F) transformation*. In the following, we will introduce an efficient "on-the-fly" time-domain approach to

calculate the scattered or radiated E and H fields in the far zone proposed by Luebber *et al.* [13]. The reader is referred to [1] for the near-to-far-field transformation in the phasor domain.

Applying the inverse Fourier transformation to (4.66) and (4.67), we can obtain the following *time-domain far-field scattering field equations*:

$$\mathbf{W}(t) = \frac{1}{4\pi rc}\frac{\partial}{\partial t}\int_{S'} \mathbf{J_s}(t + r' \cdot \hat{r}/c - r/c)dS' \tag{4.68}$$

$$\mathbf{U}(t) = \frac{1}{4\pi rc}\frac{\partial}{\partial t}\int_{S'} \mathbf{M_s}(t + r' \cdot \hat{r}/c - r/c)dS' \tag{4.69}$$

where $\mathbf{J_s} = \hat{n} \times \mathbf{H}(t)$ and $\mathbf{M_s} = -\hat{n} \times \mathbf{E}(t)$ are the equivalent EM currents in time domain. For the convenience of computing $\mathbf{W}(t)$ and $\mathbf{U}(t)$, we normally use a cuboid box as the output boundary. Additionally, since $1/r$ and r/c in (4.68) and (4.69) are constants, for convenience they are suppressed in the following equations. Next, we will explain more about the discretization of (4.68) and (4.69) wherein both differential and integral operations exist. Take the equivalent magnetic current on one sample of the output boundary as an example. Consider Yee's spatial grid (i, j, k). On one of its facets, $\Delta x \Delta z$, the equivalent magnetic current in the z direction is:

$$\mathbf{M}_S \cdot \hat{z} = (-\hat{y} \times E_x \cdot \hat{x}) \cdot \hat{z} = E_x^n \tag{4.70}$$

where E_x^n can be obtained by

$$E_x^n = \left(E_x|_{i+1/2,j,k}^n + E_x|_{i+1/2,j,k+1}^n\right)/2 \tag{4.71}$$

By applying the single-point numerical integration, an additive increment to $\boldsymbol{U}(t)$ induced by this current is expressed as:

$$\Delta\mathbf{U} = \Delta U_z \hat{z} = \frac{\Delta x \Delta z}{4\pi rc}\frac{\partial}{\partial t}\left(E_x^n \hat{z}\right)\Big|_{\text{center of the facet}} \tag{4.72}$$

which is contributable to $\mathbf{U}(t)$ after a time delay, expressed by a fractional number of time step f:

$$f = \frac{\tau}{\Delta t} = \frac{r - \mathbf{r}' \cdot \hat{r}}{c\Delta t} \tag{4.73}$$

where

$$\begin{aligned}\mathbf{r}' \cdot \hat{r} = {} & ((i-1) + 1/2)\Delta x \cos\phi_s \sin\theta_s + (j-1)\Delta y \sin\phi_s \sin\theta_s \\ & + ((k-1) + 1/2)\Delta z \cos\theta_s\end{aligned} \tag{4.74}$$

and (θ_s, ϕ_s) are the spherical coordinates of the observation point (or "field point"). The reference point is chosen as $(\Delta x, \Delta y, \Delta z)$. Approximating the partial derivative in (4.72) by central difference and by considering the delay in ΔU_z yields

$$U_z|^{n+1/2+f} \leftarrow U_z|^{n+1/2+f} + \frac{\Delta x \Delta z}{4\pi c \Delta t}\left(E_x^{n+1} - E_x^n\right) \tag{4.75}$$

In practice, we normally can only use integer array $u_z(i)$ to store the values of U_z on all the time steps. However, $n+1/2+f$ is not necessarily an integer. One way to solve this problem is to obtain an approximated ΔU by using linear interpolation of the two discrete values from the neighboring time steps of $n+1/2+f$. First, we extract the greatest integer in $n+1/2+f$ by using function "*INT*:"

$$m = INT(n+1/2+f) \tag{4.76}$$

Then, the fractional time step between the true time location of the delayed increment and the discrete sample on m can be written as

$$a = n+1/2+f-m \tag{4.77}$$

Thus, the following two equations can be used to replace (4.75) showing the effect of ΔU

$$U_z|^{m} \leftarrow U_z|^{m} + (1-a)\frac{\Delta x \Delta z}{4\pi c \Delta t}\left(E_x^{n+1} - E_x^n\right) \tag{4.78}$$

$$U_z|^{m+1} \leftarrow U_z|^{m+1} + (a)\frac{\Delta x \Delta z}{4\pi c \Delta t}\left(E_x^{n+1} - E_x^n\right) \tag{4.79}$$

Apply (4.78) and (4.79) to all the facets on the output boundary, (4.69) can be computed and the component $U_z(t)$ in the far-field is obtained. Note that the dimension M of the integer array u_z must be greater than the total number of the FDTD time step, because different parts on the output boundary experience different time delays before reaching the field point. To be specific, if the maximum distance between two points on the output boundary is $L\Delta(\Delta = c\Delta t)$, then $M = N+L$ can assure a sufficient integer array u_z to store all the U_z values. After obtaining U_x, U_y, W_x, W_y, W_z in a similar approach, we transform U and W from the Cartesian coordinate system to the spherical one, and obtain the scattered field in the far zone in the time domain by using (4.64) and (4.65). In order to obtain the scattering characteristics in the frequency domain, we need only to apply FFT to the time-domain series of the scattered field and the incident field, and apply division on each frequency point.

4.1.10 Computer Programing and Numerical Examples

In order to illustrate how to write the FDTD program, the hierarchical chart and the flow chart are presented below. As shown in Figure 4.8, level 1 includes four main functions: Reading in the geometry and material information, defining the coefficients,

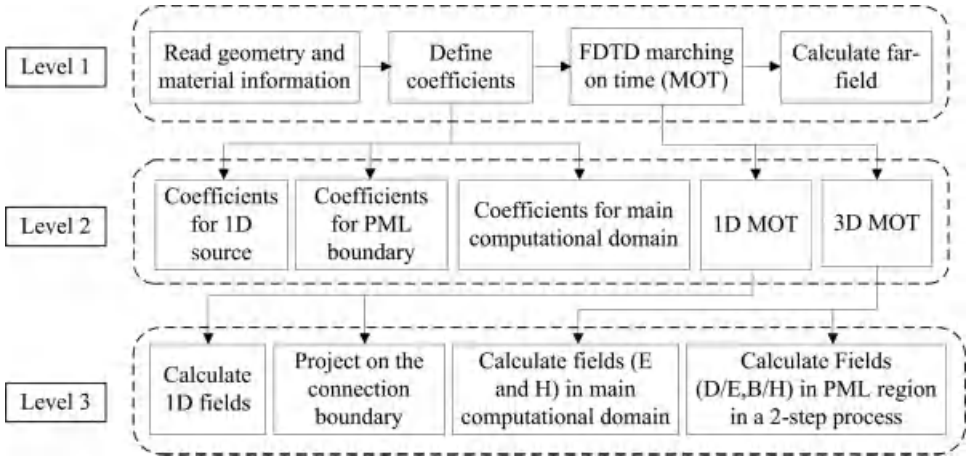
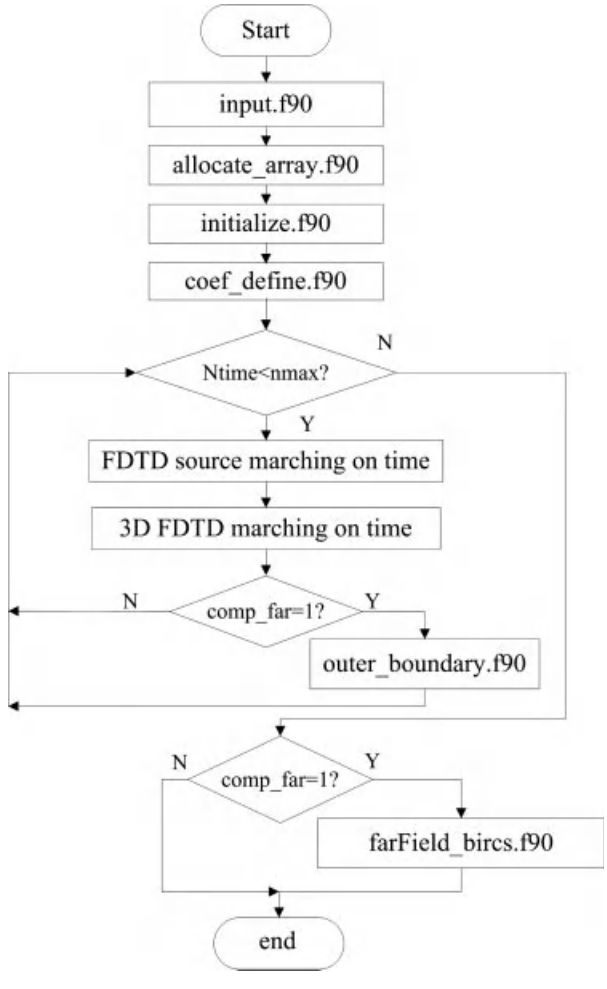


Figure 4.8 Hierarchical chart of the FDTD program.

FDTD marching on time, and calculating the far-field. The function of defining coefficients is further clarified in Level 2, which includes three parts of defining coefficients, respectively, for the 1D source, the PML region, and the main computational domain. The FDTD marching on time in the main loop also includes two parts: the 1D marching for the source, and the 3D marching for the total and scattered fields. Level 3 further explain the functions of the 1D and 3D MOT. The 1D MOT includes not only the 1D source calculation, but also the projection of this source on the connection boundary. The 3D MOT includes the calculation of fields in the main computational domain by a 1-step process and that in the PML region by using the 2-step process with the aid of auxiliary variables (D for E, and B for H). Figure 4.9 shows the flow chart of the main program fdtd_main.f90. The function of the subroutine is listed in the attached table.

In order to show the accuracy of the FDTD, we present a few numerical experiment results in the following. The absorbing boundary condition used in these simulations is the second-order Mur ABC. The scatterer in the first numerical experiment is a metallic cube. In Figure 4.10, the backscattering cross section by normal incident upon the cube is presented. The spatial and the temporal increments in the simulations are $\Delta x = \Delta y = \Delta z = \Delta = 0.01$ m, and $\Delta t = 0.5\Delta/c$, respectively. The time-domain incident wave is a Gaussian pulse given by (4.55), with $t_0 = 80\Delta t$ and $\tau = 3t_0$. The absorbing boundary and the output boundary are placed 0.1 m and 0.05 m away from the metallic surface, respectively. From the figure, it can be seen that the FDTD simulation results agree well with the measured experimental results. The scatterer in the second numerical experiment is a metallic sphere, with diameter of 1.0 m. Figure 4.11 presents the comparison of the simulated backscattering with the analytical result obtained by a Mie series. The dashed and the solid lines denote the simulations without and with contour treatment. It can be seen that the contour treatment can significantly improve



Notes for the figure

input.f90	Read the geometry and material information. Determine dt and nmax.
allocate_array.f90	Array allocation.
initialize.f90	Initialize all the variables.
coef_define.f90	Calculate the coefficients for the FDTD marching on time.
FDTD source marching on time	Calculate the 1-D source (E and H) of each time step, and project the 1-D fields onto the connection boundary.
3D FDTD marching on time	Calculate the E and H field in the whole computational domain.
outer_boundary.f90	Calculate the Fourier transformed field values on outer boundary to compute the RCS.
farField_bircs.f90	Calculate the bistatic RCS.
Index: comp_far	*1: Compute bistatic RCS, 2:Not to do so.*

Figure 4.9 Flow chart of the FDTD main program.

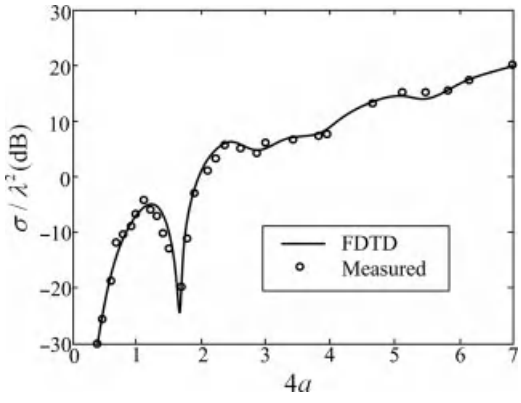


Figure 4.10 The measured and simulated backscattering cross section of the metallic cube.

the accuracy of the result, especially in the high-frequency band. In the third example, we simulated the scattering from a coated PEC sphere. The sphere has a diameter of 1.0 m. The coating material has a relative permittivity of 5.6 and thickness of 0.08 m. The simulated bistatic RCS is compared with the analytical result obtained by a Mie series in Figure 4.12. As with Figure 4.11, the dashed and the solid lines donate the simulations without and with contour treatment. The significant improvement in the accuracy by using the contour treatment is verified again.

Problems

4.13 Repeat the numerical experiments in Section 4.1.10 and obtain the FDTD result as shown in Figure 4.10.

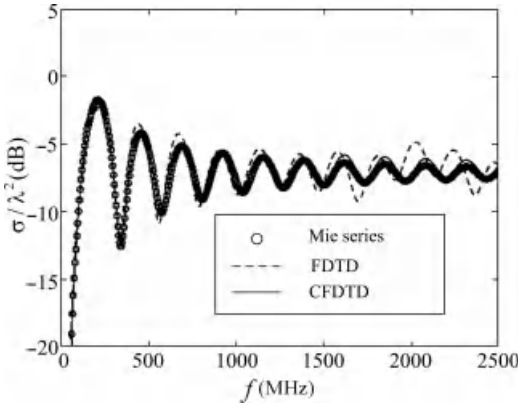


Figure 4.11 Comparison of the backscattering cross section of the metallic sphere obtained by Yee's scheme, CFDTD, and by the analytical method.

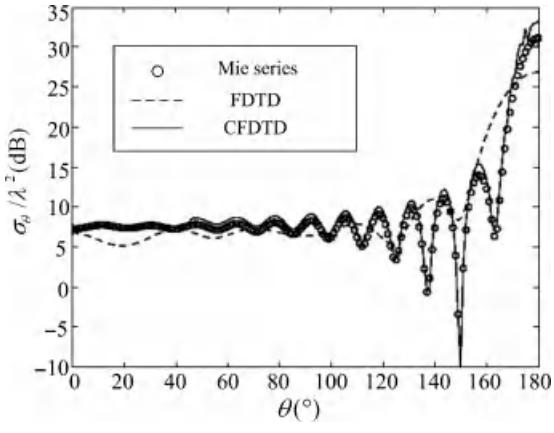


Figure 4.12 Comparison of the bistatic RCS of the sphere with cladding obtained by Yee's scheme, CFDTD, and by the analytical method.

4.14 Repeat the numerical experiments in Section 4.1.10 and obtain the FDTD result as shown in Figure 4.11.

4.15 Repeat the numerical experiments in Section 4.1.10 and obtain the FDTD result as shown in Figure 4.12.

4.2 Treatment for Special Problems

Through the above introduction of the FDTD procedure in solving the 3D scattering problem, it can be seen that the FDTD method is a simple and versatile method. After the model is set up, the implementation is not complex either. However, for specific problems, there are some useful techniques if we want to make the most of it. Most of these techniques can be found in [1]. For illustration, we only briefly introduce the techniques corresponding to three specific problems below.

4.2.1 Treatments for Thin Metallic Wires

The thin wire referred to here is the wire structure whose radius is smaller than the grid dimension in FDTD. This kind of structure is often seen in microwave engineering. For example, the wire antenna is a thin-wire structure. If the radius of the thin wire is so small that it can be omitted, then we only need to force the tangential electric field components on the thin wire to be zero in order to apply conventional Yee's iteration equations. If the radius of the wire cannot be omitted, then the size of the wire raidus needs to be taken into consideration and more manipulation is needed in FDTD equations. For example, we can choose the same treatment techniques as in the CP-FDTD introduced for conformal surfaces. Alternatively, we can use a more accurate technique as introduced below. According to the physical characteristics of the

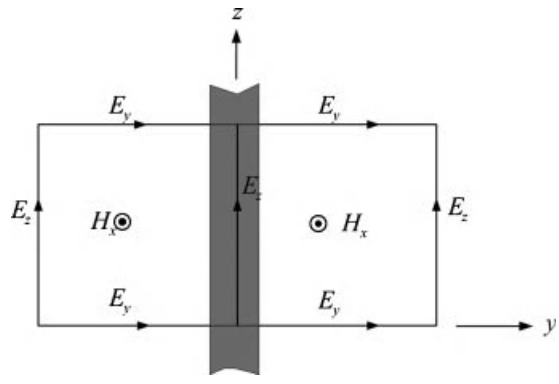


Figure 4.13 Treatment for thin wire.

scattered field in the near-zone, the circular and radial electric field components near the thin wire varies as $1/r$, where, r is the distance between the center of the wire and the field point. For example, if the wire is aligned along the z direction, as illustrated in Figure 4.13, the field around it can be expressed as

$$E_z(i\Delta x, j\Delta y, z) = 0 \tag{4.80}$$

$$E_y(i\Delta x, y, k\Delta z) = E_y(i\Delta x, (j+1/2)\Delta y, k\Delta z)\frac{\Delta y/2}{y} \tag{4.81}$$

$$H_x(i\Delta x, y, (k+1/2)\Delta z) = H_x(i\Delta x, (j+1/2)\Delta y, (k+1/2)\Delta z)\frac{\Delta y/2}{y} \tag{4.82}$$

Inserting the above expressions into the integral form of Faraday's law yields

$$\begin{aligned} &-\mu_0 \frac{H_x^{n+1/2}(i,j+1/2,k+1/2) - H_x^{n-1/2}(i,j+1/2,k+1/2)}{\Delta t}\Delta z \frac{\Delta y}{2}\int_{r_0}^{\Delta y}\frac{1}{y}dy \\ &= -\left[E_y^n(i,j+1/2,k+1) - E_y^n(i,j+1/2,k)\right]\frac{\Delta y}{2}\int_{r_0}^{\Delta y}\frac{1}{y}dy + E_z^n(i+1,j,k+1/2)\Delta z \end{aligned} \tag{4.83}$$

where r_0 is the radius of the wire. Rewriting the above expression we get

$$\begin{aligned} H_x^{n+1/2}(i,j+1/2,k+1/2) = {} & H_x^{n-1/2}(i,j+1/2,k+1/2) \\ & + \frac{\Delta t}{\mu_0 \Delta z}\left[E_y^n(i,j+1/2,k+1) - E_y^n(i,j+1/2,k)\right] \\ & - \frac{\Delta t}{\mu_0 \Delta y}\frac{2}{\ln(\Delta y/r_0)}E_z^n(i+1,j,k+1/2) \end{aligned} \tag{4.84}$$

This is the marching-on-time equation for H_x in the FDTD grid containing the thin wire. The other H components can be obtained in a similar approach. As with the contour treatment, since the thin-wire effect is fully accounted for in the MOT equations for updating H field, the conventional updating equations for E fields can be used without any modification.

4.2.2 Treatments for Dispersive Media

As was introduced in Chapter 1, in the dispersive media the constitutive parameters are functions of the working frequency. The plasma, which exists in the real world, and the PML introduced in Section 4.1.2, which exists only in numerical simulations, are all examples of dispersive media. For convenience, here we use an example with frequency-dependent permittivity is to illustrate the FDTD treatment for dispersive media. In contrast to nondispersive media, the partial derivative of the electric displacement vector $\vec{D}$ cannot be converted straightforwardly into a partial derivative of the electric field. This is because the relative permittivity is frequency dependent and hence its partial derivative with respect to time is no longer zero. In other words, only

$$\nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t} + \sigma \mathbf{E} \tag{4.85}$$

holds, but

$$\nabla \times \mathbf{H} = \varepsilon \frac{\partial \mathbf{E}}{\partial t} + \sigma \mathbf{E} \tag{4.86}$$

does not. Consequently, if we use Yee's grids to discretize (4.85), we can only obtain the updating equations for **D** by using **H**, but not the ones to update **E** by using **H**. Therefore, we need to further set up the relationship of **D** and **E** in the time domain. This relationship changes with the form of the frequency-dependent function describing the relative permittivity. Even for the same function, there are many ways to implement this. For example, we can use the piecewise-linear recursive convolution method (PLRC). Alternatively, the *auxiliary differential equation method* (ADE) can be applied to develop a simple time-marching scheme without using convolution. The PML introduced in Section 4.1.2 is one of the examples using the ADE method. In the following, we will show another example. In the medium with dielectric constant satisfying the Debye form,

$$\varepsilon(\omega) = \varepsilon_0[\varepsilon_\infty + \chi(\omega)] = \varepsilon_0\left[\varepsilon_\infty + \frac{\varepsilon_S - \varepsilon_\infty}{1 + j\omega\tau_0}\right] \tag{4.87}$$

where ε_s is the static relative permittivity, ε_∞ is the relative permittivity at infinite frequency, and τ_0 is the pole relaxation time. So the relationship between **D** and **E** in

frequency is

$$\mathbf{D}(\omega) = \varepsilon_0[\varepsilon_\infty + \chi(\omega)]\mathbf{E}(\omega) = \varepsilon_0\left[\varepsilon_\infty + \frac{\varepsilon_S - \varepsilon_\infty}{1 + j\omega\tau_0}\right]\mathbf{E}(\omega) \tag{4.88}$$

Alternatively, it can be written as

$$\mathbf{D}(\omega) + j\omega\tau_0\mathbf{D}(\omega) = \varepsilon_0\varepsilon_s\mathbf{E}(\omega) + j\omega\tau_0\varepsilon_0\varepsilon_\infty\mathbf{E}(\omega) \tag{4.89}$$

Applying the inverse Fourier transformation to (4.89) yields

$$\mathbf{D} + \tau_0\frac{\partial\mathbf{D}}{\partial t} = \varepsilon_0\varepsilon_S\mathbf{E}(\omega) + \tau_0\varepsilon_0\varepsilon_\infty\frac{\partial\mathbf{E}}{\partial t} \tag{4.90}$$

If we discretize the above equation at time $t = (n + 1/2)\Delta t$ we obtain

$$\frac{D^{n+1} + D^n}{2} + \tau_0\frac{D^{n+1} - D^n}{\Delta t} = \varepsilon_0\varepsilon_S\frac{E^{n+1} + E^n}{2} + \tau_0\varepsilon_0\varepsilon_S\frac{E^{n+1} - E^n}{\Delta t} \tag{4.91}$$

Further, we obtain the relationship between **D** and **E** in the time domain:

$$E^{n+1} = aE^n + bD^{n+1} - cD^n \tag{4.92}$$

where

$$a = \frac{\dfrac{\tau_0\varepsilon_0\varepsilon_\infty}{\Delta t} - \dfrac{\varepsilon_0\varepsilon_S}{2}}{\dfrac{\tau_0\varepsilon_0\varepsilon_\infty}{\Delta t} + \dfrac{\varepsilon_0\varepsilon_S}{2}} \tag{4.93}$$

$$b = \frac{\dfrac{\tau_0}{\Delta t} + \dfrac{1}{2}}{\dfrac{\tau_0\varepsilon_0\varepsilon_\infty}{\Delta t} + \dfrac{\varepsilon_0\varepsilon_S}{2}} \tag{4.94}$$

$$c = \frac{\dfrac{\tau_0}{\Delta t} - \dfrac{1}{2}}{\dfrac{\tau_0\varepsilon_0\varepsilon_\infty}{\Delta t} + \dfrac{\varepsilon_0\varepsilon_S}{2}} \tag{4.95}$$

Problems

4.16 Consider a Debye dispersive medium described by (4.88). Derive the ADE-FDTD time-marching scheme based on (4.91).

4.17 Consider a Debye dispersive medium described by

$$\varepsilon(\omega) = \varepsilon_0\left[\varepsilon_\infty + \sum_{p=1}^{2}\chi_p(\omega)\right] = \varepsilon_0\left[\varepsilon_\infty + \sum_{p=1}^{2}\frac{\varepsilon_{S,p}-\varepsilon_{\infty,p}}{1+j\omega\tau_p}\right]$$

where $\varepsilon_{s,p}$ $\varepsilon_{\infty,p}$ and τ_p correspond to the specific Debye pole (p). Derive the ADE-FDTD time-marching scheme using the central-difference approximation.

4.18 Consider a Lorentz dispersive medium described by

$$\varepsilon(\omega) = \varepsilon_0\left[\varepsilon_\infty + \chi_p(\omega)\right] = \varepsilon_0\left[\varepsilon_\infty + \frac{\left(\varepsilon_{S,p}-\varepsilon_{\infty,p}\right)\omega_p^2}{\omega_p^2 + 2j\omega\delta_p - \omega^2}\right]$$

where $\chi_p(\omega)$ has one single pole pare, $\varepsilon_{s,p}$ $\varepsilon_{\infty,p}$ are the relative permittivity at zero and infinite frequency due to the pole pair, and δ_p is the damping coefficient. By introducing the auxiliary variable

$$D(\omega) = \varepsilon_0\frac{\left(\varepsilon_{S,p}-\varepsilon_{\infty,p}\right)\omega_p^2}{\omega_p^2 + 2j\omega\delta_p - \omega^2}E(\omega)$$

and using the central-difference approximation

$$\frac{d^2\mathbf{D}(n\Delta t)}{dt^2} \cong \frac{\mathbf{D}|_{n-1} - 2\mathbf{D}|_n + \mathbf{D}|_{n+1}}{\Delta t^2}$$

derive the ADE-FDTD time-marching scheme for E.

4.2.3 Treatments for Lumped Elements

The lumped-element model of electronic circuits makes the simplifying assumption that each element is a finite point in space, and that the wires connecting elements are perfect conductors. The lumped-element model is valid whenever the circuit's characteristic length is much smaller than the operating wavelength. So the element's characteristics can be described by the relationship between the voltage and the current. The resistor, the capacitor, the inductor, the diode, and the triode, are all examples of lumped elements. In this section, we will discuss the modification to the Yee's formulation for the lumped-element model. Suppose the size of the lumped element is smaller than one FDTD grid. According to Sui [14], the effect of the lumped element can be accounted for in Maxwell's equations by adding a lumped electric current density term J_L:

$$\nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t} + \mathbf{J_L} \tag{4.96}$$

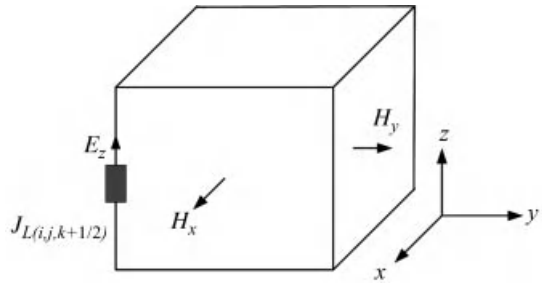


Figure 4.14 Treatment for the lumped-element model.

Without loss of generality, we assume the lumped element is z-oriented, so $\mathbf{J_L}$ is along the z-axis, as illustrated in Figure 4.14. Then, $\mathbf{J_L}$ is related to the total electric current I_L by

$$\mathbf{J_L} = \frac{I_L}{\Delta x \Delta y}\hat{z} \tag{4.97}$$

Thus, the z component in (4.96) can be discretized as

$$E_z\big|_{i,j,k+1/2}^{n+1} = E_z\Big|_{i,j,k+1/2}^{n} + \frac{\Delta t}{\varepsilon}(\nabla \times \mathbf{H})_z\big|_{i,j,k+1/2}^{n+1/2} - \frac{\Delta t}{\varepsilon \Delta x \Delta y} I_L\big|_{i,j,k+1/2}^{n+1/2} \tag{4.98}$$

Then, according to the relationship between the voltage across the lumped elements and the current through it, and the relationship between the voltage and the electric field, we can obtain the modified updating equations. For example, for the resistor R located at E_z position, the current flowing through R can be expressed as

$$I_L\big|_{i,j,k+1/2}^{n+1/2} = \frac{U_L\big|_{i,j,k+1/2}^{n+1/2}}{R} \tag{4.99}$$

where U_L is the voltage across R, which is related to the electric field by

$$\begin{aligned} U_L\big|_{i,j,k+1/2}^{n+1/2} &= \Delta z E_z\big|_{i,j,k+1/2}^{n+1/2} \\ &= \frac{\Delta z}{2}\left(E_z\big|_{i,j,k+1/2}^{n+1} + E_z\big|_{i,j,k+1/2}^{n}\right) \end{aligned} \tag{4.100}$$

Inserting (4.99) and (4.100) into (4.98) yields

$$E_z\big|_{i,j,k+1/2}^{n+1} = aE_z\big|_{i,j,k+1/2}^{n} + b(\nabla \times \mathbf{H})_z\big|_{i,j,k+1/2}^{n+1/2} \tag{4.101}$$

where

$$a = \frac{1 - \dfrac{\Delta t \Delta z}{2R\varepsilon \Delta x \Delta y}}{1 + \dfrac{\Delta t \Delta z}{2R\varepsilon \Delta x \Delta y}} \tag{4.102}$$

$$b = \frac{\dfrac{\Delta t}{\varepsilon}}{1 + \dfrac{\Delta t \Delta z}{2R\varepsilon \Delta x \Delta y}} \tag{4.103}$$

Equation 4.101 is the modified Yee algorithm for the resistor's modeling. The modified FDTD equations for other lumped elements can be obtained in a similar approach.

4.3 Comparison of the MoM, FEM and FDTD Methods

It has been clearly stated that the above three methods differ essentially in the mathematic equations that are discretized. To be more specific, MoM is to discretize the integral equations of EM problems, FEM is to discretize the variational formulations of functional for EM problems, while FDTD is to directly discretize the Maxwell's equations in the time domain. In this section, we will further discuss the difference of the numerical characteristics due to the distinct mechanisms of discretization.

First, let us check the interaction of two unknown variables x, y. In MoM, the Green's function is used to directly describe this interaction. This is a rigorous description. However, in FEM and FDTD, this interaction is passed through the interaction of a series of intermediate variables. In other words, x interacts with the neighboring unknown variable d_1, then d_1 interacts with its neighboring unknown variable d_2. The interaction is further relayed between adjacent variables through all possible passes before finally reaching y. This is an approximate description. The more intermediate variables that are involved, the larger error will be introduced. This error is the numerical dispersion introduced in Section 4.1.6. Both FEM and FDTD suffer numerical dispersion, while MoM does not. On the other hand, since any two variables in MoM interact directly, the corresponding discretized matrix is a full matrix. However, in FEM and FDTD, only neighboring variables interact directly. Therefore, the FEM matrix is sparse. The equivalent matrix of the FDTD time-marching equations is also sparse. From this analysis, it can be seen that in terms of numerical characteristics, FEM and FDTD are similar, but they are quite different from MoM. That is because the variational formulation that FEM discretizes is essentially a partial differential equation.

Next, let us examine the numerical property of the discretized equations and the solution of the three numerical methods. By using a basic FDTD method, we do not need to solve a system of equations. What we do is to update fields in the propagation of the EM waves by time-marching equations. The total time step number is determined by the situation of the propagating waves, and we cannot increase or decrease it. For example, in the time-harmonic case, as long as the electric field in the observation point reaches a stable status, we can terminate the simulation. So in the following, we will only compare the FEM and MoM methods. From the expression of the Green's function, the closer the two points are, the stronger their interaction will be. This is reflected in the MoM matrix by the relatively large absolute values in the diagonal entries. This characteristic results in much better condition number in MoM than in FEM matrices, which further leads to a much faster convergence in the MoM method if iteration is used in the solution. Since a variety of techniques such as FFT and MLFMA can be used to greatly reduce the complexity of the matrix-vector product in MoM, iteration methods currently dominate in solving the MoM discretized equations. Although the FEM matrix is sparse, we normally avoid using an iteration method due to the poor condition number. If the computer memory allows, direct methods such as the multifrontal method [15] are good choice.

Next, we are going to compare the complexity in the implementation of these three methods. It is clearly shown that in MoM implementation, complex integral equations are involved. More to the point, we need to choose the proper basis functions, treat the singularity carefully and apply various techniques to accelerate the solution process. In contrast, to implement FEM is relatively easy. We only need to pay more attention to the selection of proper basis function and proper storage manner for the sparse matrix. As to the FDTD, apparently it is the easiest one to use.

In terms of versatility, FEM and FDTD are similar and quite good. Comparatively, MoM is less versatile. For example, if we consider a scattering problem, we need to write fairly different MoM programs for metallic target, homogeneous dielectric target and in homogeneous dielectric target respectively. However, these three kinds of scattering problems can be easily solved in the same FEM or FDTD program. On the other hand, versatility is traded for high accuracy and high efficiency in MoM. Although in theory these three methods bear similar accuracy, practical numerical experiments show that MoM provides the best accuracy in the simulated result, which is followed by FEM and then by FDTD, especially for electrically large models. That is because MoM does not suffer numerical dispersion.

References

1. Taflove, A. (1995) *Computational Electrodynamics: The Finite Difference Time Domain Approach*, Artech House, Norwood, MA.
2. Mur, G. (1981) Absorbing boundary conditions for the finite-difference approximation of the time-domain electromagnetic field equations. *IEEE Transactions Electromagnetic Compatibility*, **23**, 377–382.

3. Berenger, J. (1994) A perfectly matched layer for the absorption of electromagnetic waves. *Journal of Computational Physics*, **114**(10), 185–200.
4. Chew, W.C. and Weedon, W. (1994) A 3D perfectly matched medium from modified Maxwell's equations with stretched coordinates. *Microwave and Optics Technology Letters*, **13**, 599–604.
5. Gedney, S.D. (1996) An anisotropic perfectly matched layer-absorbing medium for the truncation of FDTD lattices. *IEEE Transactions on Antennas and Propagation*, **44**(12), 1630–1639.
6. Umashankar, K. and Taflove, A. (1982) A novel method to analyze electromagnetic scattering of complex objects. *IEEE Transactions Electromagnetic Compatibility*, **24**, 397–405.
7. Wilcox, C.H. (1956) An expansion theorem for electromagnetic fields. *Communications on Pure and Applied Mathematics*, **M**, 115–134.
8. Kashiwa, T. and Fukai, I. (1990) A treatment by FDTD method of dispersive characteristics associated with electronic polarization. *Microwave and Optics Technology Letters*, **3**, 203–205.
9. Joseph, R.M., Hagness, S.C., and Taflove, A. (1991) Direct time integration of Maxwell's equations in linear dispersive media with absorption for scattering and propagation of femtosecond electromagnetic pulses. *Optics Letters*, **16**, 1412–1414.
10. Ge, D.B. and Yan, Y.B. (2002) *Finite Difference Time Domain Method for Electromagnetic Waves*, Xi-Dian University Press, Xi-An, (in Chinese).
11. Jurgens, T.G., Taflove, A., Umashankar, K., and Moore, T.G. (1992) Finite- difference time-domain modeling of curved surfaces. *IEEE Transactions on Antennas and Propagation*, **40**, 357–366.
12. Taflove, A. and Umashankar, K. (1983) Radar cross section of general three-dimensional structures. *IEEE Transactions on Electromagnetic Compatibility*, **25**, 433–440.
13. Luebbers, R.J., Kunz, K.S., Schneider, M., and Hunsberger, F. (1991) A finite-difference time-domain near zone to far zone transformation. *IEEE Transactions on Antennas and Propagation*, **39**, 429–433.
14. Sui, W., Christensen, D.A., and Durney, C.H. (1992) Extending the two dimensional FDTD method to hybrid electromagnetic systems with active and passive lumped elements. *IEEE Transactions on Microwave Theory and Techniques*, **40**(4), 724–730.
15. Liu, J.W.H. (1992) The multifrontal method for sparse matrix solution: theory and practice. *SIAM Review*, **34**, 82–109.

5

Hybrid Methods

There are many numerical methods for the computation of EM problems. These methods can be grouped roughly into two categories: high-frequency asymptotic method, and full-wave numerical method. The physical optics method (PO), geometrical optics method (GO), geometrical theory of diffraction (GTD), physical theory of diffraction (PTD), and the shooting-bouncing ray method (SBR) are all examples of the former group. MoM, FEM and FDTD are the most popular examples of the latter group. These methods are capable of independently solving a wide range of EM problems. However, these methods are distinct in terms of numerical performance. Each method has its own advantages and disadvantages. From time to time, we face the following embarrassing situation when dealing with particular practical EM problems. Although we can solve the problem by using a single method, the solution suffers either inaccuracy or low efficiency. On the other hand, we are usually able to construct a problem-tailored solution scheme by combining the advantage of different numerical methods, which provides accurate and efficient result. This hybrid method will be discussed in this chapter. Accurately speaking, the hybrid method is not a specific approach or technique. It is rather an idea towards the best numerical solution. This idea suits not only for computational electromagnetics, but also for other fields in science. However, in different scientific fields, people practice this idea in diverse ways. In CEM, two types of ways exist in the hybridization. In the first type, the computational domain is partitioned into different regions and different schemes are applied to each region. The hybrid finite element – boundary integral (FE-BI) is one of the examples. The foundation of this type of hybridization is the *equivalence principle*. In the second type, the dimension in the EM problem is separated and we apply different method for each dimension according to mode theory or variable seperation theory. The straight-line method belongs to this kind of hybrid method. In this chapter, we present the scheme combining the high-frequency asymptotic method

Essentials of Computational Electromagnetics, First Edition. Xin-Qing Sheng and Wei Song.
© 2012 John Wiley & Sons Singapore Pte. Ltd. Published 2012 by John Wiley & Sons Singapore Pte. Ltd.

and the full-wave numerical method in the first section, and the hybrid full-wave numerical methods in the second section.

5.1 Hybrid High-Frequency Asymptotic Methods and Full-Wave Numerical Methods

As we know, the high-frequency asymptotic methods in CEM have the advantage of high efficiency and low requirement on computer memory. However, usually the high-frequency method is only suitable for dealing with electrically large metallic objects. On the other hand, in principle the full-wave numerical methods can treat arbitrary objects, but usually they are not suitable for electrically very large problems due to the low efficiency and high computer memory requirement. Various techniques have been developed towards faster solutions, such as the fast multipole algorithm, the fast direct solvers for sparse matrices, and the techniques employing the fast Fourier transformation. With these techniques and with the development of computer technology such as the parallel computing systems, the capability of the full-wave numerical methods in terms of the electrical size of the target is increasing rapidly. However, there are still many practical applications that cannot be efficiently solved by solely using any one of the full-wave methods. Therefore, we still need to come up with hybrid methods for an effective solution to these problems. For example, we can combine the high-frequency asymptotic methods with the full-wave numerical methods. This idea can be traced back into the middle of the 1970s. Thiele's work [1] is one of the early works. Jakobus's work [2] and Jin's work [3–5] are the recent examples. Although they are different in schemes and skills, they share the same basic idea. In this section, we will show this basic idea and detailed procedures in the implementation by two examples.

5.1.1 Hybird Physical Optics Method and FEM

In engineering applications, we often encounter targets with cracks or cavities. The cracks or cavities usually affect considerably the scattering of the whole target. In this section, we will present the procedure to calculate the scattering from this kind of target by using a hybrid method combining the physical optics (PO) method and FEM (PO-FEM).

5.1.1.1 Idea

Figure 5.1a shows an electrically large PEC scatterer with a dielectric-filled cavity. The size of the PEC body could be tens or even hundreds of wavelengths, while the size of the dielectric-filled cavity is only of a few wavelengths. For such an object, by solely using the high-frequency asymptotic method or full-wave numerical method, we cannot get an efficient and accurate solution. Therefore, we need to look for a

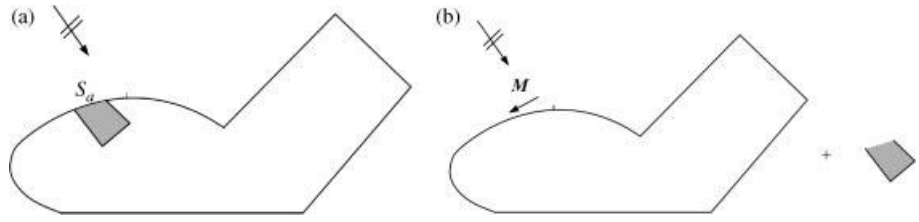


Figure 5.1 Hybrid method for solving the scattering from an electrically large PEC scatterer with a dielectric-filled cavity.

combined method. The idea here is to partition the computational domain into two parts by using the equivalent principle: the cavity and the PEC body, as shown in Figure 5.1b. Exterior to the cavity, we apply the second form of equivalent principle, as introduced in Section 1.3.1. In other words, we replace the cavity with PEC and calculate the scattering of the electrically large PEC with the coexistence of the incident wave and equivalent magnetic current **M**. Thus, the high-frequency asymptotic method can be applied. Inside the cavity, we apply FEM to obtain an accurate solution.

5.1.1.2 The Physical Optics Solution in the Region Exterior to the Cavity

According to the above analysis, the scattering field outside the cavity is made up of two parts: one is induced by the incident wave, the other is due to the equivalent magnetic current **M**. Because the scatterer now becomes an electrically large PEC body without any cavity, the high-frequency asymptotic method can be applied to calculate the scattering field induced by the incident field. In the following, we use the PO method as an example.

$$\mathbf{H}^{po} = -\int_S \mathbf{J}_{po} \times \nabla G_0 \, dS' \tag{5.1}$$

where G_0 is the Green's function in free space, and

$$\mathbf{J}_{po} = \begin{cases} 2\hat{n} \times \mathbf{H}^i(\mathbf{r}), & \mathbf{r} \in S_{lit} \\ 0, & \mathbf{r} \in S_{dark} \end{cases} \tag{5.2}$$

The scattering field arising from the equivalent magnetic current **M** can be expressed as

$$\mathbf{H}^m = -jk_0 Y_0 \int_{S_a} \underline{\underline{G}} \cdot \mathbf{M} \, dS' \tag{5.3}$$

where $\underline{\underline{G}}$ is the dyadic Green's function for the PEC body. Usually, we do not know the analytical expression. However, since the equivalent magnetic current is normally on a smooth curvature with large radius, we can approximate $\underline{\underline{G}}$ by the *half-space dyadic Green's function* $\underline{\underline{G}}_h$, that is,

$$\underline{\underline{G}}_h(\mathbf{r},\mathbf{r}') = \underline{\underline{G}}_0(\mathbf{r},\mathbf{r}') - \underline{\underline{G}}_0(\mathbf{r},\mathbf{r}'_i) + 2\hat{n}_a\hat{n}_a G_0(\mathbf{r},\mathbf{r}'_i) \tag{5.4}$$

where $\mathbf{r}'_i$ denotes the mirror image of point $\mathbf{r}'$, $\mathbf{n}_a$ denotes the outward unit normal vector of S_a, and $\underline{\underline{G}}_0$ is the dyadic Green's function of free space.

$$\underline{\underline{G}}_0(\mathbf{r},\mathbf{r}') = \left(\underline{\underline{I}} - \frac{1}{k_0^2}\nabla\nabla'\right)G_0(\mathbf{r},\mathbf{r}') \tag{5.5}$$

Thus, the total scattering field outside the scatterer is

$$\begin{aligned}\mathbf{H} &= \mathbf{H}^{po} + \mathbf{H}^{m} \\ &= \mathbf{H}^{po} - jk_0Y_0\int_{S_a}\underline{\underline{G}}\cdot\mathbf{M}\,dS'\end{aligned} \tag{5.6}$$

5.1.1.3 The FEM Solution Inside the Cavity

As is known, the equation of the field inside the cavity satisfies is equivalent to the variation of the following functional:

$$\begin{aligned}F(\mathbf{E}) = \frac{1}{2}\int_V\left[\frac{1}{\mu_r}(\nabla\times\mathbf{E})\cdot(\nabla\times\mathbf{E}) - k_0^2\varepsilon_r\mathbf{E}\cdot\mathbf{E}\right]dV \\ + jk_0Z_0\int_{S_a}\hat{n}\times\mathbf{E}\cdot\mathbf{H}\,dS\end{aligned} \tag{5.7}$$

where V denotes the interior domain of the cavity, and S_a denotes the section through which the cavity is open to free space. Substituting (5.6) into (5.7) yields

$$\begin{aligned}F(\mathbf{E}) = \frac{1}{2}\int_V\left[\frac{1}{\mu_r}(\nabla\times\mathbf{E})\cdot(\nabla\times\mathbf{E}) - k_0^2\varepsilon_r\mathbf{E}\cdot\mathbf{E}\right]dV \\ - k_0^2\int_{S_a}\hat{n}\times\mathbf{E}\cdot\left[\int_{S_a}\underline{\underline{G}}\cdot\hat{n}\times\mathbf{E}\,dS'\right]dS \\ + jk_0Z_0\int_{S_a}\hat{n}\times\mathbf{E}\cdot\mathbf{H}^{po}dS\end{aligned} \tag{5.8}$$

Mesh the domain V into tetrahedral elements, and express the fields in each element by using the vector basis functions as introduced in Chapter 4, then we can obtain the field expression as

$$\mathbf{E} = \sum_{j=1}^{N} \mathbf{N}_j E_j \tag{5.9}$$

where N denotes the total number of the unknowns, which equals the number of the edges. E_j are the interpolation unknowns, which are the tangential electric field components, and $\mathbf{N}_j$ are the vector basis functions. Substituting (5.9) into (5.8) and applying the variation principle leads to

$$[K]\{E\} = \{b\} \tag{5.10}$$

where

$$\begin{aligned} K_{ij} = & \frac{1}{2}\int_V \left[\frac{1}{\mu_r}(\nabla \times \mathbf{N}_i)\cdot(\nabla \times \mathbf{N}_j) - k_0^2 \varepsilon_r \mathbf{N}_i \cdot \mathbf{N}_j\right] dV \\ & - k_0^2 \int_{S_a} \hat{n} \times \mathbf{N}_i \cdot \left[\int_{S_a} \underline{\underline{G}} \cdot \hat{n} \times \mathbf{N}_j \, dS'\right] dS \end{aligned} \tag{5.11}$$

$$b_j = -jk_0 Z_0 \int_{S_a} \hat{n} \times \mathbf{N}_j \cdot \mathbf{H}^{po} dS \tag{5.12}$$

Solve (5.10) and we can get the field distribution of the field inside the cavity. Further, we can obtain the equivalent magnetic current $\mathbf{M}$, from which the scattering in the far-field can be calculated.

5.1.1.4 Calculation of the Scattering in the Far-Field

The scattering in the far-field is also made up of two parts: one is scattered by the incident wave, the other is due to the equivalent magnetic current $\mathbf{M}$. The former can be obtained in the same way as introduced in Section 5.1.1.2. In principle, the latter can be calculated by using (5.3). However, we need to be aware that we are not equipped with an analytical expression of the dyadic Green's function $\underline{\underline{G}}$. If we still use the half-space dyadic Green's function $\underline{\underline{G}}_h$ to approximate $\underline{\underline{G}}$, as we do in the calculation of $\mathbf{M}$, the result may suffer large error. Because by using $\underline{\underline{G}}_h$, the reflection and diffraction by the electrically large object are ignored. Although this reflection and diffraction might contribute little to the fields at the open section of the cavity, they may affect the fields in the far zone significantly. As a result, we use the reciprocity principle to solve this problem. According to the reciprocity principle, the equivalent magnetic current $\mathbf{M}$ and the field $\mathbf{H}^s$ are in the following relation:

$$\int \mathbf{H}^s \cdot \mathbf{M}_a \, dV = \int \mathbf{H}_a \cdot \mathbf{M} \, dV \tag{5.13}$$

where $\mathbf{M}_a$ denotes any magnetic current, and $\mathbf{H}_a$ is the field induced by $\mathbf{M}_a$ in the existence of an electrically large object. Choose $\mathbf{M}_a$ as a unit pulse, and put it at the observation point in the far zone, we can obtain the field induced by it at the cavity open section

$$\mathbf{H}_a = jk_0 \frac{e^{-jk_0 r}}{4\pi r} \mathbf{H}^{po} \tag{5.14}$$

Inserting (5.14) into (5.13) yields

$$E^{sc}_{\theta,\phi} = \frac{jk_0 Z_0 e^{-jk_0 r}}{4\pi r} \int_{S_a} \mathbf{M} \cdot \mathbf{H}^{po}_{v,h} \, dS \tag{5.15}$$

where $\mathbf{H}^{po}_v$ and $\mathbf{H}^{po}_h$ denote the scattered magnetic fields by the electrically large body without any cavity under incidences of vertical polarization and horizontal polarization, respectively.

5.1.1.5 Numerical Examples

In order to show the accuracy of the aforementioned hybrid method, we present the scattering calculation of a metallic cube with a slot on it. The size of the metallic cube is $5\lambda \times 5\lambda \times 5\lambda$. On the upper surface, there is a slot with length of 5λ, width of 0.2λ and depth of 0.25λ, as shown in Figure 5.2a . In Figures 5.2b and c, the numerical results calculated by using the hybrid method and MoM are compared. The good agreement indicates the satisfactory accuracy of the hybrid method.

Problems

5.1 Analogous to the 3D formulation described in Section 5.1.1, derive the 2D formulation of the hybrid PO-FE method.
5.2 Based on the above-derived formulation, write a program of the hybrid PO-FE method for scattering by a 2D large target with a cavity.
5.3 Based on the 3D formulation described in Section 5.1.1, write a program of the hybrid PO-FE method for scattering by a 3D large target with a cavity.
5.4 Compare the accuracy of methods for computing RCS between using the reciprocity principle and the half-Green's function.

5.1.2 Hybrid Physical Optics Method and Moment Method

In practical engineering, we also frequently see electrically large objects with protrusions. The protrusions also play an important role in the scattering of the whole object.

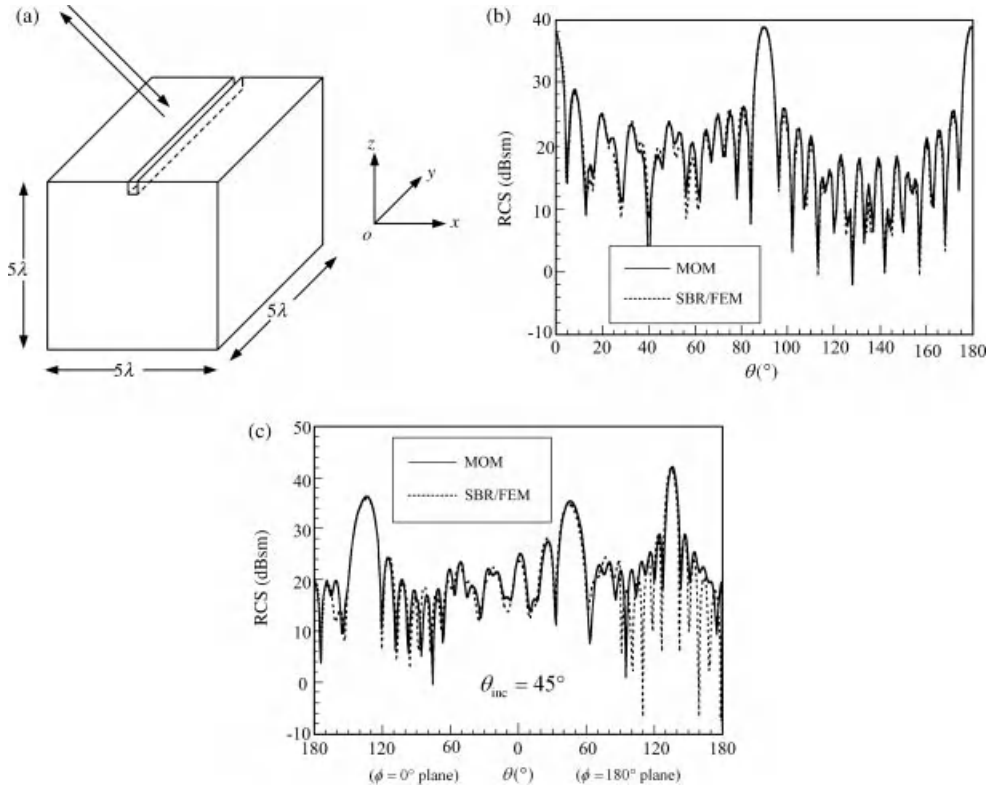


Figure 5.2 The scattering of a metallic cube with a slot calculated by using the hybrid method and MoM. (a) The model of the scatterer; (b) the monostatic RCS; (c) the bistatic RCS. (Reproduced with permission from J.M. Jin, S. Ni Sean and S.W. Lee, "Hybridization of SBR and FEM for scattering by large bodies with cracks and cavities," *IEEE Transactions on Antennas and Propagation*, **43**, 10, 1130–1139, 1995. © 1995 IEEE.)

In this section, we will introduce a scheme incorporating high-frequency asymptotic method with MoM for this kind of problem.

In Figure 5.3a, the model of an electrically large PEC object with a protrusion is illustrated. Similar to the examples in Section 5.1.1, the size of the PEC body could be tens or even hundreds of wavelengths, while the size of the protuberance is only of a few wavelengths. For such an object, we also need to combine the high-frequency asymptotic method with the full-wave numerical method in order to obtain an efficient and accurate solution. The idea here is to partition the computational domain into two parts by using the equivalent principle: the smooth PEC body, and the protuberance, as shown in Figure 5.3b. If we can obtain the dyadic Green's function $\underline{\underline{G}}$ satisfying the boundary condition of the smooth PEC body without the protrusion, then we can remove the protrusion by introducing an equivalent electric current $\mathbf{J}$.

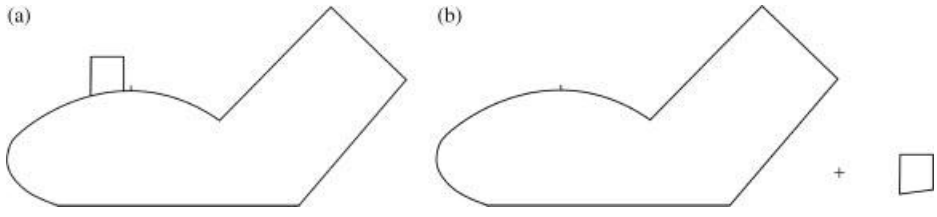


Figure 5.3 Hybrid method for calculating the scattering from an electrically large PEC object with a protrusion.

In the following discussion, the interface of the PEC body with the protrusion is referred to as the connection section. Consider the case when the incident wave impinges on the electrically large PEC body without protrusion. Denote the scattering field caused at the connection section $\mathbf{E}^s$, which can be obtained through the high-frequency asymptotic method. Then, we can solve for $\mathbf{J}$ through the following equation:

$$\left(\mathbf{E}^s - jk_0 \int_{S_p} \underline{\underline{G}} \cdot \mathbf{J} \, dS \right) \Bigg|_t = 0 \tag{5.16}$$

If we calculate $\mathbf{E}^s$ by using the PO method, exactly the same procedure can be adopted as in Section 5.1.1. Similarly, since it is hard to obtain the analytical expression for the dyadic Green's function $\underline{\underline{G}}$, we still use the half-space dyadic Green's function $\underline{\underline{G}}_h$ to approximate $\underline{\underline{G}}$, as we do in Section 5.1.1. By solving (5.16) using MoM, we can obtain the equivalent electric current $\mathbf{J}$, and then calculate the scattered field in the far zone.

As in Section 5.1.1, in order to accurately calculate the far-field scattering caused by the equivalent electric current $\mathbf{J}$, the reciprocity principle is used to obtain the following equation:

$$E^{sc}_{\theta,\phi} = -\frac{jk_0 e^{-jk_0 r}}{4\pi r} \int_{S_p} \mathbf{J} \cdot \mathbf{E}^{po}_{v,h} \, dS \tag{5.17}$$

where $\mathbf{E}^{po}_v$ and $\mathbf{E}^{po}_h$ denote the scattered electric fields by the electrically large smooth PEC body under incidences of vertical polarization and horizontal polarization, respectively.

In the following, we present one example to show the accuracy of this hybrid method. On top of a rectangular block of PEC of size $8\lambda \times 8\lambda \times 1\lambda$, we put two PEC patches with sizes of $2\lambda \times 2\lambda$ and $1\lambda \times 1\lambda$, respectively, as shown in Figure 5.4a. The numerical results by using the hybrid method and the MoM are compared in Figure 5.4b and (c). Good agreement indicates the satisfactory accuracy of the hybrid method.

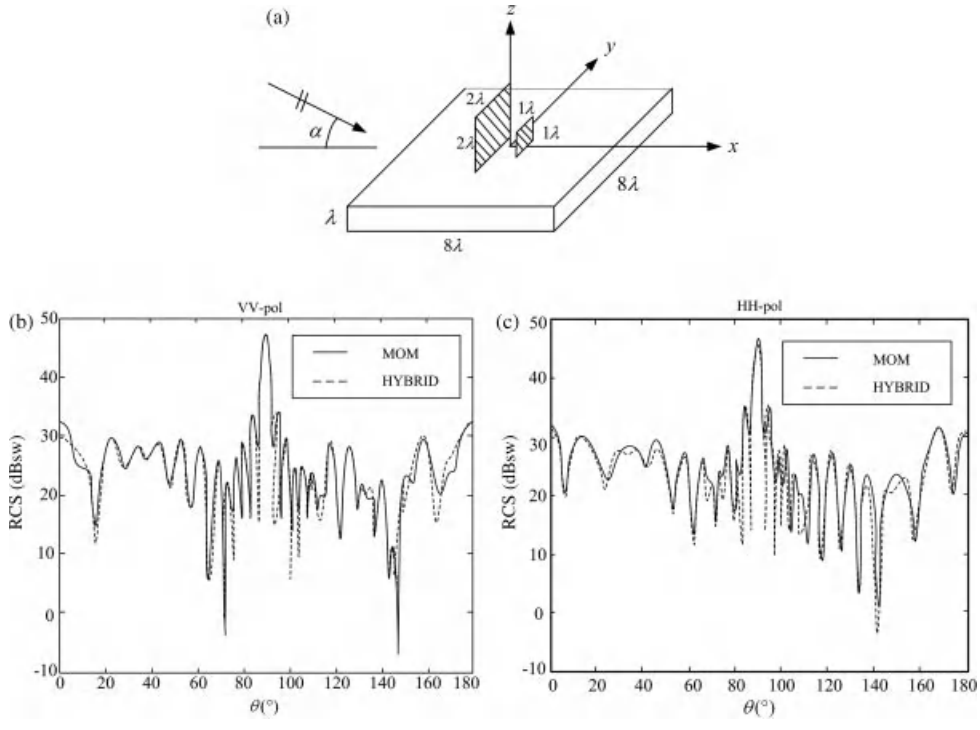


Figure 5.4 The calculated scattering by using the hybrid method and the MoM for the PEC block with two PEC patches. (a) The model of the scatterer; (b) VV polarization; (c) HH polarization. (Reproduced with permission from J.M. Jin, F. Ling, S. Carolan *et al.*, "A hybrid SBR/MoM for analysis of scattering by small protrusion on a large conducting body," *IEEE Transactions on Antennas and Propagation*, **46**, 9, 1349–1357, 1998. © 1998 IEEE.)

Problems

5.5 Analogous to the 3D formulation described in Section 5.1.2, derive the 2D formulation of the hybrid PO-MoM method

5.6 Based on the 2D formulation, write a program of the hybrid PO-MoM method for scattering by a 2D large target with a protrusion.

5.2 Hybrid Full-Wave Numerical Methods

Different full-wave methods are distinct in numerical behavior and consequently they have different ranges of applications. For example, MoM is suitable to solve for open-domain problems. A dielectric object with complex geometry can be treated accurately by using FEM. We will first consider FDTD in dealing with a wide-band problem.

Similarly, we'll go for the mode-matching method for a transmission problem. Even within the scope of one method, different algorithms have their own advantages and disadvantages. Naturally, their applications are different. For example, if we consider MoM, MFIE has the advantage of being fast in the iteration solution. However, we can only apply it to closed-domain problems. On the other hand, by using EFIE we can treat open problems such as radiation by antennas, and scattering by thin PEC patches. However, the solution suffers a slow convergence. Therefore, for complicated practical problem, we need to further integrate these methods or algorithms, in order to develop efficient and accurate schemes. In this section, we will use a few typical examples to illustrate the design and construction of these kinds of hybrid methods.

5.2.1 Hybrid FE-BI-MLFMA

Putting a coating on the metallic surface of an aircraft to absorb the scattering is a normal way in stealthy aircraft design. In order to design the coating, we need to calculate the scattering section of the composite structure. If we solely use MoM based on the volume integral equations (VIEs), the numerical efficiency is low. The reason for this is as follows. On the one hand, the majority of the composite structure is metal, in which the EM fields are null, so there is no need to include this metallic part as computational domain. On the other hand, when applying the fast Fourier transformation to accelerate the MoM based on VIEs, it is complicated to remove the metallic part from the computational domain. Alternatively, if we use FEM or FDTD, we need to truncate the computational domain. The absorbing boundary needs to be placed at a desired distance away from the coating to assure sufficient accuracy. This results in a larger computational domain and more unknowns, which consequently decreases the computational efficiency. More to the point, the absorbing boundary condition is only an approximate one, and its accuracy depends on the shape and material of the coating. Therefore, results from such methods can be used only as an estimation reference, instead of as an applicable standard. In the following, we will present a scheme combing FEM, MoM and MLFMA for an efficient solution for this problem.

5.2.1.1 Idea

Referring to Figure 5.5, the outer surface of the coating S_e separates the computational domain into two parts: the inhomogeneous dielectric part inside S_e, and the infinite free space. We apply FEM to formulate the fields inside S_e and outside the metallic boundary S_i. As for the fields outside S_e, MoM is employed. The Huygen's equivalence principle is used to connect and unify the two parts, which leads to a final solution. This is the basic idea of the hybrid method combining FEM

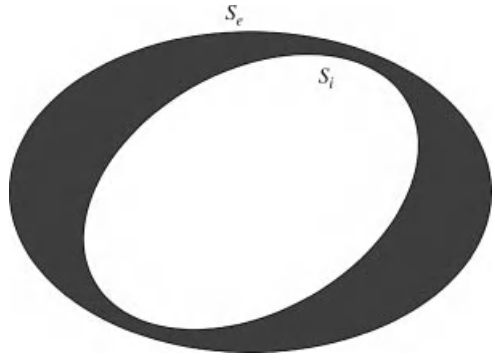


Figure 5.5 A coated scatterer.

and MoM. Historically, this hybrid method is often called the hybrid FE-BI method, where BI stands for boundary integral. The above is a description of the hybrid scheme from the view of the domain decomposition. The following will give a presentation of FE-BI from the formation of the final discretized matrix equation. Suppose the boundary condition on S_e is unknown. Then, when applying FEM to the interior field, the number of the discretized equation will be less than the number of the unknowns, because the unknowns on boundary S_e include both electric fields and magnetic fields. Only when the relationship of the electric fields and the magnetic fields on S_e is given can the linear system of equations be uniquely solved. The local boundary condition can be used in determining the above relationship. The advantage of such a treatment is that the sparseness of the FEM matrix is preserved. The disadvantage of this treatment is obvious as well. First, the local boundary condition is an approximation. Second, the outer boundary cannot be S_e. Instead, it should be some distance away from S_e. This leads to a larger computational domain. In the hybrid FE-BI, the boundary condition is formulated by the boundary integral equation. By applying the Huygens' principle, we can set up the integral equations for the electric and magnetic fields on boundary S_e. Combining this integral equation with the FEM equations we can obtain the final determined matrix equation. This treatment is accurate. However, the sparseness of the matrix is broken. In order to compensate for this, MLFMA is employed to accelerate the multiplication of the MoM matrix and vector. This is the whole idea of the hybrid finite element-boundary integral-multilevel fast multipole algorithm (FE-BI-MLFMA) to solve the scattering by complex composite targets. In the next section, the detailed formulation is presented.

5.2.1.2 Formulation

The equation that the interior electric field satisfies is equivalent to the variational formulation of the following functional

$$F(\mathbf{E}) = \frac{1}{2}\int_V \left[\frac{1}{\mu_r}(\nabla\times\mathbf{E})\cdot(\nabla\times\mathbf{E}) - k_0^2\varepsilon_r\mathbf{E}\cdot\mathbf{E}\right]dV + jk_0\int_{S_0}(\mathbf{E}\times\bar{\mathbf{H}})\cdot\hat{n}\,dS \tag{5.18}$$

where V denotes the volume between S_i and S_e, $\hat{n}$ denotes the unit normal vector pointing outwards on boundary S_e, and k_0 denotes the wave number in the free space. Define $\bar{\mathbf{H}} = Z_0\mathbf{H}$, where Z_0 is the wave impedance of the free space. The surface integral in (5.18) involves not only electric field unknowns but also magnetic field unknowns. This is because the relationship of the $\mathbf{E}$ fields and the $\mathbf{H}$ fields is still unknown. Mesh the domain V using tetrahedral elements and use the basis functions of the edge elements to discretized the above functional. Then, by taking the variational principle, we obtain

$$\begin{bmatrix} K_{II} & K_{IS} & 0 \\ K_{SI} & K_{SS} & B \end{bmatrix}\begin{Bmatrix} E_I \\ E_S \\ \bar{H}_S \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix} \tag{5.19}$$

where $\{E_I\}$ are the electric field unknowns in volume V. $\{E_S\}$ and $\{\bar{H}_S\}$ are the electric and magnetic field unknowns on boundary S_e, respectively. $K_{II}, K_{IS}, K_{SI}, K_{SS}$ and B are all sparse matrices. K_{II} and K_{SS} are symmetric, while B is antisymmetric. $K_{IS} = K_{SI}^T$ where the superscript T indicates the transpose operation. Obviously, (5.19) has more unknowns than equations, and cannot be determined until another boundary integral equation is employed to formulate the relationship of the unknown electric and magnetic fields on the boundary. This boundary integral equations can be EFIE:

$$\mathbf{M} + \hat{n}\times[\mathbf{L}(\bar{\mathbf{J}}) - \mathbf{K}(\mathbf{M})] = -\hat{n}\times\mathbf{E}^i \tag{5.20}$$

or MFIE

$$\bar{\mathbf{J}} - \hat{n}\times[\mathbf{L}(\mathbf{M}) + \mathbf{K}(\bar{\mathbf{J}})] = \hat{n}\times\mathbf{H}^i \tag{5.21}$$

The equivalent electric and magnetic currents on the boundary satisfy

$$\bar{\mathbf{J}} = \hat{n}\times\bar{\mathbf{H}} \tag{5.22}$$

$$\mathbf{M} = -\hat{n}\times\mathbf{E} \tag{5.23}$$

In order to assure the consistency of the unknowns in MoM and FEM, the equivalent electric and magnetic currents $\bar{\mathbf{J}}$ and $\mathbf{M}$ in the above equations are calculated using

the $\mathbf{E}$ and $\bar{\mathbf{H}}$ defined in FEM. Clearly, since the above FEM uses edge elements of a tetrahedron, the basis functions reduce to edge elements of a triangle on the boundary:

$$\mathbf{E} = \sum_{i=1}^{3} \mathbf{N}_i E_i \tag{5.24}$$

$$\bar{\mathbf{H}} = \sum_{i=1}^{3} \mathbf{N}_i \bar{H}_i \tag{5.25}$$

where $\mathbf{N}_i$ are the basis functions. Thus, the equivalent electric and magnetic currents in the triangular elements $\bar{\mathbf{J}}$ and $\mathbf{M}$ can be expressed as

$$\mathbf{M} = -\sum_{i=1}^{3} \hat{n} \times \mathbf{N}_i E_i \tag{5.26}$$

$$\bar{\mathbf{J}} = -\sum_{i=1}^{3} \hat{n} \times \mathbf{N}_i \bar{H}_i \tag{5.27}$$

Interestingly, $-\hat{n} \times \mathbf{N}_i$ are exactly the RWG basis functions $\mathbf{g}_i$, as we discussed in Chapter 2 which is proved in Problem 3.1.2. Therefore, the EFIE or the MFIE can be written uniformly as

$$[P]\{E_S\} + [Q]\{\bar{H}_S\} = \{b\} \tag{5.28}$$

By combining (5.19) and (5.28) we obtain

$$\begin{bmatrix} K_{II} & K_{IS} & 0 \\ K_{SI} & K_{SS} & B \\ 0 & P & Q \end{bmatrix} \begin{Bmatrix} E_I \\ E_S \\ \bar{H}_S \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \\ b \end{Bmatrix} \tag{5.29}$$

which is the final linear system of equations. With the same number of unknowns and equations, this system of equations can be solved straightforwardly.

5.2.1.3 Numerical Property of the Discretized Equation

As we know, in the coefficient matrix of (5.29), the FEM part is relatively simple. However, the MoM part results in different accuracy and efficiency if we choose different integral equations (IEs). Generally speaking there are four schemes of IEs: TE, NE, TH and NH, as shown in Chapter 2. TE is obtained through discretizing EFIE by using RWG as the test functions; NE is obtained through discretizing EFIE by using

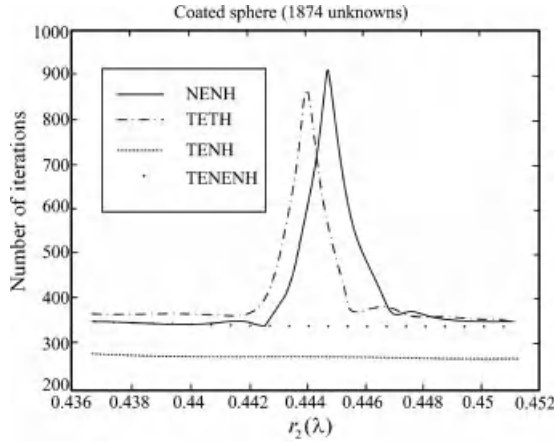


Figure 5.6 The number of iterations used for solving (5.29) by the CG method as a function of the outer radius of the spherical shell of cladding. (Reproduced with permission from X.Q. Sheng, J.M. Jin, J.M. Song *et al.*, "On the formulation of hybrid finite-element boundary-integral methods for 3D scattering," *IEEE Transactions on Antennas and Propagation*, **46**, 3, 303–311, 1998. © 1998 IEEE.)

$\hat{n} \times$ RWG as the test functions. TH and NH are obtained through discretizing *MFIE* by using RWG and $\hat{n} \times$ RWG as the test functions, respectively. These four schemes all suffer interior resonant problem. That is, at the resonant frequency, the matrix equation is singular. Evidently, the singularity in the MoM matrix part will result in the whole equation (5.29) singular. Therefore, to render a nonsingular (5.29), we cannot use directly TE, NE, TH or NH in (5.29). Alternatively, we know that in MoM, by combining the EFIE and the MFIE the interior resonance problem can be avoided, and there appear to be four schemes in the construction of the CFIE: TETH, TENH, NETH, and NENH. Further investigation shows that the CFIE using TETH or NENH schemes are still constructed in a lossless boundary. As a result, those equations still suffer the interior resonance problem. On the other hand, the TENH and NETH schemes are proper combinations that avoid the interior resonance. This statement is verified by the numerical examples in [6]. Here, we only present one figure from [6] (Figure 5.6). This figure shows the number of iterations for solving (5.29) by using the conjugate gradient method (CG) as a function of the outer radius of the coated sphere. The inner radius of the coated sphere, that is, the radius of the metallic sphere, is 0.3367λ. The relative permittivity of the coating is $\varepsilon_r = 4.0$, and the relative permeability is $\mu_r = 1.0$. As we know, by using CG to solve (5.29), the number of iterations needed is directly related to the condition number of the coefficient matrix of (5.29). A larger condition number leads to more iterations, and vice versa. From Figure 5.6, a resonance peak can be seen in both the TETH and the NENH curves, while the TENH and TENENH (TENENH is the scheme combining TE, NE and NH) curves are free of this peak. This means, for a certain electric size of the coating, the TETH and

NENH schemes result in a large condition number of the coefficient matrices in (5.29). In other words, the TETH and NENH schemes still suffer the interior resonance problem, while the TENH and TENENH do not. In this figure, the NETH curve is not provided. This is because the results of the NETH iteration hardly converge. Next, we will analyze the reason for the above numerical results and the efficiency of the above four CFIEs by examining the structures of the corresponding matrix equations. By further investigation, as was done in Chapter 2, it can be found that $[Q^{TE}]$ and $[P^{TH}]$ are diagonally strongly dominant matrices, $[Q^{TE}]$ and $[Q^{NH}]$ are diagonally weakly dominant matrices, while $[P^{TE}]$,$[Q^{TH}]$, $[Q^{NE}]$ and $[P^{NH}]$ are off-diagonally dominant matrices. This fact can be illustrated in a visual way as

$$\begin{aligned}
[Q^{TE}] = [P^{TH}] \sim \begin{bmatrix} \ddots & & \\ & 2 & \\ & & \ddots \end{bmatrix}, [P^{NE}] = -[Q^{NH}] \sim \begin{bmatrix} \ddots & & \\ & 1 & \\ & & \ddots \end{bmatrix}, \\
[P^{TE}] = -[Q^{TH}] \sim \begin{bmatrix} \ddots & & \\ & 0 & \\ & & \ddots \end{bmatrix}, [Q^{NE}] = [P^{NH}] \sim \begin{bmatrix} \ddots & & \\ & 0 & \\ & & \ddots \end{bmatrix}
\end{aligned} \tag{5.30}$$

Then, the structure of the coefficient matrix in the TETH equation is expressed as

$$[P|Q] \sim \begin{bmatrix} \ddots & & \ddots & \\ & 2 & & 2 \\ \ddots & & \ddots & \end{bmatrix} \tag{5.31}$$

The structure of the coefficient matrix in the TENH equation is illustrated as

$$[P|Q] \sim \begin{bmatrix} \ddots & & \ddots & \\ & 0 & & 3 \\ \ddots & & \ddots & \end{bmatrix} \tag{5.32}$$

The coefficient matrix in the NETH equation is illustrated as

$$[P|Q] \sim \begin{bmatrix} \ddots & & \ddots & \\ & 3 & & 0 \\ \ddots & & \ddots & \end{bmatrix} \tag{5.33}$$

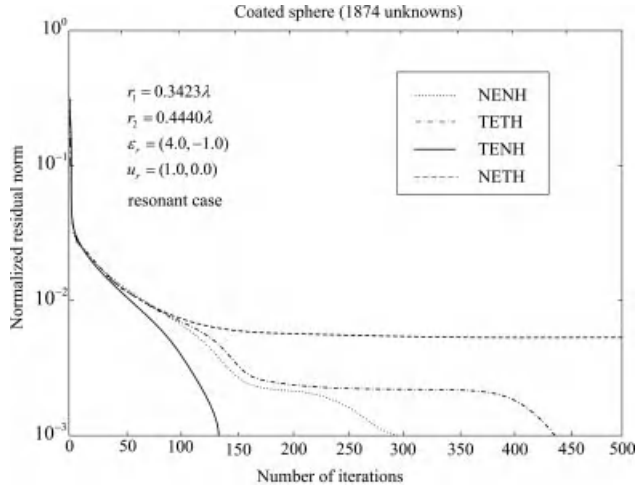


Figure 5.7 The iteration results of different CFIE schemes in solving (5.29) by using the CG method. (Reproduced with permission from X.Q. Sheng, J.M. Jin, J.M. Song *et al.*, "On the formulation of hybrid finite-element boundary-integral methods for 3D scattering," *IEEE Transactions on Antennas and Propagation*, **46**, 3, 303–311, 1998. © 1998 IEEE.)

The structure of the coefficient matrix in the NENH equation is expressed as

$$[P|Q] \sim \begin{bmatrix} \ddots & & \ddots & \\ & 1 & & 1 \\ & \ddots & & \ddots \end{bmatrix} \tag{5.34}$$

By combining with the FEM matrix structure, it can be clearly seen that the condition number of the TENH scheme is the best, which is followed by those from the TETH and NENH schemes, and the condition number of the NETH is the worst one.

This analysis is verified by the numerical experiments in [6,7] as well. Figure 5.7 presents the numerical results in [6]. It can be seen from this figure that the TENH curve demonstrated the most rapid convergence, followed by the TETH and NENH curves. As mentioned previously, the NETH curve hardly converges.

5.2.1.4 Solution of the FE-BI Matrix Equation and the Numerical Analysis

If we use a direct method to solve equation (5.29), the computational efficiency is obviously low, since the submatrices $[P]$ and $[Q]$ in (5.29) are full matrices. Alternatively, if we use the iteration method, we not only make good use of the sparseness of the FEM matrix, but the MLFMA can also be used to efficiently speed up the vector and matrices ($[P]$ and $[Q]$) product. Use N_v and N_s to denote the unknowns in

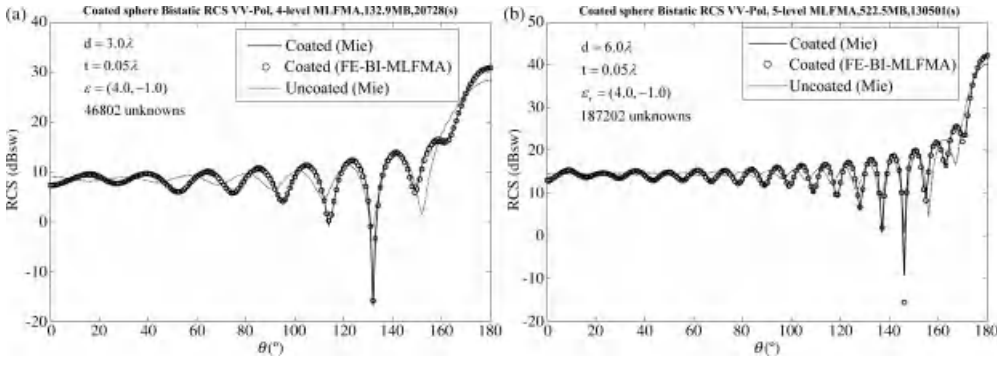


Figure 5.8 The bistatic RCS of metallic spheres (with diameter d) with one layer of cladding (with thickness of $0.05\lambda_0$, and relative permittivity $\varepsilon_r = 4.0 - j1.0$).(Reproduced with permission from X.Q. Sheng, J.M. Jin, J.M. Song *et al.*, "On the formulation of hybrid finite-element boundary-integral methods for 3D scattering," *IEEE Transactions on Antennas and Propagation*, **46**, 3, 303–311, 1998. © 1998 IEEE.)

the FEM and MoM, respectively. Through simple calculation, we know that by using iteration method (such as CG) to solve (5.29) and applying MLFMA to calculate the product of the submatrices $[P]$ and $[Q]$ with vectors, the memory requirement and the computation complexity should be $O(N_v + N_s \lg N_s)$ (N_s and N_v denote the numbers of unknowns in FEM and MOM, respectively), which can be verified by the following numerical results [8,9]. Table 5.2 listed the memory requirement and the CPU time when using this method to simulate coated spheres of different size (algorithms A and B in the table are explained in the following). From this table it can be seen that the memory requirement and the CPU time for each iteration increase almost linearly with the number of unknowns. In the following examples, the bistatic RCS of metallic spheres with one layer of coating and with two layers of coating are calculated using the FE-BI-MLFMA and the Mie series. The results are compared in Figures 5.8 and 5.9, respectively. In these two figures, the high accuracy of the FE-BI-MLFMA is demonstrated.

For convenience, the above algorithm is termed algorithm A for later reference. Although the condition number in the MoM coefficient matrix is small, the large condition number in the FEM part results in a large condition number in the coefficient matrix of the combined equation (5.29). As a result, algorithm A usually converges slowly. To solve this problem, we introduce an alternative way to solve (5.29). Precondition the sparse FEM coefficient matrix in (5.29) by using the LU decomposition method

$$[K] = \begin{bmatrix} K_{II} & K_{IS} \\ K_{SI} & K_{SS} \end{bmatrix} = [L][U] \tag{5.35}$$

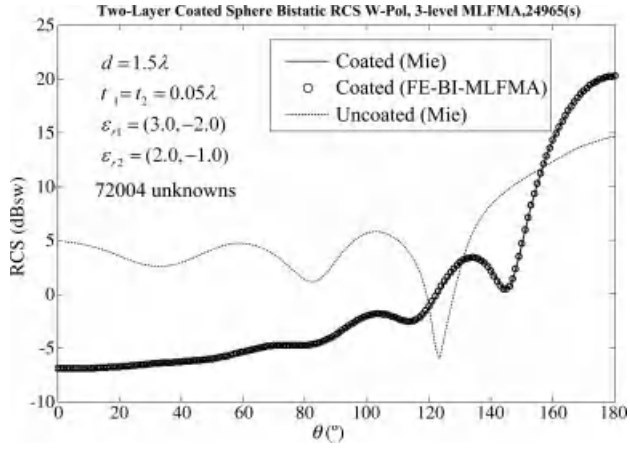


Figure 5.9 The bistatic RCS of a metallic sphere with two layers of cladding. The diameter of the metallic sphere is $1.5\lambda_0$. The thickness of each cladding shell is $0.05\lambda_0$. The EM parameters of the inner and the outer shells are $\varepsilon_{r1} = 3.0 - j2.0$, $\mu_{r1} = 2.0 - j1.0$, and $\varepsilon_{r2} = 2.0 - j1.0$, $\mu_{r2} = 3.0 - j2.0$, respectively. (Reproduced with permission from X.Q. Sheng, J.M. Jin, J.M. Song *et al.*, "On the formulation of hybrid finite-element boundary-integral methods for 3D scattering," *IEEE Transactions on Antennas and Propagation*, **46**, 3, 303–311, 1998. © 1998 IEEE.)

where $[L]$ is a lower triangular matrix, $[U]$ is an upper triangular matrix. Thus, the electric field on the boundary $\{E_S\}$ can be expressed by $\{H_S\}$:

$$\begin{Bmatrix} E_I \\ E_S \end{Bmatrix} = -[U]^{-1}[L]^{-1}\begin{bmatrix} 0 \\ B \end{bmatrix}\{H_S\} \tag{5.36}$$

For convenience, simply express $\{E_S\}$ by $\{H_S\}$ as

$$\{E_S\} = [M]\{H_S\} \tag{5.37}$$

From the following discussion it is clearly shown that we do not need to express $[M]$ explicitly. Substituting (5.37) into (5.29) we get

$$([P][M] + [Q])\{H_S\} = \{b\} \tag{5.38}$$

The numerical examples in the following show that, if we use iteration methods to solve (5.38), fast convergence can be observed. Therefore, the following algorithm (algorithm B) is proposed:

1. Use MLFMA to calculate $\{G\} \leftarrow [Q]\{H_s\}$
2. Calculate $[M]\{H_S\}$ in the following manner
   a. Calculate $\{H_s\} \leftarrow [B]\{H_s\}$

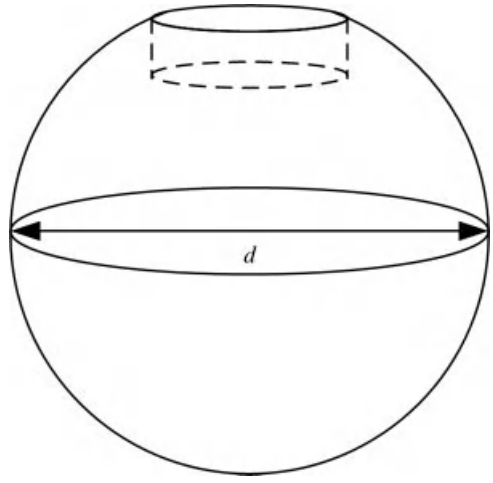


Figure 5.10 Illustration of a metallic sphere with a dielectrically filled cavity.

b. Calculate the forward substitution $\{H_s\} \leftarrow [L]^{-1}\{H_s\}$
c. Calculate the backward substitution $\{H_s\} \leftarrow [U]^{-1}\{H_s\}$.

3. Use MLFMA to calculate $\{H_s\} \leftarrow [P]\{H_s\}$
4. Calculate $\{H_s\} \leftarrow \{G\} + \{H_s\}$.

Let us examine every step in algorithm B. Generally speaking, the complexity in the LU decomposition is $O(N_v^2)$. The complexity in the forward substitution and the backward substitution are all $O(N_v^{1.5})$. Therefore, the total complexity in algorithm B is $O\left(N_v^2 + N_{ite}\left(N_v^{1.5} + N_s \lg N_s\right)\right)$, where N_{ite} is the total number of iterations. Compared with algorithm A, with computing complexity of $O(N_{ite}(N + N_s \lg N_s))$, it is hard to say which one is more efficient. This is because theoretically it is hard to estimate the iteration number in each algorithm. However, numerical experiments show that if $N_v \ll N_s$ (e.g., when the metallic body is partially coated), algorithm B is more efficient than A. The reason is that far fewer iterations are used in algorithm B than in A, and the LU decomposition consumes less CPU time. In other cases, algorithm A is more efficient than B due to the higher complexity caused by the LU decomposition, especially when large numbers of unknowns exist. It is worth noting that algorithm B suffers another disadvantage: substantial additional memory is required by the LU decomposition. Next, we will demonstrate the numerical property of these two algorithms by the calculation of scattering from two types of scatterer. The first type of scatterer is a metallic sphere with a dielectric-filled cavity, as shown in Figure 5.10. In this example, N_v is much smaller than N_s. Table 5.1 compared the number of unknowns, the memory requirement, the iteration number, and the CPU time in simulating different sizes of scatterer of

Table 5.1 Computer resources in computing the scattering from dielectric-filled metallic spheres of different size

d	$0.75\lambda_0(N_v = 235, N_s = 768)$			$1.5\lambda_0(N_v = 495, N_s = 2700)$			$3.0\lambda_0(N_v = 1851, N_s = 10\,800)$		
	N_{ite}	Memory/Mb	CPU/s	N_{ite}	Memory/Mb	CPU/s	N_{ite}	Memory/Mb	CPU/s
Algorithm A	260	3.9	63	183	12.3	262	282	48.7	2182
Algorithm B	34	4.1	8.5	17	13.0	26	62	51.6	491

this type. From the table, it can be seen that the number of iterations in algorithm *B* is much less than that in *A*. Despite the larger memory required, algorithm *B* is more computationally efficient compared with *A*. Therefore, for this type of scatterer, we can generally say algorithm *B* is better than *A*. The scatterers in the second example are metallic spheres with one layer of coating. In this case, N_v is greater than N_s. In Table 5.2, the number of unknowns, the memory requirement, the iteration number, and the CPU time in simulating different sizes of scatterer of this type are compared. Again, we can see that the number of iteration in algorithm *B* is much less than that in *A*. Despite this fact, the efficiency of algorithm *B* decreases with the increase of the scatterer size, and the memory requirement in *B* is much larger than in *A*. Therefore, for this type of scatterer, if the electrical size of the scatterer is large, then algorithm *A* is more computationally efficient than *B*. Of course we can further improve algorithm *B*. For example, instead of LU decomposition, we can use incomplete LU decomposition to precondition the matrix in the iteration.

Table 5.2 Computer resources in computing the scattering from coated spheres of different size

d	$0.75\lambda_0(N_v = 2526, N_s = 768)$			$1.5\lambda_0(N_v = 9002, N_s = 2700)$			$3.0\lambda_0(N_v = 36\,002, N_s = 10\,800)$		
	N_{ite}	Memory/Mb	CPU/s	N_{ite}	Memory/Mb	CPU/s	N_{ite}	Memory/Mb	CPU/s
Algorithm A	190	8.8	58	226	29.3	523	296	116.5	3478
Algorithm B	19	14.7	14	23	59.0	293	36	277.0	9319

5.2.1.5 Hybrid h- and p-type Multiplicative Schwarz (h-p-MUS) Preconditioned Algorithm

It is known from the above analysis that how to construct an efficient preconditioner is the key for designing fast algorithms of FE-BI-MLFMA, especially for higher-order FE-BI-MLFMA. As we known, employing higher-order elements is one of the general and robust approaches to further improve accuracy and efficiency of FE-BI-MLFMA. There are a variety of higher-order elements, as discussed in Chapter 3. The hierarchical type of higher-order elements can be easily employed by a way of using different order elements to model different subdomains, since different order elements in this type are automatically compatible. Hence, the hierarchical higher-order elements are suitable for modeling multiscale EM problems. However, compared with the interpolatory higher-order elements, the hierarchical one often results in the FEM matrix with a larger condition number. Therefore, it is more urgent to construct an efficient preconditioner for the hierarchical higher-order FE-BI-MLFMA. The following will present a so-called hybrid h-p-MUS preconditoner for the hierarchical higher-order FE-BI-MLFMA.

First, the FEM matrix part in (5.29) is rewritten by classifying unknowns into different groups corresponding to different FEM orders. Taking the hierarchical second-order FEM matrix as an example, namely, the following FEM matrix in (5.29)

$$K = \begin{bmatrix} K_{II} & K_{IS} \\ K_{SI} & K_{SS} \end{bmatrix} \tag{5.39}$$

is rewritten as

$$K = \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix} \tag{5.40}$$

where subscript "1" corresponds to the first-order FEM unknowns, "2" to the second-order FEM unknowns. Then, this rewritten FEM matrix is factorized as

$$\begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix} = \begin{bmatrix} I & 0 \\ K_{21}K_{11}^{-1} & I \end{bmatrix} \begin{bmatrix} K_{11} & 0 \\ 0 & K_{22} - K_{21}K_{11}^{-1}K_{12} \end{bmatrix} \begin{bmatrix} I & K_{11}^{-1}K_{12} \\ 0 & I \end{bmatrix} \tag{5.41}$$

Based on this factorization, the preconditioning matrix of $[K]$ can be approximated, according to the multiplicative Schwarz method, as

$$M = \begin{bmatrix} I & 0 \\ K_{21}K_{11}^{-1} & I \end{bmatrix} \begin{bmatrix} K_{11} & 0 \\ 0 & K_{22} \end{bmatrix} \begin{bmatrix} I & K_{11}^{-1}K_{12} \\ 0 & I \end{bmatrix} \tag{5.42}$$

Since the inverse of $[M]$ can be directly written out, a good preconditioner of $[K]$ can be obtained as

$$M^{-1} = \begin{bmatrix} I & -K_{11}^{-1}K_{12} \\ 0 & I \end{bmatrix} \begin{bmatrix} K_{11}^{-1} & 0 \\ 0 & K_{22}^{-1} \end{bmatrix} \begin{bmatrix} I & 0 \\ -K_{21}K_{11}^{-1} & I \end{bmatrix} \tag{5.43}$$

The above construction procedure of the second-order hierarchical FEM matrix can be extended to any order FEM matrix. Since the principle of grouping is based on the order of unknowns, this preconditioner is called a p-type multiplicative Schwarz (p-MUS) preconditioner. It is worth pointing out that (5.42) is obtained by dropping the coupling term of $-K_{21}K_{11}^{-1}K_{12}$. Thus, the weaker the coupling of the different order basis functions, the more efficient is the p-MUS preconditioner. In practice, the hierarchical basis functions are orthogonalized before constructing the p-MUS preconditioner.

It is known from (5.43) that to obtain the preconditioner, namely M^{-1}, the inverses of K_{11} and K_{22} are required. Since completely computing the inverses of K_{11} and K_{22} is still expensive, we usually compute approximations of the inverses of K_{11} and K_{22} instead. We often utilize an h-type multiplicative Schwarz (h-MUS) to obtain the approximations of the inverses of K_{11} and K_{22} instead of utilizing conventional incomplete LU factorization. Therefore, we name the so-obtained preconditioner as the h-p-MUS preconditioner.

The h-MUS preconditioning approach establishes on the Crout LU factorization shown in Figure 5.11. The dark area shows the parts of the entries being computed at the kth step. The shaded areas show the parts of the entries being accessed at the kth step. It can be seen that the Crout LU factorization computes the k-th column of L and kth row of U at step k [10]. In the implementation, we store L by columns and U by rows, thus the entries can be searched and accessed efficiently during the factorization.

The dark area shows the parts of the entries being computed at the kth step. The shaded areas show the parts of the entries being accessed at the kth step.

The ILU factorization is obtained by applying the dropping rule to the entries of the kth column of L and the kth row of U for each step k. The entries usually are dropped either by their location (positional dropping) or magnitude (threshold dropping). The positional dropping works well for regularly structured problems, but it is not designed for general cases. The threshold dropping is often used, but it is not robust. Since the threshold dropping cannot guarantee that the norm of the inverse of L or U is small, this may have some unpredictable effects to the ILU preconditioner.

To overcome this problem, a better dropping rule is presented in [11], which is to drop an entry at step k when it satisfies $\left|l_{jk}\right|\left\|e_k^T L^{-1}\right\| \leq \varepsilon$, where e_k^T denotes the kth unit vector, and ε is the ILU drop tolerance. Since the inverse of L is usually not

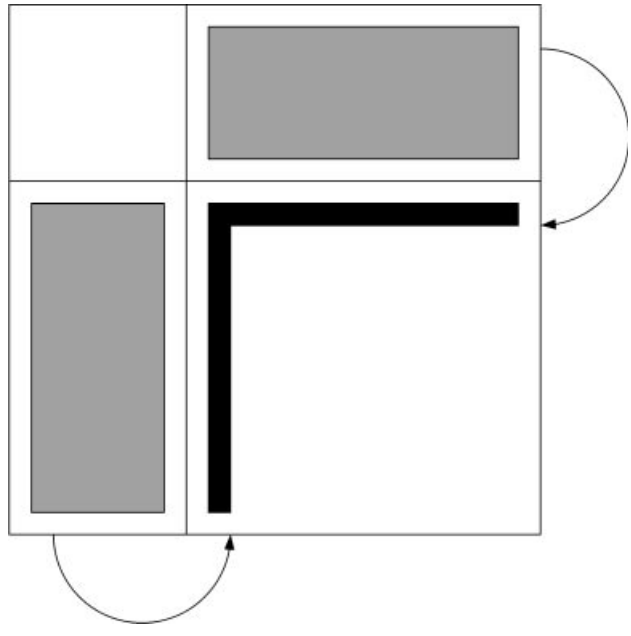


Figure 5.11 The computational pattern for the Crout factorization.

available, the next problem is how to estimate $\left\|e_k^T L^{-1}\right\|$. To this end, the estimation of $\left\|e_k^T L^{-1}\right\|$ is approximated as $\left\|e_k^T L^{-1} b\right\|_\infty$ for any testing vector **b** satisfying $\|b\|_\infty = 1$. Since e_k^T denotes the k-th unit vector, the problem to estimate $\left\|e_k^T L^{-1} b\right\|_\infty$ at step k is reduced to determine the largest kth component of the solution to the linear system $Lx = b$. The implementation adopted in this paper uses the simplest rule which amounts to selecting $b_k = \pm 1$ at each step k, in such a way as to maximize the kth component of $L^{-1}b$.

To take full advantage of the idea of the inverse-based ILU, we extend it to a multilevel version. Suppose the kth column of L and kth row of U have been calculated by the Crout factorization, we then estimate $\left\|L^{-1}\right\|$ and $\left\|U^{-1}\right\|$. If $\left\|L^{-1}\right\| < \kappa$ and $\left\|U^{-1}\right\| < \kappa$ (κ is the set bound), then the inverse-based ILU is performed, otherwise, the kth column and row will be postponed to the last column and row. This procedure is repeated through all columns and rows. Finally, all columns and rows of matrix K are categorized into two types: factorized and postponed. To be matched with the concepts of algebraic multilevel method, the factorized columns and rows can be considered as corresponding to the fine grid, whereas the postponed ones correspond to the coarse grid. For a clear description, the original matrix can be rearranged as

$$K = \begin{pmatrix} K_{FF} & K_{FC} \\ K_{CF} & K_{CC} \end{pmatrix} \tag{5.44}$$

Table 5.3 The memory requirement and CPU time for computing the dielectric sphere

Target	Unknowns FEM/MoM	Memory (GB)		CPU time (s)		Number of iterations
		FEM	MoM	h-p-MUS construction	Iteration	
Dielectric sphere	2 593 996/16 428	4.6	0.88	700	889	31

Since the inverse-based ILU on K_{FF} has been performed, the preconditioner $\tilde{K}_{FF}^{-1}$ of K_{FF} can be easily constructed. Based on the multiplicative Schwarz method, the preconditioner of K is

$$K^{-1} \approx \begin{bmatrix} I & -K_{FF}^{-1}K_{FC} \\ 0 & I \end{bmatrix} \begin{bmatrix} K_{FF}^{-1} & 0 \\ 0 & \left(K_{CC}-K_{CF}K_{FF}^{-1}K_{FC}\right)^{-1} \end{bmatrix} \begin{bmatrix} I & 0 \\ -K_{CF}K_{FF}^{-1} & I \end{bmatrix} \tag{5.45}$$

Thus, the construction of the preconditioner of K is reduced to the construction of the preconditioner of S_{CC}, where $S_{CC} = K_{CC} - K_{CF}K_{FF}^{-1}K_{FC}$. In other words, the remaining problem is to perform the h-MUS preconditioner in the coarse level. This procedure can recursively proceed to obtain a multilevel h-MUS preconditioner. To obtain a preconditioner for the whole matrix equation of (5.29), the simplest diagonal preconditioner for the BI matrix part of (5.29) is utilized to combine the h-p-MUS preconditioner to obtain the preconditioner of (5.29).

To demonstrate the accuracy and efficiency of this h-p-MUS preconditioned FE-BI-MLFMA, we compute the scattering by an electrically large dielectric sphere with diameter of 9λ, and relative permittivity is $3-4j$. In this computation, the unknowns of the BI part are 16 428 and unknowns of the FEM part are 2 593 996. The detailed computation information is list in Table 5.3. The computed bistatic VV-polarized RCS is shown in Figure 5.12 together with the Mie results, demonstrating the powerful capability of the proposed h-p-MUS preconditioned FE-BI-MLFMA.

5.2.2 Hybrid Method Combining EFIE and MFIE

Many practical structures such as the wire antennas on airplane or satellite can be considered as a thin wire on an electrically large metallic platform, as illustrated in Figure 5.13. If we simply use the moment method introduced in Chapter 2, we cannot get a solution with high numerical efficiency. The reason is as follows. Due to the thread structure, the wire antenna can only be formulated by EFIE. Hence, if we use a single integral equation to formulate the whole problem, we have to employ EFIE.

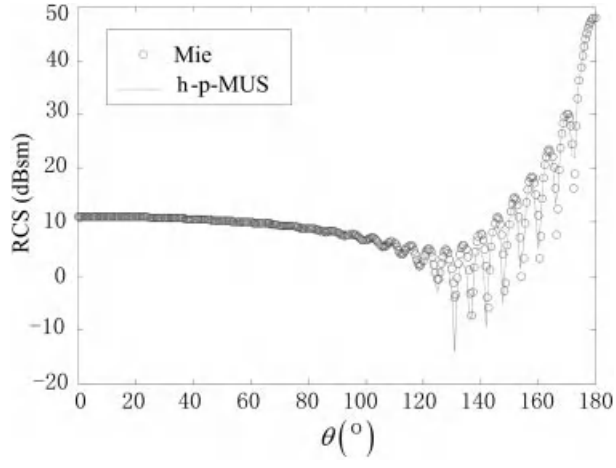


Figure 5.12 Bistatic VV-polarized RCS for a dielectric sphere with diameter of 9 wavelengths.

As we know, for electrically large problems, the matrix equation from discretizing EFIE cannot be solved by direct solvers due to their large computational complexity. It can only be solved by iterative solvers with the aid of acceleration of the matrix-vector product by MLFMA. Since EFIE generally leads to a MoM matrix with a large condition number, the iterative solution will experience slow convergence. Therefore, this kind of solution for the thin wire on an electrically large metallic platform is overall not efficient. To address this problem, a hybrid approach composed of EFIE and MFIE is proposed in [12]. The basic idea is to separate the computational domain into two parts, that is, the surface of the wire antenna S_1 and the surface of the metallic platform S_2. Construct MFIE on the electrically large S_2, and apply EFIE on S_1. Then, use an iteration method to solve the equation. The formulation is as follows. Construct EFIE on S_1:

$$\mathbf{L}_1(\mathbf{J}_1) + \mathbf{L}_2(\mathbf{J}_2) \;=\; -\mathbf{E}^i, \quad r \in S_1 \tag{5.46}$$

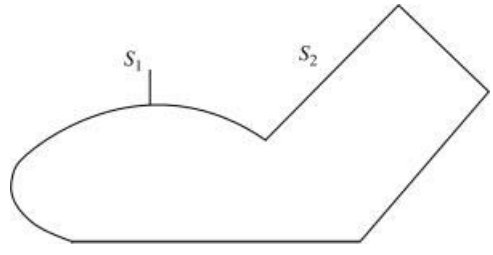


Figure 5.13 Wire antenna on an electrically large metallic platform.

where $\mathbf{J}_1$ and $\mathbf{J}_2$ are the equivalent electric currents on S_1 and S_2, respectively. Construct MFIE on S_2:

$$\mathbf{K}_1(\mathbf{J}_1) + \mathbf{K}_2(\mathbf{J}_2) \ = \ 0, \ \ r \in S_2 \tag{5.47}$$

where the operator $\mathbf{L}_i$ and $\mathbf{K}_i$ (i = 1,2) are the vector integrodifferential operators. Note that the integral domain of $\mathbf{L}_1$ and $\mathbf{K}_1$ is S_1, while that of $\mathbf{L}_2$ and $\mathbf{K}_2$ is S_2. As in the usual MoM procedure, we express the equivalent currents $\mathbf{J}_1$ and $\mathbf{J}_2$ as

$$\mathbf{J}_1 = \sum_{i=1}^{N_{S1}} \mathbf{f}_i I_{1i} \tag{5.48}$$

$$\mathbf{J}_2 = \sum_{i=1}^{N_{S2}} \mathbf{g}_i I_{2i} \tag{5.49}$$

where N_{S1} and N_{S2} denote the unknowns on S_1 and S_2, respectively. $\mathbf{f}_i$ are the basis functions on the wire antennas, which can be linear interpolation functions, for instance. $\mathbf{g}_i$ are the basis functions on the metallic platform and we recommend the use of RWG basis functions. Substituting (5.48) and (5.49) into (5.46) and (5.47), and integrating the dot product of the function with the respective testing function yields

$$[Z_{11}]_{N_{S1}\times N_{S1}}\{I_1\} + [Z_{12}]_{N_{S1}\times N_{S2}}\{I_2\} \ = \ \{V_1\} \tag{5.50}$$

$$[Z_{21}]_{N_{S2}\times N_{S1}}\{I_1\} + [Z_{22}]_{N_{S2}\times N_{S2}}\{I_2\} \ = \ \{0\} \tag{5.51}$$

The number of unknowns comprised in the linear systems of equations (5.50) and (5.51) has been clearly marked in the matrix subscript. For a neat expression, these subscripts are not shown in the following equations. According to (5.50) and (5.51), we can design an efficient iteration algorithm to calculate the equivalent electric currents I_1 and I_2. First, rewrite (5.50) and (5.51) as

$$\{I_1\} \ = \ [Z_{11}]^{-1}(\{V_1\} - [Z_{12}]\{I_2\}) \tag{5.52}$$

$$\{I_2\} \ = \ [Z_{21}]\{I_1\} + [Z_{22} + E]\{I_2\} \tag{5.53}$$

where E is the unit matrix. Then, use the PO method to estimate the surface equivalent current on the surface of the metallic platform $I_2^{(0)}$, with the superscript "0" denoting the initial value in the iteration:

$$\mathbf{J}_{po} \ = \ \begin{cases} 2\hat{n} \times \mathbf{H}^i(r), \ r \in S_{lit} \\ 0, \qquad\qquad r \in S_{dark} \end{cases} \tag{5.54}$$

where S_{lit} and S_{dark} denote the illuminated and shadow regions in PO respectively. By using (5.52), we can calculate $I_1^{(1)}$ from $I_2^{(0)}$. Then, by using (5.53)$I_2^{(1)}$ is updated

from $I_1^{(1)}$. Keep iterating until the following conditions are reached:

$$\frac{\left\| I_1^{(j+1)} - I_1^{(j)} \right\|}{\left\| I_1^{(j)} \right\|} < \varepsilon_1 \tag{5.55}$$

$$\frac{\left\| I_2^{(k+1)} - I_2^{(k)} \right\|}{\left\| I_2^{(j)} \right\|} < \varepsilon_2 \tag{5.56}$$

where ε_1 and ε_2 are predefined parameters by the user according to the accuracy requirement. $\| \cdot \|$ denotes the vector norm. Since the number of unknowns on the wire is relatively small, the computing of (5.52) takes an inconsiderable time. In fact, the product of $[Z_{22}]\{I_2\}$ in (5.53) is the most time-consuming part. We can use MLFMA to accelerate its calculation. From the above analysis, it can be seen that I_1 is updated by I_2 using a direct solution, while I_2 is updated by I_1 using iteration. Therefore, in practice, we use (5.53) to calculate I_2 for many times before using (5.52) to calculate I_1 for one time. Further investigation shows that the speed of convergence depends on two procedures: (a) the iteration procedure in calculating I_2 from I_1; (b) the iteration procedure showing the interaction of I_1 and I_2. The former depends on the MFIE prescribed by the metallic platform. Since the condition number of a MFIE is usually good, this procedure should converge fast. The latter depends on the interaction of the wire and the metallic platform. If they are weakly interacting, this iteration converges rapidly. However, when strong coupling exists between the wire and the platform, this iteration converges slowly. Under these circumstances, we can insert (5.53) into (5.52). Then, I_1 is calculated by

$$\{I_1\} = ([Z_{11}] + [Z_{12}][Z_{21}])^{-1}(\{V_1\} - [Z_{12}][Z_{22} + E]\{I_2\}) \tag{5.57}$$

In this equation, the coupling between the wire and the platform has been taken into consideration. So, the iteration should be faster than that based on (5.52).

The accuracy and efficiency of the above method has been validated by the numerical results in reference [12]. Here, we only choose one example from [12] to show the numerical performance of this bybrid EFIE-MFIE. More results and discussion can be found in the original paper. The model of this example is a wire antenna on a PEC cylinder, as shown in Figure 5.14. The radius and the height of the cylinder are 0.2778λ and 0.6111λ, respectively. The length of the wire antennas is 0.3333λ. The radiation pattern and the directivity of the antennas on the $x - y$ plane are calculated by three methods: the physical optics method (PO), the moment method based on EFIE (MM), and the hybrid EFIE-MFIE iterative method (HEMI). The results are compared in Figure 5.15. It can be seen that the accuracy of HEMI and MM are comparable, and much better than that of POHM. The computation time

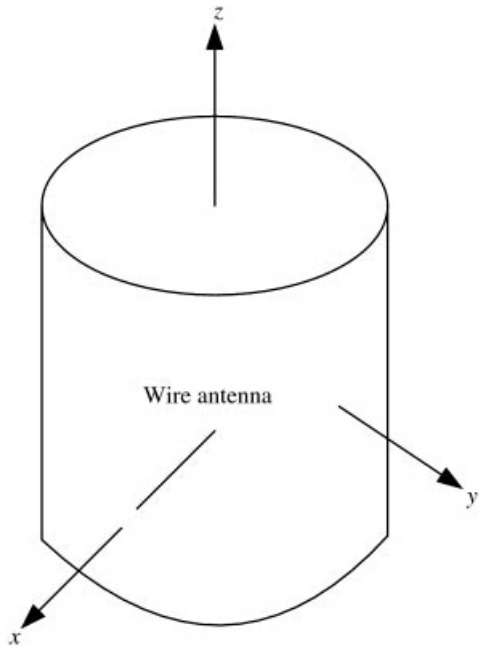


Figure 5.14 Wire antennas on an electrically large metallic cylinder.

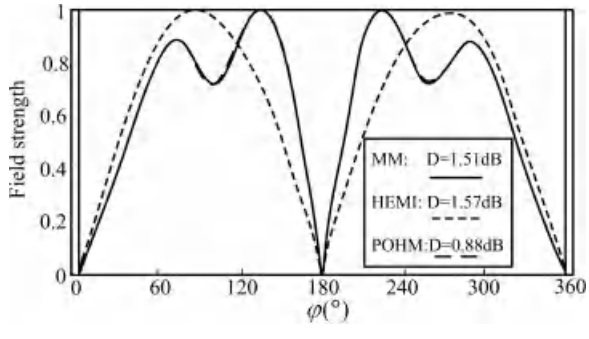


Figure 5.15 The radiation pattern and the directivity of the above antennas on the $x - y$ plane. (Reproduced with permission from R.E. Hodges and Y. Rahmat-Sammi, "An iterative current-based hybrid method for complex structures," *IEEE Transactions on Antennas and Propagation*, (**45**) 2, 265–276, 1997. © 1997 IEEE.)

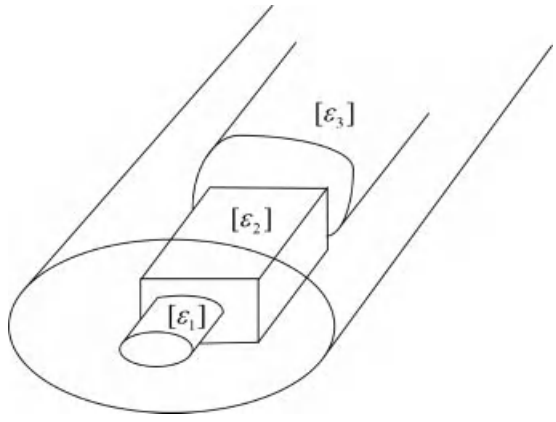


Figure 5.16 Illustration of the multistep dielectric discontinuity in a guiding structure.

used in PO, HEMI and MM simulations are 16.8 s, 55.5 s, and 655 s, respectively. These indicate that the HEMI is not only accurate but also efficient in modeling this type of structure.

Problem

5.7 Study determinants on the convergence speed of the above-described hybrid EFIE-MFIE method.

5.2.3 Hybrid Method Combining FEM and Mode-Matching Method

At the beginning of this chapter, we mentioned two types of ideas in hybridization: separate the computational domain into different regions and apply different schemes to each region; and separate the dimension in the EM problem and apply different methods for each dimension. The first type is introduced in the last two sections. In this section, we will use an example to illustrate the second idea. Figure 5.16 shows a guiding structure with a multistep dielectric discontinuity. For this problem, we can use FEM to calculate the eigenmodes of each waveguide segment first, then, use the mode-matching method to obtain the impedance transformation equation at the interface of the two segments. Finally, we calculate the scattering parameter of the whole structure according to the multimode transmission theory [13,14]. Obviously, in the above procedure, the original 3D problem has been converted to a 2D problem and the computational efficiency is increased. In the following, we will present the mathematic equations in detail. Suppose the eigenmodes of each waveguide segment are obtained by FEM [13], which is

introduced in Chapter 3. Thus, the transverse and longitudinal electric field components of the kth eigenmode can be expressed as

$$\mathbf{E}_t^{(k)} = \frac{A^{(k)}}{\beta^{(k)}} \sum_{p=1}^{M} \sum_{i=1}^{3} \mathbf{N}_i e_{tpi}^{(k)} \tag{5.58}$$

$$E_z^{(k)} = -\frac{A^{(k)}}{j} \sum_{p=1}^{M} \sum_{i=1}^{3} L_i e_{zpi}^{(k)} \tag{5.59}$$

where the superscript k denotes the kth mode and $A^{(k)}$ is the coefficient to be determined. By using the Maxwell's equation the transverse magnetic field component of the kth mode can be expressed as

$$\begin{aligned} \mathbf{H}_t &= -\frac{1}{j\omega\mu} \nabla \times \mathbf{E}|_t \\ &= \frac{A^{(k)}}{\omega\mu} \hat{i}_z \times \sum_{p=1}^{M} \sum_{i=1}^{3} \left(\mathbf{N}_i e_{tpi}^{(k)} + \nabla_t L_i e_{zpi}^{(k)} \right) \end{aligned} \tag{5.60}$$

By expressing the transverse electromagnetic field components of the kth eigenmode as

$$\mathbf{e}_t^{(k)} = A^{(k)} \sum_{p=1}^{M} \sum_{i=1}^{3} \mathbf{N}_i e_{tpi}^{(k)} \tag{5.61}$$

$$\mathbf{h}_t^{(k)} = A^{(k)} \hat{i}_z \times \sum_{p=1}^{M} \sum_{i=1}^{3} \left(\mathbf{N}_i e_{tpi}^{(k)} + \nabla_t L_i e_{zpi}^{(k)} \right), \tag{5.62}$$

the wave impedance of the kth eigenmode is obtained

$$Z_k = \frac{\omega\mu_0}{\beta_k} \tag{5.63}$$

where β_k in (5.63) is the $\beta^{(k)}$ in (5.58). The coefficient $A^{(k)}$ in (5.61) and (5.62) can be determined by using the orthogonal relationship of the waveguide eigenmode

$$\iint_{\Omega} \mathbf{e}_{ti} \times \mathbf{h}_{tj} \cdot \hat{n} \, dS = \delta_{ij}. \tag{5.64}$$

Since the eigenmodes of each waveguide segment are calculated, the fields on both sides of the segment interface can be expressed by linear combination of these modes. Express the fields at the LHS of the interface as

$$\mathbf{E}_t = \sum_i \mathbf{e}_{ti} U_i \tag{5.65}$$

$$\mathbf{H}_i = \sum_i \mathbf{h}_{ti} I_i, \tag{5.66}$$

and the fields at the RHS of the interface as

$$\bar{\mathbf{E}}_t = \sum_i \bar{\mathbf{e}}_{ti} \bar{U}_i \tag{5.67}$$

$$\bar{\mathbf{H}}_i = \sum_i \bar{\mathbf{h}}_{ti} \bar{I}_i \tag{5.68}$$

According to the continuity of the tangential fields on the interface, we can write:

$$\sum_i \mathbf{e}_{ti} U_i = \sum_i \bar{\mathbf{e}}_{ti} \bar{U}_i \tag{5.69}$$

$$\sum_i \mathbf{h}_{ti} I_i = \sum_i \bar{\mathbf{h}}_{ti} \bar{I}_i \tag{5.70}$$

Take the cross product of (5.69) with $\mathbf{h}_{ti}$ and the cross product of $\bar{\mathbf{e}}_{ti}$ with (5.70), before integrating on the whole interface. Then, according to the orthogonal relationship (5.64), we get

$$U_i = \sum_j Q_{ij} \bar{U}_j \tag{5.71}$$

$$\sum_j Q_{ji} I_j = \bar{I}_i \tag{5.72}$$

where

$$Q_{ij} = \iint_\Omega \bar{\mathbf{e}}_{ti} \times \mathbf{h}_{tj} \cdot \hat{n} \, dS \tag{5.73}$$

Writing (5.71) and (5.72) into the following matrix forms

$$\{U\} = [Q]\{\bar{U}\} \tag{5.74}$$

$$[Q]^T \{I\} = \{\bar{I}\} \tag{5.75}$$

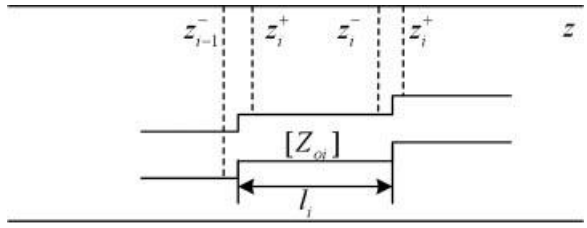


Figure 5.17 The longitudinal section of this guiding structure with multistep dielectric discontinuity.

According to the definition of the impedance of the multimode network,

$$\{U\} = [Z][I] \tag{5.76}$$

$$\{\bar{U}\} = [\bar{Z}][\bar{I}] \tag{5.77}$$

through (5.74) and (5.75) we can obtain the following impedance transformation equation at the interface of two segments:

$$[Z] = [Q][\bar{Z}][Q]^T \tag{5.78}$$

With this transformation, the scattering parameter of the whole structure can be calculated according to the multimode transmission theory. Figure 5.17 shows the longitudinal section of this guiding structure with multistep dielectric discontinuity. At the interface $z = z_i^-$, the reflection coefficient matrix from left to right is

$$[\gamma(z_i^-)] = ([Z(z_i^-)] + [Z_{oi}])^{-1}([Z(z_i^-)] - [Z_{oi}]) \tag{5.79}$$

where $[Z(z_i^-)]$ is the impedance at the interface $z = z_i^-$, which is determined by (6.76). $[Z_{oi}]$ is the characteristic impedance matrix of the ith waveguide segment. Then, at the interface $z = z_{i-1}^-$, the input impedance from left to right is

$$\left[Z\left(z_{i-1}^+\right)\right] = [Z_{oi}]\left([I] + [H_i]\left[\Upsilon(z_i^-)\right][H_i]\right)\left([I] - [H_i]\left[\Upsilon\left(Z_i^-\right)\right][H_i]\right)^{-1} \tag{5.80}$$

where $[H_i]$ is the phase-shift matrix of the ith waveguide segment. Both $[Z_{oi}]$ and $[H_i]$ are diagonal matrices, with elements

$$[Z_{oi}]_{\mathrm{mn}} = \delta_{\mathrm{mn}} \frac{\omega\mu}{\beta_{in}} \tag{5.81}$$

$$[H_i]_{\mathrm{mn}} = \delta_{mn} \exp(-\beta_{in} l_i) \tag{5.82}$$

where β_{in} is the propagation constant of the nth mode in the ith waveguide segment. By repeating the calculation of (5.78)–(5.80), we can obtain the reflection coefficient

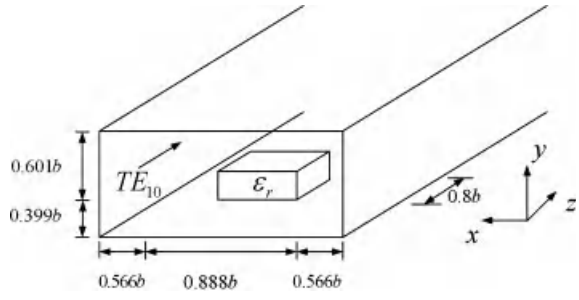


Figure 5.18 A discontinuity in a rectangular waveguide.

of the whole multistep structure. The above hybrid method is verified by the numerical results from references [13,14]. Here, we only use one example, as shown in Figure 5.18, for illustration. In Figure 5.19, the amplitude of the reflection coefficient of the dominant mode in the discontinuity structure calculated by the hybrid method and the 3D FEM are compared. Good agreement can be seen from the figure. In the hybrid method, 40 eigenmodes and 5 eigenmodes are computed for the hollow and the material-filled waveguides, respectively. At the time of this numerical experiment, the computing speed of the computer was not as fast as at present. With the same machine, for each frequency, the hybrid method took 3 minutes while the 3D FEM took nearly half an hour. Through this comparison, the efficiency of this hybrid method is clearly shown.

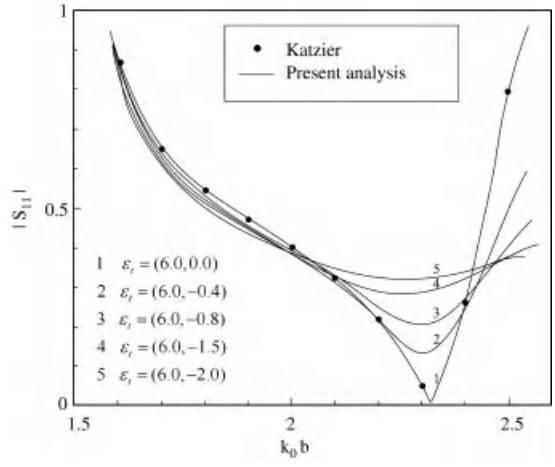


Figure 5.19 The amplitudes of the reflection coefficient of the dominant mode in the waveguide discontinuity. (Reproduced with permission from S.J. Xu and X.Q. Sheng, "Coupling of edge-element and mode-matching for multistep dielectric discontinuity in guiding structures," *IEEE Transactions on Microwave Theory and Techniques*, (**45**) 2, 284–287, 1997. © 1997 IEEE.)

Problem

5.8 Consider determinants on accuracy of the hybrid FE-MM method, and propose an approach for controlling its accuracy.

References

1. Thiele, G.A. and Newhouse, T.H. (1975) "A hybrid technique for combining moment methods with the geometrical theory of diffraction." *IEEE Transactions on Antennas and Propagation*, **23**(1), 62–69.
2. Jakobus, U. and Landstorfer, F.M. (1995) "Improved PO-MM hybrid formulation for scattering from three-dimensional perfectly conducting bodies of arbitrary shape." *IEEE Transactions on Antennas and Propagation*, **43**(2), 162–169.
3. Jin, J.M., Ni Sean, S., and Lee, S.W. (1995) "Hybridization of SBR and FEM for scattering by large bodies with cracks and cavities." *IEEE Transactions on Antennas and Propagation*, **43**(10), 1130–1139.
4. Jin, J.M., Ling, F., Carolan, S. *et al.* (1998) "A hybrid SBR/MoM for analysis of scattering by small protrusion on a large conducting body." *IEEE Transactions on Antennas and Propagation*, **46**(9), 1349–1357.
5. Sheng, X.Q. and Jin, J.M. (1997) "A hybrid FEM/SBR method to compute scattering by large bodies with small protruding scatterers." *Microwave and Optical Technology Letters*, **15**(2), 78–84.
6. Sheng, X.Q., Jin, J.M., Song, J.M. *et al.* (1998) "On the formulation of hybrid finite-element boundary-integral methods for 3D scattering." *IEEE Transactions on Antennas and Propagation*, **46** (3), 303–311.
7. Sheng, X.Q., Yung, E.K.N., Chan, C.H. *et al.* (2000) "Scattering from large bodies with cracks and cavities by fast and accurate hybrid finite-element boundary-integral method." *IEEE Transactions on Antennas and Propagation*, **48**(8), 1153–1160.
8. Sheng, X.Q. and Yung, E.K.N. (2002) "Implementation and experiments of a hybrid algorithm of the MLFMA-enhanced FE-BI method for open-region inhomogeneous electromagnetic problems." *IEEE Transactions on Antennas and Propagation*, **50**(2), 163–167.
9. Sheng, X.Q. and Yung, E.K.N. (2002) "On the computing algorithms of the hybrid FEM/MLFMA." *Microwave and Optical Technology Letters*, **33**(4), 265–268.
10. Li, N., Saad, Y., and Chow, E. (2003) "Crout versions of ILU for general sparse matrices." *SIAM Journal on Scientific Computing*, **25**(2), 716–728.
11. Bollhofer, M. and Saad, Y. (2006) "Multilevel preconditioners constructed from inverse-based ILUs." *SIAM Journal on Scientific Computing*, **27**(5), 1627–1650.
12. Hodges, R.E. and Rahmat-Sammi, Y. (1997) "An iterative current-based hybrid method for complex structures." *IEEE Transactions on Antennas and Propagation*, **45**(2), 265–276.
13. Xu, S.J. and Sheng, X.Q. (1997) "Coupling of edge-element and mode-matching for multistep dielectric discontinuity in guiding structures." *IEEE Transactions on Microwave Theory and Techniques*, **45**(2), 284–287.
14. Sheng, X.Q. and Xu, S.J. (1997) "Edge-Element Analysis for Periodic Structures." *Acta Electronica Sinica*, **25**(12), 70–73 (in Chinese).

Index

Addition theorem, of Green's function (Gegenbauer's addition theorem), 77
Area coordinates, 48, 61, 159–160
Auxiliary differential equation method (ADE), 213–4

Basis functions (BFs), 31–4
  rooftop basis functions(RBFs), 130–1
  RWG basis functions, 32–4
Bessel function, 21, 77
Boundary conditions, 3–5
  absorbing boundary condition, 4, 208
    Mur's ABC, 287
    perfectly matched layer (PML), 208, 209–215
    Sommerfeld radiation condition, 4
      the first order absorbing boundary condition, 185
      the second order absorbing boundary condition, 186
  first-type (Dirichelet) boundary condition, 4
  impedance boundary condition, 4
  second-type (Neumann) boundary condition, 4

Charge conservation, law of, 2
Connection boundary, 208
  connecting condition, 208
Contour path FDTD (CP-FDTD or CFDTD), 220–2
Constitutive relations, 3

Debye form, 235
Diagonally dominant matrix, 36
Dirac delta, 24
Discontinuity problem, 170
Discretized integral equation (DIE), 76,
  combined field integral equation (CFIE), 73
    PMCHW equation, 21, 117–121, 124–127
  electric field integral equation (EFIE), 15, 16–7, 30, 34–6, 71–3, 111–127
    combined EFIE system (TENE), 116
    combined MFIE system (THNH), 116
    exterior electric field integral equation, 16
    interior electric field integral equation, 17
    normal electric-field integral equation (N-EFIE), 35
    tangential electric-field integral equation (T-EFIE), 35
  magnetic field integral equation (MFIE), 15, 16–7, 34–6, 71–3, 111–127
    exterior magnetic field integral equation, 16

Essentials of Computational Electromagnetics, First Edition. Xin-Qing Sheng and Wei Song.

© 2012 John Wiley & Sons Singapore Pte. Ltd. Published 2012 by John Wiley & Sons Singapore Pte. Ltd.

Discretized integral equation (*Continued*)
interior magnetic field integral equation, 17
normal magnetic-field integral equation (N-MFIE), 35
tangential magnetic-field integral equation (T-MFIE), 35
single integral equation, 18–9, 122–128
volume integral equation, (VIE), 19–20
Dispersive media, 235
Duffy's method, 37, 43–73
Dyadic green's function, 24

Edge-element basis functions, 159–161
Eigenmode problem, *see* Eigenvalue problem, 153–4, 174
eigenvalue equation, *see* Eigenvalue problem, 154, 164
Eigenvalue problem, 3
Electromagnetic equations, frequency-domain version of, 6
Error, 196
Equivalence principle, 9–11, 15, 18, 243
Huygens' principle, 10, 187, 253

Far-field, 19, 86
Fast multipole method (FMM), 74–86,
aggregation, 75–9
disaggregation, 75–9
multilevel fast-multipole algorithm (MLFMA), 79–86
translation, 75–9
Finite-Difference Time-Domain method (FDTD), 207
Finite element method (FEM), 153
Finite-element time-domain method (FETD), 200–3
Fredholm integral equation of the first kind, 15, 36, 71,115
Fredholm integral equation of the second kind, 15, 36, 71,115
Functional, variation of the, 174, 186

Galerkin methods, 34
Gauss' divergence theorem, (Gauss' theorem, divergence theorem), 6, 41
Green's function, 12
dyadic Green's function, 246
half-space dyadic Green's function, 246
Gauss–Legendre formulae, 36

Higher order elements, 196–200
interpolatory basis function, 196
hierarchical basis functions, 196
High-frequency asymptotic method, 244

Integro-differential operator, 14, 122, 127, 268
Interior resonance, 7, 73–4
Interpolation polynomials, 197
Inverse Fourier transformation (IFT), 5 , 21, 24

Jacobian (Jacobian determinant), 43, 44, 46, 49, 50, 58–9, 63, 70

Layered medium, 20
Line-matching, 34
Low frequency problem, 72–3
Lumped element, 237

Marching-on-time (MOT), 217
Maxwell's equations, 1
Method of Moments (MoM), 29
Mixed-potentials expression, 24
Mode-matching method, 271
Moment, referred to MoM, 31

Near-to-far-field (N/F) transformation, 227
Node-edge element, 192–6
Numerical dispersion, *see* error, 196

Parallel technology, 100–6
Perfect electric conductor (PEC), 4, 10, 29
Perfect magnetic conductor (PMC), 4, 11
Perfectly matched layer (PML), 208, 209–215
Physical optics method (PO), 244, 245–6, 248

Point-matching, 34
Principle value integration, 38

Radar, 30
Radar cross section (RCS), 88
  echo width, 140–1
  horizontally polarized RCS, 88
  vertically polarized RCS, 88
Radiation, 146
  radiation pattern, 146, 149,150
Reciprocity theorem, 24

Scalar Helmholtz equation, 12
Scalar potential, 11–3, 25–7
Shooting-bouncing ray method (SBR), 243
Singularity, 13, 36–71
  Singularity-cancellation technique, *see* Duffy's method, 43–73
  Singularity-extraction technique, 36–43
Sommerfeld integration, 22, 27
Spurious solutions, 159
Stability, 222–3
  Courant-Friedrich-Levy (CFL) condition, (Courant stability condition), 222–3

Taylor series expansion, 38, 40, 43
Testing functions (TFs), 31, 34, 113, 123
Time-domain far-field scattering field equations, 228
Total-field/scattered-field technique (TF/SF), 208
Transmission-line equation, 23

Uniqueness theorem, 6–7, 9, 11

Variational principle *see* Functional, variation of the, 158, 164
Vector Helmholtz equation, 12
Vector partial differential equation, 1
Vector potential, 11–2, 25–6
Vector integral equation (IE), 8–9
Vector wave equation, (*also* Vector EM wave equation), 1, 8

Wave number, 8

Yee's leapfrogging scheme, (*also* leapfrog, Yee's scheme), 208, 214, 216
Yee's spatial grid, 216