1

Eigenvalue Theory

Science is spectral analysis. Art is light synthesis.
– Karl Kraus (Austrian writer and journalist, 1874–1936)

The study of eigenvalue problems can be traced back to the eighteenth century, when Swiss mathematician and physicist Leonhard Euler (1707–1783) investigated the rotational motion of a rigid body. The word "eigen" is from German and means "own" or "belonging to," and was first used by German mathematician David Hilbert (1862–1943) to characterize eigenvalues and eigenvectors in 1904. Eigenvalue problems often arise in mathematics, physics, and engineering sciences. In linear algebra, an **eigenvector** of a linear transformation is a nonzero vector that changes by a scalar factor when the linear transformation acts on it. The scalar factor is called the **eigenvalue** corresponding to the eigenvector. Geometrically, this implies that the eigenvector is not rotated after transformation. The eigenvalue problem for a differential operator often results from the boundary value problems defined in a finite region. When the defining region is unbounded, the discrete eigenvalues become a continuum. A very useful technique for studying the eigenvalue problem is to establish the Rayleigh quotient for the eigenvalues and then use the calculus of variations to investigate the properties of eigenmodes. In physics, an **eigenmode** of a system is a possible state when the system is free of excitation, which might exist in the system on its own under certain conditions, and is also called an **eigenstate** of the system. The **method of eigenfunctions** is very similar to the Fourier series expansion in signal analysis, and will be used throughout this book. The method is based on the solution of an eigenvalue problem available from the system. An arbitrary state of the system can be expressed as a linear combination of the eigenmodes, and the expansion coefficients can then be determined from the source conditions or the initial values of the system. If only one or a few eigenmodes dominate in the linear combination, this will significantly simplify the analysis of the problem.

Foundations of Antenna Radiation Theory: Eigenmode Analysis, First Edition. Wen Geyi.

Published 2023 by John Wiley & Sons, Inc.

The modal theory for a scatterer plays an important role in antenna theory and designs. The basic idea behind the modal theory is to introduce the fundamental field patterns, called **modes**, so that the fields outside the scatterer can be expanded into a linear combination of these modes. There have been several modal theories for studying electromagnetic (EM) radiation and scattering problems (exterior boundary value problems). The **singularity expansion method** (SEM) is based on the analysis in complex frequency domain and formulated by electric field integral equation [1, 2]. The **natural resonant frequencies** arise from the requirement that a nontrivial current distribution exists on a conducting scatterer free of incident fields. The corresponding field patterns are called **natural resonant modes**. The natural resonant frequencies and the modes in SEM are complex, which significantly increases the computational time and the difficulty in numerical implementations. The **eigenmode expansion method** (EEM) expands the currents and the radiated fields in terms of the eigenmodes of an integral operator [3, 4]. Same with the SEM, the eigenvalues and the eigenmodes in EEM are complex numbers. The EEM is based on the eigenfunctions of integral equations and lacks a solid mathematical foundation. The integral operator involved in EEM is not symmetric, and it is therefore hard to prove the existence and completeness of the eigenfunctions. A more useful method for the study of scattering problem is the **singular function expansion**, which was first introduced by the German mathematician Erhard Schmidt (1876–1959) in 1907 [5], and has been applied to study various scattering problems [6, 7]. The theory of **characteristic mode** is another interesting modal notion and is carried out in the real frequency domain [8–11], of which the characteristic values (eigenvalues) and the corresponding characteristic modes are all real. In general, the characteristic values range from $-\infty$ to $+\infty$, among which those of the smallest magnitudes are the most important for radiation and scattering problems. The external resonant modes correspond to the zero characteristic values, and can be determined approximately by sweeping the frequency. It is noted that all the abovementioned modal formulations depend not only on the properties of the scatterer but also on the operating frequency.

The eigenvalue problems discussed in this book are derived from waveguide, cavity resonator, and spherical waveguide, whose eigenfunctions are independent of frequency and can thus be used to expand the fields in either frequency or time domain. The importance of eigenvalue theory in mathematics and physics cannot be overstated. There have been various methods developed to calculate eigenvalues and eigenfunctions, with the most important one being the variational method based on the Rayleigh quotient [12]. This chapter provides the necessary background information for later chapters. The Maxwell equations and the solution methods for partial differential equations (PDEs) are briefly introduced. The emphasis is upon the eigenvalue theory for operators, including the matrix and the

Laplacian on scalar and vector fields. The properties of eigenfunctions are derived from the Rayleigh quotient, and the Ritz method for the numerical solution of the Rayleigh quotient is demonstrated. Also included in this chapter is the Helmholtz theorem, which states that any vector field can be decomposed into the sum of an irrotational vector field and a solenoidal vector field. Such a decomposition has interesting applications in the modal expansion of fields and is the theoretical basis of introducing scalar and vector potentials. The Helmholtz theorem indicates that a vector field is fully determined by its divergence and curl. Indeed, Maxwell equations are nothing but a couple of rules that regulate the divergences and the curls of electric and magnetic fields according to impressed and induced sources. As a generalization of Helmholtz theorem, the eigenfunctions of curl operator are discussed, in terms of which the plane-wave expansions for the fields as well as the dyadic Green's functions can be obtained.

1.1 Maxwell Equations

Maxwell equations are a set of PDEs that unify electricity and magnetism and describe how electric and magnetic fields, as the functions of space and time, are generated by charges and currents and altered by each other. They have been proved to be very successful in explaining and predicting a variety of macroscopic EM phenomena.

1.1.1 Wave Equations

The **generalized Maxwell equations** that include both electric and magnetic sources consist of two vector equations and two scalar equations:

$$\begin{aligned}
\nabla \times \mathbf{H}(\mathbf{r},t) &= \frac{\partial \mathbf{D}(\mathbf{r},t)}{\partial t} + \mathbf{J}(\mathbf{r},t),\\
\nabla \times \mathbf{E}(\mathbf{r},t) &= -\frac{\partial \mathbf{B}(\mathbf{r},t)}{\partial t} - \mathbf{J}_m(\mathbf{r},t),\\
\nabla \cdot \mathbf{D}(\mathbf{r},t) &= \rho(\mathbf{r},t),\\
\nabla \cdot \mathbf{B}(\mathbf{r},t) &= \rho_m(\mathbf{r},t).
\end{aligned} \tag{1.1}$$

In the above, $\mathbf{r}$ is the observation point of the fields in meter (m) and t is the time in second (s), $\mathbf{H}$ is the **magnetic field intensity** measured in amperes per meter (A/m), $\mathbf{B}$ is the **magnetic induction intensity** measured in tesla (Wb/m^2), $\mathbf{E}$ is **electric field intensity** measured in volts per meter (V/m), $\mathbf{D}$ is the **electric induction intensity** measured in coulombs per square meter (C/m^2), $\mathbf{J}$ is **electric current density** measured in amperes per square meter (A/m^2), ρ is the **electric charge density** measured in coulombs per cubic meter (C/m^3), $\mathbf{J}_m$

is **magnetic current density** in volts per square meter (V/m^2), and ρ_m is **magnetic charge density** in webers per cubic meter (Wb/m^3). The first equation is **Ampère's law**, and it describes how the electric field changes according to the current density and magnetic field. The positive sign in the first equation indicates that the directions of the magnetomotive force and the electric current are related by the right-hand rule. The term $\partial \mathbf{D}/\partial t$ was introduced by Maxwell in 1861 and is called **displacement current**, which is necessary for the existence of wave solutions. The second equation is **Faraday's law**, and it characterizes how the magnetic field varies according to the electric field and equivalent magnetic current density. The minus sign in the second equation indicates that the directions of electromotive force and the magnetic current are related by the left-hand rule, which is required by **Lenz's law**. In other words, when an electromotive force is generated by a change of magnetic flux, the polarity of the induced electromotive force is such that it produces a current whose magnetic field opposes the change, which produces it. The third equation is **Coulomb's law**, and it says that the electric field depends on the charge distribution and obeys the inverse square law. The last equation shows that the magnetic field also obeys the inverse square law and depends on the equivalent magnetic charge distribution. It should be understood that none of the experiments had anything to do with waves at the time when Maxwell derived his equations. Maxwell equations imply more than the experimental facts. The **continuity equation** can be derived from (1.1):

$$\nabla \cdot \mathbf{J} = -\frac{\partial \rho}{\partial t}. \tag{1.2}$$

The electric charge density ρ and the electric current density $\mathbf{J}$ in Maxwell equations are free charge density and currents and they exclude charges and currents forming part of the structure of atoms and molecules. The bound charges and currents are regarded as material, which are not included in ρ and $\mathbf{J}$. The current density usually consists of two parts: $\mathbf{J} = \mathbf{J}_{con} + \mathbf{J}_{imp}$. Here, $\mathbf{J}_{imp}$ is referred to as external or **impressed current source**, which is independent of the fields and delivers energy to electric charges in a system. The impressed current source can be of electric and magnetic type as well as of non-electric or nonmagnetic origin. $\mathbf{J}_{con} = \sigma \mathbf{E}$, where σ is the **conductivity** of the medium in siemens per meter (S/m), denotes the **conduction current** induced by the impressed source $\mathbf{J}_{imp}$. Sometimes it is convenient to introduce an **impressed electric field** $\mathbf{E}_{imp}$ defined by $\mathbf{J}_{imp} = \sigma \mathbf{E}_{imp}$. In more general situation, one may write $\mathbf{J} = \mathbf{J}_{ind} + \mathbf{J}_{imp}$, where $\mathbf{J}_{ind}$ is the **induced current** by the impressed current $\mathbf{J}_{imp}$. The continuity equation for the magnetic current $\mathbf{J}_m$ and magnetic charges ρ_m can be derived from (1.1):

$$\nabla \cdot \mathbf{J}_m = -\frac{\partial \rho_m}{\partial t}. \tag{1.3}$$

The inclusions of magnetic sources $\mathbf{J}_m$ and ρ_m make Maxwell equations more symmetric although there has been no evidence that the magnetic current and charge are physically present (Appendix D gives an explanation of why the magnetic charge does not exist). The validity of introducing such concepts in Maxwell equations is justified by the equivalence principle, i.e. they are introduced as a mathematical equivalent to EM fields. For the time-harmonic (sinusoidal) fields with a single frequency ω, the generalized Maxwell equations (1.1) reduce to

$$\begin{aligned} \nabla \times \mathbf{H}(\mathbf{r}) &= j\omega\mathbf{D}(\mathbf{r}) + \mathbf{J}(\mathbf{r}), \\ \nabla \times \mathbf{E}(\mathbf{r}) &= -j\omega\mathbf{B}(\mathbf{r}) - \mathbf{J}_m(\mathbf{r}), \\ \nabla \cdot \mathbf{D}(\mathbf{r}) &= \rho(\mathbf{r}), \\ \nabla \cdot \mathbf{B}(\mathbf{r}) &= \rho_m(\mathbf{r}), \end{aligned} \tag{1.4}$$

where all the field quantities denote the **complex amplitudes** (**phasors**), defined by

$$\mathbf{E}(\mathbf{r}, t) = \text{Re}\left[\mathbf{E}(\mathbf{r})e^{j\omega t}\right], \quad \text{etc.}$$

For brevity, the same notations will be used for both time-domain and frequency-domain quantities. The **boundary conditions** on the surface between two different media can be easily obtained as follows:

$$\begin{aligned} \mathbf{u}_n \times (\mathbf{H}_1 - \mathbf{H}_2) &= \mathbf{J}_s, \\ \mathbf{u}_n \times (\mathbf{E}_1 - \mathbf{E}_2) &= -\mathbf{J}_{ms}, \\ \mathbf{u}_n \cdot (\mathbf{D}_1 - \mathbf{D}_2) &= \rho_s, \\ \mathbf{u}_n \cdot (\mathbf{B}_1 - \mathbf{B}_2) &= \rho_{ms}, \end{aligned} \tag{1.5}$$

where $\mathbf{u}_n$ is the unit normal of the boundary directed from medium 2 to medium 1; $\mathbf{J}_s$ and ρ_s are the **surface current density** and **surface charge density**, respectively; $\mathbf{J}_{ms}$ and ρ_{ms} are the **surface magnetic current density** and **surface magnetic charge density**, respectively.

Maxwell equations (without magnetic sources) are a set of 7 equations involving 16 unknowns (i.e. five vectors **E**, **H**, **B**, **D**, **J** and one scalar ρ and the last equation of (1.1) is not independent). To determine the fields, nine more equations are needed, and they are given by the **generalized constitutive relations**:

$$\mathbf{D} = f(\mathbf{E}, \mathbf{H}), \quad \mathbf{B} = g(\mathbf{E}, \mathbf{H})$$

together with the **generalized Ohm's law**

$$\mathbf{J} = h(\mathbf{E}, \mathbf{H})$$

if the medium is conducting. The constitutive relations establish the connections between field quantities and reflect the properties of the medium, and they are totally independent of the Maxwell equations. For time-harmonic fields, the constitutive relations for a **bi-anisotropic medium** are defined by

$$\mathbf{D} = \overleftrightarrow{\boldsymbol{\varepsilon}} \cdot \mathbf{E} + \overleftrightarrow{\boldsymbol{\xi}} \cdot \mathbf{H},$$
$$\mathbf{B} = \overleftrightarrow{\boldsymbol{\zeta}} \cdot \mathbf{E} + \overleftrightarrow{\boldsymbol{\mu}} \cdot \mathbf{H},$$

where $\overleftrightarrow{\boldsymbol{\mu}}$, $\overleftrightarrow{\boldsymbol{\varepsilon}}$, $\overleftrightarrow{\boldsymbol{\xi}}$, and $\overleftrightarrow{\boldsymbol{\zeta}}$ are dyadics. The medium is called **anisotropic** if $\overleftrightarrow{\boldsymbol{\xi}} = \overleftrightarrow{\boldsymbol{\zeta}} = 0$. The medium is called **isotropic** if $\overleftrightarrow{\boldsymbol{\xi}} = \overleftrightarrow{\boldsymbol{\zeta}} = 0$ and $\overleftrightarrow{\boldsymbol{\mu}}$ and $\overleftrightarrow{\boldsymbol{\varepsilon}}$ are, respectively, degenerated to $\overleftrightarrow{\boldsymbol{\mu}} = \mu\overleftrightarrow{\mathbf{I}}$ and $\overleftrightarrow{\boldsymbol{\varepsilon}} = \varepsilon\overleftrightarrow{\mathbf{I}}$, where $\overleftrightarrow{\mathbf{I}}$ is the identity dyadic; μ and ε are referred to as **permittivity** and **permeability** of the medium, respectively.

The EM wave equations are second-order PDEs that describe the propagation of EM waves through a medium. On elimination of **E** or **H** in the generalized Maxwell equations (1.4), the **wave equations** for the time-harmonic fields in an inhomogeneous and anisotropic medium are

$$\begin{aligned}\nabla \times \overleftrightarrow{\boldsymbol{\mu}}^{-1} \cdot \nabla \times \mathbf{E}(\mathbf{r}) - \omega^2 \overleftrightarrow{\boldsymbol{\varepsilon}} \cdot \mathbf{E}(\mathbf{r}) &= -j\omega\mathbf{J}(\mathbf{r}) - \nabla \times \overleftrightarrow{\boldsymbol{\mu}}^{-1} \cdot \mathbf{J}_m,\\ \nabla \times \overleftrightarrow{\boldsymbol{\varepsilon}}^{-1} \cdot \nabla \times \mathbf{H}(\mathbf{r}) - \omega^2 \overleftrightarrow{\boldsymbol{\mu}} \cdot \mathbf{H}(\mathbf{r}) &= -j\omega\mathbf{J}_m(\mathbf{r}) + \nabla \times \overleftrightarrow{\boldsymbol{\varepsilon}}^{-1} \cdot \mathbf{J}.\end{aligned} \tag{1.6}$$

If the medium is homogeneous and isotropic, μ and ε are constants, and the wave equations (1.6) are simplified to

$$\begin{aligned}\nabla \times \nabla \times \mathbf{E} - k^2\mathbf{E} &= -j\omega\mu\mathbf{J} - \nabla \times \mathbf{J}_m,\\ \nabla \times \nabla \times \mathbf{H} - k^2\mathbf{H} &= -j\omega\varepsilon\mathbf{J}_m + \nabla \times \mathbf{J},\end{aligned} \tag{1.7}$$

where $k = \omega\sqrt{\mu\varepsilon}$ is the **wavenumber**. It can be seen that the source terms on the right-hand side of (1.7) are very complicated. To simplify the analysis, the EM potential functions can be introduced. The wave equations may be used to solve the following three different field problems:

1) EM fields in source-free region: wave propagations in space and waveguides, wave oscillation in cavity resonators, etc.
2) EM fields generated by known source distributions: antenna radiations, excitations in waveguides and cavity resonators, etc.
3) Interaction of fields and sources: wave propagation in plasma, coupling between electron beams and propagation mechanism, etc.

1.1.2 Properties of Electromagnetic Fields

A number of theorems can be derived from Maxwell equations [13–16], and they usually bring deep physical insight into the EM field problems. When applied properly, these theorems can simplify the problems dramatically.

1.1.2.1 Superposition Theorem

Superposition theorem applies to all linear systems. Suppose that the impressed current source $\mathbf{J}_{imp}$ can be expressed as a linear combination of independent impressed current sources $\mathbf{J}_{imp}^k$ ($k = 1, 2, \ldots, n$):

$$\mathbf{J}_{imp} = \sum_{k=1}^{n} a_k \mathbf{J}_{imp}^k,$$

where a_k ($k = 1, 2, \ldots, n$) are arbitrary constants. If $\mathbf{E}^k$ and $\mathbf{H}^k$ are the fields produced by the source $\mathbf{J}_{imp}^k$, the **superposition theorem** for EM fields asserts that the fields $\mathbf{E} = \sum_{k=1}^{n} a_k \mathbf{E}^k$ and $\mathbf{H} = \sum_{k=1}^{n} a_k \mathbf{H}^k$ are a solution of Maxwell equations produced by the source $\mathbf{J}_{imp}$.

1.1.2.2 Conservation of Electromagnetic Field Energy

The law of **conservation of EM field energy** is known as the **Poynting theorem**, named after the English physicist John Henry Poynting (1852–1914). It can be found from (1.1) that

$$-\mathbf{J}_{imp} \cdot \mathbf{E} - \mathbf{J}_{ind} \cdot \mathbf{E} = \nabla \cdot \mathbf{S} + \mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t} + \mathbf{H} \cdot \frac{\partial \mathbf{B}}{\partial t}, \tag{1.8}$$

where $\mathbf{J} = \mathbf{J}_{imp} + \mathbf{J}_{ind}$ has been assumed. In a region V bounded by S, the integral form of (1.8) is

$$-\int_V \mathbf{J}_{imp} \cdot \mathbf{E} dV = \int_V \mathbf{J}_{ind} \cdot \mathbf{E} dV + \int_S \mathbf{S} \cdot \mathbf{u}_n dS + \int_V \left(\mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t} + \mathbf{H} \cdot \frac{\partial \mathbf{B}}{\partial t} \right) dV, \tag{1.9}$$

where $\mathbf{u}_n$ is the unit outward normal of S, and $\mathbf{S} = \mathbf{E} \times \mathbf{H}$ is the **Poynting vector** representing the EM power-flow density measured in watts per square meter ($\mathrm{W/m^2}$). It is assumed that this explanation holds for all media. Thus, the left-hand side of the above equation stands for the power supplied by the impressed current source. The first term on the right-hand side is the work done per second by the electric field to maintain the current in the conducting part of the system. The second term denotes the EM power flowing out of S. The last term can be interpreted as the work done per second by the impressed source to establish the fields. The total field energy density w of the EM fields may be defined as follows:

$$dw = \left(\mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t} + \mathbf{H} \cdot \frac{\partial \mathbf{B}}{\partial t} \right) dt. \tag{1.10}$$

If all the sources and fields are zero at $t = -\infty$, the total field energy density can be written as

$$w = w_e + w_m, \tag{1.11}$$

where w_e and w_m are the **electric field energy density** and **magnetic field energy density**, respectively,

$$\begin{aligned} w_e &= \frac{1}{2}\mathbf{E}\cdot\mathbf{D} + \int_{-\infty}^{t} \frac{1}{2}\left(\mathbf{E}\cdot\frac{\partial \mathbf{D}}{\partial t} - \mathbf{D}\cdot\frac{\partial \mathbf{E}}{\partial t}\right)dt, \\ w_m &= \frac{1}{2}\mathbf{H}\cdot\mathbf{B} + \int_{-\infty}^{t} \frac{1}{2}\left(\mathbf{H}\cdot\frac{\partial \mathbf{B}}{\partial t} - \mathbf{B}\cdot\frac{\partial \mathbf{H}}{\partial t}\right)dt. \end{aligned} \tag{1.12}$$

Equation (1.9) can thus be written as

$$-\int_V \mathbf{J}_{imp}\cdot\mathbf{E}dV = \int_V \mathbf{J}_{ind}\cdot\mathbf{E}dV + \int_S \mathbf{S}\cdot\mathbf{u}_n dS + \frac{\partial}{\partial t}\int_V (w_e + w_m)dV. \tag{1.13}$$

In general, the field energy density w does not represent the stored field energy density in the fields: the energy temporarily located in the fields and completely recoverable when the fields are reduced to zero. The field energy density w given by (1.11) can be considered as the stored field energy density only if the medium is lossless (i.e. $\nabla\cdot\mathbf{S} = 0$). If the medium is isotropic and time-invariant, (1.12) reduces to

$$w_e = \frac{1}{2}\mathbf{E}\cdot\mathbf{D},\quad w_m = \frac{1}{2}\mathbf{H}\cdot\mathbf{B}. \tag{1.14}$$

For the time-harmonic fields, the **time averages** of Poynting vector, the field energy densities in (1.14) over one period of the sinusoidal wave $e^{j\omega t}$, denoted T, are

$$\begin{aligned} \frac{1}{T}\int_0^T \mathbf{E}\times\mathbf{H}dt &= \frac{1}{2}\,\mathrm{Re}\left(\mathbf{E}\times\overline{\mathbf{H}}\right), \\ \frac{1}{T}\int_0^T \frac{1}{2}\mathbf{E}\cdot\mathbf{D}dt &= \frac{1}{4}\,\mathrm{Re}\left(\mathbf{E}\cdot\overline{\mathbf{D}}\right), \\ \frac{1}{T}\int_0^T \frac{1}{2}\mathbf{H}\cdot\mathbf{B}dt &= \frac{1}{4}\,\mathrm{Re}\left(\mathbf{H}\cdot\overline{\mathbf{B}}\right). \end{aligned}$$

The energy balance relations for the time-harmonic fields can be derived by using complex variable analysis [17]. Let $s = \alpha + j\omega$ denote the complex frequency. For an arbitrary analytic function $f(\mathbf{r}, s)$, the Cauchy–Riemann conditions imply

$$\frac{\partial f(\mathbf{r},s)}{\partial \alpha} = -j\frac{\partial f(\mathbf{r},s)}{\partial \omega}. \tag{1.15}$$

For sufficiently small α, the analytic function has the first-order expansion

$$\begin{aligned} f(\mathbf{r},s) &= f(\mathbf{r},\alpha + j\omega) \approx f(\mathbf{r},s)\big|_{\alpha=0} + \alpha\frac{\partial f(\mathbf{r},s)}{\partial \alpha}\bigg|_{\alpha=0} \\ &= f(\mathbf{r},j\omega) - j\alpha\frac{\partial f(\mathbf{r},j\omega)}{\partial \omega}, \end{aligned} \tag{1.16}$$

where (1.15) has been used. If the Laplace transform

$$\mathbf{E}(\mathbf{r},s) = \int_{-\infty}^{\infty} \mathbf{E}(\mathbf{r},t)e^{-st}dt$$

is applied to the first two equations of the Maxwell equations (1.1) (with $\mathbf{J}_m = 0$), one may find

$$\begin{aligned} \nabla \times \mathbf{H}(\mathbf{r},s) &= \mathbf{J}(\mathbf{r},s) + s\mathbf{D}(\mathbf{r},s), \\ \nabla \times \mathbf{E}(\mathbf{r},s) &= -s\mathbf{B}(\mathbf{r},s). \end{aligned} \tag{1.17}$$

From the vector identity $\nabla \cdot (\mathbf{a} \times \mathbf{b}) = \mathbf{b} \cdot \nabla \times \mathbf{a} - \mathbf{a} \cdot \nabla \times \mathbf{b}$ and (1.17), one may obtain

$$\begin{aligned} \nabla \cdot \left[\mathbf{E}(\mathbf{r},s) \times \overline{\mathbf{H}}(\mathbf{r},s)\right] = &-\mathbf{E}(\mathbf{r},s) \cdot \overline{\mathbf{J}}(\mathbf{r},s) \\ &-\alpha\left[\overline{\mathbf{H}}(\mathbf{r},s) \cdot \mathbf{B}(\mathbf{r},s) + \mathbf{E}(\mathbf{r},s) \cdot \overline{\mathbf{D}}(\mathbf{r},s)\right] \\ &-j\omega\left[\overline{\mathbf{H}}(\mathbf{r},s) \cdot \mathbf{B}(\mathbf{r},s) - \mathbf{E}(\mathbf{r},s) \cdot \overline{\mathbf{D}}(\mathbf{r},s)\right], \end{aligned} \tag{1.18}$$

where the bar denotes complex conjugate. By use of the expansion (1.16), the first-order expansion for the electric field can be expressed by

$$\begin{aligned} \mathbf{E}(\mathbf{r},s) &= \mathbf{E}(\mathbf{r},\alpha + j\omega) \approx \mathbf{E}(\mathbf{r},j\omega) - j\alpha\frac{\partial \mathbf{E}(\mathbf{r},j\omega)}{\partial \omega} \\ &= \mathbf{E}(\mathbf{r}) - j\alpha\frac{\partial \mathbf{E}(\mathbf{r})}{\partial \omega}, \end{aligned} \tag{1.19}$$

where $\mathbf{E}(\mathbf{r})$ denotes the phasor for the time-harmonic electric field. By introducing the first-order expansions for all the field quantities into (1.18), one may arrive at

$$\begin{aligned}
&\nabla\cdot\left[\mathbf{E}(\mathbf{r})\times\overline{\mathbf{H}}(\mathbf{r})\right]+j\alpha\nabla\cdot\left[\mathbf{E}(\mathbf{r})\times\frac{\partial\overline{\mathbf{H}}(\mathbf{r})}{\partial\omega}-\frac{\partial\mathbf{E}(\mathbf{r})}{\partial\omega}\times\overline{\mathbf{H}}(\mathbf{r})\right]\\
&\quad=-\mathbf{E}(\mathbf{r})\cdot\overline{\mathbf{J}}(\mathbf{r})-j\alpha\left[\mathbf{E}(\mathbf{r})\cdot\frac{\partial\overline{\mathbf{J}}(\mathbf{r})}{\partial\omega}-\frac{\partial\mathbf{E}(\mathbf{r})}{\partial\omega}\cdot\overline{\mathbf{J}}(\mathbf{r})\right]\\
&\quad-j\omega\left[\mathbf{B}(\mathbf{r})\cdot\overline{\mathbf{H}}(\mathbf{r})-\mathbf{E}(\mathbf{r})\cdot\overline{\mathbf{D}}(\mathbf{r})\right]\\
&\quad-\alpha\left[\mathbf{B}(\mathbf{r})\cdot\overline{\mathbf{H}}(\mathbf{r})+\mathbf{E}(\mathbf{r})\cdot\overline{\mathbf{D}}(\mathbf{r})\right]\\
&\quad-\alpha\omega\left[\overline{\mathbf{H}}(\mathbf{r})\cdot\frac{\partial\mathbf{B}(\mathbf{r})}{\partial\omega}-\mathbf{B}(\mathbf{r})\cdot\frac{\partial\overline{\mathbf{H}}(\mathbf{r})}{\partial\omega}\right]\\
&\quad-\alpha\omega\left[\mathbf{E}(\mathbf{r})\cdot\frac{\partial\overline{\mathbf{D}}(\mathbf{r})}{\partial\omega}-\overline{\mathbf{D}}(\mathbf{r})\cdot\frac{\partial\mathbf{E}(\mathbf{r})}{\partial\omega}\right].
\end{aligned}\tag{1.20}$$

Comparing the coefficients of similar terms containing α on both sides of (1.20) yields the energy balance relation

$$\begin{aligned}
&\frac{1}{4}\mathbf{E}\cdot\overline{\mathbf{D}}+\frac{1}{4}\omega\left(\mathbf{E}\cdot\frac{\partial\overline{\mathbf{D}}}{\partial\omega}-\overline{\mathbf{D}}\cdot\frac{\partial\mathbf{E}}{\partial\omega}\right)+\frac{1}{4}\mathbf{B}\cdot\overline{\mathbf{H}}+\frac{1}{4}\omega\left(\overline{\mathbf{H}}\cdot\frac{\partial\mathbf{B}}{\partial\omega}-\mathbf{B}\cdot\frac{\partial\overline{\mathbf{H}}}{\partial\omega}\right)\\
&\quad=-j\frac{1}{4}\nabla\cdot\left(\mathbf{E}\times\frac{\partial\overline{\mathbf{H}}}{\partial\omega}-\frac{\partial\mathbf{E}}{\partial\omega}\times\overline{\mathbf{H}}\right)-j\frac{1}{4}\left(\mathbf{E}\cdot\frac{\partial\overline{\mathbf{J}}}{\partial\omega}-\frac{\partial\mathbf{E}}{\partial\omega}\cdot\overline{\mathbf{J}}\right),
\end{aligned}\tag{1.21}$$

and the well-known Poynting theorem for time-harmonic fields

$$-\frac{1}{2}\mathbf{E}\cdot\overline{\mathbf{J}}=\nabla\cdot\frac{1}{2}\left(\mathbf{E}\times\overline{\mathbf{H}}\right)+j2\omega\left(\frac{1}{4}\mathbf{B}\cdot\overline{\mathbf{H}}-\frac{1}{4}\mathbf{E}\cdot\overline{\mathbf{D}}\right).\tag{1.22}$$

Consequently, the complex analysis produces two energy balance relations simultaneously. In this sense, the complex analysis implies more than the real analysis. It should be noted that the Poynting theorem (1.9) in time domain and the Poynting theorem (1.22) in frequency domain are independent. This property can be used to find the stored field energies of small antenna [18]. The physical implication of the energy conservation law (1.21) becomes clear if it is decomposed into the real and imaginary parts. The real part is given by

$$w_e+w_m=\nabla\cdot\mathrm{Im}\frac{1}{4}\left(\mathbf{E}\times\frac{\partial\overline{\mathbf{H}}}{\partial\omega}-\frac{\partial\mathbf{E}}{\partial\omega}\times\overline{\mathbf{H}}\right)+\mathrm{Im}\frac{1}{4}\left(\mathbf{E}\cdot\frac{\partial\overline{\mathbf{J}}}{\partial\omega}-\frac{\partial\mathbf{E}}{\partial\omega}\cdot\overline{\mathbf{J}}\right),\tag{1.23}$$

where w_e and w_m are defined by

$$w_e=\frac{1}{4}\,\mathrm{Re}\,\mathbf{E}\cdot\overline{\mathbf{D}}+\frac{1}{4}\omega\,\mathrm{Re}\left(\mathbf{E}\cdot\frac{\partial\overline{\mathbf{D}}}{\partial\omega}-\overline{\mathbf{D}}\cdot\frac{\partial\mathbf{E}}{\partial\omega}\right),\tag{1.24}$$

$$w_m = \frac{1}{4}\,\mathrm{Re}\,\mathbf{H}\cdot\overline{\mathbf{B}} + \frac{1}{4}\omega\,\mathrm{Re}\left(\mathbf{H}\cdot\frac{\partial\overline{\mathbf{B}}}{\partial\omega} - \overline{\mathbf{B}}\cdot\frac{\partial\mathbf{H}}{\partial\omega}\right). \tag{1.25}$$

The expressions (1.24) and (1.25) were first derived by the author in [17], and they are shown to, respectively, represent the (time averaged) **stored electric field energy density** and the **stored magnetic field energy density** in an arbitrary medium. In Chapter 4, a new narrow-band approach will be introduced to rederive energy expressions (1.24) and (1.25). The stored field energies can be decomposed into the sum of two distinct parts: the dominant (nondispersive) parts defined by

$$w_e^{dom} = \frac{1}{4}\,\mathrm{Re}\,\mathbf{E}\cdot\overline{\mathbf{D}}, \tag{1.26}$$

$$w_m^{dom} = \frac{1}{4}\,\mathrm{Re}\,\mathbf{H}\cdot\overline{\mathbf{B}}, \tag{1.27}$$

and the dispersive parts defined by

$$w_e^{dis} = \frac{1}{4}\,\mathrm{Re}\,\omega\left(\mathbf{E}\cdot\frac{\partial\overline{\mathbf{D}}}{\partial\omega} - \overline{\mathbf{D}}\cdot\frac{\partial\mathbf{E}}{\partial\omega}\right), \tag{1.28}$$

$$w_m^{dis} = \frac{1}{4}\,\mathrm{Re}\,\omega\left(\mathbf{H}\cdot\frac{\partial\overline{\mathbf{B}}}{\partial\omega} - \overline{\mathbf{B}}\cdot\frac{\partial\mathbf{H}}{\partial\omega}\right), \tag{1.29}$$

which are caused by the dispersion of materials. The imaginary part of (1.21) gives

$$\begin{aligned} w_{ed} - w_{md} &= -\,\mathrm{Im}\,\pi\omega\left(\mathbf{E}\cdot\frac{\partial\overline{\mathbf{D}}}{\partial\omega} - \overline{\mathbf{D}}\cdot\frac{\partial\mathbf{E}}{\partial\omega}\right) - \mathrm{Im}\,\pi\omega\left(\overline{\mathbf{H}}\cdot\frac{\partial\mathbf{B}}{\partial\omega} - \mathbf{B}\cdot\frac{\partial\overline{\mathbf{H}}}{\partial\omega}\right) \\ &\quad - \nabla\cdot\,\mathrm{Re}\,\pi\left(\mathbf{E}\times\frac{\partial\overline{\mathbf{H}}}{\partial\omega} - \frac{\partial\mathbf{E}}{\partial\omega}\times\overline{\mathbf{H}}\right) - \mathrm{Re}\,\pi\left(\mathbf{E}\cdot\frac{\partial\overline{\mathbf{J}}}{\partial\omega} - \frac{\partial\mathbf{E}}{\partial\omega}\cdot\overline{\mathbf{J}}\right), \end{aligned} \tag{1.30}$$

where w_{ed} and w_{md} are the (average) **dissipated electric field energy density** and **dissipated magnetic field energy density**, respectively, defined by

$$w_{ed} = \mathrm{Im}\,\pi\mathbf{E}\cdot\overline{\mathbf{D}}, \tag{1.31}$$

$$w_{md} = \mathrm{Im}\,\pi\mathbf{H}\cdot\overline{\mathbf{B}}. \tag{1.32}$$

As a result, the energy balance relation (1.21) gives two expressions: one is for the sum of stored electric and magnetic field energies and the other is for the difference of the dissipated electric and magnetic field energies, both expressions being valid in an arbitrary medium. It will be shown in Chapter 4 that (1.23) logically gives the definition of stored field energy of antenna in an arbitrary medium.

Remark 1.1 The quantities $\frac{1}{4}\operatorname{Re}\mathbf{H}\cdot\overline{\mathbf{B}}$ and $\frac{1}{4}\operatorname{Re}\mathbf{E}\cdot\overline{\mathbf{D}}$ in the Poynting theorem (1.22) only represent the nondispersive parts (1.26) and (1.27). For this reason, the Poynting theorem (1.22) cannot be considered as an energy balance relation for the time-harmonic fields in a general medium. □

1.1.2.3 Equivalence Theorem

It is known that there is no answer to the question of whether field or source is primary. The equivalence theorem just indicates that the distinction between the field and source is kind of blurred. Let V be an arbitrary region bounded by S, as shown in Figure 1.1. Two sources that produce the same fields inside a region are said to be **equivalent** within that region. Similarly, two EM fields $\{\mathbf{E}_1, \mathbf{D}_1, \mathbf{H}_1, \mathbf{B}_1\}$ and $\{\mathbf{E}_2, \mathbf{D}_2, \mathbf{H}_2, \mathbf{B}_2\}$ are said to be equivalent inside a region if they both satisfy the Maxwell equations and are equal in that region. The main application of the equivalence theorem is to find equivalent sources to replace the influences of substance (the medium is homogenized), so that the formulae for retarding potentials can be used. The equivalent sources may be located inside S (equivalent volume sources) or on S (equivalent surface sources). One of the important equivalence theorems is found by American mathematician Sergei Alexander Schelkunoff (1897–1992) and English mathematician Augustus Edward Hough Love (1863–1940), whose derivation can be found in [16].

Schelkunoff-Love Equivalence Theorem: Let $\{\mathbf{E}, \mathbf{D}, \mathbf{H}, \mathbf{B}\}$ be the EM fields with source confined in S. The following surface sources on S

$$\mathbf{J}_s = \mathbf{u}_n \times \mathbf{H}, \quad \mathbf{J}_{ms} = -\mathbf{u}_n \times \mathbf{E}, \tag{1.33}$$

produce the same fields $\{\mathbf{E}, \mathbf{D}, \mathbf{H}, \mathbf{B}\}$ outside S and a zero field inside S. □

Since the sources in (1.33) produce a zero field inside S, the interior of S may be filled with a perfect electric conductor. By use of the Lorentz reciprocity theorem [see (1.36)], it can be shown that the surface electric current pressed tightly on the

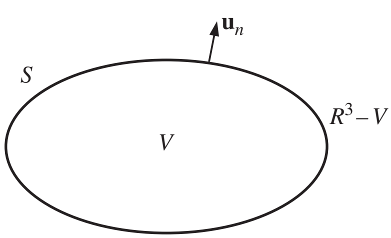

Figure 1.1 Equivalence theorem.

perfect conductor does not produce fields. As a result, only the surface magnetic current is needed in (1.33). Similarly, the interior of S may be filled with a perfect magnetic conductor, and in this case the surface magnetic current does not produce fields and only the surface electric current is needed in (1.33). In both cases, one cannot directly apply the vector potential formula even if the medium outside S is homogeneous.

1.1.2.4 Reciprocity

A linear system is said to be **reciprocal** if the response of the system with a particular load and a source is the same as the response when the source and the load are interchanged. Consider two sets of time-harmonic sources, $\mathbf{J}_1$, $\mathbf{J}_{m1}$ and $\mathbf{J}_2$, $\mathbf{J}_{m2}$, of the same frequency in the same linear medium. The fields produced by the two sources are, respectively, denoted by $\mathbf{E}_1$, $\mathbf{H}_1$ and $\mathbf{E}_2$, $\mathbf{H}_2$. The reciprocity can be stated as

$$\begin{aligned}\int_V (\mathbf{E}_2 \cdot \mathbf{J}_1 - \mathbf{H}_2 \cdot \mathbf{J}_{m1}) dV &= \int_V (\mathbf{E}_1 \cdot \mathbf{J}_2 - \mathbf{H}_1 \cdot \mathbf{J}_{m2}) dV \\ &\quad + \int_S (\mathbf{E}_1 \times \mathbf{H}_2 - \mathbf{E}_2 \times \mathbf{H}_1) \cdot \mathbf{u}_n dS,\end{aligned} \tag{1.34}$$

where V is a finite region bounded by S. If both sources are outside S, the surface integral in (1.34) is zero. If both sources are inside S, it can be shown that the surface integral is also zero by using the radiation condition. Therefore, one obtains the **Lorentz form of reciprocity**

$$\int_S (\mathbf{E}_1 \times \mathbf{H}_2 - \mathbf{E}_2 \times \mathbf{H}_1) \cdot \mathbf{u}_n dS = 0 \tag{1.35}$$

and the **Rayleigh–Carson form of reciprocity**

$$\int_V (\mathbf{E}_2 \cdot \mathbf{J}_1 - \mathbf{H}_2 \cdot \mathbf{J}_{m1}) dV = \int_V (\mathbf{E}_1 \cdot \mathbf{J}_2 - \mathbf{H}_1 \cdot \mathbf{J}_{m2}) dV. \tag{1.36}$$

If the surface S only contains the sources $\mathbf{J}_1(\mathbf{r})$ and $\mathbf{J}_{m1}(\mathbf{r})$, (1.34) becomes

$$\int_V (\mathbf{E}_2 \cdot \mathbf{J}_1 - \mathbf{H}_2 \cdot \mathbf{J}_{m1}) dV = \int_S (\mathbf{E}_2 \cdot \mathbf{u}_n \times \mathbf{H}_1 - \mathbf{H}_2 \cdot \mathbf{E}_1 \times \mathbf{u}_n) dS.$$

This is the familiar form of **Huygens' principle**. The EM reciprocity theorem (1.35) can also be generalized to an anisotropic medium.

1.2 Methods for Partial Differential Equations

Various analytic and numerical methods for the solution of PDEs have been developed [19–21]. Linear PDEs are generally solved by means of the **method of separation of variables**, the **method of Green's function**, named after the British mathematician George Green (1793–1841), and the **variational method**. Some usual trinities for PDEs are summarized in Table 1.1.

1.2.1 Method of Separation of Variables

The study of eigenvalue problems has its roots in the **method of separation of variables** or series solutions of PDEs. The basic idea of separation of variables is to seek a solution of the form of a product of functions, each of which depends on one variable only, so that the solution of original PDEs may reduce to the solution of ordinary differential equations. The latter is usually solved by the power series methods, resulting in various special functions. The method of separation of variables, also called Fourier method, was first introduced by Swiss mathematician Johann Bernoulli (1667–1748) between the years 1694 and 1697. The Helmholtz equation will be used to illustrate the procedure. The **Helmholtz equation**, named after the German physicist Hermann Ludwig Ferdinand von

Table 1.1 Some trinities for PDEs.

Trinity	Description
Three types of PDEs	Elliptical, hyperbolic, and parabolic.
Three types of problems	Boundary value problems, initial value problems, and eigenvalue problems.
Three types of boundary conditions	Dirichlet boundary condition, named after the German mathematician Johann Peter Gustav Lejeune Dirichlet (1805–1859); Neumann boundary condition, named after the German mathematician Carl Gottfried Neumann (1832–1925); and Robin boundary condition, named after the French mathematician Victor Gustave Robin (1855–1897).
Three mathematical tools	Divergence theorem, inequalities, and convergence theorems.
Three analytical methods	Method of separation of variables, method of Green's function, and variational method.
Three numerical methods	Finite element method, finite difference method, and moment method.

Helmholtz (1821–1894), also called **reduced wave equation**, is the time-independent form of wave equation, and is defined by

$$\left(\nabla^2 + k^2\right)u = 0, \tag{1.37}$$

where k is a constant. When k is zero, the Helmholtz equation reduces to the Laplace equation, named after the French mathematician Pierre-Simon marquis de Laplace (1749–1827). The Helmholtz equation is separable in 11 orthogonal coordinate systems [22]. The separated solutions form a subset of all solutions of (1.37) and can be served as a basis in terms of which all solutions of (1.37) can be expressed as a linear combination of the separated solutions.

1.2.1.1 Rectangular Coordinate System

In rectangular coordinate system, Helmholtz equation (1.37) becomes

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} + k^2 u = 0. \tag{1.38}$$

One may seek a solution in the form of product of three functions of one coordinate each

$$u = X(x)Y(y)Z(z). \tag{1.39}$$

If this is substituted into (1.38), one obtains

$$\frac{1}{X}\frac{d^2X}{dx^2} + \frac{1}{Y}\frac{d^2Y}{dy^2} + \frac{1}{Z}\frac{d^2Z}{dz^2} + k^2 = 0. \tag{1.40}$$

Since k is a constant and each term depends on one variable only and can change independently, the left-hand side of (1.40) can sum to zero for all coordinate values only if each term is a constant. Thus,

$$\begin{aligned} &\frac{d^2X}{dx^2} + k_x^2 X = 0, \\ &\frac{d^2Y}{dy^2} + k_y^2 Y = 0, \\ &\frac{d^2Z}{dz^2} + k_z^2 Z = 0, \end{aligned} \tag{1.41}$$

where k_x, k_y, and k_z are **separation constants** and satisfy

$$k_x^2 + k_y^2 + k_z^2 = k^2. \tag{1.42}$$

The solutions of (1.41) are harmonic functions, denoted by $X(k_x x)$, $Y(k_y y)$, and $Z(k_z z)$, and they are any linear combination of the following independent **harmonic functions**:

$$e^{ik_\alpha \alpha}, e^{-ik_\alpha \alpha}, \cos k_\alpha \alpha, \sin k_\alpha \alpha \ (\alpha = x, y, z). \tag{1.43}$$

Consequently, the solution (1.39) may be expressed as

$$u = X(k_x x) Y\left(k_y y\right) Z(k_z z). \tag{1.44}$$

The separation constants k_x, k_y, and k_z are also called **eigenvalues**, and they are determined by the boundary conditions. The corresponding solutions (1.44) are called **eigenfunctions** or **elementary wavefunctions**. The general solution of (1.38) can be expressed as a linear combination of the eigenfunctions. For the solutions defined in finite regions, only discrete spectra of eigenvalues are involved. The discrete spectra become a continuum for the solutions defined in infinite regions. The harmonic functions should be properly selected according to the physical properties that the solutions must have. Note that the exponential functions in (1.43) represent a travelling wave while the sine and cosine functions represent a standing wave. Also note that the separation constants are a complex number for the waves propagating in a lossy medium.

1.2.1.2 Cylindrical Coordinate System

In a cylindrical coordinate system, (1.37) can be written as

$$\frac{1}{\rho}\frac{\partial}{\partial \rho}\left(\rho \frac{\partial u}{\partial \rho}\right) + \frac{1}{\rho^2}\frac{\partial^2 u}{\partial \varphi^2} + \frac{\partial^2 u}{\partial z^2} + k^2 u = 0. \tag{1.45}$$

By the method of separation of variables, the solutions may be assumed to be of the form

$$u = R(\rho)\Phi(\varphi)Z(z). \tag{1.46}$$

Introducing (1.46) into (1.45) yields

$$\begin{aligned} &\frac{d^2 R}{d\rho^2} + \frac{1}{\rho}\frac{dR}{d\rho} + \left(\mu^2 - \frac{p^2}{\rho^2}\right)R = 0, \\ &\frac{d^2 \Phi}{d\varphi^2} + p^2 \Phi = 0, \\ &\frac{d^2 Z}{dz^2} + \beta^2 Z = 0, \end{aligned} \tag{1.47}$$

where μ, p, and β are separation constants and satisfy

$$\beta^2 + \mu^2 = k^2. \tag{1.48}$$

The first equation of (1.47) is **Bessel equation**, named after the German mathematician Friedrich Wilhelm Bessel (1784–1846), whose solutions are **Bessel functions**:

$$J_p(\mu\rho), N_p(\mu\rho), H_p^{(1)}(\mu\rho), H_p^{(2)}(\mu\rho),$$

where $J_p(\mu\rho)$ and $N_p(\mu\rho)$ are the Bessel functions of the first and second kind, $H_p^{(1)}(\mu\rho)$ and $H_p^{(2)}(\mu\rho)$ are the Bessel functions of the third and fourth kind, also called **Hankel functions** of first and second kind, respectively, named after German mathematician Hermann Hankel (1839–1873). The Bessel function of the first kind is defined by

$$J_p(\mu z) = \sum_{m=0}^{\infty} \frac{(-1)^m}{\Gamma(m+1)\Gamma(p+m+1)} \left(\frac{\mu z}{2}\right)^{p+2m}, \tag{1.49}$$

where $\Gamma(\alpha)$ is the gamma function defined by

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx, \alpha > 0.$$

If p is not an integer, a second independent solution is $J_{-p}(\mu z)$. If $p = n$ is an integer, $J_{-n}(\mu z)$ is related to $J_n(\mu z)$ by $J_{-n}(z) = (-1)^n J_n(z)$. The Bessel function of the second kind is defined by

$$N_p(\mu z) = \frac{\cos p\pi J_p(\mu z) - J_{-p}(\mu z)}{\sin p\pi}, \tag{1.50}$$

and the Bessel functions of the third (Hankel function of the first kind) and fourth kind (Hankel function of the second kind) are defined by

$$\begin{aligned} H_p^{(1)}(\mu z) &= J_p(\mu z) + jN_p(\mu z), \\ H_p^{(2)}(\mu z) &= J_p(\mu z) - jN_p(\mu z). \end{aligned} \tag{1.51}$$

The solutions of second and third equation of (1.47) are harmonic functions. Note that only $J_p(\mu\rho)$ is finite at $\rho = 0$. The separation constants μ and p are determined by the boundary conditions. For example, if the field u is finite and satisfies homogeneous Dirichlet boundary condition $u = 0$ at $\rho = a$, the separation constant μ is determined by $J_p(\mu\rho) = 0$. If the cylindrical region contains all φ from 0 to 2π, the separation constant p is usually determined by the requirement that the field is single-valued, i.e. $\Phi(0) = \Phi(2\pi)$. In this case, p must be integers. If the cylindrical region only contains a circular sector, p will be fractional numbers.

Let $R_p(\mu z) = AJ_p(\mu z) + BN_p(\mu z)$, where A and B are constants. Some typical recurrence relations for the linear combination of the Bessel functions are listed below:

$$\frac{2p}{\mu z} R_p(\mu z) = R_{p-1}(\mu z) + R_{p+1}(\mu z),$$

$$\frac{1}{\mu}\frac{d}{dz} R_p(\mu z) = \frac{1}{2}\left[R_{p-1}(\mu z) - R_{p+1}(\mu z)\right],$$

$$z\frac{d}{dz} R_p(\mu z) = pR_p(\mu z) - \mu z R_{p+1}(\mu z),$$

$$\frac{d}{dz}\left[z^p R_p(\mu z)\right] = \mu z^p R_{p-1}(\mu z),$$

$$\frac{d}{dz}\left[z^{-p} R_p(\mu z)\right] = -\mu z^{-p} R_{p+1}(\mu z).$$

The Bessel functions have the orthogonality property

$$\int_0^1 x J_p\left(\chi_{pm} x\right) J_p\left(\chi_{pn} x\right) dx = \begin{cases} 0, & m \neq n \\ \frac{1}{2}\left[J_p'\left(\chi_{pm}\right)\right]^2, & m = n \end{cases}, \tag{1.52}$$

where χ_{pn} denotes the nth positive zero of the Bessel function J_p, i.e. $J_p(\chi_{pn}) = 0$, and the prime denotes the derivative of the Bessel function with respect to its argument. Let $C[a, b]$ denote the set of continuous functions defined on the closed interval $[a, b]$. An arbitrary continuous function in $C[0, 1]$ can be expanded in terms of the Bessel functions.

Theorem 1.1 (Fourier–Bessel Expansion)

If the function $f(x) \in C[0, 1]$ and the integral $\int_0^1 \sqrt{t} f(t) dt$ exist, it has the expansion

$$f(x) = \sum_n a_n J_p\left(\chi_{pn} x\right), \quad p > -1, \tag{1.53}$$

where the expansion coefficients can be determined from (1.52):

$$a_n = \frac{2}{\left[J_p'\left(\chi_{pn}\right)\right]^2} \int_0^1 x f(x) J_p\left(\chi_{pn} x\right) dx = \frac{2}{\left[J_{p+1}\left(\chi_{pn}\right)\right]^2} \int_0^1 x f(x) J_p\left(\chi_{pn} x\right) dx.$$

□

1.2.1.3 Spherical Coordinate System

In spherical coordinate system, (1.37) can be expressed as

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial u}{\partial r}\right)+\frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial u}{\partial\theta}\right)+\frac{1}{r^2\sin\theta}\frac{\partial^2 u}{\partial\varphi^2}+k^2u=0. \tag{1.54}$$

By means of the separation of variables, one may assume

$$u=R(r)\Theta(\theta)\Phi(\varphi). \tag{1.55}$$

Substitution of (1.55) into (1.54) leads to

$$\begin{aligned}
&\frac{1}{R}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right)+k^2r^2=\beta^2,\\
&\frac{1}{\Theta\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{d\Theta}{d\theta}\right)-\frac{m^2}{\sin^2\theta}=-\beta^2,\\
&\frac{d^2\Phi}{d\varphi^2}+m^2\Phi=0,
\end{aligned} \tag{1.56}$$

where β and m are separation constants. Let $x=\cos\theta$ and $P(x)=\Theta(\theta)$. The second equation of (1.56) becomes

$$\left(1-x^2\right)\frac{d^2P}{dx^2}-2x\frac{dP}{dx}+\left(\beta^2-\frac{m^2}{1-x^2}\right)P=0. \tag{1.57}$$

This is called **Legendre equation**, named after the French mathematician Adrien-Marie Legendre (1752–1833). The points $x=\pm 1$ are singular. Equation (1.57) has two linearly independent solutions and can be expressed as a power series at $x=0$. In general, the series solution diverges at $x=\pm 1$. But if one lets $\beta^2=n(n+1)$, $n=0,1,2\ldots$, the series will be finite at $x=\pm 1$ and have finite terms. Thus, the separation constant β is determined naturally and (1.56) can be rewritten as

$$\begin{aligned}
&\frac{d}{dr}\left(r^2\frac{dR}{dr}\right)+\left[k^2r^2-n(n+1)\right]R=0,\\
&\left(1-x^2\right)\frac{d^2P}{dx^2}-2x\frac{dP}{dx}+\left[n(n+1)-\frac{m^2}{1-x^2}\right]P=0,\\
&\frac{d^2\Phi}{d\varphi^2}+m^2\Phi=0.
\end{aligned} \tag{1.58}$$

The solutions of the first equation of (1.58) are **spherical Bessel functions** of the first and second kinds, **spherical Hankel functions** of the first and second kinds, respectively, defined by

$$\begin{aligned}
j_n(kr) &= \sqrt{\frac{\pi}{2kr}} J_{n+1/2}(kr),\\
n_n(kr) &= \sqrt{\frac{\pi}{2kr}} N_{n+1/2}(kr),\\
h_n^{(1)}(kr) &= \sqrt{\frac{\pi}{2kr}} H_{n+1/2}^{(1)}(kr),\\
h_n^{(2)}(kr) &= \sqrt{\frac{\pi}{2kr}} H_{n+1/2}^{(2)}(kr).
\end{aligned} \tag{1.59}$$

Let $z_n(kr) = Aj_n(kr) + Bn_n(kr)$, where A and B are constants. The recurrence relations for the linear combination of the spherical Bessel functions are summarized as follows:

$$\begin{aligned}
&\frac{2n+1}{kr} z_n(kr) = z_{n-1}(kr) + z_{n+1}(kr),\\
&\frac{2n+1}{k}\frac{d}{dr} z_n(kr) = nz_{n-1}(kr) - (n+1)z_{n+1}(kr),\\
&\frac{d}{dr}\left[r^{n+1} z_n(kr)\right] = kr^{n+1} z_{n-1}(kr),\\
&\frac{d}{dr}\left[r^{-n} z_n(kr)\right] = -kr^{-n} z_{n+1}(kr).
\end{aligned}$$

The solutions of the second equation of (1.58) are **associated Legendre functions** of first and second kinds defined, respectively, by

$$P_n^m(x) = \frac{(1-x^2)^{m/2}}{2^n n!}\frac{d^{m+n}}{dx^{m+n}}\left(x^2-1\right)^n, m \le n, \tag{1.60}$$

and

$$Q_n^m(x) = \left(1-x^2\right)^{\frac{m}{2}}\frac{d^m}{dx^m} Q_n(x), m \le n, \tag{1.61}$$

where

$$Q_n(x) = \frac{1}{2}P_n^0(x)\ln\frac{1+x}{1-x} - \sum_{r=1}^{n}\frac{1}{r}P_{r-1}^0(x)P_{n-r}^0(x)$$

is the **Legendre function of the second kind**. The following integrations on orthogonality are useful:

1. $$\int_{-1}^{1} \frac{P_n^m(x)P_n^k(x)}{1-x^2}dx = \frac{1}{m}\frac{(n+m)!}{(n-m)!}\delta_{mk}.$$

2. $$\int_{-1}^{1} P_l^m(x)P_n^m(x)dx = \frac{2}{2l+1}\frac{(l+m)!}{(l-m)!}\delta_{ln}.$$

3. $$\int_{0}^{\pi} \left[\frac{dP_n^m(\cos\theta)}{d\theta}\frac{dP_l^m(\cos\theta)}{d\theta} + \frac{m^2}{\sin^2\theta}P_n^m(\cos\theta)P_l^m(\cos\theta)\right]\sin\theta d\theta$$
$$= \frac{2}{2n+1}\frac{(n+m)!}{(n-m)!}n(n+1)\delta_{nl}.$$

In the above, $\delta_{mn} = \begin{cases} 1, m = n \\ 0, m \neq n \end{cases}$. The solutions of the third equation of (1.58) are harmonic functions. Note that the separation constants are not related in the spherical coordinate system.

Theorem 1.2 (Completeness of Associated Legendre Functions)

Any function $f(x) \in C[-1, 1]$ satisfying the boundary conditions $f(-1) = f(1) = 0$ has the expansion

$$f(x) = \sum_{m \le n} a_n P_n^m(x),$$

where

$$a_n = \frac{2n+1}{2}\frac{(n-m)!}{(n+m)!}\int_{-1}^{1} f(x)P_n^m(x)dx.$$

□

1.2.2 Method of Green's Function

A source may be divided into a number of elementary sources. The superposition theorem indicates that the field generated by the source can be expressed as the sum of the fields generated by the elementary sources. The concept of the Green's function was first developed by the British mathematician George Green in the 1820s. Physically, Green's function represents the field produced by a point source (an elementary source). The well-known impulse response of a linear system is a typical Green's function with specified initial conditions. Green's function provides a general method to solve PDEs, by means of which the solution of a PDE can be represented by an integral defined over the source region or on a closed surface enclosing the source. Consider a PDE defined in a source region V contained in R^3 (the three-dimensional (3D) Euclidian space)

$$\hat{L}u(\mathbf{r}) = f(\mathbf{r}), \tag{1.62}$$

where $\hat{L}$ is a differential operator, $\mathbf{r} = (x, y, z) \in R^3$, and f is a known source function. The solution of the above equation can be expressed by

$$u(\mathbf{r}) = \hat{L}^{-1} f(\mathbf{r}),$$

where $\hat{L}^{-1}$ stands for the inverse of $\hat{L}$ and is often represented by an integral operator with a kernel function $G(\mathbf{r}, \mathbf{r}')$:

$$\hat{L}^{-1} f(\mathbf{r}) = -\int_V G(\mathbf{r}, \mathbf{r}') f(\mathbf{r}') dV(\mathbf{r}'). \tag{1.63}$$

If $\hat{L}$ is applied to both sides of the above equation and use is made of $\hat{L}\hat{L}^{-1} = \hat{I}$, where $\hat{I}$ is the **identity operator**, one obtains

$$\hat{L}\hat{L}^{-1} f(\mathbf{r}) = f(\mathbf{r}) = -\int_V \hat{L} G(\mathbf{r}, \mathbf{r}') f(\mathbf{r}') dV(\mathbf{r}').$$

This implies that the kernel function G satisfies

$$\hat{L} G(\mathbf{r}, \mathbf{r}') f(\mathbf{r}) = -\delta(\mathbf{r} - \mathbf{r}'), \tag{1.64}$$

where $\delta(\mathbf{r})$ is the 3D **Dirac delta function** defined symbolically by

$$\int_{R^3} \delta(\mathbf{r} - \mathbf{r}') \phi(\mathbf{r}') dV(\mathbf{r}') = \phi(\mathbf{r}) \tag{1.65}$$

with $\phi(\mathbf{r})$ being an arbitrary smooth function defined in R^3. The kernel function G is called the **fundamental solution** or **Green's function** of Eq. (1.62).

1.2.2.1 Green's Functions for Helmholtz Equation

Let $\boldsymbol{\rho} = (x, y)$ denote a point in the two-dimensional (2D) Euclidian space R^2 and $d\Omega = dxdy$ the differential area element. The fundamental solutions of Laplace and Helmholtz equations are summarized in Table 1.2. The 2D Dirac delta function $\delta(\boldsymbol{\rho})$ is defined by

$$\int_{R^2} \delta(\boldsymbol{\rho} - \boldsymbol{\rho}') \phi(\boldsymbol{\rho}') d\Omega(\boldsymbol{\rho}') = \phi(\boldsymbol{\rho}),$$

Table 1.2 Green's functions.

Equations	Green's functions
2D Laplace equation: $\nabla^2 G(\boldsymbol{\rho}, \boldsymbol{\rho}') = -\delta(\boldsymbol{\rho} - \boldsymbol{\rho}')$	$G(\boldsymbol{\rho}, \boldsymbol{\rho}') = -\frac{1}{2\pi} \ln\lvert\boldsymbol{\rho} - \boldsymbol{\rho}'\rvert$
3D Laplace equation: $\nabla^2 G(\mathbf{r}, \mathbf{r}') = -\delta(\mathbf{r} - \mathbf{r}')$	$G(\mathbf{r}, \mathbf{r}') = \frac{1}{4\pi\lvert\mathbf{r} - \mathbf{r}'\rvert}$
2D Helmholtz equation: $(\nabla^2 + k^2)G(\boldsymbol{\rho}, \boldsymbol{\rho}') = -\delta(\boldsymbol{\rho} - \boldsymbol{\rho}')$	$G(\boldsymbol{\rho}, \boldsymbol{\rho}') = \frac{1}{4j} H_0^{(2)}(k\lvert\boldsymbol{\rho} - \boldsymbol{\rho}'\rvert)$
3D Helmholtz equation: $(\nabla^2 + k^2)G(\mathbf{r}, \mathbf{r}') = -\delta(\mathbf{r} - \mathbf{r}')$	$G(\mathbf{r}, \mathbf{r}') = \frac{e^{-jk\lvert\mathbf{r} - \mathbf{r}'\rvert}}{4\pi\lvert\mathbf{r} - \mathbf{r}'\rvert}$

where $\phi(\mathbf{r})$ is an arbitrary smooth function defined in R^2. It can be seen that the Green's functions are symmetric $G(\mathbf{r}, \mathbf{r}') = G(\mathbf{r}', \mathbf{r})$.

Example 1.1 The Green's function for one-dimensional Helmholtz equation satisfies

$$\begin{cases} \dfrac{d^2 G(z, z')}{dz^2} + k^2 G(z, z') = -\delta(z - z'), \\ \lim\limits_{z \to \pm\infty} \left(\dfrac{dG}{dz} \pm jkG \right) = 0, \end{cases} \tag{1.66}$$

where the one-dimensional Dirac delta function $\delta(z)$ is defined by

$$\int_{-\infty}^{\infty} \delta(z - z')\phi(z')dz' = \phi(z)$$

and $\phi(z)$ is an arbitrary smooth function defined in the real axis R. The second equation in (1.66) denotes the radiation condition at infinity. To solve (1.66), one may let

$$G(z, z') = \begin{cases} G_1(z, z'), z < z' \\ G_2(z, z'), z > z' \end{cases}.$$

The functions G_1 and G_2 satisfy the homogeneous Helmholtz equation in the regions $z < z'$ and $z > z'$, respectively, and can be written as

$$G_1(z, z') = a_1 e^{-jk(z - z')} + b_1 e^{jk(z - z')},$$
$$G_2(z, z') = a_2 e^{-jk(z - z')} + b_2 e^{jk(z - z')},$$

where a_1, b_1, a_2, b_2 are constants to be determined. On account of the radiation conditions at $z = \pm\infty$, one may easily find $a_1 = b_2 = 0$. Thus,

$$\begin{aligned} G_1(z,z') &= b_1 e^{jk(z-z')}, \\ G_2(z,z') &= a_2 e^{-jk(z-z')}. \end{aligned} \tag{1.67}$$

The first equation of (1.66) implies that G must be continuous while its first derivative has a jump discontinuity at the source point $z = z'$:

$$\begin{aligned} &G_1(z,z') = G_2(z,z'), \\ &\frac{dG_2(z,z')}{dz} - \frac{dG_1(z,z')}{dz} = -1. \end{aligned}$$

Introducing (1.67) into the above equations yields

$$a_2 = b_1 = \frac{1}{j2k}.$$

The Green's function for one-dimensional Helmholtz equation is thus given by

$$G(z,z') = \begin{cases} \dfrac{1}{j2k} e^{jk(z-z')}, z < z' \\ \dfrac{1}{j2k} e^{-jk(z-z')}, z > z' \end{cases} = \frac{1}{j2k} e^{-jk|z-z'|}. \tag{1.68}$$

□

1.2.2.2 Dyadic Green's Functions and Integral Representations

The Green's function often appears as the kernel of an integral operator. For a scalar field, the kernel is also a scalar; but for a vector field, the kernel must be a dyadic. A dyadic is a second-order tensor formed by putting two vectors side by side. Its manipulation rules are analogous to that for matrix algebra (see Appendix B). Dyadic notation was first established by Josiah Willard Gibbs (1839–1903) in 1884. The application of dyadic Green's function in solving EM boundary value problem can be traced back to Schwinger's work in the early 1940s. Levine and Schwinger applied the dyadic Green's function to investigate the diffraction problem by an aperture in an infinite plane conducting screen [23]. In 1953, Morse and Feshbach discussed various applications of dyadic Green's functions [20]. A systematic study of dyadic Green's functions and their applications in EM engineering can be found in [24]. One of the advantages of using dyadic Green's functions is that it affords a compact formulation or solution for the field problems. Consider an electric current element in the direction of $\alpha(\alpha = x, y, z)$ located at $\mathbf{r}'$:

$$\mathbf{J}^{(\alpha)}(\mathbf{r}) = -\frac{1}{j\omega\mu}\delta(\mathbf{r}-\mathbf{r}')\mathbf{u}_\alpha,$$

which produces EM fields $\mathbf{E}^{(\alpha)}(\mathbf{r})$ and $\mathbf{H}^{(\alpha)}(\mathbf{r})$ at $\mathbf{r}$. Let

$$\begin{aligned}\mathbf{G}_e^{(\alpha)}(\mathbf{r},\mathbf{r}') &= \mathbf{E}^{(\alpha)}(\mathbf{r}),\\ \mathbf{G}_m^{(\alpha)}(\mathbf{r},\mathbf{r}') &= -j\omega\mu\mathbf{H}^{(\alpha)}(\mathbf{r}).\end{aligned} \tag{1.69}$$

Here $\mathbf{G}_e^{(\alpha)}(\mathbf{r},\mathbf{r}')$ and $\mathbf{G}_m^{(\alpha)}(\mathbf{r},\mathbf{r}')$ are, respectively, referred to as **electric and magnetic Green's functions** along direction α in free space. It follows from Maxwell equations that

$$\begin{aligned}\nabla\times\mathbf{G}_e^{(\alpha)}(\mathbf{r},\mathbf{r}') &= \mathbf{G}_m^{(\alpha)}(\mathbf{r},\mathbf{r}'),\\ \nabla\times\mathbf{G}_m^{(\alpha)}(\mathbf{r},\mathbf{r}') &= \mathbf{u}_\alpha\delta(\mathbf{r}-\mathbf{r}') + k^2\mathbf{G}_e^{(\alpha)}(\mathbf{r},\mathbf{r}').\end{aligned}$$

The dyadic functions defined by

$$\begin{aligned}\overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}') &= \sum_{\alpha=x,y,z}\mathbf{G}_e^{(\alpha)}(\mathbf{r},\mathbf{r}')\mathbf{u}_\alpha,\\ \overleftrightarrow{\mathbf{G}}_m(\mathbf{r},\mathbf{r}') &= \sum_{\alpha=x,y,z}\mathbf{G}_m^{(\alpha)}(\mathbf{r},\mathbf{r}')\mathbf{u}_\alpha,\end{aligned}$$

are, respectively, called **electric and magnetic dyadic Green's functions** in free space. Apparently,

$$\begin{aligned}\nabla\times\overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}') &= \overleftrightarrow{\mathbf{G}}_m(\mathbf{r},\mathbf{r}'),\\ \nabla\times\overleftrightarrow{\mathbf{G}}_m(\mathbf{r},\mathbf{r}') &= \overleftrightarrow{\mathbf{I}}\delta(\mathbf{r}-\mathbf{r}') + k^2\overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}'),\end{aligned} \tag{1.70}$$

where $\overleftrightarrow{\mathbf{I}}$ is the identity dyadic. Note that $\overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}')$ is a symmetric while $\overleftrightarrow{\mathbf{G}}_m(\mathbf{r},\mathbf{r}')$ is antisymmetric upon interchange of $\mathbf{r}$ and $\mathbf{r}'$:

$$\begin{aligned}\overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}') &= \overleftrightarrow{\mathbf{G}}_e(\mathbf{r}',\mathbf{r}),\\ \overleftrightarrow{\mathbf{G}}_m(\mathbf{r},\mathbf{r}') &= -\overleftrightarrow{\mathbf{G}}_m(\mathbf{r}',\mathbf{r}).\end{aligned} \tag{1.71}$$

It follows from (1.70) that

$$\begin{aligned}\nabla\times\nabla\times\overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}') - k^2\overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}') &= \overleftrightarrow{\mathbf{I}}\delta(\mathbf{r}-\mathbf{r}'),\\ \nabla\times\nabla\times\overleftrightarrow{\mathbf{G}}_m(\mathbf{r},\mathbf{r}') - k^2\overleftrightarrow{\mathbf{G}}_m(\mathbf{r},\mathbf{r}') &= \nabla\times\left[\overleftrightarrow{\mathbf{I}}\delta(\mathbf{r}-\mathbf{r}')\right].\end{aligned} \tag{1.72}$$

The free space electric dyadic Green's function $\overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}')$ may be represented by free space Green's function $G(\mathbf{r},\mathbf{r}')$:

$$\overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}') = \left(\overleftrightarrow{\mathbf{I}} + \frac{1}{k^2}\nabla\nabla\right)G(\mathbf{r},\mathbf{r}'). \tag{1.73}$$

In fact, the first equation of (1.72) may be written as

$$-\nabla^2\overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}') - k^2\overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}') + \nabla\nabla\cdot\overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}') = \overleftrightarrow{\mathbf{I}}\delta(\mathbf{r}-\mathbf{r}'). \tag{1.74}$$

Taking the divergence of the first equation of (1.72) yields

$$\nabla\cdot\overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}') = -\frac{1}{k^2}\nabla\cdot\left[\overleftrightarrow{\mathbf{I}}\delta(\mathbf{r}-\mathbf{r}')\right] = -\frac{1}{k^2}\nabla\delta(\mathbf{r}-\mathbf{r}').$$

Insertion of the above into (1.74) gives

$$\nabla^2\overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}') + k^2\overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}') = -\left(\overleftrightarrow{\mathbf{I}} + \frac{1}{k^2}\nabla\nabla\right)\delta(\mathbf{r}-\mathbf{r}').$$

Obviously, expression (1.73) for the free space electric dyadic Green's function satisfies the above equation. The free space magnetic dyadic Green's function may be expressed as

$$\overleftrightarrow{\mathbf{G}}_m(\mathbf{r},\mathbf{r}') = \nabla\times\overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}') = \nabla\times\left[G(\mathbf{r},\mathbf{r}')\overleftrightarrow{\mathbf{I}}\right] = \nabla\times\left[\overleftrightarrow{\mathbf{G}}_0(\mathbf{r},\mathbf{r}')\right],$$

where $\overleftrightarrow{\mathbf{G}}_0 = G\overleftrightarrow{\mathbf{I}}$ satisfies Helmholtz equation

$$\left(\nabla^2 + k^2\right)\overleftrightarrow{\mathbf{G}}_0(\mathbf{r},\mathbf{r}') = -\overleftrightarrow{\mathbf{I}}\delta(\mathbf{r}-\mathbf{r}'). \tag{1.75}$$

Consider the scattering problem of an incident field by an obstacle bounded by S as illustrated in Figure 1.2. Let V be the region bounded by $S + S_\infty$, where S_∞ is a surface enclosing the source and the obstacle. By letting $\mathbf{P} = \mathbf{E}$, $\mathbf{Q} = \mathbf{G}_e^{(\alpha)}$ in the vector Green's identity

$$\int_V (\mathbf{P}\cdot\nabla\times\nabla\times\mathbf{Q} - \mathbf{Q}\cdot\nabla\times\nabla\times\mathbf{P})dV = -\int_S (\mathbf{u}_n\times\mathbf{Q}\cdot\nabla\times\mathbf{P} - \mathbf{u}_n\times\mathbf{P}\cdot\nabla\times\mathbf{Q})dS$$
$$+ \int_{S_\infty} (\mathbf{u}_n\times\mathbf{Q}\cdot\nabla\times\mathbf{P} - \mathbf{u}_n\times\mathbf{P}\cdot\nabla\times\mathbf{Q})dS,$$

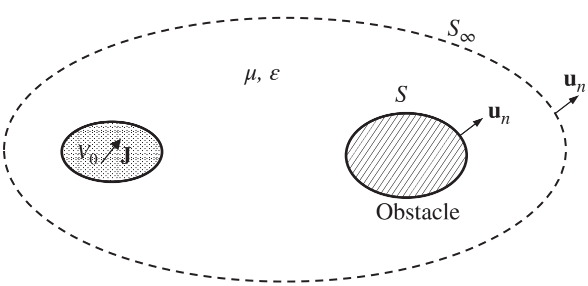

Figure 1.2 Scattering by an obstacle.

and by making use of the first Eq. (1.72) and the wave equation for the electric field

$$\nabla \times \nabla \times \mathbf{E}(\mathbf{r}) - k^2\mathbf{E}(\mathbf{r}) = -j\omega\mu\mathbf{J}(\mathbf{r}) - \nabla \times \mathbf{J}_m(\mathbf{r}), \tag{1.76}$$

one may find the integral expression for the electric field

$$\begin{aligned}\mathbf{E}(\mathbf{r}) &= -j\omega\mu \int_{V_0} \overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}') \cdot \mathbf{J}(\mathbf{r}')dV(\mathbf{r}') - \int_{V_0} \overleftrightarrow{\mathbf{G}}_m(\mathbf{r},\mathbf{r}') \cdot \mathbf{J}_m(\mathbf{r}')dV(\mathbf{r}') \\ &\quad -j\omega\mu \int_S \overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}') \cdot \mathbf{J}_s(\mathbf{r}')dS(\mathbf{r}') - \int_S \overleftrightarrow{\mathbf{G}}_m(\mathbf{r},\mathbf{r}') \cdot \mathbf{J}_{ms}(\mathbf{r}')dS(\mathbf{r}'),\end{aligned} \tag{1.77}$$

where $\mathbf{J}_s = \mathbf{u}_n \times \mathbf{H}$ and $\mathbf{J}_{ms} = -\mathbf{u}_n \times \mathbf{E}$ are, respectively, the equivalent surface electric and magnetic currents. Similarly, one may find the integral expression for the magnetic field

$$\begin{aligned}\mathbf{H}(\mathbf{r}) &= -j\omega\varepsilon \int_{V_0} \overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}') \cdot \mathbf{J}_m(\mathbf{r}')dV(\mathbf{r}') + \int_{V_0} \overleftrightarrow{\mathbf{G}}_m(\mathbf{r},\mathbf{r}') \cdot \mathbf{J}(\mathbf{r}')dV(\mathbf{r}') \\ &\quad -j\omega\varepsilon \int_S \overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}') \cdot \mathbf{J}_{ms}(\mathbf{r}')dS(\mathbf{r}') + \int_S \overleftrightarrow{\mathbf{G}}_m(\mathbf{r},\mathbf{r}') \cdot \mathbf{J}_s(\mathbf{r}')dS(\mathbf{r}').\end{aligned} \tag{1.78}$$

In the derivation of (1.77) and (1.78), the surface integral over S_∞ has been ignored since it approaches to zero due to the radiation condition as S_∞ expands to infinity. The volume integrals in (1.77) and (1.78) represent the incident fields while the surface integrals denote the scattered fields produced by the induced currents on the obstacle.

Remark 1.2 Care must be exercised when the observation point $\mathbf{r}$ is inside the source region V_0. The dyadic Green's functions in (1.77) and (1.78) become infinite when $\mathbf{r}$ approaches $\mathbf{r}'$. This singular behavior is often treated with the principal volume method by introducing an exclusion volume around $\mathbf{r}$ [25–27]. □

Remark 1.3 For many applications discussed in this book, the dyadic Green's functions may be expanded in terms of eigenfunctions of the Helmholtz equation for the vector fields. In these cases, the integral representations (1.77) and (1.78) become an infinite series. □

1.2.3 Variational Method

Variational method or **calculus of variations** is a generalization of calculus and can be traced back to the early 1730s when Swiss mathematician Leonhard Euler

(1707–1783) and French mathematician Joseph-Louis Lagrange (1736–1813) elaborated the subject. Instead of finding the extrema of a function in calculus, the calculus of variations deals with maximizing or minimizing functionals, which are often expressed as definite integrals involving functions and their derivatives. The extremal functions that make the functional stationary (i.e. attain a maximum or minimum value) can be obtained by assuming that the rate of change of the functional is zero. Let F be a map from a linear space consisting of functions into the real axis. Such a map is called a **functional**. Let v be an arbitrary function in the space. The **gradient** or **functional derivative** of F at u, denoted by $\delta F(u)/\delta u$ and used to describe the rate of change of the functional, is defined by [16]

$$\left(\frac{\delta F(u)}{\delta u},v\right)=\frac{d}{d\varepsilon}F(u+\varepsilon v)\bigg|_{\varepsilon=0}, \tag{1.79}$$

where $(\cdot,\cdot)$ is an **inner product** defined by the following rules:

1) Positive definiteness: $(u, u)\geq 0$ and $(u, u)=0$ if and only if $u=0$.
2) Hermitian property: $(u,v)=\overline{(v,u)}$.
3) Homogeneity: $(\alpha u, v)=\alpha(u, v)$.
4) Additivity: $(u+v, w)=(u, w)+(v, w)$.

Here, u, v, z are functions and α is a number. A linear space equipped with an inner product is called an **inner product space**.

Extremum Theorem: A necessary condition for a functional F to have an extremum at u is that its functional derivative vanishes at u:

$$\frac{\delta F(u)}{\delta u}=0. \tag{1.80}$$

This is referred to as **Euler–Lagrangian equation**. □

Example 1.2 An **operator** (or **transformation**) $\hat{L}$ on an inner product space is defined as a map from the inner product space to itself. An operator $\hat{L}$ is called **self-adjoint** if it satisfies

$$\left(\hat{L}u,v\right)=\left(u,\hat{L}v\right).$$

The most important functional to be dealt with in this book is the **Rayleigh quotient**, named after the English scientist Lord Rayleigh (1942–1919), defined by

$$\lambda(u)=\frac{\left(\hat{L}u,u\right)}{(u,u)}. \tag{1.81}$$

Since $\hat{L}$ is self-adjoint, λ can be shown to be real. For an arbitrary v, the functional derivative of the Rayleigh quotient can be determined from (1.79):

$$\left(\frac{\delta\lambda}{\delta u}, v\right) = \frac{d}{d\varepsilon}\lambda(u+\varepsilon v)\bigg|_{\varepsilon=0} = \frac{d}{dt}\frac{\left(\hat{L}(u+\varepsilon v), u+\varepsilon v\right)}{(u+\varepsilon v, u+\varepsilon v)}\Bigg|_{\varepsilon=0}$$
$$= \frac{1}{(u,u)} 2\mathrm{Re}\left(\hat{L}u - \lambda u, v\right).$$

If $\delta\lambda/\delta u = 0$, one obtains the eigenvalue equation

$$\hat{L}u = \lambda u. \tag{1.82}$$

This is the Lagrangian equation for the Rayleigh quotient (1.81). □

1.3 Eigenvalue Problem for Hermitian Matrix

When the dimension of the underlying vector space of a linear transformation is finite, the linear transformation is reduced to a matrix under a fixed basis. A complex square matrix $[A]$ is called **Hermitian**, named after the French mathematician Charles Hermite (1822–1901), if $[A] = [A]^H$, where the superscript H denotes the conjugate transpose of a matrix. Hermitian matrices are the complex extension of real symmetric matrices and are fundamental to mathematics, physics, and engineering sciences.

1.3.1 Properties

Let $[A]$ denote an $N \times N$ Hermitian matrix. If an $N \times 1$ matrix $[x]$ (also called a **column vector** or **vector**) and a scalar λ satisfy the equation

$$[A][x] = \lambda[x], \tag{1.83}$$

then λ is called an **eigenvalue** of $[A]$, $[x]$ is called an **eigenvector** corresponding to the eigenvalue λ, and $(\lambda, [x])$ is called an **eigenpair** of $[A]$. The eigenvalue problem (1.83) for a Hermitian matrix has the following properties, which can be found in a standard text book on matrix [e.g. 28]:

1) All the eigenvalues are real and they can be ordered such that

   $$\lambda_1 \le \lambda_2 \le \cdots \le \lambda_N,$$

   and the corresponding eigenvectors are denoted by $[x_1]$, $[x_2]$, ..., $[x_N]$.
2) The eigenfunctions corresponding to different eigenvalues are orthogonal.

3) Assume that the eigenvectors $[x_1], [x_2], \ldots, [x_N]$ are orthonormal. The matrix $[A]$ has the **spectral decomposition**

$$[A] = [U][\Lambda][U]^H = \sum_{j=1}^{N} \lambda_j [x_j] [x_j]^H,$$

where

$$[U] = [[x_1], [x_2], \ldots, [x_N]],$$
$$[\Lambda] = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_N).$$

The matrix $[U]$ is unitary $[U]^H[U] = [I]$, where $[I]$ denotes identity matrix.

The following statements are equivalent.

1) The Hermitian matrix $[A]$ is **positive**, i.e. $[x]^H[A][x] > 0$ for all $[x] \neq 0$.
2) All eigenvalues of $[A]$ are positive.
3) There exists a nonsingular $n \times n$ matrix $[R]$, such that $[A] = [R]^H[R]$.
4) There exists a nonsingular $n \times n$ matrix $[R]$, such that $[R]^H[A][R]$ is positive.

A positive matrix $[A]$ (not necessarily Hermitian) has the following properties:

1) The diagonal elements of $[A]$ are all positive.
2) $\det[A] > 0$.
3) $[A]^{-1}$ is positive.
4) The element of the largest absolute value is a diagonal element.

1.3.2 Rayleigh Quotient

The **Rayleigh quotient** for an $n \times n$ Hermitian matrix $[A]$ is defined by

$$\lambda = \frac{[x]^H[A][x]}{[x]^H[x]}. \tag{1.84}$$

Any vector $[x]$ can be expressed as a linear combination of the eigenvectors

$$[x] = \sum_{j=1}^{N} c_j [x_j] = [U][c_j], \tag{1.85}$$

where $[c_j] = [c_1, c_2, \ldots, c_N]^T$. Assume that the eigenvectors $[x_1], [x_2], \ldots, [x_N]$ of $[A]$ are orthonormal. Introducing the above expansion into (1.84), one may find

$$\begin{aligned}\lambda &= \frac{[x]^H[A][x]}{[x]^H[x]} = \frac{[c_j]^H[U]^H[A][U][c_j]}{[c_j]^H[U]^H[U][c_j]} = \frac{[c_j]^H[\Lambda][c_j]}{[c_j]^H[c_j]} \\ &= \frac{\lambda_1|c_1|^2 + \lambda_2|c_2|^2 + \cdots + \lambda_n|c_N|^2}{|c_1|^2 + |c_2|^2 + \cdots + |c_N|^2},\end{aligned} \tag{1.86}$$

which implies

$$\lambda_1 \le \frac{[x]^H[A][x]}{[x]^H[x]} \le \lambda_N. \tag{1.87}$$

As a result, the following optimization problems are established:

$$\begin{aligned}\lambda_1 &= \min \frac{[x]^H[A][x]}{[x]^H[x]}, \\ \lambda_N &= \max \frac{[x]^H[A][x]}{[x]^H[x]},\end{aligned} \tag{1.88}$$

where $[x] \in C^N$ is called **trial vector** and C^N stands for the set consisting of all complex column vectors of dimension N. To determine other eigenvalues, the trial vector $[x]$ will be assumed to be orthogonal to $[x_1]$. Thus, one finds $c_1 = 0$ in the expansion (1.85), and (1.86) reduces to

$$\lambda = \frac{\lambda_2|c_2|^2 + \cdots + \lambda_N|c_N|^2}{|c_2|^2 + \cdots + |c_N|^2} \ge \lambda_2. \tag{1.89}$$

The equality holds for $[x] = [x_2]$. The above relation suggests

$$\lambda_2 = \min_{[x]\perp[x_1]} \frac{[x]^H[A][x]}{[x]^H[x]}. \tag{1.90}$$

The second eigenvalue λ_2 can thus be found by minimizing (1.84) with the supplementary constraint that the trial vector $[x]$ must be orthogonal to the first eigenvector $[x_1]$. In general, one may write

$$\lambda_n = \min_{[x]\in[[x_1],[x_2],\ldots,[x_{n-1}]]^\perp} \frac{[x]^H[A][x]}{[x]^H[x]}, \quad 2 \le n \le N. \tag{1.91}$$

Similarly,

$$\lambda_n = \max_{[x]\in[[x_{n+1}],\ldots,[x_N]]^\perp} \frac{[x]^H[A][x]}{[x]^H[x]}, 1 \le n \le N-1. \tag{1.92}$$

1.4 Eigenvalue Problems for the Laplace Operator on Scalar Field

If the dimension of the underlying vector space of a linear transformation is infinite, the linear transformation is usually a differential operator defined on a subspace, or an integral operator defined on the whole space. In 1894, the French mathematician Jules Henri Poincaré (1854–1912) established the existence of an infinite sequence of eigenvalues and the corresponding eigenfunctions for the Laplace operator under Dirichlet boundary condition. This key result extends the eigenvector and eigenvalue theory of a square matrix and has played an important role in mathematical physics.

1.4.1 Rayleigh Quotient

Let $L^2(\Omega)$ denote the set of all functions defined on a region Ω bounded by Γ in 2D or 3D space such that $\int_{\Omega} |u|^2 d\Omega < \infty$, where $d\Omega$ denotes the area element (2D) or volume element (3D). The **inner product** and the **norm** in $L^2(\Omega)$ are, respectively, defined by

$$(u, v) = \int_{\Omega} u\bar{v}d\Omega, \quad \|u\| = (u, u)^{1/2}. \tag{1.93}$$

Two functions are said to be **orthogonal** if their inner product is zero. Consider the eigenvalue problems for the **Laplace operator** (also called **Laplacian**) $-\nabla^2$ acting on the scalar fields with three different boundary conditions:

1) **Dirichlet problem**:

$$\begin{cases} -\nabla^2 u(\mathbf{r}) = \lambda u(\mathbf{r}), \mathbf{r} \in \Omega \\ u(\mathbf{r}) = 0, \mathbf{r} \in \Gamma \end{cases}. \tag{1.94}$$

2) **Neumann problem**:

$$\begin{cases} -\nabla^2 u(\mathbf{r}) = \lambda u(\mathbf{r}), \mathbf{r} \in \Omega \\ \dfrac{\partial u(\mathbf{r})}{\partial n} = 0, \mathbf{r} \in \Gamma \end{cases}, \tag{1.95}$$

where $\partial/\partial n$ denotes the derivative in the direction normal to the boundary Γ.

3) **Robin problem**:

$$\begin{cases} -\nabla^2 u(\mathbf{r}) = \lambda u(\mathbf{r}), \mathbf{r} \in \Omega \\ \dfrac{\partial u(\mathbf{r})}{\partial n} + a(\mathbf{r})u(\mathbf{r}) = 0, \mathbf{r} \in \Gamma \end{cases}, \tag{1.96}$$

where $a(\mathbf{r})$ is a continuous function.

For any functions u_1 and u_2 satisfying one of the above boundary conditions, the Laplacian $-\nabla^2$ is **symmetric**

$$\begin{aligned}(-\nabla^2 u_1, u_2) &= \int_\Omega -\overline{u}_2 \nabla^2 u_1 d\Omega = \int_\Omega \nabla u_1 \nabla \overline{u}_2 d\Omega - \int_\Gamma \overline{u}_2 \frac{\partial u_1}{\partial n} d\Gamma \\ &= \int_\Omega \nabla u_1 \nabla \overline{u}_2 d\Omega = (u_1, -\nabla^2 u_2),\end{aligned} \tag{1.97}$$

where use has been made of the **Green's first identity**

$$\int_\Omega (u\nabla^2 v + \nabla u \nabla v) d\Omega = \int_\Gamma u \frac{\partial v}{\partial n} d\Gamma. \tag{1.98}$$

The symmetric property of the Laplacian is also easily seen from the **Green's second identity**

$$\int_\Omega (u\nabla^2 v - v\nabla^2 u) d\Omega = \int_\Gamma \left(u\frac{\partial v}{\partial n} - v\frac{\partial u}{\partial n} \right) d\Gamma. \tag{1.99}$$

Apparently,

$$(-\nabla^2 u, u) = \int_\Omega \|\nabla u\|^2 d\Omega \geq 0, \tag{1.100}$$

which implies that the Laplacian $-\nabla^2$ is **non-negative**. The three eigenvalue problems (1.94)–(1.96) have the properties:

1) All the eigenvalues are real and the corresponding eigenfunctions can be chosen to be real.
2) The eigenfunctions corresponding to different eigenvalues are orthogonal.
3) All the eigenvalues are positive for Dirichlet problem (1.94). All the eigenvalues are non-negative (positive or zero) for Neumann and Robin problems (1.95) and (1.96).

The **Rayleigh quotient** for the Laplacian $-\nabla^2$ can be obtained by taking the inner product of the first equation of (1.94) with u:

$$\lambda = \frac{(-\nabla^2 u, u)}{(u, u)}. \tag{1.101}$$

After applying Green's first identity to (1.101), the Rayleigh quotient can be written as

$$\lambda = \frac{1}{(u, u)} \left[(\nabla u, \nabla u) - \int_\Gamma \overline{u} \frac{\partial u}{\partial n} d\Gamma \right]. \tag{1.102}$$

If u satisfies the Dirichlet boundary condition, the boundary integral in (1.102) vanishes and the Rayleigh quotient reduces to

$$\lambda = \frac{(\nabla u, \nabla u)}{(u, u)}. \tag{1.103}$$

Consider the minimization of the Rayleigh quotient

$$\lambda = \min_{h|_\Gamma = 0} \frac{(\nabla h, \nabla h)}{(h, h)}, \tag{1.104}$$

where h is a smooth function (called **trial function**), satisfying the Dirichlet boundary condition $h|_\Gamma = 0$.

Theorem 1.3 If u_1 is a solution of the minimum problem (1.104) and λ_1 is the value of the minimum (i.e. the Rayleigh quotient reaches minimum value λ_1 when $h = u_1$), then λ_1 is the smallest eigenvalue of the Dirichlet problem (1.94) and u_1 its corresponding eigenfunction. □

Proof. Let v be an arbitrary function and ε be a small number. Since λ_1 is a minimum of the Rayleigh quotient (1.103), the function

$$\lambda(\varepsilon) = \frac{(\nabla(u_1 + \varepsilon v), \nabla(u_1 + \varepsilon v))}{(u_1 + \varepsilon v, u_1 + \varepsilon v)} \tag{1.105}$$

has a minimum at $\varepsilon = 0$ with

$$\lambda_1 = \frac{(\nabla u_1, \nabla u_1)}{(u_1, u_1)}. \tag{1.106}$$

By ordinary calculus, this implies $\lambda'(0) = 0$. Thus,

$$\lambda'(0) = \frac{1}{(u_1, u_1)^2}[2(\nabla u_1, \nabla v)(u_1, u_1) - 2(u_1, v)(\nabla u_1, \nabla u_1)] = 0.$$

It follows that

$$(\nabla u_1, \nabla v) - (u_1, v)\frac{(\nabla u_1, \nabla u_1)}{(u_1, u_1)} = 0. \tag{1.107}$$

By using Green's first identity (1.98), (1.106), and the Dirichlet boundary condition, one may obtain

$$\left(\nabla^2 u_1, v\right) + \lambda_1(u_1, v) = 0.$$

Since v is arbitrary, it may be concluded that λ_1 is an eigenvalue of (1.94) and its eigenfunction is u_1. It is easy to verify that λ_1 is the smallest eigenvalue of (1.94). The proof is completed.

Once the first eigenvalue λ_1 and its eigenfunction u_1 are determined, the second eigenvalue λ_2 can be found by minimizing (1.104) with the supplementary constraint that the trial function must be orthogonal to the first eigenfunction u_1. In general, u_n is the nth eigenfunction that minimizes the Rayleigh quotient (1.104) under the conditions

$$(u_n, u_1) = (u_n, u_2) = \cdots (u_n, u_{n-1}) = 0,$$

and the corresponding eigenvalues satisfy $0 < \lambda_1 \le \lambda_2 \le \cdots$.

Consider now the Neumann eigenvalue problem (1.95). Instead of using (1.104), the following minimization problem for the Rayleigh quotient will be introduced:

$$\lambda = \min \frac{(\nabla h, \nabla h)}{(h, h)}, \tag{1.108}$$

where the trial function h is not required to satisfy any boundary conditions.

Theorem 1.4 If u_1 is a solution of the minimum problem (1.108) and λ_1 is the value of the minimum, then λ_1 is the smallest eigenvalue of the Neumann problem (1.95) and u_1 its corresponding eigenfunction. □

Proof. The steps of the proof is similar to that of Theorem 1.3 until (1.107), which, after applying Green's first identity, can be written as

$$\left(\nabla^2 u_1, v\right) + \lambda_1 (u_1, v) = \int_\Gamma v \frac{\partial u_1}{\partial n} d\Gamma. \tag{1.109}$$

Since v is an arbitrary function, one can first choose v arbitrarily inside Ω and $v = 0$ on the boundary Γ. Then, (1.109) implies $\nabla^2 u_1 + \lambda_1 u_1 = 0$ inside Ω. Thus,

$$\int_\Gamma v \frac{\partial u_1}{\partial n} d\Gamma = 0$$

holds for an arbitrary function v. Now selecting $v = \partial u_1/\partial n$ yields $\partial u_1/\partial n = 0$, which is the Neumann boundary condition. The proof is completed.

The above discussion indicates that the Neumann boundary condition is naturally met after minimization. For this reason, it is called **natural boundary condition** or **free boundary condition**. It is noted that the first eigenvalue λ_1 for Neumann problem (1.95) can be zero and the corresponding eigenfunction becomes a constant. Other eigenvalues and their eigenfunctions can be determined by (1.108) in the same way as discussed for Dirichlet problem.

Remark 1.4 It can be shown that the Dirichlet boundary condition in (1.94) becomes a natural boundary condition if one uses the following Rayleigh quotient

$$\lambda = \min \frac{(\nabla h, \nabla h) - 2\int_\Gamma h \frac{\partial h}{\partial n} d\Gamma}{(h, h)}, \tag{1.110}$$

where the trial functions are not required to satisfy any boundary conditions. □

1.4.2 Properties of Eigenvalues

Let λ_j and λ'_j, respectively, denote Dirichlet and Neumann eigenvalues. The Rayleigh quotients in (1.104) and (1.108) have the same expressions except that the trial functions for λ_j satisfy extra constraints. Some important properties of eigenvalues are summarized below [29]:

1) $\lambda'_j < \lambda_j (j = 1, 2, \ldots)$.
2) As the domain increases, the eigenvalues λ_n or λ'_n decreases.
3) If Ω is a plane domain in a 2D space, the eigenvalues for the Dirichlet problem (1.94) satisfy $\lim_{n\to\infty} \frac{\lambda_n}{n} = \frac{4\pi}{A}$, where A is the area of Ω. If Ω is a solid domain in a 3D space, the eigenvalues for the Dirichlet problem (1.94) satisfy $\lim_{n\to\infty} \frac{\lambda_n^{3/2}}{n} = \frac{6\pi^2}{V}$, where V is the volume of Ω.
4) $\lim_{n\to\infty} \lambda_n = \infty$ and $\lim_{n\to\infty} \lambda'_n = \infty$.

The first property results from the consideration that additional constraints will increase the value of minimum. In waveguide theory, this property implies that the dominant mode of a hollow metal waveguide is always a TE mode. The second property is simply because the set of trial functions defined in a larger domain includes those defined in a smaller domain contained in the larger domain. In waveguide theory, this property implies that the cutoff wavenumber of a hollow metal waveguide decreases as the cross section of the waveguide increases. As the cross section becomes infinitely large, the discrete cutoff wavenumbers will become closer together and merge into the positive axis. Only the proof of the last property will be given here. To this end, one needs the following **Rellich's theorem**, named after the Austrian-German mathematician Franz Rellich (1906–1955).

Rellich's Theorem: Any sequence $\{f_n\}$ that satisfies

$$\|f_n\|^2 = \int_\Omega f_n^2 d\Omega \le c_1, \quad \|\nabla f_n\|^2 = \int_\Omega (\nabla f_n)^2 d\Omega \le c_2,$$

where c_1 and c_2 are constants, has a subsequence, still denoted by $\{f_n\}$, such that

$$\lim_{m,n\to\infty} \int_\Omega (f_m - f_n)^2 d\Omega = 0.$$

□

According to the solution procedures of the eigenfunctions described by (1.104) or (1.108), one may write

$$\lambda_n = \frac{(\nabla u_n, \nabla u_n)}{(u_n, u_n)}. \tag{1.111}$$

If λ_n is assumed to be finite as $n \to \infty$, the sequence $\{u_n|(u_n, u_n) = 1\}$ satisfies the conditions stated in Rellich's theorem. So, one can choose a subsequence of $\{u_n|(u_n, u_n) = 1\}$, still denoted by $\{u_n|(u_n, u_n) = 1\}$, such that

$$\lim_{m,n\to\infty} \int_\Omega (u_m - u_n)^2 d\Omega = 0.$$

This contradicts the fact that $\int_\Omega (u_m - u_n)^2 d\Omega = 2$. Therefore, the assumption that λ_n remains finite as $n \to \infty$ is wrong. The proof is completed.

Remark 1.5 For the eigenvalue problems defined in 2D or 3D space, the eigenvalues are numbered multiple integers. For example, if Ω is a rectangle $[0, a] \times [0, b]$ in the plane, one has

$$\lambda_n = \left(\frac{p\pi}{a}\right)^2 + \left(\frac{q\pi}{b}\right)^2, p = 1, 2, \ldots, q = 1, 2, \ldots. \tag{1.112}$$

In this case, n stands for multiple indices (p, q) and it is difficult to see how (1.112) is related to property 3. To solve this problem, one can introduce the **enumeration function** $N(\lambda)$, defined as the number of eigenvalues that do not exceed λ. Clearly $N(\lambda)$ is the number of lattice points (p, q) satisfying

$$\left(\frac{p}{a}\right)^2 + \left(\frac{q}{b}\right)^2 \le \frac{\lambda}{\pi^2}.$$

If the eigenvalues (1.112) are arranged in increasing order, it is easy to see $N(\lambda) = n$. □

1.4.3 Completeness of Eigenfunctions

A set of eigenfunctions is said to be **complete** in $L^2(\Omega)$ if any function in $L^2(\Omega)$ can be expressed as a linear combination of the eigenfunctions.

Theorem 1.5 The eigenfunctions resulted from (1.104) (Dirichlet boundary condition) and (1.108) (Neumann boundary condition) are complete in $L^2(\Omega)$. In other words, an arbitrary function $f \in L^2(\Omega)$ can be represented as a linear combination of the eigenfunctions in the sense that

$$\left\| f - \sum_{n=1}^{N} c_n u_n \right\| \to 0, \tag{1.113}$$

as $N \to \infty$. In the above, $u_n (n = 1, 2, ...)$ are the eigenfunctions obtained either from (1.104) or (1.108), which are assumed to be mutually orthogonal, and

$$c_n = \frac{(f, u_n)}{(u_n, u_n)}, \quad (n = 1, 2, ...) \tag{1.114}$$

are the (Fourier) expansion coefficients. □

Proof. Let f be a trial function. The remainder of the expansion of f

$$r_N = f - \sum_{n=1}^{N} c_n u_n$$

is also a trial function. From the orthogonality of the eigenfunctions, one may easily verify that $(r_N, u_j) = 0$ for $j \leq N$, which implies

$$\lambda_N \leq \frac{(\nabla r_N, \nabla r_N)}{(r_N, r_N)} \tag{1.115}$$

by the construction of the eigenfunctions. After expanding the numerator, one may find

$$(\nabla r_N, \nabla r_N) = (\nabla f, \nabla f) - \sum_{n=1}^{N} c_n^2 \lambda_n (u_n, u_n) \leq \|\nabla f\|^2. \tag{1.116}$$

It follows from (1.115) and (1.116) that

$$\|r_N\|^2 \leq \frac{\|\nabla f\|^2}{\lambda_N}. \tag{1.117}$$

Since $\lambda_N \to \infty$ as $N \to \infty$, one may find $\|r_N\| \to 0$ as $N \to \infty$. Hence, (1.113) is valid for a trial function f. It can be shown that an arbitrary function in $L^2(\Omega)$ can be approximated by a trial function. Thus, the proof is completed.

Remark 1.6 The convergence in the function space $L^2(\Omega)$ as demonstrated in (1.113) is usually nonuniform. Great care must be taken in performing term-by-term differentiation, integration, and limiting process of the series. In many cases, as will be seen later in the theory of waveguide and cavity resonator, the function and its derivative must be expanded separately. □

Remark 1.7 If the set of eigenfunctions $\{u_n\}$ is **orthonormal**

$$(u_m, u_n) = \delta_{mn}, \tag{1.118}$$

and satisfy the homogeneous Dirichlet condition in (1.94) or Neumann boundary condition in (1.95), one can construct a set of vector field $\{\mathbf{E}_n\}$, defined by $\mathbf{E}_n = (1/\lambda_n)\nabla u_n$, such that

$$\int_\Omega \mathbf{E}_m \cdot \mathbf{E}_n d\Omega = \delta_{mn}. \tag{1.119}$$

Equation (1.119) is a direct consequence of the Green's first identity (1.98). □

1.4.4 Differential Equations with Variable Coefficients

The main results obtained in previous sections can be generalized to the differential equations with variable coefficients

$$\frac{1}{w(\mathbf{r})}[-\nabla \cdot p(\mathbf{r})\nabla + q(\mathbf{r})]u(\mathbf{r}) = \lambda u(\mathbf{r}), \mathbf{r} \in \Omega, \tag{1.120}$$

where p and w are assumed to be positive, and q is a continuous function. Three boundary value problems similar to (1.94)–(1.96) may be introduced with the operator $-\nabla^2$ replaced by $(-\nabla \cdot p\nabla + q)/w$. The inner product and the norm are now, respectively, defined by

$$(u, v)_w = \int_\Omega w(\mathbf{r})u(\mathbf{r})\bar{v}(\mathbf{r})d\Omega, \quad \|u\|_w = \sqrt{(u, u)_w}.$$

Let $L_w^2(\Omega)$ denote the set of all functions defined on the region Ω in 2D or 3D space such that $\|u\|_w < \infty$. Then, all the eigenvalues of (1.120) under the three boundary conditions are real and the corresponding eigenfunctions can be chosen to be real, and the eigenfunctions corresponding to different eigenvalues are orthogonal. The Rayleigh quotient for the eigenvalue problem (1.120) is given by

$$\lambda = \frac{\left(\frac{1}{w}(-\nabla \cdot p\nabla + q)u, u\right)_w}{(u, u)_w}. \tag{1.121}$$

By use of the identity $\nabla \cdot (\phi \mathbf{a}) = \phi \nabla \cdot \mathbf{a} + \mathbf{a} \cdot \nabla \phi$, the above quotient can be written as

$$\lambda = \frac{\int_{\Omega} (p|\nabla u|^2 + qu^2) d\Omega - \int_{\Gamma} p \frac{\partial u}{\partial n} u d\Gamma}{(u, u)_w}. \tag{1.122}$$

For the Dirichlet and Neumann problems, the boundary term in (1.122) vanishes and the quotient reduces to

$$\lambda = \frac{\int_{\Omega} (p|\nabla u|^2 + qu^2) d\Omega}{(u, u)_w}. \tag{1.123}$$

During the minimization process for determining the eigenvalues and the eigenfunctions, the trial functions in the quotient (1.123) must satisfy the Dirichlet boundary condition for Dirichlet problem while they are free for Neumann problem. For the Robin problem, (1.122) becomes

$$\lambda = \frac{\int_{\Omega} (p|\nabla u|^2 + qu^2) d\Omega + \int_{\Gamma} pau^2 d\Gamma}{(u, u)_w}. \tag{1.124}$$

In this case, the trial functions are also free.

As a special case in one dimension, let us consider the most important eigenvalue problem in mathematical physics, the **Sturm–Liouville type**

$$\frac{1}{w(x)} \left[-\frac{d}{dx} p(x) \frac{d}{dx} + q(x) \right] u(x) = \lambda u(x), x \in (a, b), \tag{1.125}$$

where the functions p and w are smooth and positive, and q is real piecewise continuous function, all defined over a finite interval $[a, b]$. Then, the following properties can be established for the eigenvalue problem (1.125):

1) The eigenvalues are real and all the corresponding eigenfunctions can be assumed to be real.
2) The eigenfunctions of different eigenvalues are orthogonal.
3) The eigenvalue problem (1.125) has an infinite set of eigenvalues $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n \cdots$, and $\lambda_n \to \infty$ as $n \to \infty$. In addition, the corresponding eigenfunctions $\{u_n\}$ constitute a complete system in $L_w^2(a, b)$.

1.4.5 Green's Function and Spectral Representation

The eigenvalue theory can be used to solve the operator equation of the type

$$(\hat{L} - \xi)g(\mathbf{r}, \xi) = f(\mathbf{r}), \tag{1.126}$$

where $\hat{L}$ is a self-adjoint operator, ξ is a complex parameter, f is a known source function, and g is the unknown field. Suppose that $\{u_n\}$ is the complete set of the normalized eigenfunctions of the eigenvalue problem, which satisfy

$$\hat{L}u_n = \lambda_n u_n.$$

Then, both the source and the field may have the expansions

$$f = \sum_n f_n u_n, \quad g = \sum_n g_n u_n,$$

with $f_n = (f, u_n)$, $g_n = (g, u_n)$. Substituting these expansions into (1.126) gives

$$g = -\sum_n \frac{f_n}{\xi - \lambda_n} u_n. \tag{1.127}$$

Let C_R be a circle of radius R at the origin in the complex ξ-plane. Then,

$$\int_{C_R} g d\xi = -\sum_n f_n u_n \int_{C_R} \frac{d\xi}{\xi - \lambda_n}, \tag{1.128}$$

where the sum is over those eigenvalues λ_n contained within the circle. The singularities of the integrand are simple poles with residue of unity at all $\xi = \lambda_n$ within the contour. Since $\hat{L}$ is self-adjoint, the poles must lie on the real axis in the ξ-plane. Taking the limit as $R \to \infty$ and using the residue theorem, one may find

$$\lim_{R \to \infty} \int_{C_R} g d\xi = -2\pi j \sum_n f_n u_n,$$

where the sum is now over all of the eigenfunctions. Therefore, there exists a relationship between the source and the field

$$f = -\frac{1}{2\pi j} \int_C g d\xi, \tag{1.129}$$

where C is a circle at infinity obtained in the limit operation. As a special case, one may consider the Green's function defined by

$$(\hat{L} - \xi)G(\mathbf{r}, \mathbf{r}'; \xi) = \delta(\mathbf{r} - \mathbf{r}'). \tag{1.130}$$

From (1.129), it is easy to see

$$\delta(\mathbf{r}-\mathbf{r}') = -\frac{1}{2\pi j}\int_C G(\mathbf{r},\mathbf{r}';\xi)d\xi. \tag{1.131}$$

This is called the **spectral representation of the delta function** for the operator $\hat{L}$. From (1.127), the Green's function has the expansion

$$G(\mathbf{r},\mathbf{r}';\xi) = -\sum_n \frac{u_n(\mathbf{r})\overline{u}_n(\mathbf{r}')}{\xi-\lambda_n}. \tag{1.132}$$

Combining (1.131) and (1.132) yields

$$\delta(\mathbf{r}-\mathbf{r}') = \sum_n u_n(\mathbf{r})\overline{u}_n(\mathbf{r}'). \tag{1.133}$$

Equation (1.133) is the **completeness identity** of the eigenfunctions and it should be taken as a symbolic equality.

Example 1.3 Let us consider the Green's function defined by

$$\begin{cases} -\left(\dfrac{d^2}{dx^2}+\xi\right)G(x,x';\xi) = \delta(x-x') \\ G(0,x';\xi) = G(a,x';\xi) = 0 \end{cases}.$$

The Green's function is easily found to be [30]

$$G(x,x';\xi) = \frac{\sin\sqrt{\xi}x\sin\sqrt{\xi}(a-x')}{\sqrt{\xi}\sin\sqrt{\xi}a}H(x'-x) + \frac{\sin\sqrt{\xi}(a-x)\sin\sqrt{\xi}x'}{\sqrt{\xi}\sin\sqrt{\xi}a}H(x-x').$$

Considering

$$\frac{1}{2\pi j}\int_C G(x,x';\xi)d\xi = \frac{1}{2\pi j}\int_C \frac{\sin\sqrt{\xi}x\sin\sqrt{\xi}(a-x')}{\sqrt{\xi}\sin\sqrt{\xi}a}H(x'-x)d\xi + \frac{1}{2\pi j}\int_C \frac{\sin\sqrt{\xi}(a-x)\sin\sqrt{\xi}x'}{\sqrt{\xi}\sin\sqrt{\xi}a}H(x-x')d\xi$$

and

$$\begin{aligned} \operatorname*{Res}_{\xi=(n\pi/a)^2}\frac{\sin\sqrt{\xi}x\sin\sqrt{\xi}(a-x')}{\sqrt{\xi}\sin\sqrt{\xi}a} &= \left.\frac{2\sin\sqrt{\xi}x\sin\sqrt{\xi}(a-x')}{\xi^{-1/2}\sin\sqrt{\xi}a + a\cos\sqrt{\xi}a}\right|_{\xi=(n\pi/a)^2} \\ &= -\frac{2}{a}\sin\frac{n\pi x}{a}\sin\frac{n\pi x'}{a} \\ &= \operatorname*{Res}_{\xi=(n\pi/a)^2}\frac{\sin\sqrt{\xi}(a-x)\sin\sqrt{\xi}x'}{\sqrt{\xi}\sin\sqrt{\xi}a}, \end{aligned}$$

where $\operatorname*{Res}_{\xi=a} f(\xi)$ denotes the residue of $f(\xi)$ at $\xi = a$, one may find

$$\frac{1}{2\pi j}\int_C G(x,x';\xi)d\xi = -\frac{2}{a}\sum_{n=1}^{\infty} \sin\frac{n\pi x}{a}\sin\frac{n\pi x'}{a}$$

from residue theorem. It follows from (1.131) that

$$\delta(x-x') = \sum_{n=1}^{\infty}\sqrt{\frac{2}{a}}\sin\frac{n\pi x}{a}\sqrt{\frac{2}{a}}\sin\frac{n\pi x'}{a}. \tag{1.134}$$

By shifting the origin to $a/2$, the above equation can be written as

$$\begin{aligned}\delta(x-x') &= \frac{2}{a}\sum_{n=1}^{\infty}\sin\frac{n\pi}{a}\left(x-\frac{a}{2}\right)\sin\frac{n\pi}{a}\left(x'-\frac{a}{2}\right)\\ &= \frac{2}{a}\sum_{n=1}^{\infty}\left(\sin\frac{n\pi x}{a}\sin\frac{n\pi x'}{a}\cos^2\frac{n\pi}{2} + \cos\frac{n\pi x}{a}\cos\frac{n\pi x'}{a}\sin^2\frac{n\pi}{2}\right)\\ &= \frac{2}{a}\sum_{n=1}^{\infty}\sin\frac{2n\pi x}{a}\sin\frac{2n\pi x'}{a} + \frac{2}{a}\sum_{n=1}^{\infty}\cos\frac{(2n-1)\pi x}{a}\cos\frac{(2n-1)\pi x'}{a}.\end{aligned} \tag{1.135}$$

As a becomes very large, the eigenvalues $\lambda_n = (n\pi/a)^2$ become closer and closer together until they merge in the limit into a continuous spectrum (positive axis). Setting

$$\tau_n = \frac{n\pi}{a}, \Delta\tau_n = \frac{\pi}{a},$$

and letting $a \to \infty$, (1.134) can be written as

$$\begin{aligned}\delta(x-x') &= \lim_{a\to\infty}\frac{2}{\pi}\sum_{n=1}^{\infty}\sin\frac{n\pi x}{a}\sin\frac{n\pi x'}{a}\frac{\pi}{a}\\ &= \lim_{\Delta\tau_n\to\infty}\frac{2}{\pi}\sum_{n=1}^{\infty}\sin\tau_n x\sin\tau_n x'\Delta\tau_n = \frac{2}{\pi}\int_0^{\infty}\sin\tau x\sin\tau x' d\tau.\end{aligned} \tag{1.136}$$

Similarly, if a approaches to infinity, (1.135) becomes

$$\delta(x-x') = \frac{1}{\pi}\int_0^{\infty}\cos\tau(x-x')d\tau = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{j\tau(x-x')}d\tau. \tag{1.137}$$

This is the **Fourier integral representation**. □

Example 1.4 Consider the eigenvalue problem

$$-\nabla^2 u(\mathbf{r}) = \lambda u(\mathbf{r}), \mathbf{r} = (x, y, z) \in R^3.$$

The above equation has normalized plane-wave solutions

$$u_{\mathbf{k}}(\mathbf{r}) = \frac{1}{(2\pi)^{3/2}} e^{j\mathbf{k}\cdot\mathbf{r}}, \quad \mathbf{k} = (k_x, k_y, k_z),$$

which satisfy the orthonormal condition

$$\int_{R^3} u_{\mathbf{k}}(\mathbf{r}) \bar{u}_{\mathbf{k}'}(\mathbf{r}) dxdydz = \frac{1}{(2\pi)^3} \int_{R^3} e^{j(\mathbf{k}-\mathbf{k}')\cdot\mathbf{r}} dxdydz = \delta(\mathbf{k}-\mathbf{k}').$$

The corresponding eigenvalue is given by $\lambda = k^2 = |\mathbf{k}|^2$. The Green's function defined by

$$\left(\nabla^2 + k_0^2\right) G(\mathbf{r}, \mathbf{r}') = -\delta(\mathbf{r}-\mathbf{r}'), \mathbf{r}, \mathbf{r}' \in R^3$$

is easily found to be $e^{-jk_0|\mathbf{r}-\mathbf{r}'|}/4\pi|\mathbf{r}-\mathbf{r}'|$. In the expansion (1.132), one may let $\xi = k_0^2$, replace the discrete summation index n by the continuous vector $\mathbf{k}$, u_n by $u_{\mathbf{k}}$, λ_n by k^2, and the discrete summation by the integral over the whole space, to find

$$\frac{e^{-jk_0|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|} = \frac{1}{(2\pi)^3} \int_{R^3} \frac{e^{j\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')}}{k^2 - k_0^2} dk_x dk_y dk_z. \tag{1.138}$$

The above expression is the plane-wave expansion for the Green's function $e^{-jk_0|\mathbf{r}-\mathbf{r}'|}/4\pi|\mathbf{r}-\mathbf{r}'|$. □

1.5 Eigenvalue Problems for the Laplace Operator on Vector Field

A vector field can be decomposed into three scalar field components and a vector field problem can always be reduced to scalar field problems. But this procedure is not always the most effective. One can also express a vector field as the sum of the gradient of a scalar potential and the curl of a vector potential. Since the vector potential must satisfy the gauge condition, only two components of the vector potential are independent. A vector field is usually required to satisfy certain boundary conditions, which must be imposed to the scalar field components that represent the original vector field. The imposition of the boundary conditions to the scalar fields is very complicated in many cases. Therefore, one is forced to adopt a direct approach to solve vector field equation [20].

1.5.1 Rayleigh Quotient

Consider the eigenvalue problems for the Laplace operator on the vector field with the **boundary condition of electric typ**e (BCE)

$$\begin{cases} -\nabla^2 \mathbf{e}(\boldsymbol{\rho}) = \lambda \mathbf{e}(\boldsymbol{\rho}), \boldsymbol{\rho} \in \Omega \\ \text{BCE} : \mathbf{u}_n \times \mathbf{e}(\boldsymbol{\rho}) = 0, \nabla \cdot \mathbf{e}(\boldsymbol{\rho}) = 0, \boldsymbol{\rho} \in \Gamma \end{cases}, \tag{1.139}$$

and that with the **boundary condition of magnetic type** (BCM)

$$\begin{cases} -\nabla^2 \mathbf{h}(\boldsymbol{\rho}) = \lambda \mathbf{h}(\boldsymbol{\rho}), \boldsymbol{\rho} \in \Omega \\ \text{BCM} : \mathbf{u}_n \cdot \mathbf{h}(\boldsymbol{\rho}) = 0, \mathbf{u}_n \times \nabla \times \mathbf{h}(\boldsymbol{\rho}) = 0, \boldsymbol{\rho} \in \Gamma \end{cases}, \tag{1.140}$$

where the operator ∇^2 is defined by $\nabla^2 \mathbf{a} = -\nabla \times \nabla \times \mathbf{a} + \nabla \nabla \cdot \mathbf{a}$. For convenience, the domain Ω will be assumed to be a finite region in 2D space and is bounded by Γ. Let $[L^2(\Omega)]^2$ denote the **product space** $L^2(\Omega) \times L^2(\Omega)$. The **inner product** and the **norm** in $[L^2(\Omega)]^2$ are, respectively, defined by

$$(\mathbf{a}_1, \mathbf{a}_2) = \int_\Omega \mathbf{a}_1 \cdot \bar{\mathbf{a}}_2 d\Omega, \ \ \|\mathbf{a}\| = (\mathbf{a}, \mathbf{a})^{1/2}, \ \ \mathbf{a}_1, \mathbf{a}_2, \mathbf{a} \in \left[L^2(\Omega)\right]^2. \tag{1.141}$$

It is noted that the BCE implies that the tangential component as well as the normal derivative of the normal component of the vector field $\mathbf{e}$ vanish on the boundary while the BCM implies that the normal component and the normal derivative of the tangential component of the vector field $\mathbf{h}$ are zero on the boundary. For any two vector fields $\mathbf{a}_1$ and $\mathbf{a}_2$ satisfying the BCE or BCM, the Laplacian $-\nabla^2$ is symmetric

$$\begin{aligned} \left(-\nabla^2 \mathbf{a}_1, \mathbf{a}_2\right) &= \int_\Omega \nabla \times \mathbf{a}_1 \cdot \nabla \times \bar{\mathbf{a}}_2 d\Omega + \int_\Omega \nabla \cdot \mathbf{a}_1 \nabla \cdot \bar{\mathbf{a}}_2 d\Omega \\ &\quad - \int_\Gamma (\mathbf{u}_n \times \bar{\mathbf{a}}_2) \cdot \nabla \times \mathbf{a}_1 d\Gamma - \int_\Gamma (\mathbf{u}_n \cdot \bar{\mathbf{a}}_2) \nabla \cdot \mathbf{a}_1 d\Gamma \\ &= \int_\Omega (\nabla \times \mathbf{a}_1 \cdot \nabla \times \bar{\mathbf{a}}_2 + \nabla \cdot \mathbf{a}_1 \nabla \cdot \bar{\mathbf{a}}_2) d\Omega = \left(\mathbf{a}_1, -\nabla^2 \mathbf{a}_2\right) \end{aligned} \tag{1.142}$$

after applying integration by parts. The Laplacian $-\nabla^2$ is **non-negative**

$$\left(-\nabla^2 \mathbf{a}, \mathbf{a}\right) = \int_\Omega \left(|\nabla \times \mathbf{a}|^2 + |\nabla \cdot \mathbf{a}|^2\right) d\Omega \geq 0.$$

Similar to the eigenvalue problems for scalar fields, the eigenvalue problems (1.139) and (1.140) for vector fields have the properties:

1) All the eigenvalues are real and the corresponding eigenfunctions can be chosen to be real.
2) The eigenfunctions corresponding to different eigenvalues are orthogonal.
3) All the eigenvalues are positive or zero.

From now on, it will be assumed that all the eigenfunctions are real, and a shorthand notation $\mathbf{a}^2 = \mathbf{a} \cdot \mathbf{a}$ will be used. The **Rayleigh quotient** for the Laplacian $-\nabla^2$is

$$\lambda = \frac{(-\nabla^2\mathbf{a}, \mathbf{a})}{(\mathbf{a}, \mathbf{a})} = \frac{\int_\Omega (\nabla \times \mathbf{a})^2 d\Omega + \int_\Omega (\nabla \cdot \mathbf{a})^2 d\Omega - \int_\Gamma (\mathbf{u}_n \times \mathbf{a}) \cdot \nabla \times \mathbf{a} d\Gamma - \int_\Gamma (\mathbf{u}_n \cdot \mathbf{a}) \nabla \cdot \mathbf{a} d\Gamma}{\int_\Omega \mathbf{a}^2 d\Omega}. \tag{1.143}$$

The line integrals vanish if $\mathbf{a}$ satisfies the BCE or BCM. One may thus introduce the minimization problem

$$\lambda = \min_{\text{BCE or BCM}} \frac{\int_\Omega \left[(\nabla \times \mathbf{g})^2 + (\nabla \cdot \mathbf{g})^2\right] d\Omega}{\int_\Omega \mathbf{g}^2 d\Omega}, \tag{1.144}$$

where $\mathbf{g}$ is the trial function and is assumed to satisfy the BCE or BCM.

Theorem 1.6 If $\mathbf{a}_1$ is a solution of the minimum problem (1.144) and λ_1 is the value of the minimum, then λ_1 is the smallest eigenvalue of the Rayleigh quotient (1.144) and $\mathbf{a}_1$ is its corresponding eigenfunction. □

Proof. Let $\mathbf{d}$ be an arbitrary function that satisfies the BCE or BCM, and ε be a small number. Since λ_1 is a minimum of (1.144), the expression

$$\lambda(\varepsilon) = \frac{\int_\Omega \left\{[\nabla \times (\mathbf{a}_1 + \varepsilon\mathbf{d})]^2 + [\nabla \cdot (\mathbf{a}_1 + \varepsilon\mathbf{d})]^2\right\} d\Omega}{\int_\Omega (\mathbf{a}_1 + \varepsilon\mathbf{d})^2 d\Omega} \tag{1.145}$$

has a minimum at $\varepsilon = 0$ with

$$\lambda_1 = \frac{\int_\Omega \left[(\nabla \times \mathbf{a}_1)^2 + (\nabla \cdot \mathbf{a}_1)^2\right] d\Omega}{\int_\Omega \mathbf{a}_1^2 d\Omega}. \tag{1.146}$$

By ordinary calculus, this requires $\lambda'(0) = 0$. A simple calculation leads to

$$\begin{aligned} \lambda'(0) \; &= \frac{2\int_\Omega \mathbf{a}_1^2 d\Omega \int_\Omega \nabla \times \mathbf{a}_1 \cdot \nabla \times \mathbf{d} d\Omega - 2\int_\Omega \mathbf{a}_1 \cdot \mathbf{d} d\Omega \int_\Omega (\nabla \times \mathbf{a}_1)^2 d\Omega}{\left(\int_\Omega \mathbf{a}_1^2 d\Omega\right)^2} \\ &+ \frac{2\int_\Omega \mathbf{a}_1^2 d\Omega \int_\Omega \nabla \cdot \mathbf{a}_1 \nabla \cdot \mathbf{d} d\Omega - 2\int_\Omega \mathbf{a}_1 \cdot \mathbf{d} d\Omega \int_\Omega (\nabla \cdot \mathbf{a}_1)^2 d\Omega}{\left(\int_\Omega \mathbf{a}_1^2 d\Omega\right)^2}. \end{aligned}$$

The condition that $\lambda'(0)$ must be zero implies

$$\begin{aligned} &\int_\Omega \mathbf{a}_1^2 d\Omega \int_\Omega \nabla \times \mathbf{a}_1 \cdot \nabla \times \mathbf{d} d\Omega - \int_\Omega \mathbf{a}_1 \cdot \mathbf{d} d\Omega \int_\Omega (\nabla \times \mathbf{a}_1)^2 d\Omega \\ &+ \int_\Omega \mathbf{a}_1^2 d\Omega \int_\Omega \nabla \cdot \mathbf{a}_1 \nabla \cdot \mathbf{d} d\Omega - \int_\Omega \mathbf{a}_1 \cdot \mathbf{d} d\Omega \int_\Omega (\nabla \cdot \mathbf{a}_1)^2 d\Omega = 0. \end{aligned} \tag{1.147}$$

By the vector identities

$$\begin{aligned} \int_\Omega \nabla \times \mathbf{a} \cdot \nabla \times \mathbf{b} d\Omega &= \int_\Omega \mathbf{a} \cdot \nabla \times \nabla \times \mathbf{b} d\Omega + \int_\Gamma \nabla \times \mathbf{b} \cdot (\mathbf{u}_n \times \mathbf{a}) d\Gamma, \\ \int_\Omega \nabla \cdot \mathbf{a} \nabla \cdot \mathbf{b} d\Omega &= -\int_\Omega \mathbf{a} \cdot \nabla \nabla \cdot \mathbf{b} d\Omega + \int_\Gamma \nabla \cdot \mathbf{b} (\mathbf{u}_n \cdot \mathbf{a}) d\Gamma, \end{aligned} \tag{1.148}$$

(1.147) may be rewritten as

$$\int_\Omega \mathbf{d}\cdot\nabla\times\nabla\times\mathbf{a}_1 d\Omega - \int_\Omega \mathbf{d}\cdot\nabla\nabla\cdot\mathbf{a}_1 d\Omega - \int_\Omega \mathbf{a}_1\cdot\mathbf{d}d\Omega \frac{\int_\Omega \left[(\nabla\times\mathbf{a}_1)^2 + (\nabla\cdot\mathbf{a}_1)^2\right]d\Omega}{\int_\Omega \mathbf{a}_1^2 d\Omega}$$
$$+ \int_\Gamma (\mathbf{u}_n\times\mathbf{d})\cdot\nabla\times\mathbf{a}_1 d\Gamma + \int_\Gamma (\mathbf{u}_n\cdot\mathbf{d})\nabla\cdot\mathbf{a}_1 d\Gamma = 0. \tag{1.149}$$

On account of (1.146) and the same boundary conditions that both $\mathbf{d}$ and $\mathbf{e}_1$ must satisfy, the line integrals in (1.149) vanish. Thus,

$$\int_\Omega \mathbf{d}\cdot(\nabla\times\nabla\times\mathbf{a}_1 - \nabla\nabla\cdot\mathbf{a}_1 - \lambda_1\mathbf{a}_1)d\Omega = 0. \tag{1.150}$$

Since $\mathbf{d}$ is arbitrary, the above equation indicates that λ_1 is an eigenvalue of (1.139) and its eigenfunction is $\mathbf{a}_1$. It is easy to verify that λ_1 is the smallest eigenvalue of (1.139) or (1.140). The proof is completed.

1.5.2 Completeness of Vector Modal Functions

A procedure similar to the study of eigenvalue problems for scalar fields produces a set of orthogonal eigenfunctions $\{\mathbf{a}_1, \mathbf{a}_2, \ldots\}$, and the corresponding eigenvalues satisfy $0 \le \lambda_1 \le \lambda_2 \le \cdots$. The eigenfunction $\mathbf{a}_n$ is called nth **vector modal function**. From now on, it will be assumed that the vector modal functions are orthonormal

$$\int_\Omega \mathbf{a}_m\cdot\mathbf{a}_n d\Omega = \delta_{mn}. \tag{1.151}$$

Theorem 1.7

$$\lim_{n\to\infty}\lambda_n = \infty.$$

□

Proof. Since $\mathbf{a}_n$ is normalized, one may write

$$\lambda_n = \int_\Omega \left[(\nabla\times\mathbf{a}_n)^2 + (\nabla\cdot\mathbf{a}_n)^2\right]d\Omega = \int_\Omega (\nabla\times\nabla\times\mathbf{a}_n - \nabla\nabla\cdot\mathbf{a}_n)\cdot\mathbf{a}_n d\Omega \tag{1.152}$$

from (1.143). Using the vector identity $\nabla\times\nabla\times\mathbf{a}-\nabla\nabla\cdot\mathbf{a}=-\nabla^2\mathbf{a}$ and Green's first identity (1.98), the above equation may be reduced to

$$\begin{aligned}\lambda_n &= -\int_\Omega \mathbf{a}_n\cdot\nabla^2\mathbf{a}_n d\Omega = -\int_\Omega\left(a_{xn}\nabla^2 a_{xn}+a_{yn}\nabla^2 a_{yn}\right)d\Omega \\ &= \int_\Omega\left[(\nabla a_{xn})^2+\left(\nabla a_{yn}\right)^2\right]d\Omega-\int_\Gamma \mathbf{a}_n\cdot\frac{\partial\mathbf{a}_n}{\partial n}d\Gamma,\end{aligned} \tag{1.153}$$

where the decomposition $\mathbf{a}_n=a_{xn}\mathbf{u}_x+a_{yn}\mathbf{u}_y$ in the rectangular coordinate system has been used. The boundary integral term vanishes due to the boundary conditions BCE or BCM. In fact, one may write

$$\begin{aligned}\mathbf{a}_n\cdot\frac{\partial\mathbf{a}_n}{\partial n} &= [(\mathbf{a}_n\cdot\mathbf{u}_n)\mathbf{u}_n+(\mathbf{a}_n\cdot\mathbf{u}_t)\mathbf{u}_t]\cdot\left[\frac{\partial(\mathbf{a}_n\cdot\mathbf{u}_n)}{\partial n}\mathbf{u}_n+\frac{\partial(\mathbf{a}_n\cdot\mathbf{u}_t)}{\partial n}\mathbf{u}_t\right] \\ &= (\mathbf{a}_n\cdot\mathbf{u}_n)\frac{\partial(\mathbf{a}_n\cdot\mathbf{u}_n)}{\partial n}+(\mathbf{a}_n\cdot\mathbf{u}_t)\frac{\partial(\mathbf{a}_n\cdot\mathbf{u}_t)}{\partial n},\end{aligned} \tag{1.154}$$

where $\mathbf{u}_n$ and $\mathbf{u}_t$ are, respectively, the unit normal and unit tangent along the boundary Γ, and can be used to decompose the field $\mathbf{a}_n$, as illustrated in Figure 1.3.

If the field $\mathbf{a}_n$ satisfies the BCE or BCM, it is easy to show that

$$\mathbf{a}_n\cdot\mathbf{u}_t=0,\ \frac{\partial(\mathbf{a}_n\cdot\mathbf{u}_n)}{\partial n}=0,\boldsymbol{\rho}\in\Gamma, \tag{1.155}$$

and

$$\mathbf{a}_n\cdot\mathbf{u}_n=0,\ \frac{\partial(\mathbf{a}_n\cdot\mathbf{u}_t)}{\partial n}=0,\boldsymbol{\rho}\in\Gamma, \tag{1.156}$$

respectively, for the BCE and BCM. Substituting (1.155) or (1.156) into (1.154), one may find $\mathbf{a}_n\cdot\partial\mathbf{a}_n/\partial n=0$ in both cases. Thus,

$$\lambda_n=\int_\Omega\left[(\nabla a_{xn})^2+\left(\nabla a_{yn}\right)^2\right]d\Omega.\quad\left[L^2(\Omega)\right]^2. \tag{1.157}$$

If λ_n is assumed to be finite as $n\to\infty$, Rellich's theorem thus applies. As a result, one can choose a subsequence of $\{\mathbf{a}_n\}$, still denoted $\{\mathbf{a}_n\}$, such that

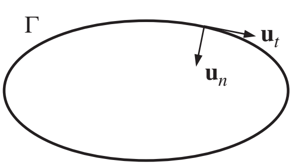

Figure 1.3 Decomposition of field along local rectangular coordinate system.

$\lim_{n,m\to\infty}\|a_{xn}-a_{xm}\|=0$. Another subsequence can be chosen from this subsequence, such that $\lim_{n,m\to\infty}\|a_{yn}-a_{ym}\|=0$. Then,

$$\lim_{n,m\to\infty}\|\mathbf{a}_n-\mathbf{a}_m\|^2=\lim_{n,m\to\infty}\left(\|a_{xn}-a_{xm}\|^2+\|a_{yn}-a_{ym}\|^2\right)=0.$$

This contradicts the fact that $\|\mathbf{a}_n-\mathbf{a}_m\|^2=2$. Consequently, the assumption that λ_n remains finite as $n\to\infty$ is invalid. The proof is completed.

Theorem 1.8 The eigenfunctions determined from (1.144) are complete in $[L^2(\Omega)]^2$. In other words, an arbitrary vector function $\mathbf{F}\in[L^2(\Omega)]^2$ can be expressed as a linear combination of the eigenfunctions $\{\mathbf{a}_n\}$ in the sense that

$$\lim_{N\to\infty}\left\|\mathbf{F}-\sum_{n=1}^{N}c_n\mathbf{a}_n\right\|\to 0, \tag{1.158}$$

where $c_n=(\mathbf{F},\mathbf{a}_n)$, $n=1,2,\ldots$, are the (Fourier) expansion coefficients. □

Proof. Let $\mathbf{F}$ be a trial function. The remainder of the expansion of $\mathbf{F}$

$$\mathbf{r}_N=\mathbf{F}-\sum_{n=1}^{N}c_n\mathbf{a}_n$$

is also a trial function. From the orthogonality of the eigenfunctions, it is easy to verify that $(\mathbf{r}_N,\mathbf{a}_j)=0$ for $j\le N$, which implies

$$\lambda_N\le\frac{\int_\Omega\left(|\nabla\times\mathbf{r}_N|^2+|\nabla\cdot\mathbf{r}_N|^2\right)d\Omega}{\int_\Omega|\mathbf{r}_N|^2d\Omega} \tag{1.159}$$

by the construction of the eigenfunctions. Expanding the integrands in the numerator, one may find

$$\begin{aligned}|\nabla\times\mathbf{r}_N|^2&=|\nabla\times\mathbf{F}|^2-2\sum_{n=1}^{N}c_n[\mathbf{F}\cdot\nabla\times\nabla\times\mathbf{a}_n+\nabla\cdot(\mathbf{F}\times\nabla\times\mathbf{a}_n)]\\&\quad+\sum_{m=1}^{N}\sum_{n=1}^{N}c_mc_n[\mathbf{a}_m\cdot\nabla\times\nabla\times\mathbf{a}_n+\nabla\cdot(\mathbf{a}_m\times\nabla\times\mathbf{a}_n)],\\(\nabla\cdot\mathbf{r}_N)^2&=(\nabla\cdot\mathbf{F})^2-2\sum_{n=1}^{N}c_n[\nabla\cdot(\mathbf{F}\nabla\cdot\mathbf{a}_n)-\mathbf{F}\cdot\nabla\nabla\cdot\mathbf{a}_n]\\&\quad+\sum_{m=1}^{N}\sum_{n=1}^{N}c_mc_n[\nabla\cdot(\mathbf{a}_m\nabla\cdot\mathbf{a}_n)-\mathbf{a}_m\cdot\nabla\nabla\cdot\mathbf{a}_n].\end{aligned}$$

Thus,

$$\int_\Omega (|\nabla \times \mathbf{r}_N|^2 + |\nabla \cdot \mathbf{r}_N|^2) d\Omega = \int_\Omega \left[|\nabla \times \mathbf{F}|^2 + |\nabla \cdot \mathbf{F}|^2\right] d\Omega - \sum_{n=1}^{N} c_n^2 \lambda_n \leq \int_\Omega \left[|\nabla \times \mathbf{F}|^2 + |\nabla \cdot \mathbf{F}|^2\right] d\Omega. \tag{1.160}$$

It follows from (1.159) and (1.160) that

$$\|\mathbf{r}_N\|^2 \leq \frac{\int_\Omega \left[|\nabla \times \mathbf{F}|^2 + |\nabla \cdot \mathbf{F}|^2\right] d\Omega}{\lambda_N}. \tag{1.161}$$

Since $\lim_{N\to\infty} \lambda_N = \infty$, it is to see $\lim_{N\to\infty} \|\mathbf{r}_N\| = 0$. Hence (1.158) is valid for a trial function $\mathbf{F}$. It can be shown that an arbitrary function in $[L^2(\Omega)]^2$ can be approximated by a trial function. The proof is completed.

Remark 1.8 In the proof of the completeness of eigenfunctions for both scalar and vector fields, it has been stated that an arbitrary function in $L^2(\Omega)$ or $[L^2(\Omega)]^2$ can be approximated by a trial function without giving the details. Serious readers may find that a proof for the existence of eigenfunctions is also missing in the preceding discussion. To prove the existence and completeness of eigenfunctions, the concept of the generalized solutions for differential equations must be used, which is beyond the scope of this book. For a rigorous treatment, please refer to [16, 31]. □

Remark 1.9 Similar to (1.110), the BCE and BCM can be made free by, respectively, introducing the minimization problems [32]

$$\lambda = \min \frac{\int_\Omega \left[(\nabla \times \mathbf{g})^2 + (\nabla \cdot \mathbf{g})^2\right] d\Omega - 2\int_\Gamma (\mathbf{u}_n \times \mathbf{g}) \cdot \nabla \times \mathbf{g} d\Gamma}{\int_\Omega \mathbf{g}^2 d\Omega}, \tag{1.162}$$

$$\lambda = \min \frac{\int_\Omega \left[(\nabla \times \mathbf{g})^2 + (\nabla \cdot \mathbf{g})^2\right] d\Omega - 2\int_\Gamma (\mathbf{u}_n \cdot \mathbf{g}) \nabla \cdot \mathbf{g} d\Gamma}{\int_\Omega \mathbf{g}^2 d\Omega}. \tag{1.163}$$

It is readily shown that the solution of (1.162) is an eigenfunction that satisfies (1.139). Let **d** be an arbitrary function and ε be a small number. Suppose that **a** is a solution of (1.162). Then, the expression

$$\lambda(\varepsilon)=\frac{\int\limits_{\Omega}\left\{[\nabla\times(\mathbf{a}+\varepsilon\mathbf{d})]^2+[\nabla\cdot(\mathbf{a}+\varepsilon\mathbf{d})]^2\right\}d\Omega-2\int\limits_{\Gamma}[\mathbf{u}_n\times(\mathbf{a}+\varepsilon\mathbf{d})]\cdot\nabla\times(\mathbf{a}+\varepsilon\mathbf{d})d\Gamma}{\int\limits_{\Omega}(\mathbf{a}+\varepsilon\mathbf{d})^2 d\Omega} \tag{1.164}$$

must satisfy $\lambda'(0)=0$, which leads to

$$\begin{aligned}
&2\int\limits_{\Omega}\mathbf{a}^2 d\Omega\int\limits_{\Omega}\nabla\times\mathbf{a}\cdot\nabla\times\mathbf{d}d\Omega-2\int\limits_{\Omega}\mathbf{a}\cdot\mathbf{d}d\Omega\int\limits_{\Omega}(\nabla\times\mathbf{a})^2 d\Omega\\
&\quad+2\int\limits_{\Omega}\mathbf{a}^2 d\Omega\int\limits_{\Omega}\nabla\cdot\mathbf{a}\nabla\cdot\mathbf{d}d\Omega-2\int\limits_{\Omega}\mathbf{a}\cdot\mathbf{d}d\Omega\int\limits_{\Omega}(\nabla\cdot\mathbf{a})^2 d\Omega\\
&\quad-\int\limits_{\Omega}\mathbf{a}^2 d\Omega\left[2\int\limits_{\Gamma}(\mathbf{u}_n\times\mathbf{a})\cdot\nabla\times\mathbf{d})d\Gamma+2\int\limits_{\Gamma}(\mathbf{u}_n\times\mathbf{d})\cdot\nabla\times\mathbf{a}d\Gamma\right]\\
&\quad+4\int\limits_{\Omega}\mathbf{a}\cdot\mathbf{d}d\Omega\int\limits_{\Gamma}(\mathbf{u}_n\times\mathbf{a})\cdot\nabla\times\mathbf{a}d\Gamma=0.
\end{aligned}$$

By means of the vector identities (1.148), one may find

$$\begin{aligned}
&\int\limits_{\Omega}\mathbf{d}\cdot\nabla\times\nabla\times\mathbf{a}d\Omega-\int\limits_{\Omega}\mathbf{d}\cdot\nabla\nabla\cdot\mathbf{a}d\Omega\\
&\quad-\int\limits_{\Omega}\mathbf{d}\cdot\mathbf{a}d\Omega\frac{\int\limits_{\Omega}(\nabla\times\mathbf{a})^2 d\Omega+\int\limits_{\Omega}(\nabla\cdot\mathbf{a})^2 d\Omega-2\int\limits_{\Gamma}\mathbf{u}_n\times\mathbf{a}\cdot\nabla\times\mathbf{a}d\Gamma}{\int\limits_{\Omega}\mathbf{a}^2 d\Omega}\\
&\quad+\int\limits_{\Gamma}\nabla\cdot\mathbf{a}(\mathbf{u}_n\cdot\mathbf{d})d\Gamma-\int\limits_{\Gamma}(\mathbf{u}_n\times\mathbf{a})\cdot\nabla\times\mathbf{d}d\Gamma=0.
\end{aligned}$$

This can be simplified to

$$\begin{aligned}&\int_\Omega \mathbf{d}\cdot[\nabla\times\nabla\times\mathbf{a}-\nabla\nabla\cdot\mathbf{a}-\lambda\mathbf{a}]d\Omega+\int_\Gamma \nabla\cdot\mathbf{a}(\mathbf{u}_n\cdot\mathbf{d})d\Gamma\\&\quad-\int_\Gamma(\mathbf{u}_n\times\mathbf{a})\cdot\nabla\times\mathbf{d}d\Gamma=0.\end{aligned}\tag{1.165}$$

Since $\mathbf{d}$ is arbitrary, it can be first chosen arbitrarily inside Ω such that $\mathbf{u}_n\cdot\ \mathbf{d}=\mathbf{u}_n\times\nabla\times\mathbf{d}=0$ on the boundary Γ. From Eq. (1.165), the following equation can be obtained:

$$\nabla\times\nabla\times\mathbf{a}-\nabla\nabla\cdot\mathbf{a}-\lambda\mathbf{a}=0,\boldsymbol{\rho}\in\Omega.$$

Thus,

$$\int_\Gamma \nabla\cdot\mathbf{a}(\mathbf{u}_n\cdot\mathbf{d})d\Gamma-\int_\Gamma(\mathbf{u}_n\times\mathbf{a})\cdot\nabla\times\mathbf{d}d\Gamma=0,\tag{1.166}$$

which is valid for an arbitrary $\mathbf{d}$. One can first choose $\mathbf{u}_n\ \cdot\ \mathbf{d}=0$ in the above equation so that

$$\int_\Gamma(\mathbf{u}_n\times\mathbf{a})\cdot\nabla\times\mathbf{d}d\Gamma=0.\tag{1.167}$$

From the expansion for curl of $\mathbf{d}$

$$\begin{aligned}\nabla\times\mathbf{d}&=\left(\frac{\partial}{\partial n}\mathbf{u}_n+\frac{\partial}{\partial t}\mathbf{u}_t\right)\times[(\mathbf{d}\cdot\mathbf{u}_n)\mathbf{u}_n+(\mathbf{d}\cdot\mathbf{u}_t)\mathbf{u}_t]\\&=-\mathbf{u}_z\frac{\partial(\mathbf{d}\cdot\mathbf{u}_t)}{\partial n}+\mathbf{u}_z\frac{\partial(\mathbf{d}\cdot\mathbf{u}_n)}{\partial t},\end{aligned}$$

it is easy to see that $\nabla\times\mathbf{d}$ can still be selected arbitrarily even though $\mathbf{u}_n\ \cdot\ \mathbf{d}=0$ has been assumed. Hence, (1.167) implies

$$\mathbf{u}_n\times\mathbf{a}=0,\boldsymbol{\rho}\in\Gamma.$$

One can also first choose $\nabla\times\mathbf{d}=0$ in (1.166) so that

$$\int_\Gamma \nabla\cdot\mathbf{a}(\mathbf{u}_n\cdot\mathbf{d})d\Gamma=0.\tag{1.168}$$

For $\mathbf{u}_n\ \cdot\ \mathbf{d}$ can still be chosen arbitrarily, one may find

$$\nabla\cdot\mathbf{a}=0,\boldsymbol{\rho}\in\Gamma.$$

Therefore, the vector field $\mathbf{a}$ is an eigenfunction that satisfies (1.139). In a similar way, it can be shown that the solution of (1.163) is an eigenfunction that satisfies (1.140). □

1.5.3 Classification of Vector Modal Functions

The vector modal function $\mathbf{a}_n$ belongs to one of the following four types:

1. $\nabla \times \mathbf{a}_n = 0, \nabla \cdot \mathbf{a}_n = 0,$
2. $\nabla \times \mathbf{a}_n \neq 0, \nabla \cdot \mathbf{a}_n = 0,$
3. $\nabla \times \mathbf{a}_n = 0, \nabla \cdot \mathbf{a}_n \neq 0,$
4. $\nabla \times \mathbf{a}_n \neq 0, \nabla \cdot \mathbf{a}_n \neq 0.$

A complete set of eigenfunctions can be constructed from the first three types [32], which is implied by the Helmholtz theorem to be discussed later. Indeed, if $\mathbf{a}_n$ belongs to type 4, two new functions may be introduced through

$$\mathbf{a}'_n = c'\nabla \times \nabla \times \mathbf{a}_n, \quad \mathbf{a}''_n = c''\nabla\nabla \cdot \mathbf{a}_n, \tag{1.169}$$

where c' and c'' are two normalizing constants. It can be verified that both $\mathbf{a}'_n$ and $\mathbf{a}''_n$ are eigenfunctions that are mutually orthogonal and satisfy (1.139). Apparently, $\mathbf{a}'_n$ and $\mathbf{a}''_n$, respectively, belong to types 2 and 3. From (1.139) and (1.169), one may find

$$\mathbf{a}_n = \frac{1}{\lambda_n}\left(\frac{1}{c'}\mathbf{a}'_n - \frac{1}{c''}\mathbf{a}''_n\right).$$

As a result, the eigenfunctions belonging to type 4 can be expressed as a linear combination of the eigenfunctions in the first three types. A complete set of eigenfunctions can be established as follows. Suppose that the nth eigenfunction $\mathbf{a}_n$ happens to fall into type 4 during the minimization process with (1.144). Instead of using $\mathbf{e}_n$ as the nth eigenfunction, one may take $\mathbf{a}'_n$ and $\mathbf{a}''_n$ as the nth and $(n+1)$th eigenfunctions, respectively. This process guarantees that all the eigenfunctions obtained with (1.144) fall into the first three types. Based on the above analysis, an arbitrary vector field $\mathbf{F}$ can be expressed as a linear combination of the eigenfunctions in the first three types and therefore can be split up into three components

$$\mathbf{F} = \mathbf{F}_L + \mathbf{F}_T + \mathbf{F}_H, \tag{1.170}$$

where

$$\begin{cases} \nabla \cdot \mathbf{F}_L = \nabla \cdot \mathbf{F} \\ \nabla \times \mathbf{F}_L = 0 \end{cases}, \begin{cases} \nabla \cdot \mathbf{F}_T = 0 \\ \nabla \times \mathbf{F}_L = \nabla \times \mathbf{F} \end{cases}, \begin{cases} \nabla \cdot \mathbf{F}_H = 0 \\ \nabla \times \mathbf{F}_H = 0 \end{cases}. \tag{1.171}$$

The components $\mathbf{F}_L$, $\mathbf{F}_T$, and $\mathbf{F}_H$ are, respectively, called **longitudinal**, **transverse**, and **harmonic**. The decomposition is very useful if $\mathbf{F}$ is unknown but $\nabla \times \mathbf{F}$ and $\nabla \cdot \mathbf{F}$ are specified. Such a decomposition is usually referred to as Helmholtz theorem, which will be discussed in Section 1.7.

It is noted that all the results obtained for eigenvalue problems (1.139) and (1.140) in 2D space are also valid in 3D space.

1.6 Ritz Method for the Solution of Eigenvalue Problem

For the solution of the eigenvalue problem

$$\hat{L}u = \lambda u, \tag{1.172}$$

where $\hat{L}$ is a **positive-bounded-below operator** $(\hat{L}u, u) \geq c\|u\|^2$ with c being a constant, one may resort to seeking the solution of the minimization problem of the Rayleigh quotient for the operator $\hat{L}$:

$$\lambda = \min \frac{(\hat{L}u, u)}{(u, u)}. \tag{1.173}$$

This is equivalent to

$$\lambda = \min_{(u,u)=1} (\hat{L}u, u). \tag{1.174}$$

The minimization problem (1.174) can be solved numerically. Let $\{u_j | j = 1, 2, \ldots, N\}$ be a set of linearly independent functions in the domain of operator $\hat{L}$, called **basis** or **trial functions**. As an approximation, the unknown function u may be expanded as follows:

$$u = \sum_{j=1}^{N} a_j u_j, \tag{1.175}$$

where a_j are the expansion coefficients to be determined. In this case, (1.174) is equivalent to

$$\begin{cases} \lambda = \min \sum_{i,j=1}^{N} (\hat{L}u_i, u_j) a_i a_j \\ \text{s.t.} \quad \sum_{i,j=1}^{N} (u_i, u_j) a_i a_j = 1. \end{cases} \tag{1.176}$$

The constrained problem (1.176) can be solved by the **method of Lagrangian multipliers**. The **Lagrangian function** for (1.176) is given by

$$L(a_i, \xi) = \sum_{i,j=1}^{N} (\hat{L}u_i, u_j) a_i a_j - \xi \sum_{i,j=1}^{N} (u_i, u_j) a_i a_j, \tag{1.177}$$

where ξ is the Lagrangian multiplier. The partial derivatives of the Lagrangian function with respect to the coefficients a_j must be zero, which results in the linear system of equations

$$\sum_{i=1}^{N} a_i \left[(\hat{L}u_i, u_j) - \xi (u_i, u_j) \right] = 0, \quad j = 1, 2, \ldots, N. \tag{1.178}$$

The existence of a nonzero solution requires that the determinant of the system (1.178) is zero

$$\begin{bmatrix} (\hat{L}u_1,u_1)-\xi(u_1,u_1) & (\hat{L}u_2,u_1)-\xi(u_2,u_1) & \cdots & (\hat{L}u_N,u_1)-\xi(u_N,u_1) \\ (\hat{L}u_1,u_2)-\xi(u_1,u_2) & (\hat{L}u_2,u_2)-\xi(u_2,u_2) & \cdots & (\hat{L}u_N,u_2)-\xi(u_N,u_2) \\ \vdots & \vdots & \ddots & \vdots \\ (\hat{L}u_1,u_N)-\xi(u_1,u_N) & (\hat{L}u_2,u_N)-\xi(u_2,u_N) & \cdots & (\hat{L}u_N,u_N)-\xi(u_N,u_N) \end{bmatrix} = 0. \tag{1.179}$$

It is easy to see that (1.179) has N roots. Substituting the roots into (1.178), the coefficients a_i can be determined up to a constant multiplier c. One can then use the constraint $(u, u) = 1$ to determine the constant c. If the set $\{u_j | j = 1, 2, ..., N\}$ is orthonormal

$$(u_i, u_j) = \delta_{ij} = \begin{cases} 1, i = j \\ 0, i \neq j \end{cases}, \tag{1.180}$$

(1.179) reduces to

$$\begin{bmatrix} (\hat{L}u_1,u_1)-\xi & (\hat{L}u_2,u_1) & \cdots & (\hat{L}u_N,u_1) \\ (\hat{L}u_1,u_2) & (\hat{L}u_2,u_2)-\xi & \cdots & (\hat{L}u_N,u_2) \\ \vdots & \vdots & \ddots & \vdots \\ (\hat{L}u_1,u_N) & (\hat{L}u_2,u_N) & \cdots & (\hat{L}u_N,u_N)-\xi \end{bmatrix} = 0. \tag{1.181}$$

The procedure described above is called **Ritz method**, named after Swiss theoretical physicist Walther Heinrich Wilhelm Ritz (1878–1909) [33].

Example 1.5 Let $\hat{L} = -\left(d^2/dx^2\right)$. The domain of the operator $\hat{L}$ consists of the smooth functions which satisfy the boundary conditions $u(0) = u(1) = 0$. It is easy to find that the eigenvalues of $\hat{L}$ are $\lambda_n = (n\pi)^2$, $n = 1, 2, 3, ...$ The Ritz method can be used to estimate the first eigenvalue λ_1. In this case, the minimization problem (1.173) is of the form

$$\lambda = \min \frac{\int_0^1 u'^2 dx}{\int_0^1 u^2 dx}.$$

Choose the trial function $u_1 = x(1-x)$ in the domain of $\hat{L}$. Substituting $u = a_1 u_1$ into the above, one may obtain $\lambda = 10$, which is good approximation to the first eigenvalue $\lambda_1 = \pi^2$. □

1.7 Helmholtz Theorems

In 1905, German mathematician Blumenthal (1876–1944) showed that every continuously differentiable vector field that vanishes at infinity can be split into an irrotational (curl free) and a solenoidal (divergence free) part. It means that a vector field **F** can be decomposed into the sum of a gradient and a curl

$$\mathbf{F}(\mathbf{r}) = -\nabla\phi(\mathbf{r}) + \nabla \times \mathbf{A}(\mathbf{r}). \tag{1.182}$$

The component $\mathbf{F}^{\parallel} = -\nabla\phi$ generated by the scalar field ϕ is irrotational (longitudinal); the component $\mathbf{F}^{\perp} = \nabla \times \mathbf{A}$ is solenoidal (transverse). Apparently, the decomposition (1.182) can be carried out in an infinite number of ways. The vector potential **A** in (1.182) is determined to within a gradient. In order for the decomposition to be unique, certain restrictions have to be placed on the vector field **F**.

1.7.1 Helmholtz Theorem for the Field in Infinite Space

By use of the identity

$$\nabla^2 \frac{1}{4\pi R} = -\delta(\mathbf{r}-\mathbf{r}'), \tag{1.183}$$

the vector field **F** defined on the infinite space can be divided into the sum of two components

$$\mathbf{F}(\mathbf{r}) = \int_{R^3} \mathbf{F}(\mathbf{r}')\delta(\mathbf{r}-\mathbf{r}')dV(\mathbf{r}') = -\nabla^2 \int_{R^3} \frac{\mathbf{F}(\mathbf{r}')}{4\pi R} dV(\mathbf{r}') = \mathbf{F}^{\parallel} + \mathbf{F}^{\perp}, \tag{1.184}$$

where

$$\begin{aligned} \mathbf{F}^{\parallel}(\mathbf{r}) &= -\nabla\nabla\cdot \int_{R^3} \frac{\mathbf{F}(\mathbf{r}')}{4\pi R} dV(\mathbf{r}'), \\ \mathbf{F}^{\perp}(\mathbf{r}) &= \nabla\times\nabla\times \int_{R^3} \frac{\mathbf{F}(\mathbf{r}')}{4\pi R} dV(\mathbf{r}') \end{aligned} \tag{1.185}$$

are referred to as the **irrotational component** and **solenoidal component** of **F**, respectively. The vector field **F** will be assumed to decrease faster than $1/r$ as $r \to \infty$. The irrotational component can further be written as

$$\mathbf{F}^{\parallel}(\mathbf{r}) = \nabla \int_{R^3} \nabla' \cdot \left[\frac{\mathbf{F}(\mathbf{r}')}{4\pi R}\right] dV(\mathbf{r}') - \nabla \int_{R^3} \frac{\nabla' \cdot \mathbf{F}(\mathbf{r}')}{4\pi R} dV(\mathbf{r}'). \tag{1.186}$$

Upon using the Gauss's theorem, the first term on the right-hand side approaches to zero so that

$$\mathbf{F}^{\parallel}(\mathbf{r}) = -\nabla \int_{R^3} \frac{\nabla' \cdot \mathbf{F}(\mathbf{r}')}{4\pi R} dV(\mathbf{r}'). \tag{1.187}$$

Similarly, the solenoidal component can be expressed as

$$\begin{aligned}\mathbf{F}^{\perp}(\mathbf{r}) &= -\nabla \times \int_{R^3} \nabla' \times \left[\frac{1}{4\pi R}\mathbf{F}(\mathbf{r}')\right] dV(\mathbf{r}') + \nabla \times \int_{R^3} \frac{\nabla' \times \mathbf{F}(\mathbf{r}')}{4\pi R} dV(\mathbf{r}') \\ &= \nabla \times \int_{R^3} \frac{\nabla' \times \mathbf{F}(\mathbf{r}')}{4\pi R} dV(\mathbf{r}'),\end{aligned} \tag{1.188}$$

after applying Gauss theorem. Hence, the following theorem is obtained.

Theorem 1.9 Any vector field **F** that decreases faster than $1/r$ as $r \to \infty$ can be expressed by

$$\mathbf{F}(\mathbf{r}) = -\nabla \int_{R^3} \frac{\nabla' \cdot \mathbf{F}(\mathbf{r}')}{4\pi R} dV(\mathbf{r}') + \nabla \times \int_{R^3} \frac{\nabla' \times \mathbf{F}(\mathbf{r}')}{4\pi R} dV(\mathbf{r}'). \tag{1.189}$$

□

This is called **Helmholtz theorem** or **Helmholtz identity**, also known as the **fundamental theorem of vector calculus**. The theorem states that a vector field that decreases rapidly at infinity is uniquely determined by its divergence and curl. An immediate consequence of Helmholtz identity is

$$\int_{R^3} |\mathbf{F}(\mathbf{r})|^2 dV(\mathbf{r}') = \int_{R^3}\int_{R^3} \frac{\nabla \cdot \mathbf{F}(\mathbf{r})\nabla' \cdot \overline{\mathbf{F}}(\mathbf{r}') + \nabla \times \mathbf{F}(\mathbf{r}) \cdot \nabla' \times \overline{\mathbf{F}}(\mathbf{r}')}{4\pi R} dV(\mathbf{r})dV(\mathbf{r}'). \tag{1.190}$$

It is noted that the Helmholtz theorem can be generalized to more complicated domains [34].

1.7.2 Helmholtz Theorem for the Field in Finite Region

For the vector field $\mathbf{F}$ defined in a finite region V bounded by S, it can also be decomposed into the sum of two components

$$\mathbf{F}(\mathbf{r}) = \int_V \mathbf{F}(\mathbf{r}')\delta(\mathbf{r}-\mathbf{r}')dV(\mathbf{r}') = -\nabla^2 \int_V \frac{\mathbf{F}(\mathbf{r}')}{4\pi R} dV(\mathbf{r}') = \mathbf{F}^{\parallel} + \mathbf{F}^{\perp}, \tag{1.191}$$

where

$$\mathbf{F}^{\parallel}(\mathbf{r}) = -\nabla\nabla\cdot \int_V \frac{\mathbf{F}(\mathbf{r}')}{4\pi R} dV(\mathbf{r}'), \mathbf{F}^{\perp}(\mathbf{r}) = \nabla\times\nabla\times \int_V \frac{\mathbf{F}(\mathbf{r}')}{4\pi R} dV(\mathbf{r}'). \tag{1.192}$$

Applying Gauss theorem, the irrotational component can further be written as

$$\begin{aligned} \mathbf{F}^{\parallel}(\mathbf{r}) &= \nabla \int_V \nabla'\cdot\left[\frac{\mathbf{F}(\mathbf{r}')}{4\pi R}\right] dV(\mathbf{r}') - \nabla \int_V \frac{\nabla'\cdot\mathbf{F}(\mathbf{r}')}{4\pi R} dV(\mathbf{r}') \\ &= -\nabla\left[\int_V \frac{\nabla'\cdot\mathbf{F}(\mathbf{r}')}{4\pi R} dV(\mathbf{r}') - \int_S \frac{\mathbf{u}_n(\mathbf{r}')\cdot\mathbf{F}(\mathbf{r}')}{4\pi R} dS(\mathbf{r}')\right], \end{aligned} \tag{1.193}$$

where $\mathbf{u}_n$ is unit outward normal of S. In the same way, the solenoidal component may be expressed by

$$\begin{aligned} \mathbf{F}^{\perp}(\mathbf{r}) &= -\nabla\times \int_V \nabla'\times\left[\frac{1}{4\pi R}\mathbf{F}(\mathbf{r}')\right] dV(\mathbf{r}') + \nabla\times \int_V \frac{\nabla'\times\mathbf{F}(\mathbf{r}')}{4\pi R} dV(\mathbf{r}') \\ &= \nabla\times\left[\int_V \frac{\nabla'\times\mathbf{F}(\mathbf{r}')}{4\pi R} dV(\mathbf{r}') - \int_S \frac{\mathbf{u}_n(\mathbf{r}')\times\mathbf{F}(\mathbf{r}')}{4\pi R} dS(\mathbf{r}')\right]. \end{aligned} \tag{1.194}$$

The vector field $\mathbf{F}$ may then be decomposed according to

$$\begin{aligned} \mathbf{F}(\mathbf{r}) &= -\nabla\left[\int_V \frac{\nabla'\cdot\mathbf{F}(\mathbf{r}')}{4\pi R} dV(\mathbf{r}') - \int_S \frac{\mathbf{u}_n(\mathbf{r}')\cdot\mathbf{F}(\mathbf{r}')}{4\pi R} dS(\mathbf{r}')\right] \\ &\quad + \nabla\times\left[\int_V \frac{\nabla'\times\mathbf{F}(\mathbf{r}')}{4\pi R} dV(\mathbf{r}') - \int_S \frac{\mathbf{u}_n(\mathbf{r}')\times\mathbf{F}(\mathbf{r}')}{4\pi R} dS(\mathbf{r}')\right]. \end{aligned} \tag{1.195}$$

Consequently, a vector field defined in a finite region is determined by its divergence and curl, as well as by its values on the boundary.

Remark 1.10 If the vector field **F** is both curl free and divergence free inside V, (1.195) reduces to

$$\mathbf{F}(\mathbf{r}) = \nabla \int_S \frac{\mathbf{u}_n(\mathbf{r}') \cdot \mathbf{F}(\mathbf{r}')}{4\pi R} dS(\mathbf{r}') - \nabla \times \int_S \frac{\mathbf{u}_n(\mathbf{r}') \times \mathbf{F}(\mathbf{r}')}{4\pi R} dS(\mathbf{r}'). \tag{1.196}$$

If S is a perfect conductor and the vector field **F** stands for the electric field, its tangential component vanishes: $\mathbf{u}_n \times \mathbf{F} = 0$. Let $\mathbf{F} = -\nabla\phi$. Since **F** is harmonic, one may write

$$\begin{cases} \nabla^2\phi = 0, \mathbf{r} \in V, \\ \mathbf{u}_n \times \nabla\phi = 0, \mathbf{r} \in S. \end{cases} \tag{1.197}$$

The boundary condition in (1.197) implies ϕ is a constant on S. As a result,

$$\int_V \nabla \cdot (\phi\nabla\phi) dV = \int_S \phi\mathbf{u}_n \cdot \nabla\phi dS = \text{const} \int_S \mathbf{u}_n \cdot \nabla\phi dS = \text{const} \int_V \nabla \cdot \nabla\phi dV = 0, \tag{1.198}$$

where use has been made of the Gauss theorem and (1.197). The left-hand side of the above equation can be written as

$$\int_V \nabla \cdot (\phi\nabla\phi) dV = \int_V (\nabla\phi \cdot \nabla\phi + \phi\nabla \cdot \nabla\phi) dV = \int_V (\nabla\phi \cdot \nabla\phi) dV.$$

This implies $\mathbf{F} = -\nabla\phi = 0$ inside V. It is noted that the derivation of (1.198) is only valid for simply connected region. If V is multiply connected, the potential function ϕ can take a different value on each boundary. As a result, a nonzero vector field **F** inside V can exist. □

1.7.3 Helmholtz Theorem for Time-Dependent Field

In order to obtain the Helmholtz theorem for a time-dependent vector field $\mathbf{F}(\mathbf{r}, t)$, one may use the retarded Green's function $G(\mathbf{r}, \mathbf{r}'; t, t')$ for the wave equation

$$\begin{cases} \left(\nabla^2 - \dfrac{1}{c^2}\dfrac{\partial^2}{\partial t^2}\right) G(\mathbf{r}, \mathbf{r}'; t, t') = -\delta(\mathbf{r} - \mathbf{r}')\delta(t - t'), \\ G(\mathbf{r}, \mathbf{r}'; t, t') = 0, t < t'. \end{cases} \tag{1.199}$$

It is easy to find the solution of the above equation

$$G(\mathbf{r}, \mathbf{r}'; t, t') = \frac{1}{4\pi R}\delta\left(t - t' - \frac{|\mathbf{r} - \mathbf{r}'|}{c}\right). \tag{1.200}$$

The time-dependent vector field $\mathbf{F}(\mathbf{r}, t)$ can be written as an integration over space and time

$$\begin{aligned}\mathbf{F}(\mathbf{r},t) &= \int_{R^3}\int_{-\infty}^{+\infty} \mathbf{F}(\mathbf{r}',t')\delta(\mathbf{r}-\mathbf{r}')\delta(t-t')dV(\mathbf{r}')dt' \\ &= -\int_{R^3}\int_{-\infty}^{+\infty} \mathbf{F}(\mathbf{r}',t')\left(\nabla^2-\frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)G(\mathbf{r},\mathbf{r}';t,t')dV(\mathbf{r}')dt'.\end{aligned}$$

Similar to the derivation of (1.189), the above equation can be expressed as

$$\begin{aligned}\mathbf{F}(\mathbf{r},t) &= -\nabla\int_{R^3}\int_{-\infty}^{+\infty} \nabla'\cdot\mathbf{F}(\mathbf{r}',t')G(\mathbf{r},\mathbf{r}';t,t')dV(\mathbf{r}')dt' \\ &+ \nabla\times\int_{R^3}\int_{-\infty}^{+\infty} \nabla'\times\mathbf{F}(\mathbf{r}',t')G(\mathbf{r},\mathbf{r}';t,t')dV(\mathbf{r}')dt' \\ &+ \frac{1}{c^2}\frac{\partial}{\partial t}\int_{R^3}\int_{-\infty}^{+\infty} \frac{\partial\mathbf{F}(\mathbf{r}',t')}{\partial t'}G(\mathbf{r},\mathbf{r}';t,t')dV(\mathbf{r}')dt'.\end{aligned} \tag{1.201}$$

Substituting (1.200) into (1.201) yields the time-domain Helmholtz theorem

$$\begin{aligned}\mathbf{F}(\mathbf{r},t) &= -\nabla\int_{R^3} \frac{\nabla'\cdot\mathbf{F}(\mathbf{r}',t-|\mathbf{r}-\mathbf{r}'|/c)}{4\pi R}dV(\mathbf{r}') \\ &+ \nabla\times\int_{R^3} \frac{\nabla'\times\mathbf{F}(\mathbf{r}',t-|\mathbf{r}-\mathbf{r}'|/c)}{4\pi R}dV(\mathbf{r}') \\ &+ \frac{1}{c^2}\frac{\partial}{\partial t}\int_{R^3} \frac{1}{4\pi R}\frac{\partial\mathbf{F}(\mathbf{r}',t-|\mathbf{r}-\mathbf{r}'|/c)}{\partial t'}dV(\mathbf{r}').\end{aligned} \tag{1.202}$$

Note that the time-domain Helmholtz theorem contains a time derivative term that may have both transverse and longitudinal components.

1.8 Curl Operator

Eigenfunctions of the curl operator are useful in expanding solenoidal vector fields and have found applications in some fields of physics. For example, in plasma physics, a magnetic field $\mathbf{B}$ is called a force-free field if $\nabla\times\mathbf{B}=\lambda\mathbf{B}$; in fluid dynamics, a velocity field $\mathbf{v}$ satisfying $\nabla\times\mathbf{v}=\lambda\mathbf{v}$ is called a Beltrami flow [35].

1.8.1 Eigenfunctions of Curl Operator

The eigenvalue problem of the curl operator has interesting applications in electromagnetics [36–38]. The eigenfunction **e** of a curl operator is defined by

$$\nabla \times \mathbf{e}(\mathbf{r}) = \lambda \mathbf{e}(\mathbf{r}), \mathbf{r} \in R^3. \tag{1.203}$$

The plane-wave solutions of Maxwell equations are widely used in field analysis for their simplicity [39]. To seek the plane-wave solution of (1.203), one may assume

$$\mathbf{e}(\mathbf{r}) = \mathbf{A} e^{j\mathbf{k}\cdot\mathbf{r}}, \tag{1.204}$$

where $\mathbf{k} = (k_x, k_y, k_z)$, and $\mathbf{A}$ is a constant vector. Introducing (1.204) into (1.203) yields

$$j\mathbf{k} \times \mathbf{A} = \lambda \mathbf{A}.$$

Explicitly this is

$$\begin{bmatrix} -\lambda & -jk_z & jk_y \\ jk_z & -\lambda & -jk_x \\ -jk_y & jk_x & -\lambda \end{bmatrix} \begin{bmatrix} A_x \\ A_y \\ A_z \end{bmatrix} = 0. \tag{1.205}$$

A nonzero solution exists if and only if the determinant of the coefficient matrix is zero

$$\det \begin{bmatrix} -\lambda & -jk_z & jk_y \\ jk_z & -\lambda & -jk_x \\ -jk_y & jk_x & -\lambda \end{bmatrix} = \lambda\left(-\lambda^2 + k^2\right) = 0,$$

where $k = |\mathbf{k}|$. Hence Eq. (1.205) has three eigenvalues $\lambda = nk (n = 0, \pm 1)$. The eigenfunctions corresponding to $\lambda = 0, k, -k$, respectively, satisfy

$$\begin{cases} k_y A_z - k_z A_y = 0 \\ k_z A_x - k_x A_z = 0 \\ k_x A_y - k_y A_x = 0 \end{cases}, \begin{cases} kA_x + jk_z A_y - jk_y A_z = 0 \\ kA_y + jk_x A_z - jk_z A_x = 0 \\ kA_z + jk_y A_x - jk_x A_y = 0 \end{cases}, \begin{cases} kA_x - jk_z A_y + jk_y A_z = 0 \\ kA_y - jk_x A_z + jk_z A_x = 0 \\ kA_z - jk_y A_x + jk_x A_y = 0 \end{cases}.$$

The orthonormal eigenfunctions can be easily found from the above equations as follows:

$$\mathbf{A}_0(\mathbf{k}) = \frac{1}{k} \begin{bmatrix} k_x \\ k_y \\ k_z \end{bmatrix}, \mathbf{A}_n(\mathbf{k}) = \frac{1}{\sqrt{2}k\left(k_x^2 + k_y^2\right)^{1/2}} \begin{bmatrix} jnkk_y - k_x k_z \\ -jnkk_x - k_y k_z \\ k_x^2 + k_y^2 \end{bmatrix}, (n = \pm 1).$$

Let $\mathbf{A}_n(\mathbf{k}) = [A_{nx}(\mathbf{k}), A_{ny}(\mathbf{k}), A_{nz}(\mathbf{k})]^T$. Then, it is readily found that

$$\begin{aligned}&\mathbf{A}_m \cdot \overline{\mathbf{A}}_n = \delta_{mn}(m, n = 0, \pm 1),\\&\sum_n A_{n\alpha}\overline{A}_{n\beta} = \delta_{\alpha\beta}(\alpha, \beta = x, y, z).\end{aligned} \tag{1.206}$$

The following orthonormal vectors may be introduced:

$$\mathbf{e}_n(\mathbf{r},\mathbf{k}) = \left[e_{nx}(\mathbf{r},\mathbf{k}), e_{ny}(\mathbf{r},\mathbf{k}), e_{nz}(\mathbf{r},\mathbf{k})\right]^T = \frac{1}{(2\pi)^{3/2}}\mathbf{A}_n(\mathbf{k})e^{j\mathbf{k}\cdot\mathbf{r}},$$

which satisfy the orthonormal conditions

$$\begin{aligned}&\int_{R^3}\mathbf{e}_m(\mathbf{r},\mathbf{k})\cdot\overline{\mathbf{e}}_n(\mathbf{r},\mathbf{k}')dxdydz = \delta_{mn}\delta(\mathbf{k}-\mathbf{k}'), (m, n = 0, \pm 1),\\&\sum_n\int_{R^3}e_{n\alpha}(\mathbf{r},\mathbf{k})\overline{e}_{n\beta}(\mathbf{r}',\mathbf{k})dk_xdk_ydk_z = \delta_{\alpha\beta}\delta(\mathbf{r}-\mathbf{r}'), (\alpha, \beta = x, y, z),\end{aligned} \tag{1.207}$$

and

$$\begin{aligned}&\nabla\times\mathbf{e}_n = nk\mathbf{e}_n,\\&\nabla\cdot\mathbf{e}_n = 0\ (n = \pm 1),\\&\nabla\cdot\mathbf{e}_0 = \frac{jk}{(2\pi)^{3/2}}e^{j\mathbf{k}\cdot\mathbf{r}}.\end{aligned} \tag{1.208}$$

The second equation of (1.207) can be expressed in a dyadic form as

$$\sum_n\int_{R^3}\mathbf{e}_n(\mathbf{r},\mathbf{k})\overline{\mathbf{e}}_n(\mathbf{r}',\mathbf{k})dk_xdk_ydk_z = \overleftrightarrow{\mathbf{I}}\delta(\mathbf{r}-\mathbf{r}'). \tag{1.209}$$

In terms of (1.209), an arbitrary vector $\mathbf{F}$ can then be expanded as follows:

$$\mathbf{F}(\mathbf{r}) = \sum_n\int_{R^3}\mathbf{e}_n(\mathbf{r},\mathbf{k})f_n(\mathbf{k})dk_xdk_ydk_z = \sum_n\mathbf{F}_n(\mathbf{r}), \tag{1.210}$$

where

$$\begin{aligned}&\mathbf{F}_n(\mathbf{r}) = \int_{R^3}\mathbf{e}_n(\mathbf{r},\mathbf{k})f_n(\mathbf{k})dk_xdk_ydk_z,\\&f_n(\mathbf{k}) = \int_{R^3}\overline{\mathbf{e}}_n(\mathbf{r},\mathbf{k})\cdot\mathbf{F}(\mathbf{r})dxdydz.\end{aligned} \tag{1.211}$$

Since $\nabla \times \mathbf{F}_0 = 0$ and $\nabla \cdot \mathbf{F}_n = 0$ ($n = \pm 1$), an arbitrary vector $\mathbf{F}$ may be decomposed into three components: one is irrotational and the other two are solenoidal. This result may be regarded as the **generalized Helmholtz theorem**. By means of (1.208), one may write

$$\mathbf{F}_0(\mathbf{r}) = \nabla\phi(\mathbf{r}), \quad \mathbf{F}_n(\mathbf{r}) = \nabla \times \mathbf{A}_n(\mathbf{r}), n = \pm 1,$$

where

$$\begin{aligned}
\phi(\mathbf{r}) &= \frac{-j}{(2\pi)^{3/2}} \int\limits_{R^3} \frac{1}{k} e^{j\mathbf{k}\cdot\mathbf{r}} f_0(\mathbf{k}) dk_x dk_y dk_z, \\
\mathbf{A}_n(\mathbf{r}) &= \frac{1}{n} \int\limits_{R^3} \frac{1}{k} \mathbf{e}_n(\mathbf{r}, \mathbf{k}) f_n(\mathbf{k}) dk_x dk_y dk_z.
\end{aligned}$$

As a result, (1.210) can be expressed by

$$\mathbf{F} = \nabla\phi + \nabla \times \mathbf{A}_{-1} + \nabla \times \mathbf{A}_1. \tag{1.212}$$

1.8.2 Plane-Wave Expansions for the Fields and Dyadic Green's Functions

Consider the solution of the generalized Maxwell equations

$$\begin{aligned}
\nabla \times \mathbf{H}(\mathbf{r}) &= \mathbf{J}(\mathbf{r}) + j\omega\varepsilon\mathbf{E}(\mathbf{r}), \\
\nabla \times \mathbf{E}(\mathbf{r}) &= -j\omega\mu\mathbf{H}(\mathbf{r}) - \mathbf{J}_m(\mathbf{r}).
\end{aligned} \tag{1.213}$$

In terms of (1.210), the fields and the sources have the following expansions:

$$\begin{aligned}
\mathbf{E}(\mathbf{r}) &= \sum_n \int\limits_{R^3} \mathbf{e}_n(\mathbf{r}, \mathbf{k}) e_n(\mathbf{k}) dk_x dk_y dk_z, \\
\mathbf{H}(\mathbf{r}) &= \sum_n \int\limits_{R^3} \mathbf{e}_n(\mathbf{r}, \mathbf{k}) h_n(\mathbf{k}) dk_x dk_y dk_z, \\
\mathbf{J}(\mathbf{r}) &= \sum_n \int\limits_{R^3} \mathbf{e}_n(\mathbf{r}, \mathbf{k}) j_n(\mathbf{k}) dk_x dk_y dk_z, \\
\mathbf{J}_m(\mathbf{r}) &= \sum_n \int\limits_{R^3} \mathbf{e}_n(\mathbf{r}, \mathbf{k}) j_{m,n}(\mathbf{k}) dk_x dk_y dk_z.
\end{aligned} \tag{1.214}$$

Upon substitution of the field and source expansions into (1.213), the expansion coefficients for the fields may be found as follows:

$$\begin{aligned} e_n(\mathbf{k}) &= -\frac{j\omega\mu j_n(\mathbf{k}) + nkj_{m,n}(\mathbf{k})}{n^2k^2 - k_0^2}, \\ h_n(\mathbf{k}) &= \frac{nkj_n(\mathbf{k}) - j\omega\varepsilon j_{m,n}(\mathbf{k})}{n^2k^2 - k_0^2}, \end{aligned} \tag{1.215}$$

where $k_0 = \omega\sqrt{\mu\varepsilon}$. The field expansions in (1.214) can be written in a compact form as

$$\begin{aligned} \mathbf{E}(\mathbf{r}) &= -j\omega\mu \int\limits_{R^3} \overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}') \cdot \mathbf{J}(\mathbf{r}')dx'dy'dz' - \int\limits_{R^3} \overleftrightarrow{\mathbf{G}}_m(\mathbf{r},\mathbf{r}') \cdot \mathbf{J}_m(\mathbf{r}')dx'dy'dz', \\ \mathbf{H}(\mathbf{r}) &= -j\omega\varepsilon \int\limits_{R^3} \overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}') \cdot \mathbf{J}_m(\mathbf{r}')dx'dy'dz' + \int\limits_{R^3} \overleftrightarrow{\mathbf{G}}_m(\mathbf{r},\mathbf{r}') \cdot \mathbf{J}(\mathbf{r}')dx'dy'dz', \end{aligned} \tag{1.216}$$

where $\overleftrightarrow{\mathbf{G}}_e$ and $\overleftrightarrow{\mathbf{G}}_e$ are the electric and magnetic dyadic Green's functions, defined by

$$\begin{aligned} \overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}') &= \sum_n \int\limits_{R^3} \frac{\mathbf{e}_n(\mathbf{r},\mathbf{k})\overline{\mathbf{e}}_n(\mathbf{r}',\mathbf{k})}{n^2k^2 - k_0^2} dk_x dk_y dk_z, \\ \overleftrightarrow{\mathbf{G}}_m(\mathbf{r},\mathbf{r}') &= \sum_n \int\limits_{R^3} \frac{nk\mathbf{e}_n(\mathbf{r},\mathbf{k})\overline{\mathbf{e}}_n(\mathbf{r}',\mathbf{k})}{n^2k^2 - k_0^2} dk_x dk_y dk_z. \end{aligned} \tag{1.217}$$

Ignoring the tedious process, the electric and magnetic dyadic Green's functions can be rewritten as

$$\begin{aligned} \overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}') &= \frac{1}{(2\pi)^3} \int\limits_{R^3} \frac{k_0^2\overleftrightarrow{\mathbf{I}} - \mathbf{k}\mathbf{k}}{k_0^2(k^2 - k_0^2)} e^{j\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')} dk_x dk_y dk_z, \\ \overleftrightarrow{\mathbf{G}}_m(\mathbf{r},\mathbf{r}') &= \frac{1}{(2\pi)^3} \int\limits_{R^3} \frac{\mathbf{k}\times\overleftrightarrow{\mathbf{I}}}{k^2 - k_0^2} j e^{j\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')} dk_x dk_y dk_z. \end{aligned} \tag{1.218}$$

These are the **plane-wave expansions** for the dyadic Green's functions. They can also be derived from the Fourier transform with respect to the position vector $\mathbf{r}$. It is noted that the expressions in (1.218) should be taken as a symbolic equality and are meaningful only when these expressions are used as the kernel of an integral operator because they may contain generalized functions [40]. Apparently,

$$\nabla \times \overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}') = \overleftrightarrow{\mathbf{G}}_m(\mathbf{r},\mathbf{r}'). \tag{1.219}$$

It follows from (1.7), (1.216), and (1.219) that the electric and magnetic dyadic Green's functions, respectively, satisfy

$$\begin{aligned}
\nabla \times \nabla \times \overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}') - k^2\overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}') &= \overleftrightarrow{\mathbf{I}}\delta(\mathbf{r}-\mathbf{r}'),\\
\nabla \times \nabla \times \overleftrightarrow{\mathbf{G}}_m(\mathbf{r},\mathbf{r}') - k^2\overleftrightarrow{\mathbf{G}}_m(\mathbf{r},\mathbf{r}') &= \nabla \times \left[\overleftrightarrow{\mathbf{I}}\delta(\mathbf{r}-\mathbf{r}')\right].
\end{aligned} \tag{1.220}$$

> The growing importance of eigenvalue theory in pure and applied mathematics, and in physics and chemistry, has drawn attention to various methods for approximate calculation of eigenvalues. Clearly it is important to develop these methods in a general and theoretical manner, if only because opportunities for particular application may otherwise be inadvertently missed.
>
> *–Sydney Henry Gould (Mathematician, 1909–1986)*

References

1 Baum, C. E., "On the singularity expansion method for the solution of electromagnetic interaction problems", *AFWL*, Interaction Notes, Vol. 88, 1971.

2 Baum, C. E., "The singularity expansion method", in *Transient Electromagnetic Fields*, edited by L. B. Felsen, New York, NY, USA, Springer-Verlag, pp. 129–179, 1976.

3 Ramm, A. G., "Mathematical foundations of the singularity and eigenmode expansion methods (SEM and EEM)", *J. Math. Anal. Appl.*, Vol. 86, pp. 562–591, 1982.

4 Marks, R. B., "The singular function expansion in time-dependent scattering", *IEEE Trans. Antennas Propagat.*, Vol. 37, pp. 1559–1565, 1989.

5 Cochran, J. A., *The Analysis of Linear Integral Equations*, New York, McGraw-Hill, 1972.

6 Inagaki, N., "Eigenfunctions of composite Hermitian operators with application to discrete and continuous radiating systems", *IEEE Trans. Antennas Propag.*, Vol. 30, No. 7, pp. 571–575, 1982.

7 Pozar, D. M., "Antenna synthesis and optimization using weighted Inagaki modes", *IEEE Trans. Antennas Propagat.*, Vol. 32, pp. 159–165, 1984.

8 Garbacz, R. J., "Modal expansions for resonance scattering phenomena", *Proc. IEEE*, Vol. 53, No. 8, pp. 856–864, 1965.

9 Garbacz, R. J. and R. Turpin, "A generalized expansion for radiated and scattered fields", *IEEE Trans. Antennas Propag.*, Vol. 19, No. 3, pp. 348–358, 1971.

10 Harrington, R. F. and J. R. Mautz, "Theory of characteristic modes for conducting bodies", *IEEE Trans. Antennas Propag.*, Vol. 19, No. 5, pp. 622–628, 1971.

11 Harrington, R. F., J. R. Mautz, and Y. Chang, “Characteristic modes for dielectric and magnetic bodies”, *IEEE Trans. Antennas Propag.*, Vol. 20, No. 2, pp. 194–198, 1972.

12 Gould, S. H., *Variational Methods for Eigenvalue Problems: An Introduction to the Methods of Rayleigh, Ritz, Weinstein, and Aronszajn*, Incorporated, Dover Publications, 1995.

13 Popvić, B. D., “Electromagnetic field theorems”, *IEE Proc.*, Vol. 128, pp. 47–63, 1981.

14 Harrington, R. F., *Time-Harmonic Electromagnetic Fields*, McGraw-Hill Book Company, Inc., 1961.

15 Balanis, C. A., *Advanced Engineering Electromagnetics*, 2nd Ed., John Wiley & Sons, 2012.

16 Geyi, W., *Foundations of Applied Electrodynamics*, New York, Wiley, 2010.

17 Geyi, W., “Stored electromagnetic field energies in general materials”, *J. Opt. Soc. Am. B*, Vol. 36, No. 4, pp. 917–925, 2019.

18 Geyi, W., “A method for the evaluation of small antenna Q”, *IEEE Trans. Antennas and Propagat.*, Vol. 51, pp. 2124–2129, 2003.

19 Courant, R. and D. Hilbert, *Methods of Mathematical Physics*, Vol. 1–2, John Wiley & Sons, 1953.

20 Morse, P. M. and H. Feshbach, *Methods of Theoretical Physics*, McGraw-Hill, 1953.

21 Gustafson, K. E., *Partial Differential Equations and Hilbert Space Methods*, John Wiley & Sons, 1987.

22 Eisenhart, L. P., “Separable systems of Stäckel”, *Ann. Math.*, Vol. 35, No. 2, pp. 284–305, 1934.

23 Levine, H. and J. Schwinger, “On the theory of electromagnetic wave diffraction by an aperture in an infinite plane conducting screen”, *Commun. Pure Appl. Math.*, Vol. 3, No. 4, pp. 355–391, 1950.

24 Tai, C.-T., *Dyadic Green Functions in Electromagnetic Theory*, IEEE Press, 1994.

25 Bladel, J. V., “Some remarks on Green’s functions for infinite spaces”, *IRE Trans. Antennas Propagat*, Vol. 9, pp. 563–566, 1961.

26 Yaghjian, A. D., “Electric dyadic Green’s functions in the source region”, *Proc. IEEE*, Vol. 68, No. 2, pp. 248–263, 1980.

27 Chew, W. C., *Waves and Fields in Inhomogeneous Media*, Van Nostrand Reinhold, 1990.

28 Horn, R. A. and C. R. Johnson, *Matrix Analysis*, 2nd Ed., Cambridge University Press, 2013.

29 Strauss, W. A., *Partial Differential Equations-An Introduction*, John Wiley & Sons, Inc., 1992.

30 Friedman, B., *Principles and Techniques of Applied Mathematics*, John Wiley & Sons, Inc., 1956.

31 Zeidler, E., *Applied Functional Analysis-Applications to Mathematical Physics*, Springer-Verlag, 1995.

32 Kurokawa, K., *An Introduction to Microwave Circuits*, New York, Academic Press, 1969.

33 Mikhlin, S. G., *Variational Methods in Mathematical Physics*, Oxford, Pergamon Press, 1964.

34 Bladel, J. V., "On Helmholtz's theorem in multiply-bounded and multiply-connected regions", *J. Frankl. Inst.*, Vol. 269, No. 6, pp. 445–462, 1960.

35 Yoshida, Z., "Eigenfunction expansions associated with the curl derivatives in cylindrical geometries: completeness of Chandrasekhar-Kendall eigenfunctions", *J. Math. Phys.*, Vol. 33, No. 4, pp. 1252–1256, 1992.

36 Good, R. H., "Particle aspect of the electromagnetic field equations", *Phys. Rev.*, Vol. 105, No. 6, pp. 1914–1919, 1957.

37 Moses, H. E., "Solutions of Maxwell equations in terms of a spinor notation: the direct and inverse problem", *Phys. Rev.*, Vol. 113, No. 6, pp. 1670–1679, 1959.

38 Moses, H. E. and R. T. Prosser, "Initial conditions sources, and currents for prescribed time-dependent acoustic and electromagnetic fields in three dimensions, part 1: the inverse initial value problem. Acoustic and electromagnetic bullets, expanding waves, and imploding waves", *IEEE Trans. Antennas and Propagat.*, Vol. 34, pp. 188–196, 1986.

39 Clemmow, P. C., *The Plane Wave Spectrum Representation of Electromagnetic Fields*, IEEE Press, 1996.

40 Chew, W. C., "Some observations on the spatial and eigenfunction representations of dyadic Green's functions", *IEEE Trans. Antennas Propag.*, Vol. 37, pp. 1322–1327, 1989.