Shu Hotta

# Mathematical Physical Chemistry

## Practical and Intuitive Methodology

Springer

Mathematical Physical Chemistry

Shu Hotta

# Mathematical Physical Chemistry

## Practical and Intuitive Methodology

Springer

Shu Hotta
Faculty of Materials Science and
Engineering
Kyoto Institute of Technology
Kyoto
Japan

ISBN 978-981-10-7670-1 ISBN 978-981-10-7671-8 (eBook)
https://doi.org/10.1007/978-981-10-7671-8

Library of Congress Control Number: 2017961496



Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

*To my wife Kazue*

# Preface

The contents of this book are based upon manuscripts prepared for both undergraduate courses of Kyoto Institute of Technology by the author entitled "Polymer Nanomaterials Engineering" and "Photonics Physical Chemistry" and a master's course lecture of Kyoto Institute of Technology by the author entitled "Solid-State Polymers Engineering."

This book is intended for graduate and undergraduate students, especially those who major in chemistry and, at the same time, wish to study mathematical physics. Readers are supposed to have basic knowledge of analysis and linear algebra. However, they are not supposed to be familiar with the theory of analytic functions (i.e., complex analysis), even though it is desirable to have relevant knowledge about it.

At the beginning, mathematical physics looks daunting to chemists, as used to be the case with myself as a chemist. The book introduces basic concepts of mathematical physics to chemists. Unlike other books related to mathematical physics, this book makes a reasonable selection of material so that students majoring in chemistry can readily understand the contents in spontaneity. In particular, we stress the importance of practical and intuitive methodology. We also expect engineers and physicists to benefit from reading this book.

In Part I and Part II, the book describes quantum mechanics and electromagnetism. Relevance between the two is well considered. Although quantum mechanics covers broad field of modern physics, in Part I we focus on a harmonic oscillator and a hydrogen (like) atom. This is because we can study and deal with many of fundamental concepts of quantum mechanics within these restricted topics. Moreover, knowledge acquired from the study of the topics can readily be extended to practical investigation of, e.g., electronic sates and vibration (or vibronic) states of molecular systems. We describe these topics both by analytic method (that uses differential equations) and operator approach (using matrix calculations). We believe that the basic concepts of quantum mechanics can be best understood by contrasting the analytical and algebraic approaches. For this reason, we give matrix representations of physical quantities whenever possible. Examples include energy eigenvalues of a quantum-mechanical harmonic oscillator and angular momenta of

a hydrogen-like atom. At the same time, these two physical systems supply us with a good opportunity to study classical polynomials, e.g., Hermite polynomials, (associated) Legendre polynomials, Laguerre polynomials, Gegenbauer polynomials, and special functions, more generally. These topics constitute one of important branches of mathematical physics. One of the basic concepts of the quantum mechanics is that a physical quantity is represented by an Hermitian operator or matrix. In this respect, the algebraic approach gives a good opportunity to get familiar with this concept. We present tangible examples for this. We also emphasize the importance of notion of Hermiticity of a differential operator. We often encounter unitary operator or unitary transformation alongside of the notion of Hermitian operator. We show several examples of the unitary operators in connection with transformation of vectors and coordinates.

Part II describes Maxwell's equations and their applications to various phenomena of electromagnetic waves. These include their propagation, reflection, and transmission in dielectric media. We restrict ourselves to treating those phenomena in dielectrics without charge. Yet, we cover a wide range of important topics. In particular, when two (or more) dielectrics are in contact with each other at a plane interface, reflection and transmission of light are characterized by various important parameters such as reflection and transmission coefficients, Brewster angles, and critical angles. We should have a proper understanding not only from the point of view of basic study, but also to make use of relevant knowledge in optical device applications such as a waveguide. In contrast to a concept of electromagnetic waves, light possesses a characteristic of light quanta. We present semiclassical and statistical approach to blackbody radiation occurring in a simplified system in relation to Part I. The physical processes are well characterized by a notion of two-level atoms. In this context, we outline the dipole radiation within the framework of the classical theory. We briefly describe how the optical processes occurring in a confined dielectric medium are related to a laser that is of great importance in fundamental science and its applications. Many of basic equations of physics are descried as second-order linear differential equations (SOLDEs). Different methods were developed and proposed to seek their solutions. One of the most important methods is that of Green's functions. We present introductory theory of the Green's functions accordingly. In this connection, we rethink the Hermiticity of a differential operator.

In Par III and Part IV, we describe algebraic structures of mathematical physics. Their understanding is useful to studies of quantum mechanics and electromagnetism whose topics are presented in Part I and Part II. Part III deals with theories of linear vector spaces. We focus on the discussion on vectors and their transformations in finite-dimensional vector spaces. Generally, we consider the vector transformations among the vector spaces of different dimensions. In this book, however, we restrict ourselves to the case of the transformation between the vector spaces of same dimension, i.e., endomorphism of the space $(V^n \to V^n)$. This is not only because this is most often the case with many of physical applications, but because the relevant operator is represented by a square matrix. Canonical forms of square matrices hold an important position in algebra. These include a triangle matrix,

diagonalizable matrix as well as a nilpotent matrix and idempotent matrix. The most general form will be Jordan canonical form. We present its essential parts in detail taking a tangible example. Next to the general discussion, we deal with an inner product space. Once an inner product is defined between any couple of vectors, the vector space is given a fruitful structure. An example is a norm (i.e., "length") of a vector. Also we gain a clear relationship between Part III and Part I. We define various operators or matrices that are important in physical applications. Examples include normal operators (or matrices) such as Hermitian operators, projection operators, and unitary operators. Once again, we emphasize the importance of the Hermitian operators. In particular, two commutable Hermitian matrices share simultaneous eigenvectors (or eigenstates) and, in this respect, such two matrices occupy a special position in quantum mechanics.

Finally, Part IV describes the essence of group theory and its chemical applications. Group theory has a broad range of applications in solid-state physics, solid-state chemistry, molecular science, etc. Nonetheless, the knowledge of group theory does not seem to have fully prevailed among chemists. We can discover an adequate reason for this in a preface to the first edition of "*Chemical Applications of Group Theory*" written by F. A. Cotton. It might well be natural that definition and statement of abstract algebra, especially group theory, sound somewhat pretentious for chemists, even though the definition of group is quite simple. Therefore, we present various examples for readers to get used to notions of group theory. Notion of mapping is important as in the case of the linear vector spaces. Aside from being additive with calculation for a vector space and multiplicative for a group, the fundamentals of calculation regulations are pretty much the same regarding the vector space and group. We describe characteristics of symmetry groups in detail partly because related knowledge is useful for molecular orbital (MO) calculations that are presented in the last Section of the book. Representation theory is probably one of the most daunting notions for chemists. Practically, however, the representation is just homomorphism that corresponds to a linear transformation in a vector space. In this context, the representation is merely denoted by a number or a matrix. Basis functions of representation correspond to basis vectors in a vector space. Grand orthogonality theorem (GOT) is a "nursery bed" of the representation theory. Therefore, readers are encouraged to understand its essence apart from the rigorous proof of the theorem. In conjunction with Part III, we present a variety of projection operators. These are very useful to practical applications in, e.g., quantum mechanics and molecular science. The final parts of the book are devoted to applications of group theory to problems of physical chemistry, especially those of quantum chemistry, more specifically molecular orbital calculations. We see how symmetry consideration, particularly use of projection operators, saves us a lot of labor. Examples include aromatic hydrocarbons and methane.

The above is the constitution of this book. Readers may start with any Part and go freely back and forth. This is because contents of many Sections are interrelated. For example, we stress the importance of Hermiticity of differential operators and matrices. Also projection operators and nilpotent matrices appear in many Sections along with their tangible applications to individual topics. Hence, readers are

recommended to carefully examine and compare the related contents throughout the book. We believe that readers, especially chemists, benefit from a writing style of this book, since it is suited to chemists who are good at intuitive understanding.

The author would like to thank many students for their valuable suggestions and discussions at the lectures. The author also wishes to thank Dr. Shin'ichi Koizumi, Springer for giving him an opportunity to write this book.

Kyoto, Japan Shu Hotta
October 2017

# Contents

# Part I
# Quantum Mechanics

Quantum mechanics is clearly distinguished from classical physics whose major pillars are Newtonian mechanics and electromagnetism established by Maxwell. Quantum mechanics was first established as a theory of atomic physics that handled microscopic world. Later on, quantum mechanics was applied to macroscopic world, i.e., cosmos. A question on how exactly quantum mechanics describes the natural world and on how far the theory can go remains yet problematic and is in dispute to this day.

Such an ultimate question is irrelevant to this monograph. Our major aim is to study a standard approach to applying Schrödinger equation to selected topics. The topics include a particle confined within a potential well, a harmonic oscillator, and a hydrogen-like atoms. Our major task rests on solving eigenvalue problems of these topics. To this end, we describe both an analytical method and algebraic (or operator) method. Focusing on these topics, we will be able to acquire various methods to tackle a wide range of quantum-mechanical problems. These problems are usually posed as an analytical equation (i.e., differential equation) or an algebraic equation. A Hamiltonian is constructed analytically or algebraically accordingly. Besides Hamiltonian, physical quantities are expressed as a differential operator or a matrix operator. In both analytical and algebraic approaches, Hermitian property (or Hermiticity) of an operator and matrix is of crucial importance. This feature will, therefore, be highlighted not only in this part but also throughout this book along with a unitary operator and matrix.

Optical transition and associated selection rules are dealt with in relation to the above topics. Those subjects are closely related to electromagnetic phenomena that are considered in Part II.

# Chapter 1
# Schrödinger Equation and Its Application

Quantum mechanics is an indispensable research tool of modern natural science that covers cosmology, atomic physics, molecular science, materials science, and so forth. The basic concept underlying quantum mechanics rests upon Schrödinger equation. The Schrödinger equation is described as a second-order linear differential equation (SOLDE). The equation is analytically solved accordingly. Alternatively, equations of the quantum mechanics are often described in terms of operators and matrices, and physical quantities are represented by those operators and matrices. Normally, they are non-commutative. In particular, the quantum-mechanical formalism requires the canonical commutation relation between position and momentum operators. One of the great characteristics of the quantum mechanics is that physical quantities must be Hermitian. This aspect is deeply related to the requirement that these quantities should be described by real numbers. We deal with the Hermiticity from both an analytical point of view (or coordinate representation) relevant to the differential equations and an algebraic viewpoint (or matrix representation) associated with the operators and matrices. Including these topics, we briefly survey the origin of Schrödinger equation and consider its implications. To get acquainted with the quantum-mechanical formalism, we deal with simple examples of the Schrödinger equation.

## 1.1 Early-Stage Quantum Theory

The Schrödinger equation is a direct consequence of discovery of quanta. It stemmed from the hypothesis of energy quanta propounded by Max Planck (1900). This hypothesis was further followed by photon (light quantum) hypothesis propounded by Albert Einstein (1905). He claimed that light is an aggregation of light quanta and that individual quanta carry an energy $E$ expressed as Planck constant $h$ multiplied by frequency of light $v$, i.e.,



S. Hotta, *Mathematical Physical Chemistry*,
https://doi.org/10.1007/978-981-10-7671-8_1

$$E = h\nu = \hbar\omega, \tag{1.1}$$

where $\hbar \equiv h/2\pi$ and $\omega = 2\pi\nu$. The quantity $\omega$ is called angular frequency with $\nu$ being frequency. The quantity $\hbar$ is said to be a reduced Planck constant.

Also, Einstein (1917) concluded that momentum of light quantum $p$ is identical to the energy of light quantum divided by light velocity in vacuum $c$. That is, we have

$$p = E/c = \hbar\omega/c = \hbar k, \tag{1.2}$$

where $k \equiv 2\pi/\lambda$ ($\lambda$ is wavelength of light in vacuum) and $k$ is called wavenumber. Using vector notation, we have

$$\boldsymbol{p} = \hbar\boldsymbol{k}, \tag{1.3}$$

where $\boldsymbol{k} \equiv \frac{2\pi}{\lambda}\boldsymbol{n}$ ($\boldsymbol{n}$: a unit vector in the direction of propagation of light) is said to be a wavenumber vector.

Meanwhile, Arthur Compton (1923) conducted various experiments where he investigated how an incident X-ray beam was scattered by matter (e.g., graphite, copper, etc.). As a result, Compton found out a systematical redshift in X-ray wavelengths as a function of scattering angles of the X-ray beam (Compton effect). Moreover, he found that the shift in wavelengths depended only on the scattering angle regardless of quality of material of a scatterer. The results can be summarized in a simple equation described as

$$\Delta\lambda = \frac{h}{m_e c}(1 - \cos\theta), \tag{1.4}$$

where $\Delta\lambda$ denotes a shift in wavelength of the scattered beam; $m_e$ is a rest mass of an electron; $\theta$ is a scattering angle of the X-ray beam (see Fig. 1.1). A quantity $\frac{h}{m_e c}$ has a dimension of length and denoted by $\lambda_e$. That is,

$$\lambda_e \equiv h/m_e c. \tag{1.5}$$

In other words, $\lambda_e$ is equal to the maximum shift in the wavelength of the scattered beam; this shift is obtained when $\theta = \pi/2$. The quantity $\lambda_e$ is called an electron Compton wavelength and has an approximate value of $2.426 \times 10^{-12}$ (m).

Let us derive (1.4) on the basis of conservation of energy and momentum. To this end, in Fig. 1.1 we assume that an electron is originally at rest. An X-ray beam is incident to the electron. Then the X-ray is scattered and the electron recoils as shown. The energy conservation reads as

$$\hbar\omega + m_e c^2 = \hbar\omega' + \sqrt{p^2c^2 + m_e^2 c^4}, \tag{1.6}$$

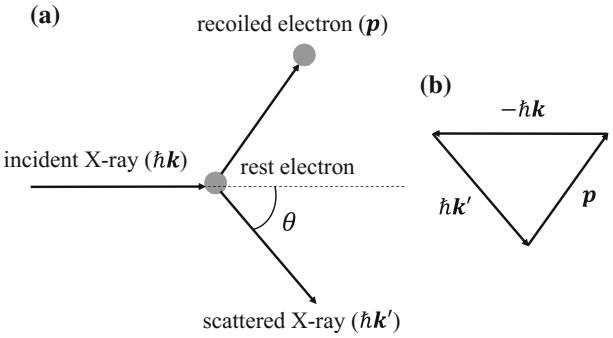


**Fig. 1.1** Scattering of an X-ray beam by an electron. **a** $\theta$ denotes a scattering angle of the X-ray beam. **b** conservation of momentum

where $\omega$ and $\omega'$ are initial and final angular frequencies of the X-ray; the second term of RHS is an energy of the electron in which $p$ is a magnitude of momentum after recoil. Meanwhile, conservation of the momentum as a vector quantity reads as

$$\hbar \boldsymbol{k} = \hbar \boldsymbol{k}' + \boldsymbol{p}, \tag{1.7}$$

where $\boldsymbol{k}$ and $\boldsymbol{k}'$ are wavenumber vectors of the X-ray before and after being scattered; $\boldsymbol{p}$ is a momentum of the electron after recoil. Note that an initial momentum of the electron is zero since the electron is originally at rest. Here, $\boldsymbol{p}$ is defined as

$$\boldsymbol{p} \equiv m\boldsymbol{u}, \tag{1.8}$$

where $\boldsymbol{u}$ is a velocity of an electron and $m$ is given by [1]

$$m = m_{\mathrm{e}} \Big/ \sqrt{1 - |\boldsymbol{u}|^2/c^2}. \tag{1.9}$$

Figure 1.1 shows that $-\hbar\boldsymbol{k}$, $\hbar\boldsymbol{k}'$, and $\boldsymbol{p}$ form a closed triangle.
From (1.6), we have

$$[m_{\mathrm{e}}c^2 + \hbar(\omega - \omega')]^2 = p^2c^2 + m_{\mathrm{e}}^2c^4. \tag{1.10}$$

Hence, we get

$$2m_{\mathrm{e}}c^2\hbar(\omega - \omega') + \hbar^2(\omega - \omega')^2 = p^2c^2. \tag{1.11}$$

From (1.7), we have

$$\boldsymbol{p}^2 = \hbar^2(\boldsymbol{k} - \boldsymbol{k}')^2 = \hbar^2\left(k^2 + k'^2 - 2kk'\cos\theta\right)$$

$$= \frac{\hbar^2}{c^2}\left(\omega^2 + \omega'^2 - 2\omega\omega' \cos\theta\right), \tag{1.12}$$

where we used the relations $\omega = ck$ and $\omega' = ck'$ with the third equality. Therefore, we get

$$p^2c^2 = \hbar^2\left(\omega^2 + \omega'^2 - 2\omega\omega' \cos\theta\right). \tag{1.13}$$

From (1.11) and (1.13), we have

$$2m_ec^2\hbar(\omega - \omega') + \hbar^2(\omega - \omega')^2 = \hbar^2\left(\omega^2 + \omega'^2 - 2\omega\omega' \cos\theta\right). \tag{1.14}$$

Equation (1.14) is simplified to the following:

$$2m_ec^2\hbar(\omega - \omega') - 2\hbar^2\omega\omega' = -2\hbar^2\omega\omega' \cos\theta.$$

That is,

$$m_ec^2(\omega - \omega') = \hbar\omega\omega'(1 - \cos\theta). \tag{1.15}$$

Thus, we get

$$\frac{\omega - \omega'}{\omega\omega'} = \frac{1}{\omega'} - \frac{1}{\omega} = \frac{1}{2\pi c}(\lambda' - \lambda) = \frac{\hbar}{m_ec^2}(1 - \cos\theta), \tag{1.16}$$

where $\lambda$ and $\lambda'$ are wavelengths of the initial and final X-ray beams, respectively. Since $\lambda' - \lambda = \Delta\lambda$, we have (1.4) from (1.16) accordingly.

We have to mention another important person, Louis Victor de Broglie (1924) in the development of quantum mechanics. Encouraged by the success of Einstein and Compton, he propounded the concept of matter wave, which was referred to as the de Broglie wave afterward. Namely, de Broglie reversed the relationship of (1.1) and (1.2) such that

$$\omega = E/\hbar, \tag{1.17}$$

and

$$k = p/\hbar \text{ or } \lambda = h/p, \tag{1.18}$$

where $p$ equals $|\boldsymbol{p}|$ and $\lambda$ is a wavelength of a corpuscular beam. This is said to be the de Broglie wavelength. In (1.18), de Broglie thought that a particle carrying an energy $E$ and momentum $p$ is accompanied by a wave that is characterized by an angular frequency $\omega$ and wavenumber $k$ (or a wavelength $\lambda = 2\pi/k$). Equation (1.18) implies that if we are able to determine the wavelength of the corpuscular beam experimentally, we can decide a magnitude of momentum accordingly.

In turn, from squares of both sides of (1.8) and (1.9), we get

$$u = \frac{p}{m_e\sqrt{1+(p/m_e c)^2}}. \tag{1.19}$$

This relation represents a velocity of particles of the corpuscular beam. If we are dealing with an electron beam, (1.19) gives the velocity of the electron beam. As a non-relativistic approximation (i.e., $p/m_e c \ll 1$), we have

$$p \approx m_e u.$$

We used a relativistic relation in the second term of RHS of (1.6), where energy of an electron $E_e$ is expressed by

$$E_e = \sqrt{p^2c^2 + m_e^2 c^4}. \tag{1.20}$$

In the meantime, deleting $u^2$ from (1.8) and (1.9), we have

$$mc^2 = \sqrt{p^2c^2 + m_e^2 c^4}.$$

Namely, we get [1]

$$E_e = mc^2. \tag{1.21}$$

The relation (1.21) is due to Einstein (1905, 1907) and is said to be the equivalence theorem of mass and energy.

If an electron is accompanied by a matter wave, that wave should be propagated with a certain phase velocity $v_p$ and a group velocity $v_g$. Thus, using (1.17) and (1.18), we have

$$\begin{aligned} v_p &= \omega/k = E_e/p = \sqrt{p^2c^2 + m_e^2c^4}/p > c, \\ v_g &= \partial\omega/\partial k = \partial E_e/\partial p = c^2 p/\sqrt{p^2c^2 + m_e^2c^4} < c, \\ v_p v_g &= c^2. \end{aligned} \tag{1.22}$$

Notice that in the above expressions, we replaced $E$ of (1.17) with $E_e$ of (1.20). The group velocity is thought to be a velocity of a wave packet and, hence, a propagation velocity of a matter wave should be identical to $v_g$. Thus, $v_g$ is considered as a particle velocity as well. In fact, $v_g$ given by (1.22) is identical to $u$ expressed in (1.19). Therefore, a particle velocity must not exceed $c$. As for photons (or light quanta), $v_p = v_g = c$ and, hence, once again we get $v_p v_g = c^2$. We will encounter the last relation of (1.22) in Part II as well.

The above discussion is a brief historical outlook of early-stage quantum theory before Erwin Schrödinger (1926) propounded his equation.

## 1.2 Schrödinger Equation

First, we introduce a wave equation expressed by

$$\nabla^2\psi = \frac{1}{v^2}\frac{\partial^2\psi}{\partial t^2}, \tag{1.23}$$

where $\psi$ is an arbitrary function of a physical quantity relevant to propagation of a wave; $v$ is a phase velocity of wave; $\nabla^2$ called Laplacian is defined below

$$\nabla^2 \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}. \tag{1.24}$$

One of special solutions for (1.24) called a plane wave is well studied and expressed as

$$\psi = \psi_0 \mathrm{e}^{i(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)}. \tag{1.25}$$

In (1.25), $\boldsymbol{x}$ denotes a position vector of a three-dimensional Cartesian coordinate and is described as

$$\boldsymbol{x} = (\boldsymbol{e}_1\boldsymbol{e}_2\boldsymbol{e}_3)\begin{pmatrix} x \\ y \\ z \end{pmatrix}, \tag{1.26}$$

where $\boldsymbol{e}_1, \boldsymbol{e}_2,$ and $\boldsymbol{e}_3$ denote basis vectors of an orthonormal base pointing to positive directions of $x$-, $y$-, and $z$-axes. Here, we make it a rule to represent basis vectors by a *row vector* and represent a coordinate or a component of a vector by a *column vector*; see Sect. 9.1.

The other way around, now we wish to seek a basic equation whose solution is described as (1.25). Taking account of (1.1)–(1.3) as well as (1.17) and (1.18), we rewrite (1.25) as

$$\psi = \psi_0 \mathrm{e}^{i\left(\frac{\boldsymbol{p}}{\hbar}\cdot\boldsymbol{x}-\frac{E}{\hbar}t\right)}, \tag{1.27}$$

where we redefine $\boldsymbol{p} = (\boldsymbol{e}_1\boldsymbol{e}_2\boldsymbol{e}_3)\begin{pmatrix} p_x \\ p_y \\ p_z \end{pmatrix}$ and $E$ as quantities associated with those of matter (electron) wave. Taking partial differentiation of (1.27) with respect to $x$, we obtain

$$\frac{\partial \psi}{\partial x} = \frac{i}{\hbar} p_x \psi_0 \mathrm{e}^{i\left(\frac{\boldsymbol{p}}{\hbar}\boldsymbol{x} - \frac{E}{\hbar}t\right)} = \frac{i}{\hbar} p_x \psi. \tag{1.28}$$

Rewriting (1.28), we have

$$\frac{\hbar}{i}\frac{\partial \psi}{\partial x} = p_x \psi. \tag{1.29}$$

Similarly, we have

$$\frac{\hbar}{i}\frac{\partial \psi}{\partial y} = p_y \psi \text{ and } \frac{\hbar}{i}\frac{\partial \psi}{\partial z} = p_z \psi. \tag{1.30}$$

Comparing both sides of (1.29), we notice that we may relate a differential operator $\frac{\hbar}{i}\frac{\partial}{\partial x}$ to $p_x$. From (1.30), similar relationship holds with the $y$ and $z$ components. That is, we have the following relations:

$$\frac{\hbar}{i}\frac{\partial}{\partial x} \leftrightarrow p_x, \frac{\hbar}{i}\frac{\partial}{\partial y} \leftrightarrow p_y, \frac{\hbar}{i}\frac{\partial}{\partial z} \leftrightarrow p_z \tag{1.31}$$

Taking partial differentiation of (1.28) once more,

$$\frac{\partial^2 \psi}{\partial x^2} = \left(\frac{i}{\hbar} p_x\right)^2 \psi_0 \mathrm{e}^{i\left(\frac{\boldsymbol{p}}{\hbar}\boldsymbol{x} - \frac{E}{\hbar}t\right)} = -\frac{1}{\hbar^2} p_x^2 \psi. \tag{1.32}$$

Hence,

$$-\hbar^2 \frac{\partial^2 \psi}{\partial x^2} = p_x^2 \psi. \tag{1.33}$$

Similarly, we have

$$-\hbar^2 \frac{\partial^2 \psi}{\partial y^2} = p_y^2 \psi \text{ and } -\hbar^2 \frac{\partial^2 \psi}{\partial z^2} = p_z^2 \psi \tag{1.34}$$

As in the above cases, we have

$$-\hbar^2 \frac{\partial^2}{\partial x^2} \leftrightarrow p_x^2, -\hbar^2 \frac{\partial^2}{\partial y^2} \leftrightarrow p_y^2, -\hbar^2 \frac{\partial^2}{\partial z^2} \leftrightarrow p_z^2, \tag{1.35}$$

Summing both sides of (1.33) and (1.34) and then dividing by $2m$, we have

$$-\frac{\hbar^2}{2m}\nabla^2 \psi = \frac{\boldsymbol{p}^2}{2m}\psi \tag{1.36}$$

and the following correspondence

$$-\frac{\hbar^2}{2m}\nabla^2 \leftrightarrow \frac{\boldsymbol{p}^2}{2m}, \tag{1.37}$$

where $m$ is the mass of a particle.

Meanwhile, taking partial differentiation of (1.27) with respect to $t$, we obtain

$$\frac{\partial \psi}{\partial t} = -\frac{i}{\hbar}E\psi_0 \mathrm{e}^{i\left(\frac{p}{\hbar}x - \frac{E}{\hbar}t\right)} = -\frac{i}{\hbar}E\psi. \tag{1.38}$$

That is,

$$i\hbar\frac{\partial \psi}{\partial t} = E\psi. \tag{1.39}$$

As the above, we get the following relationship:

$$i\hbar\frac{\partial}{\partial t} \leftrightarrow E \tag{1.40}$$

Thus, we have relationships between c-numbers (classical numbers) and q-numbers (quantum numbers, namely, operators) in (1.35) and (1.40). Subtracting (1.36) from (1.39), we get

$$i\hbar\frac{\partial \psi}{\partial t} + \frac{\hbar^2}{2m}\nabla^2\psi = \left(E - \frac{\boldsymbol{p}^2}{2m}\right)\psi. \tag{1.41}$$

Invoking the relationship on energy

$$(\text{Total energy}) = (\text{Kinetic energy}) + (\text{Potential energy}), \tag{1.42}$$

we have

$$E = \frac{\boldsymbol{p}^2}{2m} + V, \tag{1.43}$$

where $V$ is a potential energy. Thus, (1.41) reads as

$$i\hbar\frac{\partial \psi}{\partial t} + \frac{\hbar^2}{2m}\nabla^2\psi = V\psi. \tag{1.44}$$

Rearranging (1.44), we finally get

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + V\right)\psi = i\hbar\frac{\partial \psi}{\partial t}. \tag{1.45}$$

This is the Schrödinger equation, a fundamental equation of quantum mechanics. In (1.45), we define a following Hamiltonian operator $H$ as

$$H \equiv -\frac{\hbar^2}{2m}\nabla^2 + V. \tag{1.46}$$

Then we have a shorthand representation such that

$$H\psi = i\hbar\frac{\partial\psi}{\partial t}. \tag{1.47}$$

On going from (1.25) to (1.27), we realize that quantities $\boldsymbol{k}$ and $\omega$ pertinent to a field have been converted to quantities $\boldsymbol{p}$ and $E$ related to a particle. At the same time, whereas $\boldsymbol{x}$ and $t$ represent a whole space-time in (1.25), those in (1.27) are characterized as localized quantities.

From a historical point of view, we have to mention a great achievement accomplished by Werner Heisenberg (1925) who propounded matrix mechanics. The matrix mechanics is often contrasted with the wave mechanics Schrödinger initiated. Schrödinger and Pau Dirac (1926) demonstrated that wave mechanics and matrix mechanics are mathematically equivalent. Note that the Schrödinger equation is described as a non-relativistic expression based on (1.43). In fact, kinetic energy $K$ of a particle is given by [1]

$$K = \frac{m_e c^2}{\sqrt{1-(u/c)^2}} - m_e c^2$$

As a non-relativistic approximation, we get

$$K \approx m_e c^2\left[1+\frac{1}{2}\left(\frac{u}{c}\right)^2\right] - m_e c^2 = \frac{1}{2}m_e u^2 \approx \frac{p^2}{2m_e},$$

where we used $p \approx m_e u$ again as a non-relativistic approximation; also, we used

$$\frac{1}{\sqrt{1-x}} \approx 1+\frac{1}{2}x$$

when $x(>0)$ corresponding to $\left(\frac{u}{c}\right)^2$ is enough small than 1. This implies that in the above case, the group velocity $u$ of a particle is supposed to be well below light velocity $c$. Dirac (1928) formulated an equation that describes relativistic quantum mechanics (the Dirac equation).

In (1.45), $\psi$ varies as a function of $\boldsymbol{x}$ and $t$. Suppose, however, that a potential $V$ depends only upon $\boldsymbol{x}$. Then, we have

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V(\boldsymbol{x})\right]\psi(\boldsymbol{x},t) = i\hbar\frac{\partial\psi(\boldsymbol{x},t)}{\partial t}. \tag{1.48}$$

Now, let us assume that separation of variables can be done with (1.48) such that

$$\psi(\boldsymbol{x},t) = \phi(\boldsymbol{x})\xi(t). \tag{1.49}$$

Then, we have

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V(\boldsymbol{x})\right]\phi(\boldsymbol{x})\xi(t) = i\hbar\frac{\partial\phi(\boldsymbol{x})\xi(t)}{\partial t}. \tag{1.50}$$

Accordingly, (1.50) can be recast as

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V(\boldsymbol{x})\right]\phi(\boldsymbol{x})/\phi(\boldsymbol{x}) = i\hbar\frac{\partial\xi(t)}{\partial t}/\xi(t). \tag{1.51}$$

For (1.51) to hold, we must equate both sides to a constant $E$. That is, for a certain fixed point $\boldsymbol{x}_0$ we have

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V(\boldsymbol{x}_0)\right]\phi(\boldsymbol{x}_0)/\phi(\boldsymbol{x}_0) = i\hbar\frac{\partial\xi(t)}{\partial t}/\xi(t), \tag{1.52}$$

where $\phi(\boldsymbol{x}_0)$ of a numerator should be evaluated after operating $\nabla^2$, while with $\phi(\boldsymbol{x}_0)$ in a denominator, $\phi(\boldsymbol{x}_0)$ is evaluated simply replacing $\boldsymbol{x}$ in $\phi(\boldsymbol{x})$ with $\boldsymbol{x}_0$. Now, let us define a function $\Phi(\boldsymbol{x})$ such that

$$\Phi(\boldsymbol{x}) \equiv \left[-\frac{\hbar^2}{2m}\nabla^2 + V(\boldsymbol{x})\right]\phi(\boldsymbol{x})/\phi(\boldsymbol{x}). \tag{1.53}$$

Then, we have

$$\Phi(\boldsymbol{x}_0) = i\hbar\frac{\partial\xi(t)}{\partial t}/\xi(t). \tag{1.54}$$

If RHS of (1.54) varied depending on $t$, $\Phi(\boldsymbol{x}_0)$ would be allowed to have various values, but this must not be the case with our present investigation. Thus, RHS of (1.54) should take a constant value $E$. For the same reason, LHS of (1.51) should take a constant.

Thus, (1.48) or (1.51) should be separated into the following equations:

$$H\phi(\boldsymbol{x}) = E\phi(\boldsymbol{x}), \tag{1.55}$$

$$i\hbar \frac{\partial \xi(t)}{\partial t} = E\xi(t). \tag{1.56}$$

Equation (1.56) can readily be solved. Since (1.56) depends on a sole variable $t$, we have

$$\frac{d\xi(t)}{\xi(t)} = \frac{E}{i\hbar} \mathrm{d}t \text{ or } \mathrm{d}ln\xi(t) = \frac{E}{i\hbar} \mathrm{d}t. \tag{1.57}$$

Integrating (1.57) from zero to $t$, we get

$$ln \frac{\xi(t)}{\xi(0)} = \frac{Et}{i\hbar}. \tag{1.58}$$

That is,

$$\xi(t) = \xi(0) \exp(-iEt/\hbar). \tag{1.59}$$

Comparing (1.59) with (1.38), we find that the constant $E$ in (1.55) and (1.56) represents an energy of a particle (electron).

Thus, the next task we want to do is to solve an eigenvalue equation of (1.55). After solving the problem, we get a solution

$$\psi(\boldsymbol{x}, t) = \phi(\boldsymbol{x}) \exp(-iEt/\hbar), \tag{1.60}$$

where the constant $\xi(0)$ has been absorbed in $\phi(\boldsymbol{x})$. Normally, $\phi(\boldsymbol{x})$ is to be normalized after determining the functional form (vide infra).

## 1.3 Simple Applications of Schrödinger Equation

The Schrödinger equation has been expressed as (1.48). The equation is a second-order linear differential equation (SOLDE). In particular, our major interest lies in solving an eigenvalue problem of (1.55). Eigenvalues consist of points in a complex plane. Those points sometimes form a continuous domain, but we focus on the eigenvalues that comprise discrete points in the complex plane. Therefore in our studies, the eigenvalues are countable and numbered as, e.g., $\lambda_n$ $(n = 1, 2, 3, \ldots)$. An example is depicted in Fig. 1.2. Having this common belief as a background, let us first think of a simple form of SOLDE.

*Example 1.1* Let us think of a following differential equation:

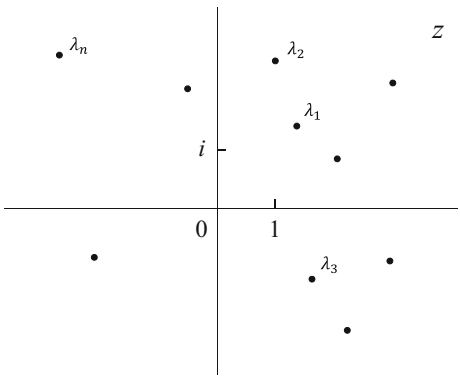


**Fig. 1.2** Eigenvalues $\lambda_n\,(n = 1, 2, 3, \ldots)$ on a complex plane

$$\frac{\mathrm{d}^2 y(x)}{\mathrm{d}x^2} + \lambda y(x) = 0, \tag{1.61}$$

where $x$ is a real variable; $y$ may be a complex function of $x$ with $\lambda$ possibly being a complex constant as well. Suppose that $y(x)$ is defined within a domain $[-L, L]\,(L > 0)$. We set boundary conditions (BCs) for (1.61) such that

$$y(L) = 0 \text{ and } y(-L) = 0\,(L > 0). \tag{1.62}$$

The BCs of (1.62) are called Dirichlet conditions. We define the following differential operator $D$ described as

$$D \equiv -\frac{\mathrm{d}^2}{\mathrm{d}x^2}. \tag{1.63}$$

Then rewriting (1.61), we have

$$Dy(x) = \lambda y(x). \tag{1.64}$$

According to a general principle of SOLDE, it has two linearly independent solutions. In the case of (1.61), we choose exponential functions for those solutions described by

$$\mathrm{e}^{ikx} \text{ and } \mathrm{e}^{-ikx} (k \neq 0).$$

This is because the above functions do not change a functional form with respect to the differentiation and we ascribe solving a differential equation to solving an algebraic equation among constants (or parameters). In the present case, $\lambda$ and $k$ are such constants.

The parameter $k$ could be a complex variable, because $\lambda$ is allowed to take a complex value as well. Linear independence of these functions is ensured from a nonvanishing Wronskian, $W$. That is,

$$W = \begin{vmatrix} e^{ikx} & e^{-ikx} \\ (e^{ikx})' & (e^{-ikx})' \end{vmatrix} = \begin{vmatrix} e^{ikx} & e^{-ikx} \\ ike^{ikx} & -ike^{-ikx} \end{vmatrix} = -ik - ik = -2ik. \tag{1.65}$$

If $k \neq 0$, $W \neq 0$. Therefore, as a general solution, we get

$$y(x) = ae^{ikx} + be^{-ikx} (k \neq 0), \tag{1.66}$$

where $a$ and $b$ are (complex) constant. We call two linearly independent solutions $e^{ikx}$ and $e^{-ikx} (k \neq 0)$ a fundamental set of solutions of a SOLDE. Inserting (1.66) into (1.61), we have

$$\left(\lambda - k^2\right)\left(ae^{ikx} + be^{-ikx}\right) = 0. \tag{1.67}$$

For (1.67) to hold with any $x$, we must have

$$\lambda - k^2 = 0 \,\text{i.e.} \lambda = k^2. \tag{1.68}$$

Using BCs (1.62), we have

$$ae^{ikL} + be^{-ikL} = 0 \text{ and } ae^{-ikL} + be^{ikL} = 0. \tag{1.69}$$

Rewriting (1.69) in a matrix form, we have

$$\begin{pmatrix} e^{ikL} & e^{-ikL} \\ e^{-ikL} & e^{ikL} \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{1.70}$$

For $a$ and $b$ in (1.70) to have nonvanishing solutions, we must have

$$\begin{vmatrix} e^{ikL} & e^{-ikL} \\ e^{-ikL} & e^{ikL} \end{vmatrix} = 0 \quad \text{i.e } e^{2ikL} - e^{-2ikL} = 0. \tag{1.71}$$

It is because if (1.71) were not zero, we would have $a = b = 0$ and $y(x) \equiv 0$. Note that with an eigenvalue problem, we must avoid having a solution that is identically zero. Rewriting (1.71), we get

$$\left(e^{ikL} + e^{-ikL}\right)\left(e^{ikL} - e^{-ikL}\right) = 0. \tag{1.72}$$

That is, we have either

$$e^{ikL} + e^{-ikL} = 0 \tag{1.73}$$

or

$$e^{ikL} - e^{-ikL} = 0. \tag{1.74}$$

In the case of (1.73), inserting this into (1.69), we have

$$e^{ikL}(a - b) = 0. \tag{1.75}$$

Therefore,

$$a = b, \tag{1.76}$$

where we used the fact that $e^{ikL}$ is a nonvanishing function for any $ikL$ (either real or complex). Similarly, in the case of (1.74), we have

$$a = -b. \tag{1.77}$$

For (1.76), from (1.66), we have

$$y(x) = a(e^{ikx} + e^{-ikx}) = 2a \cos kx. \tag{1.78}$$

With (1.77), in turn, we get

$$y(x) = a(e^{ikx} - e^{-ikx}) = 2ia \sin kx. \tag{1.79}$$

Thus, we get two linearly independent solutions (1.78) and (1.79).
Inserting BCs (1.62) into (1.78), we have

$$\cos kL = 0. \tag{1.80}$$

Hence,

$$kL = \frac{\pi}{2} + m\pi \, (m = 0, \pm 1, \pm 2, \ldots). \tag{1.81}$$

In (1.81), for instance, we have $k = \frac{\pi}{2L}$ for $m = 0$ and $k = -\frac{\pi}{2L}$ for $m = -1$. Also, we have $k = \frac{3\pi}{2L}$ for $m = 1$ and $k = -\frac{3\pi}{2L}$ for $m = -2$. These cases, however, individually give linearly dependent solutions for (1.78). Therefore, to get a set of linearly independent eigenfunctions, we may define $k$ as positive. Correspondingly, from (1.68), we get eigenvalues of

$$\lambda = (2m + 1)^2 \pi^2 / 4L^2 \, (m = 0, 1, 2, \ldots). \tag{1.82}$$

Also, inserting BCs (1.62) into (1.79), we have

$$\sin kL = 0. \tag{1.83}$$

Hence,

$$kL = n\pi \, (n = 1, 2, 3, \ldots). \tag{1.84}$$

From (1.68), we get

$$\lambda = n^2\pi^2/L^2 = (2n)^2\pi^2/4L^2 \, (n = 1, 2, 3, \ldots), \tag{1.85}$$

where we chose positive numbers $n$ for the same reason as the above. With the second equality of (1.85), we made eigenvalues easily comparable to those of (1.82). Figure 1.3 shows the eigenvalues given in both (1.82) and (1.85) in a unit of $\pi^2/4L^2$.

From (1.82) and (1.85), we find that $\lambda$ is *positive definite* (or strictly positive), and so from (1.68), we have

$$k = \sqrt{\lambda}. \tag{1.86}$$

The next step is to normalize eigenfunctions. This step corresponds to appropriate choice of a constant $a$ in (1.78) and (1.79) so that we can have

$$I = \int_{-L}^{L} y(x)^* y(x) \mathrm{d}x = \int_{-L}^{L} |y(x)|^2 \mathrm{d}x = 1. \tag{1.87}$$

That is,

$$\begin{aligned} I &= 4|a|^2 \int_{-L}^{L} \cos^2 kx \mathrm{d}x = 4|a|^2 \int_{-L}^{L} \frac{1}{2}(1 + \cos 2kx) \mathrm{d}x \\ &= 2|a|^2 \left[ x + \frac{1}{2k} \sin 2kx \right]_{-L}^{L} = 4L|a|^2. \end{aligned} \tag{1.88}$$

Combining (1.87) and (1.88), we get

$$|a| = \frac{1}{2}\sqrt{\frac{1}{L}}. \tag{1.89}$$

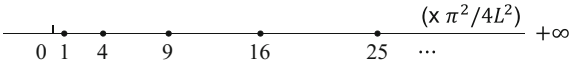


**Fig. 1.3** Eigenvalues of a differential Eq. (1.61) under boundary conditions given by (1.62). The eigenvalues are given in a unit of $\pi^2/4L^2$ on a real axis

Thus, we have

$$a = \frac{1}{2}\sqrt{\frac{1}{L}}e^{i\theta}, \tag{1.90}$$

where $\theta$ is any real number and $e^{i\theta}$ is said to be a phase factor. We usually set $e^{i\theta} \equiv 1$. Then, we have $a = \frac{1}{2}\sqrt{\frac{1}{L}}$. Thus for a normalized cosine eigenfunctions, we get

$$y(x) = \sqrt{\frac{1}{L}}\cos kx\left[kL = \frac{\pi}{2} + m\pi \ (m = 0, 1, 2, \ldots)\right] \tag{1.91}$$

that corresponds to an eigenvalue $\lambda = (2m+1)^2\pi^2/4L^2$ $(m = 0, 1, 2, \ldots)$. For another series of normalized sine functions, similarly, we get

$$y(x) = \sqrt{\frac{1}{L}}\sin kx[kL = n\pi \ (n = 1, 2, 3, \ldots)] \tag{1.92}$$

that corresponds to an eigenvalue $\lambda = (2n)^2\pi^2/4L^2$ $(n = 1, 2, 3, \ldots)$.

Notice that arranging $\lambda$ in ascending order, we have even functions and odd functions alternately as eigenfunctions corresponding to $\lambda$. Such a property is said to be *parity*. We often encounter it in quantum mechanics and related fields. From (1.61), we find that if $y(x)$ is an eigenfunction, so is $cy(x)$. That is, we should bear in mind that the eigenvalue problem is always accompanied by an indeterminate constant and that normalization of an eigenfunction does not mean the uniqueness of the solution (see Chap. 8).

Strictly speaking, we should be careful to assure that (1.81) holds on the basis of (1.80). It is because we have yet the possibility that $k$ is a complex number. To see it, we examine zeros of a cosine function that is defined in a complex domain. Here, the zeros are (complex) numbers to which the function takes zero. That is, if $f(z_0) = 0$, $z_0$ is called a zero (i.e., one of zeros) of $f(z)$. Now, we have

$$\cos z \equiv \frac{1}{2}\left(e^{iz} + e^{-iz}\right); z = x + iy \ (x, y : \text{real}). \tag{1.93}$$

Inserting $z = x + iy$ in $\cos z$ and rearranging terms, we get

$$\cos z = \frac{1}{2}[\cos x(e^{y} + e^{-y}) + i \ \sin x(e^{-y} - e^{y})]. \tag{1.94}$$

For $\cos z$ to vanish, both its real and imaginary parts must be zero. Since $e^{y} + e^{-y} > 0$ for all real numbers $y$, we must have $\cos x = 0$ for the real part to vanish, i.e.,

$$x = \frac{\pi}{2} + m\pi \, (m = 0, \pm 1, \pm 2, \ldots). \tag{1.95}$$

Note in this case that $\sin x = \pm 1 \, (\neq 0)$. Therefore, for the imaginary part to vanish, $\mathrm{e}^{-y} - \mathrm{e}^{y} = 0$. That is, we must have $y = 0$. Consequently, the zeros of $\cos z$ are real numbers. In other words, with respect to $z_0$ that satisfies $\cos z_0 = 0$, we have

$$z_0 = \frac{\pi}{2} + m\pi \, (m = 0, \pm 1, \pm 2, \ldots). \tag{1.96}$$

The above discussion equally applies to a sine function as well.

Thus, we ensure that $k$ is a nonzero real number. Eigenvalues $\lambda$ are positive definite from (1.68) accordingly. This conclusion is not fortuitous but a direct consequence of the form of a differential equation we have dealt with in combination with the BCs we imposed, i.e., the Dirichlet conditions. Detailed discussion will follow in Sects. 1.4, 8.3, and 8.4 in relation to the Hermiticity of a differential operator.

*Example 1.2* A particle confined within a potential well.

The results obtained in Example 1.1 can immediately be applied to dealing with a particle (electron) in a one-dimensional infinite potential well. In this case, (1.55) reads as

$$\frac{\hbar^2}{2m} \frac{\mathrm{d}^2 \psi(x)}{\mathrm{d}x^2} + E\psi(x) = 0, \tag{1.97}$$

where $m$ is a mass of a particle and $E$ is an energy of the particle. A potential $V$ is expressed as

$$V(x) = \begin{cases} 0 \, (-L \leq x \leq L), \\ \infty \, (-L > x; x > L). \end{cases}$$

Rewriting (1.97), we have

$$\frac{\mathrm{d}^2 \psi(x)}{\mathrm{d}x^2} + \frac{2mE}{\hbar^2} \psi(x) = 0 \tag{1.98}$$

with BCs

$$\psi(L) = \psi(-L) = 0. \tag{1.99}$$

If we replace $\lambda$ of (1.61) with $\frac{2mE}{\hbar^2}$, we can follow the procedures of Example. 1.1. That is, we put

$$E = \frac{\hbar^2 \lambda}{2m} \tag{1.100}$$

with $\lambda = k^2$ in (1.68). For $k$, we use the values of (1.81) and (1.84). Therefore, with energy eigenvalues, we get either

$$E = \frac{\hbar^2}{2m} \cdot \frac{(2m+1)^2\pi^2}{4L^2} \ (m = 0, 1, 2, \ldots), \tag{1.101}$$

to which $y(x) = \sqrt{\frac{1}{L}} \cos kx \left[kL = \frac{\pi}{2} + m\pi \ (m = 0, 1, 2, \ldots)\right]$ corresponds or

$$E = \frac{\hbar^2}{2m} \cdot \frac{(2n)^2\pi^2}{4L^2} (n = 1, 2, 3, \ldots), \tag{1.102}$$

to which $y(x) = \sqrt{\frac{1}{L}} \sin kx [kL = n\pi \ (n = 1, 2, 3, \ldots)]$ corresponds.

Since the particle behaves as a free particle within the potential well $(-L \leq x \leq L)$ and $p = \hbar k$, we obtain

$$E = \frac{p^2}{2m} = \frac{\hbar^2}{2m} k^2,$$

where

$$k = \begin{cases} (2m+1)\pi/2L & (m = 0, 1, 2, \ldots), \\ 2n\pi/2L & (n = 1, 2, 3, \ldots). \end{cases}$$

The energy $E$ is a kinetic energy of the particle.

Although in (1.97), $\psi(x) \equiv 0$ trivially holds, such a function may not be regarded as a solution of the eigenvalue problem. In fact, considering that $|\psi(x)|^2$ represents existence probability of a particle, $\psi(x) \equiv 0$ corresponds to a situation where a particle in question does not exist. Consequently, such a trivial case has physically no meaning.

## 1.4 Quantum-Mechanical Operators and Matrices

As represented by (1.55), a quantum-mechanical operator corresponds to a physical quantity. In (1.55), we connect a Hamiltonian operator to an energy (eigenvalue). Let us rephrase the situation as follows:

$$P\Psi = p\Psi. \tag{1.103}$$

In (1.103), we are viewing $P$ as an operation or measurement on a physical system that is characterized by the quantum state $\Psi$. Operating $P$ on the physical system (or state), we obtain a physical quantity $p$ relevant to $P$ as a result of the operation (or measurement).

A way to effectively achieve the above is to use a matrix and vector to represent the operation and physical state, respectively. Let us glance a little bit of matrix calculation to get used to the quantum-mechanical concept and, hence, to obtain clear understanding about it. In Part III, we will deal with matrix calculation in detail from a point of view of a general principle. At present, a (2,2) matrix suffices. Let $A$ be a (2,2) matrix expressed as

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}. \tag{1.104}$$

Let $|\psi\rangle$ be a (2,1) matrix, i.e., a column vector such that

$$|\psi\rangle = \begin{pmatrix} e \\ f \end{pmatrix}. \tag{1.105}$$

Note that operating (2,2) matrix on a (2,1) matrix produces another (2,1) matrix. Furthermore, we define an adjoint matrix $A^\dagger$ such that

$$A^\dagger = \begin{pmatrix} a^* & c^* \\ b^* & d^* \end{pmatrix}, \tag{1.106}$$

where $a^*$ is a complex conjugate of $a$. That is, $A^\dagger$ is a complex conjugate transposed matrix of $A$. Also, we define an adjoint vector $\langle\psi|$ or $|\psi\rangle^\dagger$ such that

$$\langle\psi| \equiv |\psi\rangle^\dagger = (e^* f^*). \tag{1.107}$$

In this case, $|\psi\rangle^\dagger$ also denotes a complex conjugate transpose of $|\psi\rangle$. The notation $|\psi\rangle$ and $\langle\psi|$ are due to Dirac. He named $\langle\psi|$ and $|\varphi\rangle$ a *bra* vector and *ket* vector, respectively. This naming or equivoque comes from that $\langle\psi| \cdot |\varphi\rangle = \langle\psi \mid \varphi\rangle$ forms a bracket. This is a $(1,2) \times (2,1) = (1,1)$ matrix, i.e., a c-number (including a complex number) and $\langle\psi \mid \varphi\rangle$ represent an inner product. These notations are widely used nowadays in the field of mathematics and physics.

Taking another vector $|\xi\rangle = \begin{pmatrix} g \\ h \end{pmatrix}$ and using a matrix calculation rule, we have

$$A^\dagger|\psi\rangle = \left|A^\dagger\psi\right\rangle = \begin{pmatrix} a^* & c^* \\ b^* & d^* \end{pmatrix}\begin{pmatrix} e \\ f \end{pmatrix} = \begin{pmatrix} a^*e + c^*f \\ b^*e + d^*f \end{pmatrix}. \tag{1.108}$$

According to the definition (1.107), we have

$$\left|A^\dagger\psi\right\rangle^\dagger = \left\langle A^\dagger\psi\right| = (ae^* + cf^* be^* + df^*). \tag{1.109}$$

Thus, we get

$$\left\langle A^\dagger\psi \,\middle|\, \xi \right\rangle = (ae^* + cf^* be^* + df^*)\begin{pmatrix} g \\ h \end{pmatrix} = (ag + bh)e^* + (cg + dh)f^*. \quad (1.110)$$

Similarly, we have

$$\langle \psi \mid A\xi \rangle = (e^* f^*)\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} g \\ h \end{pmatrix} = (ag + bh)e^* + (cg + dh)f^*. \quad (1.111)$$

Comparing (1.110) and (1.111), we get

$$\left\langle A^\dagger\psi \,\middle|\, \xi \right\rangle = \langle \psi \mid A\xi \rangle. \quad (1.112)$$

Also, we have

$$\langle \psi \mid A\xi \rangle^* = \langle A\xi \mid \psi \rangle. \quad (1.113)$$

Replacing $A$ with $A^\dagger$ in (1.112), we get

$$\left\langle (A^\dagger)^\dagger\psi \,\middle|\, \xi \right\rangle = \left\langle \psi \,\middle|\, A^\dagger\xi \right\rangle. \quad (1.114)$$

From (1.104) and (1.106), obviously we have

$$(A^\dagger)^\dagger = A. \quad (1.115)$$

Then from (1.114) and (1.115), we have

$$\langle A\psi \mid \xi \rangle = \left\langle \psi \,\middle|\, A^\dagger\xi \right\rangle = \langle \xi \mid A\psi \rangle^*, \quad (1.116)$$

where the second equality comes from (1.113) obtained by exchanging $\psi$ and $\xi$ there. Moreover, we have a following relation:

$$(AB)^\dagger = B^\dagger A^\dagger. \quad (1.117)$$

The proof is left for readers. Using this relation, we have

$$\langle A\psi| = |A\psi\rangle^\dagger = [A|\psi\rangle]^\dagger = |\psi\rangle^\dagger A^\dagger = \langle\psi|A^\dagger = \left\langle \psi A^\dagger \right|. \quad (1.118)$$

Making an inner product by multiplying $|\xi\rangle$ from the right of the leftmost and rightmost sides of (1.118) and using (1.116), we get

$$\langle A\psi \mid \xi\rangle = \left\langle \psi A^{\dagger} \,\middle|\, \xi \right\rangle = \left\langle \psi \,\middle|\, A^{\dagger}\xi \right\rangle.$$

This relation may be regarded as the associative law with regard to the symbol "|" of the inner product. This is equivalent to the associative law with regard to the matrix multiplication.

The results obtained above can readily be extended to a general case where $(n,n)$ matrices are dealt with.

Now, let us introduce a Hermitian operator (or matrix) $H$. When we have

$$H^{\dagger} = H, \tag{1.119}$$

$H$ is called a Hermitian matrix. Then, applying (1.112) to the Hermitian matrix $H$, we have

$$\left\langle H^{\dagger}\psi \,\middle|\, \xi \right\rangle = \langle \psi \mid H\xi\rangle = \left\langle \psi \,\middle|\, H^{\dagger}\xi \right\rangle \text{ or } \langle H\psi \mid \xi\rangle = \left\langle \psi \,\middle|\, H^{\dagger}\xi \right\rangle = \langle \psi \mid H\xi\rangle \tag{1.120}$$

Also, let us introduce a norm of a vector $|\psi\rangle$ such that

$$||\psi|| = \sqrt{\langle \psi \mid \psi\rangle}. \tag{1.121}$$

A norm is a natural extension for a notion of a "length" of a vector. The norm $||\psi||$ is zero, if and only if $|\psi\rangle = 0$ (zero vector). For, from (1.105) and (1.107), we have

$$\langle \psi \mid \psi\rangle = |e|^2 + |f|^2.$$

Therefore, $\langle \psi \mid \psi\rangle = 0 \Leftrightarrow e = f = 0,$ i.e.$|\psi\rangle = 0$.

Let us further consider an eigenvalue problem represented by our newly introduced notation. The eigenvalue equation is symbolically written as

$$H|\psi\rangle = \lambda|\psi\rangle, \tag{1.122}$$

where $H$ represents a Hermitian operator and $|\psi\rangle$ is an eigenfunction that belongs to an eigenvalue $\lambda$. Operating $\langle\psi|$ on (1.122) from the left, we have

$$\langle\psi|H|\psi\rangle = \langle\psi|\lambda|\psi\rangle = \lambda\langle\psi \mid \psi\rangle = \lambda, \tag{1.123}$$

where we assume that $|\psi\rangle$ is normalized; namely $\langle\psi \mid \psi\rangle = 1$ or $\|\psi\| = 1$. Notice that the symbol "|" in an inner product is of secondary importance. We may disregard this notation as in the case where a product notation "$\times$" is omitted by denoting $ab$ instead of $a \times b$.

Taking a complex conjugate of (1.123), we have

$$\langle \psi \mid H\psi \rangle^* = \lambda^*. \tag{1.124}$$

Using (1.116) and (1.124), we have

$$\lambda^* = \langle \psi \mid H\psi \rangle^* = \left\langle \psi \,\middle|\, H^\dagger \psi \right\rangle = \langle \psi \mid H\psi \rangle = \lambda, \tag{1.125}$$

where with the third equality we used the definition (1.119). The relation $\lambda^* = \lambda$ obviously shows that any eigenvalue $\lambda$ is real, if $H$ is Hermitian. The relation (1.125) immediately tells us that even though $|\psi\rangle$ is not an eigenfunction, $\langle \psi \mid H\psi \rangle$ is real as well, if $H$ is Hermitian. The quantity $\langle \psi \mid H\psi \rangle$ is said to be an expectation value. This value is interpreted as the most probable or averaged value of $H$ obtained as a result of operation of $H$ on a physical state $|\psi\rangle$. We sometimes denote the expectation value as

$$\langle H \rangle \equiv \langle \psi \mid H\psi \rangle, \tag{1.126}$$

where $|\psi\rangle$ is normalized. Unless $|\psi\rangle$ is not normalized, it can be normalized on the basis of (1.121) by choosing $|\Phi\rangle$ such that

$$|\Phi\rangle = |\psi\rangle / \|\psi\|. \tag{1.127}$$

Thus, we have an important consequence; if a Hermitian operator has an eigenvalue, it must be real. An expectation value of a Hermitian operator is real as well. The real eigenvalue and expectation value are a prerequisite for a physical quantity.

As discussed above, the Hermitian matrices play a central role in quantum physics. Taking a further step, let us extend the notion of Hermiticity to a function space.

In Example 1.1, we have remarked that we have finally reached a solution where $\lambda$ is a real (and positive) number, even though at the beginning we set no restriction on $\lambda$. This is because the SOLDE form (1.61) accompanied by BCs (1.62) is Hermitian, and so eigenvalues $\lambda$ are real.

In this context, we give a little bit of further consideration. We define an inner product between two functions as follows:

$$\langle g \mid f \rangle \equiv \int_a^b g(x)^* f(x) \mathrm{d}x, \tag{1.128}$$

where $g(x)^*$ is a complex conjugate of $g(x)$; $x$ is a real variable and an integration range can be either bounded or unbounded. If $a$ and $b$ are real definite numbers, $[a, b]$ is the bounded case. With the unbounded case, we have, e.g., $(-\infty, \infty)$, $(-\infty, c)$, and $(c, \infty)$, etc. where $c$ is a definite number. This notation

will appear again in Chap. 8. In (1.128), we view functions $f$ and $g$ as vectors in a function space, often referred to as a Hilbert space. We assume that any function $f$ is square-integrable, i.e., $|f|^2$ is finite. That is,

$$\int_a^b |f(x)|^2 \mathrm{d}x < \infty. \tag{1.129}$$

Using the above definition, let us calculate $\langle g \mid Df \rangle$, where $D$ was defined in (1.63). Then, using the integration by parts, we have

$$\begin{aligned}\langle g \mid Df \rangle &= \int_a^b g(x)^* \left[ -\frac{\mathrm{d}^2 f(x)}{\mathrm{d}x^2} \right] \mathrm{d}x = -[g^* f']_a^b + \int_a^b g^{*\prime} f' \mathrm{d}x \\ &= -[g^* f']_a^b + [g^{*\prime} f]_a^b - \int_a^b g^{*\prime\prime} f \mathrm{d}x = [g^{*\prime} f - g^* f']_a^b + \int_a^b (-g^{*\prime\prime} f) \mathrm{d}x \\ &= [g^{*\prime} f - g^* f']_a^b + \langle Dg \mid f \rangle.\end{aligned} \tag{1.130}$$

If we have BCs such that

$$f(b) = f(a) = 0 \text{ and } g(b)^* = g(a)^* = 0 \text{ i.e., } g(b) = g(a) = 0, \tag{1.131}$$

we get

$$\langle g \mid Df \rangle = \langle Dg \mid f \rangle. \tag{1.132}$$

In light of (1.120), (1.132) implies that $D$ is Hermitian. In (1.131), notice that the functions $f$ and $g$ satisfy the same BCs. Normally, for an operator to be Hermitian assumes this property. Thus, the Hermiticity of a differential operator is closely related to BCs of the differential equation.

Next, we consider a following inner product:

$$\langle f \mid Df \rangle = -\int_a^b f^* f'' \mathrm{d}x = -[f^* f']_a^b + \int_a^b f^{*\prime} f' \mathrm{d}x = -[f^* f']_a^b + \int_a^b |f'|^2 \mathrm{d}x. \tag{1.133}$$

Note that the definite integral of (1.133) cannot be negative. There are two possibilities for $D$ to be Hermitian according to different BCs.

(i) Dirichlet conditions: $f(b) = f(a) = 0$. If we could have $f' = 0$, $\langle f \mid Df \rangle$ would be zero. But, in that case, $f$ should be constant. If so, $f(x) \equiv 0$ according to

BCs. We must exclude this trivial case. Consequently, to avoid this situation, we must have

$$\int_a^b |f'|^2 \mathrm{d}x > 0 \text{ or } \langle f \mid Df \rangle > 0. \tag{1.134}$$

In this case, the operator $D$ is said to be positive definite. Suppose that such a positive-definite operator has an eigenvalue $\lambda$. Then, for a corresponding eigenfunction $y(x)$, we have

$$Dy(x) = \lambda y(x). \tag{1.135}$$

In this case, we state that $y(x)$ is an eigenfunction or eigenvector that corresponds (or belongs) to an eigenvalue $\lambda$. Taking an inner product of both sides, we have

$$\langle y \mid Dy \rangle = \langle y \mid \lambda y \rangle = \lambda \langle y \mid y \rangle = \lambda \|y\|^2 \text{ or } \lambda = \langle y \mid Dy \rangle / \|y\|^2. \tag{1.136}$$

Both $\langle y \mid Dy \rangle$ and $\|y\|^2$ are positive and, hence, we have $\lambda > 0$. Thus, if $D$ has an eigenvalue, it must be positive. In this case, $\lambda$ is said to be positive definite as well; see Example 1.1.

(ii) Neumann conditions: $f'(b) = f'(a) = 0$. From (1.130), $D$ is Hermitian as well. Unlike the condition (i), however, $f$ may be a nonzero constant in this case. Therefore, we are allowed to have

$$\int_a^b |f'|^2 \mathrm{d}x = 0 \text{ or } \langle f \mid Df \rangle = 0. \tag{1.137}$$

For any function, we have

$$\langle f \mid Df \rangle \geq 0. \tag{1.138}$$

In this case, the operator $D$ is said to be *nonnegative* (or *positive semi-definite*). The eigenvalue may be zero from (1.136) and, hence, is called nonnegative accordingly.

(iii) Periodic conditions: $f(b) = f(a)$ and $f'(b) = f'(a)$. We are allowed to have $\langle f \mid Df \rangle \geq 0$ as in the case of the condition (ii). Then, the operator and eigenvalues are nonnegative.

Thus, in spite of being formally the same operator, that operator behaves differently according to the different BCs. In particular, for a differential operator to be

associated with an eigenvalue of zero produces a special interest. We will encounter another illustration in Chap. 3.

## 1.5 Commutator and Canonical Commutation Relation

In quantum mechanics, it is important whether two operators $A$ and $B$ are commutable. In this context, a commutator between $A$ and $B$ is defined such that

$$[A, B] \equiv AB - BA. \tag{1.139}$$

If $[A, B] = 0$ (zero matrix), $A$ and $B$ are said to be commutable (or commutative). If $[A, B] \neq 0$, $A$ and $B$ are non-commutative. Such relationships between two operators are called commutation relation.

We have canonical commutation relation as an underlying concept of quantum mechanics. This is defined between a (canonical) coordinate $q$ and a (canonical) momentum $p$ such that

$$[q, p] = i\hbar, \tag{1.140}$$

where the presence of a unit matrix $\hat{E}$ is implied. Explicitly writing it, we have,

$$[q, p] = i\hbar\hat{E}. \tag{1.141}$$

The relations (1.140) and (1.141) are called the canonical commutation relation. On the basis of a relation $p = \frac{\hbar}{i}\frac{\partial}{\partial q}$, a brief proof for this is as follows:

$$\begin{aligned}[q, p]|\psi\rangle &= (qp - pq)|\psi\rangle = \left(q\frac{\hbar}{i}\frac{\partial}{\partial q} - \frac{\hbar}{i}\frac{\partial}{\partial q}q\right)|\psi\rangle = q\frac{\hbar}{i}\frac{\partial}{\partial q}|\psi\rangle - \frac{\hbar}{i}\frac{\partial}{\partial q}(q|\psi\rangle) \\ &= q\frac{\hbar}{i}\frac{\partial|\psi\rangle}{\partial q} - \frac{\hbar}{i}\frac{\partial q}{\partial q}|\psi\rangle - \frac{\hbar}{i}q\frac{\partial|\psi\rangle}{\partial q} = -\frac{\hbar}{i}|\psi\rangle = i\hbar|\psi\rangle\end{aligned} \tag{1.142}$$

Since $|\psi\rangle$ is an arbitrarily chosen vector, we have (1.140).

Using (1.117), we have

$$[A, B]^\dagger = (AB - BA)^\dagger = B^\dagger A^\dagger - A^\dagger B^\dagger. \tag{1.143}$$

If in (1.143) $A$ and $B$ are both Hermitian, we have

$$[A, B]^\dagger = BA - AB = -[A, B]. \tag{1.144}$$

If we have an operator $G$ such that

$$G^{\dagger} = -G, \tag{1.145}$$

$G$ is said to be anti-Hermitian. Therefore, $[A, B]$ is anti-Hermitian, if both $A$ and $B$ are Hermitian. If an anti-Hermitian operator has an eigenvalue, the eigenvalue is zero or pure imaginary. To show this, suppose that

$$G|\psi\rangle = \lambda|\psi\rangle, \tag{1.146}$$

where $G$ is an anti-Hermitian operator and $|\psi\rangle$ has been normalized. As in the case of (1.123), we have

$$\langle\psi|G|\psi\rangle = \lambda\langle\psi \mid \psi\rangle = \lambda. \tag{1.147}$$

Taking a complex conjugate of (1.147), we have

$$\langle\psi \mid G\psi\rangle^{*} = \lambda^{*}. \tag{1.148}$$

Using (1.116) and (1.145) again, we have

$$\lambda^{*} = \langle\psi \mid G\psi\rangle^{*} = \left\langle\psi \,\middle|\, G^{\dagger}\psi\right\rangle = -\langle\psi \mid G\psi\rangle = -\lambda, \tag{1.149}$$

This shows that $\lambda$ is zero or pure imaginary.

Therefore, (1.142) can be viewed as an eigenvalue equation to which any physical state $|\psi\rangle$ has a pure imaginary eigenvalue $i\hbar$ with respect to $[q, p]$. Note that both $q$ and $p$ are Hermitian (see Sect. 8.2, Example 8.3), and so $[q, p]$ is anti-Hermitian as mentioned above. The canonical commutation relation given by (1.140) is believed to underpin the uncertainty principle.

In quantum mechanics, it is of great importance whether a quantum operator is Hermitian or not. A position operator and momentum operator along with an angular momentum operator are particularly important when we constitute Hamiltonian. Let $f$ and $g$ be arbitrary functions. Let us consider, e.g., a following inner product with the momentum operator.

$$\langle g \mid pf\rangle = \int_{a}^{b} g(x)^{*} \frac{\hbar}{i}\frac{\partial}{\partial x}[f(x)]\mathrm{d}x, \tag{1.150}$$

where the domain $[a, b]$ depends on a physical system; this can be either bounded or unbounded. Performing integration by parts, we have

$$\begin{aligned}\langle g \mid pf\rangle &= \frac{\hbar}{i}[g(x)^* f(x)]_a^b - \int_a^b \frac{\partial}{\partial x}[g(x)^*]\frac{\hbar}{i}f(x)\mathrm{d}x \\ &= \frac{\hbar}{i}[g(b)^* f(b) - g(a)^* f(a)] + \int_a^b \left[\frac{\hbar}{i}\frac{\partial}{\partial x}g(x)\right]^* f(x)\mathrm{d}x.\end{aligned} \tag{1.151}$$

If we require $f(b) = f(a)$ and $g(b) = g(a)$, the first term vanishes and we get

$$\langle g \mid pf\rangle = \int_a^b \left[\frac{\hbar}{i}\frac{\partial}{\partial x}g(x)\right]^* f(x)\mathrm{d}x = \langle pg \mid f\rangle. \tag{1.152}$$

Thus, as in the case of (1.120), the momentum operator $p$ is Hermitian. Note that a position operator $q$ of (1.142) is Hermitian as a priori assumption.

Meanwhile, the angular momentum operator $L_z$ is described in a polar coordinate as follows:

$$L_z = \frac{\hbar}{i}\frac{\partial}{\partial\phi}, \tag{1.153}$$

where $\phi$ is an azimuthal angle varying from 0 to $2\pi$. The notation and implication of $L_z$ will be mentioned in Chap. 3. Similarly as the above, we have

$$\langle g \mid L_z f\rangle = \frac{\hbar}{i}[g(2\pi)^* f(2\pi) - g(0)^* f(0)] + \int_0^{2\pi} \left[\frac{\hbar}{i}\frac{\partial}{\partial\phi}g(x)\right]^* f(x)\mathrm{d}\phi. \tag{1.154}$$

Requiring an arbitrary function $f$ to satisfy a BC $f(2\pi) = f(0)$, we reach

$$\langle g \mid L_z f\rangle = \langle L_z g \mid f\rangle. \tag{1.155}$$

Note that we must have the above BC, because $\phi = 0$ and $\phi = 2\pi$ are spatially the same point. Thus, we find that $L_z$ is Hermitian as well on this condition.

On the basis of aforementioned argument, let us proceed to quantum-mechanical studies of a harmonic oscillator. Regarding the angular momentum, we will study their basic properties in Chap. 3.

## Reference

1. Møller C (1952) The theory of relativity. Oxford University Press, London

# Chapter 2
# Quantum-Mechanical Harmonic Oscillator

Quantum-mechanical treatment of a harmonic oscillator has been a well-studied topic from the beginning of the history of quantum mechanics. This topic is a standard subject in classical mechanics as well. In this chapter, first we briefly survey characteristics of a classical harmonic oscillator. From a quantum-mechanical point of view, we deal with features of a harmonic oscillator through matrix representation. We define creation and annihilation operators using position and momentum operators. A Hamiltonian of the oscillator is described in terms of the creation and annihilation operators. This enables us to easily determine energy eigenvalues of the oscillator. As a result, energy eigenvalues are found to be positive definite. Meanwhile, we express the Schrödinger equation by the coordinate representation. We compare the results with those of the matrix representation and show that the two representations are mathematically equivalent. Thus, the treatment of the quantum-mechanical harmonic oscillator supplies us with a firm ground for studying basic concepts of the quantum mechanics.

## 2.1 Classical Harmonic Oscillator

Classical Newtonian equation of a one-dimensional harmonic oscillator is expressed as

$$m\frac{\mathrm{d}^2x(t)}{\mathrm{d}t^2} = -sx(t), \tag{2.1}$$

where $m$ is a mass of an oscillator and $s$ is a spring constant. Putting $s/m = \omega^2$, we have

$$\frac{\mathrm{d}^2x(t)}{\mathrm{d}t^2} + \omega^2 x(t) = 0. \tag{2.2}$$



S. Hotta, *Mathematical Physical Chemistry*,
https://doi.org/10.1007/978-981-10-7671-8_2

In (2.2), we set $\omega$ positive, namely

$$\omega = \sqrt{s/m}, \tag{2.3}$$

where $\omega$ is called an angular frequency of the oscillator.

If we replace $\omega^2$ with $\lambda$, we have formally the same equation as (1.61). Two linearly independent solutions of (2.2) are the same as before (see Example 1.1); we have $e^{i\omega t}$ and $e^{-i\omega t}(\omega \neq 0)$ as such. Note, however, that Example 1.2 we were dealing with a quantum state related to existence probability of a particle in a potential well. In (2.2), on the other hand, we are examining a position of harmonic oscillator undergoing a force of a spring. We are thus considering a different situation.

As a general solution, we have

$$x(t) = ae^{i\omega t} + be^{-i\omega t}, \tag{2.4}$$

where $a$ and $b$ are suitable constants. Let us consider BCs different from those of Examples 1.1 or 1.2 this time. That is, we set BCs such that

$$x(0) = 0 \text{ and } x'(0) = v_0 \ (v_0 > 0). \tag{2.5}$$

Notice that (2.5) gives initial conditions (ICs). Mathematically, ICs are included in BCs (see Chap. 8). From (2.4), we have

$$x(t) = a + b = 0 \text{ and } x'(0) = i\omega(a - b) = v_0. \tag{2.6}$$

Then, we get $a = -b = v_0/2i\omega$. Thus, we get a simple harmonic motion as a solution expressed as

$$x(t) = \frac{v_0}{2i\omega}\left(e^{i\omega x} - e^{-i\omega x}\right) = \frac{v_0}{\omega}\sin \omega t. \tag{2.7}$$

From this, we have

$$E = K + V = \frac{1}{2}mv_0^2. \tag{2.8}$$

In particular, if $v_0 = 0$, $x(t) \equiv 0$. This is a solution of (2.1) that has the meaning that the particle is eternally at rest. It is physically acceptable as well. Notice also that unlike Examples 1.1 and 1.2, the solution has been determined uniquely. This is due to the different BCs.

From a point of view of a mechanical system, mathematical formulation of the classical harmonic oscillator resembles that of electromagnetic fields confined within a cavity. We return this point later in Sect. 7.6.

## 2.2 Formulation Based on an Operator Method

Now let us return to our task to find quantum-mechanical solutions of a harmonic oscillator. Potential $V$ is given by

$$V(q) = \frac{1}{2}sq^2 = \frac{1}{2}m\omega^2 q^2, \tag{2.9}$$

where $q$ is used for a one-dimensional position coordinate. Then, we have a classical Hamiltonian $H$ expressed as

$$H = \frac{p^2}{2m} + V(q) = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 q^2. \tag{2.10}$$

Following the formulation of Sect. 1.2, the Schrödinger equation as an eigenvalue equation related to energy $E$ is described as

$$\begin{aligned} &H\psi(q) = E\psi(q) \text{ or} \\ &\quad -\frac{\hbar^2}{2m}\nabla^2\psi(q) + \frac{1}{2}m\omega^2 q^2\psi(q) = E\psi(q). \end{aligned} \tag{2.11}$$

This is a SOLDE and it is well known that the SOLDE can be solved by a power series expansion method.

In the present studies, however, let us first use an operator method to solve the eigenvalue Eq. (2.11) of a one-dimensional oscillator. To this end, we use a quantum-mechanical Hamiltonian where a momentum operator $p$ is explicitly represented. Thus, the Hamiltonian reads as

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 q^2. \tag{2.12}$$

The equation of (2.12) is formally the same as (2.10). Note, however, that in (2.12) $p$ and $q$ are expressed as quantum-mechanical operators.

As in (1.126), we first examine an expectation value $\langle H\rangle$ of $H$. It is given by

$$\begin{aligned} \langle H\rangle &= \langle\psi|H\psi\rangle = \left\langle\psi\left|\frac{p^2}{2m}\psi\right.\right\rangle + \left\langle\psi\left|\frac{1}{2}m\omega^2 q^2\psi\right.\right\rangle \\ &= \frac{1}{2m}\left\langle p^\dagger\psi|p\psi\right\rangle + \frac{1}{2}m\omega^2\left\langle q^\dagger\psi|q\psi\right\rangle = \frac{1}{2m}\langle p\psi|p\psi\rangle + \frac{1}{2}m\omega^2\langle q\psi|q\psi\rangle \\ &= \frac{1}{2m}\|p\psi\|^2 + \frac{1}{2}m\omega^2\|q\psi\|^2 \geq 0, \end{aligned} \tag{2.13}$$

where again we assumed that $|\psi\rangle$ has been normalized. In (2.13), we used the notation (1.126) and the fact that both $q$ and $p$ are Hermitian. In this situation, $\langle H\rangle$ takes a nonnegative value.

In (2.13), the equality holds if and only if $|p\psi\rangle = 0$ and $|q\psi\rangle = 0$. Let us specify a vector $|\psi_0\rangle$ that satisfies these conditions such that

$$|p\psi_0\rangle = 0 \text{ and } |q\psi_0\rangle = 0. \tag{2.14}$$

Multiplying $q$ from the left on the first equation of (2.14) and multiplying $p$ from the left on the second equation, we have

$$qp|\psi_0\rangle = 0 \text{ and } pq|\psi_0\rangle = 0. \tag{2.15}$$

Subtracting the second equation of (2.15) from the first equation, we get

$$(qp - pq)|\psi_0\rangle = i\hbar|\psi_0\rangle = 0, \tag{2.16}$$

where with the first equality we used (1.140). Therefore, we would have $|\psi_0(q)\rangle \equiv 0$. This leads to the relations (2.14). That is, if and only if $|\psi_0(q)\rangle \equiv 0$, $\langle H\rangle = 0$. But, since it has no physical meaning, $|\psi_0(q)\rangle \equiv 0$ must be rejected as unsuitable for the solution of (2.11). Regarding a physically acceptable solution of (2.13), $\langle H\rangle$ must take a *positive-definite* value accordingly. Thus, on the basis of the canonical commutation relation, we restrict the range of the expectation values.

Instead of directly dealing with (2.12), it is well known to introduce following operators [1]:

$$a \equiv \sqrt{\frac{m\omega}{2\hbar}}q + \frac{i}{\sqrt{2m\hbar\omega}}p \tag{2.17}$$

and its adjoint (complex conjugate) operator

$$a^\dagger = \sqrt{\frac{m\omega}{2\hbar}}q - \frac{i}{\sqrt{2m\hbar\omega}}p. \tag{2.18}$$

Notice here again that both $q$ and $p$ are Hermitian. Using a matrix representation for (2.17) and (2.18), we have

$$\begin{pmatrix} a \\ a^\dagger \end{pmatrix} = \begin{pmatrix} \sqrt{\frac{m\omega}{2\hbar}} & \frac{i}{\sqrt{2m\hbar\omega}} \\ \sqrt{\frac{m\omega}{2\hbar}} & -\frac{i}{\sqrt{2m\hbar\omega}} \end{pmatrix} \begin{pmatrix} q \\ p \end{pmatrix} \tag{2.19}$$

Then, we have

$$\begin{aligned} a^\dagger a &= (2m\hbar\omega)^{-1}(m\omega q - ip)(m\omega q + ip) \\ &= (2m\hbar\omega)^{-1}\left[m^2\omega^2 q^2 + p^2 + im\omega(qp - pq)\right] \\ &= (\hbar\omega)^{-1}\left[\frac{1}{2}m\omega^2 q^2 + \frac{1}{2m}p^2 + \frac{1}{2}i\omega i\hbar\right] = (\hbar\omega)^{-1}\left(H - \frac{1}{2}\hbar\omega\right), \end{aligned} \tag{2.20}$$

where the second last equality comes from (1.140). Rewriting (2.20), we get

$$H = \hbar\omega a^\dagger a + \frac{1}{2}\hbar\omega. \tag{2.21}$$

Similarly, we get

$$H = \hbar\omega a a^\dagger - \frac{1}{2}\hbar\omega. \tag{2.22}$$

Subtracting (2.22) from (2.21), we have

$$0 = \hbar\omega a^\dagger a - \hbar\omega a a^\dagger + \hbar\omega. \tag{2.23}$$

That is,

$$\left[a, a^\dagger\right] = 1 \text{ or } \left[a, a^\dagger\right] = E. \tag{2.24}$$

Furthermore, using (2.21), we have

$$\left[H, a^\dagger\right] = \hbar\omega\left[a^\dagger a + \frac{1}{2}, a^\dagger\right] = \hbar\omega\left(a^\dagger a a^\dagger - a^\dagger a^\dagger a\right) = \hbar\omega a^\dagger\left[a, a^\dagger\right] = \hbar\omega a^\dagger. \tag{2.25}$$

Similarly, we get

$$[H, a] = -\hbar\omega a. \tag{2.26}$$

Next, let us calculate an expectation value of $H$. Using a normalized function $|\psi\rangle$, from (2.21), we have

$$\begin{aligned}\langle\psi|H|\psi\rangle &= \langle\psi|\hbar\omega a^\dagger a + \frac{1}{2}\hbar\omega|\psi\rangle = \hbar\omega\langle\psi|a^\dagger a|\psi\rangle + \frac{1}{2}\hbar\omega\langle\psi|\psi\rangle \\ &= \hbar\omega\langle a\psi|a\psi\rangle + \frac{1}{2}\hbar\omega = \hbar\omega\|a\psi\|^2 + \frac{1}{2}\hbar\omega \geq \frac{1}{2}\hbar\omega.\end{aligned} \tag{2.27}$$

Thus, the expectation value is equal to or larger than $\frac{1}{2}\hbar\omega$. This is consistent with that an energy eigenvalue is positive definite as mentioned above. Equation (2.27) also tells us that if we have

$$|a\psi_0\rangle = 0, \tag{2.28}$$

we get

$$\langle\psi_0|H|\psi_0\rangle = \frac{1}{2}\hbar\omega. \tag{2.29}$$

Equation (2.29) means that the smallest expectation value is $\frac{1}{2}\hbar\omega$ on the condition of (2.28). On the same condition, using (2.21), we have

$$H|\psi_0\rangle = \hbar\omega a^\dagger a|\psi_0\rangle + \frac{1}{2}\hbar\omega|\psi_0\rangle = \frac{1}{2}\hbar\omega|\psi_0\rangle. \tag{2.30}$$

Thus, $|\psi_0\rangle$ is an eigenfunction corresponding to an eigenvalue $\frac{1}{2}\hbar\omega \equiv E_0$, which is identical with the smallest expectation value of (2.29). Since this is the lowest eigenvalue, $|\psi_0\rangle$ is said to be a ground state. We ensure later that $|\psi_0\rangle$ is certainly an eligible function for a ground state.

The above method is consistent with the variational principle [2] which stipulates that under appropriate BCs an expectation value of Hamiltonian estimated with any arbitrary function is always larger than or equal to the smallest eigenvalue corresponding to the ground state.

Next, let us evaluate energy eigenvalues of the oscillator. First we have

$$H|\psi_0\rangle = \frac{1}{2}\hbar\omega|\psi_0\rangle = E_0|\psi_0\rangle. \tag{2.31}$$

Operating $a^\dagger$ on both sides of (2.31), we have

$$a^\dagger H|\psi_0\rangle = a^\dagger E_0|\psi_0\rangle. \tag{2.32}$$

Meanwhile, using (2.25), we have

$$a^\dagger H|\psi_0\rangle = \left(Ha^\dagger - \hbar\omega a^\dagger\right)|\psi_0\rangle. \tag{2.33}$$

Equating RHSs of (2.32) and (2.33), we get

$$Ha^\dagger|\psi_0\rangle = (E_0 + \hbar\omega)a^\dagger|\psi_0\rangle. \tag{2.34}$$

This implies that $a^\dagger|\psi_0\rangle$ belongs to an eigenvalue $(E_0 + \hbar\omega)$, which is larger than $E_0$ as expected. Again multiplying $a^\dagger$ on both sides of (2.34) from the left and using (2.25), we get

$$H(a^\dagger)^2|\psi_0\rangle = (E_0 + 2\hbar\omega)(a^\dagger)^2|\psi_0\rangle. \tag{2.35}$$

This implies that $(a^\dagger)^2|\psi_0\rangle$ belongs to an eigenvalue $(E_0 + 2\hbar\omega)$. Thus, repeatedly taking the above procedures, we get

$$H(a^\dagger)^n|\psi_0\rangle = (E_0 + n\hbar\omega)(a^\dagger)^n|\psi_0\rangle. \tag{2.36}$$

Thus, $(a^\dagger)^n|\psi_0\rangle$ belongs to an eigenvalue

$$E_n \equiv (E_0 + n\hbar\omega) = \left(n + \frac{1}{2}\right)\hbar\omega, \tag{2.37}$$

where $E_n$ denotes an energy eigenvalue of the $n$th excited state. The energy eigenvalues are plotted in Fig. 2.1.

Our next task is to seek normalized eigenvectors of the $n$th excited state. Let $c_n$ be a normalization constant of that state. That is, we have

$$|\psi_n\rangle = c_n(a^\dagger)^n|\psi_0\rangle, \tag{2.38}$$

where $|\psi_n\rangle$ is a normalized eigenfunction of the $n$th excited state. To determine $c_n$, let us calculate $a|\psi_n\rangle$. This includes a factor $a(a^\dagger)^n$. We have

$$\begin{aligned}
a(a^\dagger)^n &= \left(aa^\dagger - a^\dagger a\right)(a^\dagger)^{n-1} + a^\dagger a(a^\dagger)^{n-1} \\
&= \left[a, a^\dagger\right](a^\dagger)^{n-1} + a^\dagger a(a^\dagger)^{n-1} = (a^\dagger)^{n-1} + a^\dagger a(a^\dagger)^{n-1} \\
&= (a^\dagger)^{n-1} + a^\dagger\left[a, a^\dagger\right](a^\dagger)^{n-2} + (a^\dagger)^2 a(a^\dagger)^{n-2} \\
&= 2(a^\dagger)^{n-1} + (a^\dagger)^2 a(a^\dagger)^{n-2} \\
&= 2(a^\dagger)^{n-1} + (a^\dagger)^2\left[a, a^\dagger\right](a^\dagger)^{n-3} + (a^\dagger)^3 a(a^\dagger)^{n-3} \\
&= 3(a^\dagger)^{n-1} + (a^\dagger)^3 a(a^\dagger)^{n-3} \\
&= \cdots.
\end{aligned} \tag{2.39}$$

In the above procedures, we used $\left[a, a^\dagger\right] = 1$. What is implied in (2.39) is that a coefficient of $(a^\dagger)^{n-1}$ increased one by one with $a$ transferred toward the right one by one in the second term of RHS. Notice that in the second term $a$ is sandwiched such that $(a^\dagger)^m a(a^\dagger)^{n-m}$ $(m = 1, 2, \ldots)$. Finally, we have

$$a(a^\dagger)^n = n(a^\dagger)^{n-1} + (a^\dagger)^n a. \tag{2.40}$$

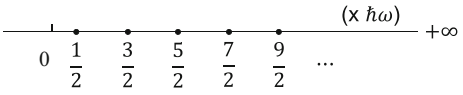


**Fig. 2.1** Energy eigenvalues of a quantum-mechanical harmonic oscillator on a real axis

Thus, we get

$$\begin{aligned} a|\psi_n\rangle &= c_n a(a^\dagger)^n|\psi_0\rangle = c_n\Big[n(a^\dagger)^{n-1} + (a^\dagger)^n a\Big]|\psi_0\rangle = c_n n(a^\dagger)^{n-1}|\psi_0\rangle \\ &= n\frac{c_n}{c_{n-1}}c_{n-1}(a^\dagger)^{n-1}|\psi_0\rangle = n\frac{c_n}{c_{n-1}}|\psi_{n-1}\rangle, \end{aligned} \tag{2.41}$$

where the third equality comes from (2.28).

Next, operating $a$ on (2.40), we get

$$\begin{aligned} a^2(a^\dagger)^n &= na(a^\dagger)^{n-1} + a(a^\dagger)^n a \\ &= n\Big[(n-1)(a^\dagger)^{n-2} + (a^\dagger)^{n-1}a\Big] + a(a^\dagger)^n a \\ &= n(n-1)(a^\dagger)^{n-2} + n(a^\dagger)^{n-1}a + a(a^\dagger)^n a. \end{aligned} \tag{2.42}$$

Operating another $a$ on (2.42), we get

$$\begin{aligned} a^3(a^\dagger)^n &= n(n-1)a(a^\dagger)^{n-2} + na(a^\dagger)^{n-1}a + a^2(a^\dagger)^n a \\ &= n(n-1)\Big[(n-2)(a^\dagger)^{n-3} + (a^\dagger)^{n-2}a\Big] + na(a^\dagger)^{n-1}a + a^2(a^\dagger)^n a \\ &= n(n-1)(n-2)(a^\dagger)^{n-3} + n(n-1)(a^\dagger)^{n-2}a + na(a^\dagger)^{n-1}a + a^2(a^\dagger)^n a \\ &= \cdots. \end{aligned} \tag{2.43}$$

To generalize the above procedures, operating $a$ on (2.40) $m(<n)$ times, we get

$$a^m(a^\dagger)^n = n(n-1)(n-2)\ldots(n-m+1)(a^\dagger)^{n-m} + f\left(a, a^\dagger\right)a, \tag{2.44}$$

where $m<n$ and $f\left(a, a^\dagger\right)$ is a polynomial of $a^\dagger$ that has a power of $a$ as a coefficient. Further operating $\langle\psi_0|$ and $|\psi_0\rangle$ from the left and right on both sides of (2.44), respectively, we have

$$\begin{aligned} \langle\psi_0|a^m(a^\dagger)^n|\psi_0\rangle &= n(n-1)(n-2)\ldots(n-m+1)\langle\psi_0|(a^\dagger)^{n-m}|\psi_0\rangle \\ &\quad + \langle\psi_0|f\left(a, a^\dagger\right)a|\psi_0\rangle \\ &= n(n-1)(n-2)\ldots(n-m+1)\langle\psi_0|(a^\dagger)^{n-m}|\psi_0\rangle. \end{aligned} \tag{2.45}$$

Note that in (2.45) $\langle\psi_0|f\left(a, a^\dagger\right)a|\psi_0\rangle$ vanishes because of (2.28).

Meanwhile, taking adjoint of (2.28), we have

$$\langle\psi_0|a^\dagger = 0. \tag{2.46}$$

Operating $a^\dagger$ $(n-m-1)$ times from the right of LHS of (2.46), we have

$$\langle\psi_0|(a^\dagger)^{n-m} = 0.$$

Further operating $|\psi_0\rangle$ from the right of the above equation, we get

$$\langle\psi_0|(a^\dagger)^{n-m}|\psi_0\rangle = 0.$$

Therefore, from (2.45), we get

$$\langle\psi_0|a^m(a^\dagger)^n|\psi_0\rangle = 0. \tag{2.47}$$

Taking adjoint of (2.47) once again, we have

$$\langle\psi_0|a^n(a^\dagger)^m|\psi_0\rangle = 0. \tag{2.48}$$

Equation (2.48) can be obtained by repeatedly using (1.117). From (2.45) and (2.48), we get

$$\langle\psi_0|a^m(a^\dagger)^n|\psi_0\rangle = 0, \text{ when } m \neq n. \tag{2.49}$$

If $m = n$, from (2.45), we get

$$\langle\psi_0|a^n(a^\dagger)^n|\psi_0\rangle = n!\langle\psi_0|\psi_0\rangle. \tag{2.50}$$

If we assume that $|\psi_0\rangle$ is normalized; i.e., $\langle\psi_0|\psi_0\rangle = 1$, (2.50) is expressed as

$$\langle\psi_0|a^n(a^\dagger)^n|\psi_0\rangle = n!. \tag{2.51}$$

From (2.51), if we put

$$|\psi_n\rangle = \frac{1}{\sqrt{n!}}(a^\dagger)^n|\psi_0\rangle, \tag{2.52}$$

then we have

$$\langle\psi_n| = \frac{1}{\sqrt{n!}}\langle\psi_0|a^n.$$

Thus, from (2.49) and (2.52), we get

$$\langle\psi_m|\psi_n\rangle = \delta_{mn}. \tag{2.53}$$

At the same time, for $c_n$ of (2.38), we get

$$c_n = \frac{1}{\sqrt{n!}}. \tag{2.54}$$

Notice here that an undetermined phase factor $\mathrm{e}^{i\theta}$ ($\theta$ : real) is intended such that

$$c_n = \frac{1}{\sqrt{n!}}\mathrm{e}^{i\theta}.$$

But, $\mathrm{e}^{i\theta}$ is usually omitted for the sake of simplicity. Thus, we have constructed a series of orthonormal eigenfunctions $|\psi_n\rangle$.

Furthermore, using (2.41) and (2.54), we get

$$a|\psi_n\rangle = \sqrt{n}|\psi_{n-1}\rangle. \tag{2.55}$$

From (2.36), we get

$$H|\psi_n\rangle = (E_0 + n\hbar\omega)|\psi_n\rangle. \tag{2.56}$$

Meanwhile, from (2.21), we have

$$H|\psi_n\rangle = \left(\hbar\omega a^\dagger a + E_0\right)|\psi_n\rangle. \tag{2.57}$$

Equating RHSs of (2.56) and (2.57), we get

$$a^\dagger a|\psi_n\rangle = n|\psi_n\rangle. \tag{2.58}$$

Thus, we find that an integer $n$ is an eigenvalue of $a^\dagger a$ when it is evaluated with respect to $|\psi_n\rangle$. For this reason, $a^\dagger a$ is called a number operator. Notice that $a^\dagger a$ is Hermitian because we have

$$(a^\dagger a)^\dagger = a^\dagger (a^\dagger)^\dagger = a^\dagger a, \tag{2.59}$$

where we used (1.115) and (1.117).

Moreover, from (2.55) and (2.58),

$$a^\dagger a|\psi_n\rangle = \sqrt{n}a^\dagger|\psi_{n-1}\rangle = n|\psi_n\rangle. \tag{2.60}$$

Thus,

$$a^\dagger|\psi_{n-1}\rangle = \sqrt{n}|\psi_n\rangle. \tag{2.61}$$

Or replacing $n$ with $n+1$, we get

$$a^{\dagger}|\psi_n\rangle = \sqrt{n+1}\,|\psi_{n+1}\rangle. \tag{2.62}$$

As implied in (2.55) and (2.62), we find that operating $a$ on $|\psi_n\rangle$ lowers an energy level by one and that operating $a^{\dagger}$ on $|\psi_n\rangle$ raises an energy level by one. For this reason, $a$ and $a^{\dagger}$ are said to be an annihilation operator and creation operator, respectively.

## 2.3 Matrix Representation of Physical Quantities

Equations (2.55) and (2.62) clearly represent the relationship between an operator and eigenfunction (or eigenvector). The relationship is characterized by

$$(\text{Matrix}) \times (\text{Vector}) = (\text{Vector}). \tag{2.63}$$

Thus, we are now in a position to construct this relation using matrices. From (2.53), we should be able to construct basis vectors using a *column vector* such that

$$|\psi_0\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ \vdots \end{pmatrix}, |\psi_1\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ \vdots \end{pmatrix}, |\psi_2\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix}, \cdots. \tag{2.64}$$

Notice that these vectors form a vector space of an infinite dimension. The orthonormal relation (2.53) can easily be checked. We represent $a$ and $a^{\dagger}$ so that (2.55) and (2.62) can be satisfied. We obtain

$$a = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & \cdots \\ 0 & 0 & \sqrt{2} & 0 & 0 & \cdots \\ 0 & 0 & 0 & \sqrt{3} & 0 & \cdots \\ 0 & 0 & 0 & 0 & 2 & \cdots \\ 0 & 0 & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \tag{2.65}$$

Similarly,

$$a^\dagger = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & \cdots \\ 1 & 0 & 0 & 0 & 0 & \cdots \\ 0 & \sqrt{2} & 0 & 0 & 0 & \cdots \\ 0 & 0 & \sqrt{3} & 0 & 0 & \cdots \\ 0 & 0 & 0 & 2 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \tag{2.66}$$

Note that neither $a$ nor $a^\dagger$ is Hermitian.

Since a determinant of the matrix of (2.19) is $-i/\hbar \neq 0$, using its inverse matrix we get

$$\begin{pmatrix} q \\ p \end{pmatrix} = \begin{pmatrix} \sqrt{\frac{\hbar}{2m\omega}} & \sqrt{\frac{\hbar}{2m\omega}} \\ \frac{1}{i}\sqrt{\frac{m\hbar\omega}{2}} & -\frac{1}{i}\sqrt{\frac{m\hbar\omega}{2}} \end{pmatrix} \begin{pmatrix} a \\ a^\dagger \end{pmatrix}. \tag{2.67}$$

That is, we have

$$q = \sqrt{\frac{\hbar}{2m\omega}}\left(a + a^\dagger\right) \text{ and } p = \frac{1}{i}\sqrt{\frac{m\hbar\omega}{2}}\left(a - a^\dagger\right). \tag{2.68}$$

With the inverse matrix, we will deal with it in Part III. Note that $q$ and $p$ are both Hermitian. Inserting (2.65) and (2.66) into (2.68), we get

$$q = \sqrt{\frac{\hbar}{2m\omega}} \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & \cdots \\ 1 & 0 & \sqrt{2} & 0 & 0 & \cdots \\ 0 & \sqrt{2} & 0 & \sqrt{3} & 0 & \cdots \\ 0 & 0 & \sqrt{3} & 0 & 2 & \cdots \\ 0 & 0 & 0 & 2 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \tag{2.69}$$

$$p = \sqrt{\frac{m\hbar\omega}{2}}i \begin{pmatrix} 0 & -1 & 0 & 0 & 0 & \cdots \\ 1 & 0 & -\sqrt{2} & 0 & 0 & \cdots \\ 0 & \sqrt{2} & 0 & -\sqrt{3} & 0 & \cdots \\ 0 & 0 & \sqrt{3} & 0 & -2 & \cdots \\ 0 & 0 & 0 & 2 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \tag{2.70}$$

Equations (2.69) and (2.70) obviously show that $q$ and $p$ are Hermitian. We can derive various physical quantities from these equations. For instance, following matrix algebra Hamiltonian $H$ can readily be calculated. The result is expressed as

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 q^2 = \frac{\hbar\omega}{2}\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 3 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 5 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 7 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 9 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \tag{2.71}$$

Looking at (2.69–2.71), we immediately realize that although neither $q$ or $p$ is diagonalized, $H$ is diagonalized. The matrix representation of (2.71) is said to be a representation that diagonalizes $H$. This representation is of great practical use. In fact, using (2.64), we get, e.g.,

$$\begin{aligned} H|\psi_2\rangle &= \frac{\hbar\omega}{2}\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 3 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 5 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 7 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 9 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix} = \frac{\hbar\omega}{2}\begin{pmatrix} 0 \\ 0 \\ 5 \\ 0 \\ 0 \\ \vdots \end{pmatrix} = \frac{5\hbar\omega}{2}\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix} \\ &= \frac{5\hbar\omega}{2}|\psi_2\rangle = \left(\frac{1}{2}+2\right)\hbar\omega|\psi_2\rangle. \end{aligned} \tag{2.72}$$

This clearly means that the second-excited state $|\psi_2\rangle$ has an eigenenergy $\left(\frac{1}{2}+2\right)\hbar\omega$. More generally, we find that $|\psi_n\rangle$ has an eigenenergy $\left(\frac{1}{2}+n\right)\hbar\omega$ as already shown in (2.36) and (2.56).

Furthermore, let us confirm the canonical commutation relation of (1.140). Using (2.68), we have

$$\begin{aligned} qp - pq &= \frac{1}{i}\sqrt{\frac{\hbar}{2m\omega}}\sqrt{\frac{m\hbar\omega}{2}}\left[\left(a+a^\dagger\right)\left(a-a^\dagger\right) - \left(a-a^\dagger\right)\left(a+a^\dagger\right)\right] \\ &= \frac{\hbar}{2i}\cdot(-2)\cdot\left(aa^\dagger - a^\dagger a\right) = i\hbar\left[a, a^\dagger\right] = i\hbar = i\hbar\widehat{E}, \end{aligned} \tag{2.73}$$

where with the second last equality we used (2.24) and the identity matrix of an infinite dimension $\widehat{E}$ is given by

$$\widehat{E} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 1 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 1 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \tag{2.74}$$

Thus, the canonical commutation relation holds with the quantum-mechanical harmonic oscillator. This can be confirmed directly from (2.69) and (2.70). The proof is left for readers as an exercise.

## 2.4 Coordinate Representation of Schrödinger Equation

The Schrödinger equation has been given in (1.55) or (2.11) as a SOLDE form. In contrast to the matrix representation, (1.55) and (2.11) are said to be coordinate representation of Schrödinger equation. Now, let us derive a coordinate representation of (2.28) to obtain an analytical solution.

On the basis of (1.31) and (2.17), $a$ is expressed as

$$a = \sqrt{\frac{m\omega}{2\hbar}}q + \frac{i}{\sqrt{2m\hbar\omega}}p = \sqrt{\frac{m\omega}{2\hbar}}q + \frac{i}{\sqrt{2m\hbar\omega}}\frac{\hbar}{i}\frac{\partial}{\partial q} = \sqrt{\frac{m\omega}{2\hbar}}q + \sqrt{\frac{\hbar}{2m\omega}}\frac{\partial}{\partial q}. \tag{2.75}$$

Thus, (2.28) reads as a following first-order linear differential equation (FOLDE):

$$\left(\sqrt{\frac{m\omega}{2\hbar}}q + \sqrt{\frac{\hbar}{2m\omega}}\frac{\partial}{\partial q}\right)\psi_0(q) = 0. \tag{2.76}$$

Or

$$\sqrt{\frac{m\omega}{2\hbar}}q\psi_0(q) + \sqrt{\frac{\hbar}{2m\omega}}\frac{\partial\psi_0(q)}{\partial q} = 0. \tag{2.77}$$

This is further reduced to

$$\frac{\partial\psi_0(q)}{\partial q} + \frac{m\omega}{\hbar}q\psi_0(q) = 0. \tag{2.78}$$

From this FOLDE form, we anticipate the following solution:

$$\psi_0(q) = N_0 \mathrm{e}^{-\alpha q^2}, \tag{2.79}$$

where $N_0$ is a normalization constant and $\alpha$ is a constant coefficient. Putting (2.79) into (2.78), we have

$$\left(-2\alpha q + \frac{m\omega}{\hbar}q\right)N_0\mathrm{e}^{-\alpha q^2} = 0. \tag{2.80}$$

Hence, we get

$$\alpha = \frac{m\omega}{2\hbar}. \tag{2.81}$$

Thus,

$$\psi_0(q) = N_0 \mathrm{e}^{-\frac{m\omega}{2\hbar}q^2}. \tag{2.82}$$

The constant $N_0$ can be determined by the normalization condition given by

$$\int_{-\infty}^{\infty} |\psi_0(q)|^2 \mathrm{d}q = 1 \text{ or } N_0^2 \int_{-\infty}^{\infty} \mathrm{e}^{-\frac{m\omega}{\hbar}q^2} \mathrm{d}q = 1. \tag{2.83}$$

Recalling the following formula:

$$\int_{-\infty}^{\infty} \mathrm{e}^{-cq^2} \mathrm{d}q = \sqrt{\frac{\pi}{c}} \ (c > 0), \tag{2.84}$$

we have

$$\int_{-\infty}^{\infty} \mathrm{e}^{-\frac{m\omega}{\hbar}q^2} \mathrm{d}q = \sqrt{\frac{\pi\hbar}{m\omega}}. \tag{2.85}$$

To get (2.84), putting $I \equiv \int_{-\infty}^{\infty} \mathrm{e}^{-cq^2} \mathrm{d}q$, we have

$$\begin{aligned} I^2 &= \left( \int_{-\infty}^{\infty} \mathrm{e}^{-cq^2} \mathrm{d}q \right) \left( \int_{-\infty}^{\infty} \mathrm{e}^{-cs^2} \mathrm{d}s \right) = \int_{-\infty}^{\infty} \mathrm{e}^{-c\left(q^2+s^2\right)} \mathrm{d}q\mathrm{d}s \\ &= \int_0^{\infty} \mathrm{e}^{-cr^2} r\mathrm{d}r \int_0^{2\pi} \mathrm{d}\theta = \frac{1}{2} \int_0^{\infty} \mathrm{e}^{-cR} \mathrm{d}R \int_0^{2\pi} \mathrm{d}\theta = \frac{\pi}{c}, \end{aligned} \tag{2.86}$$

where with the third equality we converted two-dimensional Cartesian coordinate to polar coordinate; take $q = r\cos\theta$, $s = r\sin\theta$ and convert an infinitesimal area element $\mathrm{d}q\mathrm{d}s$ to $\mathrm{d}r \cdot r\mathrm{d}\theta$. With the second last equality of (2.86), we used the variable transformation of $r^2 \to R$. Hence, we get $I = \sqrt{\frac{\pi}{c}}$.

Thus, we get

$$N_0 = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \text{ and } \psi_0(q) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \mathrm{e}^{-\frac{m\omega}{2\hbar}q^2}. \tag{2.87}$$

Also, we have

$$a^\dagger = \sqrt{\frac{m\omega}{2\hbar}}q - \frac{i}{\sqrt{2m\hbar\omega}}p = \sqrt{\frac{m\omega}{2\hbar}}q - \frac{i}{\sqrt{2m\hbar\omega}}\frac{\hbar}{i}\frac{\partial}{\partial q} = \sqrt{\frac{m\omega}{2\hbar}}q - \sqrt{\frac{\hbar}{2m\omega}}\frac{\partial}{\partial q}. \tag{2.88}$$

From (2.52), we get

$$\begin{aligned}\psi_n(q) &= \frac{1}{\sqrt{n!}}(a^\dagger)^n|\psi_0\rangle = \frac{1}{\sqrt{n!}}\left(\sqrt{\frac{m\omega}{2\hbar}}q - \sqrt{\frac{\hbar}{2m\omega}}\frac{\partial}{\partial q}\right)^n\psi_0(q) \\ &= \frac{1}{\sqrt{n!}}\left(\frac{m\omega}{2\hbar}\right)^{n/2}\left(q - \frac{\hbar}{m\omega}\frac{\partial}{\partial q}\right)^n\psi_0(q).\end{aligned} \tag{2.89}$$

Putting

$$\beta \equiv \sqrt{m\omega/\hbar} \text{ and } \xi = \beta q, \tag{2.90}$$

we rewrite (2.89) as

$$\begin{aligned}\psi_n(q) = \psi_n\left(\frac{\xi}{\beta}\right) &= \frac{1}{\sqrt{n!}}\left(\frac{m\omega}{2\hbar\beta^2}\right)^{n/2}\left(\xi - \frac{\partial}{\partial\xi}\right)^n\psi_0(q) \\ &= \frac{1}{\sqrt{n!}}\left(\frac{1}{2}\right)^{n/2}\left(\xi - \frac{\partial}{\partial\xi}\right)^n\psi_0(q) \\ &= \sqrt{\frac{1}{2^n n!}}\left(\frac{m\omega}{\pi\hbar}\right)^{1/4}\left(\xi - \frac{\partial}{\partial\xi}\right)^n \mathrm{e}^{-\frac{1}{2}\xi^2}.\end{aligned} \tag{2.91}$$

Comparing (2.81) and (2.90), we have

$$\alpha = \frac{\beta^2}{2}. \tag{2.92}$$

Moreover, putting

$$N_n \equiv \sqrt{\frac{1}{2^n n!}}\left(\frac{m\omega}{\pi\hbar}\right)^{1/4} = \sqrt{\frac{\beta}{\pi^{1/2}2^n n!}}, \tag{2.93}$$

we get

$$\psi_n(\xi/\beta) = N_n\left(\xi - \frac{\partial}{\partial\xi}\right)^n \mathrm{e}^{-\frac{1}{2}\xi^2}. \tag{2.94}$$

We have to normalize (2.94) with respect to a variable $\xi$. Since $\psi_n(q)$ has already been normalized as in (2.53), we have

$$\int_{-\infty}^{\infty} |\psi_n(q)|^2 \mathrm{d}q = 1. \tag{2.95}$$

Changing a variable $q$ to $\xi$, we have

$$\frac{1}{\beta} \int_{-\infty}^{\infty} |\psi_n(\xi/\beta)|^2 \mathrm{d}\xi = 1. \tag{2.96}$$

Let us define $\widetilde{\psi_n}(\xi)$ as being normalized with $\xi$. In other words, $\psi_n(q)$ is converted to $\widetilde{\psi_n}(\xi)$ by means of variable transformation and concomitant change in normalization condition. Then, we have

$$\int_{-\infty}^{\infty} \left|\widetilde{\psi_n}(\xi)\right|^2 \mathrm{d}\xi = 1. \tag{2.97}$$

Comparing (2.96) and (2.97), if we define $\widetilde{\psi_n}(\xi)$ as

$$\widetilde{\psi_n}(\xi) \equiv \sqrt{\frac{1}{\beta}} \psi_n(\xi/\beta), \tag{2.98}$$

$\widetilde{\psi_n}(\xi)$ should be a proper normalized function. Thus, we get

$$\widetilde{\psi_n}(\xi) = \widetilde{N_n} \left(\xi - \frac{\partial}{\partial \xi}\right)^n \mathrm{e}^{-\frac{1}{2}\xi^2} \text{ with } \widetilde{N_n} \equiv \sqrt{\frac{1}{\pi^{1/2} 2^n n!}}. \tag{2.99}$$

Meanwhile, according to a theory of classical orthogonal polynomial, the Hermite polynomials $H_n(x)$ are defined as [3]

$$H_n(x) \equiv (-1)^n \mathrm{e}^{x^2} \frac{\mathrm{d}^n}{\mathrm{d}x^n} \left(\mathrm{e}^{-x^2}\right) (n \geq 0), \tag{2.100}$$

where $H_n(x)$ is a $n$th-order polynomial. We wish to show the following relation on the basis of mathematical induction:

$$\widetilde{\psi_n}(\xi) = \widetilde{N_n} H_n(\xi) \mathrm{e}^{-\frac{1}{2}\xi^2}. \tag{2.101}$$

Comparing (2.87), (2.98), and (2.99), we make sure that (2.101) holds with $n = 0$. When $n = 1$, from (2.99), we have

$$\begin{aligned}\widetilde{\psi_1}(\xi) &= \widetilde{N_1}\left(\xi - \frac{\partial}{\partial\xi}\right)\mathrm{e}^{-\frac{1}{2}\xi^2} = \widetilde{N_1}\left[\xi\mathrm{e}^{-\frac{1}{2}\xi^2} - (-\xi)\mathrm{e}^{-\frac{1}{2}\xi^2}\right] = \widetilde{N_1}\cdot 2\xi\mathrm{e}^{-\frac{1}{2}\xi^2} \\ &= \widetilde{N_1}\left[(-1)^1\mathrm{e}^{\xi^2}\frac{\mathrm{d}}{\mathrm{d}\xi}\left(\mathrm{e}^{-\xi^2}\right)\right]\mathrm{e}^{-\frac{1}{2}\xi^2} = \widetilde{N_1}H_1(\xi)\mathrm{e}^{-\frac{1}{2}\xi^2}.\end{aligned} \tag{2.102}$$

Then, (2.101) holds with $n = 1$ as well.

Next, from supposition of mathematical induction, we assume that (2.101) holds with $n$. Then, we have

$$\begin{aligned}\widetilde{\psi_{n+1}}(\xi) &= \widetilde{N_{n+1}}\left(\xi - \frac{\partial}{\partial\xi}\right)^{n+1}\mathrm{e}^{-\frac{1}{2}\xi^2} = \frac{1}{\sqrt{2(n+1)}}\left(\xi - \frac{\partial}{\partial\xi}\right)\widetilde{N_n}\left(\xi - \frac{\partial}{\partial\xi}\right)^{n}\mathrm{e}^{-\frac{1}{2}\xi^2} \\ &= \frac{1}{\sqrt{2(n+1)}}\left(\xi - \frac{\partial}{\partial\xi}\right)\left[\widetilde{N_n}H_n(x)\mathrm{e}^{-\frac{1}{2}\xi^2}\right] \\ &= \frac{1}{\sqrt{2(n+1)}}\widetilde{N_n}\left(\xi - \frac{\partial}{\partial\xi}\right)\left\{\left[(-1)^n\mathrm{e}^{\xi^2}\frac{\mathrm{d}^n}{\mathrm{d}\xi^n}\left(\mathrm{e}^{-\xi^2}\right)\right]\mathrm{e}^{-\frac{1}{2}\xi^2}\right\} \\ &= \frac{1}{\sqrt{2(n+1)}}\widetilde{N_n}(-1)^n\left(\xi - \frac{\partial}{\partial\xi}\right)\left[\mathrm{e}^{\frac{1}{2}\xi^2}\frac{\mathrm{d}^n}{\mathrm{d}\xi^n}\left(\mathrm{e}^{-\xi^2}\right)\right] \\ &= \frac{1}{\sqrt{2(n+1)}}\widetilde{N_n}(-1)^n\left\{\xi\mathrm{e}^{\frac{1}{2}\xi^2}\frac{\mathrm{d}^n}{\mathrm{d}\xi^n}\left(\mathrm{e}^{-\xi^2}\right) - \frac{\partial}{\partial\xi}\left[\mathrm{e}^{\frac{1}{2}\xi^2}\frac{\mathrm{d}^n}{\mathrm{d}\xi^n}\left(\mathrm{e}^{-\xi^2}\right)\right]\right\} \\ &= \frac{1}{\sqrt{2(n+1)}}\widetilde{N_n}(-1)^n\left\{\xi\mathrm{e}^{\frac{1}{2}\xi^2}\frac{\mathrm{d}^n}{\mathrm{d}\xi^n}\left(\mathrm{e}^{-\xi^2}\right) - \xi\mathrm{e}^{\frac{1}{2}\xi^2}\frac{\mathrm{d}^n}{\mathrm{d}\xi^n}\left(\mathrm{e}^{-\xi^2}\right) - \mathrm{e}^{\frac{1}{2}\xi^2}\frac{\mathrm{d}^{n+1}}{\mathrm{d}\xi^{n+1}}\left(\mathrm{e}^{-\xi^2}\right)\right\} \\ &= \frac{1}{\sqrt{2(n+1)}}\widetilde{N_n}(-1)^{n+1}\mathrm{e}^{\frac{1}{2}\xi^2}\frac{\mathrm{d}^{n+1}}{\mathrm{d}\xi^{n+1}}\left(\mathrm{e}^{-\xi^2}\right) \\ &= \widetilde{N_{n+1}}\left[(-1)^{n+1}\mathrm{e}^{\xi^2}\frac{\mathrm{d}^{n+1}}{\mathrm{d}\xi^{n+1}}\left(\mathrm{e}^{-\xi^2}\right)\right]\mathrm{e}^{-\frac{1}{2}\xi^2} = \widetilde{N_{n+1}}H_{n+1}(x)\mathrm{e}^{-\frac{1}{2}\xi^2}.\end{aligned} \tag{2.103}$$

This means that (2.101) holds with $n+1$ as well. Thus, it follows that (2.101) is true of $n$ that is zero or any positive integer.

Orthogonal relation reads as

$$\int_{-\infty}^{\infty}\widetilde{\psi_m}(\xi)^*\widetilde{\psi_n}(\xi)\mathrm{d}\xi = \delta_{mn}. \tag{2.104}$$

Placing (2.98) back into the function form $\psi_n(q)$, we have

$$\psi_n(q) = \sqrt{\beta}\widetilde{\psi_n}(\beta q). \tag{2.105}$$

Using (2.101) and explicitly rewriting (2.105), we get

$$\psi_n(q) = \left(\frac{m\omega}{\hbar}\right)^{1/4} \sqrt{\frac{1}{\pi^{1/2} 2^n n!}} H_n\left(\sqrt{\frac{m\omega}{\hbar}} q\right) \mathrm{e}^{-\frac{m\omega}{2\hbar} q^2} (n = 0, 1, 2, \ldots). \tag{2.106}$$

We tabulate first several Hermite polynomials $H_n(x)$ in Table 2.1, where the index $n$ represents the highest order of the polynomials. In Table 2.1, we see that even functions and odd functions appear alternately (i.e., parity). This is the case with $\psi_n(q)$ as well, because $\psi_n(q)$ is a product of $H_n(x)$ and an even function $\mathrm{e}^{-\frac{m\omega}{2\hbar} q^2}$.

Combining (2.101) and (2.104), the orthogonal relation between $\widetilde{\psi_n}(\xi)$ $(n = 0, 1, 2, \ldots)$ can be described alternatively as [3]

$$\int_{-\infty}^{\infty} \mathrm{e}^{-\xi^2} H_m(\xi) H_n(\xi) \mathrm{d}\xi = \sqrt{\pi} 2^n n! \delta_{mn}. \tag{2.107}$$

Note that $H_m(\xi)$ is a real function, and so $H_m(\xi)^* = H_m(\xi)$. The relation (2.107) is well known as the orthogonality of Hermite polynomials with $\mathrm{e}^{-\xi^2}$ taken as a weight function [3]. Here, the weight function is a real and nonnegative function within the domain considered [e.g., $(-\infty, +\infty)$ in the present case] and independent of indices $m$ and $n$. We will deal with it again in Sect. 8.4.

The relation (2.101) and the orthogonality relationship described as (2.107) can more explicitly be understood as follows: From (2.11), we have the Schrödinger equation of a one-dimensional quantum-mechanical harmonic oscillator such that

$$-\frac{\hbar^2}{2m} \frac{\mathrm{d}^2 u(q)}{\mathrm{d}q^2} + \frac{1}{2} m\omega^2 q^2 u(q) = E u(q). \tag{2.108}$$

Changing a variable as in (2.90), we have

$$-\frac{\mathrm{d}^2 u(\xi)}{\mathrm{d}\xi^2} + \xi^2 u(\xi) = \frac{2E}{\hbar\omega} u(\xi). \tag{2.109}$$

Defining a dimensionless parameter

$$\lambda \equiv \frac{2E}{\hbar\omega}. \tag{2.110}$$

**Table 2.1** First six Hermite polynomials

| |
|---|
| $H_0(x) = 1$ |
| $H_1(x) = 2x$ |
| $H_2(x) = 4x^2 - 2$ |
| $H_3(x) = 8x^3 - 12x$ |
| $H_4(x) = 16x^4 - 48x^2 + 12$ |
| $H_5(x) = 32x^5 - 160x^3 + 120x$ |

and also defining a differential operator $D$ such that

$$D \equiv -\frac{\mathrm{d}^2}{\mathrm{d}\xi^2} + \xi^2, \tag{2.111}$$

we have a following eigenvalue equation:

$$Du(\xi) = \lambda u(\xi). \tag{2.112}$$

We further consider a following function $v(\xi)$ such that

$$u(\xi) = v(\xi)\mathrm{e}^{-\xi^2/2}. \tag{2.113}$$

Then, (2.109) is converted as follows:

$$\left[-\frac{\mathrm{d}^2 v(\xi)}{\mathrm{d}\xi^2} + 2\xi\frac{\mathrm{d}v(\xi)}{\mathrm{d}\xi}\right]\mathrm{e}^{-\frac{\xi^2}{2}} = (\lambda - 1)v(\xi)\mathrm{e}^{-\frac{\xi^2}{2}}. \tag{2.114}$$

Since $\mathrm{e}^{-\frac{\xi^2}{2}}$ does not vanish with any $\xi$, we have

$$-\frac{\mathrm{d}^2 v(\xi)}{\mathrm{d}\xi^2} + 2\xi\frac{\mathrm{d}v(\xi)}{\mathrm{d}\xi} = (\lambda - 1)v(\xi). \tag{2.115}$$

If we define another differential operator $\widetilde{D}$ such that

$$\widetilde{D} \equiv -\frac{\mathrm{d}^2}{\mathrm{d}\xi^2} + 2\xi\frac{\mathrm{d}}{\mathrm{d}\xi}, \tag{2.116}$$

we have another eigenvalue equation

$$\widetilde{D}v(\xi) = (\lambda - 1)v(\xi). \tag{2.117}$$

Meanwhile, we have a following well-known differential equation:

$$\frac{\mathrm{d}^2 H_n(\xi)}{\mathrm{d}\xi^2} - 2\xi\frac{\mathrm{d}H_n(\xi)}{\mathrm{d}\xi} + 2nH_n(\xi) = 0. \tag{2.118}$$

This equation is said to be Hermite differential equation. Using (2.116), (2.118) can be recast as an eigenvalue equation such that

$$\widetilde{D}H_n(\xi) = 2nH_n(\xi). \tag{2.119}$$

Therefore, comparing (2.115) and (2.118) and putting

$$\lambda = 2n + 1, \tag{2.120}$$

we get

$$v(\xi) = cH_n(\xi), \tag{2.121}$$

where $c$ is an arbitrary constant. Thus, using (2.113), for a solution of (2.109), we get

$$u_n(\xi) = cH_n(\xi)\mathrm{e}^{-\xi^2/2}, \tag{2.122}$$

where the solution $u(\xi)$ is indexed with $n$. From (2.110), as an energy eigenvalue, we have

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega.$$

Thus, (2.37) is recovered. A normalization constant $c$ of (2.122) can be decided as in (2.106).

As discussed above, the operator representation and coordinate representation are fully consistent.

## 2.5 Variance and Uncertainty Principle

Uncertainty principle is one of most fundamental concepts of quantum mechanics. To think of this conception on the basis of a quantum harmonic oscillator, let us introduce a variance operator [4]. Let $A$ be a physical quantity and let $\langle A\rangle$ be an expectation value as defined in (1.126). We define a variance operator as

$$\left\langle (\Delta A)^2 \right\rangle,$$

where we have

$$\Delta A \equiv A - \langle A\rangle. \tag{2.123}$$

In (2.123), we assume that $\langle A\rangle$ is obtained by operating $A$ on a certain physical state $|\psi\rangle$. Then, we have

$$\left\langle (\Delta A)^2 \right\rangle = \left\langle (A - \langle A\rangle)^2 \right\rangle = \left\langle A^2 - 2\langle A\rangle A + \langle A\rangle^2 \right\rangle = \langle A\rangle^2 - \langle A\rangle^2. \tag{2.124}$$

If $A$ is Hermitian, $\Delta A$ is Hermitian as well. This is because

$$(\Delta A)^\dagger = A^\dagger - \langle A\rangle^* = A - \langle A\rangle = \Delta A, \tag{2.125}$$

where we used the fact that an expectation value of an Hermitian operator is real. Then, $\left\langle (\Delta A)^2 \right\rangle$ is nonnegative as in the case of (2.13). Moreover, if $|\psi\rangle$ is an eigenstate of $A$, $\left\langle (\Delta A)^2 \right\rangle = 0$. Therefore, $\left\langle (\Delta A)^2 \right\rangle$ represents a measure of how large measured values are dispersed when $A$ is measured in reference to a quantum state $|\psi\rangle$. Also, we define a standard deviation $\delta A$ as

$$\delta A \equiv \sqrt{\left\langle (\Delta A)^2 \right\rangle}. \tag{2.126}$$

We have a following important theorem on a standard deviation $\delta A$ [4].

**Theorem 2.1** *Let $A$ and $B$ be Hermitian operators. If $A$ and $B$ satisfy*

$$[A, B] = ik \,(k: \text{ non-zero real number}), \tag{2.127}$$

*then we have*

$$\delta A \cdot \delta B \geq |k|/2 \tag{2.128}$$

*in reference to any quantum state $|\psi\rangle$.*

*Proof* We have

$$[\Delta A, \Delta B] = [A - \langle\psi|A|\psi\rangle, B - \langle\psi|B|\psi\rangle] = [A, B] = ik. \tag{2.129}$$

In (2.129), we used the fact that $\langle\psi|A|\psi\rangle$ and $\langle\psi|B|\psi\rangle$ are just real numbers and those commute with any operator. Next, we calculate a following quantity in relation to a real number $\lambda$:

$$\begin{aligned} \|(\Delta A + i\lambda\Delta B)|\psi\rangle\|^2 &= \langle\psi|(\Delta A - i\lambda\Delta B)(\Delta A + i\lambda\Delta)|\psi\rangle \\ &= \langle\psi|(\Delta A)^2|\psi\rangle - k\lambda + \langle\psi|(\Delta B)^2|\psi\rangle\lambda^2, \end{aligned} \tag{2.130}$$

where we used the fact that $\Delta A$ and $\Delta B$ are Hermitian. For the above quadratic form to hold with any real number $\lambda$, we have

$$(-k)^2 - 4\langle\psi|(\Delta A)^2|\psi\rangle\langle\psi|(\Delta B)^2|\psi\rangle \leq 0. \tag{2.131}$$

Thus, (2.128) will follow.

On the basis of Theorem 2.1, we find that both $\delta A$ and $\delta B$ are positive on condition that (2.127) holds. We have another important theorem.

**Theorem 2.2** *Let $A$ be an Hermitian operator. The necessary and sufficient condition for a physical state $|\psi_0\rangle$ to be an eigenstate of $A$ is $\delta A = 0$.*

*Proof* Suppose that $|\psi_0\rangle$ is a normalized eigenstate of $A$ that belongs to an eigenvalue $a$. Then, we have

$$\begin{aligned} &\langle\psi_0|A^2\psi_0\rangle = a\langle\psi_0|A\psi_0\rangle = a^2\langle\psi_0|\psi_0\rangle = a^2, \\ &\langle\psi_0|A\psi_0\rangle^2 = [a\langle\psi_0|\psi_0\rangle]^2 = a^2. \end{aligned} \tag{2.132}$$

From (2.124) and (2.126), we have

$$\langle\psi_0|(\Delta A)^2\psi_0\rangle = 0 \quad \text{i.e. } \delta A = 0. \tag{2.133}$$

Note that $\delta A$ is measured in reference to $|\psi_0\rangle$. Conversely, suppose that $\delta A = 0$. Then,

$$\delta A = \sqrt{\langle\psi_0|(\Delta A)^2\psi_0\rangle} = \sqrt{\langle\Delta A\psi_0|\Delta A\psi_0\rangle} = ||\Delta A\psi_0||, \tag{2.134}$$

where we used the fact that $\Delta A$ is Hermitian. From the definition of norm of (1.121), for $\delta A = 0$ to hold, we have

$$\Delta A\psi_0 = (A - \langle A\rangle)\psi_0 = 0 \quad \text{i.e. } A\psi_0 = \langle A\rangle\psi_0. \tag{2.135}$$

This indicates that $\psi_0$ is an eigenstate of $A$ that belongs to an eigenvalue $\langle A\rangle$. This completes the proof.

Theorem 2.1 implies that (2.127) holds with *any* physical state $|\psi\rangle$. That is, we must have $\delta A > 0$ and $\delta B > 0$, if $\delta A$ and $\delta B$ are evaluated in reference to any $|\psi\rangle$ on condition that (2.127) holds. From Theorem 2.2, in turn, it follows that eigenstates cannot exist with $A$ or $B$ under the condition of (2.127).

To explicitly show this, we take an inner product of (2.127). That is, with Hermitian operators $A$ and $B$, consider the following inner product:

$$\langle\psi|[A, B]|\psi\rangle = \langle\psi|ik|\psi\rangle \quad \text{i.e. } \langle\psi|AB - BA|\psi\rangle = ik, \tag{2.136}$$

where we assumed that $|\psi\rangle$ is arbitrarily chosen normalized vector. Suppose now that $|\psi_0\rangle$ is an eigenstate of $A$ that belongs to an eigenvalue $a$. Then, we have

$$A|\psi_0\rangle = a|\psi_0\rangle. \tag{2.137}$$

Taking an adjoint of (2.137), we get

$$\langle\psi_0|A^\dagger = \langle\psi_0|A = \langle\psi_0|a^* = a\langle\psi_0|, \tag{2.138}$$

where the last equality comes from the fact that $A$ is Hermitian. From (2.138), we would have

$$\langle\psi_0|AB-BA|\psi_0\rangle = \langle\psi_0|AB|\psi_0\rangle - \langle\psi_0|BA|\psi_0\rangle$$
$$= \langle\psi_0|aB|\psi_0\rangle - \langle\psi_0|Ba|\psi_0\rangle = a\langle\psi_0|B|\psi_0\rangle - a\langle\psi_0|B|\psi_0\rangle = 0.$$

This would imply that (2.136) does not hold with $|\psi_0\rangle$, in contradiction to (2.127), where $ik \neq 0$. Namely, we conclude that any physical state cannot be an eigenstate of $A$ on condition that (2.127) holds. Equation (2.127) is rewritten as

$$\langle\psi|BA-AB|\psi\rangle = -ik. \tag{2.139}$$

Suppose now that $|\varphi_0\rangle$ is an eigenstate of $B$ that belongs to an eigenvalue $b$. Then, we can similarly show that any physical state cannot be an eigenstate of $B$.

Summarizing the above, we restate that once we have a relation $[A,B] = ik$ $(k \neq 0)$, their representation matrix does not diagonalize $A$ or $B$. Or, once we postulate $[A,B] = ik(k \neq 0)$, we must abandon an effort to have a representation matrix that diagonalizes $A$ and $B$. In the quantum-mechanical formulation of a harmonic oscillator, we have introduced the canonical commutation relation (see Sect. 2.3) described by $[q,p] = i\hbar$ (1.140). Indeed, neither $q$ nor $p$ is diagonalized as shown in (2.69) or (2.70).

*Example 2.1* Taking a quantum harmonic oscillator as an example, we consider variance of $q$ and $p$ in reference to $|\psi_n\rangle$ $(n = 0, 1, \ldots)$. We have

$$\left\langle (\Delta q)^2 \right\rangle = \langle\psi_n|q^2|\psi_n\rangle - \langle\psi_n|q|\psi_n\rangle^2. \tag{2.140}$$

Using (2.55) and (2.62) as well as (2.68), we get

$$\begin{aligned}\langle\psi_n|q|\psi_n\rangle &= \sqrt{\frac{\hbar}{2m\omega}}\langle\psi_n|a+a^\dagger|\psi_n\rangle \\ &= \sqrt{\frac{\hbar n}{2m\omega}}\langle\psi_{n-1}|\psi_n\rangle + \sqrt{\frac{\hbar(n+1)}{2m\omega}}\langle\psi_n|\psi_{n+1}\rangle = 0,\end{aligned} \tag{2.141}$$

where the last equality comes from (2.53). We have

$$q^2 = \frac{\hbar}{2m\omega}(a+a^\dagger)^2 = \frac{\hbar}{2m\omega}\left[a^2 + \widehat{E} + 2a^\dagger a + (a^\dagger)^2\right], \tag{2.142}$$

where $\widehat{E}$ denotes a unit operator and we used (2.24) along with the following relation:

$$aa^\dagger = aa^\dagger - a^\dagger a + a^\dagger a = \left[a, a^\dagger\right] + a^\dagger a = \widehat{E} + a^\dagger a. \tag{2.143}$$

Using (2.55) and (2.62), we have

$$\langle\psi_n|q^2|\psi_n\rangle = \frac{\hbar}{2m\omega}\left[\langle\psi_n|\psi_n\rangle + 2\langle\psi_n|a^\dagger a\psi_n\rangle\right] = \frac{\hbar}{2m\omega}(2n+1), \tag{2.144}$$

where we used (2.60) with the last equality. Thus, we get

$$\left\langle(\Delta q)^2\right\rangle = \langle\psi_n|q^2|\psi_n\rangle - \langle\psi_n|q|\psi_n\rangle^2 = \frac{\hbar}{2m\omega}(2n+1).$$

Following similar procedures to those mentioned above, we get

$$\langle\psi_n|p|\psi_n\rangle = 0 \text{ and } \langle\psi_n|p^2|\psi_n\rangle = \frac{m\hbar\omega}{2}(2n+1). \tag{2.145}$$

Thus, we get

$$\left\langle(\Delta p)^2\right\rangle = \langle\psi_n|p^2|\psi_n\rangle = \frac{m\hbar\omega}{2}(2n+1).$$

Accordingly, we have

$$\delta q \cdot \delta p = \sqrt{\left\langle(\Delta q)^2\right\rangle}\cdot\sqrt{\left\langle(\Delta p)^2\right\rangle} = \frac{\hbar}{2}(2n+1) \geq \frac{\hbar}{2}. \tag{2.146}$$

The quantity $\delta q \cdot \delta p$ is equal to $\frac{\hbar}{2}$ for $n = 0$ and becomes larger with increasing $n$.

The above example gives a good illustration for Theorem 2.1. Note that putting $A = q$ and $B = p$ along with $k = \hbar$ in Theorem 2.1, we should have from (1.140)

$$\delta q \cdot \delta p \geq \frac{\hbar}{2}.$$

This is indeed the case with (2.146) for the quantum-mechanical harmonic oscillator. This example represents uncertainty principle more generally.

In relation to the aforementioned argument, we might well wonder if in Examples 1.1 and 1.2 have an eigenstate of a fixed momentum. Suppose that we chose for an eigenstate $y(x) = ce^{ikx}$, where $c$ is a constant. Then, we would have $\frac{\hbar}{i}\frac{\partial y(x)}{\partial x} = \hbar k y(x)$ and get an eigenvalue $\hbar k$ for a momentum. Nonetheless, such $y(x)$ does not satisfy the proper BCs; i.e., $y(L) = y(-L) = 0$. This is because $e^{ikx}$ never vanishes with any real numbers of $k$ or $x$ (any complex numbers of $k$ or $x$, more generally). Thus, we cannot obtain a proper solution that has an eigenstate with a fixed momentum in a confined physical system.

## References

1. Messiah A (1999) Quantum mechanics. Dover, New York
2. Stakgold I (1998) Green's functions and boundary value problems, 2nd edn. Wiley, New York
3. Lebedev NN (1972) Special functions and their applications. Dover, New York
4. Shimizu A (2003) Upgrade foundation of quantum theory. Saiensu-sha, Tokyo (in Japanese)

# Chapter 3
# Hydrogen-like Atoms

In a history of quantum mechanics, it was first successfully applied to the motion of an electron in a hydrogen atom along with a harmonic oscillator. Unlike the case of a one-dimensional harmonic oscillator we dealt with in Chap. 2 however, with a hydrogen atom we have to consider three-dimensional motion of an electron. Accordingly, it takes somewhat elaborate calculations to constitute the Hamiltonian. The calculation procedures themselves, however, are worth following to understand underlying basic concepts of the quantum mechanics. At the same time, this chapter is a treasure of special functions. In Chap. 2, we have already encountered one of them, i.e., Hermite polynomials. Here, we will deal with Legendre polynomials associated Legendre polynomials, etc. These special functions arise when we deal with a physical system having, e.g., the spherical symmetry. In a hydrogen atom, an electron is moving in a spherically symmetric Coulomb potential field produced by a proton. This topic provides us with a good opportunity to study various special functions. The related Schrödinger equation can be separated into an angular part and a radial part. The solutions of angular parts are characterized by spherical (surface) harmonics. The (associated) Legendre functions are correlated with them. The solutions of the radial part are connected to the (associated) Laguerre polynomials. The exact solutions are obtained by the product of the (associated) Legendre functions and (associated) Laguerre polynomials accordingly. Thus, to study the characteristics of hydrogen-like atoms from the quantum-mechanical perspective is of fundamental importance.

## 3.1 Introductory Remarks

The motion of the electron in hydrogen is well-known as a two-particle problem (or two-body problem) in a central force field. In that case, the coordinate system of the physical system is separated into the relative coordinates and center-of-mass coordinates. To be more specific, the coordinate separation is true of the case where



S. Hotta, *Mathematical Physical Chemistry*,
https://doi.org/10.1007/978-981-10-7671-8_3

two particles are moving under control only by a force field between the two particles without other external force fields [1].

In the classical mechanics, equation of motion is separated into two equations related to the relative coordinates and center-of-mass coordinates accordingly. Of these, a term of the potential field is only included in the equation of motion with respect to the relative coordinates.

The situation is the same with the quantum mechanics. Namely, the Schrödinger equation of motion with the relative coordinates is expressed as an eigenvalue equation that reads as

$$\left[-\frac{\hbar^2}{2\mu}\nabla^2 + V(r)\right]\psi = E\psi, \tag{3.1}$$

where $\mu$ is a reduced mass of two particles [1], i.e., an electron and a proton; $V(r)$ is a potential with $r$ being a distance between the electron and proton. In (3.1), we assume the spherically symmetric potential; i.e., the potential is expressed only as a function of the distance $r$. Moreover, if the potential is coulombic,

$$\left(-\frac{\hbar^2}{2\mu}\nabla^2 - \frac{e^2}{4\pi\varepsilon_0 r}\right)\psi = E\psi, \tag{3.2}$$

where $\varepsilon_0$ is permittivity of vacuum and $e$ is an elementary charge.

If we think of hydrogen-like atoms such as $He^+$, $Li^{2+}$, $Be^{3+}$, etc., we have an equation described as

$$\left(-\frac{\hbar^2}{2\mu}\nabla^2 - \frac{Ze^2}{4\pi\varepsilon_0 r}\right)\psi = E\psi, \tag{3.3}$$

where $Z$ is an atomic number and $\mu$ is a reduced mass of an electron and a nucleus pertinent to the atomic (or ionic) species. We start with (3.3) in this chapter.

## 3.2 Constitution of Hamiltonian

As explicitly described in (3.3), the coulombic potential has a spherical symmetry. In such a case, it will be convenient to recast (3.3) in a spherical coordinate (or polar coordinate). As the physical system is of three-dimensional, we have to consider orbital angular momentum $\boldsymbol{L}$ in Hamiltonian.

We have

$$\boldsymbol{L} = (\boldsymbol{e}_1\boldsymbol{e}_2\boldsymbol{e}_3)\begin{pmatrix} L_x \\ L_y \\ L_z \end{pmatrix}, \tag{3.4}$$

where $\boldsymbol{e}_1, \boldsymbol{e}_2,$ and $\boldsymbol{e}_3$ denote an orthonormal basis vectors in a three-dimensional Cartesian space ($\mathbb{R}^3$); $L_x$, $L_y$, and $L_z$ represent each component of $\boldsymbol{L}$. The angular momentum $\boldsymbol{L}$ is expressed in a form of determinant as

$$\boldsymbol{L} = \boldsymbol{x} \times \boldsymbol{p} = \begin{vmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3 \\ x & y & z \\ p_x & p_y & p_z \end{vmatrix},$$

where $\boldsymbol{x}$ denotes a position vector with respect to the relative coordinates $x$, $y$, and $z$. That is,

$$\boldsymbol{x} = (\boldsymbol{e}_1 \boldsymbol{e}_2 \boldsymbol{e}_3) \begin{pmatrix} x \\ y \\ z \end{pmatrix}. \tag{3.5}$$

The quantity $\boldsymbol{p}$ denotes a momentum of an electron (as a particle carrying a reduced mass $\mu$) with $p_x$, $p_y$, and $p_z$ being their components; $\boldsymbol{p}$ is denoted similar to the above.

As for each component of $\boldsymbol{L}$, we have, e.g.,

$$L_x = yp_z - zp_y. \tag{3.6}$$

To calculate $\boldsymbol{L}^2$, we estimate $L_x^2$, $L_y^2$, and $L_z^2$ separately. We have

$$\begin{aligned} L_x^2 &= (yp_z - zp_y) \cdot (yp_z - zp_y) \\ &= yp_z yp_z - yp_z zp_y - zp_y yp_z - zp_y zp_y \\ &= y^2 p_z^2 - yp_z zp_y - zp_y yp_z + z^2 p_y^2 \\ &= y^2 p_z^2 - y(zp_z - i\hbar)p_y - z(yp_y - i\hbar)p_z + z^2 p_y^2 \\ &= y^2 p_z^2 + z^2 p_y^2 - yzp_z p_y - zyp_y p_z + i\hbar(yp_y + zp_z), \end{aligned}$$

where we have used canonical commutation relation (1.140) in the third equality. In the above calculations, we used commutability of, e.g., $y$ and $p_z$; $z$ and $p_y$. For example, we have

$$[p_z, y]|\psi\rangle = \frac{\hbar}{i}\left(\frac{\partial}{\partial z}y - y\frac{\partial}{\partial z}\right)|\psi\rangle = \frac{\hbar}{i}\left(y\frac{\partial|\psi\rangle}{\partial z} - y\frac{\partial|\psi\rangle}{\partial z}\right) = 0.$$

Since $|\psi\rangle$ is arbitrarily chosen, this relation implies that $p_z$ and $y$ commute. We obtain similar relations regarding $L_y^2$ and $L_z^2$ as well. Thus, we have

$$\begin{aligned}\boldsymbol{L}^2 &= L_x^2 + L_y^2 + L_z^2 \\ &= (y^2p_z^2 + z^2p_y^2 + z^2p_x^2 + x^2p_z^2 + x^2p_y^2 + y^2p_x^2) \\ &\quad + (x^2p_x^2 - x^2p_x^2 + y^2p_y^2 - y^2p_y^2 + z^2p_z^2 - z^2p_z^2) \\ &\quad - (yzp_zp_y + zyp_yp_z + zxp_xp_z + xzp_zp_x + xyp_yp_x + yxp_xp_y) \\ &\quad + i\hbar(yp_y + zp_z + zp_z + xp_x + xp_x + yp_y) \\ &= (y^2p_z^2 + z^2p_y^2 + z^2p_x^2 + x^2p_z^2 + x^2p_y^2 + y^2p_x^2 \\ &\quad + x^2p_x^2 + y^2p_y^2 + z^2p_z^2) - (x^2p_x^2 + y^2p_y^2 + z^2p_z^2 \\ &\quad + yzp_zp_y + zyp_yp_z + zxp_xp_z + xzp_zp_x + xyp_yp_x + yxp_xp_y) \\ &\quad + i\hbar(yp_y + zp_z + zp_z + xp_x + xp_x + yp_y) \\ &= \boldsymbol{r}^2 \cdot \boldsymbol{p}^2 - \boldsymbol{r}(\boldsymbol{r} \cdot \boldsymbol{p}) \cdot \boldsymbol{p} + 2i\hbar(\boldsymbol{r} \cdot \boldsymbol{p})\end{aligned} \quad . \tag{3.7}$$

In (3.7), we are able to ease the calculations by virtue of putting a term $(x^2p_x^2 - x^2p_x^2 + y^2p_y^2 - y^2p_y^2 + z^2p_z^2 - z^2p_z^2)$. As a result, for the second term after the second to the last equality we have

$$\begin{aligned}&- (x^2p_x^2 + y^2p_y^2 + z^2p_z^2 \\ &+ yzp_zp_y + zyp_yp_z + zxp_xp_z + xzp_zp_x + xyp_yp_x + yxp_xp_y) \\ &= -\left[x(xp_x + yp_y + zp_z)p_x + y(xp_x + yp_y + zp_z)p_y + z(xp_x + yp_y + zp_z)p_z\right] \\ &= -\boldsymbol{r}(\boldsymbol{r} \cdot \boldsymbol{p}) \cdot \boldsymbol{p}.\end{aligned}$$

The calculations of $\boldsymbol{r}^2 \cdot \boldsymbol{p}^2$ [the first term of (3.7)] and $\boldsymbol{r} \cdot \boldsymbol{p}$ (in the third term) are straightforward.

In a spherical coordinate, momentum $\boldsymbol{p}$ is expressed as

$$\boldsymbol{p} = p_r\boldsymbol{e}^{(r)} + p_\theta\boldsymbol{e}^{(\theta)} + p_\phi\boldsymbol{e}^{(\phi)}, \tag{3.8}$$

where $p_r$, $p_\theta$, and $p_\phi$ are components of $\boldsymbol{p}$; $\boldsymbol{e}^{(r)}$, $\boldsymbol{e}^{(\theta)}$, and $\boldsymbol{e}^{(\phi)}$ are orthonormal basis vectors of $\mathbb{R}^3$ in the direction of increasing $r$, $\theta$, and $\phi$, respectively (see Fig. 3.1). In Fig. 3.1b, $\boldsymbol{e}^{(\phi)}$ is perpendicular to the plane shaped by the $z$-axis and a straight line of $y = x\tan\phi$. Notice that the said plane is spanned by $\boldsymbol{e}^{(r)}$ and $\boldsymbol{e}^{(\theta)}$. Meanwhile, the momentum operator is expressed as [2]

$$\begin{aligned}\boldsymbol{p} &= \frac{\hbar}{i}\boldsymbol{\nabla} \\ &= \frac{\hbar}{i}\left[\boldsymbol{e}^{(r)}\frac{\partial}{\partial r} + \boldsymbol{e}^{(\theta)}\frac{1}{r}\frac{\partial}{\partial\theta} + \boldsymbol{e}^{(\phi)}\frac{1}{r\sin\theta}\frac{\partial}{\partial\phi}\right].\end{aligned} \tag{3.9}$$

The vector notation of (3.9) corresponds to (1.31). That is, in the Cartesian coordinate, we have

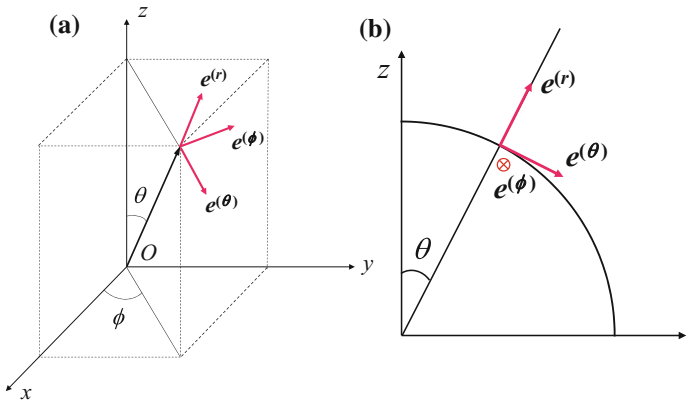


**Fig. 3.1** Spherical coordinate system and orthonormal basis set. **a** Orthonormal basis vectors $\boldsymbol{e}^{(r)}$, $\boldsymbol{e}^{(\theta)}$, and $\boldsymbol{e}^{(\phi)}$ in $\mathbb{R}^3$. **b** The basis vector $\boldsymbol{e}^{(\phi)}$ is perpendicular to the plane shaped by the $z$-axis and a straight line of $y = x\tan\phi$

$$\boldsymbol{p} = \frac{\hbar}{i}\boldsymbol{\nabla} = \frac{\hbar}{i}\left(\boldsymbol{e}_1\frac{\partial}{\partial x} + \boldsymbol{e}_2\frac{\partial}{\partial y} + \boldsymbol{e}_3\frac{\partial}{\partial z}\right),$$

where $\boldsymbol{\nabla}$ is said to be nabla (or del), a kind of differential vector operator.

Noting that

$$\boldsymbol{r} = r\boldsymbol{e}^{(r)}, \tag{3.10}$$

and using (3.9), we have

$$\boldsymbol{r}\cdot\boldsymbol{p} = \boldsymbol{r}\cdot\frac{\hbar}{i}\boldsymbol{\nabla} = r\left(\frac{\hbar}{i}\frac{\partial}{\partial r}\right). \tag{3.11}$$

Hence,

$$\boldsymbol{r}(\boldsymbol{r}\cdot\boldsymbol{p})\cdot\boldsymbol{p} = r\left[r\left(\frac{\hbar}{i}\frac{\partial}{\partial r}\right)\right]\left(\frac{\hbar}{i}\frac{\partial}{\partial r}\right) = -\hbar^2 r^2\frac{\partial^2}{\partial r^2}. \tag{3.12}$$

Thus, we have

$$\boldsymbol{L}^2 = r^2\boldsymbol{p}^2 + \hbar^2 r^2\frac{\partial^2}{\partial r^2} + 2\hbar^2 r\frac{\partial}{\partial r} = r^2\boldsymbol{p}^2 + \hbar^2\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right). \tag{3.13}$$

Therefore,

$$\boldsymbol{p}^2 = -\frac{\hbar^2}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{\boldsymbol{L}^2}{r^2}. \tag{3.14}$$

Notice here that $\boldsymbol{L}^2$ does not contain $r$ (vide infra); i.e., $\boldsymbol{L}^2$ commutes with $r^2$, and so it can freely be divided by $r^2$. Thus, the Hamiltonian $H$ is represented by

$$\begin{aligned} H &= \frac{\boldsymbol{p}^2}{2\mu} + V(r) \\ &= \frac{1}{2\mu}\left[-\frac{\hbar^2}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{\boldsymbol{L}^2}{r^2}\right] - \frac{Ze^2}{4\pi\varepsilon_0 r}. \end{aligned} \tag{3.15}$$

Thus, the Schrödinger equation can be expressed as

$$\left\{\frac{1}{2\mu}\left[-\frac{\hbar^2}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{\boldsymbol{L}^2}{r^2}\right] - \frac{Ze^2}{4\pi\varepsilon_0 r}\right\}\psi = E\psi. \tag{3.16}$$

Now, let us describe $\boldsymbol{L}^2$ in a polar coordinate. The calculation procedures are somewhat lengthy, but straightforward. First, we have

$$\left.\begin{array}{c} x = r\sin\theta\cos\phi, \\ y = r\sin\theta\sin\phi, \\ z = r\cos\theta, \end{array}\right\} \tag{3.17}$$

where we have $0 \le \theta \le \pi$ and $0 \le \phi \le 2\pi$. Rewriting (3.17) with respect to $r$, $\theta$, and $\phi$, we get

$$\left.\begin{array}{l} r = (x^2 + y^2 + z^2)^{\frac{1}{2}}, \\ \theta = \tan^{-1}\frac{(x^2+y^2)^{1/2}}{z}, \\ \phi = \tan^{-1}\frac{y}{x}. \end{array}\right\} \tag{3.18}$$

Thus, we have

$$L_z = xp_y - yp_x = -i\hbar\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right), \tag{3.19}$$

$$\frac{\partial}{\partial x} = \frac{\partial r}{\partial x}\frac{\partial}{\partial r} + \frac{\partial\theta}{\partial x}\frac{\partial}{\partial\theta} + \frac{\partial\phi}{\partial x}\frac{\partial}{\partial\phi}. \tag{3.20}$$

In turn, we have

$$\begin{aligned}
\frac{\partial r}{\partial x} &= \frac{x}{r} = \sin\theta\cos\phi, \\
\frac{\partial \theta}{\partial x} &= \frac{1}{1+(x^2+y^2)/z^2}\cdot\frac{(x^2+y^2)^{-\frac{1}{2}}\cdot 2x}{2z} = \frac{z}{x^2+y^2+z^2}\cdot\frac{x}{(x^2+y^2)^{\frac{1}{2}}} = \frac{\cos\theta\cos\phi}{r}, \\
\frac{\partial \phi}{\partial x} &= \frac{1}{1+(y^2/x^2)}\cdot y\left(-\frac{1}{x^2}\right) = -\frac{\sin\phi}{r\sin\theta}.
\end{aligned} \tag{3.21}$$

In calculating the last two equations of (3.21), we used the differentiation of an arc tangent function along with a composite function. Namely,

$$(\tan^{-1}x)' = \frac{1}{1+x^2}.$$

Inserting (3.21) into (3.20), we get

$$\frac{\partial}{\partial x} = \sin\theta\cos\phi\frac{\partial}{\partial r} + \frac{\cos\theta\cos\phi}{r}\frac{\partial}{\partial\theta} - \frac{\sin\phi}{r\sin\theta}\frac{\partial}{\partial\phi}. \tag{3.22}$$

Similarly, we have

$$\frac{\partial}{\partial y} = \sin\theta\sin\phi\frac{\partial}{\partial r} + \frac{\cos\theta\sin\phi}{r}\frac{\partial}{\partial\theta} + \frac{\cos\phi}{r\sin\theta}\frac{\partial}{\partial\phi}. \tag{3.23}$$

Inserting (3.22) and (3.23) together with (3.17) into (3.19), we get

$$L_z = -i\hbar\frac{\partial}{\partial\phi}. \tag{3.24}$$

In a similar manner, we have

$$\frac{\partial}{\partial z} = \frac{\partial r}{\partial z}\frac{\partial}{\partial r} + \frac{\partial\theta}{\partial z}\frac{\partial}{\partial\theta} = \cos\theta\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial\theta}.$$

Combining this relation with either (3.23) or (3.22), we get

$$L_x = i\hbar\left(\sin\phi\frac{\partial}{\partial\theta} + \cot\theta\cos\phi\frac{\partial}{\partial\phi}\right), \tag{3.25}$$

$$L_y = -i\hbar\left(\cos\phi\frac{\partial}{\partial\theta} - \cot\theta\sin\phi\frac{\partial}{\partial\phi}\right). \tag{3.26}$$

Now, we introduce following operators:

$$L^{(+)} \equiv L_x + iL_y \text{ and } L^{(-)} \equiv L_x - iL_y. \tag{3.27}$$

Then, we have

$$L^{(+)} = \hbar e^{i\phi}\left(\frac{\partial}{\partial\theta} + i\cot\theta\frac{\partial}{\partial\phi}\right), \quad \text{and } L^{(-)} = \hbar e^{-i\phi}\left(-\frac{\partial}{\partial\theta} + i\cot\theta\frac{\partial}{\partial\phi}\right). \tag{3.28}$$

Thus, we get

$$\begin{aligned} L^{(+)}L^{(-)} &= \hbar^2 e^{i\phi}\left(\frac{\partial}{\partial\theta} + i\cot\theta\frac{\partial}{\partial\phi}\right)e^{-i\phi}\left(-\frac{\partial}{\partial\theta} + i\cot\theta\frac{\partial}{\partial\phi}\right) \\ &= \hbar^2 e^{i\phi}\left\{e^{-i\phi}\left[-\frac{\partial^2}{\partial\theta^2} + i\left(-\frac{1}{\sin^2\theta}\right)\frac{\partial}{\partial\phi} + i\cot\theta\frac{\partial^2}{\partial\theta\partial\phi}\right]\right. \\ &\quad \left. + e^{-i\phi}\cot\theta\left(-\frac{\partial}{\partial\theta} + i\cot\theta\frac{\partial}{\partial\phi}\right) + ie^{-i\phi}\cot\theta\left(-\frac{\partial^2}{\partial\phi\partial\theta} + i\cot\theta\frac{\partial^2}{\partial\phi^2}\right)\right\} \\ &= -\hbar^2\left(\frac{\partial^2}{\partial\theta^2} + \cot\theta\frac{\partial}{\partial\theta} + i\frac{\partial}{\partial\phi} + \cot^2\theta\frac{\partial^2}{\partial\phi^2}\right). \end{aligned} \tag{3.29}$$

In the above calculation procedure, we used differentiation of a product function. For instance, we have

$$\begin{aligned} \frac{\partial}{\partial\theta}\left(i\cot\theta\frac{\partial}{\partial\phi}\right) &= i\left(\frac{\partial\cot\theta}{\partial\theta}\frac{\partial}{\partial\phi} + \cot\theta\frac{\partial^2}{\partial\theta\partial\phi}\right) \\ &= i\left[\left(-\frac{1}{\sin^2\theta}\right)\frac{\partial}{\partial\phi} + \cot\theta\frac{\partial^2}{\partial\theta\partial\phi}\right]. \end{aligned}$$

Note also that $\frac{\partial^2}{\partial\theta\partial\phi} = \frac{\partial^2}{\partial\phi\partial\theta}$. This is because we are dealing with continuous and differentiable functions.

Meanwhile, we have following commutation relations:

$$[L_x, L_y] = i\hbar L_z, [L_y, L_z] = i\hbar L_x, \quad \text{and } [L_z, L_x] = i\hbar L_y. \tag{3.30}$$

This can easily be confirmed by requiring canonical commutation relations. The derivation can routinely be performed, but we show it because the procedures include several important points. For instance, we have

$$
\begin{aligned}
[L_x, L_y] &= L_x L_y - L_y L_x \\
&= (yp_z - zp_y)(zp_x - xp_z) - (zp_x - xp_z)(yp_z - zp_y) \\
&= yp_z zp_x - yp_z xp_z - zp_y zp_x + zp_y xp_z \\
&\quad - zp_x yp_z + zp_x zp_y + xp_z yp_z - xp_z zp_y \\
&= (yp_x p_z z - zp_x yp_z) + (zp_y xp_z - xp_z zp_y) \\
&\quad + (xp_z yp_z - yp_z xp_z) + (zp_x zp_y - zp_y zp_x) \\
&= -yp_x(zp_z - p_z z) + xp_y(zp_z - p_z z) = i\hbar(xp_y - yp_x) = i\hbar L_z
\end{aligned}
$$

In the above calculations, we used the canonical commutation relation as well as commutability of, e.g., $y$ and $p_x$; $y$ and $z$; $p_x$ and $p_y$. For example, we get

$$
\left[p_x, p_y\right]|\psi\rangle = -\hbar^2\left(\frac{\partial}{\partial x}\frac{\partial}{\partial y} - \frac{\partial}{\partial y}\frac{\partial}{\partial x}\right)|\psi\rangle = -\hbar^2\left(\frac{\partial^2|\psi\rangle}{\partial x \partial y} - \frac{\partial^2|\psi\rangle}{\partial y \partial x}\right) = 0.
$$

In the above equation, we assumed that the order of differentiation with respect to $x$ and $y$ can be switched. It is because we are dealing with continuous and differentiable normal functions. Thus, $p_x$ and $p_y$ commute.

For other important commutation relations, we have

$$
[L_x, \boldsymbol{L}^2] = 0, \ [L_y, \boldsymbol{L}^2] = 0, \quad \text{and } [L_z, \boldsymbol{L}^2] = 0. \tag{3.31}
$$

With the derivation, use

$$
[A, B + C] = [A, B] + [A, C].
$$

The derivation is straightforward and it is left for readers. The relations (3.30) and (3.31) imply that a *simultaneous eigenstate* exists for $\boldsymbol{L}^2$ and one of $L_x$, $L_y$, and $L_z$. This is because $\boldsymbol{L}^2$ commute with them from (3.31), whereas $L_z$ does not commute with $L_x$ or $L_y$. The detailed argument about the simultaneous eigenstate can be seen in Part III.

Thus, we have

$$
\begin{aligned}
L^{(+)}L^{(-)} &= L_x^2 + L_y^2 + i(L_y L_x - L_x L_y) = L_x^2 + L_y^2 + i[L_y, L_x] \\
&= L_x^2 + L_y^2 + \hbar L_z.
\end{aligned}
$$

Notice here that $[L_y, L_x] = -[L_x, L_y] = -i\hbar L_z$. Hence,

$$
\boldsymbol{L}^2 = L^{(+)}L^{(-)} + L_z^2 - \hbar L_z. \tag{3.32}
$$

From (3.24), we have

$$L_z^2 = -\hbar^2 \frac{\partial^2}{\partial \phi^2}. \tag{3.33}$$

Finally, we get

$$\begin{aligned} \boldsymbol{L}^2 &= -\hbar^2 \left( \frac{\partial^2}{\partial \theta^2} + \cot\theta \frac{\partial}{\partial \theta} + \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial \phi^2} \right) \quad \text{or} \\ \boldsymbol{L}^2 &= -\hbar^2 \left[ \frac{1}{\sin\theta} \frac{\partial}{\partial \theta} \left( \sin\theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial \phi^2} \right]. \end{aligned} \tag{3.34}$$

Replacing $\boldsymbol{L}^2$ in (3.15) with that of (3.34), we have

$$H = -\frac{\hbar^2}{2\mu r^2} \left[ \frac{\partial}{\partial r} \left( r^2 \frac{\partial}{\partial r} \right) + \frac{1}{\sin\theta} \frac{\partial}{\partial \theta} \left( \sin\theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial \phi^2} \right] - \frac{Ze^2}{4\pi\varepsilon_0 r} \tag{3.35}$$

Thus, the Schrödinger equation of (3.3) takes a following form:

$$\left\{ -\frac{\hbar^2}{2\mu r^2} \left[ \frac{\partial}{\partial r} \left( r^2 \frac{\partial}{\partial r} \right) + \frac{1}{\sin\theta} \frac{\partial}{\partial \theta} \left( \sin\theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial \phi^2} \right] - \frac{Ze^2}{4\pi\varepsilon_0 r} \right\} \psi = E\psi. \tag{3.36}$$

## 3.3 Separation of Variables

If the potential is spherically symmetric (e.g., a Coulomb potential), it is well-known that the Schrödinger equations of (3.1–3.3) can be solved by a method of separation of variables. More specifically, (3.36) can be separated into two differential equations one of which only depends on a radial component $r$ and the other of which depends only upon angular components $\theta$ and $\phi$.

To apply the method of separation of variables to (3.36), let us first return to (3.15). Considering that $\boldsymbol{L}^2$ is expressed as (3.34), we assume that $\boldsymbol{L}^2$ has eigenvalues $\gamma$ (at any rate if any) and takes eigenfunctions $Y(\theta, \phi)$ (again, if any as well) corresponding to $\gamma$. That is,

$$\boldsymbol{L}^2 Y(\theta, \phi) = \gamma Y(\theta, \phi), \tag{3.37}$$

where $Y(\theta, \phi)$ is assumed to be normalized. Meanwhile,

$$\boldsymbol{L}^2 = L_x^2 + L_y^2 + L_z^2. \tag{3.38}$$

From (3.6), we have

$$L_x^\dagger = (yp_z - zp_y)^\dagger = p_z^\dagger y^\dagger - p_y^\dagger z^\dagger = p_z y - p_y z = yp_z - zp_y = L_x. \tag{3.39}$$

Note that $p_z$ and $y$ commute, so do $p_y$ and $z$. Therefore, $L_x$ is Hermitian, so is $L_x^2$. More generally if an operator $A$ is Hermitian, so is $A^n$ ($n$: a positive integer); readers, please show it. Likewise, $L_y$ and $L_z$ are Hermitian as well. Thus, $\boldsymbol{L}^2$ is Hermitian, too.

Next, we consider an expectation value of $\boldsymbol{L}^2$, i.e., $\langle \boldsymbol{L}^2 \rangle$. Let $|\psi\rangle$ be an arbitrary normalized nonzero vector (or function). Then,

$$\begin{aligned}
\langle \boldsymbol{L}^2 \rangle &\equiv \langle \psi \mid \boldsymbol{L}^2 \psi \rangle \\
&= \langle \psi \mid L_x^2 \psi \rangle + \langle \psi \mid L_y^2 \psi \rangle + \langle \psi \mid L_z^2 \psi \rangle \\
&= \langle L_x^\dagger \psi \mid L_x \psi \rangle + \langle L_y^\dagger \psi \mid L_y \psi \rangle + \langle L_z^\dagger \psi \mid L_z \psi \rangle \\
&= \langle L_x \psi \mid L_x \psi \rangle + \langle L_y \psi \mid L_y \psi \rangle + \langle L_z \psi \mid L_z \psi \rangle \\
&= \|L_x \psi\|^2 + \|L_y \psi\|^2 + \|L_z \psi\|^2 \geq 0.
\end{aligned} \tag{3.40}$$

Notice that the second last equality comes from that $L_x$, $L_y$, and $L_z$ are Hermitian. An operator that satisfies (3.40) is said to be nonnegative (see Sects. 1.4 and 2.2, etc., where we saw the calculation routines). Note also that in (3.40) the equality holds only when the following relations hold:

$$|L_x \psi\rangle = |L_y \psi\rangle = |L_z \psi\rangle = 0. \tag{3.41}$$

On this condition, we have

$$\begin{aligned}
|\boldsymbol{L}^2 \psi\rangle &= |(L_x^2 + L_y^2 + L_z^2)\psi\rangle = |L_x^2 \psi\rangle + |L_y^2 \psi\rangle + |L_z^2 \psi\rangle \\
&= L_x |L_x \psi\rangle + L_y |L_y \psi\rangle + L_z |L_z \psi\rangle = 0.
\end{aligned} \tag{3.42}$$

The eigenfunction that satisfies (3.42) and the next relation (3.43) is a simultaneous eigenstate of $L_x$, $L_y$, $L_z$, and $\boldsymbol{L}^2$. This could seem to be in contradiction to the fact that $L_z$ does not commute with $L_x$ or $L_y$. However, this is an exceptional case. Let $|\psi_0\rangle$ be the eigenfunction that satisfies both (3.41) and (3.42). Then, we have

$$|L_x \psi_0\rangle = |L_y \psi_0\rangle = |L_z \psi_0\rangle = |\boldsymbol{L}^2 \psi_0\rangle = 0. \tag{3.43}$$

As can be seen from (3.24) to (3.26) along with (3.34), the operators $L_x$, $L_y$, $L_z$, and $\boldsymbol{L}^2$ are differential operators. Therefore, (3.43) implies that $|\psi_0\rangle$ is a constant. We will come back this point later. In spite of this exceptional situation, it is impossible that all $L_x$, $L_y$, and $L_z$ as well as $\boldsymbol{L}^2$ take a whole set of eigenfunctions as simultaneous eigenstates. We briefly show this as below.

In Chap. 2, we mention that if $[A, B] = ik$, any physical state cannot be an eigenstate of $A$ or $B$. The situation is different, on the other hand, if we have a following case

$$[A, B] = iC, \tag{3.44}$$

where $A$, $B$, and $C$ are Hermitian operators. The relation (3.30) is a typical example for this. If $C|\psi\rangle = 0$ in (3.44), $|\psi\rangle$ might well be an eigenstate of $A$ and/or $B$. However, if $C|\psi\rangle = c|\psi\rangle (c \neq 0)$, $|\psi\rangle$ cannot be an eigenstate of $A$ or $B$. This can readily be shown in a fashion similar to that described in Sect. 2.5. Let us think of, e.g., $[L_x, L_y] = i\hbar L_z$. Suppose that for $^{\exists}\psi_0$ we have $L_z|\psi_0\rangle = 0$. Taking an inner product using $|\psi_0\rangle$, from (3.30) we have

$$\langle \psi_0 \mid (L_x L_y - L_y L_x)\psi_0 \rangle = 0.$$

In this case, moreover, even if we have $|L_x\psi_0\rangle = 0$ and $|L_y\psi_0\rangle = 0$, we have no inconsistency. If, on the other hand, $L_z|\psi\rangle = m|\psi\rangle (m \neq 0)$, $|\psi\rangle$ cannot be an eigenstate of $L_x$ or $L_y$ as mentioned above. Thus, we should be careful to deal with a general situation where we have $[A, B] = iC$.

In the case where $[A, B] = 0$; $AB = BA$, namely $A$ and $B$ commute, we have a different situation. This relation is equivalent to that an operator $AB - BA$ has an eigenvalue zero for any physical state $|\psi\rangle$. Yet, this statement is of less practical use. Again, regarding details we wish to make a discussion in Sect. 12.6 of Part III.

Returning to (3.40), let us replace $\psi$ with a particular eigenfunction $Y(\theta, \phi)$. Then, we have

$$\langle Y \mid \boldsymbol{L}^2 Y \rangle = \langle Y \mid \gamma Y \rangle = \gamma \langle Y \mid Y \rangle = \gamma \geq 0. \tag{3.45}$$

Again, if $\boldsymbol{L}^2$ has an eigenvalue, the eigenvalue should be nonnegative. Taking account of the coefficient $\hbar^2$ in (3.34), it is convenient to put

$$\gamma = \hbar^2 \lambda (\lambda \geq 0). \tag{3.46}$$

On ground that the solution of (3.36) can be described as

$$\psi(r, \theta, \phi) = R(r)Y(\theta, \phi), \tag{3.47}$$

the Schrödinger equation (3.16) can be rewritten as

$$\left\{ \frac{1}{2\mu} \left[ -\frac{\hbar^2}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial}{\partial r} \right) + \frac{\boldsymbol{L}^2}{r^2} \right] - \frac{Ze^2}{4\pi\varepsilon_0 r} \right\} R(r)Y(\theta, \phi) = ER(r)Y(\theta, \phi). \tag{3.48}$$

That is,

$$\frac{1}{2\mu}\left[-\frac{\hbar^2}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial R(r)}{\partial r}\right)Y(\theta,\phi)+\frac{\boldsymbol{L}^2Y(\theta,\phi)}{r^2}R(r)\right]-\frac{Ze^2}{4\pi\varepsilon_0 r}R(r)Y(\theta,\phi) = ER(r)Y(\theta,\phi). \tag{3.49}$$

Recalling (3.37) and (3.46), we have

$$\frac{1}{2\mu}\left[-\frac{\hbar^2}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial R(r)}{\partial r}\right)Y(\theta,\phi)+\frac{\hbar^2\lambda Y(\theta,\phi)}{r^2}\right]R(r)-\frac{Ze^2}{4\pi\varepsilon_0 r}R(r)Y(\theta,\phi) = ER(r)Y(\theta,\phi). \tag{3.50}$$

Dividing both sides by $Y(\theta,\phi)$, we get a SOLDE of a radial component as

$$\frac{1}{2\mu}\left[-\frac{\hbar^2}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial R(r)}{\partial r}\right)+\frac{\hbar^2\lambda}{r^2}\right]R(r)-\frac{Ze^2}{4\pi\varepsilon_0 r}R(r)=ER(r). \tag{3.51}$$

Regarding angular components $\theta$ and $\phi$, using (3.34), (3.37), and (3.46), we have

$$\boldsymbol{L}^2Y(\theta,\phi)=-\hbar^2\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right]Y(\theta,\phi)=\hbar^2\lambda Y(\theta,\phi). \tag{3.52}$$

Dividing both sides by $\hbar^2$, we get

$$-\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right]Y(\theta,\phi)=\lambda Y(\theta,\phi). \tag{3.53}$$

Notice in (3.53) that the angular part of SOLDE does not depend on a specific form of the potential.

Now, we further assume that (3.53) can be separated into a zenithal angle part $\theta$ and azimuthal angle part $\phi$ such that

$$Y(\theta,\phi)=\Theta(\theta)\Phi(\phi). \tag{3.54}$$

Then, we have

$$-\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\Theta(\theta)}{\partial\theta}\right)\Phi(\phi)+\frac{1}{\sin^2\theta}\frac{\partial^2\Phi(\phi)}{\partial\phi^2}\right]\Theta(\theta)=\lambda\Theta(\theta)\Phi(\phi). \tag{3.55}$$

Multiplying both sides by $\sin^2\theta/\Theta(\theta)\Phi(\phi)$ and arranging both the sides, we get

$$-\frac{1}{\Phi(\phi)}\frac{\partial^2\Phi(\phi)}{\partial\phi^2}=\frac{\sin^2\theta}{\Theta(\theta)}\left\{\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left[\sin\theta\frac{\partial\Theta(\theta)}{\partial\theta}\right]+\lambda\Theta(\theta)\right\}. \tag{3.56}$$

Since LHS of (3.56) depends only upon $\phi$ and RHS depends only on $\theta$, we must have

$$\text{LHS of } (3.56) = \text{RHS of } (3.56) = \eta \text{ (constant)}. \tag{3.57}$$

Thus, we have a following relation of LHS of (3.56):

$$-\frac{1}{\Phi(\phi)}\frac{\mathrm{d}^2\Phi(\phi)}{\mathrm{d}\phi^2} = \eta. \tag{3.58}$$

Putting $D \equiv -\frac{\mathrm{d}^2}{\mathrm{d}\phi^2}$, we get

$$D\Phi(\phi) = \eta\Phi(\phi). \tag{3.59}$$

The SOLDEs of (3.58) and (3.59) are formally the same as (1.61) of Sect. 1.3, where boundary conditions (BCs) are Dirichlet conditions. Unlike (1.61), however, we have to consider different BCs; i.e., the periodic BCs.

As in Example 1.1, we adopt two linearly independent solutions. That is, we have

$$\mathrm{e}^{im\phi}, \quad \text{and } \mathrm{e}^{-im\phi} \ (m \neq 0).$$

As their linear combination, we have

$$\Phi(\phi) = a\mathrm{e}^{im\phi} + b\mathrm{e}^{-im\phi}. \tag{3.60}$$

As BCs, we consider $\Phi(0) = \Phi(2\pi)$ and $\Phi'(0) = \Phi'(2\pi)$; i.e., we have

$$a + b = a\mathrm{e}^{i2\pi m} + b\mathrm{e}^{-i2\pi m}. \tag{3.61}$$

Meanwhile, we have

$$\Phi'(\phi) = aim\mathrm{e}^{im\phi} - bim\mathrm{e}^{-im\phi}. \tag{3.62}$$

Therefore, from BCs we have

$$aim - bim = aim\mathrm{e}^{i2\pi m} - bim\mathrm{e}^{-i2\pi m}.$$

Then,

$$a - b = a\mathrm{e}^{i2\pi m} - b\mathrm{e}^{-i2\pi m}. \tag{3.63}$$

From (3.61) to (3.63), we have

$$2a(1 - \mathrm{e}^{i2\pi m}) = 0, \quad \text{and } 2b(1 - \mathrm{e}^{-i2\pi m}) = 0.$$

If $a \neq 0$, we must have $m = 0, \pm 1, \pm 2, \ldots$. If $a = 0$, we must have $b \neq 0$ to avoid having $\Phi(\phi) \equiv 0$ as a solution. In that case, we have $m = 0, \pm 1, \pm 2, \ldots$ as well. Thus, it suffices to put $\Phi(\phi) = ce^{im\phi} (m = 0, \pm 1, \pm 2, \ldots)$. Therefore, as a normalized function $\tilde{\Phi}(\phi)$, we get

$$\tilde{\Phi}(\phi) = \frac{1}{\sqrt{2\pi}} e^{im\phi} (m = 0, \pm 1, \pm 2, \ldots). \tag{3.64}$$

Inserting it into (3.58), we have

$$m^2 e^{im\phi} = \eta e^{im\phi}.$$

Therefore, we get

$$\eta = m^2 \ (m = 0, \pm 1, \pm 2, \ldots). \tag{3.65}$$

From (3.56) to (3.65), we have

$$-\frac{1}{\sin\theta}\frac{d}{d\theta}\left[\sin\theta \frac{d\Theta(\theta)}{d\theta}\right] + \frac{m^2\Theta(\theta)}{\sin^2\theta} = \lambda\Theta(\theta) (m = 0, \pm 1, \pm 2, \ldots). \tag{3.66}$$

In (3.64) putting $m = 0$ as an eigenvalue, we have $\Phi(\phi) = 1/\sqrt{2\pi}$ as a corresponding eigenfunction. Unlike Examples 1.1 and 1.2, this reflects that the differential operator $-\frac{d^2}{d\phi^2}$ accompanied by the periodic BCs is a nonnegative operator that allows an eigenvalue of zero. Yet, we are uncertain of a range of $m$. To clarify this point, we consider generalized angular momentum in the next section.

## 3.4 Generalized Angular Momentum

We obtained commutation relations of (3.30) among individual angular momentum components $L_x$, $L_y$, and $L_z$. In an opposite way, we may start with (3.30) to define angular momentum. Such a quantity is called generalized angular momentum.

Let $\tilde{\boldsymbol{J}}$ be a generalized angular momentum as in the case of (3.4) such that

$$\tilde{\boldsymbol{J}} = (\boldsymbol{e}_1 \boldsymbol{e}_2 \boldsymbol{e}_3) \begin{pmatrix} \tilde{J}_x \\ \tilde{J}_y \\ \tilde{J}_z \end{pmatrix}. \tag{3.67}$$

For the sake of simple notation, let us define $\boldsymbol{J}$ as follows so that we can eliminate $\hbar$ and deal with dimensionless quantities in the present discussion:

$$\boldsymbol{J} \equiv \tilde{\boldsymbol{J}}/\hbar = (\boldsymbol{e}_1\boldsymbol{e}_2\boldsymbol{e}_3)\begin{pmatrix} \tilde{J}_x/\hbar \\ \tilde{J}_y/\hbar \\ \tilde{J}_z/\hbar \end{pmatrix} = (\boldsymbol{e}_1\boldsymbol{e}_2\boldsymbol{e}_3)\begin{pmatrix} J_x \\ J_y \\ J_z \end{pmatrix}, \tag{3.68}$$
$$\boldsymbol{J}^2 = J_x^2 + J_y^2 + J_z^2.$$

Then, we require following commutation relations:

$$[J_x, J_y] = iJ_z, [J_y, J_z] = iJ_x, \text{ and } [J_z, J_x] = iJ_y. \tag{3.69}$$

Also, we require $J_x$, $J_y$, and $J_z$ to be Hermitian. The operator $\boldsymbol{J}^2$ is Hermitian accordingly. The relations (3.69) lead to

$$[J_x, \boldsymbol{J}^2] = 0, [J_y, \boldsymbol{J}^2] = 0, \text{ and } [J_z, \boldsymbol{J}^2] = 0. \tag{3.70}$$

This can be confirmed as in the case of (3.30).

As noted above, again a simultaneous eigenstate exists for $\boldsymbol{J}^2$ and one of $J_x$, $J_y$, and $J_z$. According to the convention, we choose $\boldsymbol{J}^2$ and $J_z$ for the simultaneous eigenstate. Then, designating the eigenstate by $|\zeta, \mu\rangle$, we have

$$\boldsymbol{J}^2|\zeta, \mu\rangle = \zeta|\zeta, \mu\rangle \quad \text{and } J_z|\zeta, \mu\rangle = \mu|\zeta, \mu\rangle. \tag{3.71}$$

The implication of (3.71) is that $|\zeta, \mu\rangle$ is the simultaneous eigenstate and that $\mu$ is an eigenvalue of $J_z$ which $|\zeta, \mu\rangle$ belongs to with $\zeta$ being an eigenvalue of $\boldsymbol{J}^2$ which $|\zeta, \mu\rangle$ belongs to as well.

Since $J_z$ and $\boldsymbol{J}^2$ are Hermitian, both $\mu$ and $\zeta$ are real (see Sect. 1.4). Of these, $\zeta \geq 0$ as in the case of (3.45). We define following operators $J^{(+)}$ and $J^{(-)}$ as in the case of (3.27):

$$J^{(+)} \equiv J_x + iJ_y \quad \text{and } J^{(-)} \equiv J_x - iJ_y. \tag{3.72}$$

Then, from (3.69) to (3.70), we get

$$\left[J^{(+)}, \boldsymbol{J}^2\right] = \left[J^{(-)}, \boldsymbol{J}^2\right] = 0. \tag{3.73}$$

Also, we obtain following commutation relations:

$$\left[J_z, J^{(+)}\right] = J^{(+)}; \left[J_z, J^{(-)}\right] = -J^{(-)}; \left[J^{(+)}, J^{(-)}\right] = 2J_z. \tag{3.74}$$

From (3.70) to (3.72), we get

$$\begin{aligned} \boldsymbol{J}^2 J^{(+)}|\zeta, \mu\rangle &= J^{(+)}\boldsymbol{J}^2|\mu\rangle = \zeta J^{(+)}|\zeta, \mu\rangle, \\ \boldsymbol{J}^2 J^{(-)}|\zeta, \mu\rangle &= J^{(-)}\boldsymbol{J}^2|\zeta, \mu\rangle = \zeta J^{(-)}|\zeta, \mu\rangle. \end{aligned} \tag{3.75}$$

Equation (3.75) indicates that both $J^{(+)}|\zeta,\mu\rangle$ and $J^{(-)}|\zeta,\mu\rangle$ are eigenvectors of $\boldsymbol{J}^2$ that correspond to an eigenvalue $\zeta$.

Meanwhile, from (3.74) we get

$$\begin{aligned} J_z J^{(+)}|\zeta,\mu\rangle &= J^{(+)}(J_z+1)|\zeta,\mu\rangle = (\mu+1)J^{(+)}|\zeta,\mu\rangle, \\ J_z J^{(-)}|\zeta,\mu\rangle &= J^{(-)}(J_z-1)|\zeta,\mu\rangle = (\mu-1)J^{(-)}|\zeta,\mu\rangle. \end{aligned} \tag{3.76}$$

The relation (3.76) means that $J^{(+)}|\zeta,\mu\rangle$ is an eigenvector of $J_z$ corresponding to an eigenvalue $(\mu+1)$, while $J^{(-)}|\zeta,\mu\rangle$ is an eigenvector of $J_z$ corresponding to an eigenvalue $(\mu-1)$. This implies that $J^{(+)}$ and $J^{(-)}$ function as raising and lowering operators (or ladder operators) that have been introduced in this chapter. Thus, using undetermined constants (or phase factors) $a_\mu^{(+)}$ and $a_\mu^{(-)}$, we describe

$$J^{(+)}|\zeta,\mu\rangle = a_\mu^{(+)}|\zeta,\mu+1\rangle \quad \text{and } J^{(-)}|\zeta,\mu\rangle = a_\mu^{(-)}|\zeta,\mu-1\rangle. \tag{3.77}$$

Next, let us characterize eigenvalues $\mu$. We have

$$J_x^2 + J_y^2 = \boldsymbol{J}^2 - J_z^2. \tag{3.78}$$

Therefore,

$$(J_x^2+J_y^2)|\zeta,\mu\rangle = (\boldsymbol{J}^2 - J_z^2)|\zeta,\mu\rangle = (\zeta-\mu^2)|\zeta,\mu\rangle. \tag{3.79}$$

Since $(J_x^2+J_y^2)$ is a nonnegative operator, its eigenvalues are nonnegative as well, as can be seen from (3.40) to (3.45). Then, we have

$$\zeta - \mu^2 \geq 0. \tag{3.80}$$

Thus, for a fixed value of nonnegative $\zeta$, $\mu$ is bounded both upward and downward. We define then a maximum of $\mu$ as $j$ and a minimum of $\mu$ as $j'$. Consequently, on the basis of (3.77), we have

$$J^{(+)}|\zeta,j\rangle = 0 \quad \text{and } J^{(-)}|\zeta,j'\rangle = 0. \tag{3.81}$$

This is because we have no quantum state corresponding to $|\zeta,j+1\rangle$ or $|\zeta,j'-1\rangle$. From (3.75) to (3.81), possible numbers of $\mu$ are

$$j, j-1, j-2, \ldots, j'. \tag{3.82}$$

From (3.69) to (3.72), we get

$$J^{(-)}J^{(+)} = \boldsymbol{J}^2 - J_z^2 - J_z, \quad J^{(+)}J^{(-)} = \boldsymbol{J}^2 - J_z^2 + J_z. \tag{3.83}$$

Operating these operators on $|\zeta,j\rangle$ or $|\zeta,j'\rangle$ and using (3.81) we get

$$\begin{aligned} J^{(-)}J^{(+)}|\zeta,j\rangle &= (\boldsymbol{J}^2 - J_z^2 - J_z)|\zeta,j\rangle = (\zeta - j^2 - j)|\zeta,j\rangle = 0, \\ J^{(+)}J^{(-)}|\zeta,j'\rangle &= (\boldsymbol{J}^2 - J_z^2 + J_z)|\zeta,j'\rangle = (\zeta - j'^2 + j')|\zeta,j'\rangle = 0. \end{aligned} \tag{3.84}$$

Since $|\zeta,j\rangle \neq 0$ and $|\zeta,j'\rangle \neq 0$, we have

$$\zeta - j^2 - j = \zeta - j'^2 + j' = 0. \tag{3.85}$$

This means that

$$\zeta = j(j+1) = j'(j'-1). \tag{3.86}$$

Moreover, from (3.86) we get

$$j(j+1) - j'(j'-1) = (j+j')(j-j'+1) = 0. \tag{3.87}$$

As $j \geq j'$, $j - j' + 1 > 0$. From (3.87), therefore, we get

$$j + j' = 0 \text{ or } j = -j'. \tag{3.88}$$

Then, we conclude that the minimum of $\mu$ is $-j$. Accordingly, possible values of $\mu$ are

$$\mu = j, j-1, j-2, \ldots, -j-1, -j. \tag{3.89}$$

That is, the number $\mu$ can take is $(2j+1)$. The relation (3.239) implies that taking a positive integer $k$,

$$j - k = -j \text{ or } j = k/2. \tag{3.90}$$

In other words, $j$ is permitted to take a number zero, a positive integer, or a positive half-integer (or more precisely, half-odd-integer). For instance, if $j = 1/2$, $\mu$ can be $1/2$ or $-1/2$. When $j = 1$, $\mu$ can be $1, 0,$ or $-1$.

Finally, we have to decide undetermined constants $a_\mu^{(+)}$ and $a_\mu^{(-)}$. To this end, multiplying $\langle\zeta, \mu - 1|$ on both sides of the second equation of (3.77) from the left, we have

$$\langle\zeta, \mu-1|J^{(-)}|\zeta,\mu\rangle = a_\mu^{(-)}\langle\zeta, \mu-1 \mid \zeta, \mu-1\rangle = a_\mu^{(-)}, \tag{3.91}$$

where the second equality comes from that $|\zeta, \mu - 1\rangle$ has been normalized; i.e., $|||\zeta, \mu - 1\rangle|| = 1$. Meanwhile, taking adjoint of both sides of the first equation of (3.77), we have

$$\langle\zeta,\mu|[J^{(+)}]^\dagger = \left[a_\mu^{(+)}\right]^* \langle\zeta, \mu+1|. \tag{3.92}$$

But, from (3.72) and the fact that $J_x$ and $J_y$ are Hermitian,

$$[\boldsymbol{J}^{(+)}]^{\dagger} = \boldsymbol{J}^{(-)}. \tag{3.93}$$

Using (3.93) and replacing $\mu$ in (3.92) with $\mu - 1$, we get

$$\langle \zeta, \mu - 1 | \boldsymbol{J}^{(-)} = \left[ a_{\mu-1}^{(+)} \right]^* \langle \zeta, \mu |. \tag{3.94}$$

Furthermore, multiplying $|\zeta, \mu\rangle$ on (3.94) from the right, we have

$$\langle \zeta, \mu - 1 | \boldsymbol{J}^{(-)} | \zeta, \mu \rangle = \left[ a_{\mu-1}^{(+)} \right]^* \langle \zeta, \mu \mid \zeta, \mu \rangle = \left[ a_{\mu-1}^{(+)} \right]^*, \tag{3.95}$$

where again $|\zeta, \mu\rangle$ is assumed to be normalized. Comparing (3.91) and (3.95), we get

$$a_{\mu}^{(-)} = \left[ a_{\mu-1}^{(+)} \right]^*. \tag{3.96}$$

Taking an inner product regarding the first equation of (3.77) and its adjoint,

$$\langle \zeta, \mu | \boldsymbol{J}^{(-)} \boldsymbol{J}^{(+)} | \zeta, \mu \rangle = \left[ a_{\mu}^{(+)} \right]^* a_{\mu}^{(+)} \langle \zeta, \mu + 1 \mid \zeta, \mu + 1 \rangle = \left| a_{\mu}^{(+)} \right|^2. \tag{3.97}$$

Once again, the second equality of (3.97) results from the normalization of the vector.

Using (3.83) as well as (3.71) and (3.86), (3.97) can be rewritten as

$$\begin{aligned} &\langle \zeta, \mu | \boldsymbol{J}^2 - J_z^2 - J_z | \zeta, \mu \rangle \\ &= \langle \zeta, \mu | j(j+1) - \mu^2 - \mu | \zeta, \mu \rangle = \langle \zeta, \mu \mid \zeta, \mu \rangle (j - \mu)(j + \mu + 1) = |a_{\mu}^{(+)}|^2. \end{aligned} \tag{3.98}$$

Thus, we get

$$a_{\mu}^{(+)} = \mathrm{e}^{i\delta} \sqrt{(j - \mu)(j + \mu + 1)} \ (\delta : \text{an arbitrary real number}), \tag{3.99}$$

where $\mathrm{e}^{i\delta}$ is a phase factor. From (3.96), we also get

$$a_{\mu}^{(-)} = \mathrm{e}^{-i\delta} \sqrt{(j - \mu + 1)(j + \mu)}. \tag{3.100}$$

In (3.99) and (3.100), we routinely put $\delta = 0$ so that $a_{\mu}^{(+)}$ and $a_{\mu}^{(-)}$ can be positive numbers. Explicitly rewriting (3.77), we get

$$\begin{aligned} \boldsymbol{J}^{(+)} | \zeta, \mu \rangle &= \sqrt{(j - \mu)(j + \mu + 1)} | \zeta, \mu + 1 \rangle, \\ \boldsymbol{J}^{(-)} | \zeta, \mu \rangle &= \sqrt{(j - \mu + 1)(j + \mu)} | \zeta, \mu - 1 \rangle, \end{aligned} \tag{3.101}$$

where $j$ is a fixed given number chosen from among zero, positive integers, and positive half-integers (or half-odd-integers).

## 3.5 Orbital Angular Momentum: Operator Approach

In Sect. 3.4, we have derived various important results on angular momenta on the basis of the commutation relations (3.69) and the assumption that $J_x$, $J_y$, and $J_z$ are Hermitian. Now, let us return to the discussion on orbital angular momenta we dealt with in Sects. 3.2 and 3.3. First, we treat the orbital angular momenta via operator approach. This approach enables us to understand why a quantity $j$ introduced in Sect. 3.4 takes a value zero or positive integers with the orbital angular momenta. In the next section (Sect. 3.6), we will deal with the related issues by an analytical method.

In (3.28), we introduced differential operators $L^{(+)}$ and $L^{(-)}$. According to Sect. 3.4, we define following operators to eliminate $\hbar$ so that we can deal with dimensionless quantities:

$$\begin{aligned} \boldsymbol{M} \equiv \boldsymbol{L}/\hbar &= (\boldsymbol{e}_1 \boldsymbol{e}_2 \boldsymbol{e}_3) \begin{pmatrix} M_x \\ M_y \\ M_z \end{pmatrix}, \\ \boldsymbol{M}^2 = \boldsymbol{L}^2/\hbar^2 &= M_x^2 + M_y^2 + M_z^2. \end{aligned} \tag{3.102}$$

Hence, we have

$$M_x = L_x/\hbar, M_y = L_y/\hbar, \quad \text{and } M_z = L_z/\hbar. \tag{3.103}$$

Moreover, we define following operators:

$$M^{(+)} \equiv M_x + iM_y = L^{(+)}/\hbar = \mathrm{e}^{i\phi}\left(\frac{\partial}{\partial\theta} + i\cot\theta\frac{\partial}{\partial\phi}\right), \tag{3.104}$$

$$M^{(-)} \equiv L^{(-)}/\hbar = \mathrm{e}^{-i\phi}\left(-\frac{\partial}{\partial\theta} + i\cot\theta\frac{\partial}{\partial\phi}\right). \tag{3.105}$$

Then, we have

$$\boldsymbol{M}^2 = -\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right]. \tag{3.106}$$

Here, we execute variable transformation such that

$$\xi = \cos\theta (0 \le \theta \le \pi) \text{ or } \sin\theta = \sqrt{1-\xi^2}. \tag{3.107}$$

Noting, e.g., that

$$\frac{\partial}{\partial\theta} = \frac{\partial\xi}{\partial\theta}\frac{\partial}{\partial\xi} = -\sin\theta\frac{\partial}{\partial\xi} = -\sqrt{1-\xi^2}\frac{\partial}{\partial\xi}, \sin\theta\frac{\partial}{\partial\theta} = -\sin^2\theta\frac{\partial}{\partial\xi} = -(1-\xi^2)\frac{\partial}{\partial\xi}, \tag{3.108}$$

we get

$$\begin{aligned} \boldsymbol{M}^{(+)} &= \mathrm{e}^{i\phi}\left(-\sqrt{1-\xi^2}\frac{\partial}{\partial\xi} + i\frac{\xi}{\sqrt{1-\xi^2}}\frac{\partial}{\partial\phi}\right), \\ \boldsymbol{M}^{(-)} &= \mathrm{e}^{-i\phi}\left(\sqrt{1-\xi^2}\frac{\partial}{\partial\xi} + i\frac{\xi}{\sqrt{1-\xi^2}}\frac{\partial}{\partial\phi}\right), \\ \boldsymbol{M}^2 &= -\frac{\partial}{\partial\xi}\left[(1-\xi^2)\frac{\partial}{\partial\xi}\right] - \frac{1}{1-\xi^2}\frac{\partial^2}{\partial\phi^2}. \end{aligned} \tag{3.109}$$

Although we showed in Sect. 3.3 that $m = 0, \pm 1, \pm 2, \ldots$, the range of $m$ was unclear. The relationship between $m$ and $\lambda$ in (3.66) remains unclear so far as well. On the basis of a general approach developed in Sect. 3.4, however, we have known that the eigenvalue $\mu$ of the dimensionless $z$-component angular momentum $J_z$ is bounded with its maximum and minimum being $j$ and $-j$, respectively see (3.89), where $j$ can be zero, a positive integer, or a positive half-odd-integer. Concomitantly, the eigenvalue $\zeta$ of $\boldsymbol{J}^2$ equals $j(j+1)$.

In the present section, let us reconsider the relationship between $m$ and $\lambda$ in (3.66) in light of the knowledge obtained in Sect. 3.4. According to the custom, we replace $\mu$ in (3.89) with $m$ to have

$$m = j, j-1, j-2, \ldots, j-1, -j. \tag{3.110}$$

At the moment, we assume that $m$ can be a half-odd-integer besides zero or an integer [3].

Now, let us define notation of $Y(\theta, \phi)$ that appeared in (3.37). This function is eligible for a simultaneous eigenstate of $\boldsymbol{M}^2$ and $M_z$ and can be indexed with $j$ and $m$ as in (3.110). Then, let $Y(\theta, \phi)$ be described accordingly as

$$Y_j^m(\theta, \phi) \equiv Y(\theta, \phi). \tag{3.111}$$

From (3.54) to (3.64), we have

$$Y_j^m(\theta, \phi) \propto \mathrm{e}^{im\phi}.$$

Therefore, we get

$$\begin{aligned} M^{(+)}Y_j^m(\theta,\phi) &= \mathrm{e}^{i\phi}\left(-\sqrt{1-\xi^2}\,\frac{\partial}{\partial\xi}-\frac{m\xi}{\sqrt{1-\xi^2}}\right)Y_j^m(\theta,\phi) \\ &= -\mathrm{e}^{i\phi}\left(\sqrt{1-\xi^2}\right)^{m+1}\frac{\partial}{\partial\xi}\left[\left(\sqrt{1-\xi^2}\right)^{-m}Y_j^m(\theta,\phi)\right], \end{aligned} \tag{3.112}$$

where we used the following equation:

$$\begin{aligned} \frac{\partial}{\partial\xi}\left[\left(\sqrt{1-\xi^2}\right)\right]^{-m} &= (-m)\left(\sqrt{1-\xi^2}\right)^{-m-1}\cdot\frac{1}{2}\left(\sqrt{1-\xi^2}\right)^{-1}(-2\xi) \\ &= m\xi\left(\sqrt{1-\xi^2}\right)^{-m-2}, \\ \frac{\partial}{\partial\xi}\left[\left(\sqrt{1-\xi^2}\right)^{-m}Y_j^m(\theta,\phi)\right] &= m\xi\left(\sqrt{1-\xi^2}\right)^{-m-2}Y_j^m(\theta,\phi) \\ &\quad+\left(\sqrt{1-\xi^2}\right)^{-m}\frac{\partial Y_j^m(\theta,\phi)}{\partial\xi}. \end{aligned} \tag{3.113}$$

Similarly, we get

$$\begin{aligned} M^{(-)}Y_j^m(\theta,\phi) &= \mathrm{e}^{-i\phi}\left(\sqrt{1-\xi^2}\,\frac{\partial}{\partial\xi}-\frac{m\xi}{\sqrt{1-\xi^2}}\right)Y_j^m(\theta,\phi) \\ &= \mathrm{e}^{-i\phi}\left(\sqrt{1-\xi^2}\right)^{-m+1}\frac{\partial}{\partial\xi}\left[\left(\sqrt{1-\xi^2}\right)^{m}Y_j^m(\theta,\phi)\right]. \end{aligned} \tag{3.114}$$

Let us derive the relations where $M^{(+)}$ or $M^{(-)}$ is successively operated on $Y_j^m(\theta,\phi)$. In the case of $M^{(+)}$, using (3.109) we have

$$[M^{(+)}]^n Y_j^m(\theta,\phi) = (-1)^n\mathrm{e}^{in\phi}\left(\sqrt{1-\xi^2}\right)^{m+n}\frac{\partial^n}{\partial\xi^n}\left[\left(\sqrt{1-\xi^2}\right)^{-m}Y_j^m(\theta,\phi)\right]. \tag{3.115}$$

We confirm this relation by mathematical induction. We have (3.112) by replacing $n$ with 1 in (3.115). Namely, (3.115) holds when $n = 1$. Next, suppose that (3.115) holds with $n$. Then, using the first equation of (3.109) and noting (3.64), we have

$$
\begin{aligned}
&[M^{(+)}]^{n+1} Y_j^m(\theta,\phi) = M^{(+)}\left\{[M^{(+)}]^n Y_j^m(\theta,\phi)\right\} \\
&= \mathrm{e}^{i\phi}\left(-\sqrt{1-\xi^2}\,\frac{\partial}{\partial\xi} + i\frac{\xi}{\sqrt{1-\xi^2}}\frac{\partial}{\partial\phi}\right) \\
&\quad\times (-1)^n \mathrm{e}^{in\phi}\left(\sqrt{1-\xi^2}\right)^{m+n}\frac{\partial^n}{\partial\xi^n}\left[\left(\sqrt{1-\xi^2}\right)^{-m} Y_j^m(\theta,\phi)\right] \\
&= (-1)^n \mathrm{e}^{i(n+1)\phi}\left[-\sqrt{1-\xi^2}\,\frac{\partial}{\partial\xi} - \frac{\xi(n+m)}{\sqrt{1-\xi^2}}\right] \\
&\quad\times \left(\sqrt{1-\xi^2}\right)^{m+n}\frac{\partial^n}{\partial\xi^n}\left[\left(\sqrt{1-\xi^2}\right)^{-m} Y_j^m(\theta,\phi)\right] \\
&= (-1)^n \mathrm{e}^{i(n+1)\phi}\left\{-\sqrt{1-\xi^2}\left[(m+n)\left(\sqrt{1-\xi^2}\right)^{m+n-1}\cdot\frac{1}{2}\left(\sqrt{1-\xi^2}\right)^{-1}(-2\xi)\right]\right. \\
&\quad\times \frac{\partial^n}{\partial\xi^n}\left[\left(\sqrt{1-\xi^2}\right)^{-m} Y_j^m(\theta,\phi)\right] \\
&\quad - \sqrt{1-\xi^2}\left(\sqrt{1-\xi^2}\right)^{m+n}\frac{\partial^{n+1}}{\partial\xi^{n+1}}\left[\left(\sqrt{1-\xi^2}\right)^{-m} Y_j^m(\theta,\phi)\right] \\
&\quad \left. - \frac{\xi(n+m)}{\sqrt{1-\xi^2}}\left(\sqrt{1-\xi^2}\right)^{m+n}\frac{\partial^n}{\partial\xi^n}\left[\left(\sqrt{1-\xi^2}\right)^{-m} Y_j^m(\theta,\phi)\right]\right\} \\
&= (-1)^{n+1}\mathrm{e}^{i(n+1)\phi}\left(\sqrt{1-\xi^2}\right)^{m+(n+1)}\frac{\partial^{n+1}}{\partial\xi^{n+1}}\left[\left(\sqrt{1-\xi^2}\right)^{-m} Y_j^m(\theta,\phi)\right].
\end{aligned}
\tag{3.116}
$$

Notice that the first and third terms in the second last equality canceled each other. Thus, (3.115) certainly holds with $(n+1)$. Similarly, we have [3]

$$
[M^{(-)}]^n Y_j^m(\theta,\phi) = \mathrm{e}^{-in\phi}\left(\sqrt{1-\xi^2}\right)^{-m+n}\frac{\partial^n}{\partial\xi^n}\left[\left(\sqrt{1-\xi^2}\right)^{m} Y_j^m(\theta,\phi)\right].
\tag{3.117}
$$

Proof of (3.117) is left for readers.

From the second equation of (3.81) and (3.114) where $m$ is replaced with $-j$, we have

$$
M^{(-)} Y_j^{-j}(\theta,\phi) = \mathrm{e}^{-i\phi}\left(\sqrt{1-\xi^2}\right)^{j+1}\frac{\partial}{\partial\xi}\left[\left(\sqrt{1-\xi^2}\right)^{-j} Y_j^{-j}(\theta,\phi)\right] = 0.
\tag{3.118}
$$

This implies that $\left(\sqrt{1-\xi^2}\right)^{-j} Y_j^{-j}(\theta,\phi)$ is a constant with respect to $\xi$. We describe this as

$$\left(\sqrt{1-\xi^2}\right)^{-j} Y_j^{-j}(\theta,\phi) = c \ (c : \text{constant with respect to } \xi). \tag{3.119}$$

Meanwhile, putting $m = -j$ and $n = 2j+1$ in (3.115) and taking account of the first equation of (3.77) and the first equation of (3.81), we get

$$\begin{aligned}&[M^{(+)}]^{2j+1} Y_j^{-j}(\theta,\phi)\\ &= (-1)^{2j+1} \mathrm{e}^{i(2j+1)\phi} \left(\sqrt{1-\xi^2}\right)^{j+1} \frac{\partial^{2j+1}}{\partial \xi^{2j+1}} \left[\left(\sqrt{1-\xi^2}\right)^{j} Y_j^{-j}(\theta,\phi)\right] = 0.\end{aligned} \tag{3.120}$$

This means that

$$\left(\sqrt{1-\xi^2}\right)^{j} Y_j^{-j}(\theta,\phi) = (\text{at most a } 2j\text{-degree polynomial with } \xi). \tag{3.121}$$

Replacing $Y_j^{-j}(\theta,\phi)$ in (3.121) with that of (3.119), we get

$$\begin{aligned}&c\left(\sqrt{1-\xi^2}\right)^{j}\left(\sqrt{1-\xi^2}\right)^{j} = c(1-\xi^2)^j\\ &= (\text{at most a } 2j\text{-degree polynomial with } \xi).\end{aligned} \tag{3.122}$$

Here, if $j$ is a half-odd-integer, $c(1-\xi^2)^j$ of (3.122) cannot be a polynomial. If on the other hand $j$ is zero or a positive integer, $c(1-\xi^2)^j$ is certainly a polynomial and, to top it all, a $2j$-degree polynomial with respect to $\xi$; so is $\left(\sqrt{1-\xi^2}\right)^{j} Y_j^{-j}(\theta,\phi)$.

According to the custom, henceforth we use $l$ as zero or a positive integer instead of $j$. That is,

$$Y(\theta,\phi) \equiv Y_l^m(\theta,\phi) \ (l : \text{zero or a positive integer}). \tag{3.123}$$

At the same time, so far as the orbital angular momentum is concerned, from (3.71) to (3.86) we can identify $\zeta$ in (3.71) with $l(l+1)$. Namely, we have

$$\zeta = l(l+1).$$

Concomitantly, $m$ in (3.110) is determined as

$$m = l, l-1, l-2, \ldots 1, 0, -1, \ldots -l+1, -l. \tag{3.124}$$

Thus, as expected $m$ is zero or a positive or negative integer. Considering (3.37) and (3.46), $\zeta$ is identical with $\lambda$ in (3.46). Finally, we rewrite (3.66) such that

$$-\frac{1}{\sin\theta}\frac{\mathrm{d}}{\mathrm{d}\theta}\left[\sin\theta\frac{\mathrm{d}\Theta(\theta)}{\mathrm{d}\theta}\right] + \frac{m^2\Theta(\theta)}{\sin^2\theta} = l(l+1)\Theta(\theta), \tag{3.125}$$

where $l$ is equal to zero or positive integers and $m$ is given by (3.124).

On condition of $\xi = \cos\theta$ (3.107), defining the following function

$$P_l^m(\xi) \equiv \Theta(\theta), \tag{3.126}$$

and considering (3.109) along with (3.54), we arrive at the next SOLDE described as

$$\frac{\mathrm{d}}{\mathrm{d}\xi}\left[(1-\xi^2)\frac{\mathrm{d}P_l^m(\xi)}{\mathrm{d}\xi}\right] + \left[l(l+1) - \frac{m^2}{1-\xi^2}\right]P_l^m(\xi) = 0. \tag{3.127}$$

The SOLDE of (3.127) is well-known as the associated Legendre differential equation. The solutions $P_l^m(\xi)$ are called associated Legendre functions.

In the next section, we characterize the said equation and functions by an analytical method. Before going into details, however, we further seek characteristics of $P_l^m(\xi)$ by the operator approach.

Adopting the notation of (3.123) and putting $m = l$ in (3.112), we have

$$M^{(+)}Y_l^l(\theta,\phi) = -\mathrm{e}^{i\phi}\left(\sqrt{1-\xi^2}\right)^{l+1}\frac{\partial}{\partial\xi}\left[\left(\sqrt{1-\xi^2}\right)^{-l}Y_l^l(\theta,\phi)\right]. \tag{3.128}$$

Corresponding to (3.81), we have $M^{(+)}Y_l^l(\theta,\phi) = 0$. This implies that

$$\left(\sqrt{1-\xi^2}\right)^{-l}Y_l^l(\theta,\phi) = c \ (c : \text{constant with respect to}, \xi). \tag{3.129}$$

From (3.107) to (3.64), we get

$$Y_l^l(\theta,\phi) = \kappa_l \sin^l\theta \mathrm{e}^{il\phi}, \tag{3.130}$$

where $\kappa_l$ is another constant that depends on $l$, but is independent of $\theta$ and $\phi$. Let us seek $\kappa_l$ by normalization condition. That is,

$$\int_0^{2\pi}\mathrm{d}\phi\int_0^{\pi}\sin\theta\mathrm{d}\theta\left|Y_l^l(\theta,\phi)\right|^2 = 2\pi\cdot\left|\kappa_l\right|^2\int_0^{\pi}\sin^{2l+1}\theta\mathrm{d}\theta = 1, \tag{3.131}$$

where the integration is performed on a unit sphere. Note that an infinitesimal area element on the unit sphere is represented by $\sin\theta \mathrm{d}\theta \mathrm{d}\phi$.

We evaluate the above integral denoted as

$$I \equiv \int_0^\pi \sin^{2l+1}\theta \mathrm{d}\theta. \tag{3.132}$$

Using integration by parts,

$$\begin{aligned} I &= \int_0^\pi (-\cos\theta)' \sin^{2l}\theta \mathrm{d}\theta \\ &= [(-\cos\theta)\sin^{2l}\theta]_0^\pi + \int_0^\pi (\cos\theta)\cdot 2l \cdot \sin^{2l-1}\theta\cos\theta \mathrm{d}\theta \\ &= 2l\int_0^\pi \sin^{2l-1}\theta \mathrm{d}\theta - 2l\int_0^\pi \sin^{2l+1}\theta \mathrm{d}\theta. \end{aligned} \tag{3.133}$$

Thus, we get a recurrence relation with respect to $I$ (3.132) such that

$$I = \frac{2l}{2l+1}\int_0^\pi \sin^{2l-1}\theta \mathrm{d}\theta. \tag{3.134}$$

Repeating the above process, we get

$$I = \frac{2l}{2l+1}\cdot\frac{2l-2}{2l-1}\cdots\frac{2}{3}\int_0^\pi \sin\theta \mathrm{d}\theta = \frac{2^{2l+1}(l!)^2}{(2l+1)!}. \tag{3.135}$$

Then,

$$|\kappa_l| = \frac{1}{2^l l!}\sqrt{\frac{(2l+1)!}{4\pi}} \text{ or } \kappa_l = \frac{\mathrm{e}^{i\chi}}{2^l l!}\sqrt{\frac{(2l+1)!}{4\pi}} \ (\chi : \text{real}), \tag{3.136}$$

where $\mathrm{e}^{i\chi}$ is an undetermined constant (phase factor) that is to be determined below. Thus we get

$$Y_l^l(\theta,\phi) = \frac{\mathrm{e}^{i\chi}}{2^l l!}\sqrt{\frac{(2l+1)!}{4\pi}}\sin^l\theta\,\mathrm{e}^{il\phi}. \tag{3.137}$$

Meanwhile, in the second equation of (3.101) replacing $J^{(-)}$, $j$, and $\mu$ in (3.101) with $M^{(-)}$, $l$, and $m$, respectively, and using $Y_l^m(\theta,\phi)$ instead of $|\zeta,\mu\rangle$, we get

$$Y_l^{m-1}(\theta,\phi) = \frac{1}{\sqrt{(l-m+1)(l+m)}} M^{(-)} Y_l^m(\theta,\phi). \tag{3.138}$$

Replacing $m$ with $l$ in (3.138), we have

$$Y_l^{l-1}(\theta,\phi) = \frac{1}{\sqrt{2l}} M^{(-)} Y_l^l(\theta,\phi). \tag{3.139}$$

Operating $M^{(-)}$ $(l-m)$ times in total on $Y_l^l(\theta,\phi)$ of (3.139), we have

$$\begin{aligned} Y_l^m(\theta,\phi) &= \frac{1}{\sqrt{2l(2l-1)\ldots(l+m+1)}\sqrt{1\cdot 2\ldots(l-m)}} [M^{(-)}]^{l-m} Y_l^l(\theta,\phi) \\ &= \sqrt{\frac{(l+m)!}{(2l)!(l-m)!}} [M^{(-)}]^{l-m} Y_l^l(\theta,\phi). \end{aligned} \tag{3.140}$$

Meanwhile, putting $m = l$, $n = l - m$, and $j = l$ in (3.117), we have

$$[M^{(-)}]^{l-m} Y_l^l(\theta,\phi) = \mathrm{e}^{-i(l-m)\phi} \left(\sqrt{1-\xi^2}\right)^{-m} \frac{\partial^{l-m}}{\partial \xi^{l-m}} \left[\left(\sqrt{1-\xi^2}\right)^l Y_l^l(\theta,\phi)\right]. \tag{3.141}$$

Further replacing $[M^{(-)}]^{l-m} Y_l^l(\theta,\phi)$ in (3.140) with that of (3.141), we get

$$Y_l^m(\theta,\phi) = \sqrt{\frac{(l+m)!}{(2l)!(l-m)!}} \mathrm{e}^{-i(l-m)\phi} \left(\sqrt{1-\xi^2}\right)^{-m} \frac{\partial^{l-m}}{\partial \xi^{l-m}} \left[\left(\sqrt{1-\xi^2}\right)^l Y_l^l(\theta,\phi)\right]. \tag{3.142}$$

Finally, replacing $Y_l^l(\theta,\phi)$ in (3.142) with that of (3.137) and converting $\theta$ to $\xi$, we arrive at the following equation:

$$Y_l^m(\theta,\phi) = \frac{\mathrm{e}^{i\chi}}{2^l l!} \sqrt{\frac{(2l+1)(l+m)!}{4\pi(l-m)!}} \mathrm{e}^{im\phi} (1-\xi^2)^{-m/2} \frac{\partial^{l-m}}{\partial \xi^{l-m}} [(1-\xi^2)^l]. \tag{3.143}$$

Now, let us decide $\mathrm{e}^{i\chi}$. Putting $m = 0$ in (3.143), we have

$$Y_l^0(\theta,\phi)=\frac{\mathrm{e}^{i\chi}(-1)^l}{2^l l!(-1)^l}\sqrt{\frac{(2l+1)}{4\pi}}\frac{\partial^l}{\partial\xi^l}[(1-\xi^2)^l], \tag{3.144}$$

where we put $(-1)^l$ on both the numerator and denominator. In RHS of (3.144),

$$\frac{(-1)^l}{2^l l!}\frac{\partial^l}{\partial\xi^l}[(1-\xi^2)^l]\equiv P_l(\xi). \tag{3.145}$$

Equation (3.145) is well-known as Rodrigues formula of Legendre polynomials. We mention characteristics of Legendre polynomials in the next section. Thus,

$$Y_l^0(\theta,\phi)=\frac{\mathrm{e}^{i\chi}}{(-1)^l}\sqrt{\frac{(2l+1)}{4\pi}}P_l(\xi). \tag{3.146}$$

According to the custom [2], we require $Y_l^0(0,\phi)$ to be positive. Noting that $\theta=0$ corresponds to $\xi=1$, we have

$$Y_l^0(0,\phi)=\frac{\mathrm{e}^{i\chi}}{(-1)^l}\sqrt{\frac{(2l+1)}{4\pi}}P_l(1)=\frac{\mathrm{e}^{i\chi}}{(-1)^l}\sqrt{\frac{(2l+1)}{4\pi}}, \tag{3.147}$$

where we used $P_l(1)=1$. For this important relation, see Sect. 3.6.1. Also noting that $\left|\frac{\mathrm{e}^{i\chi}}{(-1)^l}\right|=1$, we must have

$$\frac{\mathrm{e}^{i\chi}}{(-1)^l}=1 \text{ or } \mathrm{e}^{i\chi}=(-1)^l \tag{3.148}$$

so that $Y_l^0(0,\phi)$ can be positive. Thus, (3.143) is rewritten as

$$Y_l^m(\theta,\phi)=\frac{(-1)^l}{2^l l!}\sqrt{\frac{(2l+1)(l+m)!}{4\pi(l-m)!}}\mathrm{e}^{im\phi}(1-\xi^2)^{-m/2}\frac{\partial^{l-m}}{\partial\xi^{l-m}}[(1-\xi^2)^l] \tag{3.149}$$

In Sect. 3.3, we mentioned that $|\psi_0\rangle$ in (3.43) is a constant. In fact, putting $l=m=0$ in (3.149), we have

$$Y_0^0(\theta,\phi)=\sqrt{1/4\pi}. \tag{3.150}$$

Thus, as a simultaneous eigenstate of all $L_x$, $L_y$, $L_z$, and $\boldsymbol{L}^2$ corresponding to $l=0$ and $m=0$, we have

$$|\psi_0\rangle\equiv Y_0^0(\theta,\phi).$$

The normalized functions $Y_l^m(\theta,\phi)$ described as (3.149) define simultaneous eigenfunctions of $\boldsymbol{L}^2$ (or $\boldsymbol{M}^2$) and $L_z$ (or $M_z$). Those functions are called spherical surface harmonics and frequently appear in various fields of mathematical physics.

As in the case of Sect. 2.3, matrix representation enables us to intuitively grasp the relationship between angular momentum operators and their eigenfunctions (or eigenvectors). Rewriting the relations of (3.101) so that they can meet the present purpose, we have

$$\begin{aligned} M^{(-)}|l,m\rangle &= \sqrt{(l-m+1)(l+m)}|l,m-1\rangle, \\ M^{(+)}|l,m\rangle &= \sqrt{(l-m)(l+m+1)}|l,m+1\rangle, \end{aligned} \tag{3.151}$$

where we used $l$ instead of $\zeta$ to designate the eigenstate.

Now, we know that $m$ takes $(2l+1)$ different values that correspond to each $l$. This implies that the operators can be expressed with $(2l+1, 2l+1)$ matrices. As implied in (3.151), $M^{(-)}$ takes the following form:

$$M^{(-)} = \begin{pmatrix} 0 & \sqrt{2l\cdot 1} & & & & & & & \\ & 0 & \sqrt{(2l-1)\cdot 2} & & & & & & \\ & & 0 & \ddots & & & & & \\ & & & \ddots & & & & & \\ & & & & 0 & \sqrt{(2l-k+1)\cdot k} & & & \\ & & & & & 0 & \ddots & & \\ & & & & & & \ddots & & \\ & & & & & & & 0 & \sqrt{1\cdot 2l} \\ & & & & & & & & 0 \end{pmatrix}, \tag{3.152}$$

where diagonal elements are zero and a $(k, k+1)$ element is $\sqrt{(2l-k+1)\cdot k}$. That is, nonzero elements are positioned just *above* the zero diagonal elements. Correspondingly, we have

$$M^{(+)} = \begin{pmatrix} 0 & & & & & & & & \\ \sqrt{2l\cdot 1} & 0 & & & & & & & \\ & \sqrt{(2l-1)\cdot 2} & 0 & & & & & & \\ & & \ddots & & & & & & \\ & & & 0 & & & & & \\ & & & \sqrt{(2l-k+1)\cdot k} & 0 & & & & \\ & & & & \ddots & \ddots & & & \\ & & & & & & 0 & & \\ & & & & & & \ddots & 0 & \\ & & & & & & & \sqrt{1\cdot 2l} & 0 \end{pmatrix}, \tag{3.153}$$

where again diagonal elements are zero and a $(k+1, k)$ element is $\sqrt{(2l-k+1)\cdot k}$. In this case, nonzero elements are positioned just *below* the zero

diagonal elements. Notice also that $M^{(-)}$ and $M^{(+)}$ are adjoint to each other and that these notations correspond to (2.65) and (2.66).

Basis functions $Y_l^m(\theta, \phi)$ can be represented by a column vector, as in the case of Sect. 2.3. These are denoted as follows:

$$|l,-l\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, |l,-l+1\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \ldots, |l,l-1\rangle = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \end{pmatrix}, |l,l\rangle = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 1 \end{pmatrix}, \tag{3.154}$$

where the first number $l$ in $|l,-l\rangle$, $|l,-l+1\rangle$, etc., denotes the quantum number associated with $\lambda = l(l+1)$ of (3.124) and is kept constant; the latter number denotes $m$. Note from (3.154) that the column vector whose $k$th row is 1 corresponds to $m$ such that

$$m = -l + k - 1. \tag{3.155}$$

For instance, if $k = 1$, $m = -l$; if $k = 2l+1$, $m = l$, etc.

The operator $M^{(-)}$ converts the column vector whose $(k+1)$th row is 1 to that whose $k$th row is 1. The former column vector corresponds to $|l, m+1\rangle$ and the latter corresponding to $|l, m\rangle$. Therefore, using (3.152), we get the following representation:

$$M^{(-)}|l,m+1\rangle = \sqrt{(2l-k+1)\cdot k}|l,m\rangle = \sqrt{(l-m)(l+m+1)}|l,m\rangle, \tag{3.156}$$

where the second equality is obtained by replacing $k$ with that of (3.155), i.e., $k = l+m+1$. Changing $m$ to $(m-1)$, we get the first equation of (3.151). Similarly, we obtain the second equation of (3.151) as well. That is, we have

$$M^{(+)}|l,m\rangle = \sqrt{(2l-k+1)\cdot k}|l,m+1\rangle = \sqrt{(l-m)(l+m+1)}|l,m+1\rangle. \tag{3.157}$$

From (3.32), we have

$$\boldsymbol{M}^2 = M^{(+)}M^{(-)} + M_z^2 - M_z.$$

In the above, $M^{(+)}M^{(-)}$ and $M_z$ are diagonal matrices and, hence, $M_z^2$ and $\boldsymbol{M}^2$ are diagonal matrices as well such that

$$M_z = \begin{pmatrix} -l & & & & & & \\ & -l+1 & & & & & \\ & & -l+2 & & & & \\ & & & \ddots & & & \\ & & & & k-l & & \\ & & & & & \ddots & \\ & & & & & & l \end{pmatrix},$$

$$M^{(+)}M^{(-)} = \begin{pmatrix} 0 & & & & & & \\ & 2l\cdot 1 & & & & & \\ & & (2l-1)\cdot 2 & & & & \\ & & & \ddots & (2l-k+1)\cdot k & & \\ & & & & & \ddots & \\ & & & & & & 1\cdot 2l \end{pmatrix}, \tag{3.158}$$

where $k-l$ and $(2l-k+1)\cdot k$ represent $(k+1,k+1)$ elements of $M_z$ and $M^{(+)}M^{(-)}$, respectively. Therefore, $(k+1,k+1)$ element of $\boldsymbol{M}^2$ is calculated as

$$(2l-k+1)\cdot k + (k-l)^2 - (k-l) = l(l+1).$$

As expected, $\boldsymbol{M}^2$ takes a constant value $l(l+1)$. A matrix representation is shown in (3.159) such that

$$\boldsymbol{M}^2 = \begin{pmatrix} l(l+1) & & & & & & \\ & l(l+1) & & & & & \\ & & \ddots & & & & \\ & & & l(l+1) & & & \\ & & & & \ddots & & \\ & & & & & l(l+1) & \\ & & & & & & l(l+1) \end{pmatrix}. \tag{3.159}$$

These expressions are useful to understand how the vectors of (3.154) constitute simultaneous eigenstates of $\boldsymbol{M}^2$ and $M_z$. In this situation, the matrix representation is said to diagonalize both $\boldsymbol{M}^2$ and $M_z$. In other words, the quantum states represented by (3.154) are simultaneous eigenstates of $\boldsymbol{M}^2$ and $M_z$.

The matrices (3.152) and (3.153) that represent $M^{(-)}$ and $M^{(+)}$, respectively, are said to be ladder operators or raising and lowering operators, because operating column vectors those operators convert $|m\rangle$ to $|m \mp 1\rangle$ as mentioned above. The operators $M^{(-)}$ and $M^{(+)}$ correspond to $a$ and $a^\dagger$ given in (2.65) and (2.66), respectively. All these operators are characterized by that the corresponding matrices have diagonal elements of zero and that nonvanishing elements are only positioned on "right above" or "right below" relative to the diagonal elements. These matrices are a kind of triangle matrices, and all their diagonal elements are zero. The matrices are characteristic of *nilpotent* matrices. That is, if a suitable power of a matrix is zero as a matrix, such a matrix is said to be a nilpotent matrix (see Part III). In the present case, $(2l+1)$th power of $M^{(-)}$ and $M^{(+)}$ becomes zero as a matrix.

The operator $M^{(-)}$ and $M^{(+)}$ can be described by the following shorthand representations:

$$[M^{(-)}]_{kj} = \sqrt{(2l-k+1)\cdot k}\,\delta_{k+1,j} \ (1 \le k \le 2l). \tag{3.160}$$

If $l = 0$, $M_z = M^{(+)}M^{(-)} = \boldsymbol{M}^2 = 0$. This case corresponds to $Y_0^0(\theta,\phi) = \sqrt{1/4\pi}$, and we do not need the matrix representation. Defining

$$a_k \equiv \sqrt{(2l-k+1)\cdot k},$$

we have for instance

$$\begin{aligned}\{[M^{(-)}]^2\}_{kj} &= \sum\nolimits_p a_k\delta_{k+1,p}a_p\delta_{p+1,j} = a_k a_{k+1}\delta_{k+2,j} \\ &= \sqrt{(2l-k+1)\cdot k}\sqrt{[2l-(k+1)+1]\cdot(k+1)}\,\delta_{k+2,j},\end{aligned} \tag{3.161}$$

where the summation is nonvanishing only if $p = k+1$. The factor $\delta_{k+2,j}$ implies that the elements are shifted by one toward upper right by being squared. Similarly, we have

$$[M^{(+)}]_{kj} = a_{k-1}\delta_{k,j+1} \ (1 \le k \le 2l+1). \tag{3.162}$$

In (3.158), $M^{(+)}M^{(-)}$ is represented as follows:

$$\begin{aligned}[M^{(+)}M^{(-)}]_{kj} &= \sum_p a_{k-1}\delta_{k,p+1}a_p\delta_{p+1,j} = a_{k-1}a_{j-1}\delta_{k,j} = (a_{k-1})^2\delta_{k,j} \\ &= [2l-(k-1)+1]\cdot(k-1)\delta_{k,j} = (2l-k+2)(k-1)\delta_{k,j}.\end{aligned} \tag{3.163}$$

Notice that although $a_0$ is not defined, $\delta_{1,j+1} = 0$ for any $j$, and so this causes no inconvenience. Hence, $[M^{(+)}M^{(-)}]_{kj}$ of (3.163) is well-defined with $1 \le k \le 2l+1$.

Important properties of angular momentum operators examined above are based upon the fact that those operators are ladder operators and represented by nilpotent matrices. These characteristics will further be studied in Part III.

## 3.6 Orbital Angular Momentum: Analytic Approach

In this section, our central task is to solve the associated Legendre differential equation expressed by (3.127) by an analytical method. Putting $m = 0$ in (3.127), we have

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[(1-x^2)\frac{\mathrm{d}P_l^0(x)}{\mathrm{d}x}\right] + l(l+1)P_l^0(x) = 0, \tag{3.164}$$

where we use a variable $x$ instead of $\xi$. Equation (3.164) is called Legendre differential equation, and its characteristics and solutions have been widely investigated. Hence, we put

$$P_l^0(x) \equiv P_l(x), \tag{3.165}$$

where $P_l(x)$ is said to be Legendre polynomials. We first start with Legendre differential equation and Legendre polynomials.

### *3.6.1 Spherical Surface Harmonics and Associated Legendre Differential Equation*

Let us think of a following identity according to Byron and Fuller [4]:

$$(1-x^2)\frac{\mathrm{d}}{\mathrm{d}x}(1-x^2)^l = -2lx(1-x^2)^l, \tag{3.166}$$

where $l$ is a positive integer. We differentiate both sides of (3.166) $(l+1)$ times. Here, we use the Leibniz rule about differentiation of a product function that is described by

$$\mathrm{d}^n(uv) = \sum_{m=0}^{n}\frac{n!}{m!(n-m)!}\mathrm{d}^m u \mathrm{d}^{n-m}v, \tag{3.167}$$

where

$$\mathrm{d}^m u/\mathrm{d}x^m \equiv \mathrm{d}^m u.$$

The above shorthand notation is due to Byron and Fuller [4]. We use this notation for simplicity from place to place.

Noting that the third-order and higher differentiations of $(1-x^2)$ vanish in LHS of (3.166), we have

$$\begin{aligned}\text{LHS} &= \mathrm{d}^{l+1}[(1-x^2)\mathrm{d}(1-x^2)^l] \\ &= (1-x^2)\mathrm{d}^{l+2}(1-x^2)^l - 2(l+1)x\mathrm{d}^{l+1}(1-x^2)^l \\ &\quad - l(l+1)\mathrm{d}^l(1-x^2)^l.\end{aligned}$$

Also noting that the second-order and higher differentiations of $2lx$ vanish in LHS of (3.166), we have

$$\begin{aligned}\text{RHS} &= -\mathrm{d}^{l+1}[2lx(1-x^2)^l] \\ &= -2lx\mathrm{d}^{l+1}(1-x^2)^l - 2l(l+1)\mathrm{d}^l(1-x^2)^l.\end{aligned}$$

Therefore,

$$\begin{aligned}&\text{LHS} - \text{RHS} \\ &= (1-x^2)\mathrm{d}^{l+2}(1-x^2)^l - 2x\mathrm{d}^{l+1}(1-x^2)^l + l(l+1)\mathrm{d}^l(1-x^2)^l = 0.\end{aligned}$$

We define $P_l(x)$ as

$$P_l(x) \equiv \frac{(-1)^l}{2^l l!}\frac{\mathrm{d}^l}{\mathrm{d}x^l}[(1-x^2)^l], \tag{3.168}$$

where a constant $\frac{(-1)^l}{2^l l!}$ is multiplied according to the custom so that we can explicitly represent Rodrigues formula of Legendre polynomials. Thus, from (3.164) $P_l(x)$ defined above satisfies Legendre differential equation. Rewriting it, we get

$$(1-x^2)\frac{\mathrm{d}^2 P_l(x)}{\mathrm{d}x^2} - 2x\frac{\mathrm{d}P_l(x)}{\mathrm{d}x} + l(l+1)P_l(x) = 0. \tag{3.169}$$

Or equivalently, we have

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[(1-x^2)\frac{\mathrm{d}P_l(x)}{\mathrm{d}x}\right] + l(l+1)P_l(x) = 0. \tag{3.170}$$

Returning to (3.127) and using $x$ as a variable, we rewrite (3.127) as

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[(1-x^2)\frac{\mathrm{d}P_l^m(x)}{\mathrm{d}x}\right] + \left[l(l+1) - \frac{m^2}{1-x^2}\right]P_l^m(x) = 0, \tag{3.171}$$

where $l$ is a nonnegative integer and $m$ is an integer that takes following values:

$$m = l, l-1, l-2, \ldots 1, 0, -1, \cdots -l+1, -l.$$

Deferential equations expressed as

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[p(x)\frac{\mathrm{d}y(x)}{\mathrm{d}x}\right] + c(x)y(x) = 0$$

are of particular importance. We will come back to this point in Sect. 8.3.

Since $m$ can be either positive or negative, from (3.171) we notice that $P_l^m(x)$ and $P_l^{-m}(x)$ must satisfy the same differential equation (3.171). This implies that $P_l^m(x)$ and $P_l^{-m}(x)$ are connected, i.e., linearly dependent. First, let us assume that $m \geq 0$. In the case of $m < 0$, we will examine it later soon.

According to Dennery and Krzywicki [5], we assume

$$P_l^m(x) = \kappa(1-x^2)^{m/2}C(x), \tag{3.172}$$

where $\kappa$ is a constant. Inserting (3.172) into (3.171) and rearranging the terms, we obtain

$$(1-x^2)\frac{\mathrm{d}^2C}{\mathrm{d}x^2} - 2(m+1)x\frac{\mathrm{d}C}{\mathrm{d}x} + (l-m)(l+m+1)C = 0 (0 \leq m \leq l). \tag{3.173}$$

Recall once again that if $m = 0$, the associated Legendre differential equation given by (3.127) and (3.171) is exactly identical to Legendre differential equation of (3.170). Differentiating (3.170) $m$ times, we get

$$(1-x^2)\frac{\mathrm{d}^2}{\mathrm{d}x^2}\left(\frac{\mathrm{d}^m P_l}{\mathrm{d}x^m}\right) - 2(m+1)x\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\mathrm{d}^m P_l}{\mathrm{d}x^m}\right) + (l-m)(l+m+1)\frac{\mathrm{d}^m P_l}{\mathrm{d}x^m} = 0, \tag{3.174}$$

where we used the Leibniz rule about differentiation of (3.167). Comparing (3.173) and (3.174), we find that

$$C(x) = \kappa'\frac{\mathrm{d}^m P_l}{\mathrm{d}x^m},$$

where $\kappa'$ is a constant. Inserting this relation into (3.172) and setting $\kappa\kappa' = 1$, we get

$$P_l^m(x) = (1-x^2)^{m/2}\frac{\mathrm{d}^m P_l(x)}{\mathrm{d}x^m} \quad (0 \leq m \leq l). \tag{3.175}$$

Using Rodrigues formula of (3.168), we have

$$P_l^m(x) \equiv \frac{(-1)^l}{2^l l!}(1-x^2)^{m/2}\frac{\mathrm{d}^{l+m}}{\mathrm{d}x^{l+m}}[(1-x^2)^l]. \tag{3.176}$$

Equation (3.175) defines the associated Legendre functions. Note, however, that the function form differs from literature to literature [2, 5, 6].

Amongst classical orthogonal polynomials, Gegenbauer polynomials $C_n^\lambda(x)$ often appear in the literature. The relevant differential equation is defined by

$$(1-x^2)\frac{\mathrm{d}^2}{\mathrm{d}x^2}C_n^\lambda(x) - (2\lambda+1)x\frac{\mathrm{d}}{\mathrm{d}x}C_n^\lambda(x) + n(n+2\lambda)C_n^\lambda(x) = 0 \left(\lambda > -\frac{1}{2}\right). \tag{3.177}$$

Setting $n = l - m$ and $\lambda = m + \frac{1}{2}$ in (3.177) [5], we have

$$\begin{aligned}(1-x^2)\frac{\mathrm{d}^2}{\mathrm{d}x^2}C_{l-m}^{m+\frac{1}{2}}(x) - 2(m+1)x\frac{\mathrm{d}}{\mathrm{d}x}C_{l-m}^{m+\frac{1}{2}}(x) \\ + (l-m)(l+m+1)C_{l-m}^{m+\frac{1}{2}}(x) = 0.\end{aligned} \tag{3.178}$$

Once again comparing (3.174) and (3.178), we obtain

$$\frac{\mathrm{d}^m P_l(x)}{\mathrm{d}x^m} = \text{constant} \cdot C_{l-m}^{m+\frac{1}{2}}(x)(0 \le m \le l). \tag{3.179}$$

Next, let us determine the constant appearing in (3.179). To this end, we consider a following generating function of the polynomials $C_n^\lambda(x)$ defined by [7]

$$(1-2tx+t^2)^{-\lambda} \equiv \sum_{n=0}^{\infty} C_n^\lambda(x)t^n \left(\lambda > -\frac{1}{2}\right). \tag{3.180}$$

To calculate (3.180), let us think of a following expression for $x$ and $\lambda(> -\frac{1}{2})$:

$$(1+x)^{-\lambda} = \sum_{m=0}^{\infty}\binom{-\lambda}{m}x^m, \tag{3.181}$$

where $\lambda$ is an arbitrary real number and we define $\binom{-\lambda}{m}$ as

$$\binom{-\lambda}{m} \equiv -\lambda(-\lambda-1)(-\lambda-2)\ldots(-\lambda-m+1)/m! \quad \text{and} \quad \binom{-\lambda}{0} \equiv 1. \tag{3.182}$$

Notice that (3.182) is an extension of binomial theorem. Putting $-\lambda = n$, we have

$$\binom{n}{m} = \frac{n!}{(n-m)!m!} \quad \text{and} \quad \binom{n}{0} = 1.$$

We rewrite (3.181) using gamma functions $\Gamma(z)$ such that

$$(1+x)^{-\lambda} = \sum_{m=0}^{\infty} \frac{\Gamma(-\lambda+1)}{m!\Gamma(-\lambda-m+1)} x^m, \tag{3.183}$$

where $\Gamma(z)$ is defined by integral representation as

$$\Gamma(z) = \int_0^{\infty} \mathrm{e}^{-t} t^{z-1} \mathrm{d}t (\mathrm{Re}\ z > 0). \tag{3.184}$$

Changing variables such that $t = u^2$, we have

$$\Gamma(z) = 2\int_0^{\infty} \mathrm{e}^{-u^2} u^{2z-1} \mathrm{d}u (\mathrm{Re}\ z > 0). \tag{3.185}$$

Note that the above expression is associated with the following fundamental feature of the gamma functions:

$$\Gamma(z+1) = z\Gamma(z), \tag{3.186}$$

where $z$ is any complex number.

Replacing $x$ with $-t(2x-t)$ and rewriting (3.183), we have

$$(1-2tx+t^2)^{-\lambda} = \sum_{m=0}^{\infty} \frac{\Gamma(-\lambda+1)}{m!\Gamma(-\lambda-m+1)} (-t)^m (2x-t)^m. \tag{3.187}$$

Assuming that $x$ is a real number belonging to an interval $[-1, 1]$, (3.187) holds with $t$ satisfying $|t| < 1$ [8]. The discussion is as follows: When $x$ satisfies the above condition, solving $1 - 2tx + t^2 = 0$ we get solution $t_{\pm}$ such that

$$t_{\pm} = x \pm i\sqrt{1-x^2}.$$

Defining $r$ as

$$r \equiv \min\{|t_+|, |t_-|\},$$

$(1-2tx+t^2)^{-\lambda}$, regarded as a function of $t$, is analytic in the disk $|t|<r$. But, we have

$$|t_{\pm}| = 1.$$

Thus, $(1-2tx+t^2)^{-\lambda}$ is analytic within the disk $|t|<1$ and, hence, it can be expanded in a Taylor series.

Continuing the calculation of (3.187), we have

$$\begin{aligned}
&(1-2tx+t^2)^{-\lambda} \\
&= \sum_{m=0}^{\infty} \frac{\Gamma(-\lambda+1)}{m!\Gamma(-\lambda-m+1)} (-1)^m t^m \left[ \sum_{k=0}^{m} \frac{m!}{k!(m-k)!} 2^{m-k} x^{m-k} (-1)^k t^k \right] \\
&= \sum_{m=0}^{\infty} \sum_{k=0}^{m} \frac{(-1)^{m+k}}{k!(m-k)!} \frac{\Gamma(-\lambda+1)}{\Gamma(-\lambda-m+1)} 2^{m-k} x^{m-k} t^{m+k} \\
&= \sum_{m=0}^{\infty} \sum_{k=0}^{m} \frac{(-1)^{m+k}}{k!(m-k)!} \frac{(-1)^m \Gamma(\lambda+m)}{\Gamma(\lambda)} 2^{m-k} x^{m-k} t^{m+k},
\end{aligned} \tag{3.188}$$

where the last equality results from that we rewrote gamma functions using (3.186). Replacing $(m+k)$ with $n$, we get

$$(1-2tx+t^2)^{-\lambda} = \sum_{n=0}^{\infty} \sum_{k=0}^{[n/2]} \frac{(-1)^k 2^{n-2k}}{k!(n-2k)!} \frac{\Gamma(\lambda+n-k)}{\Gamma(\lambda)} x^{n-2k} t^n, \tag{3.189}$$

where $[n/2]$ represents an integer that does not exceed $n/2$. This expression comes from a requirement that an order of $x$ must satisfy the following condition:

$$n-2k \geq 0 \text{ or } k \leq n/2. \tag{3.190}$$

That is, if $n$ is even, the maximum of $k = n/2$. If $n$ is odd, the maximum of $k = (n-1)/2$. Comparing (3.180) and (3.189), we get [8]

$$C_n^{\lambda}(x) = \sum_{k=0}^{[n/2]} \frac{(-1)^k 2^{n-2k}}{k!(n-2k)!} \frac{\Gamma(\lambda+n-k)}{\Gamma(\lambda)} x^{n-2k}. \tag{3.191}$$

Comparing (3.164) and (3.177) and putting $\lambda = 1/2$, we immediately find that the two differential equations are identical [7]. That is,

$$C_n^{1/2}(x) = P_n(x). \tag{3.192}$$

Hence, we further have

$$P_n(x) = \sum_{k=0}^{[n/2]} \frac{(-1)^k 2^{n-2k}}{k!(n-2k)!} \frac{\Gamma(\frac{1}{2}+n-k)}{\Gamma(\frac{1}{2})} x^{n-2k}. \tag{3.193}$$

Using (3.186) once again, we get [8]

$$P_n(x) = \sum_{k=0}^{[n/2]} \frac{(-1)^k (2n-2k)!}{2^n k!(n-k)!(n-2k)!} x^{n-2k}. \tag{3.194}$$

It is convenient to make a formula about a gamma function. In (3.193), $n-k>0$, and so let us think of $\Gamma(\frac{1}{2}+m)$ ($m$ : positive integer). Using (3.186), we have

$$\begin{aligned}
\Gamma\left(\frac{1}{2}+m\right) &= \left(m-\frac{1}{2}\right)\Gamma\left(m-\frac{1}{2}\right) = \left(m-\frac{1}{2}\right)\left(m-\frac{3}{2}\right)\Gamma\left(m-\frac{3}{2}\right) = \cdots \\
&= \left(m-\frac{1}{2}\right)\left(m-\frac{3}{2}\right)\cdots\left(\frac{1}{2}\right)\Gamma\left(\frac{1}{2}\right) = 2^{-m}(2m-1)(2m-3)\cdots 3\cdot 1\cdot\Gamma\left(\frac{1}{2}\right) \\
&= 2^{-m}\frac{(2m-1)!}{2^{m-1}(m-1)!}\Gamma\left(\frac{1}{2}\right) = 2^{-2m}\frac{(2m)!}{m!}\Gamma\left(\frac{1}{2}\right).
\end{aligned} \tag{3.195}$$

Notice that (3.195) still holds even if $m=0$. Inserting $n-k$ into $m$ of (3.195), we get

$$\Gamma\left(\frac{1}{2}+n-k\right) = 2^{-2(n-k)}\frac{(2n-2k)!}{(n-k)!}\Gamma\left(\frac{1}{2}\right). \tag{3.196}$$

Replacing $\Gamma\left(\frac{1}{2}+n-k\right)$ of (3.193) with RHS of the above equation, (3.194) will follow. A gamma function $\Gamma\left(\frac{1}{2}\right)$ often appears in mathematical physics. According to (3.185), we have

$$\Gamma\left(\frac{1}{2}\right) = 2\int_0^\infty \mathrm{e}^{-u^2}\mathrm{d}u = \sqrt{\pi}.$$

For the derivation of the above definite integral, see (2.86) of Sect. 2.4. From (3.184), we also have

$$\Gamma(1) = 1.$$

In relation of the discussion of Sect. 3.5, let us derive an important formula about Legendre polynomials. From (3.180) to (3.192), we get

$$(1-2tx+t^2)^{-1/2} \equiv \sum_{n=0}^{\infty} P_n(x)t^n. \tag{3.197}$$

Assuming $|t|<1$, when we put $x=1$ in (3.197), we have

$$(1-2tx+t^2)^{-\frac{1}{2}} = \frac{1}{1-t} = \sum_{n=0}^{\infty} t^n = \sum_{n=0}^{\infty} P_n(1)t^n. \tag{3.198}$$

Comparing individual coefficients of $t^n$ in (3.198), we get

$$P_n(1) = 1.$$

See the related parts of Sect. 3.5.

Now, we are in the position to determine the constant in (3.179). Differentiating (3.194) $m$ times, we have

$$\begin{aligned} \mathrm{d}^m P_l(x)/\mathrm{d}x^m &= \sum_{k=0}^{[(l-m)/2]} \frac{(-1)^k(2l-2k)!(l-2k)(l-2k-1)\ldots(l-2k-m+1)}{2^l k!(l-k)!(l-2k)!} x^{l-2k-m} \\ &= \sum_{k=0}^{[(l-m)/2]} \frac{(-1)^k(2l-2k)!}{2^l k!(l-k)!(l-2k-m)!} x^{l-2k-m}. \end{aligned} \tag{3.199}$$

Meanwhile, we have

$$C_{l-m}^{m+\frac{1}{2}}(x) = \sum_{k=0}^{[(l-m)/2]} \frac{(-1)^k 2^{l-2k-m}}{k!(l-2k-m)!} \frac{\Gamma(l+\frac{1}{2}-k)}{\Gamma(m+\frac{1}{2})} x^{l-2k-m}. \tag{3.200}$$

Using (3.195) and (3.196), we have

$$\frac{\Gamma(l+\frac{1}{2}-k)}{\Gamma(m+\frac{1}{2})} = 2^{-2(l-k-m)} \frac{(2l-2k)!}{(l-k)!} \frac{m!}{(2m)!}.$$

Therefore, we get

$$C_{l-m}^{m+\frac{1}{2}}(x) = \sum_{k=0}^{[(l-m)/2]} \frac{(-1)^k(2l-2k)!}{2^l k!(l-k)!(l-2k-m)!} \frac{2^m \Gamma(m+1)}{\Gamma(2m+1)} x^{l-2k-m}, \tag{3.201}$$

where we used $m! = \Gamma(m+1)$ and $(2m)! = \Gamma(2m+1)$. Comparing (3.199) and (3.201), we get

$$\frac{\mathrm{d}^m P_l(x)}{\mathrm{d}x^m} = \frac{\Gamma(2m+1)}{2^m \Gamma(m+1)} C_{l-m}^{m+\frac{1}{2}}(x). \tag{3.202}$$

Thus, we find that the constant appearing in (3.179) is $\frac{\Gamma(2m+1)}{2^m\Gamma(m+1)}$. Putting $m = 0$ in (3.202), we have $P_l(x) = C_l^{1/2}(x)$. Therefore, (3.192) is certainly recovered. This gives an easy checkup to (3.202).

Meanwhile, Rodrigues formula of Gegenbauer polynomials [5] is given by

$$C_n^{\lambda}(x) = \frac{(-1)^n \Gamma(n+2\lambda)\Gamma(\lambda+\frac{1}{2})}{2^n n!\Gamma(n+\lambda+\frac{1}{2})\Gamma(2\lambda)} (1-x^2)^{-\lambda+\frac{1}{2}} \frac{\mathrm{d}^n}{\mathrm{d}x^n}\left[(1-x^2)^{n+\lambda-\frac{1}{2}}\right]. \tag{3.203}$$

Hence, we have

$$C_{l-m}^{m+\frac{1}{2}}(x) = \frac{(-1)^{l-m}\Gamma(l+m+1)\Gamma(m+1)}{2^{l-m}(l-m)!\Gamma(l+1)\Gamma(2m+1)} (1-x^2)^{-m} \frac{\mathrm{d}^{l-m}}{\mathrm{d}x^{l-m}}[(1-x^2)^l]. \tag{3.204}$$

Inserting (3.204) into (3.202), we have

$$\begin{aligned} \frac{\mathrm{d}^m P_l(x)}{\mathrm{d}x^m} &= \frac{(-1)^{l-m}\Gamma(l+m+1)}{2^l(l-m)!\Gamma(l+1)} (1-x^2)^{-m} \frac{\mathrm{d}^{l-m}}{\mathrm{d}x^{l-m}}[(1-x^2)^l] \\ &= \frac{(-1)^{l-m}(l+m)!}{2^l l!(l-m)!} (1-x^2)^{-m} \frac{\mathrm{d}^{l-m}}{\mathrm{d}x^{l-m}}[(1-x^2)^l]. \end{aligned} \tag{3.205}$$

Further inserting this into (3.175), we finally get

$$P_l^m(x) = \frac{(-1)^{l-m}(l+m)!}{2^l l!(l-m)!} (1-x^2)^{-m/2} \frac{\mathrm{d}^{l-m}}{\mathrm{d}x^{l-m}}[(1-x^2)^l]. \tag{3.206}$$

When $m = 0$, we have

$$P_l^0(x) = \frac{(-1)^l}{2^l l!} \frac{\mathrm{d}^l}{\mathrm{d}x^l}[(1-x^2)^l] = P_l(x). \tag{3.207}$$

Thus, we recover the functional form of Legendre polynomials. The expression (3.206) is also meaningful for negative $m$, provided $|m| \le l$, and permits an extension of the definition of $P_l^m(x)$ given by (3.175) to negative numbers of $m$ [5].

Changing $m$ to $-m$ in (3.206), we have

$$P_l^{-m}(x) = \frac{(-1)^{l+m}(l-m)!}{2^l l!(l+m)!} (1-x^2)^{m/2} \frac{\mathrm{d}^{l+m}}{\mathrm{d}x^{l+m}}[(1-x^2)^l]. \tag{3.208}$$

Meanwhile, from (3.168) to (3.175),

$$P_l^m(x) = \frac{(-1)^l}{2^l l!}(1-x^2)^{m/2}\frac{\mathrm{d}^{l+m}}{\mathrm{d}x^{l+m}}[(1-x^2)^l] \ (0 \le m \le l). \tag{3.209}$$

Comparing (3.208) and (3.209), we get

$$P_l^{-m}(x) = \frac{(-1)^m(l-m)!}{(l+m)!}P_l^m(x) \ (-l \le m \le l). \tag{3.210}$$

Thus, as expected earlier, $P_l^m(x)$ and $P_l^{-m}(x)$ are linearly dependent.
Now, we return back to (3.149). From (3.206) we have

$$(1-x^2)^{-\frac{m}{2}}\frac{\mathrm{d}^{l-m}}{\mathrm{d}x^{l-m}}[(1-x^2)^l] = \frac{(-1)^{l-m}2^l l!(l-m)!}{(l+m)!}P_l^m(x). \tag{3.211}$$

Inserting (3.211) into (3.149) and changing the variable $x$ to $\xi$, we have

$$Y_l^m(\theta,\phi) = (-1)^m\sqrt{\frac{(2l+1)(l-m)!}{4\pi(l+m)!}}P_l^m(\xi)\mathrm{e}^{im\phi}(\xi = \cos\theta; 0 \le \theta \le \pi). \tag{3.212}$$

The coefficient $(-1)^m$ appearing (3.212) is well-known as Condon–Shortley phase [7]. Another important expression obtained from (3.210) to (3.212) is

$$Y_l^{-m}(\theta,\phi) = (-1)^m[Y_l^m(\theta,\phi)]^*. \tag{3.213}$$

Since (3.208) or (3.209) involves higher order differentiations, it would somewhat be inconvenient to find their functional forms. Here, we try to seek the convenient representation of spherical harmonics using familiar cosine and sine functions. Starting with (3.206) and applying Leibniz rule there, we have

$$\begin{aligned}
P_l^m(x) &= \frac{(-1)^{l-m}(l+m)!}{2^l l!(l-m)!}(1+x)^{-m/2}(1-x)^{-m/2} \\
&\quad \times \sum_{r=0}^{l-m}\frac{(l-m)!}{r!(l-m-r)!}\frac{\mathrm{d}^r}{\mathrm{d}x^r}[(1+x)^l]\frac{\mathrm{d}^{l-m-r}}{\mathrm{d}x^{l-m-r}}[(1-x)^l] \\
&= \frac{(-1)^{l-m}(l+m)!}{2^l l!(l-m)!}\sum_{r=0}^{l-m}\frac{(l-m)!}{r!(l-m-r)!}\frac{l!(1+x)^{l-r-\frac{m}{2}}}{(l-r)!}\frac{(-1)^{l-m-r}l!(1-x)^{m+r-\frac{m}{2}}}{(m+r)!} \\
&= \frac{l!(l+m)!}{2^l}\sum_{r=0}^{l-m}\frac{(-1)^r}{r!(l-m-r)!}\frac{(1+x)^{l-r-\frac{m}{2}}}{(l-r)!}\frac{(1-x)^{r+\frac{m}{2}}}{(m+r)!}.
\end{aligned} \tag{3.214}$$

Putting $x = \cos\theta$ in (3.214) and using a trigonometric formula, we have

$$P_l^m(x) = l!(l+m)!\sum_{r=0}^{l-m}\frac{(-1)^r}{r!(l-m-r)!}\frac{\cos^{2l-2r-m}\left(\frac{\theta}{2}\right)}{(l-r)!}\frac{\sin^{2r+m}\left(\frac{\theta}{2}\right)}{(m+r)!}. \tag{3.215}$$

Inserting this into (3.212), we get

$$Y_l^m(\theta,\phi)=l!\sqrt{\frac{(2l+1)(l+m)!(l-m)!}{4\pi}}e^{im\phi}\sum_{r=0}^{l-m}\frac{(-1)^{r+m}\cos^{2l-2r-m}\left(\frac{\theta}{2}\right)\sin^{2r+m}\left(\frac{\theta}{2}\right)}{r!(l-m-r)!(l-r)!(m+r)!}. \tag{3.216}$$

Summation domain of $r$ must be determined so that factorials of negative integers can be avoided [6]. That is,

(i) If $m\geq 0$, $0\leq r\leq l-m$; $(l-m+1)$ terms,
(ii) If $m<0$, $|m|\leq r\leq l$; $(l-|m|+1)$ terms.

For example, if we choose $l$ for $m$, putting $r=0$ in (3.216) we have

$$Y_l^l(\theta,\phi)=\sqrt{\frac{(2l+1)!}{4\pi}}e^{il\phi}\frac{(-1)^l\cos^l\left(\frac{\theta}{2}\right)\sin^l\left(\frac{\theta}{2}\right)}{l!}=\sqrt{\frac{(2l+1)!}{4\pi}}e^{il\phi}\frac{(-1)^l\sin^l\theta}{2^l l!}. \tag{3.217}$$

In particular, we have $Y_0^0(\theta,\phi)=\sqrt{\frac{1}{4\pi}}$ to recover (3.150). When $m=-l$, putting $r=l$ in (3.216) we get

$$Y_l^{-l}(\theta,\phi)=\sqrt{\frac{(2l+1)!}{4\pi}}e^{-il\phi}\frac{\cos^l\left(\frac{\theta}{2}\right)\sin^l\left(\frac{\theta}{2}\right)}{l!}=\sqrt{\frac{(2l+1)!}{4\pi}}e^{-il\phi}\frac{\sin^l\theta}{2^l l!}. \tag{3.218}$$

For instance, choosing $l=3$ and $m=\pm 3$ and using (3.217) or (3.218), we have

$$Y_3^3(\theta,\phi)=-\sqrt{\frac{35}{64\pi}}e^{i3\phi}\sin^3\theta,\ Y_3^{-3}(\theta,\phi)=\sqrt{\frac{35}{64\pi}}e^{-i3\phi}\sin^3\theta.$$

For the minus sign appearing in $Y_3^3(\theta,\phi)$ is due to the Condon–Shortley phase. For $l=3$ and $m=0$, moreover, we have

$$\begin{aligned}
Y_3^0(\theta,\phi)&=3!\sqrt{\frac{7\cdot 3!\cdot 3!}{4\pi}}\sum_{r=0}^{3}\frac{(-1)^r\cos^{6-2r}\left(\frac{\theta}{2}\right)\sin^{2r}\left(\frac{\theta}{2}\right)}{r!(3-r)!(3-r)!r!}\\
&=18\sqrt{\frac{7}{\pi}}\left[\frac{\cos^6\left(\frac{\theta}{2}\right)}{0!3!3!0!}-\frac{\cos^4\left(\frac{\theta}{2}\right)\sin^2\left(\frac{\theta}{2}\right)}{1!2!2!1!}+\frac{\cos^2\left(\frac{\theta}{2}\right)\sin^4\left(\frac{\theta}{2}\right)}{2!1!1!2!}-\frac{\sin^6\left(\frac{\theta}{2}\right)}{3!0!0!3!}\right]\\
&=18\sqrt{\frac{7}{\pi}}\left\{\frac{\cos^6\left(\frac{\theta}{2}\right)-\sin^6\left(\frac{\theta}{2}\right)}{36}+\frac{\cos^2\left(\frac{\theta}{2}\right)\sin^2\left(\frac{\theta}{2}\right)\left[\sin^2\left(\frac{\theta}{2}\right)-\cos^2\left(\frac{\theta}{2}\right)\right]}{4}\right\}\\
&=\sqrt{\frac{7}{\pi}}\left(\frac{5}{4}\cos^3\theta-\frac{3}{4}\cos\theta\right),
\end{aligned}$$

where in the last equality we used formulae of elementary algebra and trigonometric functions. At the same time, we get

$$Y_3^0(0,\phi) = \sqrt{\frac{7}{4\pi}}.$$

This is consistent with (3.147) in that $Y_3^0(0,\phi)$ is positive.

### 3.6.2 Orthogonality of Associated Legendre Functions

Orthogonality relation of functions is important. Here, we deal with it, regarding the associated Legendre functions.

Replacing $m$ with $(m-1)$ in (3.174) and using the notation introduced before, we have

$$(1-x^2)\mathrm{d}^{m+1}P_l - 2mx\mathrm{d}^m P_l + (l+m)(l-m+1)\mathrm{d}^{m-1}P_l = 0. \tag{3.219}$$

Multiplying both sides by $(1-x^2)^{m-1}$, we have

$$\begin{aligned}&(1-x^2)^m\mathrm{d}^{m+1}P_l - 2mx(1-x^2)^{m-1}\mathrm{d}^m P_l\\ &\quad + (l+m)(l-m+1)(1-x^2)^{m-1}\mathrm{d}^{m-1}P_l = 0.\end{aligned}$$

Rewriting the above equation, we get

$$\mathrm{d}[(1-x^2)^m\mathrm{d}^m P_l] = -(l+m)(l-m+1)(1-x^2)^{m-1}\mathrm{d}^{m-1}P_l. \tag{3.220}$$

Now, let us define $f(m)$ as follows:

$$f(m) \equiv \int_{-1}^{1}(1-x^2)^m(\mathrm{d}^m P_l)\mathrm{d}^m P_{l'}\mathrm{d}x \ (0 \le m \le l, l'). \tag{3.221}$$

Rewriting (3.221) as follows and integrating it by parts, we have

$$\begin{aligned}f(m) &= \int_{-1}^{1}(1-x^2)^m\mathrm{d}(\mathrm{d}^{m-1}P_l)\mathrm{d}^m P_{l'}\mathrm{d}x\\ &= [(\mathrm{d}^{m-1}P_l)(1-x^2)^m\mathrm{d}^m P_{l'}]_{-1}^{1} - \int_{-1}^{1}(\mathrm{d}^{m-1}P_l)\mathrm{d}[(1-x^2)^m\mathrm{d}^m P_{l'}]\mathrm{d}x\\ &= -\int_{-1}^{1}(\mathrm{d}^{m-1}P_l)\mathrm{d}[(1-x^2)^m\mathrm{d}^m P_{l'}]\mathrm{d}x\\ &= \int_{-1}^{1}(\mathrm{d}^{m-1}P_l)(l'+m)(l'-m+1)(1-x^2)^{m-1}\mathrm{d}^{m-1}P_{l'}\mathrm{d}x\\ &= (l'+m)(l'-m+1)f(m-1),\end{aligned} \tag{3.222}$$

where with the second equality the first term vanished and with the second last equality we used (3.220). Equation (3.222) gives a recurrence formula regarding $f(m)$. Further performing the calculation, we get

$$\begin{aligned} f(m) &= (l'+m)(l'+m-1)\cdot(l'-m+2)(l'-m+1)f(m-2) \\ &= \cdots \\ &= (l'+m)(l'+m-1)\ldots(l'+1)\cdot l'\ldots(l'-m+2)(l'-m+1)f(0) \\ &= \frac{(l'+m)!}{(l'-m)!}f(0), \end{aligned} \tag{3.223}$$

where we have

$$f(0) = \int_{-1}^{1} P_l(x)P_{l'}(x)\mathrm{d}x. \tag{3.224}$$

Note that in (3.223) a coefficient of $f(0)$ comprises $2m$ factors.

In (3.224), $P_l(x)$ and $P_{l'}(x)$ are Legendre polynomials defined in (3.165). Then, using (3.168) we have

$$f(0) = \frac{(-1)^l}{2^l l!}\frac{(-1)^{l'}}{2^{l'}l'!}\int_{-1}^{1}[\mathrm{d}^l(1-x^2)^l][\mathrm{d}^{l'}(1-x^2)^{l'}]\mathrm{d}x. \tag{3.225}$$

To evaluate (3.224), we have two cases; i.e., (i) $l \neq l'$ and (ii) $l = l'$. With the first case, assuming that $l > l'$ and taking partial integration, we have

$$\begin{aligned} I &= \int_{-1}^{1}[\mathrm{d}^l(1-x^2)^l][\mathrm{d}^{l'}(1-x^2)^{l'}]\mathrm{d}x \\ &= \left[\{\mathrm{d}^{l-1}(1-x^2)^l\}\{\mathrm{d}^{l'}(1-x^2)^{l'}\}\right]_{-1}^{1} \\ &\quad - \int_{-1}^{1}[\mathrm{d}^{l-1}(1-x^2)^l][\mathrm{d}^{l'+1}(1-x^2)^{l'}]\mathrm{d}x \end{aligned} \tag{3.226}$$

In the above, we find that the first term vanishes because it contains $(1-x^2)$ as a factor. Integrating (3.226) another $l'$ times as before, we get

$$I = (-1)^{l'+1}\int_{-1}^{1}[\mathrm{d}^{l-l'-1}(1-x^2)^l][\mathrm{d}^{2l'+1}(1-x^2)^{l'}]\mathrm{d}x. \tag{3.227}$$

In (3.227) $0 \le l - l' - 1 \le 2l$, and so $\mathrm{d}^{l-l'-1}(1-x^2)^l$ does not vanish, but $(1-x^2)^{l'}$ is an at most $2l'$-degree polynomial and, hence, $\mathrm{d}^{2l'+1}(1-x^2)^{l'}$ vanishes. Therefore,

$$f(0) = 0. \tag{3.228}$$

If $l < l'$, changing $P_l(x)$ and $P_{l'}(x)$ in (3.224), we get $f(0) = 0$ as well.
In the second case of $l = l'$, we evaluate the following integral:

$$I = \int_{-1}^{1} [\mathrm{d}^l (1-x^2)^l]^2 \mathrm{d}x. \tag{3.229}$$

Similarly integrating (3.229) by parts $l$ times, we have

$$I = (-1)^l \int_{-1}^{1} (1-x^2)^l [\mathrm{d}^{2l}(1-x^2)^l] \mathrm{d}x = (-1)^{2l}(2l)! \int_{-1}^{1} (1-x^2)^l \mathrm{d}x. \tag{3.230}$$

In (3.230), changing $x$ to $\cos\theta$, we have

$$\int_{-1}^{1} (1-x^2)^l \mathrm{d}x = \int_0^{\pi} \sin^{2l+1}\theta \mathrm{d}\theta. \tag{3.231}$$

We have already estimate this integral in (3.132) to have $\frac{2^{2l+1}(l!)^2}{(2l+1)!}$. Therefore,

$$f(0) = \frac{(-1)^{2l}(2l)!}{2^{2l}(l!)^2} \frac{2^{2l+1}(l!)^2}{(2l+1)!} = \frac{2}{2l+1}. \tag{3.232}$$

Thus, we get

$$f(m) = \frac{(l+m)!}{(l-m)!} f(0) = \frac{(l+m)!}{(l-m)!} \frac{2}{2l+1}. \tag{3.233}$$

From (3.228) to (3.233), we have

$$\int_{-1}^{1} P_l^m(x) P_{l'}^m(x) \mathrm{d}x = \frac{(l+m)!}{(l-m)!} \frac{2}{2l+1} \delta_{ll'}. \tag{3.234}$$

Accordingly, putting

$$\widetilde{P_l^m}(x) \equiv \sqrt{\frac{(2l+1)(l-m)!}{2(l+m)!}} P_l^m(x) \tag{3.235}$$

we get

$$\int_{-1}^{1} \widetilde{P_l^m}(x)\widetilde{P_{l'}^m}(x)\mathrm{d}x = \delta_{ll'}. \tag{3.236}$$

Normalized Legendre polynomials immediately follow. This is given by

$$\tilde{P}_l(x) \equiv \sqrt{\frac{2l+1}{2}} P_l(x) = \frac{(-1)^l}{2^l l!}\sqrt{\frac{2l+1}{2}}\frac{\mathrm{d}^l}{\mathrm{d}x^l}[(1-x^2)^l]. \tag{3.237}$$

Combining a normalized function (3.235) with $\frac{1}{\sqrt{2\pi}}\mathrm{e}^{im\phi}$, we recover

$$Y_l^m(\theta,\phi) = \sqrt{\frac{(2l+1)(l-m)!}{4\pi(l+m)!}} P_l^m(x)\mathrm{e}^{im\phi}(x=\cos\theta; 0\le\theta\le\pi). \tag{3.238}$$

Notice in (3.238), however, we could not determine Condon–Shortley phase $(-1)^m$; see (3.212).

Since $P_l^m(x)$ and $P_l^{-m}(x)$ are linearly dependent as noted in (3.210), the set of the associated Legendre functions cannot define a complete set of orthonormal system. In fact, we have

$$\int_{-1}^{1} P_l^m(x)P_l^{-m}(x)\mathrm{d}x = \frac{(-1)^m(l-m)!}{(l+m)!}\frac{(l+m)!}{(l-m)!}\frac{2}{2l+1} = \frac{2(-1)^m}{2l+1}. \tag{3.239}$$

This means that $P_l^m(x)$ and $P_l^{-m}(x)$ are *not* orthogonal. Thus, we need $\mathrm{e}^{im\phi}$ to constitute the complete set of orthonormal system. In other words,

$$\int_0^{2\pi}\mathrm{d}\phi\int_{-1}^{1}\mathrm{d}(\cos\theta)[Y_{l'}^{m'}(\theta,\phi)]^* Y_l^m(\theta,\phi) = \delta_{ll'}\delta_{mm'}. \tag{3.240}$$

## 3.7 Radial Wave Functions of Hydrogen-like Atoms

In Sect. 3.1 we have constructed Hamiltonian of hydrogen-like atoms. If the physical system is characterized by the central force field, the method of separation of variables into the angular part $(\theta,\phi)$ and radial $(r)$ part is successfully applied to

the problem and that method allows us to deal with the Schrödinger equation separately. The spherical surface harmonics play a central role in dealing with the differential equations related to the angular part. We studied important properties of the special functions such as Legendre polynomials and associated Legendre functions, independent of the nature of the specific central force fields such as Coulomb potential and Yukawa potential. With the Schrödinger equation pertinent to the radial part, on the other hand, its characteristics differ depending on the nature of individual force fields. Of these, the differential equation associated with the Coulomb potential gives exact (or analytical) solutions. It is well-known that the second-order differential equations are often solved by an operator representation method. Examples include its application to a quantum-mechanical harmonic oscillator and angular momenta of a particle placed in a central force field. Nonetheless, the corresponding approach to the radial equation for the electron has been less popular to date. The initial approach, however, was made by Sunakawa [3]. The purpose of this chapter rests upon further improvement of that approach.

### 3.7.1 Operator Approach to Radial Wave Functions

In Sect. 3.2, the separation of variables leaded to the radial part of the Schrödinger equation described as

$$\frac{1}{2\mu}\left[-\frac{\hbar^2}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial R(r)}{\partial r}\right)+\frac{\hbar^2\lambda}{r^2}\right]R(r)-\frac{Ze^2}{4\pi\varepsilon_0 r}R(r)=ER(r). \tag{3.51}$$

We identified $\lambda$ with $l(l+1)$ in (3.124). Thus, rewriting (3.51) and indexing $R(r)$ with $l$, we have

$$-\frac{\hbar^2}{2\mu r^2}\frac{\mathrm{d}}{\mathrm{d}r}\left[r^2\frac{\mathrm{d}R_l(r)}{\mathrm{d}r}\right]+\left[\frac{\hbar^2 l(l+1)}{2\mu r^2}-\frac{Ze^2}{4\pi\varepsilon_0 r}\right]R_l(r)=ER_l(r) \tag{3.241}$$

where $R_l(r)$ is a radial wave function parametrized with $l$; $\mu$, $Z$, $\varepsilon_0$, and $E$ denote a reduced mass of hydrogen-like atom, atomic number, permittivity of vacuum, and eigenvalue of energy, respectively. Otherwise we follow conventions.

Now, we are in position to solve (3.241). As in the cases of Chap. 2 of a quantum-mechanical harmonic oscillator and the previous section of the angular momentum operator, we present the operator formalism in dealing with radial wave functions of hydrogen-like atoms. The essential point rests upon that the radial wave functions can be derived by successively operating lowering operators on a radial wave function having a maximum allowed orbital angular momentum quantum number. The results agree with the conventional coordinate representation method based upon power series expansion that is related to associated Laguerre polynomials.

Sunakawa [3] introduced the following differential equation by suitable transformations of a variable, parameter, and function:

$$-\frac{\mathrm{d}^2\psi_l(\rho)}{\mathrm{d}\rho^2}+\left[\frac{l(l+1)}{\rho^2}-\frac{2}{\rho}\right]\psi_l(\rho)=\epsilon\,\psi_l(\rho), \tag{3.242}$$

where $\rho=\frac{Zr}{a}$, $\epsilon=\frac{2\mu}{\hbar^2}\left(\frac{a}{Z}\right)^2E$, and $\psi_l(\rho)=\rho R_l(r)$ with $a$ ($\equiv 4\pi\varepsilon_0\hbar^2/\mu e^2$) being Bohr radius of a hydrogen-like atom. Note that $\rho$ and $\epsilon$ are dimensionless quantities. The related calculations are as follows: We have

$$\frac{\mathrm{d}R_l}{\mathrm{d}r}=\frac{\mathrm{d}(\psi_l/\rho)}{\mathrm{d}\rho}\frac{\mathrm{d}\rho}{\mathrm{d}r}=\left(\frac{\mathrm{d}\psi_l}{\mathrm{d}\rho}\frac{1}{\rho}-\frac{\psi_l}{\rho^2}\right)\frac{Z}{a}.$$

Thus, we get

$$r^2\frac{\mathrm{d}R_l}{\mathrm{d}r}=\frac{\mathrm{d}\psi_l}{\mathrm{d}\rho}r-\frac{a}{Z}\psi_l,\quad \frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}R_l}{\mathrm{d}r}\right)=\frac{\mathrm{d}^2\psi_l(\rho)}{\mathrm{d}\rho^2}\rho.$$

Using the above relations, we arrive at (3.242).
Here, we define the following operators:

$$b_l\equiv\frac{\mathrm{d}}{\mathrm{d}\rho}+\left(\frac{l}{\rho}-\frac{1}{l}\right). \tag{3.243}$$

Hence,

$$b_l^\dagger=-\frac{\mathrm{d}}{\mathrm{d}\rho}+\left(\frac{l}{\rho}-\frac{1}{l}\right), \tag{3.244}$$

where the operator $b_l^\dagger$ is an adjoint operator of $b_l$. Notice that these definitions are different from those of Sunakawa [3]. The operator $\frac{\mathrm{d}}{\mathrm{d}\rho}(\equiv A)$ is *formally* an anti-Hermitian operator. We have mentioned such an operator in Sect. 1.5. The second terms of (3.243) and (3.244) are Hermitian operators, which we define as $H$. Thus, we foresee that $b_l$ and $b_l^\dagger$ can be denoted as follows:

$$b_l=A+H\quad\text{and}\quad b_l^\dagger=-A+H.$$

These representations are analogous to those appearing in the operator formalism of a quantum-mechanical harmonic oscillator. Special care, however, should be taken in dealing with the operators $b_l$ and $b_l^\dagger$. First, we should carefully examine whether $\frac{\mathrm{d}}{\mathrm{d}\rho}$ is in fact an anti-Hermitian operator. This is because for $\frac{\mathrm{d}}{\mathrm{d}\rho}$ to be anti-Hermitian, the solution $\psi_l(\rho)$ must satisfy boundary conditions in such a way

that $\psi_l(\rho)$ vanishes or takes the same value at the endpoints $\rho \to 0$ and $\infty$. Second, the coordinate system we have chosen is not Cartesian coordinate but the polar (spherical) coordinate, and so $\rho$ is defined only on a domain $\rho > 0$. We will come back to this point later.

Let us proceed on calculations. We have

$$\begin{aligned} b_l b_l^\dagger &= \left[\frac{\mathrm{d}}{\mathrm{d}\rho} + \left(\frac{l}{\rho} - \frac{1}{l}\right)\right] \cdot \left[-\frac{\mathrm{d}}{\mathrm{d}\rho} + \left(\frac{l}{\rho} - \frac{1}{l}\right)\right] \\ &= -\frac{\mathrm{d}^2}{\mathrm{d}\rho^2} + \frac{\mathrm{d}}{\mathrm{d}\rho}\left(\frac{l}{\rho} - \frac{1}{l}\right) - \left(\frac{l}{\rho} - \frac{1}{l}\right)\frac{\mathrm{d}}{\mathrm{d}\rho} + \frac{l^2}{\rho^2} - \frac{2}{\rho} + \frac{1}{l^2} \\ &= -\frac{\mathrm{d}^2}{\mathrm{d}\rho^2} - \frac{l}{\rho^2} + \frac{l^2}{\rho^2} - \frac{2}{\rho} + \frac{1}{l^2} = -\frac{\mathrm{d}^2}{\mathrm{d}\rho^2} + \frac{l(l-1)}{\rho^2} - \frac{2}{\rho} + \frac{1}{l^2}. \end{aligned} \tag{3.245}$$

Also, we have

$$b_l^\dagger b_l = -\frac{\mathrm{d}^2}{\mathrm{d}\rho^2} + \frac{l(l+1)}{\rho^2} - \frac{2}{\rho} + \frac{1}{l^2}. \tag{3.246}$$

We further define an operator $H^{(l)}$ as follows:

$$H^{(l)} \equiv -\frac{\mathrm{d}^2}{\mathrm{d}\rho^2} + \left[\frac{l(l+1)}{\rho^2} - \frac{2}{\rho}\right].$$

Then, from (3.243) to (3.244) as well as (3.245) and (3.246) we have

$$H^{(l)} = b_{l+1} b_{l+1}^\dagger + \varepsilon^{(l)} (l \geq 0), \tag{3.247}$$

where $\varepsilon^{(l)} \equiv -\frac{1}{(l+1)^2}$. Alternatively,

$$H^{(l)} = b_l^\dagger b_l + \varepsilon^{(l-1)} (l \geq 1). \tag{3.248}$$

If we put $l = n - 1$ in (3.247) with $n$ being a fixed given integer larger than $l$, we obtain

$$H^{(n-1)} = b_n b_n^\dagger + \varepsilon^{(n-1)}. \tag{3.249}$$

We evaluate the following inner product of both sides of (3.249):

$$\begin{aligned} \left\langle \chi \middle| H^{(n-1)} \middle| \chi \right\rangle &= \langle \chi | b_n b_n^\dagger | \chi \rangle + \varepsilon^{(n-1)} \langle \chi | \chi \rangle \\ &= \left\langle b_n^\dagger \chi \middle| b_n^\dagger \chi \right\rangle + \varepsilon^{(n-1)} \langle \chi | \chi \rangle \\ &= ||b_n^\dagger | \chi \rangle||^2 + \varepsilon^{(n-1)} \langle \chi | \chi \rangle \\ &\geq \varepsilon^{(n-1)} \end{aligned} \tag{3.250}$$

Here, we assume that $\chi$ is normalized. On the basis of the variational principle [9], the above expected value must take a minimum $\varepsilon^{(n-1)}$ so that $\chi$ can be an eigenfunction. To satisfy this condition, we have

$$|b_n^{\dagger}\chi> = 0. \tag{3.251}$$

In fact, if (3.251) holds, from (3.249) we have

$$H^{(n-1)}\chi = \varepsilon^{(n-1)}\chi. \tag{3.252}$$

We define such a function as below

$$\psi_{n-1}^{(n)} \equiv \chi. \tag{3.253}$$

From (3.247) to (3.248), we have the following relationship:

$$H^{(l)}b_{l+1} = b_{l+1}H^{(l+1)}(l \geq 0). \tag{3.254}$$

Meanwhile, we define the functions as shown below

$$\psi_{n-s}^{(n)} \equiv b_{n-s+1}b_{n-s+2}\ldots b_{n-1}\psi_{n-1}^{(n)}(2 \leq s \leq n). \tag{3.255}$$

With these functions $(s-1)$, operators have been operated on $\psi_{n-1}^{(n)}$. Note that if $s$ took 1, no operation of $b_l$ would take place. Thus, we find that $b_l$ functions upon the $l$-state to produce the $(l-1)$-state. That is, $b_l$ acts as an annihilation operator. For the sake of convenience, we express

$$H^{(n,s)} \equiv H^{(n-s)}. \tag{3.256}$$

Using this notation and (3.254), we have

$$\begin{aligned}
H^{(n,s)}\psi_{n-s}^{(n)} &= H^{(n,s)}b_{n-s+1}b_{n-s+2}\ldots b_{n-1}\psi_{n-1}^{(n)} \\
&= b_{n-s+1}H^{(n,s-1)}b_{n-s+2}\ldots b_{n-1}\psi_{n-1}^{(n)} \\
&= b_{n-s+1}b_{n-s+2}H^{(n,s-2)}\ldots b_{n-1}\psi_{n-1}^{(n)} \\
&\ldots \\
&= b_{n-s+1}b_{n-s+2}\ldots H^{(n,2)}b_{n-1}\psi_{n-1}^{(n)} \\
&= b_{n-s+1}b_{n-s+2}\ldots b_{n-1}H^{(n,1)}\psi_{n-1}^{(n)} \\
&= b_{n-s+1}b_{n-s+2}\ldots b_{n-1}\varepsilon^{(n-1)}\psi_{n-1}^{(n)} \\
&= \varepsilon^{(n-1)}b_{n-s+1}b_{n-s+2}\ldots b_{n-1}\psi_{n-1}^{(n)} \\
&= \varepsilon^{(n-1)}\psi_{n-s}^{(n)}.
\end{aligned} \tag{3.257}$$

Thus, total $n$ functions $\psi_{n-s}^{(n)}$ $(1 \leq s \leq n)$ belong to the same eigenvalue $\varepsilon^{(n-1)}$. Notice that the eigenenergy $E_n$ corresponding to $\varepsilon^{(n-1)}$ is given by

$$E_n = -\frac{\hbar^2}{2\mu}\left(\frac{Z}{a}\right)^2 \frac{1}{n^2}. \tag{3.258}$$

If we define $l \equiv n - s$ and take account of (3.252), total $n$ functions $\psi_l^{(n)}$ ($l$ = 0, 1, 2, … , $n$ − 1) belong to the same eigenvalue $\varepsilon^{(n-1)}$.

The quantum state $\psi_l^{(n)}$ is associated with the operators $H^{(l)}$. Thus, the solution of (3.242) has been given by functions $\psi_l^{(n)}$ parametrized with $n$ and $l$ on condition that (3.251) holds. As explicitly indicated in (3.255) and (3.257), $b_l$ lowers the parameter $l$ by one from $l$ to $l$ − 1, when it operates on $\psi_l^{(n)}$. The operator $b_0$ cannot be defined as indicated in (3.243), and so the lowest number of $l$ should be zero. Operators such as $b_l$ are known as a ladder operator (lowering operator or annihilation operator in the present case). The implication is that the successive operations of $b_l$ on $\psi_{n-1}^{(n)}$ produce various parameters $l$ as a subscript down to zero, while retaining the same integer parameter $n$ as a superscript.

### 3.7.2 *Normalization of Radial Wave Functions*

Next, we seek normalized eigenfunctions. Coordinate representation of (3.251) takes

$$-\frac{\mathrm{d}\psi_{n-1}^{(n)}}{\mathrm{d}\rho} + \left(\frac{n}{\rho} - \frac{1}{n}\right)\psi_{n-1}^{(n)} = 0. \tag{3.259}$$

The solution can be obtained as

$$\psi_{n-1}^{(n)} = c_n \rho^n \mathrm{e}^{-\rho/n}, \tag{3.260}$$

where $c_n$ is a normalization constant. This can be determined as follows:

$$\int_0^\infty |\psi_{n-1}^{(n)}|^2 \mathrm{d}\rho = 1. \tag{3.261}$$

Namely,

$$|c_n|^2 \int_0^\infty \rho^{2n} \mathrm{e}^{-2\rho/n} \mathrm{d}\rho = 1. \tag{3.262}$$

Consider the following definite integral:

$$\int_0^\infty \mathrm{e}^{-2\rho\xi}\mathrm{d}\rho = \frac{1}{2\xi}.$$

Differentiating the above integral $2n$ times with respect to $\xi$ gives

$$\int_0^\infty \rho^{2n}\mathrm{e}^{-2\rho\xi}\mathrm{d}\rho = \left(\frac{1}{2}\right)^{2n+1}(2n)!\xi^{-(2n+1)}. \tag{3.263}$$

Substituting $1/n$ into $\xi$, we obtain

$$\int_0^\infty \rho^{2n}\mathrm{e}^{-2\rho/n}\mathrm{d}\rho = \left(\frac{1}{2}\right)^{2n+1}(2n)!n^{(2n+1)}. \tag{3.264}$$

Hence,

$$c_n = \left(\frac{2}{n}\right)^{n+\frac{1}{2}}/\sqrt{(2n)!}. \tag{3.265}$$

To further normalize the other wave functions, we calculate the following inner product:

$$\left\langle \psi_l^{(n)} \,\middle|\, \psi_l^{(n)} \right\rangle = \left\langle \psi_{n-1}^{(n)} b_{n-1}^\dagger \ldots b_{l+2}^\dagger b_{l+1}^\dagger \,\middle|\, b_{l+1}b_{l+2}\ldots b_{n-1}\psi_{n-1}^{(n)} \right\rangle. \tag{3.266}$$

From (3.247) to (3.248), we have

$$b_l^\dagger b_l + \varepsilon^{(l-1)} = b_{l+1}b_{l+1}^\dagger + \varepsilon^{(l)} \ (l \geq 1). \tag{3.267}$$

Applying (3.267–3.266) repeatedly and considering (3.251), we reach the following relationship:

$$\begin{aligned} &\left\langle \psi_l^{(n)} \,\middle|\, \psi_l^{(n)} \right\rangle \\ &= \left[\varepsilon^{(n-1)} - \varepsilon^{(n-2)}\right]\left[\varepsilon^{(n-1)} - \varepsilon^{(n-3)}\right]\ldots\left[\varepsilon^{(n-1)} - \varepsilon^{(l)}\right]\left\langle \psi_{n-1}^{(n)} \,\middle|\, \psi_{n-1}^{(n)} \right\rangle. \end{aligned} \tag{3.268}$$

To show this, we use mathematical induction. We have already normalized $\psi_{n-1}^{(n)}$ in (3.261). Next, we calculate $\left\langle \psi_{n-2}^{(n)} \,\middle|\, \psi_{n-2}^{(n)} \right\rangle$ such that

$$\begin{aligned}\left\langle \psi_{n-2}^{(n)} \,\middle|\, \psi_{n-2}^{(n)} \right\rangle &= \left\langle \psi_{n-1}^{(n)} b_{n-1}^{\dagger} | b_{n-1} \psi_{n-1}^{(n)} \right\rangle = \left\langle \psi_{n-1}^{(n)} b_{n-1}^{\dagger} b_{n-1} \psi_{n-1}^{(n)} \right\rangle \\ &= \left\langle \psi_{n-1}^{(n)} \left[ b_n b_n^{\dagger} + \varepsilon^{(n-1)} - \varepsilon^{(n-2)} \right] \psi_{n-1}^{(n)} \right\rangle \\ &= \left\langle \psi_{n-1}^{(n)} b_n b_n^{\dagger} \psi_{n-1}^{(n)} \right\rangle + \left[ \varepsilon^{(n-1)} - \varepsilon^{(n-2)} \right] \left\langle \psi_{n-1}^{(n)} \,\middle|\, \psi_{n-1}^{(n)} \right\rangle \\ &= \left[ \varepsilon^{(n-1)} - \varepsilon^{(n-2)} \right] \left\langle \psi_{n-1}^{(n)} \,\middle|\, \psi_{n-1}^{(n)} \right\rangle. \end{aligned} \tag{3.269}$$

With the third equality, we used (3.267) with $l = n - 1$; with the last equality we used (3.251). Therefore, (3.268) holds with $l = n - 2$. Then, it suffices to show that assuming that (3.268) holds with $\left\langle \psi_{l+1}^{(n)} \,\middle|\, \psi_{l+1}^{(n)} \right\rangle$, it holds with $\left\langle \psi_{l}^{(n)} \,\middle|\, \psi_{l}^{(n)} \right\rangle$ as well.

Let us calculate $\left\langle \psi_{l}^{(n)} \,\middle|\, \psi_{l}^{(n)} \right\rangle$, starting with (3.266) as below:

$$\begin{aligned}\left\langle \psi_l^{(n)} | \psi_l^{(n)} \right\rangle &= \left\langle \psi_{n-1}^{(n)} b_{n-1}^{\dagger} \ldots b_{l+2}^{\dagger} b_{l+1}^{\dagger} | b_{l+1} b_{l+2} \ldots b_{n-1} \psi_{n-1}^{(n)} \right\rangle \\ &= \left\langle \psi_{n-1}^{(n)} b_{n-1}^{\dagger} \ldots b_{l+2}^{\dagger} (b_{l+1}^{\dagger} b_{l+1}) b_{l+2} \ldots b_{n-1} \psi_{n-1}^{(n)} \right\rangle \\ &= \left\langle \psi_{n-1}^{(n)} b_{n-1}^{\dagger} \ldots b_{l+2}^{\dagger} \left[ b_{l+2} b_{l+2}^{\dagger} + \varepsilon^{(l+1)} - \varepsilon^{(l)} \right] b_{l+2} \ldots b_{n-1} \psi_{n-1}^{(n)} \right\rangle \\ &= \left\langle \psi_{n-1}^{(n)} b_{n-1}^{\dagger} \ldots b_{l+2}^{\dagger} b_{l+2} (b_{l+2}^{\dagger} b_{l+2}) \ldots b_{n-1} \psi_{n-1}^{(n)} \right\rangle \\ &\quad + \left[ \varepsilon^{(l+1)} - \varepsilon^{(l)} \right] \left\langle \psi_{n-1}^{(n)} b_{n-1}^{\dagger} \ldots b_{l+2}^{\dagger} b_{l+2} \ldots b_{n-1} \psi_{n-1}^{(n)} \right\rangle \\ &= \left\langle \psi_{n-1}^{(n)} b_{n-1}^{\dagger} \ldots b_{l+2}^{\dagger} b_{l+2} (b_{l+2}^{\dagger} b_{l+2}) \ldots b_{n-1} \psi_{n-1}^{(n)} \right\rangle \\ &\quad + \left[ \varepsilon^{(l+1)} - \varepsilon^{(l)} \right] \left\langle \psi_{l+1}^{(n)} | \psi_{l+1}^{(n)} \right\rangle. \end{aligned} \tag{3.270}$$

In the next step, using $b_{l+2}^{\dagger} b_{l+2} = b_{l+3} b_{l+3}^{\dagger} + \varepsilon^{(l+2)} - \varepsilon^{(l+1)}$, we have

$$\begin{aligned}\left\langle \psi_{l}^{(n)} \,\middle|\, \psi_{l}^{(n)} \right\rangle &= \left\langle \psi_{n-1}^{(n)} b_{n-1}^{\dagger} \ldots b_{l+2}^{\dagger} b_{l+2} b_{l+3} b_{l+3}^{\dagger} b_{l+3} \ldots b_{n-1} \psi_{n-1}^{(n)} \right\rangle \\ &\quad + \left[ \varepsilon^{(l+1)} - \varepsilon^{(l)} \right] \left\langle \psi_{l+1}^{(n)} \,\middle|\, \psi_{l+1}^{(n)} \right\rangle + \left[ \varepsilon^{(l+2)} - \varepsilon^{(l+1)} \right] \left\langle \psi_{l+1}^{(n)} \,\middle|\, \psi_{l+1}^{(n)} \right\rangle. \end{aligned} \tag{3.271}$$

Thus, we find that in the first term the index number of $b_{l+3}^{\dagger}$ has been increased by one with itself transferred toward the right side. On the other hand, we notice that with the second and third terms $\varepsilon^{(l+1)} \left\langle \psi_{l+1}^{(n)} \,\middle|\, \psi_{l+1}^{(n)} \right\rangle$ cancels out. Repeating the above processes, we reach a following expression:

$$\begin{aligned}\left\langle \psi_l^{(n)} \,\middle|\, \psi_l^{(n)} \right\rangle &= \left\langle \psi_{n-1}^{(n)} b_{n-1}^\dagger \ldots b_{l+2}^\dagger b_{l+2} \ldots b_{n-1} b_{n-1}^\dagger \psi_{n-1}^{(n)} \right\rangle \\ &\quad + \left[\varepsilon^{(n-1)} - \varepsilon^{(n-2)}\right] \left\langle \psi_{l+1}^{(n)} \,\middle|\, \psi_{l+1}^{(n)} \right\rangle + \left[\varepsilon^{(n-2)} - \varepsilon^{(n-3)}\right] \left\langle \psi_{l+1}^{(n)} \,\middle|\, \psi_{l+1}^{(n)} \right\rangle + \cdots \\ &\quad + \left[\varepsilon^{(l+2)} - \varepsilon^{(l+1)}\right] \left\langle \psi_{l+1}^{(n)} \,\middle|\, \psi_{l+1}^{(n)} \right\rangle + \left[\varepsilon^{(l+1)} - \varepsilon^{(l)}\right] \left\langle \psi_{l+1}^{(n)} \,\middle|\, \psi_{l+1}^{(n)} \right\rangle \\ &= \left[\varepsilon^{(n-1)} - \varepsilon^{(l)}\right] \left\langle \psi_{l+1}^{(n)} \,\middle|\, \psi_{l+1}^{(n)} \right\rangle. \end{aligned} \tag{3.272}$$

In (3.272), the first term of RHS vanishes because of (3.251); the subsequent term produces $\left\langle \psi_{l+1}^{(n)} \,\middle|\, \psi_{l+1}^{(n)} \right\rangle$ whose coefficients have canceled one another except for $\left[\varepsilon^{(n-1)} - \varepsilon^{(l)}\right]$.

Meanwhile, from assumption of the mathematical induction we have

$$\begin{aligned}&\left\langle \psi_{l+1}^{(n)} \,\middle|\, \psi_{l+1}^{(n)} \right\rangle \\ &= \left[\varepsilon^{(n-1)} - \varepsilon^{(n-2)}\right] \left[\varepsilon^{(n-1)} - \varepsilon^{(n-3)}\right] \ldots \left[\varepsilon^{(n-1)} - \varepsilon^{(l+1)}\right] \left\langle \psi_{n-1}^{(n)} \,\middle|\, \psi_{n-1}^{(n)} \right\rangle. \end{aligned}$$

Inserting this equation into (3.272), we arrive at (3.268). In other words, we have shown that if (3.268) holds with $l = l+1$, (3.268) holds with $l = l$ as well. This completes the proof to show that (3.268) is true of $\left\langle \psi_l^{(n)} \,\middle|\, \psi_l^{(n)} \right\rangle$ with $l$ down to 0.

The normalized wave functions $\tilde{\psi}_l^{(n)}$ are expressed from (3.255) as

$$\tilde{\psi}_l^{(n)} = \kappa(n,l)^{-\frac{1}{2}} b_{l+1} b_{l+2} \ldots b_{n-1} \tilde{\psi}_{n-1}^{(n)}, \tag{3.273}$$

where $\kappa(n,l)$ is defined such that

$$\kappa(n,l) \equiv \left[\varepsilon^{(n-1)} - \varepsilon^{(n-2)}\right] \cdot \left[\varepsilon^{(n-1)} - \varepsilon^{(n-3)}\right] \ldots \left[\varepsilon^{(n-1)} - \varepsilon^{(l)}\right], \tag{3.274}$$

with $l \leq n-2$. More explicitly, we get

$$\kappa(n,l) = \frac{(2n-1)!(n-l-1)!(l!)^2}{(n+l)!(n!)^2(n^{n-l-2})^2}. \tag{3.275}$$

In particular, from (3.265) we have

$$\tilde{\psi}_{n-1}^{(n)} = \left(\frac{2}{n}\right)^{n+\frac{1}{2}} \frac{1}{\sqrt{(2n)!}} \rho^n \mathrm{e}^{-\rho/n}. \tag{3.276}$$

From (3.272), we define the following operator:

$$\tilde{b}_l \equiv [\varepsilon^{(n-1)} - \varepsilon^{(l-1)}]^{-\frac{1}{2}} b_l. \tag{3.277}$$

Then, (3.273) becomes

$$\tilde{\psi}_l^{(n)} = \tilde{b}_{l+1}\tilde{b}_{l+2}\ldots\tilde{b}_{n-1}\tilde{\psi}_{n-1}^{(n)}. \tag{3.278}$$

### 3.7.3 Associated Laguerre Polynomials

It will be of great importance to compare the functions $\tilde{\psi}_l^{(n)}$ with conventional wave functions that are expressed using associated Laguerre polynomials. For this purpose, we define the following functions $\Phi_l^{(n)}(\rho)$ such that

$$\Phi_l^{(n)}(\rho) \equiv \left(\frac{2}{n}\right)^{l+\frac{3}{2}} \sqrt{\frac{(n-l-1)!}{2n(n+l)!}} \mathrm{e}^{-\frac{\rho}{n}} \rho^{l+1} L_{n-l-1}^{2l+1}\left(\frac{2\rho}{n}\right). \tag{3.279}$$

The associated Laguerre polynomials are described as

$$L_n^{\nu}(x) = \frac{1}{n!} x^{-\nu} \mathrm{e}^{x} \frac{\mathrm{d}^n}{\mathrm{d}x^n}(x^{n+\nu}\mathrm{e}^{-x}), (\nu > -1). \tag{3.280}$$

In a form of power series expansion, the polynomials are expressed for integer $k \geq 0$ as

$$L_n^k(x) = \sum_{m=0}^{n} \frac{(-1)^m (n+k)!}{(n-m)!(k+m)!m!} x^m. \tag{3.281}$$

Notice that "Laguerre polynomials" $L_n(x)$ are defined as

$$L_n(x) \equiv L_n^0(x).$$

Hence, instead of (3.280) and (3.281), the Rodrigues formula and power series expansion of $L_n(x)$ are given by [2, 5]

$$L_n(x) = \frac{1}{n!} \mathrm{e}^{x} \frac{\mathrm{d}^n}{\mathrm{d}x^n}(x^n \mathrm{e}^{-x}),$$
$$L_n(x) = \sum_{m=0}^{n} \frac{(-1)^m n!}{(n-m)!(m!)^2} x^m.$$

The function $\Phi_l^{(n)}(\rho)$ contains multiplication factors $\mathrm{e}^{-\frac{\rho}{n}}$ and $\rho^{l+1}$. The function $L_{n-l-1}^{2l+1}\left(\frac{2\rho}{n}\right)$ is a polynomial of $\rho$ with the highest order of $\rho^{n-l-1}$. Therefore, $\Phi_l^{(n)}(\rho)$

consists of summation of terms containing $e^{-\frac{\rho}{n}}\rho^t$, where $t$ is an integer equal to 1 or larger. Consequently, $\Phi_l^{(n)}(\rho) \to 0$ when $\rho \to 0$ and $\rho \to \infty$ (vide supra). Thus, we have confirmed that $\Phi_l^{(n)}(\rho)$ certainly satisfies proper BCs mentioned earlier and, hence, the operator $\frac{d}{d\rho}$ is indeed an anti-Hermitian. To show this more explicitly, we define $D \equiv \frac{d}{d\rho}$. An inner product between arbitrarily chosen functions $f$ and $g$ is

$$\begin{aligned}\langle f \mid Dg\rangle &\equiv \int_0^\infty f^* Dg \mathrm{d}\rho \\ &= [f^* g]_0^\infty - \int_0^\infty (Df^*) g \mathrm{d}\rho \\ &= [f^* g]_0^\infty + < - Df^* | g\rangle,\end{aligned} \tag{3.282}$$

where $f^*$ is a complex conjugate of $f$. Meanwhile, from (1.112) we have

$$\langle f | Dg\rangle = \left\langle D^\dagger f | g \right\rangle \tag{3.283}$$

Therefore if the functions $f$ and $g$ vanish at $\rho \to 0$ and $\rho \to \infty$, $D^\dagger = -D$ by equating (3.282) and (3.283). This means that $D$ is anti-Hermitian. The functions $\Phi_l^{(n)}(\rho)$ we are dealing with certainly satisfy the required boundary conditions. The operator $H^{(l)}$ appearing in (3.247) and (3.248) is Hermitian accordingly. This is because

$$b_l^\dagger b_l = (-A + H)(A + H) = H^2 - A^2 - AH + HA, \tag{3.284}$$

$$\begin{aligned}(b_l^\dagger b_l)^\dagger &= (H^2)^\dagger - (A^2)^\dagger - H^\dagger A^\dagger + A^\dagger H^\dagger = (H^\dagger)^2 - (A^\dagger)^2 - H^\dagger A^\dagger + A^\dagger H^\dagger \\ &= H^2 - (-A)^2 - H(-A) + (-A)H = H^2 - A^2 + HA - AH = b_l^\dagger b_l.\end{aligned} \tag{3.285}$$

The Hermiticity is true of $b_l b_l^\dagger$ as well. Thus, the eigenvalue and eigenstate (or wave function) which belongs to that eigenvalue are physically meaningful.

Next, consider the following operation:

$$\begin{aligned}&\tilde{b}_l \Phi_l^{(n)}(\rho) \\ &= [\varepsilon^{(n-1)} - \varepsilon^{(l-1)}]^{-\frac{1}{2}} \left(\frac{2}{n}\right)^{l+\frac{3}{2}} \sqrt{\frac{(n-l-1)!}{2n(n+l)!}} \left[\frac{\mathrm{d}}{\mathrm{d}\rho} + \left(\frac{l}{\rho} - \frac{1}{l}\right)\right] \left\{ e^{-\frac{\rho}{n}} \rho^{l+1} L_{n-l-1}^{2l+1}\left(\frac{2\rho}{n}\right)\right\},\end{aligned} \tag{3.286}$$

where

$$[\varepsilon^{(n-1)} - \varepsilon^{(l-1)}]^{-\frac{1}{2}} = \frac{nl}{\sqrt{(n+l)(n-l)}}. \tag{3.287}$$

Rewriting $L_{n-l-1}^{2l+1}\left(\frac{2\rho}{n}\right)$ in a power series expansion form using (3.280) and rearranging the result, we obtain

$$\begin{aligned}
&\tilde{b}_l \Phi_l^{(n)}(\rho) \\
&= \left(\frac{2}{n}\right)^{l+\frac{3}{2}} \frac{nl}{\sqrt{(n+l)(n-l)}} \sqrt{\frac{1}{2n(n+l)!(n-l-1)!}} \left[\frac{\mathrm{d}}{\mathrm{d}\rho} + \left(\frac{l}{\rho} - \frac{1}{l}\right)\right] \\
&\quad \left[\mathrm{e}^{\rho/n} \rho^{-l} \frac{\mathrm{d}^{n-l-1}}{\mathrm{d}\rho^{n-l-1}} (\rho^{n+l} \mathrm{e}^{-2\rho/n})\right] \\
&= \left(\frac{2}{n}\right)^{l+\frac{3}{2}} \frac{nl}{\sqrt{(n+l)(n-l)}} \sqrt{\frac{1}{2n(n+l)!(n-l-1)!}} \left[\frac{\mathrm{d}}{\mathrm{d}\rho} + \left(\frac{l}{\rho} - \frac{1}{l}\right)\right] \\
&\quad \times \mathrm{e}^{-\rho/n} \sum_{m=0}^{n-l-1} \frac{(-1)^m (n+l)!}{(2l+m+1)!} \rho^{l+m+1} \left(\frac{2}{n}\right)^m \frac{(n-l-1)!}{m!(n-l-m-1)!},
\end{aligned} \tag{3.288}$$

where we used well-known Leibniz rule of higher order differentiation of a product function, i.e., $\frac{\mathrm{d}^{n-l-1}}{\mathrm{d}\rho^{n-l-1}}(\rho^{n+l}\mathrm{e}^{-2\rho/n})$. To perform further calculation, notice that $\frac{\mathrm{d}}{\mathrm{d}\rho}$ does not change a functional form of $\mathrm{e}^{-\rho/n}$, whereas it lowers the order of $\rho^{l+m+1}$ by one. Meanwhile, operation of $\frac{l}{\rho}$ lowers the order of $\rho^{l+m+1}$ by one as well. The factor $\frac{n+l}{2l}$ in Eq. (3.289) results from these calculation processes. Considering these characteristics of the operator, we get

$$\begin{aligned}
\tilde{b}_l \Phi_l^{(n)}(\rho) = \left(\frac{2}{n}\right)^{l+\frac{3}{2}} &\frac{n+l}{2l} \frac{nl}{\sqrt{(n+l)(n-l)}} \sqrt{\frac{1}{2n(n+l)!(n-l-1)!}} \mathrm{e}^{-\rho/n} \rho^l \\
&\times \left\{ \sum_{m=0}^{n-l-1} \frac{(-1)^{m+1}(n+l)!}{(2l+m+1)!} \frac{(n-l-1)!}{m!(n-l-m-1)!} \rho^{m+1} \left(\frac{2}{n}\right)^{m+1} \right. \\
&\left. + 2l \sum_{m=0}^{n-l-1} \frac{(-1)^m (n+l-1)!}{(2l+m)!} \frac{(n-l-1)!}{m!(n-l-m-1)!} \rho^m \left(\frac{2}{n}\right)^m \right\}
\end{aligned} \tag{3.289}$$

In (3.289), calculation of the part $\{\cdots\}$ next to the multiplication sign for RHS is somewhat complicated, and so we describe the outline of the calculation procedure below.

$\{\cdots\}$ of RHS of (3.289)

$$
\begin{aligned}
&= \sum_{m=1}^{n-l} \frac{(-1)^m (n+l)!}{(2l+m)!} \frac{(n-l-1)!}{(m-1)!(n-l-m)!} \rho^m \left(\frac{2}{n}\right)^m \\
&\quad + 2l \sum_{m=0}^{n-l-1} \frac{(-1)^m (n+l-1)!}{(2l+m)!} \frac{(n-l-1)!}{m!(n-l-m-1)!} \rho^m \left(\frac{2}{n}\right)^m \\
&= \sum_{m=1}^{n-l-1} \frac{(-1)^m \left(\frac{2\rho}{n}\right)^m (n-l-1)!(n+l-1)!}{(2l+m)!} \left[\frac{n+l}{(m-1)!(n-l-m)!} + \frac{2l}{m!(n-l-m-1)!}\right] \\
&\quad + (-1)^{n-l} \left(\frac{2\rho}{n}\right)^{n-l} + \frac{(n+l-1)!}{(2l-1)!} \\
&= \sum_{m=1}^{n-l-1} \frac{(-1)^m \left(\frac{2\rho}{n}\right)^m (n-l-1)!(n+l-1)!}{(2l+m)!} \cdot \frac{(m-1)!(n-l-m-1)!(2l+m)(n-l)}{(m-1)!(n-l-m)!m!(n-l-m-1)!} \\
&\quad + (-1)^{n-l} \left(\frac{2\rho}{n}\right)^{n-l} + \frac{(n+l-1)!}{(2l-1)!}. \\
&= \sum_{m=1}^{n-l-1} \frac{(-1)^m \left(\frac{2\rho}{n}\right)^m (n-l)!(n+l-1)!}{(2l+m-1)!(n-l-m)!m!} + (-1)^{n-l} \left(\frac{2\rho}{n}\right)^{n-l} + \frac{(n+l-1)!}{(2l-1)!} \\
&= (n-l)! \left[\sum_{m=1}^{n-l-1} \frac{(-1)^m \left(\frac{2\rho}{n}\right)^m (n+l-1)!}{(2l+m-1)!(n-l-m)!m!} + (-1)^{n-l} \frac{1}{(n-l)!} \left(\frac{2\rho}{n}\right)^{n-l} + \frac{(n+l-1)!}{(n-l)!(2l-1)!}\right] \\
&= (n-l)! \sum_{m=0}^{n-l} \frac{(-1)^m \left(\frac{2\rho}{n}\right)^m (n+l-1)!}{(2l+m-1)!(n-l-m)!m!} \\
&= (n-l)! L_{n-l}^{2l-1} \left(\frac{2\rho}{n}\right).
\end{aligned}
\tag{3.290}
$$

Notice that with the second equality of (3.290), the summation is divided into three terms, i.e., $1 \le m \le n-l-1$, $m = n-l$ (the highest-order term), and $m = 0$ (the lowest-order term). Note that with the second last equality of (3.290), the highest-order $(n-l)$ term and the lowest-order term (i.e., a constant) have been absorbed in a single equation, namely an associated Laguerre polynomial. Correspondingly, with the second last equality of (3.290) the summation range is extended to $0 \le m \le n-l$.

Summarizing the above results, we get

$$
\begin{aligned}
&\tilde{b}_l \Phi_l^{(n)}(\rho) \\
&\qquad = \left(\frac{2}{n}\right)^{l+\frac{3}{2}} \frac{n+l}{2l} \frac{nl(n-l)!}{\sqrt{(n+l)(n-l)}} \sqrt{\frac{1}{2n(n+l)!(n-l-1)!}} \mathrm{e}^{-\rho/n} \rho^l L_{n-l}^{2l-1} \left(\frac{2\rho}{n}\right) \\
&\qquad = \left(\frac{2}{n}\right)^{l+\frac{1}{2}} \sqrt{\frac{(n-l)!}{2n(n+l-1)!}} \mathrm{e}^{-\rho/n} \rho^l L_{n-l}^{2l-1} \left(\frac{2\rho}{n}\right) \\
&\qquad = \left(\frac{2}{n}\right)^{(l-1)+\frac{3}{2}} \sqrt{\frac{[n-(l-1)-1]!}{2n[n+(l-1)]!}} \mathrm{e}^{-\frac{\rho}{n}} \rho^{(l-1)+1} L_{n-(l-1)-1}^{2(l-1)+1} \left(\frac{2\rho}{n}\right) \equiv \Phi_{l-1}^{(n)}(\rho).
\end{aligned}
\tag{3.291}
$$

Thus, we find out that $\Phi_l^{(n)}(\rho)$ behaves exactly like $\tilde{\psi}_l^{(n)}$. Moreover, if we replace $l$ in (3.279) with $n-1$, we find

$$\Phi_{n-1}^{(n)}(\rho) = \tilde{\psi}_{n-1}^{(n)}. \tag{3.292}$$

Operating $\tilde{b}_{n-1}$ on both sides of (3.292), we get

$$\Phi_{n-2}^{(n)}(\rho) = \tilde{\psi}_{n-2}^{(n)}. \tag{3.293}$$

Likewise successively operating $\tilde{b}_l$ $(1 \le l \le n-1)$, we get

$$\Phi_l^{(n)}(\rho) = \tilde{\psi}_l^{(n)}(\rho), \tag{3.294}$$

with all allowed numbers of $l$ (i.e., $0 \le l \le n-1$). This permits us to identify

$$\Phi_l^{(n)}(\rho) \equiv \tilde{\psi}_l^{(n)}(\rho). \tag{3.295}$$

Consequently, it is clear that the parameter $n$ introduced in (3.249) is identical to a principal quantum number and that the parameter $l$ $(0 \le l \le n-1)$ is an orbital angular momentum quantum number. The functions $\Phi_l^{(n)}(\rho)$ and $\tilde{\psi}_l^{(n)}(\rho)$ are identical up to the constant $c_n$ expressed in (3.265). Note, however, that a complex constant with an absolute number of 1 (phase factor) remains undetermined, as is always the case with the eigenvalue problem.

The radial wave functions are derived from the following relationship as described earlier:

$$R_l^{(n)}(r) = \tilde{\psi}_l^{(n)}/\rho. \tag{3.296}$$

To normalize $R_l^{(n)}(r)$, we have to calculate the following integral:

$$\int_0^\infty |R_l^{(n)}(r)|^2 r^2 \mathrm{d}r = \int_0^\infty \frac{1}{\rho^2} |\tilde{\psi}_l^{(n)}|^2 \left(\frac{a}{Z}\rho\right)^2 \frac{a}{Z} \mathrm{d}\rho = \left(\frac{a}{Z}\right)^3 \int_0^\infty |\tilde{\psi}_l^{(n)}|^2 \mathrm{d}\rho = \left(\frac{a}{Z}\right)^3. \tag{3.297}$$

Accordingly, we choose the following functions $\tilde{R}_l^{(n)}(r)$ for the normalized radial wave functions:

$$\tilde{R}_l^{(n)}(r) = \sqrt{(Z/a)^3} R_l^{(n)}(r). \tag{3.298}$$

Substituting (3.296) into (3.298) and taking account of (3.279) and (3.280), we obtain

$$\tilde{R}_l^{(n)}(r) = \sqrt{\left(\frac{2Z}{an}\right)^3 \cdot \frac{(n-l-1)!}{2n(n+l)!}} \left(\frac{2Zr}{an}\right)^l \exp\left[-\left(\frac{Zr}{an}\right)\right] L_{n-l-1}^{2l+1}\left(\frac{2Zr}{an}\right). \quad (3.299)$$

Equation (3.298) is exactly the same as the normalized radial wave functions that can be obtained as the solution of (3.241) through the power series expansion. All these functions belong to the same eigenenergy $E_n = -\frac{\hbar^2}{2\mu}\left(\frac{Z}{a}\right)^2 \frac{1}{n^2}$.

In summary of this section, we have developed the operator formalism in dealing with radial wave functions of hydrogen-like atoms and seen how the operator formalism features the radial wave functions. The essential point rests upon that the radial wave functions can be derived by successively operating the lowering operators $b_l$ on $\tilde{\psi}_{n-1}^{(n)}$ that is parametrized with a principal quantum number $n$ and an orbital angular momentum quantum number $l = n - 1$. This is clearly represented by (3.278). The results agree with the conventional coordinate representation method based upon the power series expansion that leads to associated Laguerre polynomials. Thus, the operator formalism is again found to be powerful in explicitly representing the mathematical constitution of quantum-mechanical systems.

## 3.8 Total Wave Functions

Since we have obtained angular wave functions and radial wave functions, we describe normalized total wave functions $\tilde{\Lambda}_{l,m}^{(n)}$ of hydrogen-like atoms as a product of the angular part and radial part such that

$$\tilde{\Lambda}_{l,m}^{(n)} = Y_l^m(\theta, \phi)\tilde{R}_l^{(n)}(r). \quad (3.300)$$

Let us seek several tangible functional forms of hydrogen $(Z = 1)$ including angular and radial parts. For example, we have

$$\phi(1s) \equiv Y_0^0(\theta, \phi)\tilde{R}_0^{(1)}(r) = \sqrt{\frac{1}{4\pi}} a^{-3/2} \left(\frac{\tilde{\psi}_{n-1}^{(n)}}{\rho}\right) = \sqrt{\frac{1}{\pi}} a^{-3/2} \mathrm{e}^{-r/a}, \quad (3.301)$$

where we used (3.276) and (3.295).

For $\phi(2s)$, using (3.277) and (3.278) we have

$$\phi(2s) \equiv Y_0^0(\theta, \phi)\tilde{R}_0^{(2)}(r) = \frac{1}{4\sqrt{2\pi}} a^{-\frac{3}{2}} \mathrm{e}^{-\frac{r}{2a}} \left(2 - \frac{r}{a}\right). \quad (3.302)$$

For $\phi(2p_z)$, in turn, we express it as

$$\phi(2p_z) \equiv Y_1^0(\theta,\phi)\tilde{R}_1^{(2)}(r) = \sqrt{\frac{3}{4\pi}}(\cos\theta)\frac{1}{2\sqrt{6}}a^{-\frac{3}{2}}\frac{r}{a}\mathrm{e}^{-\frac{r}{2a}}$$
$$= \frac{1}{4\sqrt{2\pi}}a^{-\frac{3}{2}}\frac{r}{a}\mathrm{e}^{-\frac{r}{2a}}\cos\theta = \frac{1}{4\sqrt{2\pi}}a^{-\frac{3}{2}}\frac{r}{a}\mathrm{e}^{-\frac{r}{2a}}\frac{z}{r} = \frac{1}{4\sqrt{2\pi}}a^{-\frac{5}{2}}\mathrm{e}^{-\frac{r}{2a}}z. \quad (3.303)$$

For $\phi(2p_{x+iy})$, using (3.217) we get

$$\begin{aligned}\phi(2p_{x+iy}) &\equiv Y_1^1(\theta,\phi)\tilde{R}_1^{(2)}(r) = -\frac{1}{8\sqrt{\pi}}a^{-\frac{3}{2}}\frac{r}{a}\mathrm{e}^{-\frac{r}{2a}}\sin\theta\mathrm{e}^{i\phi}\\ &= -\frac{1}{8\sqrt{\pi}}a^{-\frac{3}{2}}\frac{r}{a}\mathrm{e}^{-\frac{r}{2a}}\frac{x+iy}{r} = -\frac{1}{8\sqrt{\pi}}a^{-\frac{5}{2}}\mathrm{e}^{-\frac{r}{2a}}(x+iy).\end{aligned} \quad (3.304)$$

In (3.304), the minus sign comes from the Condon–Shortley phase. Furthermore, we have

$$\begin{aligned}\phi(2p_{x-iy}) &\equiv Y_1^{-1}(\theta,\phi)\tilde{R}_1^{(2)}(r) = \frac{1}{8\sqrt{\pi}}a^{-\frac{3}{2}}\frac{r}{a}\mathrm{e}^{-\frac{r}{2a}}\sin\theta\mathrm{e}^{-i\phi}\\ &= \frac{1}{8\sqrt{\pi}}a^{-\frac{3}{2}}\frac{r}{a}\mathrm{e}^{-\frac{r}{2a}}\frac{x-iy}{r} = \frac{1}{8\sqrt{\pi}}a^{-\frac{5}{2}}\mathrm{e}^{-\frac{r}{2a}}(x-iy).\end{aligned} \quad (3.305)$$

Notice that the above notations $\phi(2p_{x+iy})$ and $\phi(2p_{x-iy})$ differ from the custom that uses, e.g., $\phi(2p_x)$ and $\phi(2p_y)$. We will come back to this point in Sect. 4.3.

## References

1. Schiff LI (1955) Quantum mechanics, 2nd edn. McGraw-Hill, New York
2. Arfken GB (1970) Mathematical methods for physicists, 2nd edn. Academic Press, Waltham
3. Sunakawa S (1991) Quantum mechanics. Iwanami, Tokyo (in Japanese)
4. Byron FW Jr, Fuller RW (1992) Mathematics of classical and quantum physics. Dover, New York
5. Dennery P, Krzywicki A (1996) Mathematics for physicists. Dover, New York
6. Riley KF, Hobson MP, Bence SJ (2006) Mathematical methods for physics and engineering, 3rd edn. Cambridge University Press, Cambridge
7. Arfken GB, Weber HJ, Harris FE (2013) Mathematical methods for physicists, 7th edn. Academic Press, Waltham
8. Lebedev NN (1972) Special functions and their applications. Dover, New York
9. Stakgold I (1998) Green's functions and boundary value problems, 2nd edn. Wiley, New York

# Chapter 4
# Optical Transition and Selection Rules

In Sect. 1.2, we showed the Schrödinger equation as a function of space coordinates and time. In subsequent sections, we dealt with the time-independent eigenvalue problems of a harmonic oscillator and a hydrogen-like atoms. This implies that the physical system is isolated from the outside world and that there is no interaction between the outside world and physical system we are considering. However, by virtue of the interaction, the system may acquire or lose energy, momentum, angular momentum, etc. As a consequence of the interaction, the system changes its quantum state as well. Such a change is said to be a transition. If the interaction takes place as an optical process, we are to deal with an optical transition. Of various optical transitions, the electric dipole transition is common and the most important. In this chapter, we study the optical transition of a particle confined in a potential well, a harmonic oscillator, and a hydrogen using a semiclassical approach. A question of whether the transition is allowed or forbidden is of great importance. We have a selection rule to judge it.

## 4.1 Electric Dipole Transition

We have a time-dependent Schrödinger equation described as

$$H\psi = i\hbar\frac{\partial\psi}{\partial t}. \tag{1.47}$$

Using the method of separation of variables, we obtained two equations expressed below.

$$H\phi(\boldsymbol{x}) = E\phi(\boldsymbol{x}), \tag{1.55}$$



S. Hotta, *Mathematical Physical Chemistry*,
https://doi.org/10.1007/978-981-10-7671-8_4

$$i\hbar\frac{\partial\xi(t)}{\partial t} = E\xi(t). \tag{1.56}$$

Equation (1.55) is aneigenvalue equation of energy, and (1.56) is an equation with time. So far we have focused our attention upon (1.55) taking a one-dimensional harmonic oscillator and hydrogen-like atoms as an example. In this chapter, we deal with a time-evolved Schrödinger equation and its relevance to an optical transition. The optical transition takes place according to selection rules. We mention their significance as well.

We showed that after solving the eigenvalue equation, the Schrödinger equation is expressed as

$$\psi(\boldsymbol{x},t) = \phi(\boldsymbol{x})\exp(-iEt/\hbar). \tag{1.60}$$

The probability density of the system (i.e., normally a particle such as an electron, a harmonic oscillator) residing at a certain place $\boldsymbol{x}$ at a certain time $t$ is expressed as

$$\psi^*(\boldsymbol{x},t)\psi(\boldsymbol{x},t).$$

If the Schrödinger equation is described as a form of separated variables as in the case of (1.60), the exponential factors including $t$ cancel out and we have

$$\psi^*(\boldsymbol{x},t)\psi(\boldsymbol{x},t) = \phi^*(\boldsymbol{x})\phi(\boldsymbol{x}). \tag{4.1}$$

This means that the probability density of the system depends only on spatial coordinate and is constant in time. Such a state is said to be a stationary state. That is, the system continues residing in a quantum state described by $\phi(\boldsymbol{x})$ and remains unchanged independent of time.

Next, we consider a linear combination of functions described by (1.60). That is, we have

$$\psi(\boldsymbol{x},t) = c_1\phi_1(\boldsymbol{x})\exp(-iE_1t/\hbar) + c_2\phi_2(\boldsymbol{x})\exp(-iE_2t/\hbar), \tag{4.2}$$

where the first term is pertinent to the state 1 and second term to the state 2; $c_1$ and $c_2$ are complex constants with respect to the spatial coordinates but may be weakly time-dependent. The state described by (4.2) is called a *coherent state*. The probability distribution of that state is described as

$$\psi^*(\boldsymbol{x},t)\psi(\boldsymbol{x},t) = |c_1|^2|\phi_1|^2 + |c_2|^2|\phi_2|^2 + c_1^*c_2\phi_1^*\phi_2\mathrm{e}^{-i\omega t} + c_2^*c_1\phi_2^*\phi_1\mathrm{e}^{i\omega t}, \tag{4.3}$$

where $\omega$ is expressed as

$$\omega = (E_2 - E_1)/\hbar. \tag{4.4}$$

This equation shows that the probability density of the system undergoes a sinusoidal oscillation with time. The angular frequency equals the energy difference

between the two states divided by the reduced Planck constant. If the system is a charged particle such as an electron and proton, the sinusoidal oscillation is accompanied by an oscillating electromagnetic field. Thus, the coherent state is associated with the optical transition from one state to another, when the transition is related to the charged particle.

The optical transitions result from various causes. Of these, the electric dipole transition yields the largest transition probability, and the dipole approximation is often chosen to represent the transition probability. From the point of view of optical measurements, the electric dipole transition gives the strongest absorption or emission spectral lines. The matrix elementof the electric dipole, more specifically a square of an absolute value of the matrix element, is a measure of the optical transitionprobability. Labeling the quantum states as $a$, $b$, etc. and describing the corresponding state vector as $|a\rangle, |b\rangle$, etc., the matrix element $P_{ba}$ is given by

$$P_{ba} \equiv \langle b|\boldsymbol{\varepsilon}_{\mathrm{e}} \cdot \boldsymbol{P}|a\rangle, \tag{4.5}$$

where $\boldsymbol{\varepsilon}_{\mathrm{e}}$ is a unit polarization vector of the electric field of an electromagnetic wave (i.e., light). Equation (4.5) describes the optical transition that takes place as a result of the interaction between electrons and radiation field in such a way that the interaction causes electrons in the system to change the state from $|a\rangle$ to $|b\rangle$. That interaction is represented by $\boldsymbol{\varepsilon}_{\mathrm{e}} \cdot \boldsymbol{P}$. The quantum states $|a\rangle$ and $|b\rangle$ are referred to as an initial state and final state, respectively.

The quantity $\boldsymbol{P}$ is the electric dipole moment of the system, which is defined as

$$\boldsymbol{P} \equiv e \sum_j \boldsymbol{x}_j, \tag{4.6}$$

where $e$ is an elementary charge ($e < 0$) and $\boldsymbol{x}_j$ is a position vector of the $j$th electron. Detailed description of $\boldsymbol{\varepsilon}_{\mathrm{e}}$ and $\boldsymbol{P}$ can be seen in Chap. 5. The quantity $P_{ba}$ is said to be transition dipole moment, or more precisely, transition electric dipole moment with respect to the states $|a\rangle$ and $|b\rangle$. We assume that the optical transition occurs from a quantum state $|a\rangle$ to another state $|b\rangle$. Since $P_{ba}$ is generally a complex number, $|P_{ba}|^2$ represents the transition probability.

If we adopt the coordinate representation, (4.5) is expressed by

$$P_{ba} = \int \phi_b^* \boldsymbol{\varepsilon}_{\mathrm{e}} \cdot \boldsymbol{P} \phi_a \mathrm{d}\tau, \tag{4.7}$$

where $\tau$ denotes an integral range of a space.

## 4.2 One-dimensional System

Let us apply the aforementioned general description to individual cases of Chaps. 1–3.

*Example 4.1* A particle confined in a square-well potential

This example was treated in Chap. 1. As before, we assume that a particle (i.e., electron) is confined in a one-dimensional system $[-L \leq x \leq L(L > 0)]$.

We consider the optical transition from the ground state $\phi_1(x)$ to the first excited state $\phi_2(x)$. Here, we put $L = \pi/2$ for convenience. Then, the normalized coherent state $\psi(x)$ is described as

$$\psi(x,t) = \frac{1}{\sqrt{2}}[\phi_1(x)\exp(-iE_1t/\hbar) + \phi_2(x)\exp(-iE_2t/\hbar)], \tag{4.8}$$

where we put $c_1 = c_2 = \frac{1}{\sqrt{2}}$ in (4.2). In (4.8), we have

$$\phi_1(x) = \sqrt{\frac{2}{\pi}}\cos x \text{ and } \phi_2(x) = \sqrt{\frac{2}{\pi}}\sin 2x. \tag{4.9}$$

Following (4.3), we have a following real function called a probability distribution density:

$$\psi^*(x,t)\psi(x,t) = \frac{1}{\pi}\left[\cos^2 x + \sin^2 2x + (\sin 3x + \sin x)\cos \omega t\right], \tag{4.10}$$

where $\omega$ is given by (4.4) as

$$\omega = 3\hbar/2m, \tag{4.11}$$

where $m$ is a mass of an electron. Rewriting (4.10), we have

$$\psi^*(x,t)\psi(x,t) = \frac{1}{\pi}\left[1 + \frac{1}{2}(\cos 2x - \cos 4x) + (\sin 3x + \sin x)\cos \omega t\right]. \tag{4.12}$$

Integrating (4.12) over $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, a contribution from only the first term is non-vanishing to give 1, as anticipated (because of the normalization).

Putting $t = 0$ and integrating (4.12) over a positive domain $\left[0, \frac{\pi}{2}\right]$, we have

$$\int_0^{\pi/2} \psi^*(x,0)\psi(x,0)\mathrm{d}x = \frac{1}{2} + \frac{4}{3\pi} \approx 0.924. \tag{4.13}$$

Similarly, integrating (4.12) over a negative domain $\left[-\frac{\pi}{2}, 0\right]$, we have

$$\int_{-\pi/2}^{0} \psi^*(x,0)\psi(x,0)\mathrm{d}x = \frac{1}{2} - \frac{4}{3\pi} \approx 0.076. \tag{4.14}$$

Thus, 92% of a total charge (as a probability density) is concentrated in the positive domain. Differentiation of $\psi^*(x,0)\psi(x,0)$ gives five extremals including both edges. Of these, a major maximum is located at 0.635 rad that corresponds to about 40% of $\pi/2$. This can be a measure of the transition moment. Figure 4.1 demonstrates these results (see a solid curve). Meanwhile, putting $t = \pi/\omega$ (i.e., half period), we plot $\psi^*(x,\pi/\omega)\psi(x,\pi/\omega)$. The result shows that the graph is obtained by folding back the solid curve of Fig. 4.1 with respect to the ordinate axis. Thus, we find that the charge (or the probability density) exerts a sinusoidal oscillation with an angular frequency $3\hbar/2m$ along the $x$-axis around the origin.

Let $\boldsymbol{e}_1$ be a unit vector in the positive direction of the $x$-axis. Then, the electric dipole $\boldsymbol{P}$ of the system is

$$\boldsymbol{P} = e\boldsymbol{x} = ex\boldsymbol{e}_1, \tag{4.15}$$

where $\boldsymbol{x}$ is a position vector of the electron. Let us define the matrix element of the electric dipoletransition as

$$P_{21} \equiv \langle \phi_2(x)|\boldsymbol{e}_1 \cdot \boldsymbol{P}|\phi_1(x)\rangle = \langle \phi_2(x)|ex|\phi_1(x)\rangle. \tag{4.16}$$

Notice that we only have to consider that the polarization of light is parallel to the $x$-axis. With the coordinate representation, we have

$$\begin{aligned} P_{21} &= \int_{-\pi/2}^{\pi/2} \phi_2^*(x) ex \phi_1(x) \mathrm{d}x = \int_{-\pi/2}^{\pi/2} \sqrt{\frac{2}{\pi}}(\cos x) ex \sqrt{\frac{2}{\pi}} \sin 2x \, \mathrm{d}x \\ &= e\frac{2}{\pi} \int_{-\pi/2}^{\pi/2} x \cos x \sin 2x \, \mathrm{d}x = \frac{e}{\pi} \int_{-\pi/2}^{\pi/2} x(\sin x + \sin 3x) \mathrm{d}x \\ &= \frac{e}{\pi} \int_{-\pi/2}^{\pi/2} \left[ x(-\cos x)' + x\left(-\frac{1}{3}\cos 3x\right)' \right] \mathrm{d}x = \frac{16e}{9\pi}, \end{aligned} \tag{4.17}$$

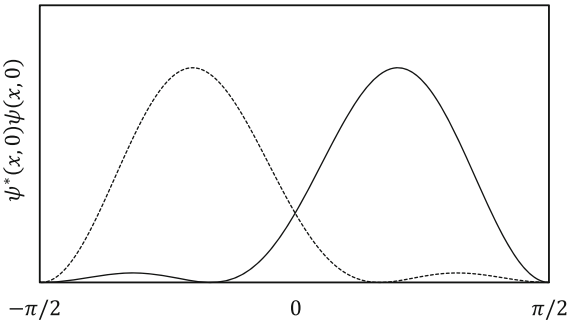


**Fig. 4.1** Probability distribution density $\psi^*(x,t)\psi(x,t)$ of a particle confined in a square-well potential. The solid curve and broken curve represent the density of $t = 0$ and $t = \pi/\omega$ (i.e., half period), respectively

where we used a trigonometric formula and integration by parts. The factor $16/9\pi$ in (4.17) is about 36% of $\pi/2$. This number is pretty good agreement with 40% that is estimated above from the major maximum of $\psi^*(x,0)\psi(x,0)$.

Note that the transition moment vanishes if the two states associated with the transition have the same parity. In other words, if these are both described by sine functions or cosine functions, the integral vanishes.

*Example 4.2* One-dimensional harmonic oscillator

Second, let us think of an optical transition regarding a harmonic oscillator that we dealt with in Chap. 2. We denote the state of the oscillator as $|n\rangle$ in place of $|\psi_n\rangle$ $(n = 0, 1, 2, \ldots)$ of Chap. 2. Then, a general expression (4.5) can be written as

$$P_{kl} = \langle k|\boldsymbol{\varepsilon}_e \cdot \boldsymbol{P}|l\rangle. \tag{4.18}$$

Since we are considering the sole one-dimensional oscillator,

$$\boldsymbol{\varepsilon}_e = \tilde{\boldsymbol{q}} \text{ and } \boldsymbol{P} = e\boldsymbol{q}, \tag{4.19}$$

where $\tilde{\boldsymbol{q}}$ is a unit vector in the positive direction of the coordinate $q$. Therefore, similarly to the above we have

$$\boldsymbol{\varepsilon}_e \cdot \boldsymbol{P} = eq. \tag{4.20}$$

That is,

$$P_{kl} = e\langle k|q|l\rangle. \tag{4.21}$$

Since $q$ is an Hermitian operator, we have

$$P_{kl}^* = e\left\langle l|q^\dagger|k\right\rangle = e\langle l|q|k\rangle = P_{lk}, \tag{4.22}$$

where we used (1. 116). Using (2. 68), we have

$$P_{kl} = e\sqrt{\frac{\hbar}{2m\omega}}\left\langle k|a + a^\dagger|l\right\rangle = \sqrt{\frac{\hbar}{2m\omega}}\left[\langle k|a|l\rangle + \left\langle k|a^\dagger|l\right\rangle\right]. \tag{4.23}$$

Taking the adjoint of (2.62) and modifying the notation, we have

$$\langle k|a = \sqrt{k+1}\langle k+1|. \tag{4.24}$$

Using (2.62) once again, we get

$$P_{kl} = e\sqrt{\frac{\hbar}{2m\omega}}\left[\sqrt{k+1}\langle k+1|l\rangle + \sqrt{l+1}\langle k|l+1\rangle\right]. \tag{4.25}$$

Using orthonormal conditions between the state vectors, we have

$$P_{kl} = e\sqrt{\frac{\hbar}{2m\omega}}\left[\sqrt{k+1}\delta_{k+1,l} + \sqrt{l+1}\delta_{k,l+1}\right]. \tag{4.26}$$

Exchanging $k$ and $l$ in the above, we get

$$P_{kl} = P_{lk}.$$

The matrix element $P_{kl}$ is symmetric with respect to indices $k$ and $l$. Notice that the first term does not vanish only when $k+1=l$. The second term does not vanish only when $k=l+1$. Therefore, we get

$$P_{k,k+1} = e\sqrt{\frac{\hbar(k+1)}{2m\omega}} \text{ and } P_{l+1,l} = e\sqrt{\frac{\hbar(l+1)}{2m\omega}}\left(\text{or } P_{k+1,k} = e\sqrt{\frac{\hbar(k+1)}{2m\omega}}\right). \tag{4.27}$$

Meanwhile, we find that the transition matrix $P$ is expressed as

$$P = eq = e\sqrt{\frac{\hbar}{2m\omega}}\left(a + a^{\dagger}\right) = e\sqrt{\frac{\hbar}{2m\omega}}\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & \cdots \\ 1 & 0 & \sqrt{2} & 0 & 0 & \cdots \\ 0 & \sqrt{2} & 0 & \sqrt{3} & 0 & \cdots \\ 0 & 0 & \sqrt{3} & 0 & 2 & \cdots \\ 0 & 0 & 0 & 2 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \tag{4.28}$$

where we used (2. 68). Note that a real Hermitian matrix is a symmetric matrix.

Practically, it is a fast way to construct a transition matrix (4.28) using (2.65) and (2.66). It is an intuitively obvious and straightforward task. Having a glance at the matrix form immediately tells us that the transition matrix elements are nonvanishing with only $(k, k+1)$ and $(k+1, k)$ positions. Whereas the $(k, k+1)$-element represents transition from the $k$th excited state to $(k-1)$-th excited state accompanied by photoemission, the $(k+1, k)$-element implies the transition from $(k-1)$-excited state to $k$th excited state accompanied by photoabsorption. The two transitions give the same transition moment. Note that zeroth excited state means the ground state; see (2.64) for basisvector representations.

We should be careful about "addresses" of the matrix accordingly. For example, $P_{0,1}$ in (4.27) represents a (1,2) element of the matrix (4.28); $P_{2,1}$ stands for a (3,2) element.

Suppose that we seek the transition dipole moments using coordinate representation. Then, we need to use (2.106) and perform definite integration. For instance, we have

$$e\int_{-\infty}^{\infty}\psi_0(q)q\psi_1(q)\mathrm{d}q$$

that corresponds to (1,2) element of (4.28). Indeed, the above integral gives $e\sqrt{\frac{\hbar}{2m\omega}}$. The confirmation is left for the readers. Nonetheless, to seek a definite integral of product of higher excited-state wave functions becomes increasingly troublesome. In this respect, the operator method described above provides us with a much better insight into complicated calculations.

Equations (4.26–4.28) imply that the electric dipole transition is allowed to occur only when the quantum number changes by one. Notice also that the transition takes place between the even function and odd function; see Table 2.1 and (2.101). Such a condition or restriction on the optical transition is called a selection rule. The former equation of (4.27) shows that the transition takes place from the upper state to the lower state accompanied by the photonemission. The latter equation, on the other hand, shows that the transition takes place from the lower state to the upper accompanied by the photonabsorption.

## 4.3 Three-dimensional System

The hydrogen-like atoms give us a typical example. Since we have fully investigated the quantum states of those atoms, we make the most of the related results.

*Example 4.3* An electron in a hydrogen atom

Unlike the one-dimensional system, we have to take account of an angular momentum in the three-dimensional system. We have already obtained explicit wave functions. Here we focus on $1s$ and $2p$ states of a hydrogen. For their normalized states, we have

$$\phi(1s)=\sqrt{\frac{1}{\pi a^3}}\mathrm{e}^{-r/a}, \tag{4.29}$$

$$\phi(2p_z)=\frac{1}{4}\sqrt{\frac{1}{2\pi a^3}}\frac{r}{a}\mathrm{e}^{-r/2a}\cos\theta, \tag{4.30}$$

$$\phi\left(2p_{x+iy}\right)=-\frac{1}{8}\sqrt{\frac{1}{\pi a^3}}\frac{r}{a}\mathrm{e}^{-r/2a}\sin\theta\mathrm{e}^{i\phi}, \tag{4.31}$$

$$\phi\left(2p_{x-iy}\right)=\frac{1}{8}\sqrt{\frac{1}{\pi a^3}}\frac{r}{a}\mathrm{e}^{-r/2a}\sin\theta\mathrm{e}^{-i\phi}, \tag{4.32}$$

where $a$denotes Bohr radius of a hydrogen. Note that a minus sign of $\phi\left(2p_{x+iy}\right)$ is due to the Condon–Shortley phase. Even though the transition probability is proportional to a square of the matrix element and so the phase factor cancels out, we describe the state vector faithfully. The energy eigenvalues are

$$E(1s) = -\frac{\hbar^2}{2\mu a^2}, E(2p_z) = E\left(2p_{x+iy}\right) = E\left(2p_{x-iy}\right) = -\frac{\hbar^2}{8\mu a^2}, \tag{4.33}$$

where $\mu$ is a reduced mass of a hydrogen. Note that the latter three states are degenerated.

First, we consider a transition between $\phi(1s)$ and $\phi(2p_z)$ states. Suppose that the normalized coherent state is described as

$$\psi(\boldsymbol{x}, t) = \frac{1}{\sqrt{2}}\{[\phi(1s)\exp[-iE(1s)t/\hbar] + \phi(2p_z)\exp[-iE(2p_z)t/\hbar]\}. \tag{4.34}$$

As before, we have

$$\begin{aligned}\psi^*(\boldsymbol{x}, t)\psi(\boldsymbol{x}, t) &= |\psi(\boldsymbol{x}, t)|^2 \\ &= \frac{1}{2}\left\{[\phi(1s)]^2 + [\phi(2p_z)]^2 + 2\phi(1s)\phi(2p_z)\cos\omega t\right\},\end{aligned} \tag{4.35}$$

where $\omega$ is given by

$$\omega = [E(2p_z) - E(1s)]/\hbar = 3\hbar/8\mu a^2. \tag{4.36}$$

In virtue of the third term that contains a $\cos\omega t$ term, the charge distribution undergoes a sinusoidal oscillation along the $z$-axis with an angular frequency described by (4.36). For instance, $\omega t = 0$ gives +1 factor to (4.35) when $t = 0$, whereas it gives $-1$ factor when $\omega t = \pi$, i.e. $t = 8\pi\mu a^2/3\hbar$.

Integrating (4.35), we have

$$\begin{aligned}\int\psi^*(\boldsymbol{x}, t)\psi(\boldsymbol{x}, t)\mathrm{d}\tau &= \int|\psi(\boldsymbol{x}, t)|^2\mathrm{d}\tau \\ &= \frac{1}{2}\int\left\{[\phi(1s)]^2 + [\phi(2p_z)]^2\right\}\mathrm{d}\tau, \\ &\quad + \cos\omega t\int\phi(1\mathrm{s})\phi(2p_z)\mathrm{d}\tau = \frac{1}{2} + \frac{1}{2} = 1\end{aligned}$$

where we used normalized functional forms of $\phi(1s)$ and $\phi(2p_z)$ together with orthogonality of them. Note that both of the functions are real.

Next, we calculate the matrix element. For simplicity, we denote the matrix element simply as $P^{(\boldsymbol{\varepsilon}_e)}$ only by designating the unit polarization vector $\boldsymbol{\varepsilon}_e$. Then, we have

$$P^{(\varepsilon_e)} = \langle \phi(1s)|\varepsilon_e \cdot \boldsymbol{P}|\phi(2p_z)\rangle, \tag{4.37}$$

where

$$\boldsymbol{P} = e\boldsymbol{x} = e(\boldsymbol{e}_1\boldsymbol{e}_2\boldsymbol{e}_3)\begin{pmatrix} x \\ y \\ z \end{pmatrix}. \tag{4.38}$$

We have three possibilities of choosing $\varepsilon_e$ out of $\boldsymbol{e}_1, \boldsymbol{e}_2,$ and $\boldsymbol{e}_3$. Choosing $\boldsymbol{e}_3$, we have

$$\begin{aligned} P^{(\boldsymbol{e}_3)}_{z,|p_z\rangle} &= e\langle \phi(1s)|z|\phi(2p_z)\rangle \\ &= \frac{e}{4\sqrt{2}\pi a^4}\int_0^\infty r^4 e^{-3r/2a}dr \int_0^\pi \cos^2\theta \sin\theta \, d\theta \int_0^{2\pi} d\phi = \frac{2^7\sqrt{2}}{3^5} ea \approx 0.745ea. \end{aligned} \tag{4.39}$$

In (4.39), we express the matrix element as $P^{(\boldsymbol{e}_3)}_{z,|p_z\rangle}$ to indicate the $z$-component of position vector and to explicitly show that $\phi(2p_z)$ state is responsible for the transition. In (4.39), we used $z = r\cos\theta$. We also used a radial part integration such that

$$\int_0^\infty r^4 e^{-3r/2a}dr = 24\left(\frac{2a}{3}\right)^5.$$

Also we changed a variable $\cos\theta \rightarrow t$ to perform the integration with respect to $\theta$. We see that a "leverage" length of the transition moment is comparable to Bohr radius $a$.

With the notation $P^{(\boldsymbol{e}_3)}_{z,|p_z\rangle}$ we need some explanation for consistency with the latter description. Equation (4.39) represents the transition from $|\phi(2p_z)\rangle$ to $|\phi(1s)\rangle$ that is accompanied by the photonemission. Thus, $|p_z\rangle$ in the notation, means that $|\phi(2p_z)\rangle$ is the initial state. In the notation, in turn, $(\boldsymbol{e}_3)$ denotes the polarization vector and $z$ represents the electric dipole.

In the case of photonabsorption where the transition occurs from $|\phi(1s)\rangle$ to $|\phi(2p_z)\rangle$, we use the following notation:

$$P^{(\boldsymbol{e}_3)}_{z,\langle p_z|} = e\langle \phi(2p_z)|z|\phi(1s)\rangle. \tag{4.40}$$

Since all the functions related to the integration are real, we have

$$P^{(\boldsymbol{e}_3)}_{z,|p_z\rangle} = P^{(\boldsymbol{e}_3)}_{z,\langle p_z|}.$$

Meanwhile, if we choose $\boldsymbol{e}_1$ for $\boldsymbol{\varepsilon}_e$ to evaluate the matrix element $P_x$, we have

$$\begin{aligned} P^{(\boldsymbol{e}_1)}_{x,|p_z\rangle} &= e\langle\phi(1s)|x|\phi(2p_z)\rangle \\ &= \frac{e}{4\sqrt{2}\pi a^4}\int_0^\infty r^4 e^{-3r/2a}\mathrm{d}r \int_0^\pi \sin^2\theta\,\cos\theta\,\mathrm{d}\theta \int_0^{2\pi}\cos\phi\,\mathrm{d}\phi = 0, \end{aligned} \tag{4.41}$$

where cos $\phi$ comes from $x = r$ sin $\theta$ cos $\phi$ and an integration of cos $\phi$ gives zero. In a similar manner, we have

$$P^{(\boldsymbol{e}_2)}_{y,|p_z\rangle} = e\langle\phi(1s)|y|\phi(2p_z)\rangle = 0. \tag{4.42}$$

Next, we estimate the matrix elements associated with $2p_x$ and $2p_y$. For this purpose, it is convenient to introduce the following complex coordinates by a unitary transformation:

$$\begin{aligned} (\boldsymbol{e}_1\boldsymbol{e}_2\boldsymbol{e}_3)\begin{pmatrix} x \\ y \\ z \end{pmatrix} &= (\boldsymbol{e}_1\boldsymbol{e}_2\boldsymbol{e}_3)\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{i}{\sqrt{2}} & \frac{i}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & -\frac{i}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} \\ &= \left(\frac{1}{\sqrt{2}}(\boldsymbol{e}_1 - i\boldsymbol{e}_2)\,\frac{1}{\sqrt{2}}(\boldsymbol{e}_1 + i\boldsymbol{e}_2)\boldsymbol{e}_3\right)\begin{pmatrix} \frac{1}{\sqrt{2}}(x+iy) \\ \frac{1}{\sqrt{2}}(x-iy) \\ z \end{pmatrix}, \end{aligned} \tag{4.43}$$

where a unitary transformation is represented by a unitary matrix defined as

$$U^\dagger U = UU^\dagger = E. \tag{4.44}$$

We will investigate details of the unitary transformation and matrix in Parts III and IV.

We define $\boldsymbol{e}_+$ and $\boldsymbol{e}_-$ as follows [1]:

$$\boldsymbol{e}_+ \equiv \frac{1}{\sqrt{2}}(\boldsymbol{e}_1 + i\boldsymbol{e}_2) \text{ and } \boldsymbol{e}_- \equiv \frac{1}{\sqrt{2}}(\boldsymbol{e}_1 - i\boldsymbol{e}_2), \tag{4.45}$$

where complex vectors $\boldsymbol{e}_+$ and $\boldsymbol{e}_-$ represent the left-circularly polarized lightand right-circularly polarized light that carry an angular momentum $\hbar$ and $-\hbar$, respectively. We will revisit the characteristics and implication of these complex vectors in Sect. 5.4.

We have

$$(\boldsymbol{e}_1\boldsymbol{e}_2\boldsymbol{e}_3)\begin{pmatrix} x \\ y \\ z \end{pmatrix} = (\boldsymbol{e}_- \boldsymbol{e}_+ \boldsymbol{e}_3)\begin{pmatrix} \frac{1}{\sqrt{2}}(x+iy) \\ \frac{1}{\sqrt{2}}(x-iy) \\ z \end{pmatrix}. \tag{4.46}$$

Note that $\boldsymbol{e}_+, \boldsymbol{e}_-,$ and $\boldsymbol{e}_3$ are orthonormal. That is,

$$\langle \boldsymbol{e}_+ | \boldsymbol{e}_+ \rangle = 1, \langle \boldsymbol{e}_+ | \boldsymbol{e}_- \rangle = 0, \text{etc.} \tag{4.47}$$

In this situation, $\boldsymbol{e}_+, \boldsymbol{e}_-,$ and $\boldsymbol{e}_3$ are said to form an orthonormal basis in a three-dimensional complex vector space (see Sect. 11.4).

Now, choosing $\boldsymbol{e}_+$ for $\boldsymbol{\varepsilon}_e$, we have [2]

$$P^{(\boldsymbol{e}_+)}_{x-iy,|p_+\rangle} \equiv e\left\langle \phi(1s)|\frac{1}{\sqrt{2}}(x-iy)|\phi\left(2p_{x+iy}\right)\right\rangle, \tag{4.48}$$

where $|p_+\rangle$ is a shorthand notation of $\phi\left(2p_{x+iy}\right)$; $x - iy$ represents a complex electric dipole. Equation (4.48) represents an optical process in which an electron causes transition from $\phi\left(2p_{x+iy}\right)$ to $\phi(1s)$ to lose an angular momentum $\hbar$, whereas the radiation field gains that angular momentum to conserve a total angular momentum $\hbar$. The notation $P^{(\boldsymbol{e}_+)}_{x-iy,|p_+\rangle}$ reflects this situation. Using the coordinate representation, we rewrite (4.48) as

$$P^{(\boldsymbol{e}_+)}_{x-iy,|p_+\rangle} = -\frac{e}{8\sqrt{2}\pi a^4}\int_0^\infty r^4 \mathrm{e}^{-3r/2a}\mathrm{d}r \int_0^\pi \sin^3\theta\,\mathrm{d}\theta \int_0^{2\pi} \mathrm{e}^{-i\phi}\mathrm{e}^{i\phi}\mathrm{d}\phi = -\frac{2^7\sqrt{2}}{3^5}ea, \tag{4.49}$$

where we used

$$x - iy = r\,\sin\,\theta\mathrm{e}^{-i\phi}. \tag{4.50}$$

In the definite integral of (4.49), $\mathrm{e}^{-i\phi}$ comes from $x - iy$, while $\mathrm{e}^{i\phi}$ comes from $\phi\left(2p_{x+iy}\right)$. Note that from (3.24) $\mathrm{e}^{i\phi}$ is an eigenfunction corresponding to an angular momentumeigenvalue $\hbar$. Notice that in (4.49), exponents $\mathrm{e}^{-i\phi}$ and $\mathrm{e}^{i\phi}$ cancel out and that an azimuthal integral is nonvanishing.

If we choose $\boldsymbol{e}_-$ for $\boldsymbol{\varepsilon}_e$, we have

$$P^{(\boldsymbol{e}_-)}_{x+iy,|p_+\rangle} = e\left\langle \phi(1s)|\frac{1}{\sqrt{2}}(x+iy)|\phi\left(2p_{x+iy}\right)\right\rangle \tag{4.51}$$

$$= \frac{e}{8\sqrt{2}\pi a^4} \int_0^\infty r^4 \mathrm{e}^{-3r/2a} \mathrm{d}r \int_0^\pi \sin^3\theta \, \mathrm{d}\theta \int_0^{2\pi} \mathrm{e}^{2i\phi} \mathrm{d}\phi = 0, \tag{4.52}$$

where we used

$$x + iy = r \sin\theta \mathrm{e}^{i\phi}.$$

With (4.52), a factor $\mathrm{e}^{2i\phi}$ results from the product $\phi\left(2p_{x+iy}\right)(x+iy)$ which renders the integral (4.51) vanishing. Note that the only difference between (4.49) and (4.52) is about the integration of $\phi$ factor. For the same reason, if we choose $\boldsymbol{e}_3$ for $\boldsymbol{\varepsilon}_\mathrm{e}$, the matrix element vanishes. Thus, with the $\phi\left(2p_{x+iy}\right)$-related matrix element, only $P^{(\boldsymbol{e}_+)}_{x-iy,|p_+\rangle}$ survives. Similarly, with the $\phi\left(2p_{x-iy}\right)$-related matrix element, only $P^{(\boldsymbol{e}_-)}_{x+iy,|p_-\rangle}$ survives. Notice that $|p_-\rangle$ is a shorthand notation of $\phi\left(2p_{x-iy}\right)$. That is, we have, e.g.,

$$\begin{aligned} P^{(\boldsymbol{e}_-)}_{x+iy,|p_-\rangle} &= e\left\langle \phi(1s)|\frac{1}{\sqrt{2}}(x+iy)|\phi\left(2p_{x-iy}\right)\right\rangle = \frac{2^7\sqrt{2}}{3^5}ea, \\ P^{(\boldsymbol{e}_+)}_{x-iy,|p_-\rangle} &= e\left\langle \phi(1s)|\frac{1}{\sqrt{2}}(x-iy)|\phi\left(2p_{x-iy}\right)\right\rangle = 0. \end{aligned} \tag{4.53}$$

Taking complex conjugate of (4.48), we have

$$\left[P^{(\boldsymbol{e}_+)}_{x-iy,|p_+\rangle}\right]^* = e\left\langle \phi\left(2p_{x+iy}\right)|\frac{1}{\sqrt{2}}(x+iy)|\phi(1s)\right\rangle = -\frac{2^7\sqrt{2}}{3^5}ea \tag{4.54}$$

Here recall (1.116) and $(x-iy)^\dagger = x+iy$. Also note that since $P^{(\boldsymbol{e}_+)}_{x-iy,|p_+\rangle}$ is real, $\left[P^{(\boldsymbol{e}_+)}_{x-iy,|p_+\rangle}\right]^*$ is real as well so that we have

$$\left[P^{(\boldsymbol{e}_+)}_{x-iy,|p_+\rangle}\right]^* = P^{(\boldsymbol{e}_+)}_{x-iy,|p_+\rangle} = P^{(\boldsymbol{e}_-)}_{x+iy,\langle p_+|}. \tag{4.55}$$

Comparing (4.48) and (4.55), we notice that the polarization vector has been switched from $\boldsymbol{e}_+$ to $\boldsymbol{e}_-$ with the allowed transition, even though the matrix element remains the same. This can be explained as follows: In (4.48), the photonemission is occurring, while the electron is causing a transition from $\phi\left(2p_{x+iy}\right)$ to $\phi(1s)$. As a result, the radiation field has gained an angular momentum by $\hbar$ during the process in which the electron has lost an angular momentum $\hbar$. In other words, $\hbar$ is transferred from the electron to the radiation field and this process results in the generation of left-circularly polarized light in the radiation field.

In (4.54), on the other hand, the reversed process takes place. That is, the photonabsorption is occurring in such a way that the electron is excited from $\phi(1s)$

to $\phi(2p_{x+iy})$. After this process has been completed, the electron has gained an angular momentum by $\hbar$, whereas the radiation field has lost an angular momentum by $\hbar$. As a result, the positive angular momentum $\hbar$ is transferred to the electron from the radiation field that involves left-circularly polarized light. This can be translated into the statement that the radiation field has gained an angular momentum by $-\hbar$. This is equivalent to the generation of right-circularly polarized light (characterized by $\boldsymbol{e}_-$) in the radiation field. In other words, the electron gains the angular momentum by $\hbar$ to compensate the change in the radiation field.

The implication of the first equation of (4.53) can be interpreted in a similar manner. Also we have

$$\left[P^{(\boldsymbol{e}_-)}_{x+iy,|p_-\rangle}\right]^* = P^{(\boldsymbol{e}_-)}_{x+iy,|p_-\rangle} = P^{(\boldsymbol{e}_+)}_{x-iy,\langle p_-|} = e\left\langle \phi(2p_{x-iy})|\frac{1}{\sqrt{2}}(x-iy)|\phi(1s)\right\rangle = \frac{2^7\sqrt{2}}{3^5}ea.$$

Notice that the inner products of (4.49) and (4.53) are real, even though operators $x+iy$ and $x-iy$ are not Hermitian. Also note that $P^{(\boldsymbol{e}_+)}_{x-iy,|p_+\rangle}$ of (4.49) and $P^{(\boldsymbol{e}_-)}_{x+iy,|p_-\rangle}$ of (4.53) have the same absolute value with minus and plus signs, respectively. The minus sign of (4.49) comes from the Condon–Shortley phase. The difference, however, is not essential, because the transition probability is proportional to $|P^{(\boldsymbol{e}_-)}_{x+iy,|p_-\rangle}|^2$ or $|P^{(\boldsymbol{e}_+)}_{x-iy,|p_+\rangle}|^2$. Some literature [3, 4] uses $-(x+iy)$ instead of $x+iy$. This is because simply of the inclusion of the Condon–Shortley phase; see (3.304).

Let us think of the coherent state that is composed of $\phi(1s)$ and $\phi(2p_{x+iy})$ or $\phi(2p_{x-iy})$. Choosing $\phi(2p_{x+iy})$, the state $\psi(\boldsymbol{x},t)$ can be given by

$$\psi(\boldsymbol{x},t) = \frac{1}{\sqrt{2}}\left[\phi(1s)\exp(-iE(1s)t/\hbar) + \phi(2p_{x+iy})\exp\left(-iE(2p_{x+iy})t/\hbar\right)\right], \tag{4.56}$$

where $\phi(1s)$ is described by (3.301) and $\phi(2p_{x+iy})$ is expressed as (3.304). Then we have

$$\begin{aligned}\psi^*(\boldsymbol{x},t)\psi(\boldsymbol{x},t) &= |\psi(\boldsymbol{x},t)|^2 \\ &= \frac{1}{2}\left\{|\phi(1s)|^2 + \left|\phi(2p_{x+iy})\right|^2 + \phi(1s)\Re(2p_{x+iy})\left[\mathrm{e}^{i(\phi-\omega t)} + \mathrm{e}^{-i(\phi-\omega t)}\right]\right\} \\ &= \frac{1}{2}\left\{|\phi(1s)|^2 + \left|\phi(2p_{x+iy})\right|^2 + 2\phi(1s)\Re(2p_{x+iy})\cos(\phi-\omega t)\right\},\end{aligned} \tag{4.57}$$

where using $\Re(2p_{x+iy})$, we denote $\phi(2p_{x+iy})$ as follows:

$$\phi\left(2p_{x+iy}\right) \equiv \Re\left(2p_{x+iy}\right)\mathrm{e}^{i\phi}. \tag{4.58}$$

That is, $\Re\left(2p_{x+iy}\right)$ represents a real component of $\phi\left(2p_{x+iy}\right)$ that depends only on $r$ and $\theta$. The third term of (4.57) implies that the existence probability density of an electron represented by $|\psi(\boldsymbol{x},t)|^2$ is rotating counterclockwise around the $z$-axis with an angular frequency of $\omega$. Similarly, in the case of $\phi\left(2p_{x-iy}\right)$, the existence probability density of an electron is rotating clockwise around the $z$-axis with an angular frequency of $\omega$.

Integrating (4.57), we have

$$\begin{aligned}\int \psi^*(\boldsymbol{x},t)\psi(\boldsymbol{x},t)\mathrm{d}\tau &= \int |\psi(\boldsymbol{x},t)|^2\mathrm{d}\tau \\ &= \frac{1}{2}\int \left\{\left|\phi(1s)\right|^2 + \left|\phi\left(2p_{x+iy}\right)\right|^2\right\}\mathrm{d}\tau \\ &\quad + \int_0^\infty r^2\mathrm{d}r \int_0^\pi \sin\theta\, d\theta\, \phi(1s)\Re\left(2p_{x+iy}\right) \\ &\quad \int_0^{2\pi} d\phi\, \cos(\phi-\omega t) = \frac{1}{2}+\frac{1}{2} = 1,\end{aligned}$$

where we used normalized functional forms of $\phi(1s)$ and $\phi\left(2p_{x+iy}\right)$; the last term vanishes because

$$\int_0^{2\pi} \mathrm{d}\phi\, \cos(\phi-\omega t) = 0.$$

This is easily shown by suitable variable transformation.

In relation to the above discussion, we often use real numbers to describe wave functions. For this purpose, we use the following unitary transformation to transform the orthonormal basis of $\mathrm{e}^{\pm im\phi}$ to $\cos m\phi$ and $\sin m\phi$. That is, we have

$$\left(\frac{1}{\sqrt{\pi}}\cos m\phi \;\; \frac{1}{\sqrt{\pi}}\sin m\phi\right) = \left(\frac{(-1)^m}{\sqrt{2\pi}}\mathrm{e}^{im\phi} \;\; \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-im\phi}\right)\begin{pmatrix}\frac{(-1)^m}{\sqrt{2}} & -\frac{(-1)^m i}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}}\end{pmatrix}, \tag{4.59}$$

where we assume that $m$ is positive so that we can appropriately take into account the Condon–Shortley phase. Alternatively, we describe it via unitary transformation as follows:

$$\left(\frac{(-1)^m}{\sqrt{2\pi}}\mathrm{e}^{im\phi}\ \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-im\phi}\right)=\left(\frac{1}{\sqrt{\pi}}\cos\ m\phi\ \frac{1}{\sqrt{\pi}}\sin\ m\phi\right)\begin{pmatrix}\frac{(-1)^m}{\sqrt{2}} & \frac{1}{\sqrt{2}}\\ \frac{(-1)^m i}{\sqrt{2}} & -\frac{i}{\sqrt{2}}\end{pmatrix},\quad(4.60)$$

In this regard, we have to be careful about normalization constants; for trigonometric functions, the constant should be $\frac{1}{\sqrt{\pi}}$, whereas for the exponential representation, the constant is $\frac{1}{\sqrt{2\pi}}$. At the same time, trigonometric functions are expressed as a linear combination of $\mathrm{e}^{im\phi}$ and $\mathrm{e}^{-im\phi}$, and so if we use the trigonometric functions, information of a magnetic quantum number is lost.

In Sect. 3.7, we showed normalized functions $\tilde{\Lambda}^{(n)}_{l,m}=Y^m_l(\theta,\phi)\tilde{R}^{(n)}_l(r)$ of the hydrogen-like atom. Noting that $Y^m_l(\theta,\phi)$ is proportional to $\mathrm{e}^{\pm im\phi}$, $\tilde{\Lambda}^{(n)}_{l,m}$ can be described using $\cos m\phi$ and $\sin m\phi$ for the basis vectors. We denote two linearly independent vectors by $\tilde{\Lambda}^{(n)}_{l,\cos m\phi}$ and $\tilde{\Lambda}^{(n)}_{l,\sin m\phi}$. Then, these vectors are expressed as

$$\left(\tilde{\Lambda}^{(n)}_{l,\cos m\phi}\tilde{\Lambda}^{(n)}_{l,\sin m\phi}\right)=\left(\tilde{\Lambda}^{(n)}_{l,m}\tilde{\Lambda}^{(n)}_{l,-m}\right)\begin{pmatrix}\frac{(-1)^m}{\sqrt{2}} & -\frac{(-1)^m i}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}}\end{pmatrix},\quad(4.61)$$

where we again assume that $m$ is positive. In chemistry and materials science, we normally use real functions of $\tilde{\Lambda}^{(n)}_{l,\cos m\phi}$ and $\tilde{\Lambda}^{(n)}_{l,\sin m\phi}$. In particular, we use the notations of, e.g., $\phi(2p_x)$ and $\phi(2p_y)$ instead of $\tilde{\Lambda}^{(2)}_{1,\cos\phi}$ and $\tilde{\Lambda}^{(2)}_{1,\sin\phi}$, respectively. In that case, we explicitly have a following form:

$$\begin{aligned}\left(\phi(2p_x)\phi(2p_y)\right)&=\left(\tilde{\Lambda}^{(2)}_{1,1}\tilde{\Lambda}^{(2)}_{1,-1}\right)\begin{pmatrix}-\frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}}\end{pmatrix}\\&=\left(\phi\left(2p_{x+iy}\right)\phi\left(2p_{x-iy}\right)\right)\begin{pmatrix}-\frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}}\end{pmatrix}\\&=\left(\frac{1}{4\sqrt{2\pi}}a^{-\frac{3}{2}}\frac{r}{a}\mathrm{e}^{-\frac{r}{2a}}\sin\ \theta\ \cos\ \phi\quad \frac{1}{4\sqrt{2\pi}}a^{-\frac{3}{2}}\frac{r}{a}\mathrm{e}^{-\frac{r}{2a}}\sin\ \theta\ \sin\ \phi\right).\end{aligned}\quad(4.62)$$

Thus, the Condon–Shortley phasefactor has been removed.

Using this expression, we calculate matrix elements of the electric dipole transition. We have

$$\begin{aligned}P^{(\boldsymbol{e}_1)}_{x,|p_x\rangle}&=e\langle\phi(1s)|x|\phi(2p_x)\rangle\\&=\frac{e}{4\sqrt{2}\pi a^4}\int_0^\infty r^4\mathrm{e}^{-3r/2a}\mathrm{d}r\int_0^\pi\sin^3\theta\,\mathrm{d}\theta\int_0^{2\pi}\cos^2\phi\,\mathrm{d}\phi=\frac{2^7\sqrt{2}}{3^5}ea.\end{aligned}\quad(4.63)$$

Thus, we obtained the same result as (4.49) apart from the minus sign. Since a square of an absolute value of the transition moment plays a role, the minus sign is again of secondary importance. With $P_{y}^{(\boldsymbol{e}_2)}$, similarly we have

$$\begin{aligned} P_{y,|p_y\rangle}^{(\boldsymbol{e}_2)} &= e\langle\phi(1s)|y|\phi(2p_y)\rangle \\ &= \frac{e}{4\sqrt{2}\pi a^4}\int_0^\infty r^4 \mathrm{e}^{-3r/2a}\mathrm{d}r \int_0^\pi \sin^3\theta\,\mathrm{d}\theta \int_0^{2\pi} \sin^2\phi\,\mathrm{d}\phi = \frac{2^7\sqrt{2}}{3^5}ea. \end{aligned} \tag{4.64}$$

Comparing (4.39), (4.63), and (4.64), we have

$$P_{z,|p_z\rangle}^{(\boldsymbol{e}_3)} = P_{x,|p_x\rangle}^{(\boldsymbol{e}_1)} = P_{y,|p_y\rangle}^{(\boldsymbol{e}_2)} = \frac{2^7\sqrt{2}}{3^5}ea.$$

In the case of $P_{z,|p_z\rangle}^{(\boldsymbol{e}_3)}, P_{x,|p_x\rangle}^{(\boldsymbol{e}_1)}$, and $P_{y,|p_y\rangle}^{(\boldsymbol{e}_2)}$, the optical transition is said to be polarized along the $z$-, $x$-, and $y$-axes, respectively, and so linearly polarized lights are relevant. Note moreover that operators $z$, $x$, and $y$ in (4.39), (4.63), and (4.64) are Hermitian and that $\phi(2p_z)$, $\phi(2p_x)$, and $\phi(2p_y)$ are real functions.

## 4.4 Selection Rules

In a three-dimensional system such as hydrogen-like atoms, quantum states of particles (i.e., electrons) are characterized by three quantum numbers; principal quantum numbers, azimuthal quantum numbers (or orbital angular momentum quantum numbers), and magnetic quantum numbers. In this section, we examine the *selection rules* for the electric dipole approximation.

Of the three quantum numbers mentioned above, angular momentum quantum numbers are denoted by $l$ and magnetic quantum numbers by $m$. First, we examine the conditions on $m$. With the angular momentum operator $\boldsymbol{L}$ and its corresponding operator $\boldsymbol{M}$, we get the following commutation relations:

$$\begin{aligned} &[M_z, x] = iy,\ [M_y, z] = ix,\ [M_x, y], = iz; \\ &[M_z, iy] = x,\ [M_y, ix] = z,\ [M_x, iz] = y; \text{etc.} \end{aligned} \tag{4.65}$$

Notice that in the upper line, the indices change cyclic like $(z, x, y)$, whereas in the lower line they change anticyclic such as $(z, y, x)$. The proof of (4.65) is left for the reader. Thus, we have, e.g.,

$$[M_z, x+iy] = x+iy,\ [M_z, x-iy] = -(x-iy), \text{etc.} \tag{4.66}$$

Putting

$$Q^+ \equiv x + iy \text{ and } Q^- \equiv x - iy,$$

we have

$$[M_z, Q^+] = Q^+, [M_z, Q^-] = -Q^-. \tag{4.67}$$

Taking an inner product of both sides of (4.67), we have

$$\begin{aligned}\langle m'|[M_z, Q^+]|m\rangle &= \langle m'|M_z Q^+ - Q^+ M_z|m\rangle = m'\langle m'|Q^+|m\rangle - m\langle m'|Q^+|m\rangle \\ &= \langle m'|Q^+|m\rangle,\end{aligned} \tag{4.68}$$

where the quantum state $|m\rangle$ is identical to $|l, m\rangle$ in (3.151). Here we need no information about $l$, and so it is omitted. Thus, we have, e.g., $M_z|m\rangle = m|m\rangle$. Taking its adjoint, we have $\langle m|M_z^\dagger = \langle m|M_z = m\langle m|$, where $M_z$ is Hermitian. These results lead to (4.68). From (4.68), we get

$$(m' - m - 1)\langle m'|Q^+|m\rangle = 0. \tag{4.69}$$

Therefore, for the matrix element $\langle m'|Q^+|m\rangle$ not to vanish, we must have

$$m' - m - 1 = 0 \text{ or } \Delta m = 1 (\Delta m \equiv m' - m).$$

This represents the selection rule with respect to the coordinate $Q^+$.
Similarly, we get

$$(m' - m + 1)\langle m'|Q^-|m\rangle = 0. \tag{4.70}$$

In this case, for the matrix element $\langle m'|Q^-|m\rangle$ not to vanish, we have

$$m' - m + 1 = 0 \text{ or } \Delta m = -1.$$

To derive (4.70), we can alternatively use the following: Taking the adjoint of (4.69), we have

$$(m' - m - 1)\langle m|Q^-|m'\rangle = 0.$$

Exchanging $m'$ and $m$, we have

$$(m - m' - 1)\langle m'|Q^-|m\rangle = 0 \text{ or } (m' - m + 1)\langle m'|Q^-|m\rangle = 0.$$

Thus, (4.70) is recovered.
Meanwhile, we have a commutation relation

$$[M_z, z] = 0. \tag{4.71}$$

Similarly, taking an inner product of both sides of (4.71), we have

$$(m' - m)\langle m'|z|m\rangle = 0.$$

Therefore, for the matrix element $\langle m'|z|m\rangle$ not to vanish, we must have

$$m' - m = 0 \text{ or } \Delta m = 0. \tag{4.72}$$

These results are fully consistent with Example 4.3 of Sect. 4.3. That is, if circularly polarized light takes part in the optical transition, $\Delta m = \pm 1$. For instance, using the present notation, we rewrite (4.48) as

$$\left\langle \phi(1s)|\frac{1}{\sqrt{2}}(x - iy)|\phi\left(2p_{x+iy}\right)\right\rangle = \frac{1}{\sqrt{2}}\langle 0|Q^-|1\rangle = -\frac{2^7\sqrt{2}}{3^5}a.$$

If linearly polarized light is related to the optical transition, we have $\Delta m = 0$.

Next, we examine the conditions on $l$. To this end, we calculate a following commutator [5]:

$$\begin{aligned}
\left[\boldsymbol{M}^2, z\right] &= \left[M_x^2 + M_y^2 + M_z^2, z\right] = \left[M_x^2, z\right] + \left[M_y^2, z\right] \\
&= M_x(M_x z - zM_x) + M_x z M_x - zM_x^2 \\
&\quad + M_y\left(M_y z - zM_y\right) + M_y z M_y - zM_y^2 \\
&= M_x[M_x, z] + [M_x, z]M_x + M_y\left[M_y, z\right] + \left[M_y, z\right]M_y \\
&= i\left(M_y x + xM_y - M_x y - yM_x\right) \\
&= i\left(M_x y - yM_x - M_y x + xM_y + 2M_y x - 2M_x y\right) \\
&= i\left(2iz + 2M_y x - 2M_x y\right) = 2i\left(M_y x - M_x y + iz\right).
\end{aligned} \tag{4.73}$$

In the above calculations, (i) we used $[M_z, z] = 0$ (with the second equality); (ii) RHS was modified so that the commutation relations can be used (the third equality); (iii) we used $-M_x y = M_x y - 2M_x y$ and $M_y x = -M_y x + 2M_y x$ so that we can use (4.65) (the second last equality). Moreover, using (4.65), (4.73) can be written as

$$\left[\boldsymbol{M}^2, z\right] = 2i\left(xM_y - M_x y\right) = 2i\left(M_y x - yM_x\right).$$

Similar results on the commutator can be obtained with $\left[\boldsymbol{M}^2, x\right]$ and $\left[\boldsymbol{M}^2, y\right]$. For further use, we give alternative relations such that

$$\begin{aligned}\left[\boldsymbol{M}^2, x\right] &= 2i(yM_z - M_y z) = 2i\left(M_z y - zM_y\right),\\ \left[\boldsymbol{M}^2, y\right] &= 2i(zM_x - M_z x) = 2i(M_x z - xM_z).\end{aligned} \tag{4.74}$$

Using (4.73), we calculate another commutator such that

$$\begin{aligned}\left[\boldsymbol{M}^2, \left[\boldsymbol{M}^2, z\right]\right] &= 2i\left\{\left[\boldsymbol{M}^2, M_y x\right] - \left[\boldsymbol{M}^2, M_x y\right] + i\left[\boldsymbol{M}^2, z\right]\right\}\\ &= 2i\left\{M_y\left[\boldsymbol{M}^2, x\right] - M_x\left[\boldsymbol{M}^2, y\right] + i\left[\boldsymbol{M}^2, z\right]\right\}\\ &= 2i\left\{2iM_y\left(yM_z - M_y z\right) - 2iM_x(M_x z - xM_z) + i\left[\boldsymbol{M}^2, z\right]\right\}\\ &= -2\left\{2\left(M_x x + M_y y + M_z z\right)M_z - 2\left(M_x^2 + M_y^2 + M_z^2\right)z + \boldsymbol{M}^2 z - z\boldsymbol{M}^2\right\}\\ &= 2\left(\boldsymbol{M}^2 z + z\boldsymbol{M}^2\right).\end{aligned} \tag{4.75}$$

In the above calculations, (i) we used $\left[\boldsymbol{M}^2, M_y\right] = \left[\boldsymbol{M}^2, M_x\right] = 0$ (with the second equality); (ii) we used (4.74) (the third equality); (iii) RHS was modified so that we can use the relation $\boldsymbol{M} \perp \boldsymbol{x}$ from the definition of the angular momentum operator, i.e., $M_x x + M_y y + M_z z = 0$ (the second last equality). We used $[M_z, z] = 0$ as well. Similar results are obtained with $x$ and $y$. That is, we have

$$\left[\boldsymbol{M}^2, \left[\boldsymbol{M}^2, x\right]\right] = 2\left(\boldsymbol{M}^2 x + x\boldsymbol{M}^2\right), \tag{4.76}$$

$$\left[\boldsymbol{M}^2, \left[\boldsymbol{M}^2, y\right]\right] = 2\left(\boldsymbol{M}^2 y + y\boldsymbol{M}^2\right). \tag{4.77}$$

Rewriting, e.g., (4.75), we have

$$\boldsymbol{M}^4 z - 2\boldsymbol{M}^2 z\boldsymbol{M}^2 + z\boldsymbol{M}^4 = 2\left(\boldsymbol{M}^2 z + z\boldsymbol{M}^2\right). \tag{4.78}$$

Using the relation (4.78) and taking inner products of both sides, we get, e.g.,

$$\left\langle l'\middle|\boldsymbol{M}^4 z - 2\boldsymbol{M}^2 z\boldsymbol{M}^2 + z\boldsymbol{M}^4\middle|l\right\rangle = \left\langle l'\middle|2\left(\boldsymbol{M}^2 z + z\boldsymbol{M}^2\right)\middle|l\right\rangle. \tag{4.79}$$

That is,

$$\left\langle l'\middle|\boldsymbol{M}^4 z - 2\boldsymbol{M}^2 z\boldsymbol{M}^2 + z\boldsymbol{M}^4\middle|l\right\rangle - \left\langle l'\middle|2\left(\boldsymbol{M}^2 z + z\boldsymbol{M}^2\right)\middle|l\right\rangle = 0.$$

Considering that both terms of LHS contain a factor $\langle l'|z|l\rangle$ in common, we have

$$\begin{aligned}&\left[l'^2(l'+1)^2 - 2l'l(l'+1)(l+1) + l^2(l+1)^2 - 2l'(l'+1) - 2l(l+1)\right]\\ &\qquad \times \langle l'|z|l\rangle = 0,\end{aligned} \tag{4.80}$$

where the quantum state $|l\rangle$ is identical to $|l, m\rangle$ in (3.151) with $m$ omitted.

To factorize the first factor of LHS of (4.80), we view it as a quartic equation with respect to $l'$. Replacing $l'$ with $-l$, we find that the first factor vanishes, and so the first factor should have a factor $(l'+l)$. Then, we factorize the first term such that.The first factor of LHS of (4.80)

$$\begin{aligned}
&= (l'+l)^2(l'-l)^2 + 2(l'+l)\left(l'^2 - l'l + l^2\right) - 2l'l(l'+l) - 2(l'+l) - (l'+l)^2 \\
&= (l'+l)\Big[(l'+l)(l'-l)^2 + 2\left(l'^2 - l'l + l^2\right) - 2l'l - 2 - (l'+l)\Big] \\
&= (l'+l)\Big[(l'+l)(l'-l)^2 + 2\left(l'^2 - 2l'l + l^2\right) - (l'+l+2)\Big] \\
&= (l'+l)\Big[(l'+l)(l'-l)^2 + 2(l'-l)^2 - (l'+l+2)\Big] \\
&= (l'+l)\Big\{(l'-l)^2[(l'+l)+2] - (l'+l+2)\Big\} \\
&= (l'+l)(l'+l+2)(l'-l+1)(l'-l-1).
\end{aligned} \tag{4.81}$$

Thus rewriting (4.80), we get

$$(l'+l)(l'+l+2)(l'-l+1)(l'-l-1)\langle l'|z|l\rangle = 0. \tag{4.82}$$

We have similar relations with respect to $\langle l'|x|l\rangle$ and $\langle l'|y|l\rangle$ because of (4.76) and (4.77). For the electric dipole transition to be allowed, among $\langle l'|x|l\rangle$, $\langle l'|y|l\rangle$, and $\langle l'|z|l\rangle$, at least one term must be nonvanishing. For this, at least one of the four factors of (4.81) should be zero. Since $l'+l+2>0$, this factor is excluded.

For $l'+l$ to vanish, we should have $l'=l=0$; notice that both $l'$ and $l$ are nonnegative integers. We must then examine this condition. This condition is equivalent to that the spherical harmonics related to the angular variables $\theta$ and $\phi$ take the form of $Y_0^0(\theta,\phi) = 1/\sqrt{4\pi}$, i.e., a constant. Therefore, the $\theta$-related integral for the matrix element $\langle l'|z|l\rangle$ only consists of a following factor:

$$\int_0^\pi \cos\theta \, \sin\theta \, \mathrm{d}\theta = \frac{1}{2}\int_0^\pi \sin 2\theta \, \mathrm{d}\theta = -\frac{1}{4}[\cos 2\theta]_0^\pi = 0,$$

where $\cos\theta$ comes from a polar coordinate $z = r\cos\theta$; $\sin\theta$ is due to an infinitesimal volume of space, i.e., $r^2 \sin\theta \, \mathrm{d}r \, \mathrm{d}\theta \, \mathrm{d}\phi$. Thus, we find that $\langle l'|z|l\rangle$ vanishes on condition that $l'=l=0$. As a polar coordinate representation, $x = r\sin\theta\cos\phi$ and $y = r\sin\theta\sin\phi$, and so the $\phi$-related integral $\langle l'|x|l\rangle$ and $\langle l'|y|l\rangle$ vanishes as well. That is,

$$\int_0^{2\pi} \cos\phi \, \mathrm{d}\phi = \int_0^{2\pi} \sin\phi \, \mathrm{d}\phi = 0.$$

Therefore, the matrix elements relevant to $l' = l = 0$ vanish with all the coordinates; i.e., we have

$$\langle l'|x|l\rangle = \langle l'|y|l\rangle = \langle l'|z|l\rangle = 0. \tag{4.83}$$

Consequently, we exclude $(l' + l)$-factor as well, when we consider a condition of the allowed transition. Thus, regarding the condition that should be satisfied with the allowed transition, from (4.82) we get

$$l' - l + 1 = 0 \text{ or } l' - l - 1 = 0. \tag{4.84}$$

Or defining $\Delta l \equiv l' - l$, we get

$$\Delta l = \pm 1. \tag{4.85}$$

Thus, for the transition to be allowed, the azimuthal quantum number must change by one.

## 4.5 Angular Momentum of Radiation [6]

In Sect. 4.3, we mentioned circularly polarized light. If the circularly polarized light acts on an electron, what can we anticipate? Here we deal with this problem within a framework of a semiclassical theory.

Let $\boldsymbol{E}$ and $\boldsymbol{H}$ be electric and magnetic fields of a left-circularly polarized light. They are expressed as

$$\boldsymbol{E} = \frac{1}{\sqrt{2}} E_0 (\boldsymbol{e}_1 + i\boldsymbol{e}_2) \exp i(kz - \omega t), \tag{4.86}$$

$$\boldsymbol{H} = \frac{1}{\sqrt{2}} H_0 (\boldsymbol{e}_2 - i\boldsymbol{e}_1) \exp i(kz - \omega t) = \frac{1}{\sqrt{2}} \frac{E_0}{\mu\nu} (\boldsymbol{e}_2 - i\boldsymbol{e}_1) \exp i(kz - \omega t). \tag{4.87}$$

Here we assume that the light is propagating in the direction of the positive $z$-axis. The electric and magnetic fields described by (4.86) and (4.87) represent the left-circularly polarized light. A synchronized motion of an electron is expected, if the electron exerts a circular motion in such a way that the motion direction of the electron is always perpendicular to the electric field and parallel to the magnetic field (see Fig. 4.2). In this situation, magnetic Lorentz force does not affect the electron motion.

Equation (4.86) can be rewritten as

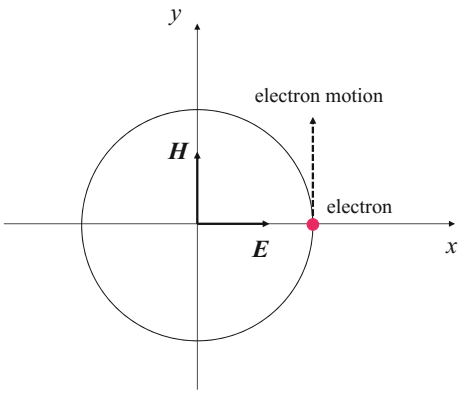


**Fig. 4.2** Synchronized motion of an electron under a left-circularly polarized light

$$\begin{aligned}\boldsymbol{E} = &\frac{1}{\sqrt{2}}E_0[\boldsymbol{e}_1\cos(kz-\omega t)-\boldsymbol{e}_2\sin(kz-\omega t)]\\ &+i\frac{1}{\sqrt{2}}E_0[\boldsymbol{e}_2\cos(kz-\omega t)+\boldsymbol{e}_1\sin(kz-\omega t)].\end{aligned} \tag{4.88}$$

Suppose that the electron exerts the circular motion in a region narrow enough around the origin and that the said electron motion is confined within the $xy$-plane that is perpendicular to the light propagation direction. Then, we can assume that $z \approx 0$ in (4.88). Ignoring $kz$ in (4.88) accordingly and taking a real part, we have

$$\boldsymbol{E} = \frac{1}{\sqrt{2}}E_0(\boldsymbol{e}_1\cos\,\omega t+\boldsymbol{e}_2\sin\,\omega t). \tag{4.89}$$

Thus, a force $\boldsymbol{F}$ exerting the electron is described by

$$\boldsymbol{F} = e\boldsymbol{E}, \tag{4.90}$$

where $e$ is an elementary charge ($e<0$). Accordingly, an equation of motion of the electron is approximated such that

$$m\ddot{\boldsymbol{x}} = e\boldsymbol{E}, \tag{4.91}$$

where $m$ is a mass of an electron and $\boldsymbol{x}$ is a position vector of the electron. With individual components of the coordinate, we have

$$m\ddot{x} = \frac{1}{\sqrt{2}}eE_0\cos\omega t \text{ and } m\ddot{y} = \frac{1}{\sqrt{2}}eE_0\sin\omega t. \tag{4.92}$$

Integrating (4.92) two times, we get

$$mx = -\frac{eE_0}{\sqrt{2}\omega^2}\cos\omega t + Ct + D, \tag{4.93}$$

where $C$ and $D$ are integration constants. Setting $x(0) = -\frac{eE_0}{\sqrt{2}m\omega^2}$ and $x'(0) = 0$, we have $C = D = 0$. Similarly, we have

$$my = -\frac{eE_0}{\sqrt{2}\omega^2}\sin\,\omega t + C't + D', \tag{4.94}$$

where $C'$ and $D'$ are integration constants. Setting $y(0) = 0$ and $y'(0) = -\frac{eE_0}{\sqrt{2}m\omega}$, we have $C' = D' = 0$. Thus, making $t$ a parameter, we get

$$x^2 + y^2 = \left(\frac{eE_0}{\sqrt{2}m\omega^2}\right)^2. \tag{4.95}$$

This implies that the electron is exerting a counterclockwise circular motion with a radius $-\frac{eE_0}{\sqrt{2}m\omega^2}$ under the influence of the electric field. This is consistent with a motion of an electron in the coherent state of $\phi(1s)$ and $\phi\left(2p_{x+iy}\right)$ as expressed in (4.57).

An angular momentum the electron has acquired is

$$\boldsymbol{L} = \boldsymbol{x} \times \boldsymbol{p} = xp_y - yp_x = \left(-\frac{eE_0}{\sqrt{2}m\omega^2}\right)\left(-\frac{meE_0}{\sqrt{2}m\omega}\right) = \frac{e^2E_0^2}{2m\omega^3}. \tag{4.96}$$

Identifying this with $\hbar$, we have

$$\frac{e^2E_0^2}{2m\omega^3} = \hbar. \tag{4.97}$$

In terms of energy, we have

$$\frac{e^2E_0^2}{2m\omega^2} = \hbar\omega. \tag{4.98}$$

Assuming a wavelength of the light is 600 nm, we need a left-circularly polarized light whose electric field is about $1.5 \times 10^{10}\,\mathrm{V/m}$.

A radius $\alpha$ of a circular motion of the electron is given by

$$\alpha = \frac{eE_0}{\sqrt{2}m\omega^2}. \tag{4.99}$$

Under the same condition as the above, $\alpha$ is estimated to be $\sim 2\,\text{Å}$.

# References

1. Jackson JD (1999) Classical electrodynamics. Wiley, New York
2. Fowles GR (1989) Introduction to modern optics, 2nd edn. Dover, New York
3. Arfken GB, Weber HJ, Harris FE (2013) Mathematical methods for physicists, 7th edn. Academic Press, Waltham
4. Chen JQ, Ping J, Wang F (2002) Group representation theory for physicists, 2nd edn. World Scientific, Singapore
5. Sunakawa S (1991) Quantum mechanics. Iwanami, Tokyo (in Japanese)
6. Rossi B (1957) Optics. Addison-Wesley, Reading, Massachusetts

# Part II
# Electromagnetism

Electromagnetism is one of pillars of modern physics, even though it belongs to classical physics along with Newtonian mechanics. Maxwell's equations describe and interpret almost all electromagnetic phenomena and are basis equations of classical physics together with Newton equation. Although after the discovery of relativistic theory Newton equation needed to be modified, Maxwell's equations did not virtually have to be changed. In this part, we treat characteristics of electromagnetic waves that are directly derived from Maxwell's equations.

Electromagnetism becomes connected to quantum mechanics, especially when we deal with emission and absorption of light. In fact, the experiments performed in connection with the blackbody radiation led to discovery of light quanta and establishment of quantum mechanics. These accounts are not only of particular interest from a historical point of view but of great importance in understanding modern physics. To understand the propagation of electromagnetic waves in a dielectric medium is important from a basic aspect of electromagnetism. Moreover, it is deeply connected to optical applications including optical devices such as waveguides and lasers.

The motion of particles as well as spatial and temporal change in, e.g., electromagnetic fields is very often described in terms of differential equations. We describe introductory methods of Green's functions in order to solve those differential equations.

# Chapter 5
# Maxwell's Equations

Maxwell's equations consist of four first-order partial differential equations. First we deal with basic properties of Maxwell's equations. Next we show how equations of electromagnetic wave motion are derived from Maxwell's equations along vector analysis. It is important to realize that the generation of the electromagnetic wave is a direct consequence of the interplay between the electric field and magnetic field that both change with time. We deal with behaviors of electromagnetic waves in dielectric media where no true charge exists. At a first glance, this restriction seems to narrow a range of application of principles of electromagnetism. In practice, however, such a situation is universalistic; topics cover a wide range of electromagnetic phenomena, e.g., light propagation in dielectrics including water, glass, polymers. Polarized properties characterize the electromagnetic waves. These include linear, circular, and elliptic polarizations. The characteristics are important both from a fundamental aspect and from the point of view of optical applications.

## 5.1 Maxwell's Equations and Their Characteristics

In this chapter, we first represent Maxwell's equations as vector forms. The equations are represented as a differential form that is consistent with a viewpoint based on "action trough medium." The equation of wave motion (or wave equation) is naturally derived from these equations.

Maxwell's equations of electromagnetism are expressed as follows:

$$\mathrm{div}\,\boldsymbol{D} = \rho, \tag{5.1}$$

$$\mathrm{div}\,\boldsymbol{B} = 0, \tag{5.2}$$



S. Hotta, *Mathematical Physical Chemistry*,
https://doi.org/10.1007/978-981-10-7671-8_5

$$\mathrm{rot}\,\boldsymbol{E} + \frac{\partial \boldsymbol{B}}{\partial t} = 0, \tag{5.3}$$

$$\mathrm{rot}\,\boldsymbol{H} - \frac{\partial \boldsymbol{D}}{\partial t} = \boldsymbol{i}. \tag{5.4}$$

In (5.3), RHS denotes a zero vector. Let us take some time to get acquainted with physical quantities with their dimension as well as basic ideas and concepts along with definitions of electromagnetism.

The quantity $\boldsymbol{D}$ is called electric flux density $\left[\frac{\mathrm{A{\cdot}s}}{\mathrm{m}^2} = \frac{\mathrm{C}}{\mathrm{m}^2}\right]$ (or electric displacement); $\rho$ is electric charge density $\left[\frac{\mathrm{C}}{\mathrm{m}^3}\right]$. We describe vector quantities $\boldsymbol{V}$ as in (3.4).

$$\boldsymbol{V} = (\boldsymbol{e}_1\boldsymbol{e}_2\boldsymbol{e}_3)\begin{pmatrix} V_x \\ V_y \\ V_z \end{pmatrix}. \tag{5.5}$$

The notation div denotes a differential operator such that

$$\mathrm{div}\,\boldsymbol{V} \equiv \frac{\partial V_x}{\partial x} + \frac{\partial V_y}{\partial y} + \frac{\partial V_z}{\partial z}. \tag{5.6}$$

Thus, the div operator converts a vector to a scalar. The quantities $\boldsymbol{D}$ and the electric field $\boldsymbol{E}$ $[\mathrm{V/m}]$ are associated with the following expression:

$$\boldsymbol{D} = \varepsilon \boldsymbol{E}, \tag{5.7}$$

where $\varepsilon$ $\left[\frac{\mathrm{C}^2}{\mathrm{Nm}^2}\right]$ is called a dielectric constant (or permittivity) of the dielectric medium. The dimension can be understood from the following Coulomb's law that describes a force exerted between two charges:

$$F = \frac{1}{4\pi\varepsilon}\frac{QQ'}{r^2}, \tag{5.8}$$

where $F$ is the force; $Q$ and $Q'$ are electric charges of the two charges; $r$ is a distance between the two charges. Equation (5.1) represents Gauss' law of electrostatics.

The electric charge of 1 C (1 [C]) is defined as follows: Suppose that two point charges having the same electric charge are placed in vacuum 1 m apart. In that situation, if a force $F$ between the two charges is

$$F = \tilde{c}^2/10^7[\mathrm{N}],$$

where $\tilde{c}$ is a light velocity, then we define the electric charge which each point charge possesses as 1 [C]. Here, note that $\tilde{c}$ is a dimensionless number related to the

light velocity in vacuum that is measured in a unit of [m/s]. Notice also that the light velocity $c$ in vacuum is *defined* as

$$c = 299,792,458\ [\mathrm{m/s}]\ (\text{exact number}).$$

Thus, we have

$$\tilde{c} = 299,792,458\ (\text{dimensionless number}).$$

Vacuum is a kind of dielectric media. From (5.8), its dielectric constant $\varepsilon_0$ is defined by

$$\varepsilon_0 = \frac{10^7}{4\pi\tilde{c}^2}\left[\frac{\mathrm{C}^2}{\mathrm{Nm}^2}\right] \approx 8.854 \times 10^{-12}\left[\frac{\mathrm{C}^2}{\mathrm{Nm}^2}\right]. \tag{5.9}$$

Meanwhile, 1 s (second) has been defined from a certain spectral line emitted from $^{133}$Cs. Thus, 1 m (meter) is *defined* by

$$(\text{distance along which light is propagated in vacuum during 1 s})/299792458.$$

The quantity $\boldsymbol{B}$ is called magnetic flux density $\left[\frac{\mathrm{V{\cdot}s}}{\mathrm{m}^2} = \frac{\mathrm{Wb}}{\mathrm{m}^2}\right]$. Equation (5.2) represents Gauss's law of magnetostatics. In contrast to (5.1), RHS of (5.2) is zero. This corresponds to the fact that although a true charge exists, a true "magnetic charge" does not exist. (More precisely, such charge has not been detected so far.) The quantity $\boldsymbol{B}$ is connected with magnetic field $\boldsymbol{H}$ by the following relation:

$$\boldsymbol{B} = \mu\boldsymbol{H}, \tag{5.10}$$

where $\mu\left[\frac{\mathrm{N}}{\mathrm{A}^2}\right]$ is said to be permeability (or magnetic permeability). Thus, $\boldsymbol{H}$ has a dimension $\left[\frac{\mathrm{A}}{\mathrm{m}}\right]$. Permeability of vacuum $\mu_0$ is defined by

$$\mu_0 = 4\pi/10^7\left[\frac{\mathrm{N}}{\mathrm{A}^2}\right]. \tag{5.11}$$

Also we have

$$\mu_0\varepsilon_0 = 1/c^2. \tag{5.12}$$

We will encounter this relation later again. Since the quantities $c$ and $\varepsilon_0$ have a defined magnitude, so does $\mu_0$ from (5.12).

We often make an issue of a relative magnitude of dielectric constant and magnetic permeability of a dielectric medium. That is, relative permittivity $\varepsilon_r$ and relative permeability $\mu_r$ are defined as follows:

$$\varepsilon_r \equiv \varepsilon/\varepsilon_0 \quad \text{and} \quad \mu_r \equiv \mu/\mu_0. \tag{5.13}$$

Note that both $\varepsilon_r$ and $\mu_r$ are dimensionless quantities. Those magnitudes are equal to 1 (in the case of vacuum) or larger than 1 (with any other dielectric media).

Equations (5.3) and (5.4) deal with the change in electric and magnetic fields with time. Of these, (5.3) represents Faraday's law of electromagnetic induction due to Michael Faraday (1831). He found that when a permanent magnet was thrust into or out of a closed circuit, the transient current flowed. Moreover, that experiment implied that even without the closed circuit, an electric field was generated around the space that changed the position relative to the permanent magnet. Equation (5.3) is easier to understand if it is rewritten as follows.

$$\operatorname{rot} \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t}. \tag{5.14}$$

That is, the electric field $\boldsymbol{E}$ is generated in such a way that the induced electric field (or induced current) tends to lessen the change in magnetic flux (or magnetic field) produced by the permanent magnet (Lenz's law). The minus sign in RHS indicates that effect.

The rot operator appearing in (5.3) and (5.4) is defined by

$$\begin{aligned} \operatorname{rot} \boldsymbol{V} = \nabla \times \boldsymbol{V} &= \begin{vmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3 \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ V_x & V_y & V_z \end{vmatrix} \\ &= \left(\frac{\partial V_z}{\partial y} - \frac{\partial V_y}{\partial z}\right)\boldsymbol{e}_1 + \left(\frac{\partial V_x}{\partial z} - \frac{\partial V_z}{\partial x}\right)\boldsymbol{e}_2 + \left(\frac{\partial V_y}{\partial x} - \frac{\partial V_x}{\partial y}\right)\boldsymbol{e}_3. \end{aligned} \tag{5.15}$$

The operator $\nabla$ has already appeared in (3.9). This operator transforms a vector to a vector. Let us think of the meaning of the rot operator. Suppose that there is a vector field that varies with time and spatial positions. Suppose also at some instant the spatial distribution of the field varies as in Fig. 5.1, where a spiral vector field $\boldsymbol{V}$ is present. For a $z$-component of rot $\boldsymbol{V}$ around the origin, we have

$$(\operatorname{rot} \boldsymbol{V})_z = \frac{\partial V_y}{\partial x} - \frac{\partial V_x}{\partial y} = \lim_{\Delta x \to 0, \Delta y \to 0} \left[\frac{(V_1)_y - (V_3)_y}{\Delta x} - \frac{(V_2)_x - (V_4)_x}{\Delta y}\right].$$

In the case of Fig. 5.1, $(V_1)_y - (V_3)_y > 0$ and $(V_2)_x - (V_4)_x < 0$ and, hence, we find that rot $\boldsymbol{V}$ has a positive $z$-component. If $V_z = 0$ and $\frac{\partial V_y}{\partial z} = \frac{\partial V_x}{\partial z} = 0$, we find from (5.15) that rot $\boldsymbol{V}$ possesses only the $z$-component. The equation $\frac{\partial V_y}{\partial z} = \frac{\partial V_x}{\partial z} = 0$ implies that the vector field $\boldsymbol{V}$ is uniform in the direction of the $z$-axis. Thus, under the above conditions, the spiral vector field $\boldsymbol{V}$ is accompanied by the rot $\boldsymbol{V}$ vector field that is directed toward the upper side of the plane of paper (i.e., the positive direction of the $z$-axis).

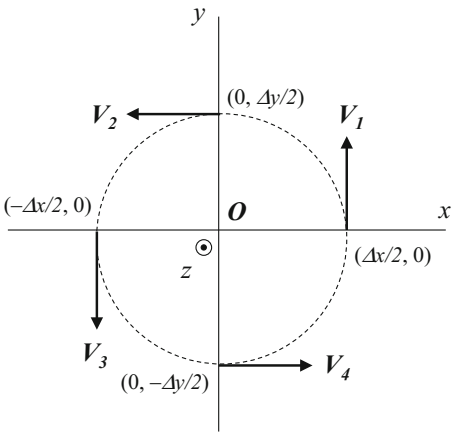


**Fig. 5.1** Schematic representation of a spiral vector field $\boldsymbol{V}$ that yields rot $\boldsymbol{V}$

Equation (5.4) can be rewritten as

$$\mathrm{rot}\,\boldsymbol{H} = \boldsymbol{i} + \frac{\partial \boldsymbol{D}}{\partial t}. \tag{5.16}$$

Notice that $\frac{\partial \boldsymbol{D}}{\partial t}$ has the same dimension as $\left[\frac{\mathrm{A}}{\mathrm{m}^2}\right]$ and is called displacement current. Without this term, we have

$$\mathrm{rot}\,\boldsymbol{H} = \boldsymbol{i}. \tag{5.17}$$

This relation is well known as Ampère's law or Ampère's circuital law (André-Marie Ampère: 1827), which determines a magnetic field yielded by a stationary current. Again with the aid of Fig. 5.1, (5.17) implies that the current given by $\boldsymbol{i}$ produces spiral magnetic field.

Now, let us think of a change in amount of charges with time in a part of three-dimensional closed space $V$ surrounded by a closed surface $S$. It is given by

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_V \rho\,\mathrm{d}V = \int_V \frac{\partial \rho}{\partial t}\mathrm{d}V = -\int_S \boldsymbol{i}\cdot\boldsymbol{n}\,\mathrm{d}S = -\int_V \mathrm{div}\,\boldsymbol{i}\,\mathrm{d}V, \tag{5.18}$$

where $\boldsymbol{n}$ is an outward-directed normal unit vector; with the last equality, we used Gauss's theorem. The Gauss's theorem is described by

$$\int_V \mathrm{div}\,\boldsymbol{i}\,\mathrm{d}V = \int_S \boldsymbol{i}\cdot\boldsymbol{n}\,\mathrm{d}S.$$

Figure 5.2 gives an intuitive diagram that explains the Gauss's theorem. The diagram shows a cross section of the closed space $V$ surrounded by a surface $S$. In

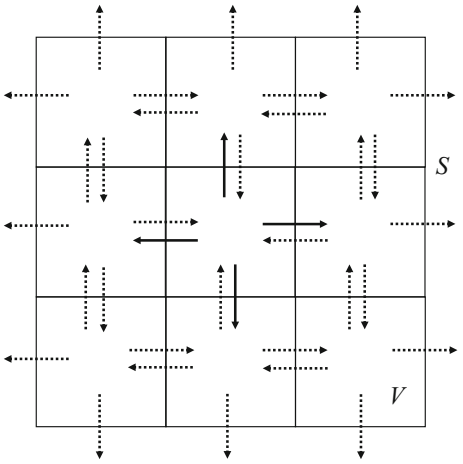


**Fig. 5.2** Intuitive diagram that explains the Gauss's theorem

this case, imagine a cube or a hexahedron as $V$. The periphery is the cross section of the closed surface accordingly. Arrows in the diagram schematically represent $\mathrm{div}\,\boldsymbol{i}$ on individual fragments; only those of the center infinitesimal fragment are shown with solid lines. The arrows of adjacent fragments cancel out each other, and only the components on the periphery are nonvanishing. Thus, the volume integration of $\mathrm{div}\,\boldsymbol{i}$ is converted to the surface integration of $\boldsymbol{i}$. Readers are referred to the appropriate literature with the vector integration [1].

Consequently, from (5.18) we have

$$\int_V \left( \frac{\partial \rho}{\partial t} + \mathrm{div}\,\boldsymbol{i} \right) \mathrm{d}V = 0. \tag{5.19}$$

Since $V$ is arbitrarily chosen, we get

$$\frac{\partial \rho}{\partial t} + \mathrm{div}\,\boldsymbol{i} = 0. \tag{5.20}$$

The relation (5.20) is called a current continuity equation. This relation represents law of conservation of charge.

Meanwhile, taking div of both sides of (5.17), we have

$$\mathrm{div}\,\mathrm{rot}\,\boldsymbol{H} = \mathrm{div}\,\boldsymbol{i}. \tag{5.21}$$

The LHS of (5.21) reads

$$\begin{aligned}\operatorname{div}\operatorname{rot}\boldsymbol{H} &= \frac{\partial}{\partial x}\left(\frac{\partial H_z}{\partial y}-\frac{\partial H_y}{\partial z}\right)+\frac{\partial}{\partial y}\left(\frac{\partial H_x}{\partial z}-\frac{\partial H_z}{\partial x}\right)+\frac{\partial}{\partial z}\left(\frac{\partial H_y}{\partial x}-\frac{\partial H_x}{\partial y}\right)\\ &= \left(\frac{\partial^2}{\partial x\partial y}-\frac{\partial^2}{\partial y\partial x}\right)H_z+\left(\frac{\partial^2}{\partial y\partial z}-\frac{\partial^2}{\partial z\partial y}\right)H_x+\left(\frac{\partial^2}{\partial z\partial x}-\frac{\partial^2}{\partial x\partial z}\right)H_y\\ &= 0.\end{aligned} \tag{5.22}$$

With the last equality of (5.22), we used the fact that if, e.g., $\frac{\partial^2 H_z}{\partial x\partial y}$ and $\frac{\partial^2 H_z}{\partial y\partial x}$ are continuous and differentiable in a certain domain $(x, y)$, $\frac{\partial^2 H_z}{\partial x\partial y}=\frac{\partial^2 H_z}{\partial y\partial x}$. That is, we assume "ordinary" functions for $H_z$, $H_x$, and $H_y$. Thus from (5.21), we have

$$\operatorname{div}\ \boldsymbol{i}=0. \tag{5.23}$$

From (5.20), we also have

$$\frac{\partial\rho(\boldsymbol{x},\,t)}{\partial t}=0, \tag{5.24}$$

where we explicitly show that $\rho$ depends upon both $\boldsymbol{x}$ and $t$. Note that $\boldsymbol{x}$ is a position vector described as (3.5). Therefore, (5.24) shows that $\rho(\boldsymbol{x},t)$ is temporally constant at a position $\boldsymbol{x}$, consistent with the stationary current.

Nevertheless, we encounter a problem when $\rho(\boldsymbol{x},t)$ is temporally varying. In other words, (5.17) goes against the charge conservation law, when $\rho(\boldsymbol{x},t)$ is temporally varying. It was James Clerk Maxwell (1861–1862) that solved the problem by introducing a concept of the displacement current. In fact, taking div of both sides of (5.16), we have

$$\operatorname{div}\operatorname{rot}\boldsymbol{H}=\operatorname{div}\boldsymbol{i}+\frac{\partial\operatorname{div}\boldsymbol{D}}{\partial t}=\operatorname{div}\boldsymbol{i}+\frac{\partial\rho}{\partial t}=0, \tag{5.25}$$

where with the first equality, we exchanged the order of differentiations with respect to $t$ and $\boldsymbol{x}$; with the second equality, we used (5.1). The last equality of (5.25) results from (5.20). In other words, in virtue of the term of $\frac{\partial\boldsymbol{D}}{\partial t}$, (5.4) is consistent with the charge conservation law. Thus, the set of Maxwell's equations (5.1)–(5.4) supply us with well-established base in natural science up until the present.

Although the set of these equations describes spatial and temporal changes in electric and magnetic fields in vacuum and matter including metal, in Part II we confine ourselves to the changes in the electric and magnetic fields in a uniform dielectric medium.

## 5.2 Equation of Wave Motion

If we further confine ourselves to the case where neither electric charge nor electric current is present in a uniform dielectric medium, we can readily obtain equations of wave motion regarding the electric and magnetic fields. That is,

$$\operatorname{div} \boldsymbol{D} = 0, \tag{5.26}$$

$$\operatorname{div} \boldsymbol{B} = 0, \tag{5.27}$$

$$\operatorname{rot} \boldsymbol{E} + \frac{\partial \boldsymbol{B}}{\partial t} = 0, \tag{5.28}$$

$$\operatorname{rot} \boldsymbol{H} - \frac{\partial \boldsymbol{D}}{\partial t} = 0. \tag{5.29}$$

The relations (5.27) and (5.28) are identical to (5.2) and (5.3), respectively.

Let us start with a formula of vector analysis. First, we introduce a grad operator. We have

$$\operatorname{grad} f = \nabla f = \frac{\partial f}{\partial x}\boldsymbol{e}_1 + \frac{\partial f}{\partial y}\boldsymbol{e}_2 + \frac{\partial f}{\partial z}\boldsymbol{e}_3.$$

That is, the grad operator transforms a scalar to a vector. We have a following formula:

$$\operatorname{rot} \operatorname{rot} \boldsymbol{V} = \operatorname{grad} \operatorname{div} \boldsymbol{V} - \nabla^2 \boldsymbol{V}. \tag{5.30}$$

The operator $\nabla^2$ has already appeared in (1.24). To show (5.30), we compare an $x$-component of both sides of (5.30). That is,

$$[\operatorname{rot} \operatorname{rot} \boldsymbol{V}]_x = \frac{\partial}{\partial y}\left(\frac{\partial V_y}{\partial x} - \frac{\partial V_x}{\partial y}\right) - \frac{\partial}{\partial z}\left(\frac{\partial V_x}{\partial z} - \frac{\partial V_z}{\partial x}\right). \tag{5.31}$$

$$\begin{aligned}\left[\operatorname{grad} \operatorname{div} \boldsymbol{V} - \nabla^2 \boldsymbol{V}\right]_x &= \frac{\partial}{\partial x}\left(\frac{\partial V_x}{\partial x} + \frac{\partial V_y}{\partial y} + \frac{\partial V_z}{\partial z}\right) - \frac{\partial^2 V_x}{\partial x^2} - \frac{\partial^2 V_x}{\partial y^2} - \frac{\partial^2 V_x}{\partial z^2} \\ &= \frac{\partial}{\partial x}\left(\frac{\partial V_y}{\partial y} + \frac{\partial V_z}{\partial z}\right) - \frac{\partial^2 V_x}{\partial y^2} - \frac{\partial^2 V_x}{\partial z^2}.\end{aligned} \tag{5.32}$$

Again assuming that $\frac{\partial^2 V_y}{\partial y \partial x} = \frac{\partial^2 V_y}{\partial x \partial y}$ and $\frac{\partial^2 V_z}{\partial z \partial x} = \frac{\partial^2 V_z}{\partial x \partial z}$, we have

$$[\operatorname{rot} \operatorname{rot} \boldsymbol{V}]_x = [\operatorname{grad} \operatorname{div} \boldsymbol{V} - \nabla^2 \boldsymbol{V}]_x. \tag{5.33}$$

Regarding $y$- and $z$-components, we have similar relations as well. Thus (5.30) holds.

Taking rot of both sides of (5.28), we have

$$\mathrm{rot}\,\mathrm{rot}\,\boldsymbol{E} + \mathrm{rot}\frac{\partial \boldsymbol{B}}{\partial t} = \mathrm{grad}\,\mathrm{div}\,\boldsymbol{E} - \nabla^2\boldsymbol{E} + \frac{\partial \mathrm{rot}\,\boldsymbol{B}}{\partial t} = -\nabla^2\boldsymbol{E} + \mu\varepsilon\frac{\partial^2\boldsymbol{E}}{\partial t^2} = 0, \quad (5.34)$$

where with the first equality, we used (5.30) and for the second equality, we used (5.7), (5.10), and (5.29). Thus we have

$$\nabla^2\boldsymbol{E} = \mu\varepsilon\frac{\partial^2\boldsymbol{E}}{\partial t^2}. \quad (5.35)$$

Similarly, from (5.29) we get

$$\nabla^2\boldsymbol{H} = \mu\varepsilon\frac{\partial^2\boldsymbol{H}}{\partial t^2}. \quad (5.36)$$

Equations (5.35) and (5.36) are called equations of wave motions for the electric and magnetic fields.

To consider implications of these equations, let us think of for simplicity a following equation in a one-dimensional space.

$$\frac{\partial^2 y(x,\,t)}{\partial x^2} = \frac{1}{v^2}\frac{\partial^2 y(x,\,t)}{\partial t^2}, \quad (5.37)$$

where $y$ is an arbitrary scalar function that depends on $x$ and $t$; $v$ is a constant. Let $f(x,\,t)$ and $g(x,\,t)$ be arbitrarily chosen functions. Then, $f(x-vt)$ and $g(x+vt)$ are two solutions of (5.37). In fact, putting $X = x - vt$, we have

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial X}\frac{\partial X}{\partial x} = \frac{\partial f}{\partial X}, \quad \frac{\partial^2 f}{\partial x^2} = \left[\frac{\partial}{\partial X}\left(\frac{\partial f}{\partial X}\right)\right]\frac{\partial X}{\partial x} = \frac{\partial^2 f}{\partial X^2}, \quad (5.38)$$

$$\frac{\partial f}{\partial t} = \frac{\partial f}{\partial X}\frac{\partial X}{\partial t} = (-v)\frac{\partial f}{\partial X}, \frac{\partial^2 f}{\partial t^2} = (-v)\left[\frac{\partial}{\partial X}\left(\frac{\partial f}{\partial X}\right)\right]\frac{\partial X}{\partial t} = (-v)^2\frac{\partial^2 f}{\partial X^2}. \quad (5.39)$$

From the second equations of (5.38) and (5.39), we recover

$$\frac{\partial^2 f}{\partial x^2} = \frac{1}{v^2}\frac{\partial^2 f}{\partial t^2}. \quad (5.40)$$

Similarly, we get

$$\frac{\partial^2 g}{\partial x^2} = \frac{1}{v^2}\frac{\partial^2 g}{\partial t^2}. \tag{5.41}$$

Therefore, as a general solution, we can take a superposition of $f(x, t)$ and $g(x, t)$. That is,

$$y(x, t) = f(x - vt) + g(x + vt). \tag{5.42}$$

The implication of (5.42) is as follows: (i) The function $f(x - vt)$ can be obtained by parallel translation of $f(x)$ by $vt$ in a positive direction of $x$-axis. In other words, $f(x - vt)$ is obtained by translating $f(x)$ by $v$ in a unit of time in a positive direction of $x$-axis, or the function represented by $f(x)$ is translated at a rate of $v$ with its form unchanged in time. (ii) The function $g(x + vt)$, on the other hand, is translated at a rate of $-v$ with its form unchanged in time as well. (iii) Thus, $y(x, t)$ of (5.42) represents two "waves", i.e., a forward wave and a backward wave. Propagation velocity of the two waves is $|v|$ accordingly. Usually, we choose a positive number for $v$, and $v$ is called a phase velocity.

Comparing (5.35) and (5.36) with (5.41), we have

$$\mu\varepsilon = 1/v^2. \tag{5.43}$$

In particular, in a vacuum we recover (5.12).

Notice that $f$ and $g$ can take *any* functional form and, hence, they are not necessarily a periodic wave. Yet, what we are mostly concerned with is a periodic wave such as sinusoidal waves. Thus, we arrive at a following functional form:

$$f(x - vt) = Ae^{i(x - vt)}, \tag{5.44}$$

where $A$ is said to be an amplitude of the wave. The constant $A$ usually takes a positive number, but it may take a complex number including a negative number. An exponent of (5.44) contains a number having a dimension [m]. To make it a dimensionless number, we multiply $(x - vt)$ by a wave number $k$ that has been introduced in (1.2) and (1.3). That is, we have

$$\tilde{f}[k(x - vt)] = Ae^{ik(x - vt)} = Ae^{i(kx - kvt)} = Ae^{i(kx - \omega t)}, \tag{5.45}$$

where $\tilde{f}$ shows the change in the functional from according to the variable transformation. In (5.45), we have

$$kv = k\lambda v = (2\pi/\lambda)\lambda v = 2\pi v = \omega, \tag{5.46}$$

where $v$ and $\omega$ are said to be frequency and angular frequency, respectively. For a three-dimensional wave $f$ of a scalar function, we have a following form:

$$f = Ae^{i(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)} = A[\cos(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t) + i\,\sin(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)], \tag{5.47}$$

where $\boldsymbol{k}$ is said to be a wave number vector. In (5.47),

$$\boldsymbol{k}^2 = k^2 = k_x^2 + k_y^2 + k_z^2. \tag{5.48}$$

Equation (5.47) is virtually identical to (1.25). When we deal with a problem of classical electromagnetism, we usually take a real part of the results after relevant calculations.

Suppose that (5.47) is a solution of a following wave equation:

$$\nabla^2 f = \frac{1}{v^2}\frac{\partial^2 f}{\partial t^2}. \tag{5.49}$$

Substituting LHS of (5.47) for $f$ of (5.49), we have

$$A\left(-k^2\right)e^{i(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)} = A\frac{1}{v^2}\left(-\omega^2\right)e^{i(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)}. \tag{5.50}$$

Comparing both sides of (5.50), we get

$$k^2v^2 = \omega^2 \quad \text{or} \quad kv = \omega. \tag{5.51}$$

Thus, we recover (5.46).

Here we introduce a unit vector $\boldsymbol{n}$ as in (1.3) whose direction parallels that of propagation of wave such that

$$\boldsymbol{k} = k\boldsymbol{n} = \frac{2\pi}{\lambda}\boldsymbol{n}, \quad \boldsymbol{n} = (\boldsymbol{e}_1\boldsymbol{e}_2\boldsymbol{e}_3)\begin{pmatrix} n_x \\ n_y \\ n_z \end{pmatrix}, \tag{5.52}$$

where $n_x$, $n_y$, and $n_z$ define direction cosines. Then (5.47) can be rewritten as

$$f = Ae^{i(k\boldsymbol{n}\cdot\boldsymbol{x}-\omega t)}, \tag{5.53}$$

where an exponent is called a phase. Suppose that the phase is fixed at zero. That is,

$$k\boldsymbol{n}\cdot\boldsymbol{x} - \omega t = 0. \tag{5.54}$$

Since $\omega = kv$, we have

$$\boldsymbol{n} \cdot \boldsymbol{x} - vt = 0. \tag{5.55}$$

Equation (5.55) defines a plane in a three-dimensional space and is called a Hesse's normal form. Figure 5.3 schematically represents a plane wave of the field $f$. The field has the same phase on a plane $P$. A solid arrow $\boldsymbol{x}$ represents an arbitrary position vector on the plane and $\boldsymbol{n}$ is a unit vector perpendicular to the plane $P$ (i.e., parallel to a normal of the plane $P$). The quantity $vt$ defines a length of a perpendicular that connects the origin $O$ and plane $P$ (i.e., the length of the perpendicular from the origin and a foot of the perpendicular) at a given time $t$.

In other words, (5.54) determines a plane in such a way that the wave $f$ has the same phase (zero) at a given time $t$ at position vectors $\boldsymbol{x}$ on the plane determined by (5.54) or (5.55). That plane is moving in the direction of $\boldsymbol{n}$ at a phase velocity $v$. From this situation, a wave $f$ described by (5.53) is called a plane wave.

A refractive index $n$ of a dielectric media is an important index that characterizes its dielectric properties. It is defined as

$$n \equiv c/v = \sqrt{\mu\varepsilon/\mu_0\varepsilon_0} = \sqrt{\mu_r\varepsilon_r}. \tag{5.56}$$

In a non-magnetic substance such as glass and polymer materials, we can assume that $\mu_r \approx 1$. Thus, we get an approximate expression as follows:

$$n \approx \sqrt{\varepsilon_r}. \tag{5.57}$$

## 5.3 Polarized Characteristics of Electromagnetic Waves

As in (5.47), we assume a similar form for a solution of (5.35) such that

$$\boldsymbol{E} = \boldsymbol{E}_0 \mathrm{e}^{i(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)} = \boldsymbol{E}_0 \mathrm{e}^{i(k\boldsymbol{n}\cdot\boldsymbol{x}-\omega t)}, \tag{5.58}$$

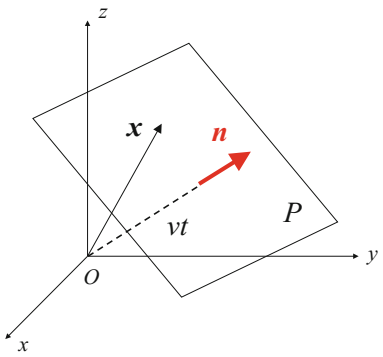


**Fig. 5.3** Schematic representation of a plane wave. The field has the same phase on a plane $P$. A solid arrow $\boldsymbol{x}$ represents an arbitrary position vector on the plane and $\boldsymbol{n}$ is a unit vector perpendicular to the plane $P$ (i.e., parallel to a normal of the plane $P$). The quantity $v$ is a phase velocity of the plane wave

where with the second equality we used (5.52). Similarly, we have

$$\boldsymbol{H} = \boldsymbol{H}_0 \mathrm{e}^{i(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)} = \boldsymbol{H}_0 \mathrm{e}^{i(k\boldsymbol{n}\cdot\boldsymbol{x}-\omega t)}. \tag{5.59}$$

In (5.58) and (5.59), $\boldsymbol{E}_0$ and $\boldsymbol{H}_0$ are constant vectors and may take a complex magnitude. Substituting (5.58) for (5.28) and using (5.10) as well as (5.43) and (5.46), we have

$$\begin{aligned} ik\boldsymbol{n} \times \boldsymbol{E}_0 \mathrm{e}^{i(k\boldsymbol{n}\cdot\boldsymbol{x}-\omega t)} &= -(-i\omega)\mu \boldsymbol{H}_0 \mathrm{e}^{i(k\boldsymbol{n}\cdot\boldsymbol{x}-\omega t)} = iv k \mu \boldsymbol{H}_0 \mathrm{e}^{i(k\boldsymbol{n}\cdot\boldsymbol{x}-\omega t)} \\ &= ik\sqrt{\mu/\varepsilon}\boldsymbol{H}_0 \mathrm{e}^{i(k\boldsymbol{n}\cdot\boldsymbol{x}-\omega t)}. \end{aligned}$$

Comparing coefficients of the exponential functions of the first and last sides, we get

$$\boldsymbol{H}_0 = \boldsymbol{n} \times \boldsymbol{E}_0 \Big/ \left(\sqrt{\mu/\varepsilon}\right). \tag{5.60}$$

Similarly, substituting (5.59) for (5.29) and using (5.7), we get

$$\boldsymbol{E}_0 = \left(\sqrt{\mu/\varepsilon}\right)\boldsymbol{H}_0 \times \boldsymbol{n}. \tag{5.61}$$

From (5.26) and (5.27), we have

$$\boldsymbol{n} \cdot \boldsymbol{E}_0 = \boldsymbol{n} \cdot \boldsymbol{H}_0 = 0. \tag{5.62}$$

This indicates that $\boldsymbol{E}$ and $\boldsymbol{H}$ are both perpendicular to $\boldsymbol{n}$; i.e., the propagation direction of the electromagnetic wave. Thus, the electromagnetic wave is characterized by a transverse wave. The fields $\boldsymbol{E}$ and $\boldsymbol{H}$ have the same phase on $P$ at an arbitrary given time. Taking account of (5.60)–(5.62), $\boldsymbol{E}$, $\boldsymbol{H}$, and $\boldsymbol{n}$ are mutually perpendicular to one another. We depict a geometry of $\boldsymbol{E}$, $\boldsymbol{H}$, and $\boldsymbol{n}$ for the electromagnetic plane wave in Fig. 5.4, where a plane $P$ is perpendicular to $\boldsymbol{n}$.

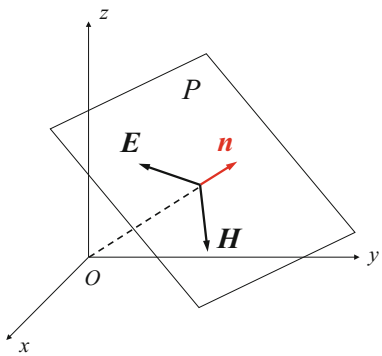


**Fig. 5.4** Mutual geometry of $\boldsymbol{E}$ and $\boldsymbol{H}$ for an electromagnetic plane wave in $P$. $\boldsymbol{E}$ and $\boldsymbol{H}$ have the same phase on $P$ at an arbitrary given time. The unit vector $\boldsymbol{n}$ is perpendicular to $P$

We find that (5.60) is not independent of (5.61). In fact, taking an outer product from the right with respect to both sides of (5.60), we have

$$\boldsymbol{H}_0 \times \boldsymbol{n} = \boldsymbol{n} \times \boldsymbol{E}_0 \times \boldsymbol{n} \Big/ \left(\sqrt{\mu/\varepsilon}\right) = \boldsymbol{E}_0 \Big/ \left(\sqrt{\mu/\varepsilon}\right),$$

where we used (5.61) with the second equality. Thus, we get

$$\boldsymbol{n} \times \boldsymbol{E}_0 \times \boldsymbol{n} = \boldsymbol{E}_0. \tag{5.63}$$

Meanwhile, vector analysis tells us that [1].

$$\boldsymbol{C} \times (\boldsymbol{A} \times \boldsymbol{B}) = \boldsymbol{A}(\boldsymbol{B} \cdot \boldsymbol{C}) - \boldsymbol{B}(\boldsymbol{C} \cdot \boldsymbol{A}).$$

In the above, putting $\boldsymbol{B} = \boldsymbol{C} = \boldsymbol{n}$, we have

$$\boldsymbol{n} \times (\boldsymbol{A} \times \boldsymbol{n}) = \boldsymbol{A}(\boldsymbol{n} \cdot \boldsymbol{n}) - \boldsymbol{n}(\boldsymbol{n} \cdot \boldsymbol{A}).$$

That is, we have

$$\boldsymbol{A} = \boldsymbol{n}(\boldsymbol{n} \cdot \boldsymbol{A}) + \boldsymbol{n} \times (\boldsymbol{A} \times \boldsymbol{n}).$$

This relation means that $\boldsymbol{A}$ can be decomposed into a component parallel to $\boldsymbol{n}$ and that perpendicular to $\boldsymbol{n}$. Equation (5.63) shows that $\boldsymbol{E}_0$ has no component parallel to $\boldsymbol{n}$. This is another confirmation that $\boldsymbol{E}$ is perpendicular to $\boldsymbol{n}$.

In (5.60) and (5.61), $\sqrt{\mu/\varepsilon}$ has a dimension [Ω]. Make sure that this can be confirmed by (5.9) and (5.11). Hence, $\sqrt{\mu/\varepsilon}$ is said to be characteristic impedance [2]. We denote it by

$$Z \equiv \sqrt{\mu/\varepsilon}. \tag{5.64}$$

Thus, we have

$$\boldsymbol{H}_0 = (\boldsymbol{n} \times \boldsymbol{E}_0)/Z.$$

In vacuum, we have

$$Z_0 = \sqrt{\mu_0/\varepsilon_0} \approx 376.7[\Omega].$$

For the electromagnetic wave to be the transverse wave means that neither $\boldsymbol{E}$ nor $\boldsymbol{H}$ has component along $\boldsymbol{n}$. Choosing the positive direction of the $z$-axis for $\boldsymbol{n}$ and ignoring components related to partial differentiation with respect to $x$ and $y$ (i.e., the component related to $\partial/\partial x$ and $\partial/\partial y$), we rewrite (5.28) and (5.29) for individual Cartesian coordinates as

$$\begin{aligned}
-\frac{\partial E_y}{\partial z}+\frac{\partial B_x}{\partial t}=0 \quad \text{or} \quad -\frac{\partial D_y}{\partial z}+\mu\varepsilon\frac{\partial H_x}{\partial t}=0,\\
\frac{\partial E_x}{\partial z}+\frac{\partial B_y}{\partial t}=0 \quad \text{or} \quad \frac{\partial D_x}{\partial z}+\mu\varepsilon\frac{\partial H_y}{\partial t}=0,\\
-\frac{\partial H_y}{\partial z}-\frac{\partial D_x}{\partial t}=0 \quad \text{or} \quad -\frac{\partial B_y}{\partial z}-\mu\varepsilon\frac{\partial x}{\partial t}=0,\\
\frac{\partial H_x}{\partial z}-\frac{\partial D_y}{\partial t}=0 \quad \text{or} \quad \frac{\partial B_x}{\partial z}-\mu\varepsilon\frac{\partial E_y}{\partial t}=0.
\end{aligned} \tag{5.65}$$

We differentiate the first equation of (5.65) with respect to $z$ to get

$$-\frac{\partial^2 E_y}{\partial z^2}+\frac{\partial^2 B_x}{\partial z\partial t}=0.$$

Also differentiating the fourth equation of (5.65) with respect to $t$ and multiplying both sides by $-\mu$, we have

$$-\mu\frac{\partial^2 H_x}{\partial t\partial z}+\mu\frac{\partial^2 D_y}{\partial t^2}=0.$$

Summing both sides of the above equations and arranging terms, we get

$$\frac{\partial^2 E_y}{\partial z^2}=\mu\varepsilon\frac{\partial^2 E_y}{\partial t^2}.$$

In a similar manner, we have

$$\frac{\partial^2 E_x}{\partial z^2}=\mu\varepsilon\frac{\partial^2 E_x}{\partial t^2}.$$

Similarly, for the magnetic field, we also get

$$\frac{\partial^2 H_x}{\partial z^2}=\mu\varepsilon\frac{\partial^2 H_x}{\partial t^2} \quad \text{and} \quad \frac{\partial^2 H_y}{\partial z^2}=\mu\varepsilon\frac{\partial^2 H_y}{\partial t^2}.$$

From the above relations, we have two plane electromagnetic waves polarized either the $x$-axis or $y$-axis.

What is implied in the above description is that as solutions of (5.35) and (5.36), we have a plane wave characterized by a specific direction defined by $\boldsymbol{E}_0$ and $\boldsymbol{H}_0$. This implies that if we observe the electromagnetic wave at a fixed point, both $\boldsymbol{E}$ and $\boldsymbol{H}$ oscillate along the mutually perpendicular directions $\boldsymbol{E}_0$ and $\boldsymbol{H}_0$. Hence, we say that the "electric wave" is polarized in the direction $\boldsymbol{E}_0$ and that the "magnetic wave" is polarized in the direction $\boldsymbol{H}_0$.

To characterize the polarization of the electromagnetic wave, we introduce following unit polarization vectors $\boldsymbol{\varepsilon}_\mathrm{e}$ and $\boldsymbol{\varepsilon}_\mathrm{m}$ (with indices e and m related to the electric and magnetic field, respectively) [3].

$$\boldsymbol{\varepsilon}_\mathrm{e} = \boldsymbol{E}_0/E_0 \quad \text{and} \quad \boldsymbol{\varepsilon}_\mathrm{m} = \boldsymbol{H}_0/H_0; H_0 = E_0/Z, \tag{5.66}$$

where $E_0$ and $H_0$ are said to be amplitude and may be again complex. We have

$$\boldsymbol{\varepsilon}_\mathrm{e} \times \boldsymbol{\varepsilon}_\mathrm{m} = \boldsymbol{n}. \tag{5.67}$$

We call $\boldsymbol{\varepsilon}_\mathrm{e}$ and $\boldsymbol{\varepsilon}_\mathrm{m}$ a unit polarization vector of the electric field and magnetic field, respectively. As noted above, $\boldsymbol{\varepsilon}_\mathrm{e}$, $\boldsymbol{\varepsilon}_\mathrm{m}$, and $\boldsymbol{n}$ constitute a right-handed system in this order and are mutually perpendicular to one another.

The phase of $\boldsymbol{E}$ in the plane wave (5.58) and that of $\boldsymbol{H}$ in (5.59) are individually the same on all the points of $P$. From the wave equations of (5.35) and (5.36), however, it is unclear whether $\boldsymbol{E}$ and $\boldsymbol{H}$ have the same phase. Suppose that $\boldsymbol{E}$ and $\boldsymbol{H}$ would have a different phase such that

$$\boldsymbol{E} = \boldsymbol{E}_0 \mathrm{e}^{i(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)} \quad \text{and}$$

$$\boldsymbol{H} = \boldsymbol{H}_0 \mathrm{e}^{i(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t+\delta)} = \boldsymbol{H}_0 \mathrm{e}^{i\delta} \mathrm{e}^{i(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)} = \tilde{\boldsymbol{H}}_0 \mathrm{e}^{i(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)},$$

where $\tilde{\boldsymbol{H}}_0 \left(= \boldsymbol{H}_0 \mathrm{e}^{i\delta}\right)$ is complex and a phase factor $\mathrm{e}^{i\delta}$ in the exponent is unknown. This factor, however, can be set at zero. To show this, let us make qualitative discussion using Fig. 5.5. Figure 5.5 shows the electric field of the plane wave at some instant as a function of phase $\Phi$. Suppose that $\Phi$ is taken in the direction of $\boldsymbol{n}$ in Fig. 5.4. Then, from (5.53) we have

$$\Phi = k\boldsymbol{n}\cdot\boldsymbol{x} - \omega t = k\boldsymbol{n}\cdot\rho\boldsymbol{n} - \omega t = k\rho - \omega t,$$

where $\rho$ is distance from the origin. Also we have

$$\boldsymbol{E} = \boldsymbol{E}_0 \mathrm{e}^{i(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)} = E_0 \boldsymbol{\varepsilon}_\mathrm{e} \mathrm{e}^{i(k\rho-\omega t)} (E_0 > 0).$$

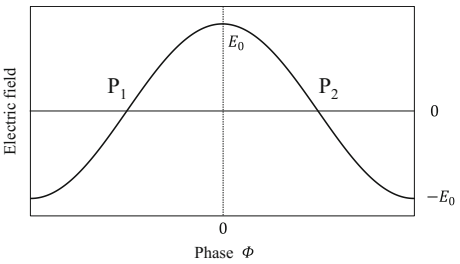


**Fig. 5.5** Electric field of a plane wave at some instant as a function of phase. The sinusoidal curve representing the electric field is shifted with time from left to right with its form unchanged

Suppose furthermore that the phase is measured at $t = 0$ and that the electric field is measured along the $\boldsymbol{\varepsilon}_e$ direction. Then, we have

$$\Phi = k\rho \quad \text{and} \quad E = |\boldsymbol{E}| = E_0 e^{ik\rho}. \tag{5.68}$$

In Fig. 5.5, a real part of $E$ is shown with $E = E_0$ at $\rho = 0$. The sinusoidal curve representing the electric field in Fig. 5.5 is shifted with time from left to right with its form unchanged.

In this situation, the electric field is to be strengthened with a negative value at a phase $P_1$ with time according to (5.68), whereas at another phase $P_2$, the electric field is to be strengthened with a positive value with time. As a result, in a region tucked between $P_1$ and $P_2$, a spiral magnetic field is generated toward the $\boldsymbol{\varepsilon}_m$ direction; i.e., upper side of a plane of paper. The magnitude of the magnetic field is expected to be maximized at a center point of $P_1P_2$ (i.e., the origin of the coordinate system) where the electric field is maximized as well. Thus, we conclude that $\boldsymbol{E}$ and $\boldsymbol{H}$ have the same phase. It is important to realize that the generation of the electromagnetic wave is a direct consequence of the interplay between the electric field and magnetic field that both change with time. It is truly based upon the nature of the electric and magnetic fields that are clearly represented in Maxwell's equations.

Next, we consider the situation where *two* electromagnetic waves are propagated in the same direction but with a different phase. Notice again that we are considering the electromagnetic wave that is propagated in a uniform and infinite dielectric media without BCs.

## 5.4 Superposition of Two Electromagnetic Waves

Let $\boldsymbol{E}_1$ and $\boldsymbol{E}_2$ be two electric waves described such that

$$\boldsymbol{E}_1 = E_1 \boldsymbol{e}_1 e^{i(kz-\omega t)} \quad \text{and} \quad \boldsymbol{E}_2 = E_2 \boldsymbol{e}_2 e^{i(kz-\omega t+\delta)},$$

where $E_1(>0)$ and $E_2(>0)$ are amplitudes and $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ represent unit polarization vectors in the direction of positive $x$-axis and $y$-axis; we assume that two waves are being propagated in the direction of the positive $z$-axis; $\delta$ is a phase difference. The total electric field $\boldsymbol{E}$ is described as the superposition of $\boldsymbol{E}_1$ and $\boldsymbol{E}_2$ such that

$$\boldsymbol{E} = \boldsymbol{E}_1 + \boldsymbol{E}_2 = E_1 \boldsymbol{e}_1 e^{i(kz-\omega t)} + E_2 \boldsymbol{e}_2 e^{i(kz-\omega t+\delta)}. \tag{5.69}$$

Note that we usually discuss the polarization characteristics of electromagnetic wave only by considering electric waves. We emphasize that an electric wave and concomitant magnetic wave share the same phase in a uniform and infinite dielectric media. A reason why the electric wave represents an electromagnetic

wave is partly because optical application is mostly made in a non-magnetic substance such as glass, water, plastics, and most of semiconductors.

Let us view temporal change of $\boldsymbol{E}$ at a fixed point $\boldsymbol{x} = 0; x = y = z = 0$. Then, taking a real part of (5.69), $x$- and $y$-components of $\boldsymbol{E}$; i.e., $E_x$ and $E_y$ are expressed as

$$E_x = E_1 \cos(-\omega t) \quad \text{and} \quad E_y = E_2 \cos(-\omega t + \delta). \tag{5.70}$$

First, let us briefly think of the case where $\delta = 0$. Eliminating $t$, we have

$$E_y = \frac{E_2}{E_1} E_x. \tag{5.71}$$

This is an equation of a straight line. The resulting electric field $\boldsymbol{E}$ is called a linearly polarized light accordingly. That is, when we are observing the electric field of the relevant light at the origin, the field is oscillating along the straight line described by (5.71) with the origin centrally located of the oscillating field. If $\delta = \pi$, we have

$$E_y = -\frac{E_2}{E_1} E_x.$$

This gives a straight line as well. Therefore, if we wish to seek the relationship between $E_x$ and $E_y$, it suffices to examine it as a function of $\delta$ in a region of $-\frac{\pi}{2} \leq \delta \leq \frac{\pi}{2}$.

(i) Case I: $E_1 \neq E_2$.

Let us consider the case where $\delta \neq 0$ in (5.70). Rewriting the second equation of (5.70) and inserting the first equation into it so that we can eliminate $t$, we have

$$E_y = E_2(\cos\omega t \cos\delta + \sin\omega t \sin\delta) = E_2\left(\cos\delta \frac{E_x}{E_1} \pm \sin\delta \sqrt{1 - \frac{E_x^2}{E_1^2}}\right).$$

Rearranging terms of the above equation, we have

$$\frac{E_y}{E_2} - (\cos\delta)\frac{E_x}{E_1} = \pm(\sin\delta)\sqrt{1 - \frac{E_x^2}{E_1^2}}. \tag{5.72}$$

Squaring both sides of (5.72) and arranging the equation, we get

$$\frac{E_x^2}{E_1^2 \sin^2\delta} - \frac{2(\cos\delta)E_x E_y}{E_1 E_2 \sin^2\delta} + \frac{E_y^2}{E_2^2 \sin^2\delta} = 1. \tag{5.73}$$

Using a matrix form, we have

$$(E_x \quad E_y)\begin{pmatrix} \frac{1}{E_1^2\sin^2\delta} & -\frac{\cos\delta}{E_1E_2\sin^2\delta} \\ -\frac{\cos\delta}{E_1E_2\sin^2\delta} & \frac{1}{E_2^2\sin^2\delta} \end{pmatrix}\begin{pmatrix} E_x \\ E_y \end{pmatrix} = 1. \tag{5.74}$$

Note that the above matrix is real symmetric. In that case, to examine properties of the matrix, we calculate its determinant along with principal minors. The principal minor means a minor with respect to a diagonal element. In this case, two principal minors are $\frac{1}{E_2^2\sin^2\delta}$ and $\frac{1}{E_1^2\sin^2\delta}$. Also we have

$$\begin{vmatrix} \frac{1}{E_1^2\sin^2\delta} & -\frac{\cos\delta}{E_1E_2\sin^2\delta} \\ -\frac{\cos\delta}{E_1E_2\sin^2\delta} & \frac{1}{E_2^2\sin^2\delta} \end{vmatrix} = \frac{1}{E_1^2E_2^2\sin^2\delta}. \tag{5.75}$$

Evidently, two principal minors as well as a determinant are all positive ($\delta \neq 0$). In this case, the (2, 2) matrix of (5.74) is said to be *positive definite*. The related discussion will be given in Part III. The positive definiteness means that in a quadratic form described by (5.74), LHS takes a positive value for *any* real number $E_x$ and $E_y$ except a unique case where $E_x = E_y = 0$, which renders LHS zero. The positive definiteness of a matrix ensures the existence of positive eigenvalues with the said matrix.

Let us consider a real symmetric (2, 2) matrix that has positive principal minors and a positive determinant in a general case. Let such a matrix $M$ be

$$M = \begin{pmatrix} a & c \\ c & b \end{pmatrix},$$

where $a, b > 0$ and $\det M > 0$; i.e., $ab - c^2 > 0$. Let a corresponding quadratic form be $Q$. Then, we have

$$Q = (x \quad y)\begin{pmatrix} a & c \\ c & b \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = ax^2 + 2cyx + by^2 = a\left[\left(\left(x + \frac{cy}{a}\right)^2 - \frac{c^2y^2}{a^2} + \frac{aby^2}{a^2}\right)\right]$$
$$= a\left[\left(x + \frac{cy}{a}\right)^2 + \frac{y^2}{a^2}(ab - c^2)\right].$$

Thus, $Q \geq 0$ for any real numbers $x$ and $y$. We seek a condition under which $Q = 0$. We readily find that with $M$ that has the above properties, only $x = y = 0$ makes $Q = 0$. Thus, $M$ is positive definite. We will deal with this issue from a more general standpoint in Part III.

In general, it is pretty complicated to seek eigenvalues and corresponding eigenvectors in the above case. Yet, we can extract important information from (5.74). The eigenvalues $\lambda$ are estimated as follows:

$$\lambda = \frac{E_1^2 + E_2^2 \pm \sqrt{(E_1^2 + E_2^2)^2 - 4E_1^2E_2^2\sin^2\delta}}{2E_1^2E_2^2\sin^2\delta}. \tag{5.76}$$

Notice that $\lambda$ in (5.76) represents two different positive eigenvalues. It is because an inside of the square root is rewritten by

$$(E_1^2 - E_2^2)^2 + 4E_1^2E_2^2\cos^2\delta > 0 (\delta \neq \pm\pi/2).$$

Also we have

$$E_1^2 + E_2^2 > \sqrt{(E_1^2 + E_2^2)^2 - 4E_1^2E_2^2\sin^2\delta}.$$

These clearly show that the quadratic form of (5.74) gives an ellipse (i.e., elliptically polarized light). Because of the presence of the second term of LHS of (5.73), both the major and minor axes of the ellipse are tilted and diverted from the $x$- and $y$-axes.

Let us inspect the ellipse described by (5.74). Inserting $E_x = E_1$ obtained at $t = 0$ in (5.70) into (5.73) and solving a quadratic equation with respect to $E_y$, we get $E_y$ as a double root such that

$$E_y = E_2 \cos\delta.$$

Similarly putting $E_y = E_2$ in (5.73), we have

$$E_x = E_1 \cos\delta.$$

These results show that an ellipse described by (5.73) or (5.74) is internally tangent to a rectangle as depicted in Fig. 5.6a. Equation (5.69) shows that the electromagnetic wave is propagated toward the positive direction of the $z$-axis. Therefore, in Fig. 5.6a, we are peeking into the oncoming wave from the bottom of a plane of paper at a certain position of $z =$ constant. We set the constant $= 0$. Then, we find that at $t = 0$ the electric field is represented by the point $P$ $(E_x = E_1, E_y = E_2\cos\delta)$; see Fig. 5.6a. From (5.70), if $\delta > 0$, $P$ traces the ellipse counterclockwise. It reaches a maximum point of $E_y = E_2$ at $t = \delta/2\omega$. Since the trace of electric field forms an ellipse as in Fig. 5.6, the associated light is said to be an elliptically polarized light. If $\delta < 0$ in (5.70), on the other hand, $P$ traces the ellipse clockwise.

In a special case of $\delta = \pi/2$, the second term of (5.73) vanishes and we have a simple form described as

$$\frac{E_x^2}{E_1^2} + \frac{E_y^2}{E_2^2} = 1. \tag{5.77}$$

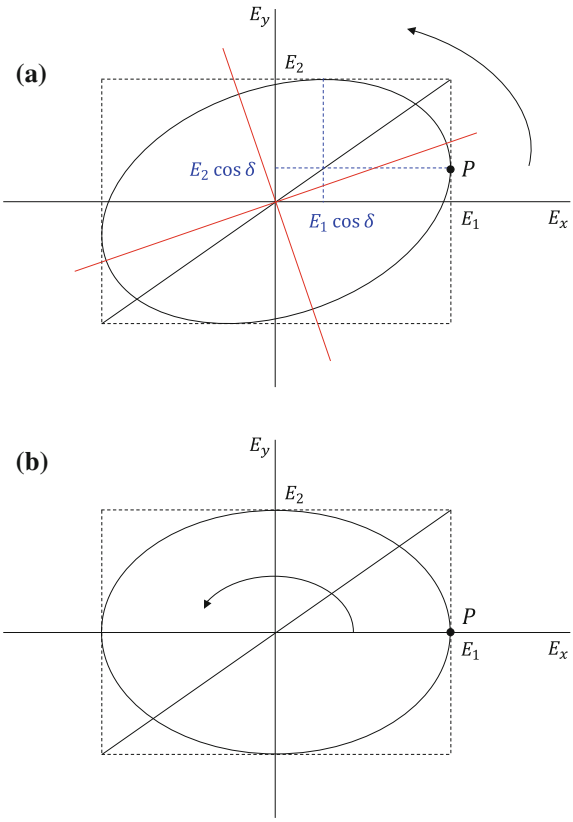


**Fig. 5.6** Trace of an electric field of an elliptically polarized light. **a** The trace is internally tangent to a rectangle of $2E_1 \times 2E_2$. In the case of $\delta > 0$, starting from $P$ at $t = 0$, the coordinate point representing the electric field traces the ellipse counterclockwise with time. **b** The trace of an elliptically polarized light for $\delta = \pi/2$

Thus, the principal axes of the ellipse coincide with the $x$- and $y$-axes. On the basis of (5.70), we see from Fig. 5.6b that starting from $P$ at $t = 0$, again the coordinate point representing the electric field traces the ellipse counterclockwise with time; see the curved arrow of Fig. 5.6b. If $\delta < 0$, the coordinate point traces the ellipse clockwise with time.

(ii) Case II: $E_1 = E_2$.

Now, let us consider a simple but important case. When $E_1 = E_2$, (5.73) is simplified to be

$$E_x^2 - 2\cos\delta E_x E_y + E_y^2 = E_1^2 \sin^2\delta. \tag{5.78}$$

Using a matrix form, we have

$$\begin{pmatrix} E_x & E_y \end{pmatrix} \begin{pmatrix} 1 & -\cos\delta \\ -\cos\delta & 1 \end{pmatrix} \begin{pmatrix} E_x \\ E_y \end{pmatrix} = E_1^2 \sin^2\delta. \tag{5.79}$$

We obtain eigenvalues $\lambda$ of the matrix of (5.79) such that

$$\lambda = 1 \pm |\cos\delta|. \tag{5.80}$$

Setting $-\frac{\pi}{2} \leq \delta \leq \frac{\pi}{2}$, we have

$$\lambda = 1 \pm \cos\delta. \tag{5.81}$$

The corresponding normalized eigenvectors $v_1$ and $v_2$ (as a column vector) are

$$v_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix} \quad \text{and} \quad v_2 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}. \tag{5.82}$$

Thus, we have a diagonalizing unitary matrix $P$ such that

$$P = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}. \tag{5.83}$$

Defining the above matrix appearing in (5.79) as $A$ such that

$$A = \begin{pmatrix} 1 & -\cos\delta \\ -\cos\delta & 1 \end{pmatrix}, \tag{5.84}$$

we obtain

$$P^{-1}AP = \begin{pmatrix} 1+\cos\delta & 0 \\ 0 & 1-\cos\delta \end{pmatrix}. \tag{5.85}$$

Notice that eigenvalues $(1+\cos\delta)$ and $(1-\cos\delta)$ are both positive as expected.

Rewriting (5.79), we have

$$\begin{aligned} &(E_x \quad E_y)PP^{-1}\begin{pmatrix} 1 & -\cos\delta \\ -\cos\delta & 1 \end{pmatrix}PP^{-1}\begin{pmatrix} E_x \\ E_y \end{pmatrix} \\ &= (E_x \quad E_y)P\begin{pmatrix} 1+\cos\delta & 0 \\ 0 & 1-\cos\delta \end{pmatrix}P^{-1}\begin{pmatrix} E_x \\ E_y \end{pmatrix} = E_1^2\sin^2\delta. \end{aligned} \tag{5.86}$$

Here, let us define new coordinates such that

$$\begin{pmatrix} u \\ v \end{pmatrix} \equiv P^{-1}\begin{pmatrix} E_x \\ E_y \end{pmatrix}. \tag{5.87}$$

This coordinate transformation corresponds to the transformation of basis vectors $(\boldsymbol{e}_1\boldsymbol{e}_2)$ such that

$$(\boldsymbol{e}_1\boldsymbol{e}_2)\begin{pmatrix}E_x\\E_y\end{pmatrix} = (\boldsymbol{e}_1\boldsymbol{e}_2)PP^{-1}\begin{pmatrix}E_x\\E_y\end{pmatrix} = (\boldsymbol{e}'_1\boldsymbol{e}'_2)\begin{pmatrix}u\\v\end{pmatrix}, \tag{5.88}$$

where new basis vectors $(\boldsymbol{e}'_1\boldsymbol{e}'_2)$ are given by

$$(\boldsymbol{e}'_1\boldsymbol{e}'_2) = (\boldsymbol{e}_1\boldsymbol{e}_2)P = \left(\frac{1}{\sqrt{2}}\boldsymbol{e}_1 - \frac{1}{\sqrt{2}}\boldsymbol{e}_2 \quad \frac{1}{\sqrt{2}}\boldsymbol{e}_1 + \frac{1}{\sqrt{2}}\boldsymbol{e}_2\right). \tag{5.89}$$

The coordinate system along with the basis vectors are depicted in Fig. 5.7. The relevant discussion will again appear in Part III.

Substituting (5.87) for (5.86) and rearranging terms, we get

$$\frac{u^2}{E_1^2(1-\cos\delta)} + \frac{v^2}{E_1^2(1+\cos\delta)} = 1. \tag{5.90}$$

Equation (5.90) indicates that a major axis and minor axis are $E_1\sqrt{1+\cos\delta}$ and $E_1\sqrt{1-\cos\delta}$, respectively. When $\delta = \pm\pi/2$, (5.90) becomes

$$\frac{u^2}{E_1^2} + \frac{v^2}{E_1^2} = 1. \tag{5.91}$$

This represents a circle. For this reason, the wave described by (5.91) is called a circularly polarized light. In (5.90) where $\delta \neq \pm\pi/2$, the wave is said to be an elliptically polarized light. Thus, we have linearly, elliptically, and circularly polarized lights depending on a magnitude of $\delta$.

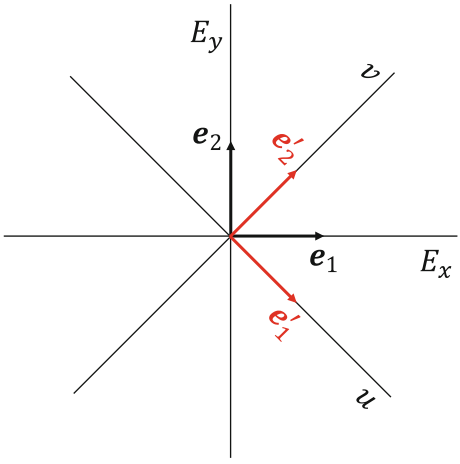


**Fig. 5.7** Relationship between the basis vectors $(\boldsymbol{e}_1\boldsymbol{e}_2)$ and $(\boldsymbol{e}'_1\boldsymbol{e}'_2)$ in the case of $E_1 = E_2$; see text

Let us closely examine characteristics of the elliptically and circularly polarized lights in the case of $E_1 = E_2$. When $t = 0$, from (5.70) we have

$$E_x = E_1 \quad \text{and} \quad E_y = E_1 \cos\delta. \tag{5.92}$$

This coordinate point corresponds to $A_1$ whose $E_x$ coordinate is $E_1$ (see Fig. 5.8a). In the case of $\Delta t = \delta/2\omega$, $E_x = E_y = E_1 \cos(\pm\delta/2)$. This point corresponds to $A_2$ in Fig. 5.8a. We have

$$\sqrt{E_x^2 + E_y^2} = \sqrt{2}E_1 \cos(\delta/2) = E_1\sqrt{1 + \cos\delta}. \tag{5.93}$$

This is equal to the major axis as anticipated. With $t = \Delta t$,

$$E_x = E_1 \cos(-\omega\Delta t) \quad \text{and} \quad E_y = E_1 \cos(-\omega\Delta t + \delta). \tag{5.94}$$

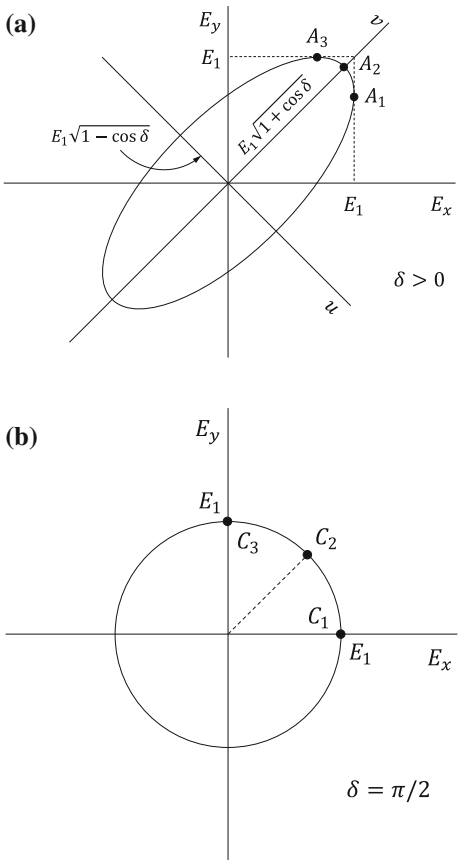


**Fig. 5.8** Polarized feature of light in the case of $E_1 = E_2$. **a** If $\delta > 0$, the electric field traces an ellipse from $A_1$ via $A_2$ to $A_3$ (see text). **b** If $\delta = \pi/2$, the electric field traces a circle from $C_1$ via $C_2$ to $C_3$ (left-circularly polarized light)

Notice that $E_y$ takes a maximum $E_1$ when $\Delta t = \delta/\omega$. Consequently, if $\delta$ takes a positive value, $E_y$ takes a maximum $E_1$ for a positive $\Delta t$, as is similarly the case with Fig. 5.6a. At that time, $E_x = E_1 \cos(-\delta) < E_1$. This point corresponds to $A_3$ in Fig. 5.8a. As a result, the electric field traces the ellipse counterclockwise with time, as in the case of Fig. 5.6. If $\delta$ takes a negative value, on the other hand, the field traces the ellipse clockwise.

If $\delta = \pm\pi/2$, in (5.94) we have

$$E_x = E_1 \cos(-\omega t) \quad \text{and} \quad E_y = E_1 \cos\left(-\omega t \pm \frac{\pi}{2}\right). \tag{5.95}$$

We examine the case of $\delta = \pi/2$ first. In this case, when $t = 0$, $E_x = E_1$ and $E_y = 0$ [Point $C_1$ in Fig. 5.8b]. If $\omega t = \pi/4$, $E_x = E_y = 1/\sqrt{2}$ [Point $C_2$ in Fig. 5.8b]. In turn, if $\omega t = \pi/2$, $E_x = 0$ and $E_y = E_1$ (Point $C_3$). Again the electric field traces the circle counterclockwise. In this situation, we see the light from above the $z$-axis. In other words, we are viewing the light against the direction of its propagation. The wave is said to be *left*-circularly polarized and have *positive* helicity. In contrast, when $\delta = -\pi/2$, starting from Point $C_1$, the electric field traces the circle clockwise. That light is said to be *right*-circularly polarized and have *negative* helicity.

With the left-circularly polarized light, (5.69) can be rewritten as

$$\boldsymbol{E} = \boldsymbol{E}_1 + \boldsymbol{E}_2 = E_1(\boldsymbol{e}_1 + i\boldsymbol{e}_2)\mathrm{e}^{i(kz-\omega t)}. \tag{5.96}$$

Therefore, a complex vector $(\boldsymbol{e}_1 + i\boldsymbol{e}_2)$ characterizes the left-circular polarization. On the other hand, $(\boldsymbol{e}_1 - i\boldsymbol{e}_2)$ characterizes the right-circular polarization. To normalize them, it is convenient to use the following vectors as in the case of Sect. 4.3 [3].

$$\boldsymbol{e}_+ \equiv \frac{1}{\sqrt{2}}(\boldsymbol{e}_1 + i\boldsymbol{e}_2) \quad \text{and} \quad \boldsymbol{e}_- \equiv \frac{1}{\sqrt{2}}(\boldsymbol{e}_1 - i\boldsymbol{e}_2), \tag{4.45}$$

In the case of $\delta = 0$, we have a linearly polarized light. For this, the points $A_1$, $A_2$, and $A_3$ coalesce to be a point on a straight line of $E_y = E_x$.

## References

1. Arfken GB, Weber HJ, Harris FE (2013) Mathematical methods for physicists, 7th edn. Academic Press, Waltham
2. Pain HJ (2005) The physics of vibrations and waves, 6th edn. Wiley, Chichester
3. Jackson JD (1999) Classical electrodynamics. Wiley, New York

# Chapter 6
# Reflection and Transmission of Electromagnetic Waves in Dielectric Media

In Chap. 5, we considered the propagation of electromagnetic waves in an infinite uniform dielectric medium. In this chapter, we think of a situation where two (or more) dielectrics are in contact with each other at a plane interface. When two dielectric media adjoin each other with an interface, propagating electromagnetic waves are partly reflected by the interface and partly transmitted beyond the interface. We deal with these phenomena in terms of characteristic impedance of the dielectric media. In the case of an oblique incidence of a wave, we categorize it into a transverse electric (TE) wave and transverse magnetic (TM) wave. If a thin plate of a dielectric is sandwiched by a couple of metal sheets, the electromagnetic wave is confined within the dielectric. In this case, the propagating mode of the wave differs from that of a wave propagating in a free space (i.e., a space filled by a three-dimensionally infinite dielectric medium). If a thin plate of a dielectric having a large refractive index is sandwiched by a couple of dielectrics with a smaller refractive index, the electromagnetic wave is also confined within the dielectric with a larger index. In this case, we have to take account of the total reflection that causes a phase change upon the reflection. We deal with such specific modes of the electromagnetic wave propagation. These phenomena are treated both from a basic aspect and from a point of view of device application. The relevant devices are called waveguides in optics.

## 6.1 Electromagnetic Fields at an Interface

We start with examining a condition of an electromagnetic field at the plane interface. Suppose that two semi-infinite dielectric media D1 and D2 are in contact with each other at a plane interface. Let us take a small rectangle $S$ that strides the interface (see Fig. 6.1). Taking a surface integral of both sides of (5.28) over the strip, we have



S. Hotta, *Mathematical Physical Chemistry*,
https://doi.org/10.1007/978-981-10-7671-8_6

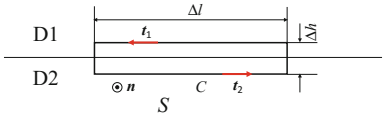


**Fig. 6.1** A small rectangle $S$ that strides an interface formed by two semi-infinite dielectric media of D1 and D2. Let a curve $C$ be a closed loop surrounding the rectangle $S$. A unit vector $\boldsymbol{n}$ is directed to a normal of $S$. Unit vectors $\boldsymbol{t}_1$ and $\boldsymbol{t}_2$ are directed to a tangential line of the interface plane

$$\int_S \mathrm{rot}\,\boldsymbol{E}\cdot\boldsymbol{n}\,\mathrm{d}S+\int_S \frac{\partial \boldsymbol{B}}{\partial t}\cdot\boldsymbol{n}\,\mathrm{d}S=0, \tag{6.1}$$

where $\boldsymbol{n}$ is a unit vector directed to a normal of $S$ as shown. Applying Stokes' theorem to the first term of (6.1), we get

$$\oint_C \boldsymbol{E}\cdot\mathrm{d}\boldsymbol{l}+\frac{\partial \boldsymbol{B}}{\partial t}\cdot\boldsymbol{n}\Delta l\Delta h=0. \tag{6.2}$$

With the line integral of the first term, $C$ is a closed loop surrounding the rectangle $S$ and $\mathrm{d}\boldsymbol{l}=\boldsymbol{t}\,\mathrm{d}l$, where $\boldsymbol{t}$ is a unit vector directed toward the tangential direction of $C$ (see $\boldsymbol{t}_1$ and $\boldsymbol{t}_2$ in Fig. 6.1). The line integration is performed such that $C$ is followed counterclockwise in the direction of $\boldsymbol{t}$.

Figure 6.2 gives an intuitive diagram that explains Stokes' theorem. The diagram shows an overview of a surface $S$ encircled by a closed curve $C$. Suppose that we have a spiral vector field $\boldsymbol{E}$ represented by arrowed circles as shown. In that case, rot $\boldsymbol{E}$ is directed toward the upper side of the plane of paper in the individual fragments. A summation of $\mathrm{rot}\,\boldsymbol{E}\cdot\boldsymbol{n}\,\mathrm{d}S$ forms a surface integral covering $S$. Meanwhile, the arrows of adjacent fragments cancel out each other and only the components on the periphery (i.e., the curve $C$) are nonvanishing (see Fig. 6.2). Thus, the surface integral of rot $\boldsymbol{E}$ is equivalent to the line integral of $\boldsymbol{E}$. Accordingly, we get Stokes' theorem described by [1]

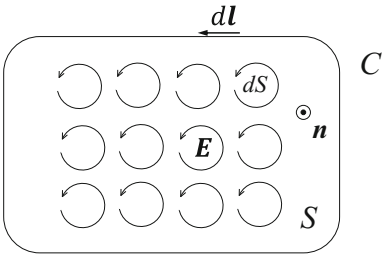


**Fig. 6.2** Diagram that intuitively explains Stokes' theorem. In the diagram, a surface $S$ is encircled by a closed curve $C$. An infinitesimal portion of $C$ is denoted by $\mathrm{d}\boldsymbol{l}$. The surface $S$ is pertinent to the surface integration. Spiral vector field $\boldsymbol{E}$ is present on and near $S$

$$\int_S \mathrm{rot}\,\boldsymbol{E}\cdot\boldsymbol{n}\,\mathrm{d}S = \oint_C \boldsymbol{E}\cdot\mathrm{d}\boldsymbol{l}. \tag{6.3}$$

Returning back to Fig. 6.1 and taking $\Delta h \to 0$, we have $\frac{\partial \boldsymbol{B}}{\partial t}\cdot\boldsymbol{n}\Delta l\Delta h \to 0$. Then, the second term of (6.2) vanishes and we get

$$\oint_C \boldsymbol{E}\cdot\mathrm{d}\boldsymbol{l} = 0.$$

This implies that

$$\Delta l\,(\boldsymbol{E}_1\cdot\boldsymbol{t}_1 + \boldsymbol{E}_2\cdot\boldsymbol{t}_2) = 0,$$

where $\boldsymbol{E}_1$ and $\boldsymbol{E}_2$ represent the electric field in the dielectrics D1 and D2 close to the interface, respectively. Considering $\boldsymbol{t}_2 = -\boldsymbol{t}_1$ and putting $\boldsymbol{t}_1 = \boldsymbol{t}$, we get

$$(\boldsymbol{E}_1 - \boldsymbol{E}_2)\cdot\boldsymbol{t} = 0, \tag{6.4}$$

where $\boldsymbol{t}$ represents a unit vector in the direction of a tangential line of the interface plane. Equation (6.4) means that the tangential components of the electric field are continuous on both sides of the interface. We obtain a similar result with the magnetic field. This can be shown by taking a surface integral of both sides of (5.29) as well. As a result, we get

$$(\boldsymbol{H}_1 - \boldsymbol{H}_2)\cdot\boldsymbol{t} = 0, \tag{6.5}$$

where $\boldsymbol{H}_1$ and $\boldsymbol{H}_2$ represent the magnetic field in D1 and D2 close to the interface, respectively. Hence, from (6.5) the tangential components of the magnetic field are continuous on both sides of the interface as well.

## 6.2 Basic Concepts Underlying Phenomena

When an electromagnetic wave is incident upon an interface of dielectrics, its reflection and transmission (refraction) take place at the interface. We address a question of how the nature of the dielectrics and the conditions dealt with in the previous section are associated with the optical phenomena. When we deal with the problem, we assume non-absorbing media. Notice that the complex wavenumber vector is responsible for an absorbing medium along with a complex index of refraction. Nonetheless, our approach is useful to discuss related problems in the absorbing media. Characteristic impedance plays a key role in the reflection and transmission of light.

We represent a field (either electric or magnetic) of the incident, reflected, and transmitted (or refracted) waves by $\boldsymbol{F}_i$, $\boldsymbol{F}_r$, and $\boldsymbol{F}_t$, respectively. We call a dielectric of the incidence side (and, hence, reflection side) D1 and another dielectric of the transmission side D2. The fields are described by

$$\boldsymbol{F}_i = F_i \boldsymbol{\varepsilon}_i \mathrm{e}^{i(\boldsymbol{k}_i \cdot \boldsymbol{x} - \omega t)}, \tag{6.6}$$

$$\boldsymbol{F}_r = F_r \boldsymbol{\varepsilon}_r \mathrm{e}^{i(\boldsymbol{k}_r \cdot \boldsymbol{x} - \omega t)}, \tag{6.7}$$

$$\boldsymbol{F}_t = F_t \boldsymbol{\varepsilon}_t \mathrm{e}^{i(\boldsymbol{k}_t \cdot \boldsymbol{x} - \omega t)}, \tag{6.8}$$

where $F_i$, $F_r$, and $F_t$ denote an amplitude of the field; $\boldsymbol{\varepsilon}_i$, $\boldsymbol{\varepsilon}_r$, and $\boldsymbol{\varepsilon}_t$ represent a unit vector of polarization direction, i.e., the direction along which the field oscillates; $\boldsymbol{k}_i$, $\boldsymbol{k}_r$, and $\boldsymbol{k}_t$ are wavenumber vectors such that $\boldsymbol{k}_i \perp \boldsymbol{\varepsilon}_i$, $\boldsymbol{k}_r \perp \boldsymbol{\varepsilon}_r$, and $\boldsymbol{k}_t \perp \boldsymbol{\varepsilon}_t$. These wavenumber vectors represent the propagation directions of individual waves. In (6.6) to (6.8), indices of $i$, $r$, and $t$ stand for incidence, reflection, and transmission, respectively.

Let $\boldsymbol{x}_s$ be an arbitrary position vector at the interface between the dielectrics. Also, let $\boldsymbol{t}$ be a unit vector paralleling the interface. Thus, tangential components of the field are described as

$$F_{i_t} = F_i (\boldsymbol{t} \cdot \boldsymbol{\varepsilon}_i) \mathrm{e}^{i(\boldsymbol{k}_i \cdot \boldsymbol{x}_s - \omega t)}, \tag{6.9}$$

$$F_{r_t} = F_r (\boldsymbol{t} \cdot \boldsymbol{\varepsilon}_r) \mathrm{e}^{i(\boldsymbol{k}_r \cdot \boldsymbol{x}_s - \omega t)}, \tag{6.10}$$

$$F_{t_t} = F_t (\boldsymbol{t} \cdot \boldsymbol{\varepsilon}_t) \mathrm{e}^{i(\boldsymbol{k}_t \cdot \boldsymbol{x}_s - \omega t)}. \tag{6.11}$$

Note that $F_{i_t}$ and $F_{r_t}$ represent the field in D1 just close to the interface and that $F_{t_t}$ denotes the field in D2 just close to the interface. Thus, in light of (6.4) and (6.5), we have

$$F_{i_t} + F_{r_t} = F_{t_t}. \tag{6.12}$$

Notice that (6.12) holds with any position $\boldsymbol{x}_s$ and any time $t$.

Let us think of elementary calculation of exponential functions or exponential polynomials and the relationship between individual coefficients and exponents. With respect to two functions $\mathrm{e}^{ikx}$ and $\mathrm{e}^{ik'x}$, we have two alternatives according to a value Wronskian takes. Here, Wronskian $W$ is expressed as

$$W = \begin{vmatrix} \mathrm{e}^{ikx} & \mathrm{e}^{ik'x} \\ (\mathrm{e}^{ikx})' & (\mathrm{e}^{ik'x})' \end{vmatrix} = -i(k - k')\, \mathrm{e}^{i(k+k')x}. \tag{6.13}$$

(i) $W \neq 0$ if and only if $k \neq k'$. In this case, $\mathrm{e}^{ikx}$ and $\mathrm{e}^{ik'x}$ are said to be linearly independent. That is, on condition of $k \neq k'$, for any $x$ we have

$$a\mathrm{e}^{ikx} + b\mathrm{e}^{ik'x} = 0 \Leftrightarrow a = b = 0. \tag{6.14}$$

(ii) $W = 0$ if $k = k'$. In that case, we have

$$a\mathrm{e}^{ikx} + b\mathrm{e}^{ik'x} = (a+b)\mathrm{e}^{ikx} = 0 \Leftrightarrow a + b = 0.$$

Notice that $\mathrm{e}^{ikx}$ never vanishes with any $x$. To conclude, if we think of an equation of an exponential polynomial

$$a\mathrm{e}^{ikx} + b\mathrm{e}^{ik'x} = 0,$$

we have two alternatives regarding the coefficients: One is a trivial case of $a = b = 0$, and the other is $a + b = 0$.

Next, with respect to $\mathrm{e}^{ik_1x}$, and $\mathrm{e}^{ik_2x}$, and $\mathrm{e}^{ik_3x}$, similarly we have

$$\begin{aligned} W &= \begin{vmatrix} \mathrm{e}^{ik_1x} & \mathrm{e}^{ik_2x} & \mathrm{e}^{ik_3x} \\ (\mathrm{e}^{ik_1x})' & (\mathrm{e}^{ik_2x})' & (\mathrm{e}^{ik_3x})' \\ (\mathrm{e}^{ik_1x})'' & (\mathrm{e}^{ik_2x})'' & (\mathrm{e}^{ik_3x})'' \end{vmatrix} \\ &= -i(k_1 - k_2)(k_2 - k_3)(k_3 - k_1)\mathrm{e}^{i(k_1+k_2+k_3)x}, \end{aligned} \tag{6.15}$$

where $W \neq 0$ if and only if $k_1 \neq k_2$, $k_2 \neq k_3$, and $k_3 \neq k_1$. That is, on this condition for any $x$ we have

$$a\mathrm{e}^{ik_1x} + b\mathrm{e}^{ik_2x} + c\mathrm{e}^{ik_3x} = 0 \Leftrightarrow a = b = c = 0. \tag{6.16}$$

If the three exponential functions are linearly dependent, at least two of $k_1$, $k_2$, and $k_3$ are equal to each other, and vice versa. On this condition, again consider a following equation of an exponential polynomial:

$$a\mathrm{e}^{ik_1x} + b\mathrm{e}^{ik_2x} + c\mathrm{e}^{ik_3x} = 0. \tag{6.17}$$

Without loss of generality, we assume that $k_1 = k_2$. Then, we have

$$a\mathrm{e}^{ik_1x} + b\mathrm{e}^{ik_2x} + c\mathrm{e}^{ik_3x} = (a+b)\mathrm{e}^{ik_1x} + c\mathrm{e}^{ik_3x} = 0.$$

If $k_1 \neq k_3$, we must have

$$a + b = 0 \quad \text{and} \quad c = 0 \tag{6.18}$$

If, on the other hand, $k_1 = k_3$, i.e., $k_1 = k_2 = k_3$, we have

$$ae^{ik_1x} + be^{ik_2x} + ce^{ik_3x} = (a+b+c)e^{ik_1x} = 0.$$

That is, we have

$$a + b + c = 0. \tag{6.19}$$

Consequently, we must have $k_1 = k_2 = k_3$ so that we can get three nonzero coefficients $a, b,$ and $c$.

Returning to (6.12), its full description is

$$F_i(\boldsymbol{t}\cdot\boldsymbol{\varepsilon}_i)e^{i(\boldsymbol{k}_i\cdot\boldsymbol{x}_s-\omega t)} + F_r(\boldsymbol{t}\cdot\boldsymbol{\varepsilon}_r)e^{i(\boldsymbol{k}_r\cdot\boldsymbol{x}_s-\omega t)} - F_t(\boldsymbol{t}\cdot\boldsymbol{\varepsilon}_t)e^{i(\boldsymbol{k}_t\cdot\boldsymbol{x}_s-\omega t)} = 0. \tag{6.20}$$

Again, (6.20) must hold with any position $\boldsymbol{x}_s$ and any time $t$. Meanwhile, for (6.20) to have a physical meaning, we should have

$$F_i(\boldsymbol{t}\cdot\boldsymbol{\varepsilon}_i) \neq 0,\ F_r(\boldsymbol{t}\cdot\boldsymbol{\varepsilon}_r) \neq 0 \text{ and } F_t(\boldsymbol{t}\cdot\boldsymbol{\varepsilon}_r) \neq 0. \tag{6.21}$$

On the basis of the above consideration, we must have the following two relations:

$$\begin{aligned} &\boldsymbol{k}_i\cdot\boldsymbol{x}_s - \omega t = \boldsymbol{k}_r\cdot\boldsymbol{x}_s - \omega t = \boldsymbol{k}_t\cdot\boldsymbol{x}_s - \omega t \quad \text{or} \\ &\boldsymbol{k}_i\cdot\boldsymbol{x}_s = \boldsymbol{k}_r\cdot\boldsymbol{x}_s = \boldsymbol{k}_t\cdot\boldsymbol{x}_s, \end{aligned} \tag{6.22}$$

and

$$\begin{aligned} &F_i(\boldsymbol{t}\cdot\boldsymbol{\varepsilon}_i) + F_r(\boldsymbol{t}\cdot\boldsymbol{\varepsilon}_r) - F_t(\boldsymbol{t}\cdot\boldsymbol{\varepsilon}_t) = 0 \quad \text{or} \\ &F_i(\boldsymbol{t}\cdot\boldsymbol{\varepsilon}_i) + F_r(\boldsymbol{t}\cdot\boldsymbol{\varepsilon}_r) = F_t(\boldsymbol{t}\cdot\boldsymbol{\varepsilon}_t). \end{aligned} \tag{6.23}$$

In this way, we are able to obtain a relation among amplitudes of the fields of incidence, reflection, and transmission. Notice that we get both the relations between exponents and coefficients at once.

First, let us consider (6.22). Suppose that the incident light ($\boldsymbol{k}_i$) is propagated in a dielectric medium D1 in parallel to the $zx$-plane and that the interface is the $xy$-plane (see Fig. 6.3). Also suppose that at the interface the light is reflected partly back to D1 and transmitted (or refracted) partly into another dielectric medium D2. In Fig. 6.3, $\boldsymbol{k}_i$, $\boldsymbol{k}_r$, and, $\boldsymbol{k}_t$ represent the incident, reflected, and transmitted lights that make an angle $\theta$, $\theta'$, and $\phi$ with the $z$-axis, respectively. Then we have

$$\boldsymbol{k}_i = (\boldsymbol{e}_1\ \boldsymbol{e}_2\ \boldsymbol{e}_3)\begin{pmatrix} k_i\ \sin\ \theta \\ 0 \\ -k_i\ \cos\ \theta \end{pmatrix}, \tag{6.24}$$

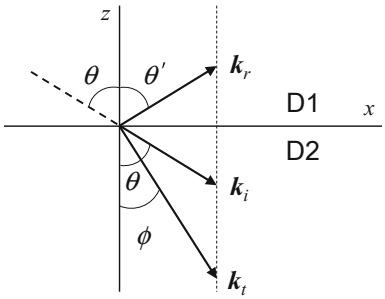


**Fig. 6.3** Geometry of the incident, reflected, and transmitted lights. We assume that the light is incident from a dielectric medium D1 toward another medium D2. The wavenumber vectors $\boldsymbol{k}_i$, $\boldsymbol{k}_r$, and $\boldsymbol{k}_t$ represent the incident, reflected, and transmitted (or refracted) lights with an angle $\theta$, $\theta'$, and $\phi$, respectively. Note here that we did not assume the equality of $\theta$ and $\theta'$ (see text)

$$\boldsymbol{x}_s = (\boldsymbol{e}_1\ \boldsymbol{e}_2\ \boldsymbol{e}_3)\begin{pmatrix} x \\ y \\ 0 \end{pmatrix}, \tag{6.25}$$

where $\theta$ is said to be an incidence angle. A plane formed by $\boldsymbol{k}_i$ and a normal to the interface is called a plane of incidence (or incidence plane). In Fig. 6.3, the $zx$-plane forms the incidence plane. From (6.24) and (6.25), we have

$$\boldsymbol{k}_i \cdot \boldsymbol{x}_s = k_i x\ \sin\ \theta, \tag{6.26}$$

$$\boldsymbol{k}_r \cdot \boldsymbol{x}_s = k_{r_x} x + k_{r_y} y, \tag{6.27}$$

$$\boldsymbol{k}_t \cdot \boldsymbol{x}_s = k_{t_x} x + k_{t_y} y, \tag{6.28}$$

where $k_i = |\boldsymbol{k}_i|$; $k_{r_x}$ and $k_{r_y}$ are $x$ and $y$ components of $\boldsymbol{k}_r$; similarly, $k_{t_x}$ and $k_{t_y}$ are $x$ and $y$ components of $\boldsymbol{k}_t$.

Since (6.22) holds with any $x$ and $y$, we have

$$k_i\ \sin\ \theta = k_{r_x} = k_{t_x}, \tag{6.29}$$

$$k_{r_y} = k_{t_y} = 0. \tag{6.30}$$

From (6.30), neither $\boldsymbol{k}_r$ nor $\boldsymbol{k}_t$ has a $y$ component. This means that $\boldsymbol{k}_i$, $\boldsymbol{k}_r$, and $\boldsymbol{k}_t$ are coplanar. That is, the incident, reflected, and transmitted waves are all parallel to the $zx$-plane. Notice that at the beginning we did not assume the coplanarity of those waves. We did not assume the equality of $\theta$ and $\theta'$ either (vide infra). From (6.29) and Fig. 6.3, however, we have

$$k_i \ \sin\ \theta = k_r \ \sin\ \theta' = k_t \ \sin\ \phi, \tag{6.31}$$

where $k_r = |\boldsymbol{k}_r|$ and $k_t = |\boldsymbol{k}_t|$; $\theta'$ and $\phi$ are said to be a reflection angle and a refraction angle, respectively. Thus, the end points of $\boldsymbol{k}_i, \boldsymbol{k}_r$, and $\boldsymbol{k}_t$ are connected on a straight line that parallels the $z$-axis. Figure 6.3 clearly shows it.

Now, we suppose that a wavelength of the electromagnetic wave in D1 is $\lambda_1$ and that in D2 is $\lambda_2$. Since the incident light and reflected light are propagated in D1, we have

$$k_i = k_r = 2\pi/\lambda_1. \tag{6.32}$$

From (6.31) and (6.32), we get

$$\sin\ \theta = \sin\ \theta'. \tag{6.33}$$

Therefore, we have either $\theta = \theta'$ or $\theta' = \pi - \theta$. since $0 < \theta, \theta' < \pi/2$, we have

$$\theta = \theta'. \tag{6.34}$$

Then, returning back to (6.31), we have

$$k_i \sin\theta = k_r \sin\theta = k_t \sin\phi. \tag{6.35}$$

This implies that the components tangential to the interface of $\boldsymbol{k}_i$, $\boldsymbol{k}_r$, and $\boldsymbol{k}_t$ are the same.

Meanwhile, we have

$$k_t = 2\pi/\lambda_2. \tag{6.36}$$

Also we have

$$c = \lambda_0 \nu, \quad v_1 = \lambda_1 \nu, \quad v_2 = \lambda_2 \nu, \tag{6.37}$$

where $v_1$ and $v_2$ are phase velocities of light in D1 and D2, respectively. Since $\nu$ is common to D1 and D2, we have

$$\begin{aligned} &c/\lambda_0 = v_1/\lambda_1 = v_2/\lambda_2 \quad \text{or} \\ &c/v_1 = \lambda_0/\lambda_1 = n_1, \quad c/v_2 = \lambda_0/\lambda_2 = n_2, \end{aligned} \tag{6.38}$$

where $\lambda_0$ is a wavelength in vacuum; $n_1$ and $n_2$ are refractive indices of D1 and D2, respectively. Combining (6.35) with (6.32), (6.36), and (6.38), we have several relations such that

$$\frac{\sin\theta}{\sin\phi} = \frac{k_t}{k_i} = \frac{\lambda_1}{\lambda_2} = \frac{n_2}{n_1} \ (\equiv n), \tag{6.39}$$

where $n$ is said to be a relative refractive index of D2 relative to D1. The relation (6.39) is called Snell's law. Notice that (6.39) reflects the kinematic aspect of light and that this characteristic comes from the exponents of (6.20).

## 6.3 Transverse Electric (TE) Waves and Transverse Magnetic (TM) Waves

On the basis of the above argument, we are now in the position to determine the relations among amplitudes of the electromagnetic fields of waves of incidence, reflection, and transmission. Notice that since we are dealing with non-absorbing media, the relevant amplitudes are real (i.e., positive or negative). In other words, when the phase is retained upon reflection, we have a positive amplitude due to $e^{i0} = 1$. When the phase is reversed upon reflection, on the other hand, we will be treating a negative amplitude due to $e^{i\pi} = -1$. Nevertheless, when we consider the total reflection, we deal with a complex amplitude (vide infra).

We start with the discussion of the vertical incidence of an electromagnetic wave before the general oblique incidence. In Fig. 6.4a, we depict electric fields $\boldsymbol{E}$ and magnetic fields $\boldsymbol{H}$ obtained at a certain moment near the interface. We index, e.g., $\boldsymbol{E}_i$ for the incident field. There we define unit polarization vectors of the electric field $\boldsymbol{\varepsilon}_i$, $\boldsymbol{\varepsilon}_r$, and $\boldsymbol{\varepsilon}_t$ as identical to be $\boldsymbol{e}_1$ (a unit vector in the direction of the $x$-axis). In (6.6), we also define $F_i$ (both electric and magnetic fields) as positive.

We have two cases about a geometry of the fields (see Fig. 6.4). The first case is that all $\boldsymbol{E}_i$, $\boldsymbol{E}_r$, and $\boldsymbol{E}_t$ are directed in the same direction (i.e., the positive direction of the $x$-axis) (see Fig. 6.4a). Another case is that although $\boldsymbol{E}_i$ and $\boldsymbol{E}_t$ are directed in the same direction, $\boldsymbol{E}_r$ is reversed (Fig. 6.4b). In this case, we define $\boldsymbol{E}_r$ as negative. Notice that $\boldsymbol{E}_i$ and $\boldsymbol{E}_t$ are always directed in the same direction and that $\boldsymbol{E}_r$ is directed either in the same direction or in the opposite direction according to the nature of the dielectrics. The situation will be discussed soon.

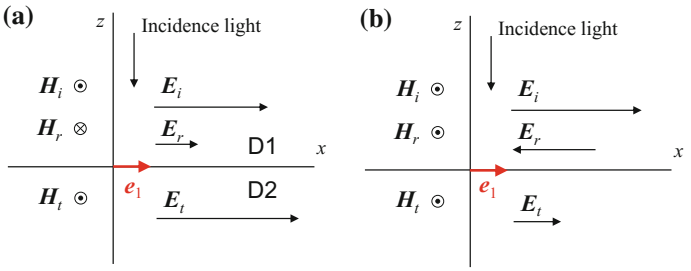


**Fig. 6.4** Geometry of the electromagnetic fields near the interface between dielectric media D1 and D2 in the case of vertical incidence. **a** All $\boldsymbol{E}_i$, $\boldsymbol{E}_r$, and $\boldsymbol{E}_t$ are directed in the same direction $\boldsymbol{e}_1$ (i.e., a unit vector in the positive direction of the $x$-axis). **b** Although $\boldsymbol{E}_i$ and $\boldsymbol{E}_t$ are directed in the same direction, $\boldsymbol{E}_r$ is reversed. In this case, we define $E_r$ as negative

Meanwhile, unit polarization vectors of the magnetic fields are determined by (5.67) for the incident, reflected, and transmitted waves. In Fig. 6.4, the magnetic fields are polarized along the $y$-axis (i.e., perpendicular to the plane of paper). The magnetic fields $\boldsymbol{H}_i$ and $\boldsymbol{H}_t$ are always directed to the same direction as in the case of the electric fields. On the other hand, if the phase of $\boldsymbol{E}_r$ is conserved, the direction of $\boldsymbol{H}_r$ is reversed and vice versa. This converse relationship with respect to the electric and magnetic fields results solely from the requirement that $\boldsymbol{E}$, $\boldsymbol{H}$, and the propagation unit vector $\boldsymbol{n}$ of light must constitute a right-handed system in this order. Notice that $\boldsymbol{n}$ is reversed upon reflection.

Next, let us consider an oblique incidence. With the oblique incidence, electromagnetic waves are classified into two special categories, i.e., transverse electric (TE) waves (or modes) or transverse magnetic (TM) waves (or modes). The TE wave is characterized by the electric field that is perpendicular to the incidence plane, whereas the TM wave is characterized by the magnetic field that is perpendicular to the incidence plane. Here, the incidence plane is a plane that is formed by the propagation direction of the incident light and the normal to the interface of the two dielectrics. Since $\boldsymbol{E}$, $\boldsymbol{H}$, and $\boldsymbol{n}$ form a right-handed system, in the TE wave $\boldsymbol{H}$ lies on the incidence plane. For the same reason, in the TM wave $\boldsymbol{E}$ lies on the incidence plane.

In a general case where a field is polarized in an arbitrary direction, that field can be formed by superimposing two fields corresponding to the TE and TM waves. In other words, if we take an arbitrary field $\boldsymbol{E}$, it can be decomposed into a component having a unit polarization vector directed perpendicular to the incidence plane and another component having the polarization vector that lies on the incidence plane. These two components are orthogonal to each other.

*Example 6.1: TE wave* In Fig. 6.5 we depict the geometry of oblique incidence of a TE wave. The $xy$-plane defines the interface of the two dielectrics, and $\boldsymbol{t}$ of (6.9) lies on that plane. The $zx$-plane defines the incidence plane. In this case, $\boldsymbol{E}$ is polarized along the $y$-axis with $\boldsymbol{H}$ polarized in the $zx$-plane. That is, regarding $\boldsymbol{E}$, we choose polarization direction $\boldsymbol{\varepsilon}_i$, $\boldsymbol{\varepsilon}_r$, and $\boldsymbol{\varepsilon}_t$ of the electric field as $\boldsymbol{e}_2$ (a unit vector toward the positive direction of the $y$-axis that is perpendicular to the plane of paper). In Fig. 6.5, the polarization direction of the electric field is denoted by a symbol $\otimes$. Therefore, we have

$$\boldsymbol{e}_2 \cdot \boldsymbol{\varepsilon}_i = \boldsymbol{e}_2 \cdot \boldsymbol{\varepsilon}_r = \boldsymbol{e}_2 \cdot \boldsymbol{\varepsilon}_t = 1. \tag{6.40}$$

For $\boldsymbol{H}$ we define the direction of unit polarization vectors $\boldsymbol{\varepsilon}_i$, $\boldsymbol{\varepsilon}_r$, and $\boldsymbol{\varepsilon}_t$ so that their direction cosine relative to the $x$-axis can be positive (see Fig. 6.5). Choosing $\boldsymbol{e}_1$ (a unit vector in the direction of the $x$-axis) for $\boldsymbol{t}$ of (6.9) with regard to $\boldsymbol{H}$, we have

$$\boldsymbol{e}_1 \cdot \boldsymbol{\varepsilon}_i = \cos\theta, \boldsymbol{e}_1 \cdot \boldsymbol{\varepsilon}_r = \cos\theta, \text{ and } \boldsymbol{e}_1 \cdot \boldsymbol{\varepsilon}_t = \cos\phi. \tag{6.41}$$

Accordingly as $\boldsymbol{H}_r$ is directed to the same direction as $\boldsymbol{\varepsilon}_r$ or the opposite direction to $\boldsymbol{\varepsilon}_r$, the amplitude is defined as positive or negative, as in the case of the vertical

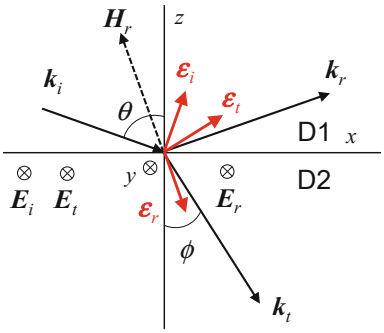


**Fig. 6.5** Geometry of the electromagnetic fields near the interface between dielectric media D1 and D2 in the case of oblique incidence of a TE wave. The electric field $\boldsymbol{E}$ is polarized along the $y$-axis (i.e., perpendicular to the plane of paper) with $\boldsymbol{H}$ polarized in the $zx$-plane. Polarization directions $\boldsymbol{\varepsilon}_i$, $\boldsymbol{\varepsilon}_r$, and $\boldsymbol{\varepsilon}_t$ are given for $\boldsymbol{H}$. To avoid complication, neither $\boldsymbol{H}$ nor $\boldsymbol{H}_t$ is shown

incidence. In Fig. 6.5, we depict the case where the amplitude $H_r$ is negative. That is, the phase of the magnetic field is reversed upon reflection and, hence, $\boldsymbol{H}_r$ is in an opposite direction to $\boldsymbol{\varepsilon}_r$ in Fig. 6.5. Note that $\boldsymbol{H}_i$ and $\boldsymbol{H}_t$ are in the same direction as $\boldsymbol{\varepsilon}_i$ and $\boldsymbol{\varepsilon}_t$, respectively. To avoid complication, neither $\boldsymbol{H}_i$ nor $\boldsymbol{H}_t$ is shown in Fig. 6.5. Applying (6.23) to both $\boldsymbol{E}$ and $\boldsymbol{H}$, we have

$$E_i + E_r = E_t, \tag{6.42}$$

$$H_i \cos\theta + H_r \cos\theta = H_t \cos\phi. \tag{6.43}$$

To derive the above equations, we choose $\boldsymbol{t} = \boldsymbol{e}_2$ with $\boldsymbol{E}$ and $\boldsymbol{t} = \boldsymbol{e}_1$ with $\boldsymbol{H}$ for (6.23). Because of the above-mentioned converse relationship with $\boldsymbol{E}$ and $\boldsymbol{H}$, we have

$$E_r H_r < 0. \tag{6.44}$$

Suppose that we carry out an experiment to determine six amplitudes in (6.42) and (6.43). Out of those quantities, we can freely choose and fix $E_i$. Then, we have five unknown amplitudes, i.e., $E_r$, $E_t$, $H_i$, $H_r$, and $H_t$. Thus, we need three more relations to determine them. Here, information about the characteristic impedance $Z$ is useful. It was defined as (5.64). From (6.6) to (6.8) as well as (6.44), we get

$$Z_1 = \sqrt{\mu_1/\varepsilon_1} = E_i/H_i = -E_r/H_r, \tag{6.45}$$

$$Z_2 = \sqrt{\mu_2/\varepsilon_2} = E_t/H_t, \tag{6.46}$$

where $\varepsilon_1$ and $\mu_1$ are permittivity and permeability of D1, respectively; $\varepsilon_2$ and $\mu_2$ are permittivity and permeability of D2, respectively. As an example, we have

$$\begin{aligned}\boldsymbol{H}_i =& \boldsymbol{n} \times \boldsymbol{E}_i/Z_1 = \boldsymbol{n} \times E_i \boldsymbol{\varepsilon}_{i,e}\, \mathrm{e}^{i(\boldsymbol{k}_i \cdot \boldsymbol{x} - \omega t)}/Z_1 = E_i \boldsymbol{\varepsilon}_{i,m}\, \mathrm{e}^{i(\boldsymbol{k}_i \cdot \boldsymbol{x} - \omega t)}/Z_1 \\ =& H_i \boldsymbol{\varepsilon}_{i,m}\, \mathrm{e}^{i(\boldsymbol{k}_i \cdot \boldsymbol{x} - \omega t)},\end{aligned} \tag{6.47}$$

where we distinguish polarization vectors of electric and magnetic fields. Note in the above discussion, however, we did not distinguish these vectors to avoid complication. Comparing coefficients of the last relation of (6.47), we get

$$E_i/Z_1 = H_i. \tag{6.48}$$

On the basis of (6.42) to (6.46), we are able to decide $E_r$, $E_t$, $H_i$, $H_r$, and $H_t$.

What we wish to determine, however, is a ratio among those quantities. To this end, dividing (6.42) and (6.43) by $E_i\,(>0)$, we define following quantities:

$$R_E^{\perp} \equiv E_r/E_i \quad \text{and} \quad T_E^{\perp} \equiv E_t/E_i, \tag{6.49}$$

where $R_E^{\perp}$ and $T_E^{\perp}$ are said to be a reflection coefficient and transmission coefficient with the electric field, respectively; the symbol $\perp$ means a quantity of the TE wave (i.e., electric field oscillating vertically with respect to the incidence plane). Thus rewriting (6.42) and (6.43) and using $R_E^{\perp}$ and $T_E^{\perp}$, we have

$$\left.\begin{aligned} R_E^{\perp} - T_E^{\perp} &= -1, \\ R_E^{\perp} \frac{\cos\theta}{Z_1} + T_E^{\perp} \frac{\cos\phi}{Z_2} &= \frac{\cos\theta}{Z_1}. \end{aligned}\right\} \tag{6.50}$$

Using Cramer's rule of matrix algebra, we have a solution such that

$$R_E^{\perp} = \frac{\begin{vmatrix} -1 & -1 \\ \frac{\cos\theta}{Z_1} & \frac{\cos\phi}{Z_2} \end{vmatrix}}{\begin{vmatrix} 1 & -1 \\ \frac{\cos\theta}{Z_1} & \frac{\cos\phi}{Z_2} \end{vmatrix}} = \frac{Z_2\ \cos\theta - Z_1\ \cos\phi}{Z_2\ \cos\theta + Z_1\ \cos\phi}, \tag{6.51}$$

$$T_E^{\perp} = \frac{\begin{vmatrix} 1 & -1 \\ \frac{\cos\theta}{Z_1} & \frac{\cos\theta}{Z_1} \end{vmatrix}}{\begin{vmatrix} 1 & -1 \\ \frac{\cos\theta}{Z_1} & \frac{\cos\phi}{Z_2} \end{vmatrix}} = \frac{2Z_2\ \cos\theta}{Z_2\ \cos\theta + Z_1\ \cos\phi}. \tag{6.52}$$

Similarly, defining

$$R_H^{\perp} \equiv H_r/H_i \quad \text{and} \quad T_H^{\perp} \equiv H_t/H_i, \tag{6.53}$$

where $R_H^{\perp}$ and $T_H^{\perp}$ are said to be a reflection coefficient and transmission coefficient with the magnetic field, respectively, we get

$$R_H^{\perp} = \frac{Z_1 \cos\phi - Z_2 \cos\theta}{Z_2 \cos\theta + Z_1 \cos\phi}, \tag{6.54}$$

$$T_H^{\perp} = \frac{2Z_1 \cos\theta}{Z_2 \cos\theta + Z_1 \cos\phi}. \tag{6.55}$$

In this case, rewrite (6.42) as a relation among $H_i$, $H_r$, and $H_t$ using (6.45) and (6.46). Derivation of (6.54) and (6.55) is left for readers. Notice also that

$$R_H^{\perp} = -R_E^{\perp}. \tag{6.56}$$

This relation can easily be derived by (6.45).

*Example 6.2: TM wave* In a manner similar to that described above, we obtain information about the TM wave. Switching a role of $\boldsymbol{E}$ and $\boldsymbol{H}$, we assume that $\boldsymbol{H}$ is polarized along the $y$-axis with $\boldsymbol{E}$ polarized in the $zx$-plane. Following the aforementioned procedures, we have

$$E_i \cos\theta + E_r \cos\theta = E_t \cos\phi, \tag{6.57}$$

$$H_i + H_r = H_t. \tag{6.58}$$

From (6.57) and (6.58), similarly we get

$$R_E^{\parallel} = \frac{Z_2 \cos\phi - Z_1 \cos\theta}{Z_1 \cos\theta + Z_2 \cos\phi}, \tag{6.59}$$

$$T_E^{\parallel} = \frac{2Z_2 \cos\theta}{Z_1 \cos\theta + Z_2 \cos\phi}. \tag{6.60}$$

Also, we get

$$R_H^{\parallel} = \frac{Z_1 \cos\theta - Z_2 \cos\phi}{Z_1 \cos\theta + Z_2 \cos\phi} = -R_E^{\parallel}, \tag{6.61}$$

$$T_H^{\parallel} = \frac{2Z_1 \cos\theta}{Z_1 \cos\theta + Z_2 \cos\phi}. \tag{6.62}$$

In Table 6.1, we list the important coefficients in relation to the reflection and transmission of electromagnetic waves along with their relationship.

In Examples 6.1 and 6.2, we have examined how the reflection and transmission coefficients vary as a function of characteristic impedance as well as incidence and refraction angles. Meanwhile, in a non-magnetic substance a refractive index $n$ can be approximated as

$$n \approx \sqrt{\varepsilon_r}, \tag{5.57}$$

**Table 6.1** Reflection and transmission coefficients of TE and TM waves

| $\perp$ Incidence (TE) | $\parallel$ Incidence (TM) |
|---|---|
| $R_E^{\perp} = \frac{Z_2 \cos\theta - Z_1 \cos\phi}{Z_2 \cos\theta + Z_1 \cos\phi}$ | $R_E^{\parallel} = \frac{Z_2 \cos\phi - Z_1 \cos\theta}{Z_1 \cos\theta + Z_2 \cos\phi}$ |
| $R_H^{\perp} = \frac{Z_1 \cos\phi - Z_2 \cos\theta}{Z_2 \cos\theta + Z_1 \cos\phi} = -R_E^{\perp}$ | $R_H^{\parallel} = \frac{Z_1 \cos\theta - Z_2 \cos\phi}{Z_1 \cos\theta + Z_2 \cos\phi} = -R_E^{\parallel}$ |
| $T_E^{\perp} = \frac{2Z_2 \cos\theta}{Z_2 \cos\theta + Z_1 \cos\phi}$ | $T_E^{\parallel} = \frac{2Z_2 \cos\theta}{Z_1 \cos\theta + Z_2 \cos\phi}$ |
| $T_H^{\perp} = \frac{2Z_1 \cos\theta}{Z_2 \cos\theta + Z_1 \cos\phi}$ | $T_H^{\parallel} = \frac{2Z_1 \cos\theta}{Z_1 \cos\theta + Z_2 \cos\phi}$ |
| $-R_E^{\perp} R_H^{\perp} + T_E^{\perp} T_H^{\perp} \frac{\cos\phi}{\cos\theta} = 1 \quad (R^{\perp} + T^{\perp} = 1)$ | $-R_E^{\parallel} R_H^{\parallel} + T_E^{\parallel} T_H^{\parallel} \frac{\cos\phi}{\cos\theta} = 1 \quad (R^{\parallel} + T^{\parallel} = 1)$ |

assuming that $\mu_r \approx 1$. In this case, we have

$$Z = \sqrt{\mu/\varepsilon} = \sqrt{\mu_r \mu_0 / \varepsilon_r \varepsilon_0} \approx Z_0 / \sqrt{\varepsilon_r} \approx Z_0 / n.$$

Using this relation, we can readily rewrite the reflection and transmission coefficients as a function of refractive indices of the dielectrics. The derivation is left for the readers.

## 6.4 Energy Transport by Electromagnetic Waves

Returning to (5.58) and (5.59), let us consider energy transport in a dielectric medium by electromagnetic waves. Let us describe their electric ($\boldsymbol{E}$) and magnetic ($\boldsymbol{H}$) fields of the electromagnetic waves in a uniform and infinite dielectric medium such that

$$\boldsymbol{E} = E\boldsymbol{\varepsilon}_{\mathrm{e}} \mathrm{e}^{i(k\boldsymbol{n}\cdot\boldsymbol{x} - \omega t)}, \tag{6.63}$$

$$\boldsymbol{H} = H\boldsymbol{\varepsilon}_{\mathrm{m}} \mathrm{e}^{i(k\boldsymbol{n}\cdot\boldsymbol{x} - \omega t)}, \tag{6.64}$$

where $\boldsymbol{\varepsilon}_{\mathrm{e}}$ and $\boldsymbol{\varepsilon}_{\mathrm{m}}$ are unit polarization vector; we assume that both $E$ and $H$ are positive. Notice again that $\boldsymbol{\varepsilon}_{\mathrm{e}}$, $\boldsymbol{\varepsilon}_{\mathrm{m}}$, and $\boldsymbol{n}$ constitute a right-handed system in this order.

The energy transport is characterized by a Poynting vector $\boldsymbol{S}$ that is described by

$$\boldsymbol{S} = \boldsymbol{E} \times \boldsymbol{H}. \tag{6.65}$$

Since $\boldsymbol{E}$ and $\boldsymbol{H}$ have a dimension $\left[\frac{\mathrm{V}}{\mathrm{m}}\right]$ and $\left[\frac{\mathrm{A}}{\mathrm{m}}\right]$, respectively, $\boldsymbol{S}$ has a dimension $\left[\frac{\mathrm{W}}{\mathrm{m}^2}\right]$. Hence, $\boldsymbol{S}$ represents an energy flow per unit time and per unit area with respect to the propagation direction. For simplicity, let us assume that the electromagnetic wave is propagating toward the $z$-direction. Then, we have

$$\boldsymbol{E} = E\boldsymbol{\varepsilon}_{\mathrm{e}}\mathrm{e}^{i(kz-\omega t)}, \tag{6.66}$$

$$\boldsymbol{H} = H\boldsymbol{\varepsilon}_{\mathrm{m}}\mathrm{e}^{i(kz-\omega t)}. \tag{6.67}$$

To seek a time-averaged energy flow toward the $z$-direction, it suffices to multiply real parts of (6.66) and (6.67) and integrate it during a period $T$ at a point of $z = 0$. Thus, a time-averaged Poynting vector $\overline{\boldsymbol{S}}$ is given by

$$\overline{\boldsymbol{S}} = \boldsymbol{e}_3 \frac{EH}{T} \int_0^T \cos^2 \omega t \, \mathrm{d}t, \tag{6.68}$$

where $T = 1/\nu = 2\pi/\omega$. Using a trigonometric formula

$$\cos^2 \omega t = \frac{1}{2}(1 + \cos 2\omega t), \tag{6.69}$$

the integration can easily be performed. Thus, we get

$$\overline{\boldsymbol{S}} = \frac{1}{2} EH\boldsymbol{e}_3. \tag{6.70}$$

Equivalently, we have

$$\overline{\boldsymbol{S}} = \frac{1}{2} \boldsymbol{E} \times \boldsymbol{H}^*. \tag{6.71}$$

Meanwhile, an energy density $W$ is given by

$$W = \frac{1}{2}(\boldsymbol{E} \cdot \boldsymbol{D} + \boldsymbol{H} \cdot \boldsymbol{B}), \tag{6.72}$$

where the first and second terms are pertinent to the electric and magnetic fields, respectively. Note in (6.72) that the dimension of $\boldsymbol{E} \cdot \boldsymbol{D}$ is $\left[\frac{\mathrm{V}}{\mathrm{m}} \cdot \frac{\mathrm{C}}{\mathrm{m}^2}\right] = \left[\frac{\mathrm{J}}{\mathrm{m}^3}\right]$ and that the dimension of $\boldsymbol{H} \cdot \boldsymbol{B}$ is $\left[\frac{\mathrm{A}}{\mathrm{m}} \cdot \frac{\mathrm{V \cdot s}}{\mathrm{m}^2}\right] = \left[\frac{\mathrm{W \cdot s}}{\mathrm{m}^3}\right] = \left[\frac{\mathrm{J}}{\mathrm{m}^3}\right]$. Using (5.7) and (5.10), we have

$$W = \frac{1}{2}\left(\varepsilon \boldsymbol{E}^2 + \mu \boldsymbol{H}^2\right). \tag{6.73}$$

As in the above case, estimating a time-averaged energy density $\overline{W}$, we get

$$\overline{W} = \frac{1}{2}\left(\frac{1}{2}\varepsilon E^2 + \frac{1}{2}\mu H^2\right) = \frac{1}{4}\varepsilon E^2 + \frac{1}{4}\mu H^2. \tag{6.74}$$

We also get this relation by integrating (6.73) over a wavelength $\lambda$ at a time of $t = 0$. Using (5.60) and (5.61), we have

$$\varepsilon E^2 = \mu H^2. \tag{6.75}$$

This implies that the energy density resulting from the electric field and that due to the magnetic field have the same value. Thus, rewriting (6.74) we have

$$\overline{W} = \frac{1}{2}\,\varepsilon E^2 = \frac{1}{2}\,\mu H^2. \tag{6.76}$$

Moreover, using (5.43), we have for an impedance

$$Z = E/H = \sqrt{\mu/\varepsilon} = \mu v \quad \text{or} \quad E = \mu v H. \tag{6.77}$$

Using this relation along with (6.75), we get

$$\overline{\boldsymbol{S}} = \frac{1}{2}\,v\varepsilon E^2 \boldsymbol{e}_3 = \frac{1}{2}\,v\mu H^2 \boldsymbol{e}_3. \tag{6.78}$$

Thus, we have various relations among amplitudes of electromagnetic waves and related physical quantities together with constant of dielectrics.

Returning to Examples 6.1 and 6.2, let us further investigate the reflection and transmission properties of the electromagnetic waves. From (6.51) to (6.55) as well as (6.59) to (6.62), we get in both the cases of TE and TM waves

$$-R_E^{\perp} R_H^{\perp} + T_E^{\perp} T_H^{\perp} \frac{\cos\phi}{\cos\theta} = 1, \tag{6.79}$$

$$-R_E^{\parallel} R_H^{\parallel} + T_E^{\parallel} T_H^{\parallel} \frac{\cos\phi}{\cos\theta} = 1. \tag{6.80}$$

In both the TE and TM cases, we define *reflectance* $R$ and *transmittance* $T$ such that

$$R \equiv -R_E R_H = R_E^2 = 2|\overline{\boldsymbol{S}}_r|/2|\overline{\boldsymbol{S}}_i| = |\overline{\boldsymbol{S}}_r|/|\overline{\boldsymbol{S}}_i|, \tag{6.81}$$

where $\overline{\boldsymbol{S}}_r$ and $\overline{\boldsymbol{S}}_i$ are time-averaged Poynting vectors of the reflected wave and incident waves, respectively. Also, we have

$$T \equiv T_E T_H \frac{\cos\phi}{\cos\theta} = \frac{2|\overline{\boldsymbol{S}}_t|}{2|\overline{\boldsymbol{S}}_i|}\frac{\cos\phi}{\cos\theta} = \frac{|\overline{\boldsymbol{S}}_t|}{|\overline{\boldsymbol{S}}_i|}\frac{\cos\phi}{\cos\theta}, \tag{6.82}$$

where $\overline{\boldsymbol{S}}_t$ is a time-averaged Poynting vector of the transmitted wave. Thus, we have

$$R + T = 1. \tag{6.83}$$

The relation (6.83) represents the energy conservation. The factor $\frac{\cos\phi}{\cos\theta}$ can be understood by Fig. 6.6 that depicts a luminous flux near the interface. Suppose that we have an incident wave with an irradiance $I\left[\frac{\mathrm{W}}{\mathrm{m}^2}\right]$ whose incidence plane is the $zx$-plane. Notice that $I$ has the same dimension as a Poynting vector.

Here, let us think of the luminous flux that is getting through a unit area (i.e., a unit length square) perpendicular to the propagation direction of the light. Then, this flux illuminates an area on the interface of a unit length (in the $y$-direction) multiplied by a length of $\frac{\cos\phi}{\cos\theta}$ (in the $x$-direction). That is, the luminous flux has been widened (or squeezed) by $\frac{\cos\phi}{\cos\theta}$ times after getting through the interface. The irradiance has been weakened (or strengthened) accordingly. Thus, to take a balance of income and outgo with respect to the luminous flux before and after getting through the interface, the transmission irradiance must be multiplied by a factor $\frac{\cos\phi}{\cos\theta}$.

## 6.5 Brewster Angles and Critical Angles

In this section and subsequent sections, we deal with non-magnetic substance as dielectrics; namely, we assume $\mu_r \approx 1$. In that case, as mentioned in Sect. 6.3 we rewrite, e.g., (6.51) and (6.59) as

$$R_E^{\perp} = \frac{\cos\theta - n\cos\phi}{\cos\theta + n\cos\phi}, \tag{6.84}$$

$$R_E^{\parallel} = \frac{\cos\phi - n\cos\theta}{\cos\phi + n\cos\theta}, \tag{6.85}$$

where $n\ (= n_2/n_1)$ is a relative refractive index of D2 relative to D1. Let us think of a condition on which $R_E^{\perp} = 0$ or $R_E^{\parallel} = 0$.

First, we consider (6.84). We have

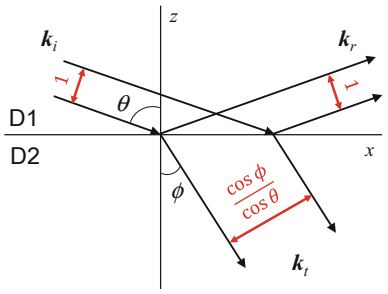


**Fig. 6.6** Luminous flux near the interface

$$\begin{aligned}[\text{numerator of}(6.\,84)] &= \cos\theta - n\,\cos\phi = \cos\theta - \frac{\sin\theta}{\sin\phi}\cos\phi \\ &= \frac{\sin\phi\,\cos\theta - \sin\theta\,\cos\phi}{\sin\phi} = \frac{\sin(\phi-\theta)}{\sin\phi},\end{aligned} \tag{6.86}$$

where with the second equality we used Snell's law; the last equality is due to trigonometric formula. Since we assume $0<\theta<\pi/2$ and $0<\phi<\pi/2$, we have $-\pi/2<\phi-\theta<\pi/2$. Therefore, if and only if $\phi-\theta=0$, $\sin(\phi-\theta)=0$. Namely, only when $\phi=\theta$, $R_E^{\perp}$ could vanish. For different dielectrics having different refractive indices, only if $\phi=\theta=0$ (i.e., a vertical incidence), we have $\phi=\theta$. But, in that case we have

$$\lim_{\phi\to 0,\,\theta\to 0}\frac{\sin(\phi-\theta)}{\sin\phi} = \frac{0}{0}.$$

This is a limit of indeterminate form. From (6.84), however, we have

$$R_E^{\perp} = \frac{1-n}{1+n}, \tag{6.87}$$

for $\phi=\theta=0$. This implies that $R_E^{\perp}$ does not vanish at $\phi=\theta=0$. Thus, $R_E^{\perp}$ never vanishes for any $\theta$ or $\phi$. Note that for this condition, naturally we have

$$R_E^{\parallel} = \frac{1-n}{1+n}.$$

This is because with $\phi=\theta=0$ we have no physical difference between TE and TM waves.

In turn, let us examine (6.85) similarly with the case of TM wave.

$$\begin{aligned}[\text{numerator of}(6.\,85)] &= \cos\phi - n\,\cos\theta = \cos\phi - \frac{\sin\theta}{\sin\phi}\cos\theta \\ &= \frac{\sin\phi\,\cos\phi - \sin\theta\,\cos\theta}{\sin\phi} = \frac{\sin(\phi-\theta)\,\cos(\phi+\theta)}{\sin\phi}.\end{aligned} \tag{6.88}$$

With the last equality of (6.88), we used a trigonometric formula. From (6.86), we know that $\frac{\sin(\phi-\theta)}{\sin\phi}$ does not vanish. Therefore, for $R_E^{\parallel}$ to vanish, we need $\cos(\phi+\theta)=0$. Since $0<\phi+\theta<\pi$, $\cos(\phi+\theta)=0$ if and only if

$$\phi+\theta=\pi/2. \tag{6.89}$$

In other words, for particular angles $\theta=\theta_{\mathrm{B}}$ and $\phi=\phi_{\mathrm{B}}$ that satisfy

$$\phi_{\mathrm{B}}+\theta_{\mathrm{B}}=\pi/2, \tag{6.90}$$

we have $R_E^{\parallel} = 0$; i.e., we do not observe a reflected wave. The particular angle $\theta_B$ is said to be the Brewster angle. For $\theta_B$, we have

$$\begin{aligned} &\sin\phi_B = \sin\left(\frac{\pi}{2} - \theta_B\right) = \cos\theta_B, \\ &n = \sin\theta_B / \sin\phi_B = \sin\theta_B / \cos\theta_B = \tan\theta_B \quad \text{or} \quad \theta_B = \tan^{-1} n, \\ &\phi_B = \tan^{-1} n^{-1}. \end{aligned} \tag{6.91}$$

Suppose that we have a parallel plate consisting a dielectric D2 of a refractive index $n_2$ sandwiched with another dielectric D1 of a refractive index $n_1$ (Fig. 6.7). Let $\theta_B$ be the Brewster angle when the TM wave is incident from D1 to D2. In the above discussion, we defined a relative refractive index $n$ of D2 relative D1 as $n = n_2/n_1$; recall (6.39). The other way around, suppose that the TM wave is incident from D2 to D1. Then, the relative refractive index of D1 relative to D2 is $n_1/n_2 = n^{-1}$. In this situation, another Brewster angle (from D2 to D1) defined as $\widetilde{\theta}_B$ is given by

$$\widetilde{\theta}_B = \tan^{-1} n^{-1}. \tag{6.92}$$

This number is, however, identical to $\phi_B$ in (6.91). Thus, we have

$$\widetilde{\theta}_B = \phi_B \tag{6.93}$$

Thus, regarding the TM wave that is propagating in D2 after getting through the interface and is to get back to D1, $\widetilde{\theta}_B = \phi_B$ is again the Brewster angle. In this way, the said TM wave is propagating from D1 to D2 and then getting back from D2 to D1 without being reflected by the two interfaces. This conspicuous feature is often utilized for an optical device.

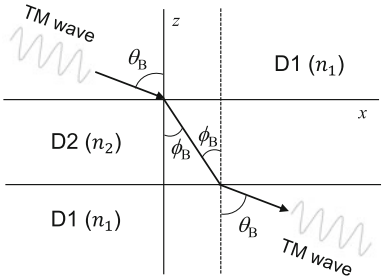


**Fig. 6.7** Diagram that explains the Brewster angle. Suppose that a parallel plate consisting of a dielectric D2 of a refractive index $n_2$ is sandwiched with another dielectric D1 of a refractive index $n_1$. The incidence angle $\theta_B$ represents the Brewster angle observed when the TM wave is incident from D1 to D2. $\phi_B$ is another Brewster angle that is observed when the TM wave is getting back from D2 to D1

If an electromagnetic wave is incident from a dielectric of a higher refractive index to that of a lower index, the total reflection takes place. This is equally the case with both TE and TM waves. For the total reflection to take place, $\theta$ should be larger than a critical angle $\theta_c$ that is defined by

$$\theta_c = \sin^{-1} n. \tag{6.94}$$

This is because at $\theta_c$ from the Snell's law we have

$$\frac{\sin\theta_c}{\sin\frac{\pi}{2}} = \sin\theta_c = \frac{n_2}{n_1}\ (\equiv n). \tag{6.95}$$

From (6.95), we have

$$\tan\theta_c = n/\sqrt{1-n^2} > n = \tan\theta_B.$$

In the case of the TM wave, therefore, we find that

$$\theta_c > \theta_B. \tag{6.96}$$

The critical angle is always larger than the Brewster angle with TM waves.

### 6.5.1 Total Reflection

In Sect. 6.2, we saw that the Snell's law results from the kinematical requirement. For this reason, we may consider it as a universal relation that can be extended to complex refraction angles. In fact, for the Snell's law to hold with $\theta > \theta_c$, we must have

$$\sin\phi > 1. \tag{6.97}$$

This needs us to extend $\phi$ to a complex domain. Putting

$$\phi = \frac{\pi}{2} + ia\ (a : \text{real},\ a \neq 0), \tag{6.98}$$

we have

$$\sin\phi \equiv \frac{1}{2i}\left(e^{i\phi} - e^{-i\phi}\right) = \frac{1}{2}\left(e^{-a} + e^{a}\right) > 1, \tag{6.99}$$

$$\cos\phi \equiv \frac{1}{2}\left(e^{i\phi} + e^{-i\phi}\right) = \frac{i}{2}\left(e^{-a} - e^{a}\right). \tag{6.100}$$

Thus, $\cos\phi$ is pure imaginary.

Now, let us consider a transmitted wave whose electric field is described as

$$\boldsymbol{E}_{\mathrm{t}} = E\boldsymbol{\varepsilon}_{\mathrm{t}}\mathrm{e}^{i(\boldsymbol{k}_t\cdot\boldsymbol{x}-\omega t)}, \tag{6.101}$$

where $\boldsymbol{\varepsilon}_{\mathrm{t}}$ is the unit polarization vector and $\boldsymbol{k}_t$ is a wavenumber vector of the transmission wave. Suppose that the incidence plane is the $zx$-plane. Then, we have

$$\boldsymbol{k}_t\cdot\boldsymbol{x} = k_{t_x}x + k_{t_z}z = xk_t\sin\phi + zk_t\cos\phi, \tag{6.102}$$

where $k_{t_x}$ and $k_{t_z}$ are $x$ and $z$ components of $\boldsymbol{k}_t$, respectively; $k_t = |\boldsymbol{k}_t|$. Putting

$$\cos\phi = ib \ (b : \text{real},\ b \neq 0), \tag{6.103}$$

we have

$$\boldsymbol{E}_{\mathrm{t}} = E\boldsymbol{\varepsilon}_{\mathrm{t}}\mathrm{e}^{i(xk_t\sin\phi + ibzk_t-\omega t)} = E\boldsymbol{\varepsilon}_{\mathrm{t}}\mathrm{e}^{i(xk_t\sin\phi-\omega t)}\mathrm{e}^{-bzk_t}. \tag{6.104}$$

With the total reflection, we must have

$$z\to\infty \Rightarrow \mathrm{e}^{-bzk_t}\to 0. \tag{6.105}$$

To meet this requirement, we have

$$b > 0. \tag{6.106}$$

Meanwhile, we have

$$\cos^2\phi = 1-\sin^2\phi = 1-\frac{\sin^2\theta}{n^2} = \frac{n^2-\sin^2\theta}{n^2}, \tag{6.107}$$

where notice that $n<1$, because we are dealing with the incidence of light from a medium with a higher refractive index to a low index medium. When we consider the total reflection, the numerator of (6.107) is negative, and so we have two choices such that

$$\cos\phi = \pm i\frac{\sqrt{\sin^2\theta-n^2}}{n}. \tag{6.108}$$

From (6.103) and (6.106), we get

$$\cos\phi = i\frac{\sqrt{\sin^2\theta-n^2}}{n}. \tag{6.109}$$

Hence, inserting (6.109) into (6.84) we have for the TE wave

$$R_E^{\perp} = \frac{\cos\theta - i\sqrt{\sin^2\theta - n^2}}{\cos\theta + i\sqrt{\sin^2\theta - n^2}}. \tag{6.110}$$

Then, we have

$$R^{\perp} = -R_E^{\perp}(R_H^{\perp})^* = R_E^{\perp}(R_E^{\perp})^* = \frac{\cos\theta - i\sqrt{\sin^2\theta - n^2}}{\cos\theta + i\sqrt{\sin^2\theta - n^2}} \cdot \frac{\cos\theta + i\sqrt{\sin^2\theta - n^2}}{\cos\theta - i\sqrt{\sin^2\theta - n^2}} = 1. \tag{6.111}$$

As for the TM wave, substituting (6.109) for (6.85) we have

$$R_E^{\parallel} = \frac{-n^2\cos\theta + i\sqrt{\sin^2\theta - n^2}}{n^2\cos\theta + i\sqrt{\sin^2\theta - n^2}}. \tag{6.112}$$

In this case, we also get

$$R^{\parallel} = \frac{-n^2\cos\theta + i\sqrt{\sin^2\theta - n^2}}{n^2\cos\theta + i\sqrt{\sin^2\theta - n^2}} \cdot \frac{-n^2\cos\theta - i\sqrt{\sin^2\theta - n^2}}{n^2\cos\theta - i\sqrt{\sin^2\theta - n^2}} = 1. \tag{6.113}$$

The relations (6.111) and (6.113) ensure that the energy flow gets back to a higher refractive index medium.

Thus, the total reflection is characterized by the complex reflection coefficient expressed as (6.110) and (6.112) as well as a reflectance of 1. From (6.110) and (6.112), we can estimate a change in a phase of the electromagnetic wave that takes place by virtue of the total reflection. For this purpose, we put

$$R_E^{\perp} \equiv \mathrm{e}^{i\alpha} \quad \text{and} \quad R_E^{\parallel} \equiv \mathrm{e}^{i\beta}. \tag{6.114}$$

Rewriting (6.110), we have

$$R_E^{\perp} = \frac{\cos^2\theta - (\sin^2\theta - n^2) - 2i\cos\theta\sqrt{\sin^2\theta - n^2}}{1 - n^2}. \tag{6.115}$$

At a critical angle $\theta_c$, from (6.95) we have

$$\sin\theta_c = n. \tag{6.116}$$

Therefore, we have

$$1 - n^2 = \cos^2\theta_c. \tag{6.117}$$

Then, as expected, we get

$$R_E^{\perp}|_{\theta=\theta_c} = 1. \tag{6.118}$$

Note, however, that at $\theta = \pi/2$ (i.e., grazing incidence) we have

$$R_E^{\perp}|_{\theta=\pi/2} = -1. \tag{6.119}$$

From (6.115), an argument $\alpha$ in a complex plane is given by

$$\tan\alpha = -\frac{2\cos\theta\sqrt{\sin^2\theta - n^2}}{\cos^2\theta - (\sin^2\theta - n^2)}. \tag{6.120}$$

The argument $\alpha$ defines a phase shift upon the total reflection. Considering (6.115) and (6.118), we have

$$\alpha|_{\theta=\theta_c} = 0.$$

Since $1 - n^2 > 0$ (*i.e.*, $n < 1$) and in the total reflection region $\sin^2\theta - n^2 > 0$, the imaginary part of $R_E^{\perp}$ is negative for any $\theta$ (i.e., 0 to $\pi/2$). On the other hand, the real part of $R_E^{\perp}$ varies from 1 to $-1$, as is evidenced from (6.118) and (6.119). At $\theta_0$ that satisfies a following condition:

$$\sin\theta_0 = \sqrt{\frac{1+n^2}{2}}, \tag{6.121}$$

the real part is zero. Thus, the phase shift $\alpha$ varies from 0 to $-\pi$ as indicated in Fig. 6.8. Comparing (6.121) with (6.116) and taking into account $n < 1$, we have

$$\theta_c < \theta_0 < \pi/2.$$

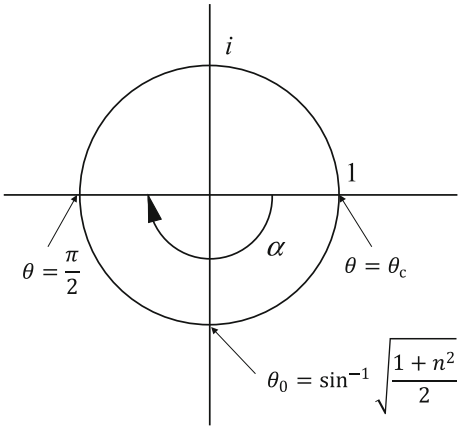


**Fig. 6.8** Phase shift $\alpha$ defined in a complex plane for the total reflection of TE wave. The number $n$ denotes a relative refractive index of D2 relative to D1. At a critical angle $\theta_c$, $\alpha = 0$

Similarly, we estimate the phase change for a TM wave. Rewriting (6.112), we have

$$\begin{aligned} R_E^{\parallel} &= \frac{-n^4\cos^2\theta + \sin^2\theta - n^2 + 2i\,n^2\cos\theta\sqrt{\sin^2\theta - n^2}}{n^4\cos^2\theta + \sin^2\theta - n^2} \\ &= \frac{-n^4\cos^2\theta + \sin^2\theta - n^2 + 2i\,n^2\cos\theta\sqrt{\sin^2\theta - n^2}}{(1-n^2)\left(\sin^2\theta - n^2\cos^2\theta\right)}. \end{aligned} \tag{6.122}$$

Then, we have

$$R_E^{\parallel}|_{\theta=\theta_c} = -1. \tag{6.123}$$

Also at $\theta = \pi/2$ (i.e., grazing incidence), we have

$$R_E^{\parallel}|_{\theta=\pi/2} = 1. \tag{6.124}$$

From (6.122), an argument $\beta$ is given by

$$\tan\beta = \frac{2\,n^2\cos\theta\sqrt{\sin^2\theta - n^2}}{-n^4\cos^2\theta + \sin^2\theta - n^2}. \tag{6.125}$$

Considering (6.122) and (6.123), we have

$$\beta|_{\theta=\theta_c} = \pi.$$

In the total reflection region, we have

$$\sin^2\theta - n^2\cos^2\theta > n^2 - n^2\cos^2\theta = n^2\left(1 - \cos^2\theta\right) > 0.$$

Therefore, the denominator of (6.122) is positive and, hence, the imaginary part of $R_E^{\parallel}$ is positive as well for any  (i.e., 0 to $\pi/2$). From (6.123) and (6.124), on the other hand, the real part of $R_E^{\parallel}$ in (6.122) varies from $-1$ to 1. At $\widetilde{\theta}_0$ that satisfies a following condition:

$$\cos\widetilde{\theta}_0 = \sqrt{\frac{1-n^2}{1+n^4}}, \tag{6.126}$$

the real part of $R_E^{\parallel}$ is zero. Once again, we have

$$\theta_c < \widetilde{\theta}_0 < \pi/2.$$

Thus, the phase $\beta$ varies from $\pi$ to 0 as depicted in Fig. 6.9.

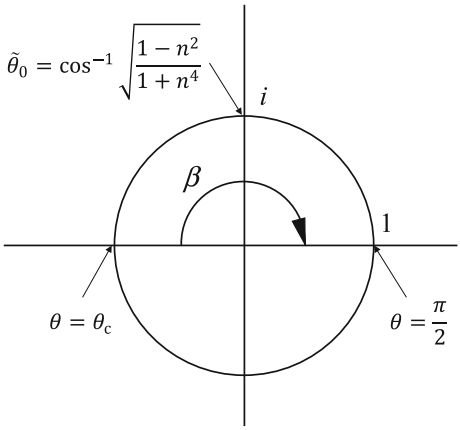


**Fig. 6.9** Phase shift $\beta$ for the total reflection of TM wave. At a critical angle $\theta_c$, $\beta = \pi$

## 6.6 Several Remarks on Analytic Functions

In Sect. 6.6, we mentioned somewhat peculiar features of complex trigonometric functions such as $\sin\phi > 1$ in light of real functions. Even though we do not make systematic description of analytic functions (or complex analysis), we add several remarks on them. Readers are referred to appropriate literature with the theory of analytic functions [2].

Suppose that a mathematical function is defined on a real domain (i.e., a real number line). Consider whether that function is differentiable at a certain point $x_0$ of the real number line. On this occasion, we can approach $x_0$ only from two directions, i.e., from the side of $x < x_0$ (from the left) or from the side of $x > x_0$ (from the right) (see Fig. 6.10a). Meanwhile, suppose that a mathematical function is defined on a complex domain (i.e., a complex plane). Also, consider whether the function is differentiable at a certain point $z_0$ of the complex plane. In this case, we can approach $z_0$ from continuously varying directions (see Fig. 6.10b) where only four directions are depicted.

Suppose now that (i) a function $f(z)$ is single-valued and that (ii) $f(z)$ is differentiable, in other words, the following finite limit exists

$$\left.\frac{df}{dz}\right|_{z_0} = \lim_{\Delta z \to 0} \frac{f(z_0 + \Delta z) - f(z_0)}{\Delta z} \tag{6.127}$$

regardless of the way (or direction) one approaches the point $z = z_0$. Notice that four such ways (or directions) are shown in Fig. 6.10b. Then, $f(z)$ is said to be analytic at $z_0$. This definition naturally implies that the differentiability of $f(z)$ at $z_0$ demands the differentiability at all points of the "neighborhood" of $z_0$. A point where $f(z)$ is analytic is called a regular point of $f(z)$. Otherwise, the point is called a singular point of $f(z)$. If $f(z)$ is analytic on the entire complex plane, $f(z)$ is called an entire function.

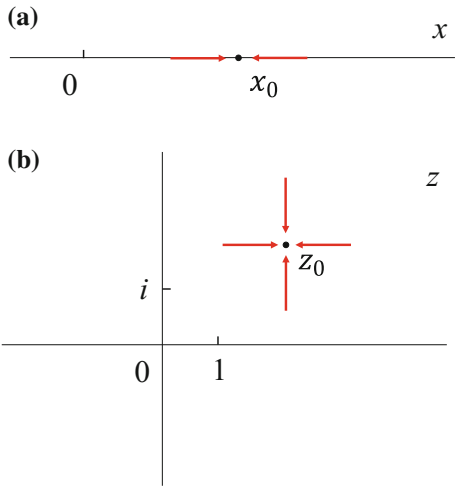


**Fig. 6.10** Ways a number (real or complex) is approached. **a** Two ways a number $x_0$ is approached in a real number line. **b** Various ways a number $z_0$ is approached in a complex plane. Here, only four ways are indicated

Usually a complex variable $z$ is given by

$$z = x + iy,$$

where $x$ and $y$ are real. Consider a flowing function

$$g(z) = x + iy = z.$$

Then, $g(z)$ is an entire function. On the other hand, consider a following function $h(z)$ such that

$$h(z) = 2x + iy = z + x.$$

Then, the derivative (6.127) with $h(z)$ varies depending on a way $z_0$ is approached. For instance, think of

$$\left.\frac{\mathrm{d}h}{\mathrm{d}z}\right|_0 = \lim_{z\to 0}\frac{h(0+z)-h(0)}{z} = \lim_{x\to 0,\, y\to 0}\frac{2x+iy}{x+iy} = \lim_{x\to 0,\, y\to 0}\frac{2x^2+y^2-ixy}{x^2+y^2}.$$

Suppose that the differentiation is taken along a straight line in a complex plane represented by $iy = (ik)x$ $(k, x, y : \text{real})$. Then, we have

$$\left.\frac{\mathrm{d}h}{\mathrm{d}z}\right|_0 = \lim_{x\to 0,\, y\to 0}\frac{2x^2+k^2x^2-ikx^2}{x^2+k^2x^2} = \lim_{x\to 0,\, y\to 0}\frac{2+k^2-ik}{1+k^2} = \frac{2+k^2-ik}{1+k^2}.$$

However, this means that $\left.\frac{\mathrm{d}h}{\mathrm{d}z}\right|_0$ takes varying values depending upon $k$. Namely, $\left.\frac{\mathrm{d}h}{\mathrm{d}z}\right|_0$ cannot uniquely be defined but depends on different ways to approach the

origin of the complex plane. Thus, we find that the derivative takes different values depending on straight lines along which the differentiation is taken. In other words, $h(z)$ is not differentiable or analytic at $z = 0$.

Reflecting the nature of the differentiability, we have an integral representation of an analytic function. This is well-known as *Cauchy's integral formula*. It is described as

$$f(z) = \frac{1}{2\pi i} \oint_C \frac{f(\xi)}{\xi - z} \, \mathrm{d}\xi, \tag{6.128}$$

where $f(z)$ is analytic within a complex domain encircled by a closed contour $C$ (see Fig. 6.11) and contour integration along $C$ is taken in the counterclockwise direction. We also have a following salient property with respect to (higher-order) derivatives such that

$$\frac{\mathrm{d}^n f(z)}{\mathrm{d}z^n} = \frac{n!}{2\pi i} \oint_C \frac{f(\xi)}{(\xi - z)^{n+1}} \, \mathrm{d}\xi. \tag{6.129}$$

We have a further outstanding theorem as follows. Equation (6.129) implies that an analytic function is infinitely differentiable and that the derivatives of all order of an analytic function are again analytic. These prominent properties arise partly from the aforementioned stringent requirement on the differentiability of a function of a complex variable.

A following theorem is important and intriguing with the analytic functions.

**Theorem 6.1** (Cauchy–Liouville Theorem) *A bounded entire function must be a constant.*

*Proof* Using (6.129), we consider the first derivative of an entire function $f(z)$ described as

$$\frac{\mathrm{d}f}{\mathrm{d}z} = \frac{1}{2\pi i} \oint_C \frac{f(\xi)}{(\xi - z)^2} \, \mathrm{d}\xi.$$

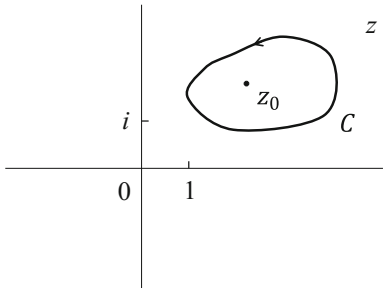


**Fig. 6.11** Complex domain encircled by a closed contour $C$ and contour integration along $C$. The integration is taken in the counterclockwise direction

Since $f(z)$ is an entire function, we can arbitrarily choose a large enough circle of radius $R$ centered at $z$ for a closed contour $C$. On the circle, we have

$$\xi = z + Re^{i\theta},$$

where $\theta$ is a real number changing from 0 to $2\pi$. Then, the above equation can be rewritten as

$$\frac{\mathrm{d}f}{\mathrm{d}z} = \frac{1}{2\pi i}\int_0^{2\pi} \frac{f(\xi)}{(R\mathrm{e}^{i\theta})^2} iR\mathrm{e}^{i\theta}\,\mathrm{d}\theta = \frac{1}{2\pi R}\int_0^{2\pi} \frac{f(\xi)}{\mathrm{e}^{i\theta}}\,\mathrm{d}\theta.$$

Taking an absolute value of both sides, we have

$$\left|\frac{\mathrm{d}f}{\mathrm{d}z}\right| \le \frac{1}{2\pi R}\int_0^{2\pi} |f(\xi)|\,\mathrm{d}\theta \le \frac{M}{2\pi R}\int_0^{2\pi} \mathrm{d}\theta = \frac{M}{R},$$

where $M$ is the maximum of $|f(\xi)|$. As $R \to \infty$, $\left|\frac{\mathrm{d}f}{\mathrm{d}z}\right|$ tends to be zero. This implies that $f(z)$ is constant. This completes the proof.

At the first glance, Cauchy–Liouville Theorem looks astonishing in terms of theory of real analysis. It is because we are too familiar with $-1 \le \sin x \le 1$ for any real number $x$. Note that $\sin x$ is bounded in a real domain. In fact, as (6.98) and (6.99) show, $\sin\phi \to \infty$ with $a \to \infty$. This simple example clearly shows that $\sin\phi$ is an unbounded function in a complex domain. As a very familiar example of a bounded entire functions, we take

$$f(z) = \cos z^2 + \sin z^2 \equiv 1,$$

which is defined in an entire complex plane.

As a matter of course, in Sect. 6.6 a complex angle $\phi$ should be determined experimentally from (6.109).

## 6.7 Waveguide Applications

There are many optical devices based upon light propagation. Among them, waveguide devices utilize the total reflection. We explain their operation principle.

Suppose that we have a thin plate (usually said to be a slab) comprising a dielectric medium that infinitely spreads two-dimensionally and that the plate is sandwiched with another dielectric (or maybe air or vacuum) or metal. In this situation, electromagnetic waves are confined within the slab. Moreover, only under a restricted condition those waves are allowed to propagate in parallel to the slab

plane. Such electromagnetic waves are usually called propagation modes or simply "modes." An optical device thus designed is called a *waveguide*. These modes are characterized by repeated total reflection during the propagation. Another mode is an evanescent mode. Because of the total reflection, the energy transport is not allowed to take place vertically to the interface of two dielectrics and the evanescent mode is thought to be propagated very close to the interface.

### 6.7.1 TE and TM Waves in a Waveguide

In a waveguide configuration, propagating waves are classified into TE and TM modes. Quality of materials that constitute a waveguide largely governs the propagation modes within the waveguide.

Figure 6.12 depicts a cross section of a slab waveguide. We assume that the electromagnetic wave is propagated toward the positive direction of the $z$-axis and that a waveguide infinitely spreads toward the $z$- and $x$-axis. Suppose that the said waveguide is spatially confined toward the $y$-axis. Let the thickness of the waveguide be $d$. From a point of view of material that shapes a waveguide, waveguides are classified into two types. (i) Electromagnetic waves are completely confined within the waveguide layer. This case typically happens when a dielectric forming the waveguide is sandwiched between a couple of metal layers (Fig. 6.12a). This is because the electromagnetic wave is not allowed to exist or propagate inside the metal.

(ii) Electromagnetic waves are not completely confined within the waveguide. This case happens when the dielectric of the waveguide is sandwiched by a couple of other dielectrics. We distinguish this case as the *total internal reflection* from the above case (i). We further describe it in Sect. 6.8.2. For the total internal reflection to take place, the refractive index of the waveguide must be higher than those of other dielectrics (Fig. 6.12b). The dielectric of the waveguide is called core layer and the other dielectric is called clad layer. In this case, electromagnetic waves are allowed to propagate inside of the clad layer, even though the region is confined very close to the interface between the clad and core layers. Such electromagnetic

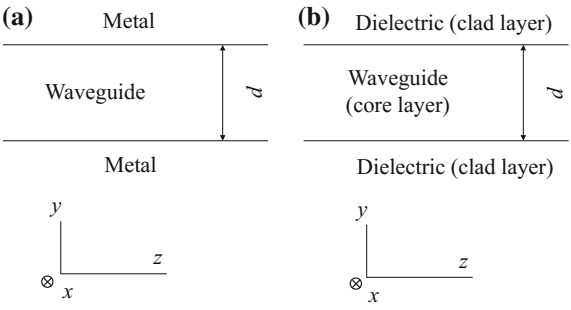


**Fig. 6.12** Cross section of a slab waveguide comprising a dielectric medium. **a** A waveguide is sandwiched between a couple of metal layers. **b** A waveguide is sandwiched between a couple of layers consisting of another dielectric called clad layer. The sandwiched layer is called core layer

waves are said to be an evanescent wave. According to these two cases (i) and (ii), we have different conditions under which the allowed modes can exist.

Now, let us return to Maxwell's equations. We have introduced the equations of wave motion (5.35) and (5.36) from Maxwell's equations (5.28) and (5.29) along with (5.7) and (5.10). One of their simplest solutions is a plane wave described by (5.53). The plane wave is characterized by that the wave has the same phase on an infinitely spreading plane perpendicular to the propagation direction (characterized by a wavevector $\boldsymbol{k}$). In a waveguide, however, the electromagnetic field is confined with respect to the direction parallel to the normal to the slab plane (i.e., the direction of the $y$-axis in Fig. 6.12). Consequently, the electromagnetic field can no longer have the same phase in that direction. Yet, as solutions of equations of wave motion, we can have a solution that has the same phase with the direction of the $x$-axis. Bearing in mind such a situation, let us think of Maxwell's equations in relation to the equations of wave motion.

Ignoring components related to partial differentiation with respect to $x$ (i.e., the component related to $\partial/\partial x$) and rewriting (5.28) and (5.29) for individual Cartesian coordinates, we have [3]

$$\frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z} + \frac{\partial B_x}{\partial t} = 0, \tag{6.130}$$

$$\frac{\partial E_x}{\partial z} + \frac{\partial B_y}{\partial t} = 0, \tag{6.131}$$

$$-\frac{\partial E_x}{\partial y} + \frac{\partial B_z}{\partial t} = 0, \tag{6.132}$$

$$\frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z} - \frac{\partial D_x}{\partial t} = 0, \tag{6.133}$$

$$\frac{\partial H_x}{\partial z} - \frac{\partial D_y}{\partial t} = 0, \tag{6.134}$$

$$-\frac{\partial H_x}{\partial y} - \frac{\partial D_z}{\partial t} = 0. \tag{6.135}$$

Of the above equations, we collect those pertinent to $E_x$ and differentiate (6.131), (6.132), and (6.133) with respect to $z$, $y$, and $t$, respectively, to get

$$\frac{\partial^2 E_x}{\partial z^2} + \frac{\partial^2 B_y}{\partial z \partial t} = 0, \tag{6.136}$$

$$\frac{\partial^2 E_x}{\partial y^2} - \frac{\partial^2 B_z}{\partial y \partial t} = 0, \tag{6.137}$$

$$\frac{\partial^2 H_z}{\partial t \partial y} - \frac{\partial^2 H_y}{\partial t \partial z} - \frac{\partial^2 D_x}{\partial t^2} = 0. \tag{6.138}$$

Multiplying (6.138) by $\mu$ and further adding (6.136) and (6.137) to it and using (5.7), we get

$$\frac{\partial^2 E_x}{\partial y^2} + \frac{\partial^2 E_x}{\partial z^2} = \mu\varepsilon \frac{\partial^2 E_x}{\partial t^2}. \tag{6.139}$$

This is a two-dimensional equation of wave motion. In a similar manner, from (6.130), (6.134), and (6.135), we have for the magnetic field

$$\frac{\partial^2 H_x}{\partial y^2} + \frac{\partial^2 H_x}{\partial z^2} = \mu\varepsilon \frac{\partial^2 H_x}{\partial t^2}. \tag{6.140}$$

Equations (6.139) and (6.140) are two-dimensional wave equations with respect to the $y$- and $z$-coordinates. With the direction of the $x$-axis, a propagating wave has the same phase. Suppose that we have plane wave solutions for them as in the case of (5.58) and (5.59). Then, we have

$$\begin{aligned} \boldsymbol{E} &= \boldsymbol{E}_0 \mathrm{e}^{i(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)} = \boldsymbol{E}_0 \mathrm{e}^{i(k\boldsymbol{n}\cdot\boldsymbol{x}-\omega t)}, \\ \boldsymbol{H} &= \boldsymbol{H}_0 \mathrm{e}^{i(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)} = \boldsymbol{H}_0 \mathrm{e}^{i(k\boldsymbol{n}\cdot\boldsymbol{x}-\omega t)}. \end{aligned} \tag{6.141}$$

Note that a plane wave expressed by (5.58) and (5.59) is propagated uniformly in a dielectric medium. In a waveguide, on the other hand, the electromagnetic waves undergo repeated (total) reflections from the two boundaries positioned either side of the waveguide, while being propagated.

In a three-dimensional version, the wavenumber vector has three components $k_x$, $k_y$, and $k_z$ as expressed in (5.48). In (6.141), in turn, $\boldsymbol{k}$ has $y$ and $z$ components such that

$$\boldsymbol{k}^2 = k^2 = k_y^2 + k_z^2. \tag{6.142}$$

Equations (6.139) and (6.140) can be rewritten as

$$\frac{\partial^2 E_x}{\partial(\pm y)^2} + \frac{\partial^2 E_x}{\partial(\pm z)^2} = \mu\varepsilon \frac{\partial^2 E_x}{\partial(\pm t)^2}, \quad \frac{\partial^2 H_x}{\partial(\pm y)^2} + \frac{\partial^2 H_x}{\partial(\pm z)^2} = \mu\varepsilon \frac{\partial^2 H_x}{\partial(\pm t)^2}.$$

Accordingly, we have four wavenumber vector components

$$k_y = \pm|k_y| \quad \text{and} \quad k_z = \pm|k_z|.$$

Figure 6.13 indicates this situation where an electromagnetic wave can be propagated within a slab waveguide in either one direction out of four choices of

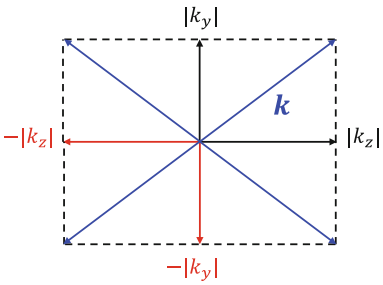


**Fig. 6.13** Four possible propagation directions $\boldsymbol{k}$ of an electromagnetic wave in a slab waveguide

$\boldsymbol{k}$. In this section, we assume that the electromagnetic wave is propagated toward the positive direction of the $z$-axis, and so we define $k_z$ as positive. On the other hand, $k_y$ can be either positive or negative. Thus, we get

$$k_z = k \sin\theta \quad \text{and} \quad k_y = \pm k \cos\theta. \tag{6.143}$$

Figure 6.14 shows the geometries of the electromagnetic waves within the slab waveguide. The slab plane is parallel to the $zx$-plane. Let the positions of the two interfaces of the slab waveguide be

$$y = 0 \quad \text{and} \quad y = d. \tag{6.144}$$

That is, we assume that the thickness of the waveguide is $d$.

Since (6.139) describes a wave equation for only one component $E_x$, (6.139) is suited for representing a TE wave. In (6.140), in turn, a wave equation is given only for $H_x$, and hence, it is suited for representing a TM wave. With the TE wave, the electric field oscillates parallel to the slab plane and vertical to the propagation direction. With the TM wave, in turn, the magnetic field oscillates parallel to the slab plane and vertical to the propagation direction. In a general case, electromagnetic waves in a slab waveguide are formed by superposition of TE and TM waves. Notice that Fig. 6.14 is applicable to both TE and TM waves.

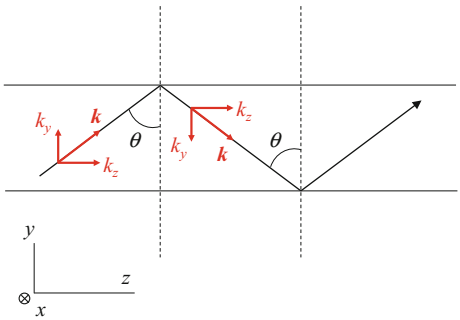


**Fig. 6.14** Geometries and propagation of the electromagnetic waves in a slab waveguide

Let us further proceed with the waveguide analysis. The electric field $\boldsymbol{E}$ within the waveguide is described by superposition of the incident and reflected waves. Using the first equation of (6.141) and (6.143), we have

$$\boldsymbol{E}(z,y) = \boldsymbol{\varepsilon}_{\mathrm{e}} E_{+}\, \mathrm{e}^{i(kz\,\sin\theta + ky\,\cos\theta - \omega t)} + \boldsymbol{\varepsilon}'_{\mathrm{e}} E_{-} \mathrm{e}^{i(kz\,\sin\theta - ky\,\cos\theta - \omega t)}, \tag{6.145}$$

where $E_{+}$ ($E_{-}$) and $\boldsymbol{\varepsilon}_{\mathrm{e}}$ $\left(\boldsymbol{\varepsilon}'_{\mathrm{e}}\right)$ represent an amplitude and unit polarization vector of the incident (reflected) waves, respectively. The vector $\boldsymbol{\varepsilon}_{\mathrm{e}}$ (or $\boldsymbol{\varepsilon}'_{\mathrm{e}}$) is defined in (5.67).

Equation (6.145) is common to both the cases of TE and TM waves. From now, we consider the TE mode case. Suppose that the slab waveguide is sandwiched with a couple of metal sheet of high conductance. Since the electric field must be absent inside the metal, the electric field at the interface must be zero owing to the continuity condition of a tangential component of the electric field. Thus, we require the following condition should be met with (6.145):

$$\begin{aligned} \boldsymbol{t}\cdot\boldsymbol{E}(z,0) = 0 &= \boldsymbol{t}\cdot\boldsymbol{\varepsilon}_{\mathrm{e}} E_{+}\, \mathrm{e}^{i(kz\,\sin\theta - \omega t)} + \boldsymbol{t}\cdot\boldsymbol{\varepsilon}'_{\mathrm{e}} E_{-} \mathrm{e}^{i(kz\,\sin\theta - \omega t)} \\ &= \left(\boldsymbol{t}\cdot\boldsymbol{\varepsilon}_{\mathrm{e}} E_{+} + \boldsymbol{t}\cdot\boldsymbol{\varepsilon}'_{\mathrm{e}} E_{-}\right) \mathrm{e}^{i(kz\,\sin\theta - \omega t)} \end{aligned} \tag{6.146}$$

Therefore, since $\mathrm{e}^{i(kz\,\sin\theta - \omega t)}$ never vanishes, we have

$$\boldsymbol{t}\cdot\boldsymbol{\varepsilon}_{\mathrm{e}} E_{+} + \boldsymbol{t}\cdot\boldsymbol{\varepsilon}'_{\mathrm{e}} E_{-} = \boldsymbol{0}, \tag{6.147}$$

where $\boldsymbol{t}$ is a tangential unit vector at the interface.

Since $\boldsymbol{E}$ is polarized along the $x$-axis, setting $\boldsymbol{\varepsilon}_{\mathrm{e}} = \boldsymbol{\varepsilon}'_{\mathrm{e}} = \boldsymbol{e}_1$ and taking $\boldsymbol{t}$ as $\boldsymbol{e}_1$, we get

$$E_{+} + E_{-} = 0.$$

This means that the reflection coefficient of the electric field is $-1$. Denoting $E_{+} = -E_{-} \equiv E_0$ ($> 0$), we have

$$\begin{aligned} \boldsymbol{E} &= \boldsymbol{e}_1 E_0 \left[\mathrm{e}^{i(kz\,\sin\theta + ky\,\cos\theta - \omega t)} - \mathrm{e}^{i(kz\,\sin\theta - ky\,\cos\theta - \omega t)}\right] \\ &= \boldsymbol{e}_1 E_0 \left[\left(\mathrm{e}^{iky\,\cos\theta} - \mathrm{e}^{-iky\,\cos\theta}\right) \mathrm{e}^{i(kz\,\sin\theta - \omega t)}\right] \\ &= \boldsymbol{e}_1 2iE_0\, \sin(ky\,\cos\theta)\, \mathrm{e}^{i(kz\,\sin\theta - \omega t)}. \end{aligned} \tag{6.148}$$

Requiring the electric field to vanish at another interface of $y = d$, we have

$$\boldsymbol{E}(z,d) = \boldsymbol{0} = \boldsymbol{e}_1 2iE_0\, \sin(kd\,\cos\theta)\, \mathrm{e}^{i(kz\,\sin\theta - \omega t)}.$$

Note that in terms of the boundary conditions we are thinking of Dirichlet conditions (see Sects. 1.3 and 8.3). In this case, we have nodes for the electric field at the interface between metal and a dielectric. For this condition to be satisfied, we must have

$$kd\cos\theta = m\pi \ (m = 1, 2, \cdots). \tag{6.149}$$

From (6.149), we have a following condition for $m$:

$$m \leq kd/\pi. \tag{6.150}$$

Meanwhile, we have

$$k = nk_0, \tag{6.151}$$

where $n$ is a refractive index of a dielectric that shapes the slab waveguide; the quantity $k_0$ is a wavenumber of the electromagnetic wave in vacuum. The index $n$ is given by

$$n = c/v, \tag{6.152}$$

where $c$ and $v$ are light velocity in vacuum and the dielectric media, respectively. Here, $v$ is meant as a velocity in an infinitely spreading dielectric. Thus, $\theta$ is allowed to take several (or more) numbers depending upon $k$, $d$, and $m$.

Since in the $z$-direction no specific boundary conditions are imposed, we have propagating modes in that direction characterized by a propagation constant (vide infra). Looking at (6.148), we notice that $k\sin\theta$ plays a role of a wavenumber in a free space. For this reason, a quantity $\beta$ defined as

$$\beta = k\sin\theta = nk_0\sin\theta \tag{6.153}$$

is said to be a propagation constant. From (6.149) and (6.153), we get

$$\beta = \left(k^2 - \frac{m^2\pi^2}{d^2}\right)^{1/2}. \tag{6.154}$$

Thus, the allowed TE waves indexed by $m$ are called TE *modes* and represented as $\mathrm{TE}_m$. The phase velocity $v_p$ is given by

$$v_p = \omega/\beta. \tag{6.155}$$

Meanwhile, the group velocity $v_g$ is given by

$$v_g = \frac{\mathrm{d}\omega}{\mathrm{d}\beta} = \left(\frac{\mathrm{d}\beta}{\mathrm{d}\omega}\right)^{-1}. \tag{6.156}$$

Using (6.154) and noting that $k^2 = \omega^2/v^2$, we get

$$v_g = v^2\beta/\omega. \tag{6.157}$$

Thus, we have

$$v_p v_g = v^2. \tag{6.158}$$

Note that in (1.22) of Sect. 1.1 we saw a relationship similar to (6.158).

The characteristics of TM waves can be analyzed in a similar manner by examining the magnetic field $H_x$. In that case, the reflection coefficient of the magnetic field is +1 and we have antinodes for the magnetic field at the interface. Concomitantly, we adopt Neumann conditions as the boundary conditions (see Sects. 1.3 and 8.3). Regardless of the difference in the boundary conditions, however, discussion including (6.149) to (6.158) applies to the analysis of TM waves. Once $H_x$ is determined, $E_y$ and $E_z$ can be determined as well from (6.134) and (6.135).

### 6.7.2 Total Internal Reflection and Evanescent Waves

If a slab waveguide shaped by a dielectric is sandwiched by a couple of another dielectric (Fig. 6.12b), the situation differs from a metal waveguide (Fig. 6.12a) we encountered in Sect. 6.8.1. Suppose in Fig. 6.12b that the former dielectric D1 of a refractive index $n_1$ is sandwiched by the latter dielectric D2 of a refractive index $n_2$. Suppose that an electromagnetic wave is being propagated from D1 toward D2. Then, we must have

$$n_1 > n_2 \tag{6.159}$$

so that the total internal reflection can take place at the interface of D1 and D2. In this case, the dielectrics D1 and D2 act as a core layer and a clad layer, respectively.

The biggest difference between the present waveguide and the previous one is that unlike the previous case, the total internal reflection occurs in the present case. Concomitantly, an evanescent wave is present in the clad layer very close to the interface.

First, let us estimate the conditions that are satisfied so that an electromagnetic wave can be propagated within a waveguide. Figure 6.15 depicts a cross section of the waveguide where the light is propagated in the direction of $\boldsymbol{k}$. In Fig. 6.15, suppose that we have a normal N to the plane of paper at P. Then, N and a straight line XY shape a plane NXY. Also, suppose that a dielectric fills a semi-infinite space situated below NXY. Further, suppose that there is another virtual plane N′X′Y′ that is parallel with NXY as shown. Here N′ is parallel to N. The separation of the two parallel planes is $d$. We need the virtual plane N′X′Y′ just to estimate an optical path difference (or phase difference, more specifically) between two waves, i.e., a propagating wave and a reflected wave.

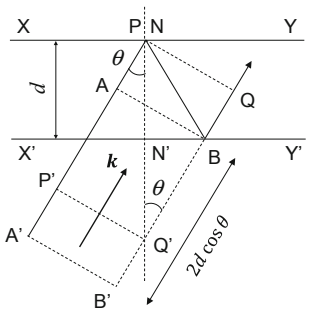


**Fig. 6.15** Cross section of the waveguide where the light is propagated in the direction of $\boldsymbol{k}$. A dielectric fills a semi-infinite space situated below NXY. We suppose another virtual plane N′X′Y′ that is parallel with NXY. We need the plane N′X′Y′ to estimate an optical path difference (or phase difference)

Let $\boldsymbol{n}$ be a unit vector in the direction of $\boldsymbol{k}$; i.e.,

$$\boldsymbol{n} = \boldsymbol{k}/|\boldsymbol{k}| = \boldsymbol{k}/k. \tag{6.160}$$

Then, the electromagnetic wave is described as

$$\boldsymbol{E} = \boldsymbol{E}_0 \mathrm{e}^{i(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)} = \boldsymbol{E}_0 \mathrm{e}^{i(k\boldsymbol{n}\cdot\boldsymbol{x}-\omega t)}. \tag{6.161}$$

Suppose we take a coordinate system such that

$$\boldsymbol{x} = r\boldsymbol{n} + s\boldsymbol{u} + t\boldsymbol{v}, \tag{6.162}$$

where $\boldsymbol{u}$ and $\boldsymbol{v}$ represent unit vectors in the direction perpendicular to $\boldsymbol{n}$. Then, (6.161) can be expressed by

$$\boldsymbol{E} = \boldsymbol{E}_0 \mathrm{e}^{i(kr-\omega t)}. \tag{6.163}$$

Suppose that the wave is propagated starting from a point A to P and reflected at P. Then, the wave is further propagated to B and reflected again to reach Q. The wave front is originally at AB and finally at PQ. Thus, the Z-shaped optical path length APBQ is equal to a separation between A′B′ and PQ. Notice that the separation between A′B′ and P′Q′ is taken so that it is equal to that between AB and PQ. The geometry of Fig. 6.15 implies that two waves starting from AB and A′B′ at once reach PQ again at once.

We find the separation between AB and A′B′ is

$$2d\,\cos\theta.$$

Let us tentatively call these waves Wave-AB and Wave-A′B′ and describe their electric fields as $\boldsymbol{E}_{\mathrm{AB}}$ and $\boldsymbol{E}_{\mathrm{A'B'}}$, respectively. Then, we denote

$$\boldsymbol{E}_{\mathrm{AB}} = \boldsymbol{E}_0 \mathrm{e}^{i(kr-\omega t)}, \tag{6.164}$$

$$\boldsymbol{E}_{\mathrm{A'B'}} = \boldsymbol{E}_0 \mathrm{e}^{i[k(r+2d\,\cos\theta)-\omega t]}, \tag{6.165}$$

where $k$ is a wavenumber in the dielectric. Note that since $\boldsymbol{E}_{\mathrm{A'B'}}$ gets behinds $\boldsymbol{E}_{\mathrm{AB}}$, a plus sign appears in the first term of the exponent. Therefore, the phase difference between the two waves is

$$2kd\,\cos\theta. \tag{6.166}$$

Now, let us come back to the actual geometry of the waveguide. That is, the core layer of thickness $d$ is sandwiched by a couple of clad layers (Fig. 6.12b). In this situation, the wave $\boldsymbol{E}_{\mathrm{AB}}$ experiences the total internal reflection two times, which we ignored in the above discussion of the metal waveguide. Since the total internal reflection causes a complex phase shift, we have to take account of this effect. The phase shift was defined as $\alpha$ of (6.114) for a TE mode and $\beta$ for a TM mode. Notice that in Fig. 6.15 the electric field oscillates perpendicularly to the plane of paper with the TE mode, whereas it oscillates in parallel with the plane of paper with the TM mode. For both the cases, the electric field oscillates perpendicularly to $\boldsymbol{n}$. Consequently, the phase shift due to these reflections has to be added to (6.166). Thus, for the phase commensuration to be obtained, the following condition must be satisfied:

$$2kd\,\cos\theta + 2\delta = 2m\pi\ (m = 0, 1, 2, \cdots), \tag{6.167}$$

where $\delta$ is either $\delta_{\mathrm{TE}}$ or $\delta_{\mathrm{TM}}$ defined below according to the case of the TE wave and TM wave, respectively. For a practical purpose, (6.167) is dealt with by a numerical calculation, e.g., to design an optical waveguide.

Unlike (6.149), what is the most important with (6.167) is that the condition $m = 0$ is permitted because of $\delta < 0$ (see just below).

For convenience and according to the custom, we adopt a phase shift notation other than that defined in (6.114). With the TE mode, the phase is retained upon reflection at the critical angle, and so we identify $\alpha$ with an additional component $\delta_{\mathrm{TE}}$. In the TM case, on the other hand, the phase is reversed upon reflection at the critical angle (i.e., a $\pi$ shift occurs). Since this $\pi$ shift has been incorporated into $\beta$, it suffices to consider only an additional component $\delta_{\mathrm{TM}}$. That is, we have

$$\delta_{\mathrm{TE}} \equiv \alpha \quad \text{and} \quad \delta_{\mathrm{TM}} \equiv \beta - \pi. \tag{6.168}$$

We rewrite (6.110) as

$$\begin{aligned} &R_E^{\perp} = \mathrm{e}^{i\alpha} = \mathrm{e}^{i\delta_{\mathrm{TE}}} \equiv \frac{a\mathrm{e}^{-i\sigma}}{a\mathrm{e}^{i\sigma}} = \mathrm{e}^{-2i\sigma} \quad \text{and} \\ &\delta_{\mathrm{TE}} = -2\sigma\ (\sigma > 0), \end{aligned} \tag{6.169}$$

where we have

$$ae^{i\sigma} = \cos\theta + i\sqrt{\sin^2\theta - n^2}. \tag{6.170}$$

Therefore,

$$\tan\sigma = -\tan\frac{\delta_{\mathrm{TE}}}{2} = \frac{\sqrt{\sin^2\theta - n^2}}{\cos\theta}. \tag{6.171}$$

Meanwhile, rewriting (6.112) we have

$$R_E^{\parallel} = e^{i\beta} = e^{i(\delta_{\mathrm{TM}}+\pi)} \equiv -\frac{be^{-i\tau}}{be^{i\tau}} = -e^{-2i\tau} = e^{i(-2\tau+\pi)}, \tag{6.172}$$

where

$$be^{i\tau} = n^2\cos\theta + i\sqrt{\sin^2\theta - n^2}. \tag{6.173}$$

Note that the minus sign with the second last equality in (6.172) is due to the phase reversal upon reflection. From (6.172), we may put

$$\beta = -2\tau + \pi.$$

Comparing this with the second equation of (6.168), we get

$$\delta_{\mathrm{TM}} = -2\tau \ (\tau > 0). \tag{6.174}$$

Consequently, we get

$$\tan\tau = -\tan\frac{\delta_{\mathrm{TM}}}{2} = \frac{\sqrt{\sin^2\theta - n^2}}{n^2\cos\theta}. \tag{6.175}$$

Finally, the additional phase change $\delta_{\mathrm{TE}}$ and $\delta_{\mathrm{TM}}$ upon the total reflection is given by [4]

$$\delta_{\mathrm{TE}} = -2\ \tan^{-1}\frac{\sqrt{\sin^2\theta - n^2}}{\cos\theta} \quad \text{and} \quad \delta_{\mathrm{TM}} = -2\ \tan^{-1}\frac{\sqrt{\sin^2\theta - n^2}}{n^2\cos\theta}. \tag{6.176}$$

We emphasize that in (6.176) both $\delta_{\mathrm{TE}}$ and $\delta_{\mathrm{TM}}$ are negative quantities. This phase shift has to be included in (6.167) as a negative quantity $\delta$.

At a first glance, (6.176) seems to differ largely from (6.120) and (6.125). Nevertheless, noting that a trigonometric formula

$$\tan 2x = \frac{2\ \tan x}{1 - \tan^2 x}$$

and remembering that $\delta_{\mathrm{TM}}$ in (6.168) includes $\pi$ arising from the phase reversal, we find that both the relations are virtually identical.

Evanescent waves are drawing a large attention in the field of basic physics and applied device physics. If the total internal reflection is absent, $\phi$ is real. But, under the total internal reflection, $\phi$ is pure imaginary. The electric field of the evanescent wave is described as

$$\begin{aligned} \boldsymbol{E}_{\mathrm{t}} &= E\boldsymbol{\varepsilon}_{\mathrm{t}}\mathrm{e}^{i(k_t z\,\sin\phi + k_t y\,\cos\phi - \omega t)} = E\boldsymbol{\varepsilon}_{\mathrm{t}}\mathrm{e}^{i(k_t z\,\sin\phi + k_t y i b - \omega t)} \\ &= E\boldsymbol{\varepsilon}_{\mathrm{t}}\mathrm{e}^{i(k_t z\,\sin\phi - \omega t)}\mathrm{e}^{-k_t y b}. \end{aligned} \tag{6.177}$$

In (6.177), a unit polarization vector $\boldsymbol{\varepsilon}_{\mathrm{t}}$ is either perpendicular to the plane of paper of Fig. 6.15 (the TE case) or in parallel to it (the TM case). Notice that the coordinate system is different from that of (6.104). The quantity $k_t \sin\phi$ is the propagation constant. Let $v_p^{(s)}$ and $v_p^{(e)}$ be a phase velocity of the electromagnetic wave in the slab waveguide (i.e., core layer) and evanescent wave in the clad layer, respectively. Then, in virtue of Snell's law we have

$$v_1 < v_p^{(s)} = \frac{v_1}{\sin\theta} = \frac{\omega}{k\,\sin\theta} = \frac{\omega}{k_t\,\sin\phi} = v_p^{(e)} = \frac{v_2}{\sin\phi} < v_2, \tag{6.178}$$

where $v_1$ and $v_2$ are light velocity in a free space filled by the dielectric D1 and D2, respectively. For this, we used a relation described as

$$\omega = v_1 k = v_2 k_t. \tag{6.179}$$

We also used Snell's law with the third equality. Notice that $\sin\phi > 1$ in the evanescent region and that $k_t \sin\phi$ is a propagation constant in the clad layer. Also, note that $v_p^{(s)}$ is equal to $v_p^{(e)}$ and that these phase velocities are in between the two velocities of the free space. Thus, the evanescent waves must be present, accompanying propagating waves that undergo the total internal reflections in a slab waveguide.

As remarked in (6.105), the electric field of evanescent waves decays exponentially with increasing $z$. This implies that the evanescent waves exist only in the clad layer very close to an interface of core and clad layers.

## 6.8 Stationary Waves

So far, we have been dealing with propagating waves either in a free space or in a waveguide. If the dielectric shaping the waveguide is confined in another direction, the propagating waves show specific properties. Examples include optical fibers.

In this section, we consider a situation where the electromagnetic wave is propagating in a dielectric medium and reflected by a "wall" formed by metal or

another dielectric. In such a situation, the original wave (i.e., a forward wave) causes interference with the backward wave and a stationary wave is formed as a consequence of the interference.

To approach this issue, we deal with superposition of two waves that have different phases and different amplitudes. To generalize the problem, let $\psi_1$ and $\psi_2$ be two cosine functions described as

$$\psi_1 = a_1 \cos b_1 \quad \text{and} \quad \psi_2 = a_2 \cos b_2. \tag{6.180}$$

Their addition is expressed as

$$\psi = \psi_1 + \psi_2 = a_1 \cos b_1 + a_2 \cos b_2. \tag{6.181}$$

Here, we wish to unify (6.181) as a single cosine (or sine) function. To this end, we modify a description of $\psi_2$ such that

$$\begin{aligned} \psi_2 &= a_2 \cos[b_1 + (b_2 - b_1)] \\ &= a_2[\cos b_1 \cos(b_2 - b_1) - \sin b_1 \sin(b_2 - b_1)] \\ &= a_2[\cos b_1 \cos(b_1 - b_2) + \sin b_1 \sin(b_1 - b_2)]. \end{aligned} \tag{6.182}$$

Then, the addition is described by

$$\begin{aligned} \psi &= \psi_1 + \psi_2 \\ &= [a_1 + a_2 \cos(b_1 - b_2)] \cos b_1 + a_2 \sin(b_1 - b_2) \sin b_1. \end{aligned} \tag{6.183}$$

Putting $R$ such that

$$\begin{aligned} R &= \sqrt{[a_1 + a_2 \cos(b_1 - b_2)]^2 + a_2^2 \sin^2(b_1 - b_2)} \\ &= \sqrt{a_1^2 + a_2^2 + 2a_1 a_2 \cos(b_1 - b_2)}, \end{aligned} \tag{6.184}$$

we get

$$\psi = R \cos(b_1 - \theta), \tag{6.185}$$

where $\theta$ is expressed by

$$\tan \theta = \frac{a_2 \sin(b_1 - b_2)}{a_1 + a_2 \cos(b_1 - b_2)}. \tag{6.186}$$

Figure 6.16 represents a geometrical diagram in relation to the superposition of two waves having different amplitudes ($a_1$ and $a_2$) and different phases ($b_1$ and $b_2$) [5].

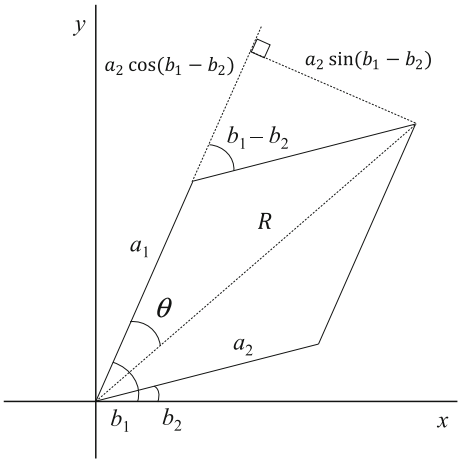


**Fig. 6.16** Geometrical diagram in relation to the superposition of two waves having different amplitudes ($a_1$ and $a_2$) and different phases ($b_1$ and $b_2$)

To apply (6.185) to the superposition of two electromagnetic waves that are propagating forward and backward to collide head-on with each other, we change the variables such that

$$b_1 = kx - \omega t \quad \text{and} \quad b_2 = -kx - \omega t, \tag{6.187}$$

where the former equation represents a forward wave, whereas the latter a backward wave. Then, we have

$$b_1 - b_2 = 2kx. \tag{6.188}$$

Equation (6.185) is rewritten as

$$\psi(x,t) = R\cos(kx - \omega t - \theta), \tag{6.189}$$

with

$$R = \sqrt{a_1^2 + a_2^2 + 2a_1a_2\cos 2kx} \tag{6.190}$$

and

$$\tan\theta = \frac{a_2\sin 2kx}{a_1 + a_2\cos 2kx}. \tag{6.191}$$

Equation (6.189) looks simple, but since both $R$ and $\theta$ vary as a function of $x$, the situation is somewhat complicated unlike a simple sinusoidal wave. Nonetheless, when $x$ takes a special value, (6.189) is expressed by a simple function form. For example, at $t = 0$,

$$\psi(x,0) = (a_1 + a_2)\cos kx.$$

This corresponds to (i) of Fig. 6.17. If $t = T/2$ (where $T$ is a period, i.e., $T = 2\pi/\omega$), we have

$$\psi(x, T/2) = -(a_1 + a_2)\cos kx.$$

This corresponds to (iii) of Fig. 6.17. But, the waves described by (ii) or (iv) do not have a simple function form.

We characterize Fig. 6.17. If we have

$$2kx = n\pi \quad \text{or} \quad x = n\lambda/4 \tag{6.192}$$

with $\lambda$ being a wavelength, then $\theta = 0$ or $\pi$, and so $\theta$ can be eliminated. This situation occurs with every quarter period of a wavelength. Let us put $t = 0$ and examine how the superposed wave looks like. For instance, putting $x = 0$, $x = \lambda/4$, and $x = \lambda/2$ we have

$$\psi(0, 0) = |a_1 + a_2|,\ \psi(\lambda/4, 0) = \psi(3\lambda/4, 0) = 0,\ \psi(\lambda/2, 0) = -|a_1 + a_2|, \tag{6.193}$$

respectively. Notice that in Fig. 6.17 we took $a_1, a_2 > 0$. At another instant $t = T/4$, we have similarly

$$\begin{aligned} &\psi(0, T/4) = 0,\ \psi(\lambda/4, T/4) = |a_1 - a_2|,\ \psi(\lambda/2, T/4) = 0, \\ &\psi(\lambda/4, T/4) = -|a_1 - a_2|. \end{aligned} \tag{6.194}$$

Thus, the waves that vary with time are characterized by two dram-shaped envelopes that have extremals $|a_1 + a_2|$ and $|a_1 - a_2|$ or those $-|a_1 + a_2|$ and $-|a_1 - a_2|$. An important implication of Fig. 6.17 is that no node is present in the superposed wave. In other words, there is no instant $t_0$ when $\psi(x, t_0) = 0$ for any $x$. From the aspect of energy transport of electromagnetic waves, if $|a_1| > |a_2|$ (where $a_1$ and $a_2$ represent an amplitude of the forward and backward waves, respectively), the net energy flow takes place in the traveling direction of the forward wave. If, on the other hand, $|a_1| > |a_2|$, the net energy flow takes place in the traveling direction of the backward wave. In this respect, think of Poynting vectors.

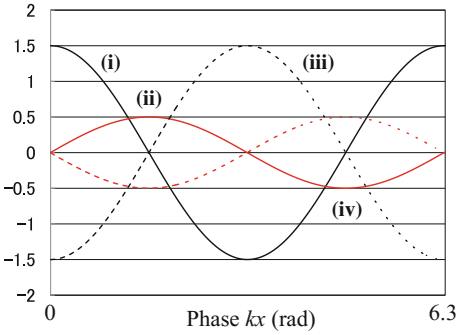


**Fig. 6.17** Superposition of two sinusoidal waves. In (6.189) and (6.190), we put $a_1 = 1$, $a_2 = 0.5$, with (i) $t = 0$. (ii) $t = T/4$. (iii) $t = T/2$. (iv) $t = 3T/4$. $\psi(x, t)$ is plotted as a function of phase $kx$

In case $|a_1| = |a_2|$, the situation is particularly simple. No net energy flow takes place in this case. Correspondingly, we observe nodes. Such waves are called stationary waves. Let us consider this simple situation for an electromagnetic wave that is incident perpendicularly to the interface between two dielectrics (one of them may be a metal).

Returning back to (5.58) and (5.66), we describe two electromagnetic waves that are propagating in the positive and negative directions of the $z$-axis such that

$$\boldsymbol{E}_1 = E_1 \boldsymbol{\varepsilon}_\mathrm{e} \mathrm{e}^{i(kz-\omega t)} \quad \text{and} \quad \boldsymbol{E}_2 = E_2 \boldsymbol{\varepsilon}_\mathrm{e} \mathrm{e}^{i(-kz-\omega t)}, \tag{6.195}$$

where $\boldsymbol{\varepsilon}_\mathrm{e}$ is a unit polarization vector arbitrarily fixed so that it can be parallel to the interface, i.e., wall (i.e., perpendicular to the $z$-axis). The situation is depicted in Fig. 6.18. Notice that in Fig. 6.18 $\boldsymbol{E}_1$ represents the forward wave (i.e., incident wave) and $\boldsymbol{E}_2$ the backward wave (i.e., wave reflected at the interface). Thus, a superposed wave is described as

$$\boldsymbol{E} = \boldsymbol{E}_1 + \boldsymbol{E}_2. \tag{6.196}$$

Taking account of the reflection of an electromagnetic wave perpendicularly incident on a wall, let us consider following two cases:

(i) Syn-phase:

The phase of the electric field is retained upon reflection. We assume that $E_1 = E_2\ (> 0)$. Then, we have

$$\begin{aligned} \boldsymbol{E} &= E_1 \boldsymbol{\varepsilon}_\mathrm{e} \left[ \mathrm{e}^{i(kz-\omega t)} + \mathrm{e}^{i(-kz-\omega t)} \right] = E_1 \boldsymbol{\varepsilon}_\mathrm{e} \mathrm{e}^{-i\omega t} \left( \mathrm{e}^{ikz} + \mathrm{e}^{-ikz} \right) \\ &= 2E_1 \boldsymbol{\varepsilon}_\mathrm{e} \mathrm{e}^{-i\omega t} \cos kz. \end{aligned} \tag{6.197}$$

In (6.197), we put $z = 0$ at the interface for convenience. Taking a real part of (6.197), we have

$$\boldsymbol{E} = 2E_1 \boldsymbol{\varepsilon}_\mathrm{e} \cos \omega t \cos kz. \tag{6.198}$$

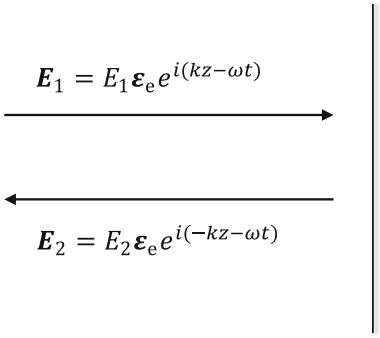


**Fig. 6.18** Superposition of electric fields of forward (or incident) wave $\boldsymbol{E}_1$ and backward (or reflected) wave $\boldsymbol{E}_2$

Note that in (6.198) variables $z$ and $t$ have been separated. This implies that we have a stationary wave. For this case to be realized, the characteristic impedance of the dielectric of the incident wave side should be smaller enough than that of the other side [see (6.51) and (6.59)]. In other words, the dielectric constant of the incident side should be large enough. We have nodes at positions that satisfy

$$-kz = \frac{\pi}{2} + m\pi \ (m = 0, 1, 2, \cdots) \quad \text{or} \quad -z = \frac{1}{4}\lambda + \frac{m}{2}\lambda. \tag{6.199}$$

Note that we are thinking of the stationary wave in the region of $z<0$. Equation (6.199) indicates that nodes are formed at a quarter wavelength from the interface and every half wavelength from it. The node means the position where no electric field is present.

Meanwhile, antinodes are observed at positions

$$-kz = m\pi \ (m = 0, 1, 2, \cdots) \quad \text{or} \quad -z = +\frac{m}{2}\lambda.$$

Thus, the nodes and antinodes alternate with every quarter wavelength.

(ii) Anti-phase:

The phase of the electric field is reversed. We assume that $E_1 = -E_2\ (>0)$. Then, we have

$$\begin{aligned} \boldsymbol{E} &= E_1\varepsilon_e\left[e^{i(kz-\omega t)} - e^{i(-kz-\omega t)}\right] = E_1\varepsilon_e e^{-i\omega t}\left(e^{ikz} - e^{-ikz}\right) \\ &= 2iE_1\varepsilon_e e^{-i\omega t}\sin kz. \end{aligned} \tag{6.200}$$

Taking a real part of (6.197), we have

$$\boldsymbol{E} = 2E_1\varepsilon_e \sin\omega t \sin kz. \tag{6.201}$$

In (6.201), variables $z$ and $t$ have been separated as well. For this case to be realized, the characteristic impedance of the dielectric of the incident wave side should be larger enough than that of the other side. In other words, the dielectric constant of the incident side should be small enough. Practically, this situation can easily be attained choosing a metal of high reflectance for the wall material. We have nodes at positions that satisfy

$$-kz = m\pi \ (m = 0, 1, 2, \cdots) \quad \text{or} \quad -z = \frac{m}{2}\lambda. \tag{6.202}$$

The nodes are formed at the interface and every half wavelength from it. As in the case of the syn-phase, the antinodes take place with the positions shifted by a quarter wavelength relative to the nodes.

If there is another interface at say $-z = L\,(>0)$, the wave goes back and forth many times. If an absolute value of the reflection coefficient at the interface is high enough (i.e., close to the unity), attenuation of the wave is ignorable. For both the syn-phase and anti-phase cases, we must have

$$kL = m\pi \ (m = 1, 2, \cdots) \quad \text{or} \quad L = \frac{m}{2}\lambda \tag{6.203}$$

so that stationary waves can stand stable. For a practical purpose, an optical device having such an interface is said to be a *resonator*. Various geometries and constitutions of the resonator are proposed in combination with various dielectrics including semiconductors. Related discussion can be seen in Chap. 7.

## References

1. Arfken GB, Weber HJ, Harris FE (2013) Mathematical methods for physicists, 7th edn. Academic Press, Waltham
2. Rudin W (1987) Real and complex analysis, 3rd edn. McGraw-Hill, New York
3. Smith FG, King TA, Wilkins D (2007) Optics and photonics, 2nd edn. Wiley, Chichester
4. Born M, Wolf E (2005) Principles of optics, 7th edn. Cambridge University Press, Cambridge
5. Pain HJ (2005) The physics of vibrations and waves, 6th edn. Wiley, Chichester

# Chapter 7
# Light Quanta: Radiation and Absorption

So far we discussed propagation of light and its reflection and transmission (or refraction) at an interface of dielectric media. We described characteristics of light from the point of view of an electromagnetic wave. In this chapter, we describe properties of light in relation to quantum mechanics. To this end, we start with Planck's law of radiation that successfully reproduced experimental results related to a blackbody radiation. Before this law had been established, Rayleigh–Jeans law failed to explain the experimental results at a high frequency region of radiation (the ultraviolet catastrophe). The Planck's law of radiation led to the discovery of light quanta. Einstein interpreted Planck's law of radiation on the basis of a model of two-level atoms. This model includes so-called Einstein $A$ and $B$ coefficients that are important in optics applications, especially lasers. We derive these coefficients from a classical point of view based on a dipole oscillation. We also consider a close relationship between electromagnetic waves confined in a cavity and a motion of a harmonic oscillator.

## 7.1 Blackbody Radiation

Historically, the relevant theory was first propounded by Max Planck and then Albert Einstein as briefly discussed in Chap. 1. The theory was developed on the basis of the experiments called cavity radiation or blackbody radiation. Here, however, we wish to derive Planck's law of radiation on the assumption of the existence of quantum harmonic oscillators.

As discussed in Chap. 2, the ground state of a quantum harmonic oscillator has an energy $\frac{1}{2}\hbar\omega$. Therefore, we measure energies of the oscillator in reference to that state. Let $N_0$ be the number of oscillators (i.e., light quanta) present in the ground state. Then, according to Boltzmann distribution law, the number of oscillators of the first excited state $N_1$ is



S. Hotta, *Mathematical Physical Chemistry*,
https://doi.org/10.1007/978-981-10-7671-8_7

$$N_1 = N_0 \mathrm{e}^{-\hbar\omega/k_\mathrm{B}T}, \tag{7.1}$$

where $k_\mathrm{B}$ is Boltzmann constant and $T$ is absolute temperature. Let $N_j$ be the number of oscillators of the $j$th excited state. Then we have

$$N_j = N_0 \mathrm{e}^{-j\hbar\omega/k_\mathrm{B}T}. \tag{7.2}$$

Let $N$ be the total number of oscillators in the system. Then, we get

$$\begin{aligned} N &= N_0 + N_0 \mathrm{e}^{-\hbar\omega/k_\mathrm{B}T} + \cdots + N_0 \mathrm{e}^{-j\hbar\omega/k_\mathrm{B}T} + \cdots \\ &= N_0 \sum_{j=0}^{\infty} \mathrm{e}^{-j\hbar\omega/k_\mathrm{B}T}. \end{aligned} \tag{7.3}$$

Let $E$ be a total energy of the oscillator system in reference to the ground state. That is, we put a ground state energy at zero. Then we have

$$\begin{aligned} E &= 0 \cdot N_0 + N_0 \hbar\omega \mathrm{e}^{-\hbar\omega/k_\mathrm{B}T} + \cdots + N_0 j\hbar\omega \mathrm{e}^{-j\hbar\omega/k_\mathrm{B}T} + \cdots \\ &= N_0 \sum_{j=0}^{\infty} j\hbar\omega \mathrm{e}^{-j\hbar\omega/k_\mathrm{B}T}. \end{aligned} \tag{7.4}$$

Therefore, an average energy of oscillators $\overline{E}$ is

$$\overline{E} = \frac{E}{N} = \hbar\omega \frac{\sum_{j=0}^{\infty} j\mathrm{e}^{-j\hbar\omega/k_\mathrm{B}T}}{\sum_{j=0}^{\infty} \mathrm{e}^{-j\hbar\omega/k_\mathrm{B}T}}. \tag{7.5}$$

Putting $x \equiv \mathrm{e}^{-\hbar\omega/k_\mathrm{B}T}$ [1], we have

$$\overline{E} = \hbar\omega \frac{\sum_{j=0}^{\infty} jx^j}{\sum_{j=0}^{\infty} x^j}. \tag{7.6}$$

Since $x < 1$, we have

$$\sum_{j=0}^{\infty} jx^j = \sum_{j=0}^{\infty} jx^{j-1} \cdot x = \left[\frac{\mathrm{d}}{\mathrm{d}x}\left(\sum_{j=0}^{\infty} x^j\right)\right] x = \left[\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{1-x}\right)\right] x = \frac{x}{(1-x)^2}. \tag{7.7}$$

$$\sum_{j=0}^{\infty} x^j = \frac{1}{1-x}. \tag{7.8}$$

Therefore, we get

$$\overline{E} = \frac{\hbar\omega x}{1-x} = \frac{\hbar\omega \mathrm{e}^{-\hbar\omega/k_\mathrm{B}T}}{1-\mathrm{e}^{-\hbar\omega/k_\mathrm{B}T}} = \frac{\hbar\omega}{\mathrm{e}^{\hbar\omega/k_\mathrm{B}T}-1}. \tag{7.9}$$

The function

$$\frac{1}{\mathrm{e}^{\hbar\omega/k_\mathrm{B}T}-1}. \tag{7.10}$$

is a form of Bose–Einstein distribution functions; more specifically, it is called the Bose–Einstein distribution function for photons today.

If $(\hbar\omega/k_\mathrm{B}T) \ll 1$, $\mathrm{e}^{\hbar\omega/k_\mathrm{B}T} \approx 1 + (\hbar\omega/k_\mathrm{B}T)$. Therefore, we have

$$\overline{E} \approx k_\mathrm{B}T. \tag{7.11}$$

Thus, the relation (7.9) asymptotically agrees with a classical theory. In other words, according to the classical theory related to law of equipartition of energy, energy of $k_\mathrm{B}T/2$ is distributed to each of two degrees of freedom of motion, i.e., a kinetic energy and a potential energy of a harmonic oscillator.

## 7.2 Planck's Law of Radiation and Mode Density of Electromagnetic Waves

Researcher at the time tried to seek the relationship between the energy density inside the cavity and (angular) frequency of radiation. To reach the relationship, let us introduce a concept of mode density of electromagnetic waves related to the blackbody radiation. We define the mode density $D(\omega)$ as the number of *modes* of electromagnetic waves per unit volume per unit angular frequency. We refer to the electromagnetic waves having allowed specific angular frequencies and polarization as modes. These modes must be described as *linearly independent* functions.

Determination of the mode density is related to boundary conditions (BCs) imposed on a physical system. We already dealt with this problem in Chaps. 2, 3, and 6. These BCs often appear when we find solutions of differential equations. Let us consider a following wave equation:

$$\frac{\partial^2\psi}{\partial x^2} = \frac{1}{v^2}\frac{\partial^2\psi}{\partial t^2}. \tag{7.12}$$

According to the method of separation of variables, we put

$$\psi(x,t) = X(x)T(t). \tag{7.13}$$

Substituting (7.13) for (7.12) and dividing both sides by $X(x)T(t)$, we have

$$\frac{1}{X}\frac{\mathrm{d}^2X}{\mathrm{d}x^2} = \frac{1}{v^2}\frac{1}{T}\frac{\mathrm{d}^2T}{\mathrm{d}t^2} = -k^2, \tag{7.14}$$

where $k$ is an undetermined (possibly complex) constant. For the $x$ component, we get

$$\frac{\mathrm{d}^2X}{\mathrm{d}x^2} + k^2X = 0. \tag{7.15}$$

Remember that $k$ is supposed to be a complex number for the moment (see Example 1.1). Modifying Example 1.1 a little bit such that (7.15) is posed in a domain $[0, L]$ and imposing the Dirichlet conditions such that

$$X(0) = X(L) = 0, \tag{7.16}$$

we find a solution of

$$X(x) = a \sin kx, \tag{7.17}$$

where $a$ is a constant. The constant $k$ can be determined to satisfy the BCs; i.e.,

$$kL = m\pi \text{ or } k = m\pi/L \ (m = 1, 2, \cdots). \tag{7.18}$$

Thus, we get real numbers for $k$. Then, we have a solution

$$T(t) = b \sin kvt = b \sin \omega t. \tag{7.19}$$

The overall solution is then

$$\psi(x, t) = c \sin kx \sin \omega t. \tag{7.20}$$

This solution has already appeared in Chap. 6 as a stationary solution. The readers are encouraged to derive these results.

In a three-dimensional case, we have a wave equation

$$\frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2} = \frac{1}{v^2}\frac{\partial^2\psi}{\partial t^2}. \tag{7.21}$$

In this case, we also assume

$$\psi(x, t) = X(x)Y(y)Z(z)T(t). \tag{7.22}$$

Similarly, we get

$$\frac{1}{X}\frac{\mathrm{d}^2X}{\mathrm{d}x^2}+\frac{1}{Y}\frac{\mathrm{d}^2Y}{\mathrm{d}x^2}+\frac{1}{Z}\frac{\mathrm{d}^2Z}{\mathrm{d}x^2}=\frac{1}{v^2}\frac{1}{T}\frac{\mathrm{d}^2T}{\mathrm{d}t^2}=-k^2. \tag{7.23}$$

Putting

$$\frac{1}{X}\frac{\mathrm{d}^2X}{\mathrm{d}x^2}=-k_x^2, \frac{1}{Y}\frac{\mathrm{d}^2Y}{\mathrm{d}x^2}=-k_y^2, \frac{1}{Z}\frac{\mathrm{d}^2Z}{\mathrm{d}x^2}=-k_z^2, \tag{7.24}$$

we have

$$k_x^2+k_y^2+k_z^2=k^2. \tag{7.25}$$

Then, we get a stationary wave solution as in the one-dimensional case such that

$$\psi(x,t)=c\,\sin k_x x\,\sin k_y y\,\sin k_z z\,\sin \omega t. \tag{7.26}$$

The BCs to be satisfied with $X(x), Y(y),$ and $Z(z)$ are

$$k_xL=m_x\pi,\ k_yL=m_y\pi,\ k_zL=m_z\pi\ \left(m_x,m_y,m_z=1,2,\cdots\right). \tag{7.27}$$

Returning to the main issue, let us deal with the mode density. Think of a cube of each side of $L$ that is placed as shown in Fig. 7.1. Calculating $k$, we have

$$k=\sqrt{k_x^2+k_y^2+k_z^2}=\frac{\pi}{L}\sqrt{m_x^2+m_y^2+m_z^2}=\frac{\omega}{c}, \tag{7.28}$$

where we assumed that the inside of a cavity is vacuum and, hence, the propagation velocity of light is $c$. Rewriting (7.28), we have

$$\sqrt{m_x^2+m_y^2+m_z^2}=\frac{L\omega}{\pi c}. \tag{7.29}$$

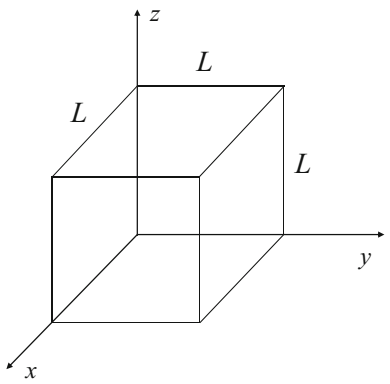


**Fig. 7.1** Cube of each side of $L$. We use this simple model to estimate mode density

The number $m_x, m_y$, and $m_z$ represents allowable modes in the cavity; the set of $(m_x, m_y, m_z)$ specifies individual modes. Note that $m_x, m_y$, and $m_z$ are all positive integers. If for instance $-m_x$ were allowed, this would produce $\sin(-k_x x) = -\sin k_x x$; but this function is linearly dependent on $\sin k_x x$. Then, a mode indexed by $-m_x$ should not be regarded as an independent mode. Given a $\omega$, a set $(m_x, m_y, m_z)$ that satisfies (7.29) corresponds to each mode. Therefore, the number of modes that satisfies a following expression

$$\sqrt{m_x^2 + m_y^2 + m_z^2} \leq \frac{L\omega}{\pi c} \tag{7.30}$$

represents those corresponding to angular frequencies equal to or less than given $\omega$.

Each mode has one-to-one correspondence with the lattice indexed by $(m_x, m_y, m_z)$. Accordingly, if $m_x, m_y, m_z \gg 1$, the number of allowed modes approximately equals one-eighth of a volume of a sphere having a radius of $\frac{L\omega}{\pi c}$. Let $N_L$ be the number of modes whose angular frequencies equal to or less than $\omega$. Recalling that there are two independent modes having the same index $(m_x, m_y, m_z)$ but mutually orthogonal polarities, we have

$$N_L = \frac{4\pi}{3}\left(\frac{L\omega}{\pi c}\right)^3 \cdot \frac{1}{8} \cdot 2 = \frac{L^3\omega^3}{3\pi^2 c^3}. \tag{7.31}$$

Consequently, the mode density $D(\omega)$ is expressed as

$$D(\omega)\,\mathrm{d}\omega = \frac{1}{L^3}\frac{\mathrm{d}N_L}{\mathrm{d}\omega}\mathrm{d}\omega = \frac{\omega^2}{\pi^2 c^3}\mathrm{d}\omega, \tag{7.32}$$

where $D(\omega)\,\mathrm{d}\omega$ represents the number of modes per unit volume whose angular frequencies range $\omega$ and $\omega + \mathrm{d}\omega$.

Now, we introduce a function $\rho(\omega)$ as an energy density per unit angular frequency. Then, combining $D(\omega)$ with (7.9), we get

$$\rho(\omega) = D(\omega)\frac{\hbar\omega}{\mathrm{e}^{\hbar\omega/k_\mathrm{B}T} - 1} = \frac{\hbar\omega^3}{\pi^2 c^3}\frac{1}{\mathrm{e}^{\hbar\omega/k_\mathrm{B}T} - 1}. \tag{7.33}$$

The relation (7.33) is called Planck's law of radiation. Notice that $\rho(\omega)$ has a dimension $[\mathrm{Jm}^{-3}\mathrm{s}]$.

To solve (7.15) under the Dirichlet conditions (7.16) is pertinent to analyzing an electric field within a cavity surrounded by a metal husk, because the electric field must be absent at an interface between the cavity and metal. The problem, however, can equivalently be solved using the magnetic field. This is because at the interface the reflection coefficient of electric field and magnetic field have a reversed sign (see Chap. 6). Thus, given an equation for the magnetic field, we may use the Neumann condition. This condition requires differential coefficients to vanish at the boundary (i.e., the interface between the cavity and metal). In a similar manner to the above, we get

$$\psi(x,t) = c\,\cos kx\,\cos \omega t. \tag{7.34}$$

By imposing BCs, again we have (7.18) that leads to the same result as the above.

We may also impose the periodic BCs. This type of equation has already been treated in Chap. 3. In that case, we have a solution of

$$\mathrm{e}^{ikx} \text{ and } \mathrm{e}^{-ikx}.$$

The BCs demand that $\mathrm{e}^0 = 1 = \mathrm{e}^{ikL}$. That is,

$$kL = 2\pi m \quad (m = 0, \pm 1, \pm 2, \cdots). \tag{7.35}$$

Notice that $\mathrm{e}^{ikx}$ and $\mathrm{e}^{-ikx}$ are linearly independent and, hence, minus sign for $m$ is permitted. Correspondingly, we have

$$N_L = \frac{4\pi}{3}\left(\frac{L\omega}{2\pi c}\right)^3 \cdot 2 = \frac{L^3\omega^3}{3\pi^2 c^3}. \tag{7.36}$$

In other words, here we have to consider a whole volume of a sphere of a *half* radius of the previous case. Thus, we reach the same conclusion as before.

If the average energy of an oscillator were described by (7.11), we would obtain a following description of $\rho(\omega)$ such that

$$\rho(\omega) = D(\omega) k_\mathrm{B} T = \frac{\omega^2}{\pi^2 c^3} k_\mathrm{B} T. \tag{7.37}$$

This relation is well-known as Rayleigh–Jeans law, but (7.37) disagreed with experimental results in that according to Rayleigh–Jeans law, $\rho(\omega)$ diverges toward infinity as $\omega$ goes to infinity. The discrepancy between the theory and experimental results was referred to as "ultraviolet catastrophe." Planck's law of radiation described by (7.33), on the other hand, reproduces the experimental results well.

## 7.3 Two-Level Atoms

Although Planck established Planck's law of radiation, researchers at that time hesitated in professing the existence of light quanta. It was Einstein that derived Planck's law by assuming two-level atoms in which light quanta play a role.

His assumption comprises the following three postulates: (i) The physical system to be addressed comprises so-called hypothetical "two-level" atoms that have only two energy levels. If two-level atoms absorb a light quantum, a ground-state electron is excited up to a higher-level (i.e., the stimulated absorption). (ii) The higher-level electron may spontaneously lose its energy and return back to the

ground state (the spontaneous emission). (iii) The higher-level electron may also lose its energy and return back to the ground state. Unlike (ii), however, the excited electron has to be stimulated by being irradiated by light quanta having an energy corresponding to the energy difference between the ground state and excited state (the stimulated emission). Figure 7.2 schematically depicts the optical processes of the Einstein model.

Under those postulates, Einstein dealt with the problem probabilistically. Suppose that the ground state and excited state have energies $E_1$ and $E_2$. Einstein assumed that light quanta having an energy equaling $E_2 - E_1$ take part in all the above three transition processes. He also propounded the idea that the light quanta have an energy that is proportional to its (angular) frequency. That is, he thought that the following relation should hold:

$$\hbar\omega_{21} = E_2 - E_1, \tag{7.38}$$

where $\omega_{21}$ is an angular frequency of light that takes part in the optical transitions. For the time being, let us follow Einstein's postulates.

(i) Stimulated absorption: This process is simply said to be an "absorption." Let $W_a\,[s^{-1}]$ be the transition probability that the electron absorbs a light quantum to be excited to the excited state. $W_a$ is described as

$$W_a = N_1 B_{21} \rho\,(\omega_{21}), \tag{7.39}$$

where $N_1$ is the number of atoms occupying the ground state; $B_{21}$ is a proportional constant; $\rho\,(\omega_{21})$ is due to (7.33). Note that in (7.39) we used $\omega_{21}$ instead of $\omega$ in (7.33). The coefficient $B_{21}$ is called Einstein $B$ coefficient; more specifically one of Einstein $B$ coefficients. Namely, $B_{21}$ is pertinent to the transition from the ground state to excited state.

(ii) Emission processes: The processes include both the spontaneous and stimulated emissions. Let $W_e\,[s^{-1}]$ be the transition probability that the electron emits a light quantum and returns back to the ground state. $W_e$ is described as

$$W_e = N_2 B_{12} \rho\,(\omega_{21}) + N_2 A_{12}, \tag{7.40}$$

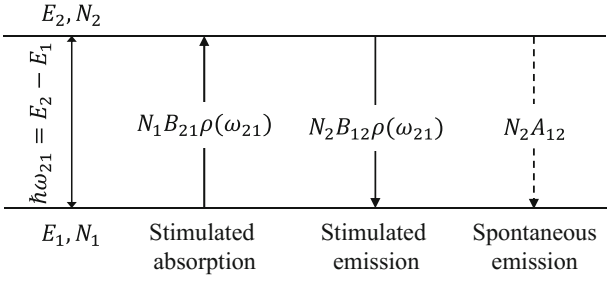


**Fig. 7.2** Optical processes of Einstein two-level atom model

where $N_2$ is the number of atoms occupying the excited state; $B_{12}$ and $A_{12}$ are proportional constants. The coefficient $A_{12}$ is called Einstein $A$ coefficient relevant to the spontaneous emission. The coefficient $B_{12}$ is associated with the stimulated emission and also called Einstein $B$ coefficient together with $B_{21}$. Here, $B_{12}$ is pertinent to the transition from the excited state to ground state.

Now, we have

$$B_{12} = B_{21}. \tag{7.41}$$

The reasoning for this is as follows: The coefficients $B_{12}$ and $B_{21}$ are proportional to the matrix elements pertinent to the optical transition. Let $T$ be an operator associated with the transition. Then, a matrix element is described using an inner product notation of Chap. 1 by

$$B_{21} = \langle \psi_2 | T | \psi_1 \rangle, \tag{7.42}$$

where $\psi_1$ and $\psi_2$ are initial and final states of the system in relation to the optical transition. As a good approximation, we use $e\boldsymbol{r}$ for $T$ (dipole approximation), where $e$ is an elementary charge and $\boldsymbol{r}$ is a position operator (see Chap. 1). If (7.42) represents the absorption process (i.e., the transition from the ground state to excited state), the corresponding emission process should be described as a reversed process by

$$B_{12} = \langle \psi_1 | T | \psi_2 \rangle. \tag{7.43}$$

Notice that in (7.43) $\psi_2$ and $\psi_1$ are initial and final states.
Taking complex conjugate of (7.42), we have

$$B_{21}^* = \langle \psi_1 | T^\dagger | \psi_2 \rangle, \tag{7.44}$$

where $T^\dagger$ is an operator adjoint to $T$ (see Chap. 1). With an Hermitian operator $H$, from Sect. 1.4 we have

$$H^\dagger = H. \tag{1.119}$$

Since $T$ is also Hermitian, we have

$$T^\dagger = T. \tag{7.45}$$

Thus, we get

$$B_{21}^* = B_{12}. \tag{7.46}$$

But, as in the cases of Sects. 4.2 and 4.3, $\psi_1$ and $\psi_2$ can be represented as real functions. Then, we have

$$B_{21}^{*} = B_{21} = B_{12}.$$

That is, we assume that the matrix $B$ is real symmetric. In the case of two-level atoms, as a matrix form we get

$$B = \begin{pmatrix} 0 & B_{12} \\ B_{12} & 0 \end{pmatrix}. \tag{7.47}$$

Compare (7.47) with (4.28).
Now, in the thermal equilibrium, we have

$$W_{\mathrm{e}} = W_{\mathrm{a}}. \tag{7.48}$$

That is,

$$N_2 B_{21}\rho(\omega_{21}) + N_2 A_{12} = N_1 B_{21}\rho(\omega_{21}), \tag{7.49}$$

where we used (7.41) for LHS. Assuming Boltzmann distribution law, we get

$$\frac{N_2}{N_1} = \exp[-(E_2 - E_1)/k_{\mathrm{B}}T]. \tag{7.50}$$

Here if moreover we assume (7.38), we get

$$\frac{N_2}{N_1} = \exp(-\hbar\omega_{21}/k_{\mathrm{B}}T). \tag{7.51}$$

Combing (7.49) and (7.51), we have

$$\exp(-\hbar\omega_{21}/k_{\mathrm{B}}T) = \frac{B_{21}\rho(\omega_{21})}{B_{21}\rho(\omega_{21}) + A_{12}}. \tag{7.52}$$

Solving (7.52) with respect to $\rho(\omega_{12})$, we finally get

$$\rho(\omega_{21}) = \frac{A_{12}}{B_{21}} \cdot \frac{\exp(-\hbar\omega_{21}/k_{\mathrm{B}}T)}{1 - \exp(-\hbar\omega_{21}/k_{\mathrm{B}}T)} = \frac{A_{12}}{B_{21}} \cdot \frac{1}{\exp(\hbar\omega_{21}/k_{\mathrm{B}}T) - 1}. \tag{7.53}$$

Assuming that

$$\frac{A_{12}}{B_{21}} = \frac{\hbar\omega_{21}^3}{\pi^2 c^3}, \tag{7.54}$$

we have

$$\rho(\omega_{21}) = \frac{\hbar\omega_{21}^3}{\pi^2 c^3} \cdot \frac{1}{\exp(\hbar\omega_{21}/k_{\mathrm{B}}T) - 1}. \tag{7.55}$$

This is none other than Planck's law of radiation.

## 7.4 Dipole Radiation

In (7.54), we only know the ratio of $A_{12}$ to $B_{21}$. To have a good knowledge of these Einstein coefficients, we briefly examine a mechanism of the dipole radiation. The electromagnetic radiation results from an accelerated motion of a dipole.

A dipole moment $\boldsymbol{p}(t)$ is defined as a function of time $t$ by

$$\boldsymbol{p}(t) = \int \boldsymbol{x}'\rho(\boldsymbol{x}', t)\mathrm{d}\boldsymbol{x}', \tag{7.56}$$

where $\boldsymbol{x}'$ is a position vector in a Cartesian coordinate; an integral is taken over a whole three-dimensional space; $\rho$ is a charge density appearing in (5.1). If we consider a system comprising point charges, integration can immediately be carried out to yield

$$\boldsymbol{p}(t) = \sum_i q_i \boldsymbol{x}_i, \tag{7.57}$$

where $q_i$ is a charge of each point charge $i$ and $\boldsymbol{x}_i$ is a position vector of the point charge $i$. From (7.56) and (7.57), we find that $\boldsymbol{p}(t)$ depends on how we set up the coordinate system. However, if a total charge of the system is zero, $\boldsymbol{p}\,(t)$ does not depend on the coordinate system. Let $\boldsymbol{p}(t)$ and $\boldsymbol{p}'(t)$ be a dipole moment viewed from the frame $O$ and $O'$, respectively (see Fig. 7.3). Then we have

$$\boldsymbol{p}'(t) = \sum_i q_i \boldsymbol{x}_i' = \sum_i q_i\,(\boldsymbol{x}_0 + \boldsymbol{x}_i) = \left(\sum_i q_i\right)\boldsymbol{x}_0 + \sum_i q_i \boldsymbol{x}_i = \sum_i q_i\,\boldsymbol{x}_i = \boldsymbol{p}(t). \tag{7.58}$$

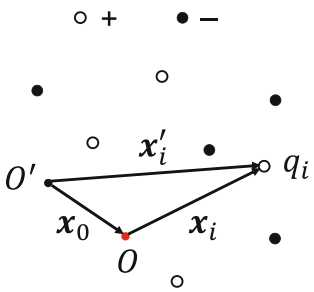


**Fig. 7.3** Dipole moment viewed from the frame $O$ or $O'$

Notice that with the third equality the first term vanishes because the total charge is zero. The system comprising two point charges that have an opposite charge $(\pm q)$ is particularly simple but very important. In that case, we have

$$\boldsymbol{p}(t) = q\boldsymbol{x}_1 + (-q)\boldsymbol{x}_2 = q(\boldsymbol{x}_1 - \boldsymbol{x}_2) = q\widetilde{\boldsymbol{x}}. \tag{7.59}$$

Here, we assume that $q > 0$ according to the custom and, hence, $\widetilde{\boldsymbol{x}}$ is a vector directing from the minus point charge to the plus charge.

Figure 7.4 displays geometry of an oscillating dipole and electromagnetic radiation from it. Figure 7.4a depicts the dipole. It is placed at the origin of the coordinate system and assumed to be of an atomic or molecular scale in extension; we regard a center of the dipole as the origin. Figure 7.4b represents a large-scale geometry of the dipole and surrounding space of it. For the electromagnetic radiation, an accelerated motion of the dipole is of primary importance. The electromagnetic fields produced by $\ddot{\boldsymbol{p}}(t)$ vary as the inverse of $r$, where $r$ is a macroscopic distance between the dipole and observation point. Namely, $r$ is much larger compared to the dipole size.

There are other electromagnetic fields that result from the dipole moment. The fields result from $\boldsymbol{p}(t)$ and $\dot{\boldsymbol{p}}(t)$. Strictly speaking, we have to include those quantities that are responsible for the electromagnetic fields associated with the dipole radiation. Nevertheless, the fields produced by $\boldsymbol{p}(t)$ and $\dot{\boldsymbol{p}}(t)$ vary as a function of the inverse cube and inverse square of $r$, respectively. Therefore, the surface integral of the square of the fields associated with $\boldsymbol{p}(t)$ and $\dot{\boldsymbol{p}}(t)$ asymptotically reaches zero with enough large $r$ with respect to a sphere enveloping the dipole. Regarding $\ddot{\boldsymbol{p}}(t)$, on the other hand, the surface integral of the square of the fields remains finite even with enough large $r$. For this reason, we refer to the spatial region where $\ddot{\boldsymbol{p}}(t)$ does not vanish as a *wave zone*.

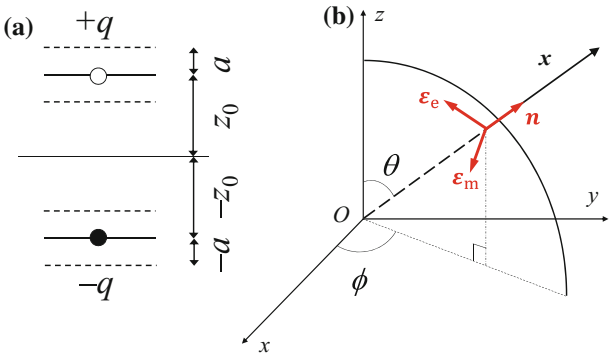


**Fig. 7.4** Electromagnetic radiation from an accelerated motion of a dipole. **a** A dipole placed at the origin of the coordinate system is executing harmonic oscillation along the $z$-direction around an equilibrium position. **b** Electromagnetic radiation from a dipole in a wave zone. $\boldsymbol{\varepsilon}_e$ and $\boldsymbol{\varepsilon}_m$ are unit polarization vectors of the electric field and magnetic field, respectively. $\boldsymbol{\varepsilon}_e$, $\boldsymbol{\varepsilon}_m$, and $\boldsymbol{n}$ form a right-handed system

Suppose that a dipole placed at the origin of the coordinate system is executing harmonic oscillation along the $z$-direction around an equilibrium position (see Fig. 7.4). Motion of two charges having plus and minus signs is described by

$$\boldsymbol{z}_+ = z_0\,\boldsymbol{e}_3 + a\mathrm{e}^{i\omega t}\boldsymbol{e}_3 \ \ (z_0, a > 0), \tag{7.60}$$

$$\boldsymbol{z}_- = -z_0\,\boldsymbol{e}_3 - a\mathrm{e}^{i\omega t}\boldsymbol{e}_3, \tag{7.61}$$

where $\boldsymbol{z}_+$ and $\boldsymbol{z}_-$ are position vectors of a plus charge and minus charge, respectively; $z_0$ and $-z_0$ are equilibrium positions of each charge; $a$ is an amplitude of the harmonic oscillation; $\omega$ is an angular frequency of the oscillation. Then, accelerations of the charges are given by

$$\boldsymbol{a}_+ \equiv \ddot{\boldsymbol{z}}_+ = -a\omega^2\mathrm{e}^{i\omega t}\boldsymbol{e}_3, \tag{7.62}$$

$$\boldsymbol{a}_- \equiv \ddot{\boldsymbol{z}}_- = a\omega^2\mathrm{e}^{i\omega t}\boldsymbol{e}_3. \tag{7.63}$$

Meanwhile, we have

$$\boldsymbol{p}(t) = q\boldsymbol{z}_+ + (-q)\boldsymbol{z}_- = q(\boldsymbol{z}_+ - \boldsymbol{z}_-) \quad (q > 0). \tag{7.64}$$

Therefore,

$$\ddot{\boldsymbol{p}}(t) = q(\ddot{\boldsymbol{z}}_+ - \ddot{\boldsymbol{z}}_-) = -2qa\omega^2\mathrm{e}^{i\omega t}\boldsymbol{e}_3. \tag{7.65}$$

The quantity $\ddot{\boldsymbol{p}}(t)$, i.e., the second derivative of $\boldsymbol{p}(t)$ with respect to time produces the electric field described by [2]

$$\boldsymbol{E} = -\ddot{\boldsymbol{p}}/4\pi\varepsilon_0 c^2 r = -qa\omega^2\mathrm{e}^{i\omega t}\boldsymbol{e}_3/2\pi\varepsilon_0 c^2 r, \tag{7.66}$$

where $r$ is a distance between the dipole and observation point. In (7.66), we ignored a term proportional to inverse square and cube of $r$ for the aforementioned reason.

As described in (7.66), the strength of the radiation electric field in the wave zone measured at a point away from the oscillating dipole is proportional to a component of the vector of the acceleration motion of the dipole [i.e., $\ddot{\boldsymbol{p}}(t)$]. The radiation electric field lies in the direction perpendicular to a line connecting the observation point and the point of the dipole (Fig. 7.4). Let $\boldsymbol{\varepsilon}_\mathrm{e}$ be a unit polarization vector of the electric field in that direction, and let $\boldsymbol{E}^\perp$ be the radiation electric field. Then, we have

$$\boldsymbol{E}^\perp = -qa\omega^2\mathrm{e}^{i\omega t}(\boldsymbol{e}_3 \cdot \boldsymbol{\varepsilon}_\mathrm{e})\boldsymbol{\varepsilon}_\mathrm{e}/2\pi\varepsilon_0 c^2 r = -qa\omega^2\mathrm{e}^{i\omega t}\boldsymbol{\varepsilon}_\mathrm{e}\sin\theta/2\pi\varepsilon_0 c^2 r\,. \tag{7.67}$$

As shown in Sect. 5.3, $(\boldsymbol{\varepsilon}_e \cdot \boldsymbol{e}_3)\,\boldsymbol{\varepsilon}_e$ in (7.67) "extracts" from $\boldsymbol{e}_3$ a vector component parallel to $\boldsymbol{\varepsilon}_e$. Such an operation is said to be a projection of a vector. The related discussion can be seen in Part III.

It takes a time of $r/c$ for the emitted light from the charge to arrive at the observation point. Consequently, the acceleration of the charge has to be measured at the time when the radiation leaves the charge. Let $t$ be the instant when the electric field is measured at the measuring point. Then, it follows that the radiation leaves the charge at a time of $t - r/c$. Thus, the electric field relevant to the radiation that can be observed far enough away from the oscillating charge is described as [2]

$$\boldsymbol{E}^{\perp}(\boldsymbol{x}, t) = -\frac{qa\omega^2 \mathrm{e}^{i\omega\left(t-\frac{r}{c}\right)} \sin\theta}{2\pi\varepsilon_0 c^2 r}\,\boldsymbol{\varepsilon}_e. \tag{7.68}$$

The radiation electric field must necessarily be accompanied by a magnetic field. Writing the radiation magnetic field as $\boldsymbol{H}^{\perp}(\boldsymbol{x}, t)$, we have [3]

$$\begin{aligned}\boldsymbol{H}^{\perp}(\boldsymbol{x}, t) &= -\frac{1}{c\mu_0} \cdot \frac{qa\omega^2 \mathrm{e}^{i\omega\left(t-\frac{r}{c}\right)} \sin\theta}{2\pi\varepsilon_0 c^2 r}\boldsymbol{n} \times \boldsymbol{\varepsilon}_e = -\frac{qa\omega^2 \mathrm{e}^{i\omega\left(t-\frac{r}{c}\right)} \sin\theta}{2\pi c r}\boldsymbol{n} \times \boldsymbol{\varepsilon}_e \\ &= -\frac{qa\omega^2 \mathrm{e}^{i\omega\left(t-\frac{r}{c}\right)} \sin\theta}{2\pi c r}\boldsymbol{\varepsilon}_m,\end{aligned} \tag{7.69}$$

where $\boldsymbol{n}$ represents a unit vector in the direction parallel to a line connecting the observation point and the dipole. The $\boldsymbol{\varepsilon}_m$ is a unit polarization vector of the magnetic field as defined by (5.67). From the above, we see that the radiation electromagnetic waves in the wave zone are transverse waves that show the properties the same as those of electromagnetic waves in a free space.

Now, let us evaluate a time-averaged energy flux from an oscillating dipole. Using (6.71), we have

$$\begin{aligned}\overline{\boldsymbol{S}}(\theta) &= \frac{1}{2}\boldsymbol{E} \times \boldsymbol{H}^* = \frac{1}{2} \cdot \frac{qa\omega^2 \mathrm{e}^{i\omega\left(t-\frac{r}{c}\right)} \sin\theta}{2\pi\varepsilon_0 c^2 r} \cdot \frac{qa\omega^2 \mathrm{e}^{-i\omega\left(t-\frac{r}{c}\right)} \sin\theta}{2\pi c r}\boldsymbol{n} \\ &= \frac{\omega^4 \sin^2\theta}{8\pi^2\varepsilon_0 c^3 r^2}(qa)^2\boldsymbol{n}.\end{aligned} \tag{7.70}$$

If we are thinking of an atom or a molecule in which the dipole consists of an electron and a positive charge that compensates it, $q$ is replaced with $-e$ $(e<0)$. Then (7.70) reads as

$$\overline{\boldsymbol{S}}(\theta) = \frac{\omega^4 \sin^2\theta}{8\pi^2\varepsilon_0 c^3 r^2}(ea)^2\boldsymbol{n}. \tag{7.71}$$

Let us relate the above argument to Einstein $A$ and $B$ coefficients. Since we are dealing with an isolated dipole, we might well suppose that the radiation comes from the spontaneous emission. Let $P$ be a total power of emission from the oscillating dipole that gets through a sphere of radius $r$. Then we have

$$\begin{aligned} P &= \int \overline{\boldsymbol{S}}(\theta) \cdot \boldsymbol{n} \, \mathrm{d}S = \int_0^{2\pi} \mathrm{d}\phi \int_0^{\pi} \overline{\boldsymbol{S}}(\theta) r^2 \sin\theta \, \mathrm{d}\theta \\ &= \frac{\omega^4}{8\pi^2 \varepsilon_0 c^3} (ea)^2 \int_0^{2\pi} \mathrm{d}\phi \int_0^{\pi} \sin^3\theta \, \mathrm{d}\theta. \end{aligned} \tag{7.72}$$

Changing $\cos\theta$ to $t$, the integral $I \equiv \int_0^{\pi} \sin^3\theta \, \mathrm{d}\theta$ can be converted into

$$I = \int_{-1}^{1} (1 - t^2) \, \mathrm{d}t = 4/3. \tag{7.73}$$

Thus, we have

$$P = \frac{\omega^4}{3\pi\varepsilon_0 c^3} (ea)^2. \tag{7.74}$$

A probability of the spontaneous emission is given by $N_2 A_{12}$. Since we are dealing with a single dipole, $N_2$ can be put 1. Accordingly, an expected power of emission is $A_{12} \hbar\omega_{21}$. Replacing $\omega$ in (7.74) with $\omega_{21}$ in (7.55) and equating $A_{12} \hbar\omega_{21}$ to $P$, we get

$$A_{12} = \frac{\omega_{21}^3}{3\pi\varepsilon_0 c^3 \hbar} (ea)^2. \tag{7.75}$$

From (7.54), we also get

$$B_{12} = \frac{\pi}{3\varepsilon_0 \hbar^2} (ea)^2. \tag{7.76}$$

In order to relate these results to those of quantum mechanics, we may replace $a^2$ in the above expressions with matrix elements of the position operator $\boldsymbol{r}$. That is, representing $|1\rangle$ and $|2\rangle$ as the quantum states of the ground and excited states of a two-level atom, we define $\langle 1|\boldsymbol{r}|2\rangle$ as the matrix element of the position operator. Relating $|\langle 1|\boldsymbol{r}|2\rangle|^2$ to $a^2$, we get

$$A_{12} = \frac{\omega_{21}^3 e^2}{3\pi\varepsilon_0 c^3 \hbar} |\langle 1|\boldsymbol{r}|2\rangle|^2 \quad \text{and} \quad B_{12} = \frac{\pi e^2}{3\varepsilon_0 \hbar^2} |\langle 1|\boldsymbol{r}|2\rangle|^2. \tag{7.77}$$

From (7.77), we have

$$B_{12} = \frac{\pi e^2}{3\varepsilon_0 \hbar^2} \langle 1|\boldsymbol{r}|2\rangle \langle 1|\boldsymbol{r}|2\rangle^* = \frac{\pi e^2}{3\varepsilon_0 \hbar^2} \langle 1|\boldsymbol{r}|2\rangle \langle 2|\boldsymbol{r}^\dagger|1\rangle = \frac{\pi e^2}{3\varepsilon_0 \hbar^2} \langle 1|\boldsymbol{r}|2\rangle \langle 2|\boldsymbol{r}|1\rangle, \tag{7.78}$$

where with the second equality we used (1.116); the last equality comes from that $\boldsymbol{r}$ is Hermitian. Meanwhile, we have

$$B_{21} = \frac{\pi e^2}{3\varepsilon_0 \hbar^2} |\langle 2|\boldsymbol{r}|1\rangle|^2 = \frac{\pi e^2}{3\varepsilon_0 \hbar^2} \langle 2|\boldsymbol{r}|1\rangle \left\langle 1\left|\boldsymbol{r}^\dagger\right|2\right\rangle = \frac{\pi e^2}{3\varepsilon_0 \hbar^2} \langle 2|\boldsymbol{r}|1\rangle \langle 1|\boldsymbol{r}|2\rangle. \tag{7.79}$$

Hence, we recover (7.41).

## 7.5 Lasers

A concept of the two-level atoms proposed by Einstein is universal and independent of materials and can be utilized for some particular purposes. Actually, in later years many researchers tried to validate that concept and verified its validity. After basic researches of 1950s and 1960s, fundamentals were put into practical use as various optical devices. Typical examples are masers and lasers, abbreviations for "microwave amplification by stimulated emission of radiation" and "light amplification by stimulated emission of radiation." Of these, lasers are common and particularly important nowadays. On the basis of universality of the concept, a lot of materials including semiconductors and dyes are used in gaseous, liquid, and solid states.

Let us consider a rectangular parallelepiped of a laser medium with a length $L$ and cross-sectional area $S$ (Fig. 7.5). Suppose there are $N$ two-level atoms in the rectangular parallelepiped such that

$$N = N_1 + N_2, \tag{7.80}$$

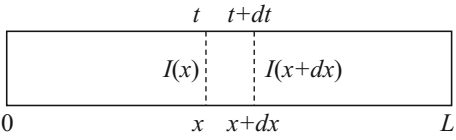


**Fig. 7.5** Rectangular parallelepiped of a laser medium with a length $L$ and a cross-sectional area $S$ (not shown). $I(x)$ denotes irradiance at a point of $x$ from the origin

where $N_1$ and $N_2$ represent the number of atoms occupying the ground and excited states, respectively. Suppose that a light is propagated from the left of the rectangular parallelepiped and entering it. Then, we expect that three processes occur simultaneously. One is a stimulated absorption and others are stimulated emission and spontaneous emission. After these process, an increment $\mathrm{d}E$ in photon energy of the total system (i.e., the rectangular parallelepiped) during $\mathrm{d}t$ is described by

$$\mathrm{d}E = \{N_2 \left[B_{21}\rho\left(\omega_{21}\right) + A_{12}\right] - N_1 B_{21}\rho\left(\omega_{21}\right)\}\hbar\omega_{21}\,\mathrm{d}t. \tag{7.81}$$

In light of (7.39) and (7.40), a dimensionless quantity $\mathrm{d}E/\hbar\omega_{21}$ represents a number of effective events of photons emission that have occurred during $\mathrm{d}t$. Since in lasers the stimulated emission is dominant, we shall forget about the spontaneous emission and rewrite (7.81) as

$$\begin{aligned}\mathrm{d}E &= \{N_2 B_{21}\rho\left(\omega_{21}\right) - N_1 B_{21}\rho\left(\omega_{21}\right)\}\hbar\omega_{21}\,\mathrm{d}t \\ &= B_{21}\rho\left(\omega_{21}\right)\left(N_2 - N_1\right)\hbar\omega_{21}\,\mathrm{d}t.\end{aligned} \tag{7.82}$$

Under a thermal equilibrium, we have $N_2 < N_1$ on the basis of Boltzmann distribution law, and so $\mathrm{d}E < 0$. In this occasion, therefore, the photon energy decreases. For the light amplification to take place, therefore, we must have a following condition:

$$N_2 > N_1. \tag{7.83}$$

This energy distribution is called inverted distribution or population inversion. Thus, the laser oscillation is a typical non-equilibrium phenomenon. To produce the population inversion, we need an external exciting source using an electrical or optical device.

The essence of lasers rests upon the fact that stimulated emission produces a photon that possesses a wave vector ($\boldsymbol{k}$) and a polarization ($\boldsymbol{\varepsilon}$) both the same as those of an original photon. For this reason, the laser light is monochromatic and highly directional. To understand the fundamental mechanism underlying the related phenomena, interested readers are encouraged to seek appropriate literature of quantum theory of light for further reading [4].

To make a discussion simple and straightforward, let us assume that the light is incident parallel to the long axis of the rectangular parallelepiped. Then, the stimulated emission produces light to be propagated in the same direction. As a result, an irradiance $I$ measured in that direction is described as

$$I = \frac{E}{SL}\cdot c'. \tag{7.84}$$

Note that the light velocity in the laser medium $c'$ is given by

$$c' = c/n, \tag{7.85}$$

where $n$ is a refractive index of the laser medium. Taking an infinitesimal of both sides of (7.84), we have

$$\begin{aligned} \mathrm{d}I &= \mathrm{d}E\frac{c'}{SL} = B_{21}\rho\left(\omega_{21}\right)\left(N_2 - N_1\right)\hbar\omega_{21}\frac{c'}{SL}\,\mathrm{d}t \\ &= B_{21}\rho\left(\omega_{21}\right)\tilde{N}\hbar\omega_{21}c'\,\mathrm{d}t\,, \end{aligned} \tag{7.86}$$

where $\tilde{N} = (N_2 - N_1)/SL$ denotes a "net" density of atoms that occupy the excited state.

The energy density $\rho\left(\omega_{21}\right)$ can be written as

$$\rho\left(\omega_{21}\right) = I\left(\omega_{21}\right)g\left(\omega_{21}\right)/c', \tag{7.87}$$

where $I\left(\omega_{12}\right)\left[\mathrm{Js}^{-1}\mathrm{m}^{-2}\right]$ represents an intensity of radiation; $g\left(\omega_{21}\right)$ is a gain function [s]. The gain function is a measure that shows how favorably (or unfavorably) the transition takes place at the said angular frequency $\omega_{12}$. This is normalized in the emission range such that

$$\int_0^\infty g(\omega)\,\mathrm{d}\omega = 1.$$

The quantity $I(\omega_{21})$ is an energy flux that gets through per unit area per unit time. This flux corresponds to an energy contained in a long and thin rectangular parallelepiped of a length $c'$ and a unit cross-sectional area. To obtain $\rho(\omega_{21})$, $I(\omega_{12})$ should be divided by $c'$ in (7.87) accordingly. Using (7.87) and replacing $c'\mathrm{d}t$ with a distance $\mathrm{d}x$ and $I(\omega_{12})$ with $I(x)$ as a function of $x$, we rewrite (7.86) as

$$\mathrm{d}I(x) = \frac{B_{21}g(\omega_{21})\widetilde{N}\hbar\omega_{21}}{c'}I(x)\,\mathrm{d}x. \tag{7.88}$$

Dividing (7.88) by $I(x)$ and integrating both sides, we have

$$\int_{I_0}^{I}\frac{\mathrm{d}I(x)}{I(x)} = \int_{I_0}^{I}\mathrm{d}\,\ln I(x) = \frac{B_{21}g\left(\omega_{21}\right)\widetilde{N}\hbar\omega_{21}}{c'}\int_0^x \mathrm{d}x, \tag{7.89}$$

where $I_0$ is an irradiance of light at an instant when the light is entering the laser medium from the left. Thus, we get

$$I(x) = I_0 \exp\left[\frac{B_{21} g(\omega_{21}) \widetilde{N} \hbar \omega_{21}}{c'} x\right]. \tag{7.90}$$

The Eq. (7.90) shows that an irradiance of the laser light is augmented exponentially along the path of the laser light. In (7.90), denoting an exponent as $G$

$$G \equiv \frac{B_{21} g(\omega_{21}) \widetilde{N} \hbar \omega_{21}}{c'}, \tag{7.91}$$

we get

$$I(x) = I_0 \exp Gx.$$

The constant $G$ is said to be a gain constant. This is an index that indicates the laser performance. Large numbers $B_{21}$, $g(\omega_{21})$, and $\widetilde{N}$ yield a high performance of the laser.

In Sects. 6.9 and 7.2, we sought conditions for electromagnetic waves to cause constructive interference. In a one-dimensional dielectric medium, the condition is described as

$$kL = m\pi \quad \text{or} \quad m\lambda = 2L \ (m = 1, 2, \cdots), \tag{7.92}$$

where $k$ and $\lambda$ denote a wave number and wavelength in the dielectric medium, respectively. Indexing $k$ and $\lambda$ with $m$ that represents a mode, we have

$$k_m L = m\pi \quad \text{or} \quad m\lambda_m = 2L \ (m = 1, 2, \cdots). \tag{7.93}$$

This condition can be expressed by different manners such that

$$\omega_m = 2\pi \nu_m = 2\pi c'/\lambda_m = 2\pi c/n\lambda_m = m\pi c/nL. \tag{7.94}$$

It is often the case that if the laser is a long and thin rod, rectangular parallelepiped, etc., we see that sharply resolved and regularly spaced spectral lines are observed in emission spectrum. These lines are referred to as longitudinal multimodes. The separation between two neighboring emission lines is referred to as the *free spectral range* [2]. If adjacent emission lines are clearly resolved so that the free spectral range can easily be recognized, we can derive useful information from the laser oscillation spectra (vide infra).

Rewriting (7.94) as, e.g.,

$$\omega_m n = \frac{\pi c}{L} m, \tag{7.95}$$

and taking differential (or variation) of both sides, we get

$$n\delta\omega_m + \omega_m\delta n = n\delta\omega_m + \omega_m \frac{\delta n}{\delta\omega_m}\delta\omega_m = \left(n + \omega_m \frac{\delta n}{\delta\omega_m}\right)\delta\omega_m = \frac{\pi c}{L}\delta m. \quad (7.96)$$

Therefore, we get

$$\delta\omega_m = \frac{\pi c}{L}\left(n + \omega_m \frac{\delta n}{\delta\omega_m}\right)^{-1}\delta m. \quad (7.97)$$

The Eq. (7.97) premises the wavelength dispersion of a refractive index of a laser medium. Here, the wavelength dispersion means that the refractive index varies as a function of wavelengths of light in a matter. The laser materials often have a considerably large dispersion, and relevant information is indispensable.

From (7.97), we find that

$$n_g \equiv n + \omega_m \frac{\delta n}{\delta\omega_m} \quad (7.98)$$

plays a role of refractive index when the laser material has a wavelength dispersion. The quantity $n_g$ is said to be a *group refractive index* (or group index). Thus, (7.97) is rewritten as

$$\delta\omega = \frac{\pi c}{L n_g}\delta m, \quad (7.99)$$

where we omitted the index $m$ of $\omega_m$. Rewriting (7.98) as a relation of continuous quantities and using differentiation instead of variation, we have [2]

$$n_g = n + \omega\frac{\mathrm{d}n}{\mathrm{d}\omega} \quad \text{or} \quad n_g = n - \lambda\frac{\mathrm{d}n}{\mathrm{d}\lambda}. \quad (7.100)$$

To derive the second equation of (7.100), we used following relations: Namely, taking a variation of $\lambda\omega = 2\pi\mathrm{c}$, we have

$$\omega\mathrm{d}\lambda + \lambda\mathrm{d}\omega = 0 \quad \text{or} \quad \frac{\omega}{\mathrm{d}\omega} = -\frac{\lambda}{\mathrm{d}\lambda}.$$

Several formulae or relation equations were proposed to describe the wavelength dispersion. One of famous and useful formula among them is Sellmeier's dispersion formula [5]. As an example, the Sellmeier's dispersion formula can be described as

$$n = \sqrt{A + \frac{B}{1 - \left(\frac{C}{\lambda}\right)^2}}, \quad (7.101)$$

where $A$, $B$, and $C$ are appropriate constants with $A$ and $B$ being dimensionless and $C$ having a dimension [m]. In an actual case, it would be difficult to determine $n$ analytically. However, if we are able to obtain well-resolved spectra, $\delta m$ can be put as 1 and $\delta\omega_m$ can be determined from the free spectral range. Expressing it as $\delta\omega_{\mathrm{FSR}}$ from (7.99) we have

$$\delta\omega_{\mathrm{FSR}} = \frac{\pi c}{L n_g} \quad \text{or} \quad n_g = \frac{\pi c}{L(\delta\omega_{\mathrm{FSR}})}. \tag{7.102}$$

Thus, one can determine $n_g$ as a function of wavelengths.

As an tangible example, Fig. 7.6 [6] displays a broadband emission spectra of a crystal consisting of an organic semiconductor AC'7. As another example, Fig. 7.7 [6] displays a laser oscillation spectrum of an AC'7 crystal. The structural formula of AC'7 is shown in Fig. 7.8 together with other related organic semiconductors. Once we choose an empirical formula of the wavelength dispersion [e.g., (7.101)], we can determine constants of that empirical formula by comparing it with experimentally decided data. For such data, laser oscillation spectra (Fig. 7.7) were used in addition to the broadband emission spectra. It is because the laser oscillation spectra are essentially identical to Fig. 7.6 in a sense that both the broadband and laser emission lines gave the same free spectral range. Inserting (7.101) into (7.100) and expressing $n_g$ as a function of $\lambda$, Yamao et al. got a following expression [6]:

$$n_g = \frac{A\left[1 - \left(\frac{C}{\lambda}\right)^2\right]^2 + B}{\sqrt{\left[1 - \left(\frac{C}{\lambda}\right)^2\right]^3}\sqrt{A\left[1 - \left(\frac{C}{\lambda}\right)^2\right] + B}}. \tag{7.103}$$

Determining optimum constants $A$, $B$, and $C$, a set of these constants yields a reliable dispersion formula in (7.101). Numerical calculations can be utilized effectively. The procedures are as follows: (i) Tentatively choosing probable

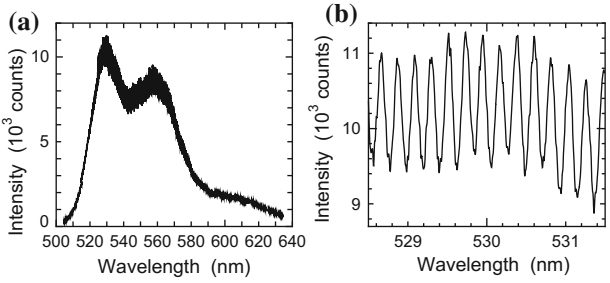


**Fig. 7.6** Broadband emission spectra of an organic semiconductor crystal AC'7. **a** Full spectrum. **b** Enlarged profile of the spectrum around 530 nm. Reproduced from Yamao et al. [6], with the permission of AIP Publishing. http://doi.org/10.1063/1.3634117

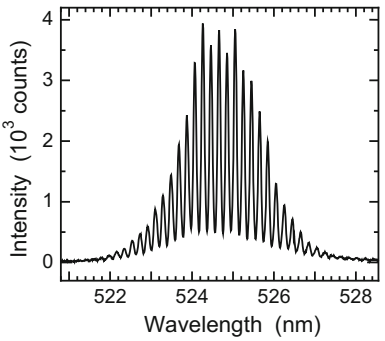


**Fig. 7.7** Laser oscillation spectrum of an organic semiconductor crystal AC'7. Reproduced from Yamao et al. [6], with the permission of AIP Publishing. http://doi.org/10.1063/1.3634117

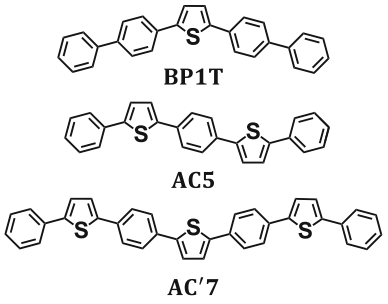


**Fig. 7.8** Structural formulae of several organic semiconductors BP1T, AC5, and AC'7

numbers for $A$, $B$, and $C$ for (7.103), $n_g$ can be expressed as a function of $\lambda$. (ii) The resulting fitting curve is then compared with $n_g$ data experimentally decided from (7.102). After this procedure, one can choose another set of $A$, $B$, and $C$ and again compare the fitting curve with the experimental data. (iii) This procedure can be repeated many times through iterative numerical computations of (7.103) using different sets of $A$, $B$, and $C$.

Thus, we should be able to adjust and determine better and better combination of $A$, $B$, and $C$ so that the refined function (7.103) can reproduce the experimental results as precise as one pleases. At the same time, we can determine the most reliable combination of $A$, $B$, and $C$ with the dispersion formula of (7.101). Figure 7.9 [6] shows several examples of the wavelength dispersion for organic semiconductor crystals. Optimized constants $A$, $B$, and $C$ of (7.101) are listed in Table 7.1 [6].

The formulae (7.101) and (7.103) along with associated procedures to determine the constants $A$, $B$, and $C$ are expected to apply to various laser and light-emitting materials consisting of semiconducting inorganic and organic materials.

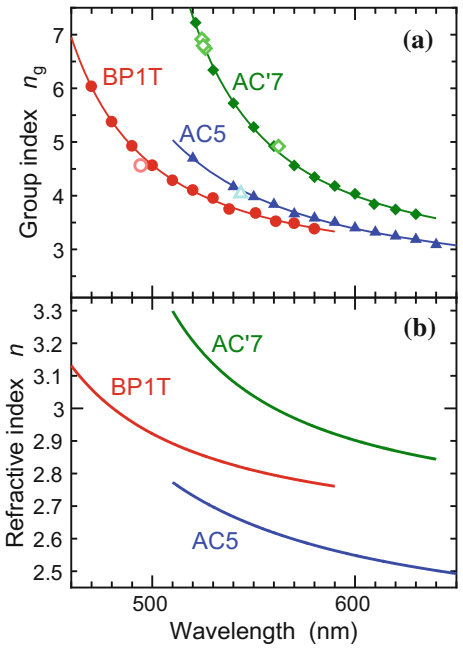


**Fig. 7.9** Examples of the wavelength dispersion of **a** group indices and **b** refractive indices for several organic semiconductor crystals. Reproduced from Yamao et al. [6], with the permission of AIP Publishing. http://doi.org/10.1063/1.3634117

**Table 7.1** Optimized constants of $A$, $B$, and $C$ for Sellmeier's dispersion formula (7.101) with several organic semiconductor crystals

| Material | $A$ | $B$ | $C$ (nm) |
|---|---|---|---|
| BP1T | 5.7 | 1.04 | 397 |
| AC5 | 3.9 | 1.44 | 402 |
| AC'7 | 6.0 | 1.06 | 452 |

Reproduced from Yamao et al. [6] with the permission of AIP Publishing. http://doi.org/10.1063/1.3634117

## 7.6 Mechanical System

As outlined above, two-level atoms have distinct characteristics in connection with lasers. Electromagnetic waves similarly confined within a one-dimensional cavity exhibit related properties and above all have many features in common with a harmonic oscillator [7].

We have already described several features and properties of the harmonic oscillator (Chap. 2). Meanwhile, we have briefly discussed formation of electromagnetic stationary waves (Chap. 6). There are several resemblance and correspondence between the harmonic oscillator and electromagnetic stationary waves when we view them as mechanical systems. The point is that in a harmonic oscillator the position and momentum are in-quadrature relationship; i.e., their phase difference is $\pi/2$. For the electromagnetic stationary waves, electric field and magnetic field are in-quadrature relationship as well.

In Chap. 6, we examine the conditions under which stationary waves are formed. In a dielectric medium both sides of which are equipped with metal layers (or mirrors), the electric field is described as

$$\boldsymbol{E} = 2E_1\boldsymbol{\varepsilon}_{\mathrm{e}} \sin\ \omega t\ \sin kz. \tag{6.201}$$

In (6.201), we assumed that forward and backward electromagnetic waves are propagating in the direction of the $z$-axis. Here, we assume that the interfaces (or walls) are positioned at $z = 0$ and $z = L$. Within a domain $[0, L]$, the two waves form a stationary wave. Since this expression assumed two waves, the electromagnetic energy was doubled. To normalize the energy so that a single wave is contained, the amplitude $E_1$ of (6.201) should be divided by $\sqrt{2}$. Therefore, we think of a following description for $\boldsymbol{E}$:

$$\boldsymbol{E} = \sqrt{2}E\boldsymbol{e}_1 \sin\omega t\ \sin kz \quad \text{or} \quad E_x = \sqrt{2}E\ \sin\omega t\ \sin kz, \tag{7.104}$$

where we designated the polarization vector as a direction of the $x$-axis. At the same time, we omitted the index from the amplitude. Thus, from the second equation of (5.65) we have

$$\frac{\partial H_y}{\partial t} = -\frac{1}{\mu}\frac{\partial E_x}{\partial z} = -\frac{\sqrt{2}Ek}{\mu}\sin\ \omega t\ \cos kz. \tag{7.105}$$

Note that this equation appears in the second equation of (6.131) as well. Integrating both sides of (7.105), we get

$$H_y = \frac{\sqrt{2}Ek}{\mu\omega}\cos\omega t\ \cos kz.$$

Using a relation $\omega = vk$, we have

$$H_y = \frac{\sqrt{2}E}{\mu v}\cos\omega t\ \cos kz + C,$$

where $v$ is a light velocity in the dielectric medium and $C$ is an integration constant. Removing $C$ and putting

$$H \equiv \frac{E}{\mu v}, \tag{7.106}$$

we have

$$H_y = \sqrt{2}H\ \cos\omega t\ \cos kz. \tag{7.107}$$

Using a vector expression, we have

$$\boldsymbol{H} = \boldsymbol{e}_2 \sqrt{2} H \cos \omega t \cos kz.$$

Thus, $\boldsymbol{E}$ ($\parallel \boldsymbol{e}_1$), $\boldsymbol{H}$ ($\parallel \boldsymbol{e}_2$), and $\boldsymbol{n}$ ($\parallel \boldsymbol{e}_3$) form the right-handed system in this order. As noted in Sects. 6.8 and 6.9, at the interface (or wall) the electric field and magnetic field form nodes and antinodes, respectively. Namely, the two fields are in-quadrature.

Let us calculate electromagnetic energy of the dielectric medium within a cavity. In the present case, the cavity is meant as the dielectric sandwiched by a couple of metal layer. We have

$$W = W_\mathrm{e} + W_\mathrm{m}, \tag{7.108}$$

where $W$ is the total electromagnetic energy; $W_\mathrm{e}$ and $W_\mathrm{m}$ are electric and magnetic energies, respectively. Let the length of the cavity be $L$. Then the energy per unit cross-sectional area is described as

$$W_\mathrm{e} = \frac{\varepsilon}{2} \int_0^L \boldsymbol{E}^2 \mathrm{d}z \quad \text{and} \quad W_\mathrm{m} = \frac{\mu}{2} \int_0^L \boldsymbol{H}^2 \mathrm{d}z. \tag{7.109}$$

Performing integration, we get

$$\begin{aligned} W_\mathrm{e} &= \frac{\varepsilon}{2} L E^2 \sin^2 \omega t, \\ W_\mathrm{m} &= \frac{\mu}{2} L H^2 \cos^2 \omega t = \frac{\mu}{2} L \left( \frac{E}{\mu v} \right)^2 \cos^2 \omega t = \frac{\varepsilon}{2} L E^2 \cos^2 \omega t, \end{aligned} \tag{7.110}$$

where we used $1/v^2 = \varepsilon \mu$ with the last equality. Thus, we have

$$W = \frac{\varepsilon}{2} L E^2 = \frac{\mu}{2} L H^2. \tag{7.111}$$

Representing an energy per unit volume as $\widetilde{W_\mathrm{e}}$, $\widetilde{W_\mathrm{m}}$, and $\widetilde{W}$, we have

$$\widetilde{W_\mathrm{e}} = \frac{\varepsilon}{2} E^2 \sin^2 \omega t, \quad \widetilde{W_\mathrm{m}} = \frac{\varepsilon}{2} E^2 \cos^2 \omega t, \quad \widetilde{W} = \frac{\varepsilon}{2} E^2 = \frac{\mu}{2} H^2. \tag{7.112}$$

In Chap. 2, we treated motion of a harmonic oscillator. There, we had

$$x(t) = \frac{v_0}{\omega} \sin \omega t = x_0 \sin \omega t. \tag{2.7}$$

Here, we have defined an amplitude of the harmonic oscillation as $x_0$ ( $> 0$)

$$x_0 \equiv \frac{v_0}{\omega}.$$

Then, momentum is described as

$$p(t) = m\dot{x}(t) = m\omega x_0 \cos \omega t.$$

Defining

$$p_0 \equiv m\omega x_0,$$

we have

$$p(t) = p_0 \cos \omega t.$$

Let a kinetic energy and potential energy of the oscillator be $K$ and $V$, respectively. Then, we have

$$\begin{aligned} K &= \frac{1}{2m} p(t)^2 = \frac{1}{2m} p_0^2 \cos^2 \omega t\, x(t), \\ V &= \frac{1}{2} m\omega^2 x(t)^2 = \frac{1}{2} m\omega^2 x_0^2 \sin^2 \omega t\, x(t), \\ W &= K + V = \frac{1}{2m} p_0^2 = \frac{1}{2} m v_0^2 = \frac{1}{2} m\omega^2 x_0^2. \end{aligned} \tag{7.113}$$

Comparing (7.112) and (7.113), we recognize the following relationship in energy between the electromagnetic fields and harmonic oscillator motion [7]:

$$\sqrt{m}\omega x_0 \leftrightarrow \sqrt{\varepsilon} E \quad \text{and} \quad p_0/\sqrt{m} \leftrightarrow \sqrt{\mu} H. \tag{7.114}$$

Thus, there is an elegant contradistinction between the dynamics of electromagnetic fields in cavity and motion of a harmonic oscillator. In fact, quantum electromagnetism is based upon the treatment of a quantum harmonic oscillator introduced in Chap. 2.

## References

1. Moore WJ (1955) Physical chemistry, 3rd edn. Prentice-Hall, Englewood Cliffs
2. Smith FG, King TA, Wilkins D (2007) Optics and photonics, 2nd edn. Wiley, Chichester
3. Sunakawa S (1965) Theoretical electromagnetism. Kinokuniya, Tokyo (in Japanese)
4. Loudon R (2000) The quantum theory of light, 3rd edn. Oxford University Press, Oxford
5. Born M, Wolf E (2005) Principles of optics, 7th edn. Cambridge University Press, Cambridge
6. Yamao T, Okuda Y, Makino Y, Hotta S (2011) Dispersion of the refractive indices of thiophene/phenylene co-oligomer single crystals. J Appl Phys 110(5): 053113/7 pages
7. Fox M (2006) Quantum optics. Oxford University Press, Oxford

# Chapter 8
# Introductory Green's Functions

In this chapter, we deal with various properties and characteristics of differential equations, especially first-order linear differential equations (FOLDEs) and second-order linear differential equations (SOLDEs). These differential equations are characterized by differential operators and boundary conditions (BCs). Of these, differential operators appearing in SOLDEs are particularly important. Under appropriate conditions, the said operators can be converted to Hermitian operators. The SOLDEs associated to classical orthogonal polynomials play a central role in many fields of mathematical physics including quantum mechanics and electromagnetism. We study the general principle of SOLDEs in relation to several specific SOLDEs we have studied in Part I and examine general features of an eigenvalue problem and an initial value problem (IVP). In this context, Green's functions provide a powerful tool for solving SOLDEs. For a practical purpose, we deal with actual construction of Green's functions. In Sect. 6.9, we dealt with steady-state characteristics of electromagnetic waves in dielectrics in terms of propagation, reflection, and transmission. When we consider transient characteristics of electromagnetic and optical phenomena, we often need to deal with SOLDEs having constant coefficients. This is well known in connection with a motion of a damped harmonic oscillator. In the latter part of this chapter, we treat the initial value problem of a SOLDE of this type.

## 8.1 Second-Order Linear Differential Equations (SOLDEs)

A general form of $n$-th-order linear differential equations has the following form:

$$a_n(x)\frac{\mathrm{d}^n u}{\mathrm{d}x^n} + a_{n-1}(x)\frac{\mathrm{d}^{n-1}u}{\mathrm{d}x^{n-1}} + \cdots + a_1(x)\frac{\mathrm{d}u}{\mathrm{d}x} + a_0(x)u = d(x). \tag{8.1}$$



S. Hotta, *Mathematical Physical Chemistry*,
https://doi.org/10.1007/978-981-10-7671-8_8

If $\mathrm{d}(x) = 0$, the differential equation is said to be homogeneous; otherwise, it is called inhomogeneous. Equation (8.1) is a linear function of $u$ and its derivatives. Likewise, we have a SOLDE such that

$$a(x)\frac{\mathrm{d}^2u}{\mathrm{d}x^2} + b(x)\frac{\mathrm{d}u}{\mathrm{d}x} + c(x)u = d(x). \tag{8.2}$$

In (8.2), we assume that the variable $x$ is real. The equation can be solved under appropriate boundary conditions (BCs). A general form of BCs is described as

$$B_1(u) = \alpha_1 u(a) + \beta_1 \left.\frac{\mathrm{d}u}{\mathrm{d}x}\right|_{x=a} + \gamma_1 u(b) + \delta_1 \left.\frac{\mathrm{d}u}{\mathrm{d}x}\right|_{x=b} = \sigma_1, \tag{8.3}$$

$$B_2(u) = \alpha_2 u(a) + \beta_2 \left.\frac{\mathrm{d}u}{\mathrm{d}x}\right|_{x=a} + \gamma_2 u(b) + \delta_2 \left.\frac{\mathrm{d}u}{\mathrm{d}x}\right|_{x=b} = \sigma_2, \tag{8.4}$$

where $\alpha_1$, $\beta_1$, $\gamma_1$, $\delta_1$ $\sigma_1$, etc., are real constants; $u(x)$ is defined in an interval $[a, b]$, where $a$ and $b$ can be infinity (i.e., $\pm\infty$). The LHS of $B_1(u)$ and $B_2(u)$ are referred to as boundary functionals [1, 2]. If $\sigma_1 = \sigma_2 = 0$, the BCs are called homogeneous; otherwise, the BCs are said to be inhomogeneous. In combination with the inhomogeneous equation expressed as (8.2), Table 8.1 summarizes characteristics of SOLDEs. We have four types of SOLDEs according to homogeneity and inhomogeneity of equations and BCs.

Even though SOLDEs are mathematically tractable, yet it is not easy necessarily to solve them depending upon the nature of $a(x)$, $b(x)$, and $c(x)$ of (8.2). Nonetheless, if those functions are constant coefficients, it can readily be solved. We will deal with SOLDEs of that type in great deal later. Suppose that we find two linearly independent solutions $u_1(x)$ and $u_2(x)$ of a following homogeneous equation:

$$a(x)\frac{\mathrm{d}^2u}{\mathrm{d}x^2} + b(x)\frac{\mathrm{d}u}{\mathrm{d}x} + c(x)u = 0. \tag{8.5}$$

Then, any solution $u(x)$ of (8.5) can be expressed as their linear combination such that

$$u(x) = c_1 u_1(x) + c_2 u_2(x), \tag{8.6}$$

where $c_1$ and $c_2$ are some arbitrary constants.

**Table 8.1** Characteristics of SOLDEs

| | Type I | Type II | Type III | Type IV |
|---|---|---|---|---|
| Equation | Homogeneous | Homogeneous | Inhomogeneous | Inhomogeneous |
| Boundary conditions | Homogeneous | Inhomogeneous | Homogeneous | Inhomogeneous |

In general, suppose that there are arbitrarily chosen $n$ functions; i.e., $f_1(x), f_2(x), \cdots, f_n(x)$. Suppose a following equation with those functions:

$$a_1 f_1(x) + a_2 f_2(x) + \cdots + a_n f_n(x) = 0, \tag{8.7}$$

where $a_1, a_2, \cdots,$ and $a_n$ are constants. If $a_1 = a_2 = \cdots = a_n = 0$, (8.7) always holds. In this case, (8.7) is said to be a trivial linear relation. If $f_1(x), f_2(x), \cdots,$ and $f_n(x)$ satisfy a non-trivial linear relation, $f_1(x), f_2(x), \cdots, f_n(x)$ are said to be linearly dependent. That is, the non-trivial expression means that in (8.7) at least one of $a_1, a_2, \cdots,$ and $a_n$ is nonzero. Suppose that $a_n \neq 0$. Then, from (8.7), $f_n(x)$ is expressed as

$$f_n(x) = -\frac{a_1}{a_n} f_1(x) - \frac{a_2}{a_n} f_2(x) - \cdots - \frac{a_{n-1}}{a_n} f_{n-1}(x). \tag{8.8}$$

If $f_1(x), f_2(x), \cdots, f_n(x)$ are not linearly dependent, they are called linearly independent. In other words, the statement that $f_1(x), f_2(x), \cdots, f_n(x)$ are linearly independent is equivalent to that (8.7) holds if and only if $a_1 = a_2 = \cdots = a_n = 0$. We will have relevant discussion in Part III.

Now suppose that with the above two linearly independent $u_1(x)$ and $u_2(x)$ we have

$$a_1 u_1(x) + a_2 u_2(x) = 0. \tag{8.9}$$

Differentiating (8.9), we have

$$a_1 \frac{\mathrm{d}u_1(x)}{\mathrm{d}x} + a_2 \frac{\mathrm{d}u_2(x)}{\mathrm{d}x} = 0. \tag{8.10}$$

Expressing (8.9) and (8.10) in a matrix form, we get

$$\begin{pmatrix} u_1(x) & u_2(x) \\ \frac{\mathrm{d}u_1(x)}{\mathrm{d}x} & \frac{\mathrm{d}u_2(x)}{\mathrm{d}x} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = 0. \tag{8.11}$$

Thus, that $u_1(x)$ and $u_2(x)$ are linearly independent is equivalent to that the following expression holds:

$$\begin{vmatrix} u_1(x) & u_2(x) \\ \frac{\mathrm{d}u_1(x)}{\mathrm{d}x} & \frac{\mathrm{d}u_2(x)}{\mathrm{d}x} \end{vmatrix} \equiv W(u_1, u_2) \neq 0, \tag{8.12}$$

where $W(u_1, u_2)$ is called Wronskian of $u_1(x)$ and $u_2(x)$. In fact, if $W(u_1, u_2) = 0$, then we have

$$u_1 \frac{\mathrm{d}u_2}{\mathrm{d}x} - u_2 \frac{\mathrm{d}u_1}{\mathrm{d}x} = 0. \tag{8.13}$$

This implies that there is a functional relationship between $u_1$ and $u_2$. In fact, if we can express as $u_2 = u_2(u_1(x))$, then $\frac{du_2}{dx} = \frac{du_2}{du_1}\frac{du_1}{dx}$. That is,

$$u_1\frac{du_2}{dx} = u_1\frac{du_2}{du_1}\frac{du_1}{dx} = u_2\frac{du_1}{dx} \quad \text{or} \quad u_1\frac{du_2}{du_1} = u_2 \quad \text{or} \quad \frac{du_2}{u_2} = \frac{du_1}{u_1}, \tag{8.14}$$

where the second equality of the first equation comes from (8.13). The third equation can easily be integrated to yield

$$\ln\frac{u_2}{u_1} = c \quad \text{or} \quad \frac{u_2}{u_1} = e^c \quad \text{or} \quad u_2 = e^c u_1. \tag{8.15}$$

Equation (8.15) shows that $u_1(x)$ and $u_2(x)$ are linearly dependent. It is easy to show if $u_1(x)$ and $u_2(x)$ are linearly dependent, $W(u_1, u_2) = 0$. Thus, we have a following statement:

Two functions are linearly dependent. $\Leftrightarrow$ $W(u_1, u_2) = 0$.

Then, as the contraposition of this statement, we have

Two functions are linearly independent. $\Leftrightarrow$ $W(u_1, u_2) \neq 0$.

On the other hand, suppose that we have another solution $u_3(x)$ for (8.5) besides $u_1(x)$ and $u_2(x)$. Then, we have

$$\begin{aligned} a(x)\frac{d^2u_1}{dx^2} + b(x)\frac{du_1}{dx} + c(x)u_1 &= 0, \\ a(x)\frac{d^2u_2}{dx^2} + b(x)\frac{du_2}{dx} + c(x)u_2 &= 0, \\ a(x)\frac{d^2u_3}{dx^2} + b(x)\frac{du_3}{dx} + c(x)u_3 &= 0. \end{aligned} \tag{8.16}$$

Again rewriting (8.16) in a matrix form, we have

$$\begin{pmatrix} \frac{d^2u_1}{dx^2} & \frac{du_1}{dx} & u_1 \\ \frac{d^2u_2}{dx^2} & \frac{du_2}{dx} & u_2 \\ \frac{d^2u_3}{dx^2} & \frac{du_3}{dx} & u_3 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = 0. \tag{8.17}$$

A necessary and sufficient condition to obtain a non-trivial solution (i.e., a solution besides $a = b = c = 0$) is that [1]

$$\begin{vmatrix} \frac{d^2u_1}{dx^2} & \frac{du_1}{dx} & u_1 \\ \frac{d^2u_2}{dx^2} & \frac{du_2}{dx} & u_2 \\ \frac{d^2u_3}{dx^2} & \frac{du_3}{dx} & u_3 \end{vmatrix} = 0. \tag{8.18}$$

Note here that

$$\begin{vmatrix} \frac{\mathrm{d}^2 u_1}{\mathrm{d}x^2} & \frac{\mathrm{d}u_1}{\mathrm{d}x} & u_1 \\ \frac{\mathrm{d}^2 u_2}{\mathrm{d}x^2} & \frac{\mathrm{d}u_2}{\mathrm{d}x} & u_2 \\ \frac{\mathrm{d}^2 u_3}{\mathrm{d}x^2} & \frac{\mathrm{d}u_3}{\mathrm{d}x} & u_3 \end{vmatrix} = - \begin{vmatrix} u_1 & u_2 & u_3 \\ \frac{\mathrm{d}u_1}{\mathrm{d}x} & \frac{\mathrm{d}u_2}{\mathrm{d}x} & \frac{\mathrm{d}u_3}{\mathrm{d}x} \\ \frac{\mathrm{d}^2 u_1}{\mathrm{d}x^2} & \frac{\mathrm{d}^2 u_2}{\mathrm{d}x^2} & \frac{\mathrm{d}^2 u_3}{\mathrm{d}x^2} \end{vmatrix} \equiv -W(u_1, u_2, u_3), \tag{8.19}$$

where $W(u_1, u_2, u_3)$ is Wronskian of $u_1(x)$, $u_2(x)$, and $u_3(x)$. In the above relation, we used the fact that a determinant of a matrix is identical to that of its transposed matrix and that a determinant of a matrix changes the sign after permutation of row vectors. To be short, a necessary and sufficient condition to get a non-trivial solution is that $W(u_1, u_2, u_3)$ vanishes.

This implies that $u_1(x)$, $u_2(x)$, and $u_3(x)$ are linearly dependent. However, we have assumed that $u_1(x)$ and $u_2(x)$ are linearly independent, and so (8.18) and (8.19) mean that $u_3(x)$ must be described as a linear combination of $u_1(x)$ and $u_2(x)$. That is, we have no third linearly independent solution. Consequently, the general solution of (8.5) must be given by (8.6). In this sense, $u_1(x)$ and $u_2(x)$ are said to be a fundamental set of solutions of (8.5).

Next, let us consider the inhomogeneous equation of (8.2). Suppose that $u_\mathrm{p}(x)$ is a particular solution of (8.2). Let us think of a following function $v(x)$ such that:

$$u(x) = v(x) + u_\mathrm{p}(x). \tag{8.20}$$

Substituting (8.20) for (8.2), we have

$$a(x)\frac{\mathrm{d}^2 v}{\mathrm{d}x^2} + b(x)\frac{\mathrm{d}v}{\mathrm{d}x} + c(x)v + a(x)\frac{\mathrm{d}^2 u_\mathrm{p}}{\mathrm{d}x^2} + b(x)\frac{\mathrm{d}u_\mathrm{p}}{\mathrm{d}x} + c(x)u_\mathrm{p} = d(x). \tag{8.21}$$

Therefore, we have

$$a(x)\frac{\mathrm{d}^2 v}{\mathrm{d}x^2} + b(x)\frac{\mathrm{d}v}{\mathrm{d}x} + c(x)v = 0.$$

But, $v(x)$ can be described by a linear combination of $u_1(x)$ and $u_2(x)$ as in the case of (8.6). Hence, the general solution of (8.2) should be expressed as

$$u(x) = c_1 u_1(x) + c_2 u_2(x) + u_\mathrm{p}(x), \tag{8.22}$$

where $c_1$ and $c_2$ are arbitrary (complex) constants.

## 8.2 First-Order Linear Differential Equations (FOLDEs)

In a discussion that follows, first-order linear differential equations (FOLDEs) supply us with useful information. A general form of FOLDEs is expressed as

$$a(x)\frac{\mathrm{du}}{\mathrm{d}x} + b(x)u = d(x)\ [a(x) \neq 0]. \tag{8.23}$$

An associated boundary condition is given by a boundary functional $B(u)$ such that

$$B(u) = \alpha u(a) + \beta u(b) = \sigma, \tag{8.24}$$

where $\alpha$, $\beta$, and $\sigma$ are real constants; $u(x)$ is defined in an interval $[a, b]$. If in (8.23) $d(x) \equiv 0$, (8.23) can readily be integrated to yield a solution. Let us multiply both sides of (8.23) by $w(x)$. Then, we have

$$w(x)a(x)\frac{\mathrm{d}u}{\mathrm{d}x} + w(x)b(x)u = w(x)d(x). \tag{8.25}$$

We define $p(x)$ as

$$p(x) \equiv w(x)a(x), \tag{8.26}$$

where $w(x)$ is called a *weight function*. As mentioned in Sect. 2.3, the weight function is a real and nonnegative function within the domain considered. Here we suppose that

$$\frac{\mathrm{d}p(x)}{\mathrm{d}x} = w(x)b(x). \tag{8.27}$$

Then, (8.25) can be rewritten as

$$\frac{\mathrm{d}}{\mathrm{d}x}[p(x)u] = w(x)d(x). \tag{8.28}$$

Thus, we can immediately integrate (8.28) to obtain a solution

$$u = \frac{1}{p(x)}\left[\int^x w(x')\mathrm{d}(x')\mathrm{d}x' + C\right], \tag{8.29}$$

where $C$ is an arbitrary integration constant.

To seek $w(x)$, from (8.26) and (8.27) we have

$$p' = (wa)' = wb = wa\left(\frac{b}{a}\right). \tag{8.30}$$

This can easily be integrated for $wa$ to be expressed as

$$wa = C' \exp\left(\int \frac{b}{a} \mathrm{d}x\right) \text{ or } w = \frac{C'}{a} \exp\left(\int \frac{b}{a} \mathrm{d}x\right), \tag{8.31}$$

where $C'$ is an arbitrary integration constant. The quantity $\frac{C'}{a}$ must be nonnegative so that $w$ can be nonnegative.

*Example 8.1* Let us think of a following FOLDE within an interval $[a, b]$; i.e., $a \leq x \leq b$.

$$\frac{\mathrm{d}u}{\mathrm{d}x} + xu = x. \tag{8.32}$$

A boundary condition is set such that

$$u(a) = \sigma. \tag{8.33}$$

Notice that (8.33) is obtained by setting $\alpha = 1$ and $\beta = 0$ in (8.24). Following the above argument, we obtain a solution described as

$$u = \frac{1}{p(x)} \left[\int_a^x w(x')d(x')dx' + u(a)p(a)\right]. \tag{8.34}$$

Also, we have

$$p(x) = w(x) = \exp(\int_a^x x' \mathrm{d}x') = \exp\left[\frac{1}{2}\left(x^2 - a^2\right)\right]. \tag{8.35}$$

The integration of RHS can be performed as follows:

$$\begin{aligned}\int_a^x w(x')d(x')\mathrm{d}x' &= \int_a^x x' \exp\left[\frac{1}{2}\left(x'^2 - a^2\right)\right]dx' = \exp\left(-\frac{a^2}{2}\right) \int_{a^2/2}^{x^2/2} \exp t \mathrm{d}t \\ &= \exp\left(-\frac{a^2}{2}\right)\left[\exp\left(\frac{x^2}{2}\right) - \exp\left(\frac{a^2}{2}\right)\right] = \exp\left(-\frac{a^2}{2}\right)\exp\left(\frac{x^2}{2}\right) - 1 = p(x) - 1,\end{aligned} \tag{8.36}$$

where with the second equality we used an integration by substitution of $\frac{1}{2}x'^2 \to t$. Considering (8.33) and putting $p(a) = 1$, (8.34) is rewritten as

$$\begin{aligned}u &= \frac{1}{p(x)}[p(x) - 1 + \sigma] = 1 + \frac{\sigma - 1}{p(x)} \\ &= 1 + (\sigma - 1)\exp\left[\frac{a^2 - x^2}{2}\right].\end{aligned} \tag{8.37}$$

where with the last equality we used (8.35).

Notice that when $\sigma = 1$, $u(x) \equiv 1$. This is because $u(x) \equiv 1$ is certainly a solution of (8.32) and satisfies the BC of (8.33). Uniqueness of a solution imposes this strict condition upon (8.37).

In the next two examples, we make general discussions.

*Example 8.2* Let us consider a following differential operator:

$$L_x = \frac{\mathrm{d}}{\mathrm{d}x}. \tag{8.38}$$

We think of a following identity using an integration by parts:

$$\int_a^b \mathrm{d}x \varphi^* \left(\frac{\mathrm{d}}{\mathrm{d}x}\psi\right) + \int_a^b \mathrm{d}x \left(\frac{\mathrm{d}}{\mathrm{d}x}\varphi^*\right)\psi = [\varphi^*\psi]_a^b. \tag{8.39}$$

Rewriting this, we get

$$\int_a^b \mathrm{d}x \varphi^* \left(\frac{\mathrm{d}}{\mathrm{d}x}\psi\right) - \int_a^b \mathrm{d}x \left[-\frac{\mathrm{d}}{\mathrm{d}x}\varphi\right]^* \psi = [\varphi^*\psi]_a^b. \tag{8.40}$$

Looking at (8.40), we notice that LHS comprises a difference between two integrals, while RHS referred to as a boundary term (or surface term) does not contain an integral.

Recalling the expression (1.128) and defining

$$\widetilde{L_x} \equiv -\frac{\mathrm{d}}{\mathrm{d}x}$$

in (8.40), we have

$$\langle \varphi | L_x \psi \rangle - \left\langle \widetilde{L_x}\varphi | \psi \right\rangle = [\varphi^*\psi]_a^b. \tag{8.41}$$

Here, RHS of (8.41) needs to vanish so that we can have

$$\langle \varphi | L_x \psi \rangle = \left\langle \widetilde{L_x}\varphi | \psi \right\rangle. \tag{8.42}$$

Meanwhile, adopting the expression (1.112) with respect to an adjoint operator, we have a following expression such that

$$\langle \varphi | L_x \psi \rangle = \left\langle L_x^\dagger \varphi | \psi \right\rangle. \tag{8.43}$$

Comparing (8.42) and (8.43) and considering that $\varphi$ and $\psi$ are arbitrary functions, we have $\widetilde{L_x} = L_x^\dagger$. Thus, as an operator adjoint to $L_x$, we get

$$L_x^\dagger = \widetilde{L_x} = -\frac{\mathrm{d}}{\mathrm{d}x} = -L_x.$$

Notice that only if the surface term vanishes, the adjoint operator $L_x^\dagger$ can appropriately be defined. We will encounter a similar expression again in Part III.

We add that if with an operator $A$, we have a relation described by

$$A^\dagger = -A, \tag{8.44}$$

the operator $A$ is said to be anti-Hermitian. We have already encountered such an operator in Sect. 1.5.

Let us then examine on what condition the surface term vanishes. The RHS of (8.40) and (8.41) is given by

$$\varphi^*(b)\psi(b) - \varphi^*(a)\psi(a).$$

For this term to vanish, we should have

$$\varphi^*(b)\psi(b) = \varphi^*(a)\psi(a) \quad \text{or} \quad \frac{\varphi^*(b)}{\varphi^*(a)} = \frac{\psi(a)}{\psi(b)}.$$

If $\psi(b) = 2\psi(a)$, then we should have $\varphi(b) = \frac{1}{2}\varphi(a)$ for the surface term to vanish. Recalling (8.24), the above conditions are expressed as

$$B(\psi) = 2\psi(a) - \psi(b) = 0, \tag{8.45}$$

$$B'(\varphi) = \frac{1}{2}\varphi(a) - \varphi(b) = 0. \tag{8.46}$$

The boundary functional $B'(\varphi)$ is said to be adjoint to $B(\psi)$. The two boundary functionals are admittedly different. If, however, we set $\psi(b) = \psi(a)$, then we should have $\varphi(b) = \varphi(a)$ for the surface term to vanish. That is,

$$B(\psi) = \psi(a) - \psi(b) = 0, \tag{8.47}$$

$$B'(\varphi) = \varphi(a) - \varphi(b) = 0. \tag{8.48}$$

Thus, the two functionals are identical and $\psi$ and $\varphi$ satisfy homogeneous BCs with respect to these functionals.

As discussed above, a FOLDE is characterized by its differential operator as well as a BC (or boundary functional). This is similarly the case with SOLDEs as well.

*Example 8.3* Next, let us consider a following differential operator:

$$L_x = \frac{1}{i}\frac{\mathrm{d}}{\mathrm{d}x}. \tag{8.49}$$

As in the case of Example 8.2, we have

$$\int_a^b \mathrm{d}x\varphi^*\left(\frac{1}{i}\frac{\mathrm{d}}{\mathrm{d}x}\psi\right) - \int_a^b \mathrm{d}x\left[\frac{1}{i}\frac{\mathrm{d}}{\mathrm{d}x}\varphi\right]^*\psi = \frac{1}{i}[\varphi^*\psi]_a^b. \tag{8.50}$$

Also rewriting (8.50) using an inner product notation, we get

$$\langle\varphi|L_x\psi\rangle - \langle L_x\varphi|\psi\rangle = \frac{1}{i}[\varphi^*\psi]_a^b, \tag{8.51}$$

Apart from the factor $\frac{1}{i}$, RHS of (8.51) is again given by

$$\varphi^*(b)\psi(b) - \varphi^*(a)\psi(a).$$

Repeating a discussion similar to Example 8.2, when the surface term vanishes, we get

$$\langle\varphi|L_x\psi\rangle = \langle L_x\varphi|\psi\rangle. \tag{8.52}$$

Comparing (8.43) and (8.52), we have

$$\langle L_x\varphi|\psi\rangle = \left\langle L_x^\dagger\varphi|\psi\right\rangle. \tag{8.53}$$

Again considering that $\varphi$ and $\psi$ are arbitrary functions, we get

$$L_x^\dagger = L_x. \tag{8.54}$$

As in (8.54), if the differential operator is identical to its adjoint operator, such an operator is called *self-adjoint*. On the basis of (1.119), $L_x$ would apparently be Hermitian. However, we have to be careful to assure that $L_x$ is Hermitian. For a differential operator to be Hermitian, (i) the said operator must be self-adjoint. (ii) The two boundary functionals adjoint to each other must be identical. In other words, $\psi$ and $\varphi$ must satisfy the same homogeneous BCs with respect to these functionals. In this example, we must have the same boundary functionals as those described by (8.47) and (8.48). If and only if the conditions (i) and (ii) are satisfied, the operator is said to be Hermitian. It seems somewhat a formal expression. Nonetheless, satisfaction of these conditions is also the case with second-order differential operators so that these operators can be Hermitian. In fact, SOLDEs we studied in Part I are essentially dealt with within the framework of the aforementioned formalism.

## 8.3 Second-Order Differential Operators

The second-order differential operators are the most common operators and frequently treated in mathematical physics. The general differential operators are described as

$$L_x = a(x)\frac{d^2}{dx^2} + b(x)\frac{d}{dx} + c(x), \tag{8.55}$$

where $a(x)$, $b(x)$, and $c(x)$ can in general be complex functions of a real variable $x$. Let us think of following identities [1]:

$$\begin{aligned} v^* a\frac{\mathrm{d}^2 u}{\mathrm{d}x^2} - u\frac{\mathrm{d}^2(av^*)}{\mathrm{d}x^2} &= \frac{\mathrm{d}}{\mathrm{d}x}\left[av^*\frac{\mathrm{d}u}{\mathrm{d}x} - u\frac{\mathrm{d}(av^*)}{\mathrm{d}x}\right], \\ v^* b\frac{\mathrm{d}u}{\mathrm{d}x} + u\frac{\mathrm{d}(bv^*)}{\mathrm{d}x} &= \frac{\mathrm{d}}{\mathrm{d}x}[buv^*], \\ v^* cu - ucv^* &= 0. \end{aligned} \tag{8.56}$$

Summing both sides of (8.56), we have an identity

$$\begin{aligned} &v^*\left\{a\frac{\mathrm{d}^2 u}{\mathrm{d}x^2} + b\frac{\mathrm{d}u}{\mathrm{d}x} + cu\right\} - u\left\{\frac{\mathrm{d}^2(av^*)}{\mathrm{d}x^2} - \frac{\mathrm{d}(bv^*)}{\mathrm{d}x} + cv^*\right\} \\ &= \frac{\mathrm{d}}{\mathrm{d}x}\left[av^*\frac{\mathrm{d}u}{\mathrm{d}x} - u\frac{\mathrm{d}(av^*)}{\mathrm{d}x}\right] + \frac{\mathrm{d}}{\mathrm{d}x}[buv^*]. \end{aligned} \tag{8.57}$$

Hence, following the expressions of Sect. 8.2, we define $L_x^\dagger$ such that

$$[L_x^\dagger v]^* \equiv \frac{\mathrm{d}^2(av^*)}{\mathrm{d}x^2} - \frac{\mathrm{d}(bv^*)}{\mathrm{d}x} + cv^*. \tag{8.58}$$

Taking a complex conjugate of both sides, we get

$$L_x^\dagger v = \frac{\mathrm{d}^2(a^*v)}{\mathrm{d}x^2} - \frac{\mathrm{d}(b^*v)}{\mathrm{d}x} + c^*v. \tag{8.59}$$

Considering the differential of a product function, we have as $L_x^\dagger$

$$L_x^\dagger = a^*\frac{\mathrm{d}^2}{\mathrm{d}x^2} + \left(2\frac{\mathrm{d}a^*}{\mathrm{d}x} - b^*\right)\frac{\mathrm{d}}{\mathrm{d}x} + \frac{\mathrm{d}^2 a^*}{\mathrm{d}x^2} - \frac{\mathrm{d}b^*}{\mathrm{d}x} + c^*. \tag{8.60}$$

Replacing (8.57) with (8.55) and (8.59), we have

$$v^*(L_x u) - [L_x^\dagger v]^* u = \frac{\mathrm{d}}{\mathrm{d}x}\left[av^*\frac{\mathrm{d}u}{\mathrm{d}x} - u\frac{\mathrm{d}(av^*)}{\mathrm{d}x} + buv^*\right]. \tag{8.61}$$

Assuming that the relevant SOLDE is defined in $[r, s]$ and integrating (8.61) within that interval, we get

$$\int_r^s \mathrm{d}x\left[v^*(L_x u) - [L_x^\dagger v]^* u\right] = \left[av^*\frac{\mathrm{d}u}{\mathrm{d}x} - u\frac{\mathrm{d}(av^*)}{\mathrm{d}x} + buv^*\right]_r^s \tag{8.62}$$

Using the definition of an inner product described in (1.128) and rewriting (8.62), we have

$$\langle v|L_x u\rangle - \left\langle L_x^\dagger v|u\right\rangle = \left[av^*\frac{\mathrm{d}u}{\mathrm{d}x} - u\frac{\mathrm{d}(av^*)}{\mathrm{d}x} + buv^*\right]_r^s.$$

Here if RHS of the above (i.e., the surface term of the above expression) vanishes, we get

$$\langle v|L_x u\rangle = \left\langle L_x^\dagger v|u\right\rangle.$$

We find that this notation is consistent with (1.112).

Bearing in mind this situation, let us seek a condition under which the differential operator $L_x$ is Hermitian. Suppose here that $a(x)$, $b(x)$, and $c(x)$ are all real and that

$$\frac{\mathrm{d}a(x)}{\mathrm{d}x} = b(x).$$

Then, instead of (8.60), we have

$$L_x^\dagger = a(x)\frac{\mathrm{d}^2}{\mathrm{d}x^2} + b(x)\frac{\mathrm{d}}{\mathrm{d}x} + c(x) = L_x. \tag{8.63}$$

Thus, we are successful in constituting a self-adjoint operator $L_x$. In that case, (8.62) can be rewritten as

$$\int_r^s \mathrm{d}x[v^*(L_x u) - (L_x v)^* u] = \left[a\left(v^*\frac{\mathrm{d}u}{\mathrm{d}x} - u\frac{\mathrm{d}v^*}{\mathrm{d}x}\right)\right]_r^s. \tag{8.64}$$

Notice that $b(x)$ is eliminated from (8.64). If RHS of (8.64) vanishes, we get

$$\langle v|L_x u\rangle = \langle L_x v|u\rangle.$$

This notation is consistent with (1.119) and the Hermiticity of $L_x$ becomes well defined.

If we do not have the condition of $\frac{\mathrm{d}a(x)}{\mathrm{d}x} = b(x)$, how can we deal with the problem? The answer is that following the procedures in Sect. 8.2, we can convert $L_x$ to a self-adjoint operator by multiplying $L_x$ by a weight function $w(x)$ introduced in (8.26), (8.27), and (8.31). Replacing $a(x)$, $b(x)$, and $c(x)$ with $w(x)a(x)$, $w(x)b(x)$, and $w(x)c(x)$, respectively, in the identity (8.57), we rewrite (8.57) as

$$\begin{aligned} &v^*\left\{aw\frac{\mathrm{d}^2u}{\mathrm{d}x^2} + bw\frac{\mathrm{d}u}{\mathrm{d}x} + cwu\right\} - u\left\{\frac{\mathrm{d}^2(awv^*)}{\mathrm{d}x^2} - \frac{\mathrm{d}(bwv^*)}{\mathrm{d}x} + cwv^*\right\} \\ &= \frac{\mathrm{d}}{\mathrm{d}x}\left[awv^*\frac{\mathrm{d}u}{\mathrm{d}x} - u\frac{\mathrm{d}(awv^*)}{\mathrm{d}x}\right] + \frac{\mathrm{d}}{\mathrm{d}x}[bwuv^*]. \end{aligned} \tag{8.65}$$

Let us calculate $\{\cdots\}$ of the second term for LHS of (8.65). Using (8.26) and (8.27), we have

$$\begin{aligned} \frac{\mathrm{d}^2(awv^*)}{\mathrm{d}x^2} - \frac{\mathrm{d}(bwv^*)}{\mathrm{d}x} + cwv^* &= [(aw)'v^* + awv^{*'}]' - [\Big(bw\Big)'v^* + bwv^{*'}\Big] + cwv^* \\ &= \Big[bwv^* + awv^{*'}\Big]' - \Big[\Big(bw\Big)'v^* + bwv^{*'}\Big] + cwv^* \\ &= (bw)'v^* + bwv^{*'} + (aw)'v^{*'} + awv^{*''} - (bw)'v^* - bwv^{*'} + cwv^* \\ &= (aw)'v^{*'} + awv^{*''} + cwv^* = bwv^{*'} + awv^{*''} + cwv^* \\ &= w\Big(av^{*''} + bv^{*'} + cv^*\Big) = w\Big(a^*v^{*''} + b^*v^{*'} + c^*v^*\Big) \\ &= w(av'' + bv' + cv)^*. \end{aligned} \tag{8.66}$$

The second last equality of (8.66) is based on the assumption that $a(x)$, $b(x)$, and $c(x)$ are real functions. Meanwhile, for RHS of (8.65), we have

$$\begin{aligned} &\frac{\mathrm{d}}{dx}\left[awv^*\frac{\mathrm{d}u}{\mathrm{d}x} - u\frac{\mathrm{d}(awv^*)}{\mathrm{d}x}\right] + \frac{\mathrm{d}}{\mathrm{d}x}[bwuv^*] \\ &= \frac{\mathrm{d}}{\mathrm{d}x}\left[awv^*\frac{\mathrm{d}u}{\mathrm{d}x} - u(aw)'v^* - uaw\frac{\mathrm{d}v^*}{\mathrm{d}x}\right] + \frac{\mathrm{d}}{\mathrm{d}x}[bwuv^*] \\ &= \frac{\mathrm{d}}{\mathrm{d}x}\left[awv^*\frac{\mathrm{d}u}{\mathrm{d}x} - ubwv^* - uaw\frac{\mathrm{d}v^*}{\mathrm{d}x}\right] + \frac{\mathrm{d}}{\mathrm{d}x}[bwuv^*] \\ &= \frac{\mathrm{d}}{\mathrm{d}x}\left[aw\left(v^*\frac{\mathrm{d}u}{\mathrm{d}x} - u\frac{\mathrm{d}v^*}{\mathrm{d}x}\right)\right] = \frac{\mathrm{d}}{\mathrm{d}x}\left[p\left(v^*\frac{\mathrm{d}u}{\mathrm{d}x} - u\frac{\mathrm{d}v^*}{\mathrm{d}x}\right)\right]. \end{aligned} \tag{8.67}$$

With the last equality of (8.67), we used (8.26). Using (8.66) and (8.67), we rewrite (8.65) once again as

$$v^* w\left[a\frac{\mathrm{d}^2 u}{\mathrm{d}x^2}+b\frac{\mathrm{d}u}{\mathrm{d}x}+cu\right]-uw\left[a\frac{\mathrm{d}^2 v}{\mathrm{d}x^2}+b\frac{\mathrm{d}v}{\mathrm{d}x}+cv\right]^*=\frac{\mathrm{d}}{\mathrm{d}x}\left[p\left(v^*\frac{\mathrm{d}u}{\mathrm{d}x}-u\frac{\mathrm{d}v^*}{\mathrm{d}x}\right)\right]. \tag{8.68}$$

Then, integrating (8.68) from $r$ to $s$, we finally get

$$\int_r^s \mathrm{d}x w(x)\{v^*(L_x u)-[L_x v]^* u\}=\left[p(v^*\frac{\mathrm{d}u}{\mathrm{d}x}-u\frac{\mathrm{d}v^*}{\mathrm{d}x})\right]_r^s. \tag{8.69}$$

The relations (8.69) along with (8.62) are called the generalized Green's identity.

We emphasize that as far as the coefficients $a(x)$, $b(x)$, and $c(x)$ in (8.55) are real functions, the associated differential operator $L_x$ can be converted to a self-adjoint form following the procedures of (8.66) and (8.67).

In the above, LHS of the original homogeneous differential equation (8.5) is rewritten as

$$a(x)w(x)\frac{\mathrm{d}^2 u}{\mathrm{d}x^2}+b(x)w(x)\frac{\mathrm{d}u}{\mathrm{d}x}+c(x)w(x)u=\frac{\mathrm{d}}{\mathrm{d}x}\left[p(x)\frac{\mathrm{d}u}{\mathrm{d}x}\right]+cw(x)u,$$

Rewriting this, we have

$$L_x u=\frac{1}{w(x)}\frac{\mathrm{d}}{\mathrm{d}x}\left[p(x)\frac{\mathrm{d}u}{\mathrm{d}x}\right]+cu\quad[w(x)>0], \tag{8.70}$$

where we have

$$p(x)=a(x)w(x)\text{ and }\frac{\mathrm{d}p(x)}{\mathrm{d}x}=b(x)w(x). \tag{8.71}$$

When the differential operator $L_x$ is defined as (8.70), $L_x$ is said to be self-adjoint with respect to a weight function of $w(x)$.

Now we examine boundary functionals. The homogeneous adjoint boundary functionals are described as follows:

$$B_1^\dagger(u)=\alpha_1 v^*(a)+\beta_1\left.\frac{\mathrm{d}v^*}{\mathrm{d}x}\right|_{x=a}+\gamma_1 v^*(b)+\delta_1\left.\frac{\mathrm{d}v^*}{\mathrm{d}x}\right|_{x=b}=0, \tag{8.72}$$

$$B_2^\dagger(u)=\alpha_2 v^*(a)+\beta_2\left.\frac{\mathrm{d}v^*}{\mathrm{d}x}\right|_{x=a}+\gamma_2 v^*(b)+\delta_2\left.\frac{\mathrm{d}v^*}{\mathrm{d}x}\right|_{x=b}=0. \tag{8.73}$$

In (8.3), putting $\alpha_1 = 1$ and $\beta_1 = \gamma_1 = \delta_1 = 0$, we have

$$B_1(u) = u(a) = \sigma_1. \tag{8.74}$$

Also putting $\gamma_2 = 1$ and $\alpha_2 = \beta_2 = \delta_2 = 0$, we have

$$B_2(u) = u(b) = \sigma_2. \tag{8.75}$$

Further putting

$$\sigma_1 = \sigma_2 = 0, \tag{8.76}$$

we also get homogeneous BCs of

$$B_1(u) = B_2(u) = 0; \quad \text{i.e.} \quad u(a) = u(b) = 0. \tag{8.77}$$

For RHS of (8.69) to vanish, it suffices to define $B_1^\dagger(u)$ and $B_2^\dagger(u)$ such that

$$B_1^\dagger(u) = v^*(a) \quad \text{and} \quad B_2^\dagger(u) = v^*(b). \tag{8.78}$$

Then, homogeneous adjoint BCs read as

$$v^*(a) = v^*(b) = 0 \quad i.e. \quad v(a) = v(b) = 0. \tag{8.79}$$

In this manner, we can readily construct the homogeneous adjoint BCs the same as those of (8.77) so that $L_x$ can be Hermitian.

We list several prescriptions of typical BCs below.

(i) $u(a) = u(b) = 0$ (Dirichlet conditions),

(ii) $\left.\frac{\mathrm{d}u}{\mathrm{d}x}\right|_{x=a} = \left.\frac{\mathrm{d}u}{\mathrm{d}x}\right|_{x=b} = 0$ (Neumann conditions),

(iii) $u(a) = u(b)$ and $\left.\frac{\mathrm{d}u}{\mathrm{d}x}\right|_{x=a} = \left.\frac{\mathrm{d}u}{\mathrm{d}x}\right|_{x=b}$ (periodicconditions). (8.80)

Yet, care should be taken when handling RHS of (8.69); i.e., the surface terms. It is because conditions (i) to (iii) are not necessary but sufficient conditions for the surface terms to vanish. Such conditions are not limited to them. Meanwhile, we often have to deal with the nonvanishing surface terms. In that case, we have to start with (8.62) instead of (8.69).

In Sect. 8.2, we mentioned the definition of Hermiticity of the differential operator in such a way that the said operator is self-adjoint and that homogeneous BCs and homogeneous adjoint BCs are the same. In light of the above argument, however, we may relax the conditions for a differential operator to be Hermitian. This is particularly the case when $p(x) = a(x)w(x)$ in (8.69) vanishes at both the endpoints. We will encounter such a situation in Sect. 8.7.

## 8.4 Green's Functions

Having aforementioned discussions, let us proceed with studies of Green's functions for SOLDEs. Though minimum, we have to mention a bit of formalism.

Given $L_x$ defined by (8.55), let us assume

$$L_x u(x) = d(x) \tag{8.81}$$

under homogeneous BCs with an inhomogeneous term $d(x)$ being an arbitrary function. We also assume that (8.81) is well defined in a domain $[r, s]$. The numbers $r$ and $s$ can be infinity. Suppose simultaneously that we have

$$L_x^\dagger v(x) = h(x) \tag{8.82}$$

under homogeneous adjoint BCs [1, 2] with an inhomogeneous term $h(x)$ being an arbitrary function as well.

Let us describe the above relations as

$$L|u\rangle = |d\rangle \quad \text{and} \quad L^\dagger|v\rangle = |h\rangle. \tag{8.83}$$

Suppose that there is an inverse operator $L^{-1} \equiv G$ such that

$$GL = LG = E, \tag{8.84}$$

where $E$ is an identity operator. Operating $G$ on (8.83), we have

$$GL|u\rangle = E|u\rangle = |u\rangle = G|d\rangle. \tag{8.85}$$

This implies that (8.81) has been solved and the solution is given by $G|d\rangle$. Since an inverse operation to differentiation is integration, $G$ is expected to be an integral operator.

We have

$$\langle x|LG|y\rangle = L_x\langle x|G|y\rangle = L_x G(x, y). \tag{8.86}$$

Meanwhile, using (8.84), we get

$$\langle x|LG|y\rangle = \langle x|E|y\rangle = \langle x|y\rangle. \tag{8.87}$$

Using a weight function $w(x)$, we generalize an inner product of (1.128) such that

$$\langle g|f\rangle \equiv \int_r^s w(x) g(x)^* f(x) \mathrm{d}x. \tag{8.88}$$

As we expand an arbitrary vector using basis vectors, we "expand" an arbitrary function $|f\rangle$ using basis vectors $|x\rangle$. Here, we are treating real numbers as if they formed continuous innumerable basis vectors on a real number line (see Fig. 8.1). Thus, we could expand $|f\rangle$ in terms of $|x\rangle$ such that

$$|f\rangle = \int_r^s \mathrm{d}x w(x) f(x) |x\rangle. \tag{8.89}$$

In (8.89), we considered $f(x)$ as if it were an expansion coefficient. The following notation would be reasonable accordingly:

$$f(x) \equiv \langle x|f\rangle. \tag{8.90}$$

In (8.90), $f(x)$ can be viewed as coordinate representation of $|f\rangle$. Thus, from (8.89), we get

$$\langle x'|f\rangle = f(x') = \int_r^s \mathrm{d}x w(x) f(x) \langle x'|x\rangle. \tag{8.91}$$

Alternatively, we have

$$f(x') = \int_r^s \mathrm{d}x f(x) \delta(x - x'). \tag{8.92}$$

This comes from a property of the $\delta$ function [1] described as

$$\int_r^s \mathrm{d}x f(x) \delta(x) = f(0). \tag{8.93}$$

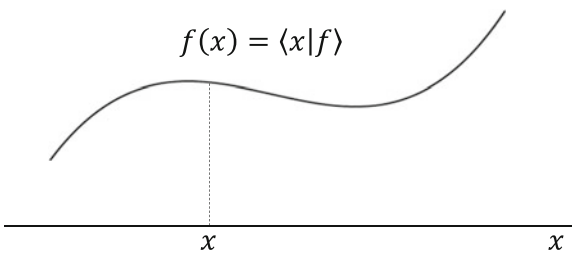


**Fig. 8.1** Function $|f\rangle$ and its coordinate representation $f(x)$

Comparing (8.91) and (8.92), we have

$$w(x)\langle x'|x\rangle = \delta(x-x') \quad \text{or} \quad \langle x'|x\rangle = \frac{\delta(x-x')}{w(x)} = \frac{\delta(x'-x)}{w(x)}. \tag{8.94}$$

Thus comparing (8.86) and (8.87) and using (8.94), we get

$$L_x G(x,y) = \frac{\delta(x-y)}{w(x)}. \tag{8.95}$$

In a similar manner, we also have

$$L_x^\dagger g(x,y) = \frac{\delta(x-y)}{w(x)}. \tag{8.96}$$

To arrive at (8.96), we start the discussion assuming an operator $(L_x^\dagger)^{-1}$ such that $(L_x^\dagger)^{-1} \equiv g$ with $gL^\dagger = L^\dagger g = E$.

The function $G(x,y)$ is called a Green's function and $g(x,y)$ is said to be an adjoint Green's function. Handling of Green's functions and adjoint Green's functions is based upon (8.95) and (8.96), respectively. As (8.81) is defined in a domain $r \le x \le s$, (8.95) and (8.96) are defined in a domain $r \le x \le s$ and $r \le y \le s$. Notice that except for the point $x = y$ we have

$$L_x G(x,y) = 0 \quad \text{and} \quad L_x^\dagger g(x,y) = 0. \tag{8.97}$$

That is, $G(x,y)$ and $g(x,y)$ satisfy the homogeneous equation with respect to the variable $x$. Accordingly, we require $G(x,y)$ and $g(x,y)$ to satisfy the same homogeneous BCs with respect to the variable $x$ as those imposed upon $u(x)$ and $v(x)$ of (8.81) and (8.82), respectively [1].

The relation (8.88) can be obtained as follows: Operating $\langle g|$ on (8.89), we have

$$\langle g|f\rangle = \int_r^s \mathrm{d}x w(x) f(x) \langle g|x\rangle = \int_r^s w(x) g(x)^* f(x) \mathrm{d}x, \tag{8.98}$$

where for the last equality we used

$$\langle g|x\rangle = \langle x|g\rangle^* = g(x)^*. \tag{8.99}$$

For this, see (1.113) where $A$ is replaced with an identity operator $E$ with regard to a complex conjugate of an inner product of two vectors. Also see (11.2) of Sect. 11.1.

If in (8.69) the surface term (i.e., RHS) vanishes under appropriate conditions, e.g., (8.80), we have

$$\int_r^s \mathrm{d}x w(x)\{v^*(L_x u) - [L_x^\dagger v]^* u\} = 0, \tag{8.100}$$

which is called Green's identity. Since (8.100) is derived from identities (8.56), (8.100) is an identity as well (as a terminology of Green's identity shows). Therefore, (8.100) must hold with any functions $u$ and $v$ so far as they satisfy homogeneous BCs. Thus, replacing $v$ in (8.100) with $g(x,y)$ and using (8.96) together with (8.81), we have

$$\begin{aligned}
&\int_r^s \mathrm{d}x w(x)\left\{g^*(x,y)[L_x u(x)] - [L_x^\dagger g(x,y)]^* u(x)\right\} \\
&= \int_r^s \mathrm{d}x w(x)\left\{g^*(x,y)d(x) - \left[\frac{\delta(x-y)}{w(x)}\right]^* u(x)\right\} \\
&= \int_r^s dx w(x) g^*(x,y) d(x) - u(y) = 0,
\end{aligned} \tag{8.101}$$

where with the second last equality we used a property of the $\delta$ functions. Also notice that $\frac{\delta(x-y)}{w(x)}$ is a real function. Rewriting (8.101), we get

$$u(y) = \int_r^s \mathrm{d}x w(x) g^*(x,y) d(x). \tag{8.102}$$

Similarly, replacing $u$ in (8.100) with $G(x,y)$ and using (8.95) together with (8.82), we have

$$v(y) = \int_r^s \mathrm{d}x w(x) G^*(x,y) h(x). \tag{8.103}$$

Replacing $u$ and $v$ in (8.100) with $G(x,q)$ and $g(x,t)$, respectively, we have

$$\int_r^s \mathrm{d}x w(x)\left\{g^*(x,t)[L_x G(x,q)] - \left[L_x^\dagger g(x,t)\right]^* G(x,q)\right\} = 0 \tag{8.104}$$

Notice that we have chosen $q$ and $t$ for the second argument $y$ in (8.95) and (8.96), respectively. Inserting (8.95) and (8.96) into the above equation after changing arguments, we have

$$\int_r^s \mathrm{d}x w(x)\left\{g^*(x,t)\frac{\delta(x-q)}{w(x)} - \left[\frac{\delta(x-t)}{w(x)}\right]^* G(x,q)\right\} = 0. \tag{8.105}$$

Thus, we get

$$g^*(q,t) = G(t,q) \quad \text{or} \quad g(q,t) = G^*(t,q). \tag{8.106}$$

This implies that $G^*(t,q)$ must satisfy the adjoint BCs with respect to the second argument $q$. Inserting (8.106) into (8.102), we get

$$u(y) = \int_r^s \mathrm{d}x w(x) G(y,x) d(x). \tag{8.107}$$

Or exchanging the arguments $x$ and $y$, we have

$$u(x) = \int_r^s \mathrm{d}x w(y) G(x,y) d(y). \tag{8.108}$$

Similarly, using (8.103) into (8.106), we get

$$v(y) = \int_r^s \mathrm{d}x w(x) g(y,x) h(x). \tag{8.109}$$

Or, we have

$$v(x) = \int_r^s \mathrm{d}y w(y) g(x,y) h(y). \tag{8.110}$$

Equations (8.107–8.110) clearly show that homogeneous equations [given by putting $\mathrm{d}(x) = h(x) = 0$] have a trivial solution $u(x) \equiv 0$ and $v(x) \equiv 0$ under homogeneous BCs. Note that it is always the case when we are able to construct a Green's function. This in turn implies that we can construct a Green's function if the differential operator is accompanied by *initial conditions*. Conversely, if the homogeneous equation has a non-trivial solution under homogeneous BCs, Eqs. (8.107–8.110) will not work.

If the differential operator $L$ in (8.81) is Hermitian, according to the associated remarks of Sect. 8.2, we must have $L_x = L_x^\dagger$ and $u(x)$ and $v(x)$ of (8.81) and (8.82) must satisfy the same homogeneous BCs. Consequently, in the case of an Hermitian operator, we should have

$$G(x,y) = g(x,y). \tag{8.111}$$

From (8.106) and (8.111), if the operator is Hermitian, we get

$$G(x,y) = G^*(y,x). \tag{8.112}$$

In Sect. 8.3, we assume that the coefficients $a(x)$, and $b(x)$, and $c(x)$ are real to assure that $L_x$ is Hermitian [1]. On this condition, $G(x,y)$ is real as well (vide infra). Then, we have

$$G(x,y) = G(y,x). \tag{8.113}$$

That is, $G(x,y)$ is real symmetric with respect to the arguments $x$ and $y$.

To be able to apply Green's functions to practical use, we will have to estimate a behavior of the Green's function near $x = y$. This is because in light of (8.95) and (8.96), there is a "jump" at $x = y$.

When we deal with a case where a self-adjoint operator is relevant, using a function $p(x)$ of (8.69), we have

$$\frac{a(x)}{p(x)}\frac{\partial}{\partial x}\left(p\frac{\partial G}{\partial x}\right) + c(x)G = \frac{\delta(x-y)}{w(x)}. \tag{8.114}$$

Multiplying both sides by $\frac{p(x)}{a(x)}$, we have

$$\frac{\partial}{\partial x}\left(p\frac{\partial G}{\partial x}\right) = \frac{p(x)}{a(x)}\frac{\delta(x-y)}{w(x)} - \frac{p(x)c(x)}{a(x)}G(x,y). \tag{8.115}$$

Using a property of the $\delta$ function expressed by

$$f(x)\delta(x) = f(0)\delta(x) \quad \text{or} \quad f(x)\delta(x-y) = f(y)\delta(x-y) \tag{8.116}$$

we have

$$\frac{\partial}{\partial x}\left(p\frac{\partial G}{\partial x}\right) = \frac{p(y)}{a(y)}\frac{\delta(x-y)}{w(y)} - \frac{p(x)c(x)}{a(x)}G(x,y). \tag{8.117}$$

Integrating (8.117) with respect to $x$, we get

$$p\frac{\partial G(x,y)}{\partial x} = \frac{p(y)}{a(y)w(y)}\theta(x-y) - \int_r^x \mathrm{d}t\frac{p(t)c(t)}{a(t)}G(t,y) + C, \tag{8.118}$$

where $C$ is a constant. The function $\theta(x-y)$ is defined by

$$\theta(x) = \begin{cases} 1\,(x > 0) \\ 0\,(x < 0). \end{cases} \tag{8.119}$$

Note that we have

$$\frac{\mathrm{d}\theta(x)}{\mathrm{d}x} = \delta(x). \tag{8.120}$$

In RHS of (8.118), the first term has a discontinuity at $x = y$ because of $\theta(x - y)$, whereas the second term is continuous with respect to $y$. Thus, we have

$$\begin{aligned} &\lim_{\varepsilon \to +0} \left[ p(y+\varepsilon) \frac{\partial G(x,y)}{\partial x}\bigg|_{x=y+\varepsilon} - p(y-\varepsilon) \frac{\partial G(x,y)}{\partial x}\bigg|_{x=y-\varepsilon} \right] \\ &\quad = \lim_{\varepsilon \to +0} \frac{p(y)}{a(y)w(y)} [\theta(+\varepsilon) - \theta(-\varepsilon)] = \frac{p(y)}{a(y)w(y)}. \end{aligned} \tag{8.121}$$

Since $p(y)$ is continuous with respect to the argument $y$, this factor drops off and we get

$$\lim_{\varepsilon \to +0} \left[ \frac{\partial G(x,y)}{\partial x}\bigg|_{x=y+\varepsilon} - \frac{\partial G(x,y)}{\partial x}\bigg|_{x=y-\varepsilon} \right] = \frac{1}{a(y)w(y)}. \tag{8.122}$$

Thus, $\frac{\partial G(x,y)}{\partial x}$ is accompanied by a discontinuity at $x = y$ by a magnitude of $\frac{1}{a(y)w(y)}$. Since RHS of (8.122) is continuous with respect to the argument $y$, integrating (8.122) again with respect to $x$, we find that $G(x, y)$ is continuous at $x = y$. These properties of $G(x, y)$ are useful to calculate Green's functions in practical use. We will encounter several examples in next sections.

Suppose that there are two Green's functions that satisfy the same homogeneous BCs. Let $G(x, y)$ and $\tilde{G}(x, y)$ be such functions. Then, we must have

$$L_x G(x,y) = \frac{\delta(x-y)}{w(x)} \quad \text{and} \quad L_x \tilde{G}(x,y) = \frac{\delta(x-y)}{w(x)}. \tag{8.123}$$

Subtracting both sides of (8.123), we have

$$L_x \left[ \tilde{G}(x,y) - G(x,y) \right] = 0. \tag{8.124}$$

In virtue of the linearity of BCs, $G(x, y) - \tilde{G}(x, y)$ must satisfy the same homogeneous BCs as well. But, (8.124) is a homogeneous equation, and so we must have a trivial solution from the aforementioned constructability of the Green's function such that

$$G(x,y) - \tilde{G}(x,y) \equiv 0 \quad \text{or} \quad G(x,y) \equiv \tilde{G}(x,y). \tag{8.125}$$

This obviously indicates that a Green's function should be unique.

We have assumed in Sect. 8.3 that the coefficients $a(x)$, and $b(x)$, and $c(x)$ are real. Therefore, taking complex conjugate of (8.95), we have

$$L_x G(x,y)^* = \frac{\delta(x-y)}{w(x)}. \tag{8.126}$$

Notice here that both $\delta(x-y)$ and $w(x)$ are real functions. Subtracting (8.95) from (8.126), we have

$$L_x[G(x,y)^* - G(x,y)] = 0.$$

Again, from the uniqueness of the Green's function, we get $G(x,y)^* = G(x,y)$; i.e., $G(x,y)$ is real accordingly. This is independent of specific structures of $L_x$. In other words, so far as we are dealing with real coefficients $a(x)$, $b(x)$, and $c(x)$, $G(x,y)$ is real whether or not $L_x$ is self-adjoint.

## 8.5 Construction of Green's Functions

So far we dealt with homogeneous boundary conditions (BCs) with respect to a differential equation

$$a(x)\frac{\mathrm{d}^2 u}{\mathrm{d}x^2} + b(x)\frac{\mathrm{d}u}{\mathrm{d}x} + c(x)u = d(x), \tag{8.2}$$

where coefficients $a(x)$, $b(x)$, and $c(x)$ are real. In this case, if $d(x) \equiv 0$ in (8.108), namely the SOLDE is homogeneous equation, we have a solution $u(x) \equiv 0$ on the basis of (8.108). If, on the other hand, we have inhomogeneous boundary conditions (BCs), additional terms appear on RHS of (8.108) in both the cases of homogeneous and inhomogeneous equations. In this section, we examine how we can deal with this problem.

Following the remarks made in Sect. 8.3, we start with (8.62) or (8.69). If we deal with a self-adjoint or Hermitian operator, we can apply (8.69) to the problem. In a more general case where the operator is not self-adjoint, (8.62) is useful. In this respect, in Sect. 8.6, we have a good opportunity for this.

In Sect. 8.3, we mentioned that we may relax the definition of Hermiticity of the differential operator in the case where the surface term vanishes. Meanwhile, we should bear in mind that the Green's functions and adjoint Green's functions are constructed using homogeneous BCs regardless of whether we are concerned with a homogeneous equation or inhomogeneous equation. Thus, even if the surface terms do not vanish, we may regard the differential operator as Hermitian. This is because

we deal with essentially the same Green's function to solve a problem with both the cases of homogeneous equation and inhomogeneous equations (vide infra). Notice also that whether or not RHS vanishes, we are to use the same Green's function [1]. In this sense, we do not have to be too strict with the definition of Hermiticity.

Now, suppose that with an Hermitian operator $L_x$ we are given

$$\int_r^s \mathrm{d}x w(x)\left\{v^*(L_x u) - [L_x^\dagger v]^* u\right\} = \left[p(x)\left(v^*\frac{\mathrm{d}u}{\mathrm{d}x} - u\frac{\mathrm{d}v^*}{\mathrm{d}x}\right)\right]_r^s, \tag{8.69}$$

where $w(x) > 0$ in a domain $[r, s]$ and $p(x)$ is a real function. Note that since (8.69) is an identity by an appropriate choice of $w(x)$, we insert $g(x, y)$ into $v$. Then from (8.97) and (8.106), we have

$$\int_r^s dx w(x)\left\{G(y,x)d(x) - \left[\frac{\delta(x-y)}{w(x)}\right]^* u\right\}$$

$$= \left[p(x)\left\{G(y,x)\frac{\mathrm{d}u(x)}{\mathrm{d}x} - u(x)\frac{\partial G(y,x)}{\partial x}\right\}\right]_{x=r}^s. \tag{8.127}$$

Using a property of the $\delta$ function, we get

$$u(y) = \int_r^s \mathrm{d}x w(x)G(y,x)d(x) - \left[p(x)\left\{G(y,x)\frac{\mathrm{d}u(x)}{\mathrm{d}x} - u(x)\frac{\partial G(y,x)}{\partial x}\right\}\right]_{x=r}^s. \tag{8.128}$$

The differential operator is Hermitian according to the discussion of Sect. 8.3. Hence, we have

$$G(x, y) = G(y, x). \tag{8.113}$$

The function $G(x, y)$ satisfies homogeneous BCs. Hence, if we assume, e.g., the Dirichlet BCs, we have

$$G(r, y) = G(s, y) = 0. \tag{8.129}$$

Since $G(x, y)$ is symmetric with respect to arguments $x$ and $y$, from (8.129) we get

$$G(y, r) = G(y, s) = 0. \tag{8.130}$$

Thus, the first term of the surface terms of (8.128) is eliminated to yield

$$u(y) = \int_r^s \mathrm{d}x w(x) G(y,x) d(x) + \left[ p(x) u(x) \left( \frac{\partial G(y,x)}{\partial x} \right) \right]_{x=r}^{s}.$$

Exchanging the arguments $x$ and $y$, we get

$$u(x) = \int_r^s \mathrm{d}y w(y) G(x,y) d(y) + \left[ p(y) u(y) \left( \frac{\partial G(x,y)}{\partial y} \right) \right]_{y=r}^{s}. \tag{8.131}$$

Then, (i) substituting surface terms of $u(s)$ and $u(r)$ that are associated with the inhomogeneous BCs described as

$$B_1(u) = \sigma_1 \quad \text{and} \quad B_2(u) = \sigma_2. \tag{8.132}$$

and (ii) calculating $\frac{\partial G(x,y)}{\partial y}|_{y=r}$ and $\frac{\partial G(x,y)}{\partial y}|_{y=s}$, we will be able to obtain a unique solution. Once again, notice that (8.131) is used for a differential equation rendered self-adjoint by means of a weight function $w(x)$.

On the basis of the general discussion of Sect. 8.4 and this section, we are in the position to construct the Green's functions. Except for the points of $x = y$, the Green's function $G(x, y)$ must satisfies the following differential equation:

$$L_x G(x,y) = 0, \tag{8.133}$$

where $L_x$ is given by

$$L_x = a(x) \frac{\mathrm{d}^2}{\mathrm{d}x^2} + b(x) \frac{\mathrm{d}}{\mathrm{d}x} + c(x). \tag{8.55}$$

The differential equation $L_x u = d(x)$ is defined within an interval $[r, s]$, where $r$ may be $-\infty$ and $s$ may be $+\infty$.

From now on, we regard $a(x)$, $b(x)$, and $c(x)$ as real functions. From (8.133), we expect the Green's function to be described as a linear combination of a fundamental set of solutions $u_1(x)$ and $u_2(x)$. Here the set of fundamental solutions is given by two linearly independent solutions of a homogeneous equation $L_x u = 0$. Then, we should be able to express $G(x, y)$ as a combination of $F_1(x, y)$ and $F_2(x, y)$ that are described as

$$\begin{aligned} F_1(x,y) &= c_1 u_1(x) + c_2 u_2(x) \quad \text{for} \quad r \le x < y, \\ F_2(x,y) &= d_1 u_1(x) + d_2 u_2(x) \quad \text{for} \quad y < x \le s, \end{aligned} \tag{8.134}$$

where $c_1, c_2, d_1,$ and $d_2$ are arbitrary (complex) constants to be determined later. These constants are given as a function of $y$. The combination has to be made such that

$$G(x,y) = \begin{cases} F_1(x,y) \text{ for } r \le x < y, \\ F_2(x,y) \text{ for } y < x \le s. \end{cases} \tag{8.135}$$

Thus using $\theta(x)$ function defined as (8.119), we describe $G(x,y)$ as

$$G(x,y) = F_1(x,y)\theta(y-x) + F_2(x,y)\theta(x-y). \tag{8.136}$$

Notice that $F_1(x,y)$ and $F_2(x,y)$ are "ordinary" functions and that $G(x,y)$ is not, because $G(x,y)$ contains the $\theta(x)$ function.

If we have

$$F_2(x,y) = F_1(y,x), \tag{8.137}$$

$$G(x,y) = F_1(x,y)\theta(y-x) + F_1(y,x)\theta(x-y). \tag{8.138}$$

Hence, we get

$$G(x,y) = G(y,x). \tag{8.139}$$

From (8.113), $L_x$ is Hermitian. Suppose that $F_1(x,y) = (x-r)(y-s)$ and $F_2(x,y) = (x-s)(y-r)$. Then, (8.137) is satisfied and, hence, if we can construct the Green's function from $F_1(x,y)$ and $F_2(x,y)$, $L_x$ should be Hermitian. However, if we had, e.g., $F_1(x,y) = x - r$ and $F_2(x,y) = y - s$, $G(x,y) \neq G(y,x)$, and so $L_x$ would not be Hermitian.

The Green's functions must satisfy the homogeneous BCs. That is,

$$B_1(G) = B_2(G) = 0. \tag{8.140}$$

Also, we require continuity condition of $G(x,y)$ at $x = y$ and discontinuity condition of $\frac{\partial G(x,y)}{\partial x}$ at $x = y$ described by (8.122). Thus, we have four conditions including BCs and continuity and discontinuity conditions to be satisfied by $G(x,y)$. Thus, we can determine four constants $c_1, c_2, d_1,$ and $d_2$ by the four conditions.

Now, let us inspect further details about the Green's functions by an example.

*Example 8.4* Let us consider a following differential equation

$$\frac{\mathrm{d}^2 u}{\mathrm{d}x^2} + u = 1. \tag{8.141}$$

We assume that a domain of the argument $x$ is $[0, L]$. We set boundary conditions such that

$$u(0) = \sigma_1 \quad \text{and} \quad u(L) = \sigma_2. \tag{8.142}$$

Thus, if at least one of $\sigma_1$ and $\sigma_2$ is not zero, we are dealing with an inhomogeneous differential equation under inhomogeneous BCs.

Next, let us seek conditions that the Green's function satisfies. We also seek a fundamental set of solutions of a homogeneous equation described by

$$\frac{\mathrm{d}^2 u}{\mathrm{d}x^2} + u = 0. \tag{8.143}$$

This is obtained by putting $a = c = 1$ and $b = 0$ in a general form of (8.5) with a weight function being unity. The differential equation (8.143) is therefore self-adjoint according to the argument of Sect. 8.3. A fundamental set of solutions are given by

$\mathrm{e}^{ix}$ and $\mathrm{e}^{-ix}$.

Then, we have

$$\begin{aligned} F_1(x,y) &= c_1\mathrm{e}^{ix} + c_2\mathrm{e}^{-ix} \quad \text{for} \quad 0 \le x < y. \\ F_2(x,y) &= d_1\mathrm{e}^{ix} + d_2\mathrm{e}^{-ix} \quad \text{for} \quad y < x \le L. \end{aligned} \tag{8.144}$$

The functions $F_1(x,y)$ and $F_2(x,y)$ must satisfy the following BCs such that

$$F_1(0,y) = c_1 + c_2 = 0 \quad \text{and} \quad F_2(L,y) = d_1\mathrm{e}^{iL} + d_2\mathrm{e}^{-iL} = 0. \tag{8.145}$$

Thus, we have

$$F_1(x,y) = c_1(\mathrm{e}^{ix} - \mathrm{e}^{-ix}), F_2(x,y) = d_1(\mathrm{e}^{ix} - \mathrm{e}^{2iL}\mathrm{e}^{-ix}). \tag{8.146}$$

Therefore, at $x = y$, we have

$$c_1(\mathrm{e}^{iy} - \mathrm{e}^{-iy}) = d_1(\mathrm{e}^{iy} - \mathrm{e}^{2iL}\mathrm{e}^{-iy}). \tag{8.147}$$

Discontinuity condition of (8.122) is equivalent to

$$\frac{\partial F_2(x,y)}{\partial x}\Big|_{x=y} - \frac{\partial F_1(x,y)}{\partial x}\Big|_{x=y} = 1. \tag{8.148}$$

This is because both $F_1(x,y)$ and $F_2(x,y)$ are ordinary functions and supposed to be differentiable at any $x$. The relation (8.148) then reads as

$$id_1(\mathrm{e}^{iy} + \mathrm{e}^{2iL}\mathrm{e}^{-iy}) - ic_1(\mathrm{e}^{iy} + \mathrm{e}^{-iy}) = 1. \tag{8.149}$$

From (8.147) and (8.149), using Cramer's rule, we have

$$c_1 = \frac{\begin{vmatrix} 0 & -e^{iy} + e^{2iL}e^{-iy} \\ i & -e^{iy} - e^{2iL}e^{-iy} \end{vmatrix}}{\begin{vmatrix} e^{iy} - e^{-iy} & -e^{iy} + e^{2iL}e^{-iy} \\ e^{iy} + e^{-iy} & -e^{iy} - e^{2iL}e^{-iy} \end{vmatrix}} = \frac{i(e^{iy} - e^{2iL}e^{-iy})}{2(1 - e^{2iL})}, \tag{8.150}$$

$$d_1 = \frac{\begin{vmatrix} e^{iy} - e^{-iy} & 0 \\ e^{iy} + e^{-iy} & i \end{vmatrix}}{\begin{vmatrix} e^{iy} - e^{-iy} & -e^{iy} + e^{2iL}e^{-iy} \\ e^{iy} + e^{-iy} & -e^{iy} - e^{2iL}e^{-iy} \end{vmatrix}} = \frac{i(e^{iy} - e^{-iy})}{2(1 - e^{2iL})}. \tag{8.151}$$

Substituting these parameters for (8.146), we get

$$F_1(x,y) = \frac{\sin x(e^{2iL}e^{-iy} - e^{iy})}{1 - e^{2iL}}, F_2(x,y) = \frac{\sin y(e^{2iL}e^{-ix} - e^{ix})}{1 - e^{2iL}}. \tag{8.152}$$

Making a denominator real, we have

$$F_1(x,y) = \frac{\sin x[\cos(y - 2L) - \cos y]}{2\sin^2 L}, F_2(x,y) = \frac{\sin y[\cos(x - 2L) - \cos x]}{2\sin^2 L}. \tag{8.153}$$

Using the $\theta(x)$ function, we get

$$G(x,y) = \frac{\sin x[\cos(y - 2L) - \cos y]}{2\sin^2 L}\theta(y - x) + \frac{\sin y[\cos(x - 2L) - \cos x]}{2\sin^2 L}\theta(x - y). \tag{8.154}$$

Thus, $G(x,y) = G(y,x)$ as expected. Notice, however, that if $L = n\pi(n = 1, 2, \cdots)$ the Green's function cannot be defined as an ordinary function even if $x \neq y$. We return to this point later.

The solution for (8.141) under the homogeneous BCs is then described as

$$\begin{aligned} u(x) &= \int_0^L dx G(x,y) \\ &= \frac{\cos(x - 2L) - \cos x}{2\sin^2 L}\int_0^x \sin y dy + \frac{\sin x}{2\sin^2 L}\int_x^L [\cos(y - 2L) - \cos y] dy. \end{aligned} \tag{8.155}$$

This can readily be integrated to yield solution for the inhomogeneous equation such that

$$u(x) = \frac{\cos(x-2L) - \cos x - 2\sin L \sin x - \cos 2L + 1}{2\sin^2 L} = \frac{\cos(x-2L) - \cos x - 2\sin L \sin x + 2\sin^2 L}{2\sin^2 L} \tag{8.156}$$

Next, let us consider the surface term. This is given by the second term of (8.131). We get

$$\frac{\partial F_1(x,y)}{\partial x}|_{y=L} = \frac{\sin x}{\sin L}, \frac{\partial F_2(x,y)}{\partial x}|_{y=0} = \frac{\cos(x-2L) - \cos x}{2\sin^2 L}. \tag{8.157}$$

Therefore, with the inhomogeneous BCs, we have the following solution for the inhomogeneous equation:

$$u(x) = \frac{\cos(x-2L) - \cos x - 2\sin L \sin x + 2\sin^2 L}{2\sin^2 L} + \frac{2\sigma_2 \sin L \sin x + \sigma_1[\cos x - \cos(x-2L)]}{2\sin^2 L}, \tag{8.158}$$

where the second term is the surface term. If $\sigma_1 = \sigma_2 = 1$, we have

$$u(x) \equiv 1.$$

Looking at (8.141), we find that $u(x) \equiv 1$ is certainly a solution for (8.141) with inhomogeneous BCs of $\sigma_1 = \sigma_2 = 1$. The uniqueness of the solution then ensures that $u(x) \equiv 1$ is a sole solution under the said BCs.

From (8.154), we find that $G(x, y)$ has a singularity at $L = n\pi (n : \text{integers})$. This is associated with the fact that a homogenous equation (8.143) has a non-trivial solution, e.g., $u(x) = \sin x$ under homogeneous BCs $u(0) = u(L) = 0$. The present situation is essentially the same as that of Example 1.1 of Sect. 1.3. In other words, when $\lambda = 1$ in (1. 61), the form of a differential equation is identical to (8.143) with virtually the same Dirichlet conditions. The point is that (8.143) can be viewed as a homogeneous equation and, at the same time, as an eigenvalue equation. In such a case, a Green's function approach will fail.

## 8.6 Initial Value Problems (IVPs)

### *8.6.1 General Remarks*

The IVPs are frequently appeared in mathematical physics. The relevant conditions are dealt with as BCs in the theory of differential equations. With boundary functionals $B_1(u)$ and $B_2(u)$ of (8.3) and (8.4), setting $\alpha_1 = \beta_2 = 1$ and other coefficients as zero, we get

$$B_1(u) = u(p) = \sigma_1 \quad \text{and} \quad B_2(u) = \frac{\mathrm{d}u}{\mathrm{d}x}|_{x=p} = \sigma_2. \tag{8.159}$$

In the above, note that we choose $[r, s]$ for a domain of $x$. The points $r$ and $s$ can be infinity as before. Any point $p$ within the domain $[r, s]$ may be designated as a special point on which the BCs (8.159) are imposed. The initial conditions are particularly prominent among BCs. This is because the conditions are set at one point of the argument. This special condition is usually called initial conditions. In this section, we investigate fundamental characteristics of IVPs.

Suppose that we have

$$u(p) = \frac{\mathrm{d}u}{\mathrm{d}x}|_{x=p} = 0, \tag{8.160}$$

with homogeneous BCs. Given a differential operator $L_x$ defined as (8.55), i.e.,

$$L_x = a(x)\frac{\mathrm{d}^2}{\mathrm{d}x^2} + b(x)\frac{\mathrm{d}}{\mathrm{d}x} + c(x), \tag{8.55}$$

let a fundamental set of solutions be $u_1(x)$ and $u_2(x)$ for

$$L_x u(x) = 0. \tag{8.161}$$

A general solution $u(x)$ for (8.161) is given by a linear combination of $u_1(x)$ and $u_2(x)$ such that

$$u(x) = c_1 u_1(x) + c_2 u_2(x), \tag{8.162}$$

where $c_1$ and $c_2$ are arbitrary (complex) constants. Suppose that we have homogeneous BCs expressed by (8.160). Then, we have

$$u(p) = c_1 u_1(p) + c_2 u_2(p) = 0,$$

$$u'(p) = c_1 u_1'(p) + c_2 u_2'(p) = 0.$$

Rewriting it in a matrix form, we have

$$\begin{pmatrix} u_1(p) & u_2(p) \\ u_1'(p) & u_2'(p) \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = 0.$$

Since the matrix represents Wronskian of a fundamental set of solutions $u_1(x)$ and $u_2(x)$, its determinant never vanishes at any point $p$. That is, we have

$$\begin{vmatrix} u_1(p) & u_2(p) \\ u_1'(p) & u_2'(p) \end{vmatrix} \neq 0. \tag{8.163}$$

Then, we necessarily have $c_1 = c_2 = 0$. From (8.162), we have a trivial solution

$$u(x) \equiv 0$$

under the initial conditions as homogeneous BCs. Thus, as already discussed a Green's function can always be constructed for IVPs.

To seek the Green's functions for IVPs, we return back to the generalized Green's identity described as

$$\int_r^s \mathrm{d}x \left[ v^*(L_x u) - [L_x^\dagger v]^* u \right] = \left[ a v^* \frac{\mathrm{d}u}{\mathrm{d}x} - u \frac{\mathrm{d}(a v^*)}{\mathrm{d}x} + b u v^* \right]_r^s. \tag{8.62}$$

For the surface term (RHS) to vanish, for homogeneous BCs, we have, e.g.,

$$u(s) = \frac{\mathrm{d}u}{\mathrm{d}x}|_{x=s} = 0 \quad \text{and} \quad v(r) = \frac{\mathrm{d}v}{\mathrm{d}x}|_{x=r} = 0,$$

for the two sets of BCs adjoint to each other. Obviously, these are not identical simply because the former is determined at $s$ and the latter is determined at a different point $r$. For this reason, the operator $L_x$ is not Hermitian, even though it is formally self-adjoint. In such a case, we would rather use $L_x$ directly than construct a self-adjoint operator because we cannot make the operator Hermitian either way.

Hence, unlike the precedent sections, we do not need a weight function $w(x)$. Or, we may regard $w(x) \equiv 1$. Then, we reconsider the conditions which the Green's functions should satisfy. On the basis of the general consideration of Sect. 8.4, especially (8.86), (8.87), and (8.94), we have [2]

$$L_x G(x, y) = \langle x|y\rangle = \delta(x - y). \tag{8.164}$$

Therefore, we have

$$\frac{\partial^2 G(x,y)}{\partial x^2} + \frac{b(x)}{a(x)} \frac{\partial G(x,y)}{\partial x} + \frac{c(x)}{a(x)} G(x,y) = \frac{\delta(x-y)}{a(x)}.$$

Integrating or integrating by parts the above equation, we get

$$\frac{\partial G(x,y)}{\partial x} + \left[ \frac{b(x)}{a(x)} G(x,y) \right]_{x_0}^x - \int_{x_0}^x \left[ \frac{b(\xi)}{a(\xi)} \right]' G(\xi, y) \mathrm{d}\xi + \int_{x_0}^x \frac{c(\xi)}{a(\xi)} G(\xi, y) \mathrm{d}\xi = \frac{\theta(x-y)}{a(y)}.$$

Noting that the functions other than $\frac{\partial G(x,y)}{\partial x}$ and $\frac{\theta(x-y)}{a(y)}$ are continuous, as before we have

$$\lim_{\varepsilon \to +0} \left[ \left. \frac{\partial G(x,y)}{\partial x} \right|_{x=y+\varepsilon} - \left. \frac{\partial G(x,y)}{\partial x} \right|_{x=y-\varepsilon} \right] = \frac{1}{a(y)}. \tag{8.165}$$

### 8.6.2 Green's Functions for IVPs

From a practical point of view, we may set $r = 0$ in (8.62). Then, we can choose a domain for $[0, s]$ (for $s > 0$) or $[s, 0]$ (for $s < 0$) with (8.2). For simplicity, we use $x$ instead of $s$. We consider two cases of $x > 0$ and $x < 0$.

(i) Case I ($x > 0$): Let $u_1(x)$ and $u_2(x)$ be a fundamental set of solutions. We define $F_1(x, y)$ and $F_2(x, y)$ as before such that

$$\begin{array}{ll} F_1(x,y) = c_1 u_1(x) + c_2 u_2(x) & \text{for} \quad 0 \le x < y, \\ F_2(x,y) = d_1 u_1(x) + d_2 u_2(x) & \text{for} \quad 0 < y < x. \end{array} \tag{8.166}$$

As before, we set

$$G(x,y) = \begin{cases} F_1(x,y) \text{ for } 0 \le x < y, \\ F_2(x,y) \text{ for } 0 < y < x. \end{cases}$$

Homogeneous BCs are defined as

$$u(0) = 0 \quad \text{and} \quad u'(0) = 0.$$

Correspondingly, we have

$$F_1(0,y) = 0 \quad \text{and} \quad F_1'(0,y) = 0.$$

This is translated into

$$c_1 u_1(0) + c_2 u_2(0) = 0 \quad \text{and} \quad c_1 u_1'(0) + c_2 u_2'(0) = 0.$$

In a matrix form, we get

$$\begin{pmatrix} u_1(0) & u_2(0) \\ u_1'(0) & u_2'(0) \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = 0.$$

As mentioned above, since $u_1(x)$ and $u_2(x)$ are a fundamental set of solutions, we have $c_1 = c_2 = 0$. Hence, we get

$$F_1(x,y) = 0.$$

From the continuity and discontinuity conditions (8.165) imposed upon the Green's functions, we have

$$d_1 u_1(y) + d_2 u_2(y) = 0 \quad \text{and} \quad d_1 u_1'(y) + d_2 u_2'(y) = 1/a(y). \tag{8.167}$$

As before, we get

$$d_1 = -\frac{u_2(y)}{a(y)W(u_1(y), u_2(y))} \quad \text{and} \quad d_2 = \frac{u_1(y)}{a(y)W(u_1(y), u_2(y))}.$$

where $W(u_1(y), u_2(y))$ is Wronskian of $u_1(y)$ and $u_2(y)$. Thus, we get

$$F_2(x, y) = \frac{u_2(x)u_1(y) - u_1(x)u_2(y)}{a(y)W(u_1(y), u_2(y))}. \tag{8.168}$$

(ii) Case II $(x<0)$: Next, we think of the case as below:

$$\begin{array}{ll} F_1(x, y) = c_1 u_1(x) + c_2 u_2(x) & \text{for} \quad y<x\leq 0, \\ F_2(x, y) = d_1 u_1(x) + d_2 u_2(x) & \text{for} \quad x<y<0. \end{array} \tag{8.169}$$

Similarly proceeding as the above, we have $c_1 = c_2 = 0$. Also, we get

$$F_2(x, y) = \frac{u_1(x)u_2(y) - u_2(x)u_1(y)}{a(y)W(u_1(y), u_2(y))}. \tag{8.170}$$

Here notice that the sign is reversed in (8.170) relative to (8.168). This is because on the discontinuity condition, instead of (8.167) we have to have

$$d_1 u_1'(y) + d_2 u_2'(y) = -1/a(y).$$

This results from the fact that magnitude relationship between the arguments $x$ and $y$ has been reversed in (8.169) relative to (8.166).

Summarizing the above argument, (8.168) is obtained in the domain $0\leq y<x$; (8.170) is obtained in the domain $x<y<0$. Noting this characteristic, we define a function such that

$$\Theta(x, y) \equiv \theta(x-y)\theta(y) - \theta(y-x)\theta(-y). \tag{8.171}$$

Notice that

$$\Theta(x, y) = -\Theta(-x, -y).$$

That is, $\Theta(x, y)$ is antisymmetric with respect to the origin.

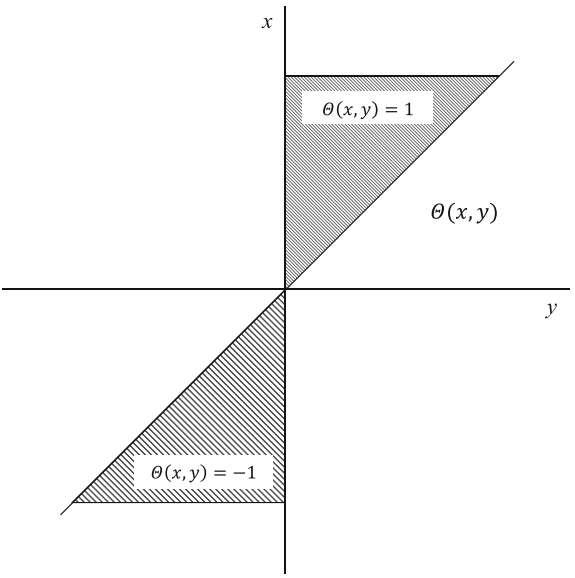


**Fig. 8.2** Graph of a function $\Theta(x, y)$. $\Theta(x, y) = 1$ or $-1$ in hatched areas, otherwise $\Theta(x, y) = 0$

Figure 8.2 shows a feature of $\Theta(x, y)$. If the "initial point" is taken at $x = a$, we can use $\Theta(x - a, y - a)$ instead; see Fig. 8.3. The function is described as

$$\Theta(x - a, y - a) = \theta(x - y)\theta(y - a) - \theta(y - x)\theta(a - y).$$

Note that $\Theta(x - a, y - a)$ can be obtained by shifting $\Theta(x, y)$ toward the positive direction of the $x$- and $y$-axes by $a$ ($a$ can be either positive or negative; in Fig. 8.3, we assume $a > 0$). Using the $\Theta(x, y)$ function, the Green's function is described as

$$G(x, y) = \frac{u_2(x)u_1(y) - u_1(x)u_2(y)}{a(y)W(u_1(y), u_2(y))}\Theta(x, y). \tag{8.172}$$

Defining a function $\mathcal{F}$ such that

$$\mathcal{F}(x, y) \equiv \frac{u_2(x)u_1(y) - u_1(x)u_2(y)}{a(y)W(u_1(y), u_2(y))}, \tag{8.173}$$

we have

$$G(x, y) = \mathcal{F}(x, y)\Theta(x, y). \tag{8.174}$$

Notice that

$$\Theta(x, y) \neq \Theta(y, x) \quad \text{and} \quad G(x, y) \neq G(y, x). \tag{8.175}$$

It is therefore obvious that the differential operator is not Hermitian.

### 8.6.3 Estimation of Surface Terms

To include the surface term of the inhomogeneous case, we use (8.62).

$$\int_r^s \mathrm{d}x\left[v^*(L_x u) - [L_x^\dagger v]^* u\right] = \left[av^* \frac{\mathrm{d}u}{\mathrm{d}x} - u\frac{\mathrm{d}(av^*)}{\mathrm{d}x} + buv^*\right]_r^s. \tag{8.62}$$

As before, we set $r = 0$ in (8.62). Also, we classify (8.62) into two cases according as $s > 0$ or $s < 0$.

(i) Case I $(x > 0, y > 0)$: Equation (8.62) reads as

$$\int_0^\infty \mathrm{d}x\left[v^*(L_x u) - [L_x^\dagger v]^* u\right] = \left[av^* \frac{\mathrm{d}u}{\mathrm{d}x} - u\frac{\mathrm{d}(av^*)}{\mathrm{d}x} + buv^*\right]_0^\infty. \tag{8.176}$$

As before, inserting

$$g(x, y) = G^*(y, x) = G(y, x) \tag{8.177}$$

into $v$ of (8.176) and arranging terms, we have

$$u(y) = \int_0^\infty \mathrm{d}x G(y, x) d(x) - \left[a(x)G(y, x)\frac{\mathrm{d}u(x)}{\mathrm{d}x} - u(x)\frac{\mathrm{d}a(x)}{\mathrm{d}x}G(y, x) - u(x)a(x)\frac{\partial G(y, x)}{\partial x} + b(x)u(x)G(y, x)\right]_{x=0}^\infty. \tag{8.178}$$

Note that in the above we used $L_x^\dagger g(x, y) = \delta(x - y)$. In Fig. 8.4, we depict a domain of $G(y, x)$ in which $G(y, x)$ does not vanish. Notice that we get the domain by folding back that of $\Theta(x, y)$ (see Fig. 8.2) relative to a straight line $y = x$. Thus, we find that $G(y, x)$ vanishes at $x > y$. So does $\frac{\partial G(y,x)}{\partial x}$; see Fig. 8.4. Namely, the second term of RHS of (8.178) vanishes at $x = \infty$. In other words, $g(x, y)$ and $G(y, x)$ must satisfy the adjoint BCs; i.e.,

$$g(\infty, y) = G(y, \infty) = 0. \tag{8.179}$$

At the same time, the upper limit of integration range of (8.178) can be set at $y$. Noting the above, we have

(the first term) of (8.178)

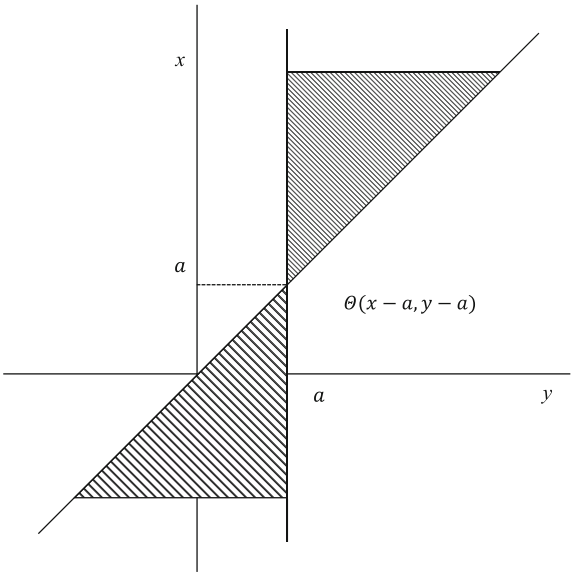


**Fig. 8.3** Graph of a function $\Theta(x-a,y-a)$. We assume $a>0$

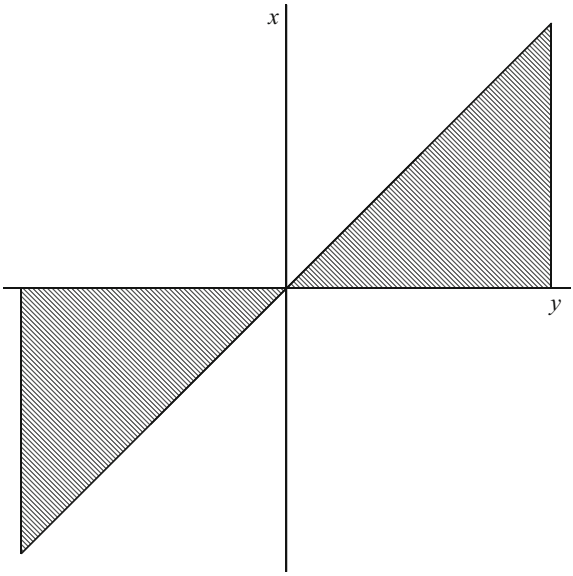


**Fig. 8.4** Domain of $G(y,x)$. Areas for which $G(y,x)$ does not vanish are hatched

$$= \int_0^y \mathrm{d}x G(y,x) d(x). \tag{8.180}$$

Also with the second term of (8.178), we get
(the second term) of (8.178) =

$$+ \left[ a(x)G(y,x)\frac{\mathrm{d}u(x)}{\mathrm{d}x} - u(x)\frac{\mathrm{d}a(x)}{\mathrm{d}x}G(y,x) - u(x)a(x)\frac{\partial G(y,x)}{\partial x} + b(x)u(x)G(y,x) \right]_{x=0}$$
$$= a(0)G(y,0)\frac{du(0)}{\mathrm{d}x} - u(0)\frac{\mathrm{d}a(0)}{\mathrm{d}x}G(y,0) - u(0)a(0)\frac{\partial G(y,x)}{\partial x}\bigg|_{x=0} + b(0)u(0)G(y,0). \tag{8.181}$$

If we substitute inhomogeneous BCs

$$u(0) = \sigma_1 \quad \text{and} \quad \frac{\mathrm{d}u}{\mathrm{d}x}\bigg|_{x=0} = \sigma_2. \tag{8.182}$$

for (8.181) along with other appropriate values, we should be able to get a unique solution as

$$u(y) = \int_0^y \mathrm{d}x G(y,x) d(x) + \left[ \sigma_2 a(0) - \sigma_1 \frac{\mathrm{d}a(0)}{\mathrm{d}x} + \sigma_1 b(0) \right] G(y,0) - \sigma_1 a(0)\frac{\partial G(y,x)}{\partial x}\bigg|_{x=0} \tag{8.183}$$

Exchanging arguments $x$ and $y$, we get

$$u(x) = \int_0^x \mathrm{d}y G(x,y) d(y) + \left[ \sigma_2 a(0) - \sigma_1 \frac{\mathrm{d}a(0)}{\mathrm{d}y} + \sigma_1 b(0) \right] G(x,0) - \sigma_1 a(0)\frac{\partial G(x,y)}{\partial y}\bigg|_{y=0} \tag{8.184}$$

Here, we consider that $\Theta(x,y) = 1$ in this region and use (8.174). Meanwhile, from (8.174), we have

$$\frac{\partial G(x,y)}{\partial y} = \frac{\partial \mathcal{F}(x,y)}{\partial y}\Theta(x,y) + \frac{\partial \Theta(x,y)}{\partial y}\mathcal{F}(x,y). \tag{8.185}$$

In the second term,

$$\begin{aligned}\frac{\partial \Theta(x,y)}{\partial y} &= \frac{\partial \theta(x-y)}{\partial y}\theta(y) + \theta(x-y)\frac{\partial \theta(y)}{\partial y} - \frac{\partial \theta(y-x)}{\partial y}\theta(-y) - \theta(y-x)\frac{\partial \theta(-y)}{\partial y} \\ &= -\delta(x-y)\theta(y) + \theta(x-y)\delta(y) - \delta(y-x)\theta(-y) + \theta(y-x)\delta(-y) \\ &= -\delta(x-y)[\theta(y)+\theta(-y)] + [\theta(x-y)+\theta(y-x)]\delta(y) \\ &= -\delta(x-y)[\theta(y)+\theta(-y)] + [\theta(x)+\theta(-x)]\delta(y) = -\delta(x-y)+\delta(y),\end{aligned} \tag{8.186}$$

where we used

$$\theta(x) + \theta(-x) \equiv 1 \tag{8.187}$$

as well as

$$f(y)\delta(y) = f(0)\delta(y) \tag{8.188}$$

and

$$\delta(-y) = \delta(y). \tag{8.189}$$

However, the function $-\delta(x-y)+\delta(y)$ is of secondary importance. It is because in (8.184) we may choose $[\varepsilon, x-\varepsilon](\varepsilon > 0)$ for the domain $y$ and put $\varepsilon \to +0$ after the integration and other calculations related to the surface terms. Therefore, $-\delta(x-y)+\delta(y)$ in (8.186) virtually vanishes.

Thus, we can express (8.185) as

$$\frac{\partial G(x,y)}{\partial y} = \frac{\partial \mathcal{F}(x,y)}{\partial y}\Theta(x,y) = \frac{\partial \mathcal{F}(x,y)}{\partial y}.$$

Then, finally we reach

$$\begin{aligned}u(x) = &\int_0^x \mathrm{d}y \mathcal{F}(x,y)d(y) + \left[\sigma_2 a(0) - \sigma_1 \frac{\mathrm{d}a(0)}{\mathrm{d}y} + \sigma_1 b(0)\right]\mathcal{F}(x,0) \\ &- \sigma_1 a(0)\frac{\partial \mathcal{F}(x,y)}{\partial y}|_{y=0}.\end{aligned} \tag{8.190}$$

(ii) Case II $(x<0, y<0)$: Similarly as the above, Eq. (8.62) reads as

$$\begin{aligned}u(y) = &\int_{-\infty}^0 \mathrm{d}x G(y,x)d(x) \\ &- \left[a(x)G(y,x)\frac{\mathrm{d}u(x)}{\mathrm{d}x} - u(x)\frac{\mathrm{d}a(x)}{\mathrm{d}x}G(y,x) - u(x)a(x)\frac{\partial G(y,x)}{\partial x} + b(x)u(x)G(y,x)\right]_{-\infty}^0\end{aligned}$$

Similarly as mentioned above, the lower limit of integration range is $y$. Considering both $G(y,x)$ and $\frac{\partial G(y,x)}{\partial x}$ vanish at $x<y$ (see Fig. 8.4), we have

$$\begin{aligned} u(y) &= \int_y^0 \mathrm{d}x G(y,x) d(x) - \left[ a(x) G(y,x) \frac{\mathrm{d}u(x)}{\mathrm{d}x} - u(x) \frac{\mathrm{d}a(x)}{\mathrm{d}x} G(y,x) \right. \\ &\quad \left. - u(x) a(x) \frac{\partial G(y,x)}{\partial x} + b(x) u(x) G(y,x) \right]_{x=0} \\ &= - \int_0^y \mathrm{d}x G(y,x) d(x) - \left[ \sigma_2 a(0) - \sigma_1 \frac{\mathrm{d}a(0)}{\mathrm{d}x} + \sigma_1 b(0) \right] G(y,0) + \sigma_1 a(0) \left. \frac{\partial G(y,x)}{\partial x} \right|_{x=0} \end{aligned} \tag{8.191}$$

Comparing (8.191) with (8.183), we recognize that the sign of RHS of (8.191) has been reversed relative to RHS of (8.183). This is also the case after exchanging arguments $x$ and $y$. Note, however, $\Theta(x,y) = -1$ in the present case. As a result, two minus signs cancel and (8.191) takes exactly the same expression as (8.183). Proceeding with calculations similarly, for both Cases I and II we arrive at a unified solution represented by (8.190) throughout a domain $(-\infty, +\infty)$.

### 8.6.4 Examples

To deepen understanding of Green's functions, we deal with tangible examples of the IVP below.

*Example 8.5* Let us consider a following inhomogeneous differential equation

$$\frac{\mathrm{d}^2 u}{\mathrm{d}x^2} + u = 1. \tag{8.192}$$

Note that (8.192) is formally the same differential equation of (8.141). We may encounter (8.192) when we are observing a motion of a charged harmonic oscillator that is placed under a static electric field. We assume that a domain of the argument $x$ is a whole range of real numbers. We set boundary conditions such that

$$u(0) = \sigma_1 \quad \text{and} \quad u'(0) = \sigma_2. \tag{8.193}$$

As in the case of Example 8.4, a fundamental set of solutions are given by

$$\mathrm{e}^{ix} \quad \text{and} \quad \mathrm{e}^{-ix}. \tag{8.194}$$

Therefore, following (8.173), we get

$$\mathcal{F}(x,y) = \sin(x-y). \tag{8.195}$$

Also following (8.190), we have

$$\begin{aligned} u(x) &= \int_0^x \mathrm{d}y\ \sin(x-y) + \sigma_2 \sin x - \sigma_1 \cos(x-y)\big|_{y=0} \\ &= 1 - \cos x + \sigma_2 \sin x + \sigma_1 \cos x. \end{aligned} \tag{8.196}$$

In particular, if we choose $\sigma_1 = 1$ and $\sigma_2 = 0$, we have

$$u(x) \equiv 1. \tag{8.197}$$

This also ensures that this is a unique solution under the inhomogeneous BCs described as $\sigma_1 = 1$ and $\sigma_2 = 0$.

**Example 8.6: Damped oscillator** If a harmonic oscillator undergoes friction, the oscillator exerts damped oscillation. Such an oscillator is said to be a damped oscillator. The damped oscillator is often dealt with when we think of bound electrons in a dielectric medium that undergo an effect of a dynamic external field varying with time. This is the case when the electron is placed in an alternating electric field or an electromagnetic wave.

An equation of motion of the damped oscillator is described as

$$m\frac{\mathrm{d}^2 u}{\mathrm{d}t^2} + r\frac{\mathrm{d}u}{\mathrm{d}x} + ku = d(x), \tag{8.198}$$

where $m$ is a mass of an electron; $r$ is a damping constant; $k$ is a spring constant of the damped oscillator. To seek a fundamental set of solutions of a homogeneous equation described as

$$m\frac{\mathrm{d}^2 u}{\mathrm{d}t^2} + r\frac{\mathrm{d}u}{\mathrm{d}x} + ku = 0,$$

putting

$$u = \mathrm{e}^{i\rho t} \tag{8.199}$$

and inserting it to (8.198), we have

$$\left(-\rho^2 + \frac{r}{m} i\rho + \frac{k}{m}\right)\mathrm{e}^{i\rho t} = 0.$$

Since $\mathrm{e}^{i\rho t}$ does not vanish, we have

$$-\rho^2 + \frac{r}{m} i\rho + \frac{k}{m} = 0. \tag{8.200}$$

We call this equation a characteristic quadratic equation. We have three cases for the solution of a quadratic equation of (8.200). Solving (8.200), we get

$$\rho = \frac{ir}{2m} \mp \sqrt{-\frac{r^2}{4m^2} + \frac{k}{m}}. \tag{8.201}$$

(i) Equation (8.201) gives two pure imaginary roots; i.e., $-\frac{r^2}{4m^2} + \frac{k}{m} < 0$ (an over damping). (ii) The equation has double roots; $-\frac{r^2}{4m^2} + \frac{k}{m} = 0$ (a critical damping). (iii) The equation has two complex roots; $-\frac{r^2}{4m^2} + \frac{k}{m} > 0$ (a weak damping). Of these, Case (iii) is characterized by an oscillating solution and has many applications in mathematical physics. For the Case (i) and (ii), on the other hand, we do not have an oscillating solution.

Case (i):

The characteristic roots are given by

$$\rho = \frac{ir}{2m} \mp i\sqrt{\frac{r^2}{4m^2} - \frac{k}{m}}.$$

Therefore, we have a fundamental set of solutions described by

$$u(t) = \exp\left(-\frac{rt}{2m}\right) \exp\left(\mp\sqrt{\frac{r^2}{4m^2} - \frac{k}{m}}t\right).$$

Then, a general solution is given by

$$u(t) = \exp\left(-\frac{rt}{2m}\right)\left[a \exp\left(\sqrt{\frac{r^2}{4m^2} - \frac{k}{m}}t\right) + b \exp\left(-\sqrt{\frac{r^2}{4m^2} - \frac{k}{m}}t\right)\right].$$

Case (ii):

The characteristic roots are given by

$$\rho = \frac{ir}{2m}.$$

Therefore, one of the solutions is

$$u_1(t) = \exp\left(-\frac{rt}{2m}\right).$$

Another solution $u_2(t)$ is given by

$$u_2(t) = c\frac{\partial u_1(t)}{\partial(i\rho)} = c't\,\exp\Big(-\frac{rt}{2m}\Big),$$

where $c$ and $c'$ are appropriate constants. Thus, general solution is given by

$$u(t) = a\,\exp\Big(-\frac{rt}{2m}\Big) + bt\,\exp\Big(-\frac{rt}{2m}\Big).$$

The most important and interesting feature emerges as a "damped oscillator" in the next Case (iii) in many fields of natural science. We are particularly interested in this case.

Case (iii):

Suppose that the damping is relatively weak such that the characteristic equation has two complex roots. Let us examine further details of this case following the prescriptions of IVPs. We divide (8.198) by $m$ for the sake of easy handling of the differential equation such that

$$\frac{\mathrm{d}^2u}{\mathrm{d}t^2} + \frac{r}{m}\frac{\mathrm{d}u}{\mathrm{d}x} + \frac{k}{m}u = \frac{1}{m}d(x).$$

Putting

$$\omega \equiv \sqrt{-\frac{r^2}{4m^2} + \frac{k}{m}}, \tag{8.202}$$

we get a fundamental set of solutions described as

$$u(t) = \exp\Big(-\frac{rt}{2m}\Big)\exp(\mp i\omega t), \tag{8.203}$$

Given BCs, following (8.172) we get as a Green's function

$$\begin{aligned} G(t,\tau) &= \frac{u_2(t)u_1(\tau) - u_1(t)u_2(\tau)}{W(u_1(\tau), u_2(\tau))}\Theta(t,\tau) \\ &= \frac{1}{\omega}\mathrm{e}^{-\frac{r}{2m}(t-\tau)}\sin\omega(t-\tau)\Theta(t,\tau). \end{aligned} \tag{8.204}$$

where $u_1(t) = \exp\big(-\frac{rt}{2m}\big)\exp(i\omega t)$ and $u_2(t) = \exp\big(-\frac{rt}{2m}\big)\exp(-i\omega t)$.

We examine whether $G(t,\tau)$ is eligible for the Green's function as follows:

$$\begin{aligned} \frac{\mathrm{d}G}{\mathrm{d}t} = \frac{1}{\omega}\Big[\Big(-\frac{r}{2m}\Big)\mathrm{e}^{-\frac{r}{2m}(t-\tau)}\sin\omega(t-\tau) + \omega\mathrm{e}^{-\frac{r}{2m}(t-\tau)}\cos\omega(t-\tau)\Big]\Theta(t,\tau) \\ + \frac{1}{\omega}\mathrm{e}^{-\frac{r}{2m}(t-\tau)}\sin\omega(t-\tau)[\delta(t-\tau)\theta(\tau) + \delta(\tau-t)\theta(-\tau)]. \end{aligned} \tag{8.205}$$

The second term of (8.205) vanishes because $\sin\omega(t-\tau)\delta(t-\tau)=0\cdot\delta(t-\tau)=0$. Thus,

$$\begin{aligned}\frac{\mathrm{d}^2G}{\mathrm{d}t^2}=&\frac{1}{\omega}\left[\left(-\frac{r}{2m}\right)^2\mathrm{e}^{-\frac{r}{2m}(t-\tau)}\sin\omega(t-\tau)-\frac{r}{m}\omega\mathrm{e}^{-\frac{r}{2m}(t-\tau)}\cos\omega(t-\tau)\right.\\&\left.-\omega^2e^{-\frac{r}{2m}(t-\tau)}\sin\omega(t-\tau)\right]\Theta(t,\tau)\\&+\frac{1}{\omega}\left[\left(-\frac{r}{2m}\right)e^{-\frac{r}{2m}(t-\tau)}\sin\omega(t-\tau)+\omega\mathrm{e}^{-\frac{r}{2m}(t-\tau)}\cos\omega(t-\tau)\right]\\&\times[\delta(t-\tau)\theta(\tau)+\delta(\tau-t)\theta(-\tau)]\end{aligned}\tag{8.206}$$

In the last term using the property of the $\delta$ function and $\theta$ function, we get $\delta(t-\tau)$. Note here that

$$\begin{aligned}&e^{-\frac{r}{2m}(t-\tau)}\cos\omega(t-\tau)[\delta(t-\tau)\theta(\tau)+\delta(\tau-t)\theta(-\tau)]\\&=e^{-\frac{r}{2m}0}(\cos\omega\cdot0)\{\delta(t-\tau)[\theta(\tau)+\theta(-\tau)]\}\\&=\delta(t-\tau)[\theta(\tau)+\theta(-\tau)]=\delta(t-\tau).\end{aligned}$$

Thus, rearranging (8.206), we get

$$\begin{aligned}\frac{\mathrm{d}^2G}{\mathrm{d}t^2}=&\delta(t-\tau)+\frac{1}{\omega}\left\{-\left[\left(-\frac{r}{2m}\right)^2+\omega^2\right]\mathrm{e}^{-\frac{r}{2m}(t-\tau)}\sin\omega(t-\tau)\right.\\&\left.-\frac{r}{m}\left[\left(-\frac{r}{2m}\right)\mathrm{e}^{-\frac{r}{2m}(t-\tau)}\sin\omega(t-\tau)+\omega\mathrm{e}^{-\frac{r}{2m}(t-\tau)}\cos\omega(t-\tau)\right]\right\}\Theta(t,\tau)\\=&\delta(t-\tau)-\frac{k}{m}G-\frac{r}{m}\frac{\mathrm{d}G}{\mathrm{d}t},\end{aligned}\tag{8.207}$$

where we used (8.202) for the last equality. Rearranging (8.207) once again, we have

$$\frac{\mathrm{d}^2G}{\mathrm{d}t^2}+\frac{r}{m}\frac{\mathrm{d}G}{\mathrm{d}t}+\frac{k}{m}G=\delta(t-\tau).\tag{8.208}$$

Defining the following operator

$$L_t\equiv\frac{\mathrm{d}^2}{\mathrm{d}t^2}+\frac{r}{m}\frac{\mathrm{d}}{\mathrm{d}t}+\frac{k}{m},\tag{8.209}$$

we get

$$L_tG=\delta(t-\tau).\tag{8.210}$$

Note that this expression is consistent with (8.164). Thus, we find that (8.210) satisfies the condition (8.123) of the Green's function, where the weight function is identified with unity.

Now, suppose that a sinusoidally changing external field $e^{i\Omega t}$ influences the motion of the damped oscillator. Here, we assume that an amplitude of the external field is unity. Then, we have

$$\frac{d^2u}{dt^2} + \frac{r}{m}\frac{du}{dx} + \frac{k}{m}u = \frac{1}{m}e^{i\Omega t}. \tag{8.211}$$

Thus, as a solution of the homogeneous boundary conditions [i.e., $u(0) = \dot{u}(0) = 0$], we get

$$u(t) = \frac{1}{m\omega}\int_0^t e^{-\frac{r}{2m}(t-\tau)} e^{i\Omega t} \sin\omega(t-\tau)d\tau, \tag{8.212}$$

where $t$ is an arbitrary positive or negative time. Equation (8.212) shows that with the real part we have an external field $\cos\Omega t$ and that with the imaginary part we have an external field $\sin\Omega t$. To calculate (8.212), we use

$$\sin\omega(t-\tau) = \frac{1}{2i}\left[e^{i\omega(t-\tau)} - e^{-i\omega(t-\tau)}\right]. \tag{8.213}$$

Then, the equation can readily be solved by integration of exponential functions, even though we have to do somewhat lengthy (but straightforward) calculations.

Thus for the real part (i.e., the external field is $\cos\Omega t$), we get a solution

$$\begin{aligned} Cu(t) = {} & \frac{1}{m}\left(\frac{r}{m}\right)\Omega\sin\Omega t + \frac{1}{m}\left(\frac{r}{2m}\right)^2\cos\Omega t - \frac{1}{m}\left(\Omega^2 - \omega^2\right)\cos\Omega t \\ & + \frac{1}{m}e^{-\frac{r}{2m}t}\left[\left(\Omega^2 - \omega^2\right)\cos\omega t - \frac{r}{2m}\frac{1}{\omega}\left(\Omega^2 + \omega^2\right)\sin\omega t\right. \\ & \left. - \left(\frac{r}{2m}\right)^2\cos\omega t - \left(\frac{r}{2m}\right)^3\frac{1}{\omega}\sin\omega t\right], \end{aligned} \tag{8.214}$$

where $C$ is a constant [i.e., a constant denominator of $u(t)$] expressed as

$$C = \left(\Omega^2 - \omega^2\right)^2 + \frac{r^2}{2m^2}\left(\Omega^2 + \omega^2\right) + \left(\frac{r}{2m}\right)^4. \tag{8.215}$$

For the imaginary part (i.e., the external field is $\sin\Omega t$), we get

$$Cu(t) = -\frac{1}{m}\left(\Omega^2 - \omega^2\right)\sin\Omega t - \frac{1}{m}\left(\frac{r}{m}\right)\Omega\cos\Omega t + \frac{1}{m}\left(\frac{r}{2m}\right)^2\sin\Omega t + \frac{1}{m}\mathrm{e}^{-\frac{r}{2m}t}\left[\frac{\Omega}{\omega}\left(\Omega^2 - \omega^2\right)\sin\omega t + \frac{r}{m}\Omega\cos\omega t + \left(\frac{r}{2m}\right)^2\frac{\Omega}{\omega}\sin\omega t\right]. \tag{8.216}$$

In Fig. 8.5, we show an example that depicts the positions of a damped oscillator as a function of time. In Fig. 8.5a, an amplitude of an envelope gradually diminishes with time. An enlarged diagram near the origin (Fig. 8.5b) clearly reflects the

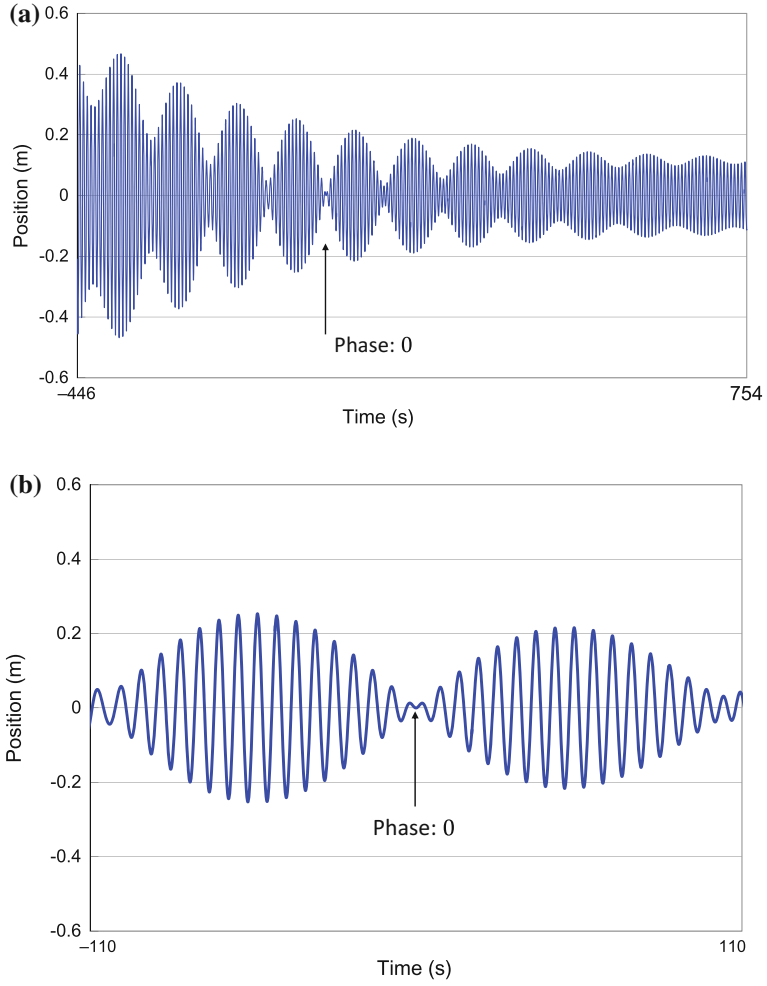


**Fig. 8.5** Example of a damped oscillation as a function of $t$. **a** Overall profile. **b** Profile enlarged near the origin

initial conditions $u(0) = \dot{u(0)} = 0$. In Fig. 8.5, we put $m = 1[\mathrm{kg}]$, $\Omega = 1\left[\frac{1}{s}\right]$, $\omega = 0.94\left[\frac{1}{s}\right]$, and $r = 0.006\left[\frac{\mathrm{kg}}{\mathrm{s}}\right]$. In the above calculations, if $\frac{r}{m}$ is small enough (i.e., damping is small enough), the third order and fourth order of $\frac{r}{m}$ may be ignored and the approximation is precise enough.

In the case of inhomogeneous BCs, given $\sigma_1 = u(0)$ and $\sigma_2 = \dot{u(0)}$, we can decide additional terms $S(t)$ using (8.190) such that

$$\begin{aligned} S(t) = &\left(\sigma_2 + \sigma_1 \frac{r}{m}\right) \frac{1}{\omega} \mathrm{e}^{-\frac{r}{2m}t} \sin \omega t \\ &- \frac{\sigma_1}{\omega}\left(\frac{r}{2m} e^{-\frac{r}{2m}t} \sin \omega t - \omega \mathrm{e}^{-\frac{r}{2m}t} \cos \omega t\right). \end{aligned} \tag{8.217}$$

This term arises from (8.190). Thus, from (8.212) and (8.217), $u(t) + S(t)$ gives a unique solution for the SOLDE with inhomogeneous BCs. Notice that $S(t)$ does not depend on the external field.

## 8.7 Eigenvalue Problems

We often encounter eigenvalue problems in mathematical physics. Of these, those related to Hermitian differential operators have particularly interesting and important features. The eigenvalue problems we have considered in Part I are typical illustrations. Here, we investigate general properties of the eigenvalue problems.

Returning to the case of homogeneous BCs, we consider a following homogeneous SOLDE:

$$a(x)\frac{\mathrm{d}^2 u}{\mathrm{d}x^2} + b(x)\frac{\mathrm{d}u}{\mathrm{d}x} + c(x)u = 0. \tag{8.5}$$

Defining a following differential operator $L_x$ such that

$$L_x = a(x)\frac{\mathrm{d}^2}{\mathrm{d}x^2} + b(x)\frac{\mathrm{d}}{\mathrm{d}x} + c(x), \tag{8.55}$$

we have a homogeneous equation

$$L_x u(x) = 0. \tag{8.218}$$

Putting a constant $-\lambda$ instead of $c(x)$, we have

$$a(x)\frac{\mathrm{d}^2 u}{\mathrm{d}x^2} + b(x)\frac{\mathrm{d}u}{\mathrm{d}x} - \lambda u = 0. \tag{8.219}$$

If we define a differential operator $L_x$ such that

$$L_x = a(x)\frac{\mathrm{d}^2}{\mathrm{d}x^2} + b(x)\frac{\mathrm{d}}{\mathrm{d}x} - \lambda, \tag{8.220}$$

we have a homogeneous equation

$$L_x u = 0 \tag{8.221}$$

to express (8.219). Instead, if we define a differential operator $L_x$ such that

$$L_x = a(x)\frac{\mathrm{d}^2}{\mathrm{d}x^2} + b(x)\frac{\mathrm{d}}{\mathrm{d}x},$$

we have the same homogeneous equation

$$L_x u = \lambda u \tag{8.222}$$

to express (8.219).

Equations (8.221) and (8.222) are essentially the same except that the expression is different. The expression using (8.222) is familiar to us as an eigenvalue equation. The difference between (8.5) and (8.219) is that whereas $c(x)$ in (8.5) is a given *fixed* function, $\lambda$ in (8.219) is constant, but may be *varied* according to the solution of $u(x)$. One of the most essential properties of the eigenvalue problem that is posed in the form of (8.222) is that its solution is not uniquely determined as already studied in various cases of Part I. Remember that the methods based upon the Green's function are valid for a problem to which a homogeneous differential equation has a trivial solution (i.e., identically zero) under homogeneous BCs. In contrast to this situation, even though the eigenvalue problem is basically posed as a homogeneous equation under homogeneous BCs, non-trivial solutions are expected to be obtained. In this respect, we have seen that in Part I we rejected a trivial solution (i.e., identically zero) because of no physical meaning.

As exemplified in Part I, the eigenvalue problems that appear in mathematical physics are closely connected to the Hermiticity of (differential) operators. This is because in many cases an eigenvalue is required to be real. We have already examined how we can convert a differential operator to the self-adjoint form. That is, if we define $p(x)$ as in (8.26), we have the self-adjoint operator as described in (8.70). As a symbolic description, we have

$$w(x)L_x u = \frac{\mathrm{d}}{\mathrm{d}x}\left[p(x)\frac{\mathrm{d}u}{\mathrm{d}x}\right] + c(x)w(x)u. \tag{8.223}$$

In the same way, multiplying both sides of (8.222) by $w(x)$, we get

$$w(x)L_x u = \lambda w(x)u. \tag{8.224}$$

For instance, Hermite differential equation that has already appeared as (2.118) in Sect. 2.3 is described as

$$\frac{\mathrm{d}^2u}{\mathrm{d}x^2} - 2x\frac{\mathrm{d}u}{\mathrm{d}x} + 2nu = 0. \tag{8.225}$$

If we express (8.225) as in (8.224), multiplying $\mathrm{e}^{-x^2}$ on both sides of (8.225), we have

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\mathrm{e}^{-x^2}\frac{\mathrm{d}u}{\mathrm{d}x}\right) + 2n\mathrm{e}^{-x^2}u = 0. \tag{8.226}$$

Notice that the differential operator has been converted to a self-adjoint form according to (8.31) that defines a real and positive weight function $\mathrm{e}^{-x^2}$ in the present case. The domain of the Hermite differential equation is $(-\infty, +\infty)$ at the endpoints (i.e., $\pm\infty$) of which the surface term of RHS of (8.69) approaches zero sufficiently rapidly in virtue of $\mathrm{e}^{-x^2}$.

In Sects. 3.4, and 3.5, in turn, we dealt with the (associated) Legendre differential equation (3.127) for which the relevant differential operator is self-adjoint. The surface term corresponding to (8.69) vanishes. It is because $\left(1-\xi^2\right)$ vanishes at the endpoints $\xi = \cos\theta = \pm 1$ (i.e., $\theta = 0$ or $\pi$) from (3.107). Thus, the Hermiticity is automatically ensured for the (associated) Legendre differential equation as well as Hermite differential equation. In those cases, even though the differential equations do not satisfy any particular BCs, the Hermiticity is yet ensured.

In the theory of differential equations, the aforementioned properties of the Hermitian operators have been fully investigated as the so-called Strum–Liouville system (or problem) in the form of a homogeneous differential equation. The related differential equations are connected to classical orthogonal polynomials having personal names such as Hermite, Laguerre, Jacobi, Gegenbauer, Legendre, Tchebichef. These equations frequently appear in quantum mechanics and electromagnetism as typical examples of Strum–Liouville system. They can be converted to differential equations by multiplying an original form by a weight function. The resulting equations can be expressed as

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[a(x)w(x)\frac{\mathrm{d}Y_n(x)}{\mathrm{d}x}\right] + \lambda_n w(x)Y_n(x) = 0, \tag{8.227}$$

where $Y_n(x)$ is a collective representation of classical orthogonal polynomials. Equation (8.226) is an example. Conventionally, a following form is adopted instead of (8.227):

$$\frac{1}{w(x)}\frac{\mathrm{d}}{\mathrm{d}x}\left[a(x)w(x)\frac{\mathrm{d}Y_n(x)}{\mathrm{d}x}\right]+\lambda_n Y_n(x)=0, \tag{8.228}$$

where we put $(aw)'=bw$. That is, the differential equation is originally described as

$$a(x)\frac{\mathrm{d}^2Y_n(x)}{\mathrm{d}x^2}+b(x)\frac{\mathrm{d}Y_n(x)}{\mathrm{d}x}+\lambda_n Y_n(x)=0. \tag{8.229}$$

In the case of Hermite polynomials, for instance, $a(x)=1$ and $w(x)=\mathrm{e}^{-x^2}$. Since we have $(aw)'=-2x\mathrm{e}^{-x^2}=bw$, we can put $b=-2x$. Examples including this case are tabulated in Table 8.2. The eigenvalues $\lambda_n$ are associated with real numbers that characterize the individual physical systems. The related fields have wide applications in many branches of natural science.

After having converted the operator to the self-adjoint form, i.e., $L_x=L_x^\dagger$, instead of (8.100), we have

$$\int_r^s \mathrm{d}x w(x)\{v^*(L_x u)-[L_x v]^* u\}=0. \tag{8.230}$$

Rewriting it, we get

$$\int_r^s \mathrm{d}x v^*[w(x)L_x u]=\int_r^s \mathrm{d}x[w(x)L_x v]^* u. \tag{8.231}$$

If we use an inner product notation described by (8.88), we get

$$\langle v|L_x u\rangle=\langle L_x v|u\rangle. \tag{8.232}$$

Here let us think of two eigenfunctions $\psi_i$ and $\psi_j$ that belong to an eigenvalue $\lambda_i$ and $\lambda_j$, respectively. That is,

$$w(x)L_x\psi_i=\lambda_i w(x)\psi_i \quad \text{and} \quad w(x)L_x\psi_j=\lambda_j w(x)\psi_j. \tag{8.233}$$

Inserting $\psi_i$ and $\psi_j$ into $u$ and $v$, respectively, in (8.232), we have

$$\langle\psi_j|L_x\psi_i\rangle=\langle\psi_j|\lambda_i\psi_i\rangle=\lambda_i\langle\psi_j|\psi_i\rangle=\lambda_j^*\langle\psi_j|\psi_i\rangle=\langle\lambda_j\psi_j|\psi_i\rangle=\langle L_x\psi_j|\psi_i\rangle. \tag{8.234}$$

With the second and third equalities, we have used a rule of the inner product (see Parts I and III). Therefore, we get

**Table 8.2** Classical polynomials and their related SOLDEs

| Name of the polynomial | SOLDE form | Weight function: $w(x)$ | Domain |
|---|---|---|---|
| Hermite: $H_n(x)$ | $\frac{\mathrm{d}^2}{\mathrm{d}x^2}H_n(x) - 2x\frac{\mathrm{d}}{\mathrm{d}x}H_n(x) + 2nH_n(x) = 0$ | $\mathrm{e}^{-x^2}$ | $(-\infty, +\infty)$ |
| Laguerre: $L_n^v(x)$ | $x\frac{\mathrm{d}^2}{\mathrm{d}x^2}L_n^v(x) + (v+1-x)\frac{\mathrm{d}}{\mathrm{d}x}L_n^v(x) + nL_n^v(x) = 0$ | $x^v\mathrm{e}^{-x}(v > -1)$ | $[0, +\infty)$ |
| Gegenbauer: $C_n^\lambda(x)$ | $(1-x^2)\frac{\mathrm{d}^2}{\mathrm{d}x^2}C_n^\lambda(x) - (2\lambda+1)x\frac{\mathrm{d}}{\mathrm{d}x}C_n^\lambda(x) + n(n+2\lambda)C_n^\lambda(x) = 0$ | $(1-x^2)^{\lambda-\frac{1}{2}}$ $(\lambda > -\frac{1}{2})$ | $[-1, +1]$ |
| Legendre: $P_n(x)$ | $(1-x^2)\frac{\mathrm{d}^2}{\mathrm{d}x^2}P_n(x) - 2x\frac{\mathrm{d}}{\mathrm{d}x}P_n(x) + n(n+1)P_n(x) = 0$ | 1 | $[-1, +1]$ |

$$\left(\lambda_i - \lambda_j^*\right)\langle\psi_j|\psi_i\rangle = 0. \tag{8.235}$$

Putting $i = j$ in (8.235), we get

$$\left(\lambda_i - \lambda_i^*\right)\langle\psi_i|\psi_i\rangle = 0. \tag{8.236}$$

An inner product $\langle\psi_i|\psi_i\rangle$ vanishes if and only if $|\psi_i\rangle \equiv 0$; see inner product calculation rules of Sect. 11.1. However, $|\psi_i\rangle \equiv 0$ is not acceptable as a physical state. Therefore, we must have $\langle\psi_i|\psi_i\rangle \neq 0$. Thus, we get

$$\lambda_i - \lambda_i^* = 0 \quad \text{or} \quad \lambda_i = \lambda_i^*. \tag{8.237}$$

The relation (8.237) obviously indicates that $\lambda_i$ is real; i.e., we find that eigenvalues of an Hermitian operator are real. If $\lambda_i \neq \lambda_j = \lambda_j^*$, from (8.235) we get

$$\langle\psi_j|\psi_i\rangle = 0. \tag{8.238}$$

That is, $|\psi_j\rangle$ and $|\psi_i\rangle$ are orthogonal to each other.

We often encounter related orthogonality relationship between vectors and functions. We saw several cases in Part I and will see other cases in Part III.

## References

1. Dennery P, Krzywicki A (1996) Mathematics for physicists. Dover, New York
2. Stakgold I (1998) Green's functions and boundary value problems, 2nd edn. Wiley, New York

# Part III
# Linear Vector Spaces

In this part, we treat vectors and their transformations in linear vector spaces so that we can address various aspects of mathematical physics systematically but intuitively. We outline general principles of linear vector spaces mostly from an algebraic point of view. Starting with abstract definition and description of vectors, we deal with their transformation in a vector space using a matrix. An inner product is a central concept in the theory of a linear vector space so that two vectors can be associated with each other to yield a scalar. Unlike many of books of linear algebra and linear vector spaces, however, we describe canonical forms of matrices before considering the inner product. This is because we can treat the topics in light of the abstract theory of matrices and vector space without a concept of the inner product. Of the canonical forms of matrices, Jordan canonical form is of paramount importance. We study how it is constructed providing a tangible example.

In relation to the inner product space, normal operators such as Hermitian operators and unitary operators frequently appear in quantum mechanics and electromagnetism. From a general aspect, we revisit the theory of Hermitian operators that often appeared in both Parts I and II.

# Chapter 9
# Vectors and Their Transformation

In this chapter, we deal with the theory of finite-dimensional linear vector spaces. Such vector spaces are spanned by a finite number of linearly independent vectors, namely basis vectors. In conjunction with developing an abstract concept and theory, we mention a notion of mapping among mathematical elements. A linear transformation of a vector is a special kind of mapping. In particular, we focus on endomorphism within a $n$-dimensional vector space $V^n$. Here, the endomorphism is defined as a linear transformation: $V^n \rightarrow V^n$. The endomorphism is represented by a $(n,n)$ square matrix. This is most often the case with physical and chemical applications, when we deal with matrix algebra. In this book, we focus on this type of transformation.

A non-singular matrix plays an important role in the endomorphism. In this connection, we consider its inverse matrix and determinant. All these fundamental concepts supply us with a sufficient basis for better understanding of the theory of the linear vector spaces. Through these processes, we should be able to get acquainted with connection between algebraic and analytical approaches and gain a broad perspective on various aspects of mathematical physics and related fields.

## 9.1 Vectors

From both fundamental and practical points of view, it is desirable to define linear vector spaces in an abstract way. Suppose $V$ is a set of elements denoted by $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$, etc., called vectors. The set $V$ is a linear vector space (or simply a vector space), if a sum $\boldsymbol{a}+\boldsymbol{b} \in V$ is defined for any pair of vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ and the elements of $V$ satisfy the following mathematical relations:

$$(\boldsymbol{a}+\boldsymbol{b})+\boldsymbol{c} = \boldsymbol{a}+(\boldsymbol{b}+\boldsymbol{c}), \tag{9.1}$$

$$\boldsymbol{a}+\boldsymbol{b} = \boldsymbol{b}+\boldsymbol{a}, \tag{9.2}$$



S. Hotta, *Mathematical Physical Chemistry*,
https://doi.org/10.1007/978-981-10-7671-8_9

$$\boldsymbol{a}+\boldsymbol{0}=\boldsymbol{a}, \tag{9.3}$$

$$\boldsymbol{a}+(-\boldsymbol{a})=\boldsymbol{0}. \tag{9.4}$$

For the above, $\boldsymbol{0}$ is called the zero vector. Furthermore, for $\boldsymbol{a} \in V$, $c\boldsymbol{a} \in V$ is defined ($c$ is a complex number called a scalar) and we assume the following relations among vectors and scalars:

$$(cd)\boldsymbol{a}=c(d\boldsymbol{a}), \tag{9.5}$$

$$1\boldsymbol{a}=\boldsymbol{a}, \tag{9.6}$$

$$c(\boldsymbol{a}+\boldsymbol{b})=c\boldsymbol{a}+c\boldsymbol{b}, \tag{9.7}$$

$$(c+d)\boldsymbol{a}=c\boldsymbol{a}+d\boldsymbol{a}. \tag{9.8}$$

On the basis of the above relations, we can construct the following expression called a linear combination:

$$c_1\boldsymbol{a}_1+c_2\boldsymbol{a}_2+\cdots+c_n\boldsymbol{a}_n.$$

If this linear combination is equated to zero, we obtain

$$c_1\boldsymbol{a}_1+c_2\boldsymbol{a}_2+\cdots+c_n\boldsymbol{a}_n=\boldsymbol{0}. \tag{9.9}$$

If (9.9) holds only in the case where every $c_i=0(1\leq i\leq n)$, the vectors $\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_n$ are said to be linearly independent. In this case, the relation represented by (9.9) is said to be trivial. If the relation is non-trivial (i.e., ${}^{\exists}c_i\neq 0$), those vectors are said to be linearly dependent.

If in the vector space $V$ the maximum number of linearly independent vectors is $n$, $V$ is said to be an $n$-dimensional vector space and sometimes denoted by $V^n$. In this case, any vector $\boldsymbol{x}$ of $V^n$ is expressed uniquely as a linear combination of linearly independent vectors such that

$$\boldsymbol{x}=x_1\boldsymbol{a}_1+x_2\boldsymbol{a}_2+\cdots+x_n\boldsymbol{a}_n. \tag{9.10}$$

Suppose $\boldsymbol{x}$ is denoted by

$$\boldsymbol{x}=x_1'\boldsymbol{a}_1+x_2'\boldsymbol{a}_2+\cdots+x_n'\boldsymbol{a}_n. \tag{9.11}$$

Subtracting both sides of (9.11) from (9.10), we obtain

$$\boldsymbol{0}=(x_1-x_1')\boldsymbol{a}_1+(x_2-x_2')\boldsymbol{a}_2+\cdots+(x_n-x_n')\boldsymbol{a}_n. \tag{9.12}$$

Linear independence of the vectors $\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_n$ implies $x_n-x_n'=0$; i.e., $x_n=x_n'(1\leq i\leq n)$. These $n$ linearly independent vectors are referred to as basis vectors.

A vector space that has a finite number of basis vectors is called finite-dimensional; otherwise, it is infinite dimensional.

Alternatively, we express (9.10) as

$$\boldsymbol{x} = (\boldsymbol{a}_1 \cdots \boldsymbol{a}_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}. \tag{9.13}$$

A set of coordinates $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ is called a column vector (or a numerical vector) that indicates an "address" of the vector $\boldsymbol{x}$ with respect to the basis vectors $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n$. Any vector in $V^n$ can be expressed as a linear combination of the basis vectors, and hence, we say that $V^n$ is spanned by $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n$. This is represented as

$$V^n = \text{Span}\,\{\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n\}. \tag{9.14}$$

Let us think of a subset $W$ of $V^n$ (i.e., $W \subset V^n$). If the following relations hold for $W$, $W$ is said to be a (linear) subspace of $V^n$.

$$\begin{aligned} &\boldsymbol{a}, \boldsymbol{b} \in W \Rightarrow \boldsymbol{a} + \boldsymbol{b} \in W, \\ &\boldsymbol{a} \in W \Rightarrow c\boldsymbol{a} \in W. \end{aligned} \tag{9.15}$$

These two relations ensure that the relations of (9.1) to (9.8) hold for $W$ as well. The dimension of $W$ is equal to or smaller than $n$. For instance, $W = \text{Span}\ \{\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_r\} (r \leq n)$ is a subspace of $V^n$. If $r = n$, $W = V^n$. Suppose that there are two subspaces $W_1 = \text{Span}\,\{\boldsymbol{a}_1\}$ and $W_2 = \text{Span}\ \{\boldsymbol{a}_2\}$. Note that in this case, $W_1 \cup W_2$ is not a subspace, because $W_1 \cup W_2$ does not contain $\boldsymbol{a}_1 + \boldsymbol{a}_2$. However, a set $U$ defined by

$$U = \{\boldsymbol{x} = \boldsymbol{x}_1 + \boldsymbol{x}_2; {}^\forall\boldsymbol{x}_1 \in W_1, {}^\forall\boldsymbol{x}_2 \in W_2\} \tag{9.16}$$

is a subspace of $V^n$. We denote this subspace by $W_1 + W_2$.

To show this is in fact a subspace, suppose that $\boldsymbol{x}, \boldsymbol{y} \in W_1 + W_2$. Then, we may express $\boldsymbol{x} = \boldsymbol{x}_1 + \boldsymbol{x}_2$ and $\boldsymbol{y} = \boldsymbol{y}_1 + \boldsymbol{y}_2$, where $\boldsymbol{x}_1, \boldsymbol{y}_1 \in W_1$; $\boldsymbol{x}_2, \boldsymbol{y}_2 \in W_2$. We have $\boldsymbol{x} + \boldsymbol{y} = (\boldsymbol{x}_1 + \boldsymbol{y}_1) + (\boldsymbol{x}_2 + \boldsymbol{y}_2)$, where $\boldsymbol{x}_1 + \boldsymbol{y}_1 \in W_1$ and $\boldsymbol{x}_2 + \boldsymbol{y}_2 \in W_2$ because both $W_1$ and $W_2$ are subspaces. Therefore, $\boldsymbol{x} + \boldsymbol{y} \in W_1 + W_2$. Meanwhile, with any scalar $c$, $c\boldsymbol{x} = c\boldsymbol{x}_1 + c\boldsymbol{x}_2 \in W_1 + W_2$. By definition (9.15), $W_1 + W_2$ is a subspace accordingly. Suppose here $\boldsymbol{x}_1 \in W_1$. Then, $\boldsymbol{x}_1 = \boldsymbol{x}_1 + \boldsymbol{0} \in W_1 + W_2$. Then, $W_1 \subset W_1 + W_2$. Similarly, we have $W_2 \subset W_1 + W_2$. Thus, $W_1 + W_2$ contains both $W_1$ and $W_2$. Conversely, let $W$ be an arbitrary subspace that contains both $W_1$ and $W_2$. Then, we have ${}^\forall\boldsymbol{x}_1 \in W_1 \subset W$ and ${}^\forall\boldsymbol{x}_2 \in W_2 \subset W$ and, hence, we have $\boldsymbol{x}_1 + \boldsymbol{x}_2 \in W$ by definition (9.15). But, from (9.16), $W_1 + W_2 = \{\boldsymbol{x}_1 + \boldsymbol{x}_2; {}^\forall\boldsymbol{x}_1$

$\in W_1, {}^{\forall}\boldsymbol{x}_2 \in W_2\}$. Hence, $W_1 + W_2 \subset W$. Consequently, any subspace necessarily contains $W_1 + W_2$. This implies that $W_1 + W_2$ is the *smallest* subspace that contains both $W_1$ and $W_2$.

*Example 9.1* Consider a three-dimensional Cartesian space $\mathbb{R}^3$ (Fig. 9.1). We regard the $xy$-plane and $yz$-plane as a subspace $W_1$ and $W_2$, respectively, and $\mathbb{R}^3 = W_1 + W_2$. In Fig. 9.1a, a vector $\overrightarrow{OB}$ (in $\mathbb{R}^3$) is expressed as $\overrightarrow{OA} + \overrightarrow{AB}$ (i.e., a sum of a vector in $W_1$ and that in $W_2$). Alternatively, the same vector $\overrightarrow{OB}$ can be expressed as $\overrightarrow{OA'} + \overrightarrow{A'B}$. On the other hand, we can designate a subspace in a different way; i.e., in Fig. 9.1b, the $z$-axis is chosen for a subspace $W_3$ instead of $W_2$. We have $\mathbb{R}^3 = W_1 + W_3$ as well. In this case, however, $\overrightarrow{OB}$ is uniquely expressed as $\overrightarrow{OB} = \overrightarrow{OP} + \overrightarrow{PB}$. Notice that in Fig. 9.1a, $W_1 \cap W_2 = \text{Span}\{\boldsymbol{e}_2\}$,

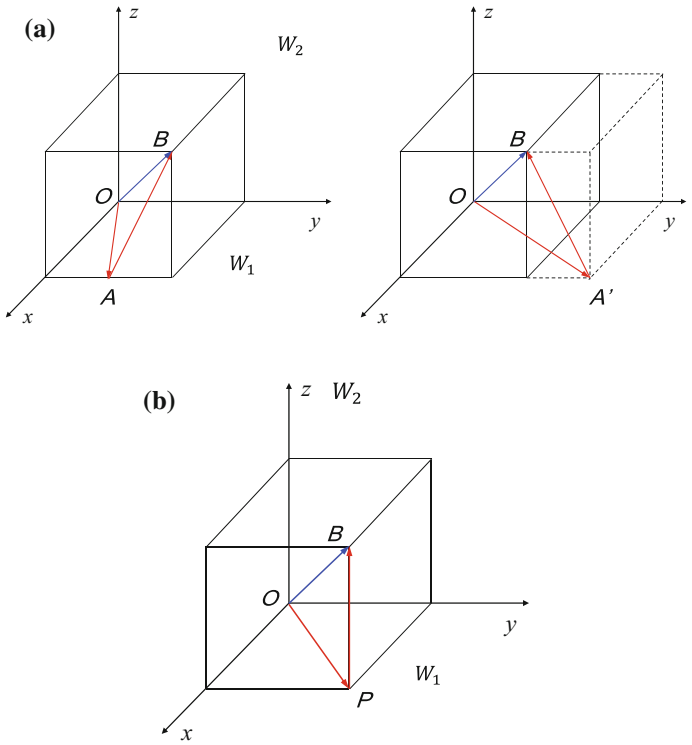


**Fig. 9.1** Decomposition of a vector in a three-dimensional Cartesian space $\mathbb{R}^3$ into two subspaces. **a** $\mathbb{R}^3 = W_1 + W_2$; $W_1 \cap W_2 = \text{Span}\{\boldsymbol{e}_2\}$, where $\boldsymbol{e}_2$ is a unit vector in the positive direction of the $y$-axis. **b** $\mathbb{R}^3 = W_1 + W_3$; $W_1 \cap W_3 = \{\boldsymbol{0}\}$

where $\boldsymbol{e}_2$ is a unit vector in the positive direction of the $y$-axis. In Fig. 9.1b, on the other hand, we have $W_1 \cap W_3 = \{\boldsymbol{0}\}$.

We can generalize this example to the following theorem.

**Theorem 9.1** *Let $W_1$ and $W_2$ be subspaces of $V$ and $V = W_1 + W_2$. Then, a vector $\boldsymbol{x}$ in $V$ is uniquely expressed as*

$$\boldsymbol{x} = \boldsymbol{x}_1 + \boldsymbol{x}_2,\ \boldsymbol{x}_1 \in W_1, \boldsymbol{x}_2 \in W_2,$$

if and only if $W_1 \cap W_2 = \{\boldsymbol{0}\}$.

*Proof* Suppose $W_1 \cap W_2 = \{\boldsymbol{0}\}$ and $\boldsymbol{x} = \boldsymbol{x}_1 + \boldsymbol{x}_2 = \boldsymbol{x}_1' + \boldsymbol{x}_2'$, $\boldsymbol{x}_1, \boldsymbol{x}_1' \in W_1, \boldsymbol{x}_2, \boldsymbol{x}_2' \in W_2$. Then, $\boldsymbol{x}_1 - \boldsymbol{x}_1' = \boldsymbol{x}_2' - \boldsymbol{x}_2$. LHS belongs to $W_1$, and RHS belongs to $W_2$. Both sides belong to $W_1 \cap W_2$ accordingly. Hence, from the supposition, both the sides should be equal to a zero vector. Therefore, $\boldsymbol{x}_1 = \boldsymbol{x}_1'$, $\boldsymbol{x}_2 = \boldsymbol{x}_2'$. This implies that $\boldsymbol{x}$ is expressed uniquely as $\boldsymbol{x} = \boldsymbol{x}_1 + \boldsymbol{x}_2$. Conversely, suppose the vector representation $(\boldsymbol{x} = \boldsymbol{x}_1 + \boldsymbol{x}_2)$ is unique and $\boldsymbol{x} \in W_1 \cap W_2$. Then, $\boldsymbol{x} = \boldsymbol{x} + \boldsymbol{0} = \boldsymbol{0} + \boldsymbol{x}; \boldsymbol{x}, \boldsymbol{0} \in W_1$ and $\boldsymbol{x}, \boldsymbol{0} \in W_2$. Uniqueness of the representation implies that $\boldsymbol{x} = \boldsymbol{0}$. Consequently, $W_1 \cap W_2 = \{\boldsymbol{0}\}$ follows.

In case $W_1 \cap W_2 = \{\boldsymbol{0}\}$, $V = W_1 + W_2$ is said to be a direct sum of $W_1$ and $W_2$ or we say that $V$ is decomposed into a direct sum of $W_1$ and $W_2$. We symbolically denote this by

$$V = W_1 \oplus W_2. \tag{9.17}$$

In this case, the following equality holds:

$$\dim\ V = \dim\ W_1 + \dim\ W_2, \tag{9.18}$$

where "dim" stands for dimension of the vector space considered. To prove (9.18), we suppose that $V$ is a $n$-dimensional vector space and that $W_1$ and $W_2$ are spanned by $r_1$ and $r_2$ linearly independent vectors, respectively, such that

$$W_1 = \text{Span}\ \left\{\boldsymbol{e}_1^{(1)}, \boldsymbol{e}_2^{(1)}, \cdots, \boldsymbol{e}_{r_1}^{(1)}\right\} \quad \text{and} \quad W_2 = \text{Span}\ \left\{\boldsymbol{e}_1^{(2)}, \boldsymbol{e}_2^{(2)}, \cdots, \boldsymbol{e}_{r_2}^{(2)}\right\}. \tag{9.19}$$

This is equivalent to that dimension of $W_1$ and $W_2$ is $r_1$ and $r_2$, respectively. If $V = W_1 + W_2$ (here we do not assume that the summation is a direct sum), we have

$$V = \text{Span}\ \left\{\boldsymbol{e}_1^{(1)}, \boldsymbol{e}_2^{(1)}, \cdots, \boldsymbol{e}_{r_1}^{(1)}, \boldsymbol{e}_1^{(2)}, \boldsymbol{e}_2^{(2)}, \cdots, \boldsymbol{e}_{r_2}^{(2)}\right\}. \tag{9.20}$$

Then, we have $n \leq r_1 + r_2$. This is almost trivial. Suppose $r_1 + r_2 < n$. Then, these $(r_1 + r_2)$ vectors cannot span $V$, but we need additional vectors for all vectors including the additional vectors to span $V$. Thus, we should have $n \leq r_1 + r_2$ accordingly. That is,

$$\dim V \leq \dim W_1 + \dim W_2. \tag{9.21}$$

Now, let us assume that $V = W_1 \oplus W_2$. Then, $\boldsymbol{e}_i^{(2)} (1 \leq i \leq r_2)$ must be linearly independent of $\boldsymbol{e}_1^{(1)}, \boldsymbol{e}_2^{(1)}, \cdots, \boldsymbol{e}_{r_1}^{(1)}$. If not, $\boldsymbol{e}_i^{(2)}$ could be described as a linear combination of $\boldsymbol{e}_1^{(1)}, \boldsymbol{e}_2^{(1)}, \cdots, \boldsymbol{e}_{r_1}^{(1)}$. But, this would imply that $\boldsymbol{e}_i^{(2)} \in W_1$, i.e., $e_i^{(2)} \in W_1 \cap W_2$, in contradiction to that we have $V = W_1 \oplus W_2$. It is because $W_1 \cap W_2 = \{\boldsymbol{0}\}$ by assumption. Likewise, $\boldsymbol{e}_j^{(1)} (1 \leq j \leq r_1)$ is linearly independent of $\boldsymbol{e}_1^{(2)}, \boldsymbol{e}_2^{(2)}, \cdots, \boldsymbol{e}_{r_2}^{(2)}$. Hence, $\boldsymbol{e}_1^{(1)}, \boldsymbol{e}_2^{(1)}, \cdots, \boldsymbol{e}_{r_1}^{(1)}, \boldsymbol{e}_1^{(2)}, \boldsymbol{e}_2^{(2)}, \cdots, \boldsymbol{e}_{r_2}^{(2)}$ must be linearly independent. Thus, $n \geq r_1 + r_2$. This is because in the vector space $V$ we may well have additional vector(s) that are independent of the above $(r_1 + r_2)$ vectors. Meanwhile, $n \leq r_1 + r_2$ from the above. Consequently, we must have $n = r_1 + r_2$. Thus, we have proven that

$$V = W_1 \oplus W_2 \Rightarrow \dim V = \dim W_1 + \dim W_2.$$

Conversely, suppose $n = r_1 + r_2$. Then, any vector $\boldsymbol{x}$ in $V$ is expressed uniquely as

$$\boldsymbol{x} = \left(a_1 \boldsymbol{e}_1^{(1)} + \cdots + a_{r_1} \boldsymbol{e}_{r_1}^{(1)}\right) + \left(b_1 \boldsymbol{e}_1^{(2)} + \cdots + b_{r_2} \boldsymbol{e}_{r_2}^{(2)}\right). \tag{9.22}$$

The vector described by the first term is contained in $W_1$ and that described by the second term in $W_2$. Both the terms are again expressed uniquely. Therefore, we get $V = W_1 \oplus W_2$. This is a proof of

$$\dim V = \dim W_1 + \dim W_2 \Rightarrow V = W_1 \oplus W_2.$$

The above statements are summarized as the following theorem:

**Theorem 9.2** *Let $V$ be a vector space and let $W_1$ and $W_2$ be subspaces of $V$. Also, suppose $V = W_1 + W_2$. Then, the following relation holds*:

$$\dim V \leq \dim W_1 + \dim W_2. \tag{9.21}$$

Furthermore, we have

$$\dim V = \dim W_1 + \dim W_2, \tag{9.23}$$

if and only if $V = W_1 \oplus W_2$.

Theorem 9.2 is readily extended to the case where there are three or more subspaces. That is, having $W_1, W_2, \cdots, W_m$ so that $V = W_1 + W_2 \cdots + W_m$, we obtain the following relation:

$$\dim V \leq \dim W_1 + \dim W_2 + \cdots + W_m. \tag{9.24}$$

The equality of (9.24) holds if and only if $V = W_1 \oplus W_2 \oplus \cdots \oplus W_m$.

In light of Theorem 9.2, Example 9.1 says that $3 = \dim \mathbb{R}^3 < 2+2 = \dim W_1 + \dim W_2$. But, $\dim \mathbb{R}^3 = 2+1 = \dim W_1 + \dim W_3$. Therefore, we have $\mathbb{R}^3 = W_1 \oplus W_2$.

## 9.2 Linear Transformations of Vectors

In the previous section, we introduced vectors and their calculation rules in a linear vector space. It is natural and convenient to relate a vector to another vector, as a function $f$ relates a number (either real or complex) to another number such that $y = f(x)$, where $x$ and $y$ are certain two numbers. A linear transformation from the vector space $V$ to another vector space $W$ is a mapping $A : V \to W$ such that

$$A(c\boldsymbol{a} + d\boldsymbol{b}) = cA(\boldsymbol{a}) + dA(\boldsymbol{b}). \tag{9.25}$$

We will briefly discuss the concepts of mapping at the end of this section.

It is convenient to define addition of the linear transformations. It is defined as

$$(A+B)\boldsymbol{a} = A\boldsymbol{a} + B\boldsymbol{a}, \tag{9.26}$$

where $\boldsymbol{a}$ is any vector in $V$.

Since (9.25) is a broad but abstract definition, we begin with a well-known simple example of rotation of a vector within a $xy$-plane (Fig. 9.2). We denote an arbitrary position vector $\boldsymbol{x}$ in the $xy$-plane by

$$\begin{aligned} \boldsymbol{x} &= x\boldsymbol{e}_1 + y\boldsymbol{e}_2. \\ &= (\boldsymbol{e}_1\boldsymbol{e}_2)\begin{pmatrix} x \\ y \end{pmatrix}, \end{aligned} \tag{9.27}$$

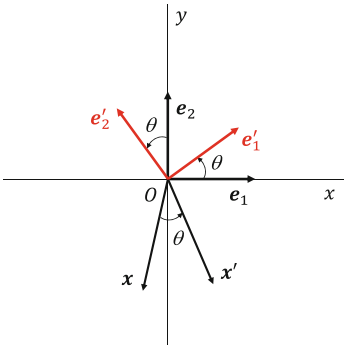


**Fig. 9.2** Rotation of a vector within a $xy$-plane

where $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ are unit basis vectors in the $xy$-plane, and $x$ and $y$ are coordinates of the vector $\boldsymbol{x}$ in reference to $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$. The expression (9.27) is consistent with (9.13). The rotation represented in Fig. 9.2 is an example of a linear transformation. We call this rotation $R$. According to the definition,

$$R(x\boldsymbol{e}_1 + y\boldsymbol{e}_2) = R(\boldsymbol{x}) = xR(\boldsymbol{e}_1) + yR(\boldsymbol{e}_2). \tag{9.28}$$

Putting $R(x\boldsymbol{e}_1 + y\boldsymbol{e}_2) = \boldsymbol{x}'$, $R(\boldsymbol{e}_1) = \boldsymbol{e}'_1$, and $R(\boldsymbol{e}_2) = \boldsymbol{e}'_2$,

$$\begin{aligned} \boldsymbol{x}' &= x\boldsymbol{e}'_1 + y\boldsymbol{e}'_2, \\ &= (\boldsymbol{e}'_1\boldsymbol{e}'_2)\begin{pmatrix} x \\ y \end{pmatrix}. \end{aligned} \tag{9.29}$$

From Fig. 9.2, we readily obtain

$$\begin{aligned} \boldsymbol{e}'_1 &= \boldsymbol{e}_1 \cos\,\theta + \boldsymbol{e}_2 \sin\,\theta, \\ \boldsymbol{e}'_2 &= -\boldsymbol{e}_1 \sin\,\theta + \boldsymbol{e}_2 \cos\,\theta. \end{aligned} \tag{9.30}$$

Using a matrix representation,

$$(\boldsymbol{e}'_1\boldsymbol{e}'_2) = (\boldsymbol{e}_1\boldsymbol{e}_2)\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}. \tag{9.31}$$

Substituting (9.30) into (9.29), we obtain

$$\boldsymbol{x}' = (x\cos\,\theta - y\sin\,\theta)\boldsymbol{e}_1 + (x\sin\,\theta + y\cos\,\theta)\boldsymbol{e}_2. \tag{9.32}$$

Meanwhile, $\boldsymbol{x}'$ can be expressed relative to the original basis vectors $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$.

$$\boldsymbol{x}' = x'\boldsymbol{e}_1 + y'\boldsymbol{e}_2. \tag{9.33}$$

Comparing (9.32) and (9.33), uniqueness of the representation ensures that

$$\begin{aligned} x' &= x\cos\,\theta - y\sin\,\theta, \\ y' &= x\sin\,\theta + y\cos\,\theta. \end{aligned} \tag{9.34}$$

Using a matrix representation once again,

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}. \tag{9.35}$$

Further combining (9.29) and (9.31), we get

$$\boldsymbol{x}' = R(\boldsymbol{x}) = (\boldsymbol{e}_1\boldsymbol{e}_2)\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}. \tag{9.36}$$

The above example demonstrates that the linear transformation $R$ has a (2,2) matrix representation shown in (9.36). Moreover, this example obviously shows that if a vector is expressed as a linear combination of the basis vectors, the "coordinates" (represented by a column vector) can be transformed as well by the same matrix.

Regarding an abstract $n$-dimensional linear vector space $V^n$, the linear vector transformation $A$ is given by

$$A(\boldsymbol{x}) = (\boldsymbol{e}_1 \cdots \boldsymbol{e}_n) \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \tag{9.37}$$

where $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots,$ and $\boldsymbol{e}_n$ are basis vectors and $x_1, x_2, \cdots,$ and $x_n$ are the corresponding coordinates of a vector $\boldsymbol{x} = \sum_{i=1}^{n} x_i \boldsymbol{e}_i$. We assume that the transformation is a mapping $A$: $V^n \to V^n$ (i.e., endomorphism). In this case, the transformation is represented by an $(n,n)$ matrix. Note that the matrix operates on the basis vectors from the right and that it operates on the coordinates (i.e., a column vector) from the left. In (9.37), we often omit a parenthesis to simply write $A\boldsymbol{x}$.

Here we mention matrix notation for later convenience. We often identify a linear vector transformation with its representation matrix and denote both transformation and matrix by $A$. On this occasion, we write

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}, \; A = (A)_{ij} = \left(a_{ij}\right), \text{ etc.}, \tag{9.38}$$

where with the second expression $(A)_{ij}$ and $\left(a_{ij}\right)$ represent the matrix $A$ itself; for we frequently use indexed matrix notations such as $A^{-1}$, $A^{\dagger}$, $\widetilde{A}$. The notation (9.38) can conveniently be used in such cases. Note moreover that $a_{ij}$ represents the matrix $A$ as well.

Equation (9.37) has duality such that the matrix $A$ operates either on the basis vectors or coordinates. This can explicitly be written as

$$\begin{aligned} A(x) &= \left[ (\boldsymbol{e}_1 \cdots \boldsymbol{e}_n) \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \right] \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \\ &= (\boldsymbol{e}_1 \cdots \boldsymbol{e}_n) \left[ \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right]. \end{aligned}$$

That is, we assume the associative law with the above expression. Making summation representation, we have

$$\begin{aligned} A(x) &= \left(\sum_{k=1}^{n} \boldsymbol{e}_k a_{k1} \cdots \sum_{k=1}^{n} \boldsymbol{e}_k a_{kn}\right) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (\boldsymbol{e}_1 \cdots \boldsymbol{e}_n) \begin{pmatrix} \sum_{l=1}^{n} a_{1l} x_l \\ \vdots \\ \sum_{l=1}^{n} a_{nl} x_l \end{pmatrix}. \\ &= \sum_{k=1}^{n} \sum_{l=1}^{n} \boldsymbol{e}_k a_{kl} x_l \\ &= \sum_{l=1}^{n} \left(\sum_{k=1}^{n} \boldsymbol{e}_k a_{kl}\right) x_l = \sum_{k=1}^{n} \left(\sum_{l=1}^{n} a_{kl} x_l\right) \boldsymbol{e}_k. \end{aligned} \tag{9.39}$$

That is, the above equation can be viewed in either of two ways, i.e., coordinate transformation with fix vectors or vector transformation with fixed coordinates.

Also, Eq. (9.37) can formally be written as

$$A(\boldsymbol{x}) = (\boldsymbol{e}_1 \cdots \boldsymbol{e}_n) A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (\boldsymbol{e}_1 A \cdots \boldsymbol{e}_n A) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \tag{9.40}$$

where we assumed that the distributive law holds with operation of $A$ on $(\boldsymbol{e}_1 \cdots \boldsymbol{e}_n)$. Meanwhile, if in (9.37) we put $x_i = 1, x_j = 0\,(j \neq i)$, from (9.39) we get

$$A(\boldsymbol{e}_i) = \sum_{k=1}^{n} \boldsymbol{e}_k a_{ki}.$$

Therefore, (9.39) can be rewritten as

$$A(\boldsymbol{x}) = \begin{pmatrix} A(\boldsymbol{e}_1) & \cdots & A(\boldsymbol{e}_n) \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}. \tag{9.41}$$

Since $x_i\,(1 \leq i \leq n)$ can arbitrarily be chosen, comparing (9.40) and (9.41) we have

$$\boldsymbol{e}_i A = A(\boldsymbol{e}_i)\,(1 \leq i \leq n). \tag{9.42}$$

The matrix representation is unique in reference to the same basis vectors. Suppose that there is another matrix representation of the transformation $A$ such that

$$A(\boldsymbol{x}) = (\boldsymbol{e}_1 \cdots \boldsymbol{e}_n) \begin{pmatrix} a'_{11} & \cdots & a'_{1n} \\ \vdots & \ddots & \vdots \\ a'_{n1} & \cdots & a'_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}. \tag{9.43}$$

Subtracting (9.43) from (9.37), we obtain

$$(\boldsymbol{e}_1 \cdots \boldsymbol{e}_n) \left[ \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} - \begin{pmatrix} a'_{11} & \cdots & a'_{1n} \\ \vdots & \ddots & \vdots \\ a'_{n1} & \cdots & a'_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right]$$
$$= (\boldsymbol{e}_1 \cdots \boldsymbol{e}_n) \begin{pmatrix} \sum_{k=1}^{n} (a_{1k} - a'_{1k}) x_k \\ \vdots \\ \sum_{k=1}^{n} (a_{nk} - a'_{nk}) x_k \end{pmatrix} = \boldsymbol{0}.$$

On the basis of the linear dependence of the basis vectors, $\sum_{k=1}^{n} (a_{ik} - a'_{ik}) x_k = 0 \, (1 \leq i \leq n)$. This relationship holds for any arbitrarily and independently chosen complex numbers $x_i \, (1 \leq i \leq n)$. Therefore, we must have $a_{ik} = a'_{ik} (1 \leq i, k \leq n)$, meaning that the matrix representation of $A$ is unique with regard to fixed basis vectors.

Nonetheless, if a set of vectors $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots$, and $\boldsymbol{e}_n$ does not constitute basis vectors (i.e., those vectors are linearly dependent), the aforementioned uniqueness of the matrix representation loses its meaning. For instance, in $V^2$ take vectors $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ such that $\boldsymbol{e}_1 = \boldsymbol{e}_2$ (i.e., the two vectors are linearly dependent) and let the transformation matrix be $B = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$. This means that $\boldsymbol{e}_1$ should be $\boldsymbol{e}_1$ after the transformation. At the same time, the vector $\boldsymbol{e}_2 (= \boldsymbol{e}_1)$ should be converted to $2\boldsymbol{e}_2 (= 2\boldsymbol{e}_1)$. It is impossible except for the case of $\boldsymbol{e}_1 = \boldsymbol{e}_2 = \boldsymbol{0}$. The above matrix $B$ in its own right is an object of matrix algebra, of course.

Putting $\boldsymbol{a} = \boldsymbol{b} = \boldsymbol{0}$ and $c = d = 1$ in the definition of (9.25) for the linear transformation, we obtain

$$A(\boldsymbol{0}) = A(\boldsymbol{0}) + A(\boldsymbol{0}).$$

Combining this relation with (9.4) gives

$$A(\boldsymbol{0}) = \boldsymbol{0}. \tag{9.44}$$

Then do we have a vector $\boldsymbol{u} \neq \boldsymbol{0}$ for which $A(\boldsymbol{u}) = \boldsymbol{0}$? An answer is yes. This is because if a (2,2)-matrix of $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ is chosen for $R$, we get a linear transformation such that

$$(\boldsymbol{e}_1'\boldsymbol{e}_2') = (\boldsymbol{e}_1\boldsymbol{e}_2)\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = (\boldsymbol{e}_1\boldsymbol{0}).$$

That is, we have $R(\boldsymbol{e}_2) = \boldsymbol{0}$.

In general, vectors $\boldsymbol{x}$ $(\in V)$ satisfying $A(\boldsymbol{x}) = \boldsymbol{0}$ form a subspace in a vector space $V$. This is because $A(\boldsymbol{x}) = \boldsymbol{0}$ and $A(\boldsymbol{y}) = \boldsymbol{0} \Rightarrow A(\boldsymbol{x}+\boldsymbol{y}) = A(\boldsymbol{x}) + A(\boldsymbol{y}) = \boldsymbol{0}$, $A(c\boldsymbol{x}) = cA(\boldsymbol{x}) = \boldsymbol{0}$. We call this subspace of a maximum dimension a *null-space* and represent it as Ker $A$, where Ker stands for "kernel." In other words, Ker $A = A^{-1}(\boldsymbol{0})$. Note that this symbolic notation does not ensure the existence of the inverse transformation $A^{-1}$ (vide infra) but represents a set comprising elements $\boldsymbol{x}$ that satisfy $A(\boldsymbol{x}) = \boldsymbol{0}$. In the above example, Ker $R$ = Span$\{\boldsymbol{e}_2\}$.

$A(V^n)$ (here $V^n$ is a vector space considered and we assume that $A$ is an endomorphism of $V^n$) also forms a subspace in $V^n$. In fact, for any $\boldsymbol{x}, \boldsymbol{y} \in V^n$, we have $A(\boldsymbol{x}), A(\boldsymbol{y}) \in A(V^n) \Rightarrow A(\boldsymbol{x}) + A(\boldsymbol{y}) = A(\boldsymbol{x}+\boldsymbol{y}) \in A(V^n); cA(\boldsymbol{x}) = A(c\boldsymbol{x}) \in A(V^n)$. Obviously, $A(V^n) \subset V^n$. The subspace $A(V^n)$ is said to be an image of the transformation $A$ and sometimes denoted by Im $A$. We have a so-called dimension theorem expressed as follows:

**Theorem 9.3** Dimension theorem *Let $V^n$ be a linear vector space of dimension $n$. Also, let $A$ be an endomorphism: $V^n \to V^n$. Then, we have*

$$\dim V^n = \dim A(V^n) + \dim \operatorname{Ker} A. \tag{9.45}$$

The number of dim $A(V^n)$ is said to be a rank of the linear transformation $A$. That is, we write

$$\dim A(V^n) = \operatorname{rank} A.$$

Also, the number of dim Ker $A$ is said to be a *nullity* of the linear transformation $A$. That is, we have

$$\dim \text{ Ker } A = \text{nullity } A.$$

Thus, (9.45) can be written succinctly as

$$\dim V^n = \operatorname{rank} A + \operatorname{nullity} A.$$

*Proof* Let $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots$, and $\boldsymbol{e}_n$ be basis vectors of $V^n$. First, assume

$$A(\boldsymbol{e}_1) = A(\boldsymbol{e}_2) = \cdots = A(\boldsymbol{e}_n) = \boldsymbol{0}.$$

This implies that nullity $A = n$. Then, $A\left(\sum_{i=1}^n x_i\boldsymbol{e}_i\right) = \sum_{i=1}^n x_i A(\boldsymbol{e}_i) = \boldsymbol{0}$. Since $x_i$ is arbitrarily chosen, the expression means that $A(\boldsymbol{x}) = \boldsymbol{0}$ for $^\forall\boldsymbol{x} \in V^n$. This implies $A = 0$. That is, rank $A = 0$. Thus, (9.45) certainly holds.

To proceed with proof of the theorem, we think of a linear combination $\sum_{i=1}^{n} c_i \boldsymbol{e}_i$. Next, assume that Ker $A = \text{Span}\{\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_v\} (v < n)$; dim Ker $A = v$. After $A$ is operated on the above linear combination, we are left with $\sum_{i=v+1}^{n} c_i A(\boldsymbol{e}_i)$. We put

$$\sum_{i=1}^{n} c_i A(\boldsymbol{e}_i) = \sum_{i=v+1}^{n} c_i A(\boldsymbol{e}_i) = \boldsymbol{0}. \tag{9.46}$$

Suppose that the $(n - v)$ vectors $A(\boldsymbol{e}_i)(v + 1 \leq i \leq n)$ are linearly dependent. Then without loss of generality, we can assume $c_{v+1} \neq 0$. Dividing (9.46) by $c_{v+1}$, we obtain

$$\begin{aligned}
&A(\boldsymbol{e}_{v+1}) + \frac{c_{v+2}}{C_{v+1}} A(\boldsymbol{e}_{v+2}) + \cdots + \frac{c_n}{C_{v+1}} A(\boldsymbol{e}_n) = \boldsymbol{0}, \\
&A(\boldsymbol{e}_{v+1}) + A(\frac{c_{v+2}}{C_{v+1}} \boldsymbol{e}_{v+2}) + \cdots + A(\frac{c_n}{C_{v+1}} \boldsymbol{e}_n) = \boldsymbol{0}, \\
&A(\boldsymbol{e}_{v+1} + \frac{c_{v+2}}{C_{v+1}} \boldsymbol{e}_{v+2} + \cdots + \frac{c_n}{C_{v+1}} \boldsymbol{e}_n) = \boldsymbol{0}.
\end{aligned}$$

Meanwhile, the $(v + 1)$ vectors $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_v$ and $\boldsymbol{e}_{v+1} + \frac{c_{v+2}}{C_{v+1}} \boldsymbol{e}_{v+2} + \cdots + \frac{c_n}{C_{v+1}} \boldsymbol{e}_n$ are linearly independent, because $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots,$ and $\boldsymbol{e}_n$ are basis vectors of $V^n$. This would imply that the dimension of Ker $A$ is $v + 1$, but this is in contradiction to Ker $A = \text{Span}\{\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_v\}$. Thus, the $(n - v)$ vectors $A(\boldsymbol{e}_i)(v + 1 \leq i \leq n)$ should be linearly independent.

Let $V^{n-v}$ be described as

$$V^{n-v} = \text{Span } \{A(\boldsymbol{e}_{v+1}), A(\boldsymbol{e}_{v+2}), \cdots, A(\boldsymbol{e}_n)\}. \tag{9.47}$$

Then, $V^{n-v}$ is a subspace of $V^n$, and so dim $A(V^n) \geq n - v = \dim V^{n-v}$. Meanwhile, from (9.46), we have

$$A\left(\sum_{i=v+1}^{n} c_i \boldsymbol{e}_i\right) = \boldsymbol{0}.$$

From the above discussion, however, this relation holds if and only if $c_{v+1} = \cdots = c_n = 0$. This implies that Ker $A \cap V^{n-v} = \{\boldsymbol{0}\}$. Then, we have

$$V^{n-v} + \text{Ker}\, A = V^{n-v} \oplus \text{Ker}\, A.$$

Meanwhile, from Theorem 9.2, we have

$$\dim[V^{n-v} \oplus \text{Ker}\, A] = \dim V^{n-v} + \dim \text{Ker}\, A = (n - v) + v = n = \dim V^n.$$

Thus, we must have dim $A(V^n) = n - v = \dim V^{n-v}$. Since $V^{n-v}$ is a subspace of $V^n$ and $V^{n-v} \subset A(V^n)$ from (9.47), $V^{n-v} = A(V^n)$.

To conclude, we get

$$\begin{aligned} &\dim A(V^n) + \dim \operatorname{Ker} A = \dim V^n, \\ &V^n = A(V^n) \oplus \operatorname{Ker} A. \end{aligned} \tag{9.48}$$

This completes the proof.

Comparing Theorem 9.3 with Theorem 9.2, we find that Theorem 9.3 is a special case of Theorem 9.2. Equations (9.45) and (9.48) play an important role in the theory of linear vector space.

As an exercise, we have a following example:

*Example 9.2* Let $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3, \boldsymbol{e}_4$ be basis vectors of $V^4$. Let $A$ be an endomorphism of $V^4$ and described by

$$A = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix}.$$

We have

$$(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3, \boldsymbol{e}_4) \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix} = (\boldsymbol{e}_1 + \boldsymbol{e}_3, \boldsymbol{e}_2 + \boldsymbol{e}_4, \boldsymbol{e}_1 + \boldsymbol{e}_2 + \boldsymbol{e}_3 + \boldsymbol{e}_4, \boldsymbol{0}).$$

That is,

$$\begin{aligned} &A(\boldsymbol{e}_1) = \boldsymbol{e}_1 + \boldsymbol{e}_3,\ A(\boldsymbol{e}_2) = \boldsymbol{e}_2 + \boldsymbol{e}_4,\ A(\boldsymbol{e}_3) = \boldsymbol{e}_1 + \boldsymbol{e}_2 + \boldsymbol{e}_3 + \boldsymbol{e}_4, \\ &A(\boldsymbol{e}_4) = \boldsymbol{0}. \end{aligned}$$

We have

$$A(-\boldsymbol{e}_1 - \boldsymbol{e}_2 + \boldsymbol{e}_3) = -A(\boldsymbol{e}_1) - A(\boldsymbol{e}_2) + A(\boldsymbol{e}_3) = \boldsymbol{0}.$$

Then, we find

$$A\left(V^4\right) = \operatorname{Span}\ \{\boldsymbol{e}_1 + \boldsymbol{e}_3, \boldsymbol{e}_2 + \boldsymbol{e}_4\} \quad \operatorname{Ker} A = \operatorname{Span}\{-\boldsymbol{e}_1 - \boldsymbol{e}_2 + \boldsymbol{e}_3, \boldsymbol{e}_4\}. \tag{9.49}$$

For any $\boldsymbol{x} \in V^n$, using scalar $c_i (1 \leq i \leq 4)$, we have

$$\begin{aligned} \boldsymbol{x} &= c_1\boldsymbol{e}_1 + c_2\boldsymbol{e}_2 + c_3\boldsymbol{e}_3 + c_2\boldsymbol{e}_4 \\ &= \frac{1}{2}(c_1 + c_3)(\boldsymbol{e}_1 + \boldsymbol{e}_3) + \frac{1}{2}(2c_2 + c_3 - c_1)(\boldsymbol{e}_2 + \boldsymbol{e}_4) \\ &\quad + \frac{1}{2}(c_3 - c_1)(-\boldsymbol{e}_1 - \boldsymbol{e}_2 + \boldsymbol{e}_3) + \frac{1}{2}(c_1 - 2c_2 - c_3 + 2c_4)\boldsymbol{e}_4. \end{aligned} \tag{9.50}$$

Thus, $\boldsymbol{x}$ has been uniquely represented as (9.50) with respect to basis vectors $(\boldsymbol{e}_1+\boldsymbol{e}_3)$, $(\boldsymbol{e}_2+\boldsymbol{e}_4)$, $(-\boldsymbol{e}_1-\boldsymbol{e}_2+\boldsymbol{e}_3)$, and $\boldsymbol{e}_4$. The linear independence of these vectors can easily be checked by equating (9.50) with zero. We also confirm that

$$V^4 = A\left(V^4\right) \oplus \text{Ker}\ \ A.$$

Linear transformation is a kind of mapping. Figure 9.3 depicts the concept of mapping. Suppose two sets of $X$ and $Y$. The mapping $f$ is a correspondence between an element $x(\in X)$ and $y(\in Y)$. The set $f(X)(\subset Y)$ is said to be a range of $f$. (i) The mapping $f$ is *injective*: If $x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2)$. (ii) The mapping is *surjective*: $f(X) = Y$. For $^{\forall}y \in Y$ corresponding element(s) $x \in X$ exist(s). (iii) The mapping is *bijective*: If the mapping $f$ is both injective and surjective, it is said to be bijective (or reversible mapping or invertible mapping). A mapping that is not invertible is said to be a non-invertible mapping.

If the mapping $f$ is bijective, a unique element $^{\exists}x \in X$ exists for $^{\forall}y \in Y$ such that $f(x) = y$. In terms of solving an equation, we say that with any given $y$ we can find a unique solution $x$ to the equation $f(x) = y$. In this case, $x$ is said to be an inverse element to $y$ and this is denoted by $x = f^{-1}(y)$. The mapping $f^{-1}$ is called an inverse mapping. If the linear transformation is relevant, the mapping is said to be an inverse transformation.

Here we focus on a case where both $X$ and $Y$ form a vector space and the mapping is an endomorphism. Regarding the linear transformation $A$: $V^n \to V^n$ (i.e., an endomorphism of $V^n$), we have a following important theorem:

**Theorem 9.4** *Let* $A$: $V^n \to V^n$ *be an endomorphism of* $V^n$. *A necessary and sufficient condition for the existence of an inverse transformation to A (i.e.,* $A^{-1}$*) is* $A^{-1}(\mathbf{0}) = \{\mathbf{0}\}$.

*Proof* Suppose $A^{-1}(\mathbf{0}) = \{\mathbf{0}\}$. Then $A(\boldsymbol{x}_1) = A(\boldsymbol{x}_2) \Leftrightarrow A(\boldsymbol{x}_1 - \boldsymbol{x}_2) = \mathbf{0} \Leftrightarrow \boldsymbol{x}_1 - \boldsymbol{x}_2 = \mathbf{0}$; i.e., $\boldsymbol{x}_1 = \boldsymbol{x}_2$. This implies that the transformation $A$ is injective. Other way round, suppose that $A$ is injective. If $A^{-1}(\mathbf{0}) \neq \{\mathbf{0}\}$, there should be $\boldsymbol{b}(\neq \mathbf{0})$ with which $A(\boldsymbol{b}) = \mathbf{0}$. This is, however, in contradiction to that $A$ is injective. Then, we must have $A^{-1}(\mathbf{0}) = \{\mathbf{0}\}$. Thus, $A^{-1}(\mathbf{0}) = \{\mathbf{0}\} \Leftrightarrow A$ is injective.

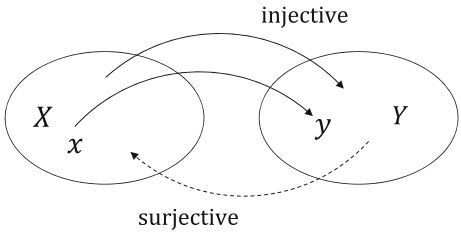


**Fig. 9.3** Concept of mapping from a set $X$ to another set $Y$

Meanwhile, $A^{-1}(\mathbf{0}) = \{\mathbf{0}\} \Leftrightarrow \dim A^{-1}(\mathbf{0}) = 0 \Leftrightarrow \dim A(V^n) = n$ (due to Theorem 9.3); i.e., $A(V^n) = V^n$. Then $A^{-1}(\mathbf{0}) = \{\mathbf{0}\} \Leftrightarrow A$ is surjective. Combining this with the above-mentioned statement, we have $A^{-1}(\mathbf{0}) = \{\mathbf{0}\} \Leftrightarrow A$ is bijective. This statement is equivalent to that an inverse transformation exists.

In the proof of Theorem 9.4, to show that $A^{-1}(\mathbf{0}) = \{\mathbf{0}\} \Leftrightarrow A$ is surjective, we have used the dimension theorem (Theorem 9.3), for which the relevant vector space is finite (i.e., $n$-dimensional). In other words, that $A$ is surjective is equivalent to that $A$ is injective with a finite-dimensional vector space, and vice versa. To conclude, so far as we are thinking of the endomorphism of a finite-dimensional vector space, if we can show it is either injective or surjective, the other necessarily follows and, hence, the mapping is bijective.

## 9.3 Inverse Matrices and Determinants

The existence of the inverse transformation plays a particularly important role in the theory of linear vector spaces. The inverse transformation is a linear transformation. Let $\boldsymbol{x}_1 = A^{-1}(\boldsymbol{y}_1), \boldsymbol{x}_2 = A^{-1}(\boldsymbol{y}_2)$. Also, we have $A(c_1\boldsymbol{x}_1 + c_2\boldsymbol{x}_2) = c_1A(\boldsymbol{x}_1) + c_2A(\boldsymbol{x}_2) = c_1\boldsymbol{y}_1 + c_2\boldsymbol{y}_2$. Thus, $c_1\boldsymbol{x}_1 + c_2\boldsymbol{x}_2 = A^{-1}(c_1\boldsymbol{y}_1 + c_2\boldsymbol{y}_2) = c_1A^{-1}(\boldsymbol{y}_1) + c_2A^{-1}(\boldsymbol{y}_2)$, showing that $A^{-1}$ is a linear transformation. As already mentioned, a matrix that represents a linear transformation $A$ is uniquely determined with respect to fixed basis vectors. This should be the case with $A^{-1}$ accordingly. We have an important theorem for this.

**Theorem 9.5 [1]** *The necessary and sufficient condition for the matrix* $A^{-1}$ *that represents the inverse transformation to A to exist is that* $\det A \neq 0$ (“*det*” *means a determinant*). *Here the matrix A represents the linear transformation A. The matrix* $A^{-1}$ *is uniquely determined and given by*

$$(A^{-1})_{ij} = (-1)^{i+j}(M)_{ji}/(\det A), \tag{9.51}$$

where $(M)_{ij}$ is the minor of $\det A$ corresponding to the element $A_{ij}$.

*Proof* First, we suppose that the matrix $A^{-1}$ exists so that it satisfies the following relation

$$\sum_{k=1}^{n} (A^{-1})_{ik}(A)_{kj} = \delta_{ij} (1 \leq i, j \leq n) \tag{9.52}$$

On this condition, suppose that $\det A = 0$. From the properties of determinants, this implies that one of the columns of $A$ (let it be the $m$th column) can be expressed as a linear combination of the other columns of $A$ such that

$$A_{km} = \sum_{j \neq m} A_{kj} c_j. \tag{9.53}$$

Putting $i = m$ in (9.52), multiplying by $c_j$, and summing over $j \neq m$, we get

$$\sum_{k=1}^{n} (A^{-1})_{mk} \sum_{j \neq m} A_{kj} c_j = \sum_{j \neq m} \delta_{mj} c_j = 0. \tag{9.54}$$

From (9.52) and (9.53), on the other hand, we obtain

$$\sum_{k=1}^{n} (A^{-1})_{mk} \sum_{j \neq m} A_{kj} c_j = \left\{ (A^{-1})_{ik} (A)_{km} \right\}_{i=m} = 1. \tag{9.55}$$

There is the inconsistency between (9.54) and (9.55). The inconsistency resulted from the supposition that the matrix $A^{-1}$ exists. Therefore, we conclude that if $\det A = 0$, $A^{-1}$ does not exist. Taking contraposition to the above statement, we say that if $A^{-1}$ exists, $\det A \neq 0$. Suppose next that $\det A \neq 0$. In this case, on the basis of the well-established result, a unique $A^{-1}$ exists and it is given by (9.51). This completes the proof.

Summarizing the characteristics of the endomorphism within a finite-dimensional vector space, we have

injective $\Leftrightarrow$ surjective $\Leftrightarrow$ bijective $\Leftrightarrow$ $\det A \neq 0$.

Let a matrix $A$ be

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}.$$

The determinant of a matrix$A$ is denoted by $\det A$ or by

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}.$$

The determinant is defined as

$$\det A \equiv \sum_{\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{pmatrix}} \varepsilon(\sigma) a_{1 i_1} a_{2 i_2} a_{n i_n}, \tag{9.56}$$

where $\sigma$ means permutation among 1, 2, …, $n$ and $\varepsilon(\sigma)$ denotes a sign of $+$ (in the case of even permutations) or $-$ (for odd permutations).

We deal with triangle matrices for future discussion. It is denoted by

$$T = \begin{pmatrix} a_{11} & & * \\ & \ddots & \\ 0 & & a_{nn} \end{pmatrix}, \tag{9.57}$$

where an asterisk (*) means that upper right off-diagonal elements can take any complex numbers (including zero). A large zero shows that all the lower left off-diagonal elements are zero. Its determinant is given by

$$\det T = a_{11}a_{22}\cdots a_{nn}. \tag{9.58}$$

In fact, focusing on $a_{ni_n}$, we notice that only if $i_n = n$, $a_{ni_n}$ does not vanish. Then, we get

$$\det A \equiv \sum_{\sigma=\begin{pmatrix} 1 & 2 & \cdots & n-1 \\ i_1 & i_2 & \cdots & i_{n-1} \end{pmatrix}} \varepsilon(\sigma) a_{1i_1} a_{2i_2} a_{n-1i_{n-1}} a_{nn}. \tag{9.59}$$

Repeating this process, we finally obtain (9.58).

The endomorphic linear transformation can be described succinctly as

$$A(\boldsymbol{x}) = \boldsymbol{y}, \tag{9.60}$$

where we have vectors such that $\boldsymbol{x} = \sum_{i=1}^{n} x_i \boldsymbol{e}_i$ and $\boldsymbol{y} = \sum_{i=1}^{n} y_i \boldsymbol{e}_i$. In reference to the same set of basis vectors $(\boldsymbol{e}_1 \cdots \boldsymbol{e}_n)$ and using a matrix representation, we have

$$A(\boldsymbol{x}) = (\boldsymbol{e}_1 \cdots \boldsymbol{e}_n) \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (\boldsymbol{e}_1 \cdots \boldsymbol{e}_n) \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}. \tag{9.61}$$

From the unique representation of a vector in reference to the basis vectors, (9.61) is simply expressed as

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}. \tag{9.62}$$

With a shorthand notation, we have

$$y_i = \sum_{k=1}^{n} a_{ik} x_k (1 \le i \le n). \tag{9.63}$$

From the above discussion, for the linear transformation $A$ to be bijective, $\det A \neq 0$. In terms of the system of linear equations, we say that for (9.62) to have a unique solution $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ for a given $\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$, we must have $\det A \neq 0$. Conversely, $\det A = 0$ is equivalent to that (9.62) has indefinite solutions or has no solution. As far as the matrix algebra is concerned, (9.61) is symbolically described by omitting a parenthesis as

$$A\boldsymbol{x} = \boldsymbol{y}. \tag{9.64}$$

However, when the vector transformation is explicitly taken into account, the full representation of (9.61) should be borne in mind.

The relations (9.60) and (9.64) can be considered as a set of simultaneous equations. A necessary and sufficient condition for (9.64) to have a unique solution $\boldsymbol{x}$ for a given $\boldsymbol{y}$ is $\det A \neq 0$. In that case, the solution $\boldsymbol{x}$ of (9.64) can be symbolically expressed as

$$\boldsymbol{x} = A^{-1}\boldsymbol{y}, \tag{9.65}$$

where $A^{-1}$ represents an inverse matrix of $A$.

*Example 9.3* Think of three-dimensional rotation by $\theta$ in $\mathbb{R}^3$ around the $z$-axis. The relevant transformation matrix is

$$R = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

As $\det R = 1 \neq 0$, the transformation is bijective. This means that for $^{\forall}\boldsymbol{y} \in \mathbb{R}^3$, there is always a corresponding $\boldsymbol{x} \in \mathbb{R}^3$. This $\boldsymbol{x}$ can be found by solving $R\boldsymbol{x} = \boldsymbol{y}$; i.e., $\boldsymbol{x} = R^{-1}\boldsymbol{y}$. Putting $\boldsymbol{x} = x\boldsymbol{e}_1 + y\boldsymbol{e}_2 + z\boldsymbol{e}_3$ and $\boldsymbol{y} = x'\boldsymbol{e}_1 + y'\boldsymbol{e}_2 + z'\boldsymbol{e}_3$, a matrix representation is given by

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = R^{-1}\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}.$$

Thus, $\boldsymbol{x}$ can be obtained by rotating $\boldsymbol{y}$ by $-\theta$.

*Example 9.4* Think of a following matrix that represents a linear transformation $P$:

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{9.66}$$

This matrix transforms a vector $x = x\boldsymbol{e}_1 + y\boldsymbol{e}_2 + z\boldsymbol{e}_3$ into $\boldsymbol{y} = x\boldsymbol{e}_1 + y\boldsymbol{e}_2$ as follows:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix}. \tag{9.67}$$

In this example, we are thinking of an endomorphism $P$: $\mathbb{R}^3 \to \mathbb{R}^3$. Geometrically, it can be viewed as in Fig. 9.4. Let us think of (9.67) from a point of view of solving a system of linear equations and newly consider the next equation. In other words, we are thinking of finding $x$, $y$, and $z$ with given $a$, $b$, and $c$ in (9.68).

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}. \tag{9.68}$$

If $c = 0$, we can readily find a solution of $x = a$, $y = b$, but $z$ can be any (complex) number; we have thus indefinite solutions. If $c \neq 0$, we have no solution. The former situation reflects the fact that the transformation represented by $P$ is not injective. Meanwhile, the latter reflects the fact that the transformation is not surjective. Remember that as $\det P = 0$, the transformation is not injective or surjective.

## 9.4 Basis Vectors and Their Transformations

In the previous sections, we show that a vector is uniquely represented as a column vector in reference to a set of the fixed basis vectors. The representation, however, will be changed under a different set of basis vectors.

First, let us think of a linear transformation of a set of basis vectors $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots$, and $\boldsymbol{e}_n$. The transformation matrix $A$ representing a linear transformation $A$ is defined as follows:

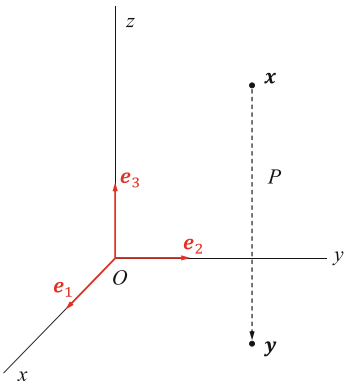


**Fig. 9.4** Example of an endomorphism $P$: $\mathbb{R}^3 \to \mathbb{R}^3$

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}.$$

Notice here that we often denote both a linear transformation and its corresponding matrix by the same character. After the transformation, suppose that the resulting vectors are given by $\boldsymbol{e}'_1, \boldsymbol{e}'_2$, and $\boldsymbol{e}'_n$. This is explicitly described as

$$(\boldsymbol{e}_1 \cdots \boldsymbol{e}_n) \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} = (\boldsymbol{e}'_1 \cdots \boldsymbol{e}'_n). \tag{9.69}$$

With a shorthand notation, we have

$$\boldsymbol{e}'_i = \sum_{k=1}^{n} a_{ki} \boldsymbol{e}_k (1 \leq i \leq n). \tag{9.70}$$

Care should be taken not to confuse (9.70) with (9.63). Here, a set of vectors $\boldsymbol{e}'_1, \boldsymbol{e}'_2$, and $\boldsymbol{e}'_n$ may or may not be linearly independent. Let us operate both sides from the left on $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ and equate both the sides to zero. That is,

$$(\boldsymbol{e}_1 \cdots \boldsymbol{e}_n) \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (\boldsymbol{e}'_1 \cdots \boldsymbol{e}'_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \mathbf{0}.$$

Since $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots$, and $\boldsymbol{e}_n$ are the basis vectors, we get

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = 0. \tag{9.71}$$

Meanwhile, we must have $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = 0$ so that $\boldsymbol{e}'_1, \boldsymbol{e}'_2$, and $\boldsymbol{e}'_n$ can be linearly independent (i.e., so as to be a set of basis vectors). But this means that (9.71) has such a unique (and trivial) solution and, hence, det $A \neq 0$. If conversely det $A \neq 0$, (9.71) has a unique trivial solution and $\boldsymbol{e}'_1, \boldsymbol{e}'_2$, and $\boldsymbol{e}'_n$ are linearly independent. Thus, a necessary and sufficient condition for $\boldsymbol{e}'_1, \boldsymbol{e}'_2$ and $\boldsymbol{e}'_n$ to be a set of basis vectors is det $A \neq 0$. If det $A = 0$, (9.71) has indefinite solutions (including a trivial

solution) and $\boldsymbol{e}'_1, \boldsymbol{e}'_2$, and $\boldsymbol{e}'_n$ are linearly dependent, and vice versa. In case $\det A \neq 0$, an inverse matrix $A^{-1}$ exists and so we have

$$(\boldsymbol{e}_1 \cdots \boldsymbol{e}_n) = \left(\boldsymbol{e}'_1 \cdots \boldsymbol{e}'_n\right)A^{-1}. \tag{9.72}$$

In the previous steps, we see how the linear transformation (and the corresponding matrix representation) converts a set of basis vectors $(\boldsymbol{e}_1 \cdots \boldsymbol{e}_n)$ to another set of basis vectors $\left(\boldsymbol{e}'_1 \cdots \boldsymbol{e}'_n\right)$. Is this possible then to find a suitable transformation between two sets of arbitrarily chosen basis vectors? The answer is yes. This is because any vector can be expressed uniquely by a linear combination of any set of basis vectors. A whole array of such linear combinations uniquely defines a transformation matrix between the two sets of basis vectors as expressed in (9.69) and (9.72). The matrix has nonzero determinant and has an inverse matrix. A concept of the transformation between basis vectors is important and very often used in various fields of natural science.

*Example 9.5* We revisit Example 9.2. The relation (9.49) tells us that the basis vectors of $A(V^4)$ and those of Ker$A$ span $V^4$ in total. Therefore, in light of the above argument, there should be a linear transformation $R$ between the two sets of vectors; i.e., $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3, \boldsymbol{e}_4$ and $\boldsymbol{e}_1 + \boldsymbol{e}_3, \boldsymbol{e}_2 + \boldsymbol{e}_4, -\boldsymbol{e}_1 - \boldsymbol{e}_2 + \boldsymbol{e}_3, \boldsymbol{e}_4$. Moreover, the matrix $R$ associated with the linear transformation must be non-singular (i.e., $\det R \neq 0$). In fact, we find that $R$ is expressed as

$$R = \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}.$$

This is because we have a following relation between the two sets of basis vectors:

$$(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3, \boldsymbol{e}_4)\begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} = (\boldsymbol{e}_1 + \boldsymbol{e}_3, \boldsymbol{e}_2 + \boldsymbol{e}_4, -\boldsymbol{e}_1 - \boldsymbol{e}_2 + \boldsymbol{e}_3, \boldsymbol{e}_4).$$

We have $\det R = 2 \neq 0$ as expected.

Next, let us consider successive linear transformations of vectors. Again, we assume that the transformations are endomorphism: $V^n \to V^n$. We have to take into account transformations of basis vectors along with the targeted vectors. First, we choose a transformation by a non-singular matrix (having a nonzero determinant) for the subsequent transformation to have a unique matrix representation (vide supra). The vector transformation by $P$ is expressed as

$$P(\boldsymbol{x}) = (\boldsymbol{e}_1 \cdots \boldsymbol{e}_n) \begin{pmatrix} p_{11} & \cdots & p_{1n} \\ \vdots & \ddots & \vdots \\ p_{n1} & \cdots & p_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \tag{9.73}$$

where the non-singular matrix $P$ represents the transformation $P$. Notice here that the transformation and its matrix are represented by the same $P$. As mentioned in Sect. 9.2, the matrix $P$ can be operated either from the right on the basis vectors or from the left on the column vector. We explicitly write

$$\begin{aligned} P(\boldsymbol{x}) &= \left[ (\boldsymbol{e}_1 \cdots \boldsymbol{e}_n) \begin{pmatrix} p_{11} & \cdots & p_{1n} \\ \vdots & \ddots & \vdots \\ p_{n1} & \cdots & p_{nn} \end{pmatrix} \right] \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \\ &= \left(\boldsymbol{e}'_1 \cdots \boldsymbol{e}'_n\right) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \end{aligned} \tag{9.74}$$

where $\left(\boldsymbol{e}'_1 \cdots \boldsymbol{e}'_n\right) = (\boldsymbol{e}_1 \cdots \boldsymbol{e}_n)P$ [here $P$ is the non-singular matrix defined in (9.73)]. Alternatively, we have

$$\begin{aligned} P(\boldsymbol{x}) &= (\boldsymbol{e}_1 \cdots \boldsymbol{e}_n) \left[ \begin{pmatrix} p_{11} & \cdots & p_{1n} \\ \vdots & \ddots & \vdots \\ p_{n1} & \cdots & p_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right]. \\ &= (\boldsymbol{e}_1 \cdots \boldsymbol{e}_n) \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}, \end{aligned} \tag{9.75}$$

where

$$\begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix} = \begin{pmatrix} p_{11} & \cdots & p_{1n} \\ \vdots & \ddots & \vdots \\ p_{n1} & \cdots & p_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}. \tag{9.76}$$

Equation (9.75) gives a column vector representation regarding the vector that has been obtained by the transformation $P$ and is viewed in reference to the basis vectors $(\boldsymbol{e}_1 \cdots \boldsymbol{e}_n)$. Combining (9.74) and (9.75), we get

$$P(\boldsymbol{x}) = \left(\boldsymbol{e}'_1 \cdots \boldsymbol{e}'_n\right) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (\boldsymbol{e}_1 \cdots \boldsymbol{e}_n) \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}. \tag{9.77}$$

We further make another linear transformation $A : V^n \to V^n$. In this case, a corresponding matrix $A$ may be non-singular (i.e., $\det A \neq 0$) or singular ($\det A = 0$). We have to distinguish the matrix representations of the two cases. It is because the matrix representations are uniquely defined in reference to an individual set of basis vectors; see (9.37) and (9.43). Let us denote the matrices $A_O$ and $A'$ with respect to the basis vectors $(\boldsymbol{e}_1 \cdots \boldsymbol{e}_n)$ and $\left(\boldsymbol{e}'_1 \cdots \boldsymbol{e}'_n\right)$, respectively. Then, $A[P(\boldsymbol{x})]$ can be described in two different ways as follows:

$$A[P(\boldsymbol{x})] = \left(\boldsymbol{e}'_1 \cdots \boldsymbol{e}'_n\right) A' \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (\boldsymbol{e}_1 \cdots \boldsymbol{e}_n) A_O \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}. \tag{9.78}$$

This can be rewritten in reference to a linearly independent set of vectors $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$ as

$$A[P(\boldsymbol{x})] = [(\boldsymbol{e}_1 \cdots \boldsymbol{e}_n) PA'] \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (\boldsymbol{e}_1 \cdots \boldsymbol{e}_n) \left[ A_O P \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right]. \tag{9.79}$$

As (9.79) is described for a vector $\boldsymbol{x} = \sum_{i=1}^{n} x_i \boldsymbol{e}_i$ arbitrarily chosen in $V^n$, we get

$$PA' = A_O P. \tag{9.80}$$

Since $P$ is non-singular, we finally obtain

$$A' = P^{-1} A_O P. \tag{9.81}$$

We can see (9.79) from a point of view of successive linear transformations. When the subsequent operation is viewed in reference to the basis vectors $\boldsymbol{e}'_1, \cdots, \boldsymbol{e}'_n$ newly reached by the precedent transformation, we make it a rule to write the relevant subsequent operator $A'$ *from the right*. In the case, where the subsequent operation is viewed in reference to the original basis vectors, on the other hand, we write the subsequent operator $A_O$ *from the left*. Further, discussion and examples can be seen in Part IV.

We see (9.81) in a different manner. Suppose we have

$$A(\boldsymbol{x}) = (\boldsymbol{e}_1 \cdots \boldsymbol{e}_n) A_O \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}. \tag{9.82}$$

Note that since the transformation $A$ has been performed in reference to the basis vectors $(\boldsymbol{e}_1 \cdots \boldsymbol{e}_n)$, $A_O$ should be used for the matrix representation. This is rewritten as

$$A(\boldsymbol{x}) = (\boldsymbol{e}_1 \cdots \boldsymbol{e}_n) P P^{-1} A_O P P^{-1} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}. \tag{9.83}$$

Meanwhile, any vector$\boldsymbol{x}$ in $V^n$ can be written as

$$\begin{aligned} \boldsymbol{x} &= (\boldsymbol{e}_1 \cdots \boldsymbol{e}_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \\ &= (\boldsymbol{e}_1 \cdots \boldsymbol{e}_n) P P^{-1} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (\boldsymbol{e}'_1 \cdots \boldsymbol{e}'_n) P^{-1} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (\boldsymbol{e}'_1 \cdots \boldsymbol{e}'_n) \begin{pmatrix} \widetilde{x}_1 \\ \vdots \\ \widetilde{x}_n \end{pmatrix} \end{aligned} \tag{9.84}$$

In (9.84), we put

$$(\boldsymbol{e}_1 \cdots \boldsymbol{e}_n) P = (\boldsymbol{e}'_1 \cdots \boldsymbol{e}'_n), \; P^{-1} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \widetilde{x}_1 \\ \vdots \\ \widetilde{x}_n \end{pmatrix}. \tag{9.85}$$

Equation (9.84) give a column vector representation regarding the *same* vector $\boldsymbol{x}$ viewed in reference to the basis set of vectors $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$ or $\boldsymbol{e}'_1, \cdots, \boldsymbol{e}'_n$. Equation (9.85) should not be confused with (9.76). That is, (9.76) relates the two coordinates $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ and $\begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$ of *different vectors* before and after the transformation viewed in reference to *the same set of basis vectors*. The relation (9.85), on the other hand, relates two coordinates $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ and $\begin{pmatrix} \widetilde{x}_1 \\ \vdots \\ \widetilde{x}_n \end{pmatrix}$ of *the same vector* viewed in reference to *different set of basis vectors*. Thus, from (9.83), we have

$$A(\boldsymbol{x}) = (\boldsymbol{e}'_1 \cdots \boldsymbol{e}'_n) P^{-1} A_O P \begin{pmatrix} \widetilde{x}_1 \\ \vdots \\ \widetilde{x}_n \end{pmatrix}. \tag{9.86}$$

Meanwhile, viewing the transformation $A$ in reference to the basis vectors $(\boldsymbol{e}'_1 \cdots \boldsymbol{e}'_n)$, we have

$$A(\boldsymbol{x}) = (\boldsymbol{e}'_1 \cdots \boldsymbol{e}'_n)A' \begin{pmatrix} \widetilde{x}_1 \\ \vdots \\ \widetilde{x}_n \end{pmatrix}. \tag{9.87}$$

Equating (9.86) and (9.87),

$$A' = P^{-1}A_O P. \tag{9.88}$$

Thus, (9.81) is recovered.

The relations expressed by (9.81) and (9.88) are called a similarity transformation on $A$. The matrices $A_0$ and $A'$ are said to be similar to each other. As mentioned earlier, if $A_0$ (and hence $A'$) is non-singular, the linear transformation $A$ produces a set of basis vectors other than $\boldsymbol{e}'_1, \cdots, \boldsymbol{e}'_n$, say $\boldsymbol{e}''_1, \cdots, \boldsymbol{e}''_n$. We write this symbolically as

$$(\boldsymbol{e}_1 \cdots \boldsymbol{e}_n)PA = (\boldsymbol{e}'_1 \cdots \boldsymbol{e}'_n)A = (\boldsymbol{e}''_1 \cdots \boldsymbol{e}''_n). \tag{9.89}$$

Therefore, such $A$ defines successive transformations of the basis vectors in conjunction with $P$ defined in (9.73). The successive transformations and resulting basis vectors supply us with important applications in the field of group theory. Topics will be dealt with in Part IV.

## Reference

1. Dennery P, Krzywicki A (1996) Mathematics for physicists. Dover, New York

# Chapter 10
# Canonical Forms of Matrices

In Sect. 9.4, we saw that the transformation matrices are altered depending on the basis vectors we choose. Then a following question arises. Can we convert a (transformation) matrix to as simple a form as possible by similarity transformation (s)? In Sect. 9.4, we have also shown that if we have two sets of basis vectors in a linear vector space $V^n$ we can always find a non-singular transformation matrix between the two. In conjunction with the transformation of the basis vectors, the matrix undergoes similarity transformation. It is our task in this chapter to find a simple form or a specific form (i.e., canonical form) of a matrix as a result of the similarity transformation. For this purpose, we should first find eigenvalue(s) and corresponding eigenvector(s) of the matrix. Depending upon the nature of matrices, we get various canonical forms of matrices such as a triangle matrix and a diagonal matrix. Regarding any form of matrices, we can treat these matrices under a unified form called the Jordan canonical form.

## 10.1 Eigenvalues and Eigenvectors

An eigenvalue problem is one of important subjects of the theory of linear vector spaces. Let $A$ be a linear transformation on $V^n$. The resulting matrix gets to several different kinds of canonical forms of matrices. A typical example is a diagonal matrix. To reach a satisfactory answer, we start with so-called an eigenvalue problem.

Suppose that after the transformation of $\boldsymbol{x}$ we have

$$A(\boldsymbol{x}) = \alpha \boldsymbol{x}, \tag{10.1}$$

where $\alpha$ is a certain (complex) number. Then we say that $\alpha$ is an eigenvalue and that $\boldsymbol{x}$ is an eigenvector that corresponds to the eigenvalue $\alpha$. Using a notation of (9.37) of Sect. 9.2, we have



S. Hotta, *Mathematical Physical Chemistry*,
https://doi.org/10.1007/978-981-10-7671-8_10

$$A(\boldsymbol{x}) = (\boldsymbol{e}_1 \ldots \boldsymbol{e}_n) \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \alpha(\boldsymbol{e}_1 \ldots \boldsymbol{e}_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

From linear dependence of $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots,$ and $\boldsymbol{e}_n$, we simply write

$$\begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \alpha \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

If we identify $\boldsymbol{x}$ with $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ at fixed basis vectors $(\boldsymbol{e}_1 \ldots \boldsymbol{e}_n)$, we may naturally rewrite (10.1) as

$$A\boldsymbol{x} = \alpha\boldsymbol{x}. \tag{10.2}$$

If $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ belong to the eigenvalue $\alpha$, so does $\boldsymbol{x}_1 + \boldsymbol{x}_2$ and $c\boldsymbol{x}_1$ ($c$ is an appropriate complex number). Therefore, all the eigenvectors belonging to the eigenvalue $\alpha$ along with $\mathbf{0}$ (a zero vector) form a subspace of $A$ (within $V^n$) corresponding to the eigenvalue $\alpha$.

Strictly speaking, we should use terminologies such as a "proper" (or ordinary) eigenvalue, eigenvector, eigenspace to distinguish them from a "generalized" eigenvalue, eigenvector, eigenspace, etc. We will return to this point later. Further rewriting (10.2), we have

$$(A - \alpha E)\boldsymbol{x} = \mathbf{0}, \tag{10.3}$$

where $E$ is a $(n, n)$ unit matrix. Equations (10.2) and (10.3) are said to be an eigenvalue equation (or eigenequation). In (10.2) or (10.3) $\boldsymbol{x} = \mathbf{0}$ always holds (as a trivial solution). Consequently, for $\boldsymbol{x} \neq \mathbf{0}$ to be a solution we must have

$$|A - \alpha E| = 0, \tag{10.4}$$

In (10.4), $|A - \alpha E|$ stands for $\det(A - \alpha E)$.

Now let us define the following polynomial:

$$f_A(x) = |xE - A| = \begin{vmatrix} x - a_{11} & \ldots & -a_{1n} \\ \vdots & \ddots & \vdots \\ -a_{n1} & \ldots & x - a_{nn} \end{vmatrix}. \tag{10.5}$$

A necessary and sufficient condition for $\alpha$ to be an eigenvalue is that $\alpha$ is a root of $f_A(x) = 0$. The function $f_A(x)$ is said to be a characteristic polynomial and we call $f_A(x) = 0$ a characteristic equation. This is an $n$th order polynomial. Putting

$$f_A(x) = x^n + a_1 x^{n-1} + \cdots + a_n, \tag{10.6}$$

we have

$$a_1 = -(a_{11} + a_{22} + \cdots + a_{nn}) \equiv -\mathrm{tr}A, \tag{10.7}$$

where tr stands for "trace" that is a summation of diagonal elements. Moreover,

$$a_n = (-1)^n |A|. \tag{10.8}$$

The characteristic equation $f_A(x) = 0$ has $n$ roots including possible multiple roots. Let those roots be $\alpha_1, \ldots, \alpha_n$ (some of them may be identical). Then we have

$$f_A(x) = \prod_{i=1}^{n} (x - \alpha_i). \tag{10.9}$$

Furthermore, according to relations between roots and coefficients we get

$$\alpha_1 + \cdots + \alpha_n = -a_1 = \mathrm{tr}A, \tag{10.10}$$

$$\alpha_1 \ldots \alpha_n = (-1)^n a_n = |A|. \tag{10.11}$$

The characteristic equation $f_A(x)$ is invariant under a similarity transformation. In fact,

$$\begin{aligned} f_{P^{-1}AP}(x) &= \left|xE - P^{-1}AP\right| = \left|P^{-1}\,(xE - A)P\right| \\ &= |P|^{-1}|xE - A||P| = |xE - A| = f_A(x). \end{aligned} \tag{10.12}$$

This leads to invariance of the trace under a similarity transformation. That is,

$$\mathrm{tr}\left(P^{-1}AP\right) = \mathrm{tr}A. \tag{10.13}$$

Let us think of a following triangle matrix:

$$T = \begin{pmatrix} a_{11} & & * \\ & \ddots & \\ 0 & & a_{nn} \end{pmatrix}, \tag{10.14}$$

The matrix of this type is thought to be one of canonical forms of matrices. Its characteristic equation $f_T(x)$ is

$$f_T(x) = |xE - T| = \begin{vmatrix} x - a_{11} & * & * \\ 0 & \ddots & * \\ 0 & 0 & x - a_{nn} \end{vmatrix}. \tag{10.15}$$

Therefore, we get

$$f_T(x) = \prod_{i=1}^{n} (x - a_{ii}), \tag{10.16}$$

where we used (9.58). From (10.14) and (10.16), we find that eigenvalues of a triangle matrix are given by its diagonal elements accordingly. Our immediate task will be to examine whether and how a given matrix is converted to a triangle matrix through a similarity transformation. A following theorem is important.

**Theorem 10.1** *Every* $(n, n)$ *square matrix can be converted to a triangle matrix by similarity transformation* [1].

*Proof* A triangle matrix is either an "upper" triangle matrix [to which all the lower left off-diagonal elements are zero; see (10.14)] or a "lower" triangle matrix (to which all the upper right off-diagonal elements are zero). In the present case, we show the proof for the upper triangle matrix. Regarding the lower triangle matrix, the theorem is proven in a similar manner.

The proof is due to mathematical induction. First we show that the theorem is true of a (2, 2) matrix. Suppose that one of eigenvalues of $A_2$ is $\alpha_1$ and that its corresponding eigenvector is $\boldsymbol{x}_1$. Then we have

$$A_2 \boldsymbol{x}_1 = \alpha_1 \boldsymbol{x}_1, \tag{10.17}$$

where we assume that $\boldsymbol{x}_1$ represents a column vector. Let a non-singular matrix be

$$P_1 = (\boldsymbol{x}_1 \, \boldsymbol{p}_1), \tag{10.18}$$

where $\boldsymbol{p}_1$ is another column vector chosen in such a way that $\boldsymbol{x}_1$ and $\boldsymbol{p}_1$ are linearly independent so that $P_1$ can be a non-singular matrix. Then, we have

$$P_1^{-1} A_2 P_1 = (P_1^{-1} A_2 \boldsymbol{x}_1 P_1^{-1} A_2 \boldsymbol{p}_1). \tag{10.19}$$

Meanwhile,

$$\boldsymbol{x}_1 = (\boldsymbol{x}_1 \, \boldsymbol{p}_1) \begin{pmatrix} 1 \\ 0 \end{pmatrix} = P_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Hence, we have

$$P_1^{-1}A_2\boldsymbol{x}_1 = \alpha_1 P_1^{-1}\boldsymbol{x}_1 = \alpha_1 P_1^{-1}P_1\begin{pmatrix}1\\0\end{pmatrix} = \alpha_1\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}\alpha_1\\0\end{pmatrix}. \tag{10.20}$$

Thus (10.19) can be rewritten as

$$P_1^{-1}A_2P_1 = \begin{pmatrix}\alpha_1 & *\\ 0 & *\end{pmatrix}. \tag{10.21}$$

This shows that Theorem 10.1 is true of a (2, 2) matrix $A_2$.

Now let us show that Theorem 10.1 holds in a general case, i.e., for a ($n$, $n$) square matrix $A_n$. Let $\alpha_n$ be one of eigenvalues of $A_n$. On the basis of the argument of the (2, 2) matrix case, suppose that after a suitable similarity transformation by a non-singular matrix $\widetilde{P}$ we have

$$\widetilde{A_n} = (\widetilde{P})^{-1}A_n\widetilde{P}. \tag{10.22}$$

Then, we can describe $\widetilde{A_n}$ as

$$\widetilde{A_n} = \begin{pmatrix}\alpha_n & x_1 & \cdots & x_{n-1}\\ 0 & & & \\ \vdots & & A_{n-1} & \\ 0 & & & \end{pmatrix}, \tag{10.23}$$

In (8.23), $\alpha_n$ is one of eigenvalues of $A_n$. To show that (10.23) is valid, we have a similar argument as in the case of a (2, 2) matrix. That is, we set $\widetilde{P}$ such that

$$\widetilde{P} = (\boldsymbol{a}_n\, \boldsymbol{p}_1\, \boldsymbol{p}_2 \ldots \boldsymbol{p}_{n-1}),$$

where $\boldsymbol{a}_n$ is an eigenvector corresponding to $\alpha_n$ and $\widetilde{P}$ is a non-singular matrix formed by $n$ linearly independent column vectors $\boldsymbol{a}_n$, $\boldsymbol{p}_1$, $\boldsymbol{p}_2$, ... and $\boldsymbol{p}_{n-1}$. Then we have

$$\begin{aligned}\widetilde{A_n} &= (\widetilde{P})^{-1}A_n\widetilde{P}\\ &= \left((\widetilde{P})^{-1}A_n\boldsymbol{a}_n\,(\widetilde{P})^{-1}A_n\boldsymbol{p}_1\,(\widetilde{P})^{-1}A_n\boldsymbol{p}_2 \ldots (\widetilde{P})^{-1}A_n\boldsymbol{p}_{n-1}\right).\end{aligned} \tag{10.24}$$

The vector $\boldsymbol{a}_n$ can be expressed as

$$\boldsymbol{a}_n = (\boldsymbol{a}_n\, \boldsymbol{p}_1\, \boldsymbol{p}_2 \ldots \boldsymbol{p}_{n-1})\begin{pmatrix}1\\0\\0\\\vdots\\0\end{pmatrix} = \widetilde{P}\begin{pmatrix}1\\0\\0\\\vdots\\0\end{pmatrix}.$$

Therefore, we have

$$(\widetilde{P})^{-1}A_n\boldsymbol{a}_n = (\widetilde{P})^{-1}\alpha_n\boldsymbol{a}_n = (\widetilde{P})^{-1}\alpha_n\widetilde{P}\begin{pmatrix}1\\0\\0\\\vdots\\0\end{pmatrix} = \alpha_n\begin{pmatrix}1\\0\\0\\\vdots\\0\end{pmatrix} = \begin{pmatrix}\alpha_n\\0\\0\\\vdots\\0\end{pmatrix}.$$

Thus, from (10.24) we see that one can express a matrix form of $\widetilde{A_n}$ as (10.23).

By a hypothesis of the mathematical induction, we assume that there exists a $(n-1, n-1)$ non-singular square matrix $P_{n-1}$ and an upper triangle matrix $\Delta_{n-1}$ such that

$$P_{n-1}^{-1}A_{n-1}P_{n-1} = \Delta_{n-1}. \tag{10.25}$$

Let us define a following matrix

$$P_n = \begin{pmatrix} d & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & P_{n-1} & \\ 0 & & & \end{pmatrix},$$

where $d \neq 0$. The $P_n$ is $(n, n)$ non-singular square matrix; remember that $\det P_n = d(\det P_{n-1}) \neq 0$. Operating $P_n$ on $\widetilde{A_n}$ from the right, we have

$$\begin{pmatrix} \alpha_n & x_1 & \cdots & x_{n-1} \\ 0 & & & \\ \vdots & & A_{n-1} & \\ 0 & & & \end{pmatrix}\begin{pmatrix} d & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & P_{n-1} & \\ 0 & & & \end{pmatrix} = \begin{pmatrix} \alpha_n d & \boldsymbol{x}_{n-1}^T P_{n-1} \\ 0 & \\ \vdots & A_{n-1}P_{n-1} \\ 0 & \end{pmatrix}, \tag{10.26}$$

where $\boldsymbol{x}_{n-1}^T$ is a transpose of a column vector $\begin{pmatrix}x_1\\\vdots\\x_{n-1}\end{pmatrix}$. Therefore, $\boldsymbol{x}_{n-1}^T P_{n-1}$ is a $(1, n-1)$ matrix (i.e., a row vector). Meanwhile, we have

$$\begin{pmatrix} d & 0 & 0 & 0 \\ 0 & & & \\ \vdots & & P_{n-1} & \\ 0 & & & \end{pmatrix}\begin{pmatrix} \alpha_n & \boldsymbol{x}_{n-1}^T P_{n-1}/d \\ 0 & \\ \vdots & \Delta_{n-1} \\ 0 & \end{pmatrix} = \begin{pmatrix} \alpha_n d & \boldsymbol{x}_{n-1}^T P_{n-1} \\ 0 & \\ \vdots & P_{n-1}\Delta_{n-1} \\ 0 & \end{pmatrix}. \tag{10.27}$$

From the assumption of (10.25), we have

$$\text{LHS of } (10.26) = \text{LHS of } (10.27). \tag{10.28}$$

That is,

$$\widetilde{A_n} P_n = P_n \Delta_n. \tag{10.29}$$

In (10.29), we define $\Delta_n$ such that

$$\Delta_n = \begin{pmatrix} \alpha_n & \boldsymbol{x}_{n-1}^T P_{n-1}/d \\ 0 & \\ \vdots & \Delta_{n-1} \\ 0 & \end{pmatrix}, \tag{10.30}$$

which has appeared in LHS of (10.27). As $\Delta_{n-1}$ is a triangle matrix from the assumption, $\Delta_n$ is a triangle matrix as well. Combining (10.22) and (10.29), we finally get

$$\Delta_n = (\widetilde{P} P_n)^{-1} A_n \widetilde{P} P_n. \tag{10.31}$$

Notice that $\widetilde{P} P_n$ is a non-singular matrix, and so $\widetilde{P} P_n P_n^{-1} \widetilde{P}^{-1} = E$. Hence, $P_n^{-1} \widetilde{P}^{-1} = (\widetilde{P} P_n)^{-1}$. The equation obviously shows that $A_n$ has been converted to a triangle matrix $\Delta_n$. This completes the proof.

Equation (10.31) implies that eigenvectors are disposed on diagonal positions of a triangle matrix. Triangle matrices can further undergo a similarity transformation.

*Example 10.1* Let us think of a following triangle matrix $A$:

$$A = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}. \tag{10.32}$$

Eigenvalues of $A$ are 2 and 1. Remember that diagonal elements of a triangle matrix give eigenvalues. According to (10.20), a vector $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ can be chosen for an eigenvector (as a column vector representation) corresponding to the eigenvalue 2. Another eigenvector can be chosen to be $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$. This is because for an eigenvalue 1, we obtain

$$\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \boldsymbol{x} = \boldsymbol{0}$$

as an eigenvalue equation $(A - E)\boldsymbol{x} = \boldsymbol{0}$. Therefore, with an eigenvector $\begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$ corresponding to the eigenvalue 1, we get $c_1 + c_2 = 0$. Therefore, we have $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$ as a simple form of the eigenvector. Hence, putting $P = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$, similarity transformation is carried out such that

$$P^{-1}AP = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}. \tag{10.33}$$

This is a simple example of matrix diagonalization.

Regarding the eigenvalue/eigenvector problems, we have another important theorem.

**Theorem 10.2** *Eigenvectors corresponding to different eigenvalues of A are linearly independent.*

*Proof* We prove the theorem by mathematical induction. Let $\alpha_1$ and $\alpha_2$ be two different eigenvalues of a matrix $A$ and let $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ be eigenvectors corresponding to $\alpha_1$ and $\alpha_2$, respectively.

Let us think of a following equation:

$$c_1\boldsymbol{a}_1 + c_2\boldsymbol{a}_2 = \boldsymbol{0}. \tag{10.34}$$

Suppose that $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ are linearly dependent. Then, without loss of generality we can put $c_1 \neq 0$. Accordingly, we get

$$\boldsymbol{a}_1 = -\frac{c_2}{c_1}\boldsymbol{a}_2. \tag{10.35}$$

Operating $A$ from the left of (10.35), we have

$$\alpha_1\boldsymbol{a}_1 = -\frac{c_2}{c_1}\alpha_2\boldsymbol{a}_2 = \alpha_2\boldsymbol{a}_1. \tag{10.36}$$

With the second equality, we have used (10.35). From (10.36), we have

$$(\alpha_1 - \alpha_2)\boldsymbol{a}_1 = \boldsymbol{0}. \tag{10.37}$$

As $\alpha_1 \neq \alpha_2$, $\alpha_1 - \alpha_2 \neq 0$. This implies $\boldsymbol{a}_1 = \boldsymbol{0}$, in contradiction to that $\boldsymbol{a}_1$ is a eigenvector. Thus, $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ must be linearly independent.

Next we assume that Theorem 10.2 is true of the case where we have $(n-1)$ eigenvalues $\alpha_1, \alpha_2, \ldots,$ and $\alpha_{n-1}$ that are different from one another and corresponding eigenvectors $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots,$ and $\boldsymbol{a}_{n-1}$ are linearly independent. Let us think of a following equation:

$$c_1\boldsymbol{a}_1 + c_2\boldsymbol{a}_2 + \cdots + c_{n-1}\boldsymbol{a}_{n-1} + c_n\boldsymbol{a}_n = \boldsymbol{0}, \tag{10.38}$$

where $\boldsymbol{a}_n$ is an eigenvector corresponding to an eigenvalue $\alpha_n$. Suppose here that $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots,$ and $\boldsymbol{a}_n$ are linearly dependent. If $c_n = 0$, we have

$$c_1\boldsymbol{a}_1 + c_2\boldsymbol{a}_2 + \cdots + c_{n-1}\boldsymbol{a}_{n-1} = \boldsymbol{0}. \tag{10.39}$$

But from the linear independence of $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots,$ and $\boldsymbol{a}_{n-1}$, we have $c_1 = c_2 = \cdots = c_{n-1} = 0$. Thus, it follows that all the eigenvectors $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots,$ and $\boldsymbol{a}_n$ are linearly independent. However, this is in contradiction to the assumption. We should therefore have $c_n \neq 0$. Accordingly, we get

$$\boldsymbol{a}_n = -\left(\frac{c_1}{c_n}\boldsymbol{a}_1 + \frac{c_2}{c_n}\boldsymbol{a}_2 + \cdots + \frac{c_{n-1}}{c_n}\boldsymbol{a}_{n-1}\right). \tag{10.40}$$

Operating $A$ from the left of (10.38) again, we have

$$\alpha_n\boldsymbol{a}_n = -\left(\frac{c_1}{c_n}\alpha_1\boldsymbol{a}_1 + \frac{c_2}{c_n}\alpha_2\boldsymbol{a}_2 + \cdots + \frac{c_{n-1}}{c_n}\alpha_{n-1}\boldsymbol{a}_{n-1}\right). \tag{10.41}$$

Here we think of two cases of (i) $\alpha_n \neq 0$ and (ii) $\alpha_n = 0$.

(i) Case I: $\alpha_n \neq 0$.

Multiplying both sides of (10.40) by $\alpha_n$ we have

$$\alpha_n a_n = -\left(\frac{c_1}{c_n}\alpha_n\boldsymbol{a_1} + \frac{c_2}{c_n}\alpha_n\boldsymbol{a_2} + \cdots + \frac{c_{n-1}}{c_n}\alpha_n\boldsymbol{a_{n-1}}\right). \tag{10.42}$$

Subtracting (10.42) from (10.41), we get

$$\boldsymbol{0} = \frac{c_1}{c_n}(\alpha_n - \alpha_1)\boldsymbol{a}_1 + \frac{c_2}{c_n}(\alpha_n - \alpha_2)\boldsymbol{a}_2 + \cdots + \frac{c_{n-1}}{c_n}(\alpha_n - \alpha_{n-1})\boldsymbol{a}_{n-1}. \tag{10.43}$$

Since we assume that eigenvalue are different from one another, $\alpha_n \neq \alpha_1$, $\alpha_n \neq \alpha_2, \ldots, \alpha_n \neq \alpha_{n-1}$. This implies that $c_1 = c_2 = \cdots = c_{n-1} = 0$. From (10.40), we have $\boldsymbol{a}_n = \boldsymbol{0}$. This is, however, in contradiction to that $\boldsymbol{a}_n$ is an eigenvector. This means that our original supposition that $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots,$ and $\boldsymbol{a}_n$ are linearly dependent was wrong. Thus, the eigenvectors $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots,$ and $\boldsymbol{a}_n$ should be linearly independent.

(ii) Case II: $\alpha_n = 0$.

Suppose again that $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots,$ and $\boldsymbol{a}_n$ are linearly dependent. Since as before $c_n \neq 0$, we get (10.40) and (10.41) once again. Putting $\alpha_n = 0$ in (10.41) we have

$$\boldsymbol{0} = -\left(\frac{c_1}{c_n}\alpha_1\boldsymbol{a}_1 + \frac{c_2}{c_n}\alpha_2\boldsymbol{a}_2 + \cdots + \frac{c_{n-1}}{c_n}\alpha_{n-1}\boldsymbol{a}_{n-1}\right). \tag{10.44}$$

Since eigenvalues are different, we should have $\alpha_1 \neq 0$, $\alpha_2 \neq 0, \ldots,$ and $\alpha_{n-1} \neq 0$. Then, considering that $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots,$ and $\boldsymbol{a}_{n-1}$ are linearly independent, for (10.44) to hold we must have $c_1 = c_2 = \cdots = c_{n-1} = 0$. But from (10.40), we have $\boldsymbol{a}_n = \boldsymbol{0}$, again in contradiction to that $\boldsymbol{a}_n$ is an eigenvector. Thus, the eigenvectors $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots,$ and $\boldsymbol{a}_n$ should be linearly independent. These complete the proof.

## 10.2 Eigenspaces and Invariant Subspaces

Equations (10.21), (10.30), and (10.31) show that if we adopt an eigenvector as one of basis vectors, the matrix representation of the linear transformation $A$ in reference to such basis vectors is obtained so that the leftmost column is zero except for the left top corner on which an eigenvalue is positioned. (Note that if the said eigenvector is zero, the leftmost column is a zero vector.) Meanwhile, neither Theorem 10.1 nor Theorem 10.2 tells about multiplicity of eigenvalues. If the eigenvalues have multiplicity, we have to think about different aspects. This is a major issue of this section.

Let us start with a discussion of invariant subspaces. Let $A$ be a $(n, n)$ square matrix. If a subspace $W$ in $V^n$ is characterized by

$$\boldsymbol{x} \in W \Rightarrow A\boldsymbol{x} \in W,$$

$W$ is said to be invariant with respect to $A$ (or simply $A$-invariant) or an invariant subspace in $V^n = \text{Span}\{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n\}$. Suppose that $\boldsymbol{x}$ is an eigenvector of $A$ and that its corresponding eigenvalue is $\alpha$. Then, Span $\{\boldsymbol{x}\}$ is an invariant subspace of $V^n$. It is because $A(c\boldsymbol{x}) = cA\boldsymbol{x} = c\alpha\boldsymbol{x} = \alpha(c\boldsymbol{x})$ and $c\boldsymbol{x}$ is again an eigenvector belonging to $\alpha$. Suppose that $\dim W = m (m \leq n)$ and that $W = \text{Span}\ \{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m\}$. If $W$ is $A$-invariant, $A$ causes a linear transformation within $W$. In that case, expressing $A$ in reference to $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m, \boldsymbol{a}_{m+1}, \ldots, \boldsymbol{a}_n$, we have

$$(\boldsymbol{a}_1\boldsymbol{a}_2 \ldots \boldsymbol{a}_m\boldsymbol{a}_{m+1} \ldots \boldsymbol{a}_n)A = (\boldsymbol{a}_1\boldsymbol{a}_2 \ldots \boldsymbol{a}_m\boldsymbol{a}_{m+1} \ldots \boldsymbol{a}_n)\begin{pmatrix} A_m & * \\ 0 & * \end{pmatrix}, \tag{10.45}$$

where $A_m$ is a $(m, m)$ square matrix and "zero" denotes a $(n - m, m)$ zero matrix. Notice that the transformation $A$ makes the remaining $(n - m)$ basis vectors $\boldsymbol{a}_{m+1}, \boldsymbol{a}_{m+2}, \ldots,$ and $\boldsymbol{a}_n$ in $V^n$ be converted to a linear combination of $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots,$ and $\boldsymbol{a}_n$. The triangle matrix $\Delta_n$ given in (10.30) and (10.31) is an example to which $A_m$ is a (1, 1) matrix (i.e., simply a complex number).

Let us examine properties of the $A$-invariant subspace still further. Let $\boldsymbol{a}$ be any vector in $V^n$ and think of following $(n + 1)$ vectors [2].

$$\boldsymbol{a}, A\boldsymbol{a}, A^2\boldsymbol{a}, \ldots, A^n\boldsymbol{a}.$$

These vectors are linearly dependent, since there are at most $n$ linearly independent vectors in $V^n$. These vectors span a subspace in $V^n$. Let us call this subspace $M$; i.e,.

$$M \equiv \text{Span}\{\boldsymbol{a}, A\boldsymbol{a}, A^2\boldsymbol{a}, \ldots, A^n\boldsymbol{a}\},$$

Consider the following equation:

$$c_0\boldsymbol{a} + c_1A\boldsymbol{a} + c_2A^2\boldsymbol{a} + \cdots + c_nA^n\boldsymbol{a} = \boldsymbol{0}. \tag{10.46}$$

Not all $c_i(0 \leq i \leq n)$ are zero, because the vectors are linearly dependent. Suppose that $c_n \neq 0$. Then, from (10.46) we have

$$A^n\boldsymbol{a} = -\frac{1}{c_n}\left(c_0\boldsymbol{a} + c_1A\boldsymbol{a} + c_2A^2\boldsymbol{a} + \cdots + c_{n-1}A^{n-1}\boldsymbol{a}\right).$$

Operating $A$ on the above equation from the left, we have

$$A^{n+1}\boldsymbol{a} = -\frac{1}{c_n}\left(c_0A\boldsymbol{a} + c_1A^2\boldsymbol{a} + c_2A^3\boldsymbol{a} + \cdots + c_{n-1}A^n\boldsymbol{a}\right).$$

Thus, $A^{n+1}\boldsymbol{a}$ is contained in $M$. That is, we have $A^{n+1}\boldsymbol{a} \in \mathrm{Span}\left\{\boldsymbol{a}, A\boldsymbol{a}, A^2\boldsymbol{a}, \ldots, A^n\boldsymbol{a}\right\}$. Next, suppose that $c_n = 0$. Then, at least one of $c_i(0 \leq i \leq n-1)$ is not zero. Suppose that $c_{n-1} \neq 0$. From (10.46), we have

$$A^{n-1}\boldsymbol{a} = -\frac{1}{c_{n-1}}\left(c_0\boldsymbol{a} + c_1A\boldsymbol{a} + c_2A^2\boldsymbol{a} + \cdots + c_{n-2}A^{n-2}\boldsymbol{a}\right).$$

Operating $A^2$ on the above equation from the left, we have

$$A^{n+1}\boldsymbol{a} = -\frac{1}{c_{n-1}}\left(c_0A^2\boldsymbol{a} + c_1A^3\boldsymbol{a} + c_2A^4\boldsymbol{a} + \cdots + c_{n-2}A^n\boldsymbol{a}\right).$$

Again, $A^{n+1}\boldsymbol{a}$ is contained in $M$. Repeating similar procedures, we reach

$$c_0\boldsymbol{a} + c_1A\boldsymbol{a} = \boldsymbol{0}.$$

If $c_1 = 0$, then we must have $c_0 \neq 0$. If so, $\boldsymbol{a} = \boldsymbol{0}$. This is impossible, however, because we should have $\boldsymbol{a} \neq \boldsymbol{0}$. Then, we have $c_1 \neq 0$ and, hence,

$$A\boldsymbol{a} = -\frac{c_0}{c_1}\boldsymbol{a}.$$

Operating once again $A^n$ on the above equation from the left, we have

$$A^{n+1}\boldsymbol{a} = -\frac{c_0}{c_1}A^n\boldsymbol{a}.$$

Thus, once again $A^{n+1}\boldsymbol{a}$ is contained in $M$.

In the above discussion, we get $AM \subset M$. Further operating $A$, $A^2M \subset AM \subset M$, $A^3M \subset A^2M \subset AM \subset M, \ldots$. Then we have

$$\boldsymbol{a}, A\boldsymbol{a}, A^2\boldsymbol{a}, \ldots, A^n\boldsymbol{a}, A^{n+1}\boldsymbol{a}, \ldots \in \text{Span}\left\{\boldsymbol{a}, A\boldsymbol{a}, A^2\boldsymbol{a}, \ldots, A^n\boldsymbol{a}\right\}.$$

That is, $M$ is an $A$-invariant subspace. We also have

$$m \equiv \dim M \leq \dim V^n = n.$$

There are $m$ basis vectors in $M$, and so representing $A$ in a matrix form in reference to the $n$ basis vectors of $V^n$ including these $m$ vectors, we have

$$A = \begin{pmatrix} A_m & * \\ 0 & * \end{pmatrix}. \tag{10.47}$$

Note again that in (10.45) $V^n$ is spanned by the $m$ basis vectors in $M$ together with other $(n - m)$ linearly independent vectors.

We happen to encounter a situation where two subspaces $W_1$ and $W_2$ are at once $A$-invariant. Here we can take basis vectors $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots,$ and $\boldsymbol{a}_m$ for $W_1$ and $\boldsymbol{a}_{m+1}, \boldsymbol{a}_{m+2}, \ldots,$ and $\boldsymbol{a}_n$ for $W_2$. In reference to such $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots,$ and $\boldsymbol{a}_n$ as basis vector of $V^n$, we have

$$A = \begin{pmatrix} A_m & 0 \\ 0 & A_{n-m} \end{pmatrix}, \tag{10.48}$$

where $A_{n-m}$ is a $(n - m, n - m)$ square matrix and "zero" denotes either a $(n - m, m)$ or $(m, n - m)$ zero matrix. Alternatively, we denote

$$A = A_m \oplus A_{n-m}.$$

In this situation, the matrix $A$ is said to be reducible.

As stated above, $A$ causes a linear transformation within both $W_1$ and $W_2$. In other words, $A_m$ and $A_{n-m}$ cause a linear transformation within $W_1$ and $W_2$ in reference to $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m$ and $\boldsymbol{a}_{m+1}, \boldsymbol{a}_{m+2}, \ldots, \boldsymbol{a}_n$, respectively. In this case, $V^n$ can be represented as a direct sum of $W_1$ and $W_2$ such that

$$\begin{aligned} V^n &= W_1 \oplus W_2 \\ &= \text{Span}\{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m\} \oplus \text{Span}\{\boldsymbol{a}_{m+1}, \boldsymbol{a}_{m+2}, \ldots, \boldsymbol{a}_n\}. \end{aligned} \tag{10.49}$$

This is because $\text{Span}\{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m\} \cap \text{Span}\{\boldsymbol{a}_{m+1}, \boldsymbol{a}_{m+2}, \ldots, \boldsymbol{a}_n\} = \{\boldsymbol{0}\}$. In fact, if the two subspaces possess $\boldsymbol{x}(\neq \boldsymbol{0})$ in common, $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$ become linearly dependent, in contradiction.

The vector space $V^n$ may well further be decomposed into subspaces with a lower dimension. For further discussion, we need a following theorem and a concept of a generalized eigenvector and generalized eigenspace.

**Theorem 10.3:** ***Hamilton–Cayley Theorem*** **[3]** *Let $f_A(x)$ be the characteristic polynomial pertinent to a linear transformation $A : V^n \to V^n$. Then $f_A(A)(\boldsymbol{x}) = 0$ for $^\forall\boldsymbol{x} \in V^n$. That is, Ker $f_A(A) = V^n$.*

*Proof* To prove the theorem, we use the following well-known property of determinants. Namely, let A be a $(n, n)$ square matrix expressed as

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}.$$

Let $\widetilde{A}$ be the cofactor matrix of $A$, namely

$$\widetilde{A} = \begin{pmatrix} \Delta_{11} & \cdots & \Delta_{1n} \\ \vdots & \ddots & \vdots \\ \Delta_{n1} & \cdots & \Delta_{nn} \end{pmatrix},$$

where $\Delta_{ij}$ is a cofactor of $a_{ij}$. Then

$$\widetilde{A}^T A = A\widetilde{A}^T = |A|E, \tag{10.50}$$

where $\widetilde{A}^T$ is a transposed matrix of $\widetilde{A}$. We now apply (10.50) to the characteristic polynomial.

$$(\widetilde{xE - A})^T(xE - A) = (xE - A)(\widetilde{xE - A})^T = |xE - A|E = f_A(A)E, \tag{10.51}$$

where $\widetilde{xE - A}$ is the cofactor matrix of $xE - A$. Let the cofactor of the $(i, j)$-element of $(xE - A)$ be $\Delta_{ij}$. Note in this case that $\Delta_{ij}$ is an at most $(n - 1)$th order polynomial of $x$. Let us put accordingly

$$\Delta_{ij} = b_{ij,0}x^{n-1} + b_{ij,1}x^{n-2} + \cdots + b_{ij,n-1}. \tag{10.52}$$

Also put $B_k = \left(b_{ij,k}\right)$.
Then we have

$$\begin{aligned} \widetilde{xE - A} &= \begin{pmatrix} \Delta_{11} & \cdots & \Delta_{1n} \\ \vdots & \ddots & \vdots \\ \Delta_{n1} & \cdots & \Delta_{nn} \end{pmatrix} \\ &= \begin{pmatrix} b_{11,0}x^{n-1} + \cdots + b_{11,n-1} & \cdots & b_{1n,0}x^{n-1} + \cdots + b_{1n,n-1} \\ \vdots & \ddots & \vdots \\ b_{n1,0}x^{n-1} + \cdots + b_{n1,n-1} & \cdots & b_{nn,0}x^{n-1} + \cdots + b_{nn,n-1} \end{pmatrix} \\ &= x^{n-1}\begin{pmatrix} b_{11,0} & \cdots & b_{1n,0} \\ \vdots & \ddots & \vdots \\ b_{n1,0} & \cdots & b_{nn,0} \end{pmatrix} + \cdots + \begin{pmatrix} b_{11,0} & \cdots & b_{1n,n-1} \\ \vdots & \ddots & \vdots \\ b_{n1,0} & \cdots & b_{nn,n-1} \end{pmatrix} \\ &= x^{n-1}B_0 + x^{n-2}B_1 + \cdots + B_{n-1}. \end{aligned} \tag{10.53}$$

Thus, we get

$$|xE - A|E = (xE - A)(x^{n-1}B_0^T + x^{n-2}B_1^T + \cdots + B_{n-1}^T). \tag{10.54}$$

Replacing $x$ with $A$, we have

$$f_A(A) = (A - A)\left(A^{n-1}B_0^T + A^{n-2}B_1^T + \cdots + B_{n-1}^T\right) = 0. \tag{10.55}$$

This means that $f_A(A)(\boldsymbol{x}) = 0$ for $^{\forall}\boldsymbol{x} \in V^n$.

In relation to Hamilton–Cayley theorem, we consider an important concept of a minimal polynomial.

**Definition 10.1** Let $f(x)$ be a polynomial of $x$ with scalar coefficients such that $f(A) = 0$, where $A$ is a $(n, n)$ matrix. Among $f(x)$, a lowest-order polynomial with the highest-order coefficient of one is said to be a minimal polynomial. We denote it by $\varphi_A(x)$; i.e., $\varphi_A(A) = 0$.

We have an important property for this. Namely, a minimal polynomial $\varphi_A(x)$ is a divisor of $f(A)$. In fact, suppose that we have

$$f(x) = g(x)\varphi_A(x) + h(x).$$

Inserting $A$ into $x$, we have

$$f(A) = g(A)\varphi_A(A) + h(A) = h(A) = 0.$$

From the above equation, $h(x)$ should be a polynomial whose order is lower than that of $\varphi_A(A)$. But the presence of such $h(x)$ is in contradiction to the definition of the minimal polynomial. This implies that $h(x) \equiv 0$. Thus, we get

$$f(x) = g(x)\varphi_A(x).$$

That is, $\varphi_A(x)$ must be a divisor of $f(A)$.

Suppose that $\varphi'_A(x)$ is another minimal polynomial. If the order of $\varphi'_A(x)$ is lower than that of $\varphi_A(x)$, we can choose $\varphi'_A(x)$ for a minimal polynomial from the beginning. Thus, we assume that $\varphi_A(x)$ and $\varphi'_A(x)$ are of the same order. We have

$$f(x) = g(x)\varphi_A(x) = g'(x)\varphi'_A(x).$$

Then, we have

$$\varphi_A(x) \,/\, \varphi'_A(x) = g'(x) \,/\, g(x) = c,$$

where $c$ is a constant, because $\varphi_A(x)$ and $\varphi'_A(x)$ are of the same order. But $c$ should be one, since the highest-order coefficient of the minimal polynomial is one. Thus, $\varphi_A(x)$ is uniquely defined.

## 10.3 Generalized Eigenvectors and Nilpotent Matrices

Equation (10.4) ensures that an eigenvalue is accompanied by an eigenvector. Therefore, if a matrix $A : V^n \to V^n$ has different $n$ eigenvalues without multiple roots, the vector space $V^n$ is decomposed to a direct sum of one-dimensional subspaces spanned by those individual linearly independent eigenvectors (see discussion of Sect. 10.1). Thus, we have

$$\begin{aligned} V^n &= W_1 \oplus W_2 \oplus \cdots \oplus W_n \\ &= \text{Span } \{\boldsymbol{a}_1\} \oplus \text{Span}\{\boldsymbol{a}_2\} \oplus \cdots \oplus \text{Span}\{\boldsymbol{a}_n\}, \end{aligned} \tag{10.56}$$

where $\boldsymbol{a}_i (1 \leq i \leq n)$ are eigenvectors corresponding to different $n$ eigenvalues. The situation, however, is not always simple. Even though a matrix has eigenvalues of multiple roots, we have yet a simple case as shown in a next example.

*Example 10.2* Let us think of a following matrix $A : V^3 \to V^3$.

$$A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}. \tag{10.57}$$

The matrix has a triple root 2. As can easily be seen below, individual eigenvectors $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, and $\boldsymbol{a}_3$ form basis vectors of each invariant subspace, indicating that $V^3$ can be decomposed to a direct sum of the three invariant subspaces as in (10.56).

$$(\boldsymbol{a}_1\boldsymbol{a}_2\boldsymbol{a}_3)\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} = (2\boldsymbol{a}_1 2\boldsymbol{a}_2 2\boldsymbol{a}_3). \tag{10.58}$$

Let us think of another simple example.

*Example 10.3* Let us think of a linear transformation $A : V^2 \to V^2$ expressed as

$$A(\boldsymbol{x}) = (\boldsymbol{a}_1\boldsymbol{a}_2)\begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \tag{10.59}$$

where $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ are basis vectors and for $^\forall \boldsymbol{x} \in V^2, \boldsymbol{x} = x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2$. This example has a double root 3. We have

$$(\boldsymbol{a}_1\boldsymbol{a}_2)\begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix} = (3\boldsymbol{a}_1\boldsymbol{a}_1 + 3\boldsymbol{a}_2). \tag{10.60}$$

Thus, $\text{Span}\{\boldsymbol{a}_1\}$ is $A$-invariant, but this is not the case with $\text{Span}\{\boldsymbol{a}_2\}$. This implies that $\boldsymbol{a}_1$ is an eigenvector corresponding to an eigenvalue 3 but $\boldsymbol{a}_2$ is not. Detailed discussion about matrices of this kind follows below.

Nilpotent matrices play an important role in matrix algebra. These matrices are defined as follows.

**Definition 10.2** Let $N$ be a linear transformation in a vector space $V^n$. Suppose that we have

$$N^k = 0 \text{ and } N^{k-1} \neq 0, \tag{10.61}$$

where $N$ is a $(n, n)$ square matrix and $k(\leq 2)$ is a certain natural number. Then, $N$ is said to be a nilpotent matrix of an order $k$ or a $k$th order nilpotent matrix. If (10.61) holds with $k = 1$, $N$ is a zero matrix.

Nilpotent matrices have following properties:

(i) Eigenvalues of a nilpotent matrix are zero. Let $N$ be a $k$th order nilpotent matrix. Suppose that

$$N\boldsymbol{x} = \alpha\boldsymbol{x}, \tag{10.62}$$

where $\alpha$ is an eigenvalue and $\boldsymbol{x}(\neq \boldsymbol{0})$ is its corresponding eigenvector. Operating $N$ $(k-1)$ more times from the left of both sides of (10.62), we have

$$N^k\boldsymbol{x} = \alpha^k\boldsymbol{x}. \tag{10.63}$$

Meanwhile, $N^k = 0$ by definition, and so $\alpha^k = 0$, namely $\alpha = 0$.
Conversely, suppose that eigenvalues of a $(n, n)$ matrix $N$ are zero. From Theorem 10.1, via a suitable similarity transformation with $P$ we have

$$P^{-1}NP = \widetilde{N},$$

where $\widetilde{N}$ is a triangle matrix. Then, using (10.12) and (10.16) we have

$$f_{P^{-1}NP}(x) = f_{\widetilde{N}}(x) = f_N(x) = x^n.$$

From Theorem 10.3, we have

$$f_N(N) = N^n = 0. \tag{10.64}$$

Namely, $N$ is a nilpotent matrix. In a trivial case, we have $N = 0$ (zero matrix). By Definition 10.2, we have $N^k = 0$ with a $k$th nilpotent $(n, n)$ matrix $N$. Then, its minimal polynomial is $\varphi_N(x) = x^k(k \leq n)$.

(ii) A nilpotent matrix $N$ is not diagonalizable (except for a zero matrix). Suppose that $N$ is diagonalizable. Then $N$ can be diagonalized by a non-singular matrix $P$ such that

$$P^{-1}NP = 0.$$

The above equation holds, because $N$ only takes eigenvalues of zero. Operating $P$ from the left of the above equation and $P^{-1}$ from the right, we have

$$N = 0.$$

This means that $N$ would be a zero matrix, in contradiction. Thus, a nilpotent matrix $N$ is not diagonalizable.

*Example 10.4* Let $N$ be a matrix of a following form:

$$N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Then, we can easily check that $N^2 = 0$. Therefore, $N$ is a nilpotent matrix of a second order. Note that $N$ is an upper triangle matrix, and so eigenvalues are given by diagonal elements. In the present case, the eigenvalue is zero (as a double root), consistent with the aforementioned property.

With a nilpotent matrix of an order $k$, we have at least one vector $\boldsymbol{x}$ such that $N^{k-1}\boldsymbol{x} \neq 0$. Here we add that a zero transformation $A$ (or matrix) can be defined as

$$A = 0 \Leftrightarrow A\boldsymbol{x} = \boldsymbol{0} \text{ for } {}^{\forall}\boldsymbol{x} \in V^n.$$

Taking contraposition of this, we have

$$A \neq 0 \Leftrightarrow A\boldsymbol{x} \neq \boldsymbol{0} \text{ for } {}^{\exists}\boldsymbol{x} \in V^n.$$

In relation to a nilpotent matrix, we have a following important theorem.

**Theorem 10.4** *If $N$ is a kth order nilpotent matrix, then for ${}^{\exists}\boldsymbol{x} \in V$ we have following linearly independent k vectors*:

$$\boldsymbol{x}, N\boldsymbol{x}\ N^2\boldsymbol{x}, \ldots, N^{k-1}\boldsymbol{x}.$$

*Proof* Let us think of a following equation:

$$\sum_{i=0}^{k-1} c_i N^i \boldsymbol{x} = \boldsymbol{0}. \tag{10.65}$$

Multiplying (8.64) by $N^{k-1}$ and using $N^k = 0$, we get

$$c_0 N^{k-1}\boldsymbol{x} = \boldsymbol{0}. \tag{10.66}$$

This implies that $c_0 = 0$. Thus, we are left with

$$\sum_{i=1}^{k-1} c_i N^i \boldsymbol{x} = \boldsymbol{0}. \tag{10.67}$$

Multiplying (8.66) next by $N^{k-2}$, we get similarly

$$c_1 N^{k-1}\boldsymbol{x} = \mathbf{0}, \tag{10.68}$$

implying that $c_1 = 0$. Continuing this procedure, we finally get $c_i = 0 (0 \leq i \leq k-1)$. This completes the proof.

This immediately tells us that for a $k$th order nilpotent matrix $N : V^n \to V^n$, we should have $k \leq n$. This is because the number of linearly independent vectors is at most $n$.

In Example 10.4, $N$ causes a transformation of basis vectors in $V^2$ such that

$$(\boldsymbol{e_1}\ \boldsymbol{e_2})N = (\boldsymbol{e_1}\ \boldsymbol{e_2})\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = (\mathbf{0}\ \boldsymbol{e_1}).$$

That is, $N\boldsymbol{e_2} = \boldsymbol{e_1}$. Then, linearly independent two vectors $\boldsymbol{e_2}$ and $N\boldsymbol{e_2}$ correspond to the case of Theorem 10.4.

So far we have shown simple cases where matrices can be diagonalized via similarity transformation. This is equivalent to that the relevant vector space is decomposed to a direct sum of (invariant) subspaces. Nonetheless, if a characteristic polynomial of the matrix has multiple root(s), it remains uncertain whether the vector space is decomposed to such a direct sum. To answer this question, we need a following lemma.

**Lemma 10.1** Let $f_1(x), f_2(x), \ldots,$ and $f_s(x)$ be polynomials without a common factor. Then we have $s$ polynomials $M_1(x), M_2(x), \ldots, M_s(x)$ that satisfy the following relation:

$$M_1(x)f_1(x) + M_2(x)f_2(x) + \cdots + M_s(x)f_s(x) = 1. \tag{10.69}$$

*Proof* Let $M_i(x)$ $(1 \leq i \leq s)$ be arbitrarily chosen polynomials and deal with a set of $g(x)$ that can be expressed as

$$g(x) = M_1(x)f_1(x) + M_2(x)f_2(x) + \cdots + M_s(x)f_s(x). \tag{10.70}$$

Let a whole set of $g(x)$ be $\mathfrak{H}$. Then $\mathfrak{H}$ has following two properties:

(i) $g_1(x),\ g_2(x) \in \mathfrak{H} \Rightarrow g_1(x) + g_2(x) \in \mathfrak{H}$,
(ii) $g(x) \in \mathfrak{H}, M(x)$: an arbitrarily chosen polynomial $\Rightarrow M(x)g(x) \in \mathfrak{H}$.

Now let the lowest-order polynomial out of the set $\mathfrak{H}$ be $g_0(x)$. Then $^{\forall}g(x)(\in \mathfrak{H})$ are a multiple of $g_0(x)$. Suppose that dividing $g(x)$ by $g_0(x)$, we have

$$g(x) = M(x)g_0(x) + h(x), \tag{10.71}$$

where $h(x)$ is a certain polynomial. Since $g(x),\ g_0(x) \in \mathfrak{H}$, we have $h(x) \in \mathfrak{H}$ from the above properties (i) and (ii). If $h(x) \neq 0$, the order of $h(x)$ is lower than that of $g_0(x)$ from (10.71). This is, however, in contradiction to the definition of $g_0(x)$. Therefore, $h(x) = 0$. Thus,

$$g(x) = M(x)g_0(x). \tag{10.72}$$

This implies that $\mathfrak{H}$ is identical to a whole set of polynomials comprising multiples of $g_0(x)$. In particular, polynomials $f_i(x)$ $(1 \leq i \leq s) \in \mathfrak{H}$. To show this, put $M_i(x) = 1$ with other $M_j(x) = 0$ $(j \neq i)$. Hence, the polynomial $g_0(x)$ should be a common factor of $f_i(x)$. Meanwhile, $g_0(x) \in \mathfrak{H}$, and so by virtue of (10.70) we have

$$g_0(x) = M_1(x)f_1(x) + M_2(x)f_2(x) + \cdots + M_s(x)f_s(x). \tag{10.73}$$

On assumption, the greatest common factor of $f_1(x), f_2(x), \ldots,$ and $f_s(x)$ is 1. This implies that $g_0(x) = 1$. Thus, we finally get (10.69) and complete the proof.

## 10.4 Idempotent Matrices and Generalized Eigenspaces

In Sect. 10.1, we have shown that eigenvectors corresponding to different eigenvalues of $A$ are linearly independent. Also if those eigenvalues do not possess multiple roots, the vector space comprises a direct sum of the subspaces of corresponding eigenvectors. However, how do we have to treat a situation where eigenvalues possess multiple roots? Even in this case there is at least one eigenvector that corresponds to the eigenvalue. To adequately address the question, we need a concept of generalized eigenvectors and generalized eigenspaces.

For this purpose, we extend and generalize the concept of eigenvectors. For a certain natural number $l$, if a vector $\boldsymbol{x}$ $(\in V^n)$ satisfies a following relation, $\boldsymbol{x}$ is said to be a generalized eigenvector of rank $l$ that corresponds to an eigenvalue.

$$(A - \alpha E)^l \boldsymbol{x} = \boldsymbol{0},$$
$$(A - \alpha E)^{l-1} \boldsymbol{x} \neq \boldsymbol{0}.$$

Thus (10.3) implies that an eigenvector of (10.3) may be said to be a generalized eigenvector of rank 1. When we need to distinguish it from generalized eigenvectors of rank $l(l \geq 2)$, we call it a "proper" eigenvector. Thus far we have only been concerned with the proper eigenvectors. We have a following theorem related to the generalized eigenvectors.

In this section, idempotent operators play a role. The definition is simple.

**Definition 10.3** An operator $A$ is said to be idempotent if $A^2 = A$.

From this simple definition, we can draw several important pieces of information. Let $A$ be an idempotent operator that operates on $V^n$. Let $\boldsymbol{x}$ be an arbitrary vector in $V^n$. Then, $A^2\boldsymbol{x} = A(A\boldsymbol{x}) = A\boldsymbol{x}$. That is,

$$A(A\boldsymbol{x} - \boldsymbol{x}) = \boldsymbol{0}.$$

Then, we have

$$A\boldsymbol{x} = \boldsymbol{x} \quad \text{or} \quad A\boldsymbol{x} = \boldsymbol{0}. \tag{10.74}$$

From Theorem 9.3 (dimension theorem), we have

$$V^n = A(V^n) \oplus \operatorname{Ker} A,$$

where

$$\begin{aligned} A(V^n) &= \{\boldsymbol{x};\ \boldsymbol{x} \in V^n,\ A\boldsymbol{x} = \boldsymbol{x}\}, \\ \operatorname{Ker} A &= \{\boldsymbol{x};\ \boldsymbol{x} \in V^n,\ A\boldsymbol{x} = \boldsymbol{0}\}. \end{aligned} \tag{10.75}$$

Thus, we find that $A$ decomposes $V^n$ into a direct sum of $A(V^n)$ and $\operatorname{Ker} A$. Conversely, we can readily verify that if there exists an operator $A$ that satisfies (10.75), such $A$ must be an idempotent operator. The verification is left for readers.

Meanwhile, we have

$$(E - A)^2 = E - 2A + A^2 = E - 2A + A = E - A.$$

Hence, $E - A$ is an idempotent matrix as well. Moreover, we have

$$A(E - A) = (E - A)A = 0.$$

Putting $E - A \equiv B$ and following a procedure similar to the above

$$B\boldsymbol{x} - \boldsymbol{x} = (E - A)\boldsymbol{x} - \boldsymbol{x} = \boldsymbol{x} - A\boldsymbol{x} - \boldsymbol{x} = -A\boldsymbol{x}.$$

Therefore, $B(V^n) = \{\boldsymbol{x};\ \boldsymbol{x} \in V^n,\ B\boldsymbol{x} = \boldsymbol{x}\}$ is identical to $\operatorname{Ker} A$. Writing

$$W_A = \{\boldsymbol{x};\ \boldsymbol{x} \in V^n,\ A\boldsymbol{x} = \boldsymbol{x}\},\ W_{\overline{A}} = \{\boldsymbol{x};\ \boldsymbol{x} \in V^n,\ A\boldsymbol{x} = \boldsymbol{0}\}, \tag{10.76}$$

we get

$$W_A = AV^n,\ W_{\overline{A}} = (E - A)V^n = BV^n = V^n - AV^n = \operatorname{Ker} A.$$

That is,

$$V^n = W_A + W_{\overline{A}}.$$

Suppose that $^{\exists}\boldsymbol{u} \in W_A \cap W_{\overline{A}}$. Then, from (10.76) $A\boldsymbol{u} = \boldsymbol{u}$ and $A\boldsymbol{u} = \boldsymbol{0}$, namely $\boldsymbol{u} = \boldsymbol{0}$, and so $V$ is a direct sum of $W_A$ and $W_{\overline{A}}$. That is,

$$V^n = W_A \oplus W_{\overline{A}}.$$

Notice that if we consider an identity operator $E$ as an idempotent operator, we are thinking of a trivial case. That is, $V^n = V^n \oplus \{\mathbf{0}\}$.

The result can immediately be extended to the case where more idempotent operators take part in the vector space. Let us define operators such that

$$A_1 + A_2 + \cdots + A_s = E,$$

where $E$ is a $(n, n)$ identity matrix. Also

$$A_i A_j = A_i \delta_{ij}.$$

Moreover, we define $W_i$ such that

$$W_i = A_i V^n = \{\boldsymbol{x};\ \boldsymbol{x} \in V^n,\ A_i \boldsymbol{x} = \boldsymbol{x}\}. \tag{10.77}$$

Then

$$V^n = W_1 \oplus W_2 \oplus \cdots \oplus W_s. \tag{10.78}$$

In fact, suppose that $^{\exists}\boldsymbol{x}(\in V^n) \in W_i, W_j\ (i \neq j)$. Then $A_i \boldsymbol{x} = \boldsymbol{x} = A_j \boldsymbol{x}$. Operating $A_j\ (j \neq i)$ from the left, $A_j A_i \boldsymbol{x} = A_j \boldsymbol{x} = A_j A_j \boldsymbol{x}$. That is, $\mathbf{0} = \boldsymbol{x} = A_j \boldsymbol{x}$, implying that $W_i \cap W_j = \{\mathbf{0}\}$. Meanwhile,

$$\begin{aligned} V^n &= (A_1 + A_2 + \cdots + A_s) V^n = A_1 V^n + A_2 V^n + \cdots + A_s V^n \\ &= W_1 + W_2 + \cdots + W_s. \end{aligned} \tag{10.79}$$

As $W_i \cap W_j = \{\mathbf{0}\}\ (i \neq j)$, (10.79) is a direct sum. Thus, (10.78) will follow.

*Example 10.5* Think of the following transformation $A$:

$$(\boldsymbol{e_1}\ \boldsymbol{e_2}\ \boldsymbol{e_3}\ \boldsymbol{e_4})A = (\boldsymbol{e_1}\ \boldsymbol{e_2}\ \boldsymbol{e_3}\ \boldsymbol{e_4}) \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = (\boldsymbol{e_1}\ \boldsymbol{e_2}\ \mathbf{0}\ \mathbf{0}).$$

Put $\boldsymbol{x} = \sum_{i=1}^{4} x_i \boldsymbol{e_i}$. Then, we have

$$A(\boldsymbol{x}) = \sum_{i=1}^{4} x_i A(\boldsymbol{e_i}) = \sum_{i=1}^{2} x_i \boldsymbol{e_i} = W_A,$$

$$(E - A)(\boldsymbol{x}) = \boldsymbol{x} - A(\boldsymbol{x}) = \sum_{i=3}^{4} x_i \boldsymbol{e_i} = W_{\overline{A}}.$$

In the above,

$$(\boldsymbol{e_1}\ \boldsymbol{e_2}\ \boldsymbol{e_3}\ \boldsymbol{e_4})(E-A) = (\boldsymbol{e_1}\ \boldsymbol{e_2}\ \boldsymbol{e_3}\ \boldsymbol{e_4})\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = (\boldsymbol{0}\ \boldsymbol{0}\ \boldsymbol{e_3}\ \boldsymbol{e_4}).$$

Thus, we have

$$W_A = \text{Span}\ \{\boldsymbol{e_1},\ \boldsymbol{e_2}\},\ W_{\overline{A}} = \text{Span}\ \{\boldsymbol{e_3},\ \boldsymbol{e_4}\},\ V^4 = W_A \oplus W_{\overline{A}}.$$

The properties of idempotent matrices can easily be checked. It is left for readers.

Using the aforementioned idempotent operators, let us introduce the following theorem.

**Theorem 10.5 [3]** *Let $A$ be a linear transformation $V^n \to V^n$. Suppose that a vector $\boldsymbol{x}$ $(\in V^n)$ satisfies the following relation*:

$$(A - \alpha E)^l \boldsymbol{x} = \boldsymbol{0}, \tag{10.80}$$

*where $l$ is an enough large natural number. Then a set comprising $\boldsymbol{x}$ forms an $A$-invariant subspace that corresponds to an eigenvalue $\alpha$. Let $\alpha_1, \alpha_2, \ldots, \alpha_s$ be eigenvalues of $A$ different from one another. Then $V^n$ is decomposed to a direct sum of the $A$-invariant subspaces that correspond individually to $\alpha_1, \alpha_2, \ldots, \alpha_s$. This is succinctly expressed as follows*:

$$V^n = \widetilde{W_{\alpha_1}} \oplus \widetilde{W_{\alpha_2}} \oplus \cdots \oplus \widetilde{W_{\alpha_s}}. \tag{10.81}$$

*Here $\widetilde{W}_{\alpha_i}$ $(1 \le i \le s)$ is given by*

$$W_{\alpha_i} = \left\{\boldsymbol{x};\ \boldsymbol{x} \in V^n,\ (A - \alpha_i E)^{l_i}\boldsymbol{x} = \boldsymbol{0}\right\}, \tag{10.82}$$

*where $l_i$ is an enough large natural number. If multiplicity of $\alpha_i$ is $n_i$, dim $\widetilde{W}_{\alpha_i} = n_i$.*

*Proof* Let us define the aforementioned $A$-invariant subspaces as $\widetilde{W}_{\alpha_k}$ $(1 \le k \le s)$. Let $f_A(x)$ be a characteristic polynomial of $A$. Factorizing $f_A(x)$ into a product of powers of first-degree polynomials, we have

$$f_A(x) = \prod_{i=1}^{s} (x - \alpha_i)^{n_i}, \tag{10.83}$$

where $n_i$ is a multiplicity of $\alpha_i$. Let us put

$$f_i(x) = f_A(x)/(x - \alpha_i)^{n_i} = \prod_{j \ne i}^{s} (x - \alpha_j)^{n_j}. \tag{10.84}$$

Then $f_1(x), f_2(x), \ldots,$ and $f_s(x)$ do not have a common factor. Consequently, Lemma 10.1 tells us that there are polynomials $M_1(x), M_2(x), \ldots,$ and $M_s(x)$ such that

$$M_1(x)f_1(x) + \cdots + M_s(x)f_s(x) = 1. \tag{10.85}$$

Replacing $x$ with a matrix $A$, we get

$$M_1(A)f_1(A) + \cdots + M_s(A)f_s(A) = E. \tag{10.86}$$

Or defining $M_i(A)f_i(A) \equiv A_i$

$$A_1 + A_2 + \cdots + A_s = E, \tag{10.87}$$

where $E$ is a $(n, n)$ identity matrix. Moreover, we have

$$A_iA_j = A_i\delta_{ij}. \tag{10.88}$$

In fact, if $i \neq j$,

$$\begin{aligned} A_iA_j &= M_i(A)f_i(A)M_j(A)f_j(A) = M_i(A)M_j(A)f_i(A)f_j(A) \\ &= M_i(A)M_j(A)\prod_{k\neq i}^{s}(A-\alpha_k)^{n_k}\prod_{l\neq j}^{s}(A-\alpha_l)^{n_l} \\ &= M_i(A)M_j(A)f_A(A)\prod_{k\neq i,j}^{s}(A-\alpha_k)^{n_k} \\ &= 0. \end{aligned} \tag{10.89}$$

The second equality results from the fact that $M_j(A)$ and $f_i(A)$ are commutable since both are polynomials of $A$. The last equality follows from Hamilton–Cayley Theorem. On the basis of (10.87) and (10.89),

$$A_i = A_iE = A_i(A_1 + A_2 + \cdots + A_s) = A_i^2. \tag{10.90}$$

Thus, we find that $A_i$ is an idempotent matrix.

Next, let us show that using $A_i$ determined above, $A_iV^n$ is identical to $\widetilde{W}_{\alpha_i}$ $(1 \leq i \leq s)$. To this end, we define $W_i$ such that

$$W_i = A_iV^n = \{\boldsymbol{x};\ \boldsymbol{x} \in V^n,\ A_i\boldsymbol{x} = \boldsymbol{x}\}. \tag{10.91}$$

We have

$$(x-\alpha_i)^{n_i}f_i(x) = f_A(x). \tag{10.92}$$

Therefore, from Hamilton–Cayley Theorem we have

$$(A - \alpha_i E)^{n_i} f_i(A) = 0. \tag{10.93}$$

Operating $M_i(A)$ from the left and from the fact that $M_i(A)$ commutes with $(A - \alpha_i E)^{n_i}$, we get

$$(A - \alpha_i E)^{n_i} A_i = 0, \tag{10.94}$$

where we used $M_i(A) f_i(A) = A_i$. Operating both sides of this equation on $V^n$, furthermore, we have

$$(A - \alpha_i E)^{n_i} A_i V^n = 0.$$

This means that

$$A_i V^n \subset \widetilde{W}_{\alpha_i} \ (1 \leq i \leq s). \tag{10.95}$$

Conversely, suppose that $\boldsymbol{x} \in \widetilde{W}_{\alpha_i}$. Then $(A - \alpha E)^l \boldsymbol{x} = \boldsymbol{0}$ holds for a certain natural number $l$. If $M_i(x) f_i(x)$ were divided out by $x - \alpha_i$, LHS of (10.85) would be divided out by $x - \alpha_i$ as well, leading to the contradiction. Thus, it follows that $(x - \alpha_i)^l$ and $M_i(x) f_i(x)$ do not have a common factor. Consequently, Lemma 10.1 ensures that we have polynomials $M(x)$ and $N(x)$ such that

$$M(x)(x - \alpha_i)^l + N(x) M_i(x) f_i(x) = 1.$$

Hence,

$$M(A)(A - \alpha_i E)^l + N(A) M_i(A) f_i(A) = E. \tag{10.96}$$

Operating both sides of (10.96) on $\boldsymbol{x}$, we get

$$M(A)(A - \alpha_i E)^l \boldsymbol{x} + N(A) M_i(A) f_i(A) \boldsymbol{x} = N(A) A_i \boldsymbol{x} = \boldsymbol{x}. \tag{10.97}$$

Notice that the first term of (10.97) vanishes from (10.82).
As $A_i$ is a polynomial of $A$, it commutes with $N(A)$. Hence, we have

$$\boldsymbol{x} = A_i [N(A) \boldsymbol{x}] \in A_i V^n. \tag{10.98}$$

Thus, we get

$$\widetilde{W}_{\alpha_i} \subset A_i V^n \ (1 \leq i \leq s). \tag{10.99}$$

From (10.95) and (10.99), we conclude that

$$\widetilde{W}_{\alpha_i} = A_i V^n \ (1 \leq i \leq s). \tag{10.100}$$

In other words, $W_i$ defined as (10.91) is identical to $\widetilde{W}_{\alpha_i}$ that is defined as (10.82). Thus, we have

$$V^n = W_1 \oplus W_2 \oplus \cdots \oplus W_s.$$

or

$$V^n = \widetilde{W_{\alpha_1}} \oplus \widetilde{W_{\alpha_2}} \oplus \cdots \oplus \widetilde{W_{\alpha_s}}. \tag{10.101}$$

This completes the former half of the proof. With the latter half, the proof is as follows.

Suppose that $\dim \widetilde{W}_{\alpha_i} = n_i'$. In parallel to the decomposition of $V^n$ to the direct sum of (10.81), $A$ can be reduced as

$$A \sim \begin{pmatrix} A^{(1)} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & A^{(s)} \end{pmatrix}, \tag{10.102}$$

where $A^{(i)}$ $(1 \leq i \leq s)$ is a $\left(n_i', n_i'\right)$ matrix and a symbol $\sim$ indicates that $A$ has been transformed by suitable similarity transformation. The matrix $A^{(i)}$ represents a linear transformation that $A$ causes to $\widetilde{W}_{\alpha_i}$. We denote a $n_i'$ order identity matrix by $E_{n_i'}$. Equation (10.82) implies that the matrix represented by

$$N_i = A^{(i)} - \alpha_i E_{n_i'}. \tag{10.103}$$

is a nilpotent matrix. The order of an nilpotent matrix is at most $n_i'$ (vide supra) and, hence, $l_i$ can be $n_i'$. With $N_i$, we have

$$\begin{aligned} f_{N_i}(x) &= \left|xE_{n_i'} - N_i\right| = \left|xE_{n_i'} - \left[A^{(i)} - \alpha_i E_{n_i'}\right]\right| = \left|xE_{n_i'} - A^{(i)} + \alpha_i E_{n_i'}\right| \\ &= \left|(x+\alpha_i)E_{n_i'} - A^{(i)}\right| = x^{n_i'}. \end{aligned} \tag{10.104}$$

The last equality is because eigenvalues of a nilpotent matrix are all zero. Meanwhile,

$$f_{A^{(i)}}(x) = \left|xE_{n_i'} - A^{(i)}\right| = f_{N_i}(x - \alpha_i) = (x - \alpha_i)^{n_i'}. \tag{10.105}$$

Equation (10.105) implies that

$$f_A(x) = \prod_{i=1}^{s} f_{A^{(i)}}(x) = \prod_{i=1}^{s} (x - \alpha_i)^{n_i'} = \prod_{i=1}^{s} (x - \alpha_i)^{n_i}. \tag{10.106}$$

The last equality comes from (10.83). Thus, $n_i' = n_i$. These procedures complete the proof. At the same time, we may equate $l_i$ in (10.82) to $n_i$.

Theorem 10.1 shows that any square matrix can be converted to a (upper) triangle matrix by a similarity transformation. Theorem 10.5 demonstrates that the matrix can further be segmented according to individual eigenvalues. Considering Theorem 10.1 again, $A^{(i)}$ $(1 \leq i \leq s)$ can be described as an upper triangle matrix by

$$A^{(i)} \sim \begin{pmatrix} \alpha_i & \dots & * \\ \vdots & \ddots & \vdots \\ 0 & \dots & \alpha_i \end{pmatrix}. \tag{10.107}$$

Therefore, denoting $N^{(i)}$ such that

$$N^{(i)} = A^{(i)} - \alpha_i E_{n_i} = \begin{pmatrix} 0 & \dots & * \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{pmatrix}, \tag{10.108}$$

we find that $N^{(i)}$ is nilpotent. This is because all the eigenvalues of $N^{(i)}$ are zero. From (10.108), we have

$$[N^{(i)}]^{\mu_i} = 0 \ (\mu_i \leq n_i).$$

## 10.5 Decomposition of Matrix

To investigate canonical forms of matrices, it would be convenient if a matrix can be decomposed into appropriate forms. To this end, the following definition is important.

**Definition 10.4** A matrix similar to a diagonal matrix is said to be semi-simple.

In the above definition, if a matrix is related to another matrix by similarity transformation, those matrices are said to be similar to each other. When we have two matrices $A$ and $A'$, we express it by $A \sim A'$ as stated above. This relation satisfies the equivalence law. That is,

(i) $A \sim A$, (ii) $A \sim A' \Rightarrow A' \sim A$, (iii) $A \sim A', A' \sim A'' \Rightarrow A \sim A''$.

Readers, check this. We have a following important theorem with the matrix decomposition.

**Theorem 10.6 [3]** *Any* $(n, n)$ *square matrix* $A$ *is expressed uniquely as*

$$A = S + N, \tag{10.109}$$

*where* $S$ *is semi-simple* and $N$ *is nilpotent*; $S$ *and* $N$ *are commutable, i.e.,* $SN = NS$. *Furthermore,* $S$ *and* $N$ *are polynomials of* $A$ *with scalar coefficients*.

*Proof* Using (10.86) and (10.87), we write

$$S = \alpha_1 A_1 + \cdots + \alpha_s A_s = \sum_{i=1}^{s} \alpha_i M_i(A) f_i(A). \tag{10.110}$$

Then, Eq. (10.110) is a polynomial of $A$. From Theorems 10.1 and 10.5, $A^{(i)}$ $(1 \leq i \leq s)$ in (10.102) is characterized by that $A^{(i)}$ is a triangle matrix whose eigenvalues $\alpha_i$ $(1 \leq i \leq s)$ are positioned on diagonal positions and that the order of $A^{(i)}$ is identical to the multiplicity of $\alpha_i$. Since $A_i$ $(1 \leq i \leq s)$ is an idempotent matrix, it should be diagonalized through similarity transformation (see Sect. 10.7). In fact, $S$ is transformed via similarity transformation the same as (10.102) into

$$S \sim \begin{pmatrix} \alpha_1 E_{n_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \alpha_s E_{n_s} \end{pmatrix}, \tag{10.111}$$

where $E_{n_i}$ $(1 \leq i \leq s)$ is an identity matrix of an order $n_i$ that is identical to the multiplicity of $\alpha_i$. This expression is equivalent to e.g.,

$$A_1 \sim \begin{pmatrix} E_{n_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}$$

in (10.110). Thus, $S$ is obviously semi-simple. Putting $N = A - S$, $N$ is described after the above transformation as

$$N \sim \begin{pmatrix} N^{(1)} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & N^{(s)} \end{pmatrix}, \quad N^{(i)} = A^{(i)} - \alpha_i E_{n_i}. \tag{10.112}$$

Since each $N^{(i)}$ is nilpotent as stated in Sect. 10.4, $N$ is nilpotent as well. Also (10.112) is a polynomial of $A$ as in the case of $S$. Therefore, $S$ and $N$ are commutable.

To prove the uniqueness of the decomposition, we show the following:

(i) Let $S$ and $S'$ be commutable semi-simple matrices. Then, those matrices are simultaneously diagonalized. That is, with a certain non-singular matrix $P$, $P^{-1}SP$ and $P^{-1}S'P$ are diagonalized at once. Hence, $S \pm S'$ is semi-simple as well.

(ii) Let $N$ and $N'$ be commutable nilpotent matrices. Then, $N \pm N'$ is nilpotent as well.

(iii) A matrix both semi-simple and nilpotent is zero matrix.

(i) Let different eigenvalues of $S$ be $\alpha_1, \ldots, \alpha_s$. Then, since $S$ is semi-simple, a vector space $V^n$ is decomposed into a direct sum of eigenspaces $W_{\alpha_i}$ $(1 \leq i \leq s)$. That is, we have

$$V^n = W_{\alpha_1} \oplus \cdots \oplus W_{\alpha_s}.$$

Since $S$ and $S'$ are commutable, with $^{\exists}\boldsymbol{x} \in W_{\alpha_i}$ we have $SS'\boldsymbol{x} = S'S\boldsymbol{x} = S'(\alpha_i\boldsymbol{x}) = \alpha_i S'\boldsymbol{x}$. Hence, we have $S'\boldsymbol{x} \in W_{\alpha_i}$. Namely, $W_{\alpha_i}$ is $S'$-invariant. Therefore, if we adopt the basis vectors $\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n\}$ with respect to the direct sum decomposition, we get

$$S \sim \begin{pmatrix} \alpha_1 E_{n_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \alpha_s E_{n_s} \end{pmatrix}, \quad S' \sim \begin{pmatrix} S'_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & S'_s \end{pmatrix}.$$

Since $S'$ is semi-simple, $S'_i$ $(1 \leq i \leq s)$ must be semi-simple as well. Here, let $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ be original basis vectors before the basis vector transformation and let $P$ be a representation matrix of the said transformation. Then, we have

$$(\boldsymbol{e}_1 \ldots \boldsymbol{e}_n)P = (\boldsymbol{a}_1 \ldots \boldsymbol{a}_n).$$

Thus, we get

$$P^{-1}SP = \begin{pmatrix} \alpha_1 E_{n_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \alpha_s E_{n_s} \end{pmatrix}, \quad P^{-1}S'P = \begin{pmatrix} S'_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & S'_s \end{pmatrix}.$$

This means that both $P^{-1}SP$ and $P^{-1}S'P$ are diagonal. That is, $P^{-1}SP \pm P^{-1}S'P = P^{-1}(S \pm S')P$ is diagonal, indicating that $S \pm S'$ is semi-simple as well.

(ii) Suppose that $N^{\nu} = 0$ and $N'^{\nu'} = 0$. From the assumption, $N$ and $N'$ are commutable. Consequently, using binomial theorem we have

$$(N \pm N')^m = N^m \pm mN^{m-1}N' + \cdots + (\pm 1)^i \frac{m!}{i!(m-i)!} N^{m-i}N'^i + \cdots + (\pm 1)^m N'^m. \tag{10.113}$$

Putting $m = v + v' - 1$, if $i \geq v'$, $N'^i = 0$ from the supposition. If $i < v'$, we have $m - i > m - v' = v + v' - 1 - v' = v - 1$; i.e., $m - i \geq v$. Therefore, $N^{m-i} = 0$. Consequently, we have $N^{m-i}N'^i = 0$ with any $i$ in (10.113). Thus, we get $(N \pm N')^m = 0$, indicating that $N \pm N'$ is nilpotent.

(iii) Let $S$ be a semi-simple and nilpotent matrix. We describe $S$ as

$$S \sim \begin{pmatrix} \alpha_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \alpha_n \end{pmatrix}, \tag{10.114}$$

where some of $\alpha_i \, (1 \leq i \leq n)$ may be identical. Since $S$ is nilpotent, all $\alpha_i \, (1 \leq i \leq n)$ is zero. We have then $S \sim 0$; i.e., $S = 0$ accordingly.
Now, suppose that a matrix $A$ is decomposed differently from (10.109). That is, we have

$$A = S + N = S' + N' \quad \text{or} \quad S - S' = N' - N. \tag{10.115}$$

From the assumption, $S'$ and $N'$ are commutable. Moreover, since $S, S', N$, and $N'$ are described by a polynomial of $A$, they are commutable with one another. Hence, from (i) and (ii) along with the second equation of (10.115), $S - S'$ and $N' - N$ are both semi-simple and nilpotent at once. Consequently, from (iii) $S - S' = N' - N = 0$. Thus, we finally get $S = S'$ and $N = N'$. That is, the decomposition is unique.

These complete the proof.

Theorem 10.6 implies that the matrix decomposition of (10.109) is unique. On the basis of Theorem 10.6, we investigate Jordan canonical forms of matrices in the next section.

## 10.6 Jordan Canonical Form

Once the vector space $V^n$ has been decomposed to a direct sum of generalized eigenspaces with the matrix reduced in parallel, we are able to deal with individual eigenspaces $\widetilde{W}_{\alpha_i} \, (1 \leq i \leq s)$ and the corresponding $A^{(i)} \, (1 \leq i \leq s)$ separately.

### 10.6.1 Canonical Form of Nilpotent Matrix

To avoid complication of notation, we think of a following example where we assume a $(n, n)$ matrix that operates on $V^n$: Let the nilpotent matrix be $N$: Suppose that $N^{v-1} \neq 0$ and $N^v = 0$ $(1 \leq v \leq n)$. If $v = 1$, the nilpotent matrix $N$ is a zero matrix. Notice that since a characteristic polynomial is described by $f_N(x) = x^n$, $f_N(N) = N^n = 0$ from Hamilton–Cayley Theorem. Let $W^{(i)}$ be given such that

$$W^{(i)} = \left\{\boldsymbol{x};\ \boldsymbol{x} \in V^n,\ N^i\boldsymbol{x} = \boldsymbol{0}\right\}.$$

Then we have

$$V^n = W^{(v)} \supset W^{(v-1)} \supset \cdots \supset W^{(1)} \supset W^{(0)} \equiv \{\boldsymbol{0}\}. \tag{10.116}$$

Note that when $v = 1$, we have trivially $V^n = W^{(v)} \supset W^{(0)} \equiv \{\boldsymbol{0}\}$. Let us put $\dim W^{(i)} = m_i$, $m_i - m_{i-1} = r_i (1 \leq i \leq v)$, $m_0 \equiv 0$. Then we can add $r_v$ linearly independent vectors $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots,$ and $\boldsymbol{a}_{r_v}$ to the basis vectors of $W^{(v-1)}$ so that those $r_v$ vectors can be basis vectors of $W^{(v)}$. Unless $N = 0$ (i.e., zero matrix), we must have at least one such vector; from the supposition with $^{\exists}\boldsymbol{x} \neq \boldsymbol{0}$ we have $N^{v-1}\boldsymbol{x} \neq \boldsymbol{0}$, and so $\boldsymbol{x} \notin W^{(v-1)}$. At least one such vector $\boldsymbol{x}$ is present and it is eligible for a basis vector of $W^{(v)}$. Hence, $r_v \geq 1$ and we have

$$W^{(v)} = \text{Span}\ \{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_{r_v}\} \oplus W^{(v-1)}. \tag{10.117}$$

Note that (10.117) is expressed as a direct sum. Meanwhile, $N\boldsymbol{a}_1, N\boldsymbol{a}_2, \ldots, N\boldsymbol{a}_{r_v} \in W^{(v-1)}$. In fact, suppose that $\boldsymbol{x} \in W^{(v)}$, i.e., $N^v\boldsymbol{x} = N^{v-1}(N\boldsymbol{x}) = \boldsymbol{0}$. That is, $N\boldsymbol{x} \in W^{(v-1)}$.

According to a similar reasoning made above, we have

$$\text{Span}\{N\boldsymbol{a}_1, N\boldsymbol{a}_2, \ldots, N\boldsymbol{a}_{r_v}\} \cap W^{(v-2)} = \{\boldsymbol{0}\}.$$

Moreover, these $r_v$ vectors $N\boldsymbol{a}_1, N\boldsymbol{a}_2, \ldots,$ and $N\boldsymbol{a}_{r_v}$ are linearly independent. Suppose that

$$c_1 N\boldsymbol{a}_1 + c_2 N\boldsymbol{a}_2 + \cdots + c_{r_v} N\boldsymbol{a}_{r_v} = \boldsymbol{0}. \tag{10.118}$$

Operating $N^{v-2}$ from the left, we have

$$N^{v-1}(c_1\boldsymbol{a}_1 + c_2\boldsymbol{a}_2 + \cdots + c_{r_v}\boldsymbol{a}_{r_v}) = \boldsymbol{0}.$$

This would imply that $c_1\boldsymbol{a}_1 + c_2\boldsymbol{a}_2 + \cdots + c_{r_v}\boldsymbol{a}_{r_v} \in W^{(v-1)}$. On the basis of the above argument, however, we must have $c_1 = c_2 = \cdots = c_{r_v} = 0$. In other words,

if $c_i\,(1 \le i \le r_v)$ were nonzero, we would have $\boldsymbol{a}_i \in W^{(v-1)}$, in contradiction. From (10.118), this means linear independence of $N\boldsymbol{a}_1, N\boldsymbol{a}_2, \ldots,$ and $N\boldsymbol{a}_{r_v}$.

As $W^{(v-1)} \supset W^{(v-2)}$, we may well have additional linearly independent vectors within the basis vectors of $W^{(v-1)}$. Let those vectors be $\boldsymbol{a}_{r_v+1}, \ldots, \boldsymbol{a}_{r_{v-1}}$. Here we assume that the number of such vectors is $r_{v-1} - r_v$. We have $r_{v-1} - r_v \ge 0$ accordingly. In this way, we can construct basis vectors of $W^{(v-1)}$ by including $\boldsymbol{a}_{r_v+1}, \ldots,$ and $\boldsymbol{a}_{r_{v-1}}$ along with $N\boldsymbol{a}_1, N\boldsymbol{a}_2, \ldots,$ and $N\boldsymbol{a}_{r_v}$. As a result, we get

$$W^{(v-1)} = \text{Span}\,\{N\boldsymbol{a}_1, \ldots, N\boldsymbol{a}_{r_v}, \boldsymbol{a}_{r_v+1}, \ldots, \boldsymbol{a}_{r_{v-1}}\} \oplus W^{(v-2)}.$$

We can repeat these processes to construct $W^{(v-2)}$ such that

$$W^{(v-2)} = \text{Span}\,\left\{N^2\boldsymbol{a}_1, \ldots, N^2\boldsymbol{a}_{r_v}, N\boldsymbol{a}_{r_v+1}, \ldots, N\boldsymbol{a}_{r_{v-1}}, \boldsymbol{a}_{r_{v-1}+1}, \ldots, \boldsymbol{a}_{r_{v-2}}\right\} \oplus W^{(v-3)}.$$

For $W^{(i)}$, furthermore, we have

$$W^{(i)} = \text{Span}\,\left\{N^{v-i}\boldsymbol{a}_1, \ldots, N^{v-i}\boldsymbol{a}_{r_v}, N^{v-i-1}\boldsymbol{a}_{r_v+1}, \ldots, N^{v-i-1}\boldsymbol{a}_{r_{v-1}}, \ldots, \boldsymbol{a}_{r_{i+1}+1}, \ldots, \boldsymbol{a}_{r_i}\right\} \oplus W^{(i-1)}. \tag{10.119}$$

Further repeating the procedures, we exhaust all the $n$ basis vectors of $W^{(v)} = V^n$. These vectors are given as follows:

$$N^k\boldsymbol{a}_{r_{i+1}+1}, \ldots, N^k\boldsymbol{a}_{r_i}\ (1 \le i \le v; 0 \le k \le i-1).$$

At the same time, we have

$$0 \equiv r_{v+1} < 1 \le r_v \le r_{v-1} \le \cdots \le r_1. \tag{10.120}$$

Table 10.1 [3] shows the resulting structure of these basis vectors pertinent to Jordan blocks. In Table 10.1, if laterally counting basis vectors, from the top we have $r_v, r_{v-1}, \ldots, r_1$ vectors. Their sum is $n$. This is the same number as that vertically counted. The dimension $n$ of the vector space $V^n$ is thus given by

$$n = \sum_{i=1}^{v} r_i = \sum_{i=1}^{v} i(r_i - r_{i+1}). \tag{10.121}$$

Let us examine the structure of Table 10.1 more closely. More specifically, let us inspect the $i$-layered structures of $(r_i - r_{i+1})$ vectors. Picking up a vector from among $\boldsymbol{a}_{r_{i+1}+1}, \ldots, \boldsymbol{a}_{r_i}$, we call it $\boldsymbol{a}_\rho$. Then, we get following set of vectors $\boldsymbol{a}_\rho, N\boldsymbol{a}_\rho\, N^2\boldsymbol{a}_\rho, \ldots, N^{i-1}\boldsymbol{a}_\rho$ in the $i$-layered structure. These $i$ vectors are displayed

**Table 10.1** Jordan blocks and structure of basis vectors for a nilpotent matrix

| | $r_v \downarrow$ | | | | | | |
|---|---|---|---|---|---|---|---|
| $r_v \leftarrow$ | $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{r_v}$ | $r_{v-1} - r_v \downarrow$ | | | | | |
| $r_{v-1} \leftarrow$ | $N\boldsymbol{a}_1, \ldots, N\boldsymbol{a}_{r_v}$ | $\boldsymbol{a}_{r_v+1}, \ldots, \boldsymbol{a}_{r_{v-1}}$ | | | | | |
| $\ldots$ | $N^2\boldsymbol{a}_1, \ldots, N^2\boldsymbol{a}_{r_v}$ | $N\boldsymbol{a}_{r_v+1}, \ldots, N\boldsymbol{a}_{r_{v-1}}$ | | | | | |
| $\ldots$ | $\ldots$ | $\ldots$ | $\ldots$ | $r_i - r_{i+1} \downarrow$ | | | |
| $r_i \leftarrow$ | $N^{v-i}\boldsymbol{a}_1, \ldots, N^{v-i}\boldsymbol{a}_{r_v}$ | $N^{v-i-1}\boldsymbol{a}_{r_v+1}, \ldots, N^{v-i-1}\boldsymbol{a}_{r_{v-1}}$ | $\cdots$ | $\boldsymbol{a}_{r_{i+1}+1}, \ldots, \boldsymbol{a}_{r_i}$ | | | |
| $r_{i-1} \leftarrow$ | $N^{v-i-1}\boldsymbol{a}_1, \ldots, N^{v-i-1}\boldsymbol{a}_{r_v}$ | $N^{v-i-2}\boldsymbol{a}_{r_v+1}, \ldots, N^{v-i-2}\boldsymbol{a}_{r_{v-1}}$ | $\cdots$ | $N\boldsymbol{a}_{r_{i+1}+1}, \ldots, N\boldsymbol{a}_{r_i}$ | | | |
| $\ldots$ | $\ldots$ | $\ldots$ | $\ldots$ | $\ldots$ | | $r_2 - r_3 \downarrow$ | |
| $r_2 \leftarrow$ | $N^{v-2}\boldsymbol{a}_1, \ldots, N^{v-2}\boldsymbol{a}_{r_v}$ | $N^{v-3}\boldsymbol{a}_{r_v+1}, \ldots, N^{v-3}\boldsymbol{a}_{r_{v-1}}$ | $\cdots$ | $N^{i-2}\boldsymbol{a}_{r_{i+1}+1}, \ldots, N^{i-2}\boldsymbol{a}_{r_i}$ | $\cdots$ | $\boldsymbol{a}_{r_3+1}, \ldots, \boldsymbol{a}_{r_2}$ | $r_1 - r_2 \downarrow$ |
| $r_1 \leftarrow$ | $N^{v-1}\boldsymbol{a}_1, \ldots, N^{v-1}\boldsymbol{a}_{r_v}$ | $N^{v-2}\boldsymbol{a}_{r_v+1}, \ldots, N^{v-2}\boldsymbol{a}_{r_{v-1}}$ | $\cdots$ | $N^{i-1}\boldsymbol{a}_{r_{i+1}+1}, \ldots, N^{i-1}\boldsymbol{a}_{r_i}$ | $\cdots$ | $N\boldsymbol{a}_{r_3+1}, \ldots, N\boldsymbol{a}_{r_2}$ | $\boldsymbol{a}_{r_2+1}, \ldots, \boldsymbol{a}_{r_1}$ |
| $n = \sum_{k=1}^{v} r_k$ | $vr_v$ | $(v-1)(r_{v-1} - r_v)$ | $\ldots$ | $i(r_i - r_{i+1})$ | $\ldots$ | $2(r_2 - r_3)$ | $r_1 - r_2$ |

Adapted from Satake I (1974) Linear algebra (Mathematics Library 1: in Japanese), with the permission of Shokabo Co., Ltd., Tokyo

"vertically" in Table 10.1. These vectors are linearly independent (see Theorem 10.4) and form an $i$-dimensional $N$-invariant subspace; i.e., $\mathrm{Span}\{N^{i-1}\boldsymbol{a}_\rho, N^{i-2}\boldsymbol{a}_\rho, \ldots, N\boldsymbol{a}_\rho, \boldsymbol{a}_\rho\}$, where $r_{i+1}+1 \le \rho \le r_i$. Matrix representation of the linear transformation $N$ with respect to the set of these $i$ vectors is

$$\begin{aligned}&(N^{i-1}\boldsymbol{a}_\rho N^{i-2}\boldsymbol{a}_\rho \ldots N\boldsymbol{a}_\rho \boldsymbol{a}_\rho)N \\ &= (N^{i-1}\boldsymbol{a}_\rho N^{i-2}\boldsymbol{a}_\rho \ldots N\boldsymbol{a}_\rho \boldsymbol{a}_\rho)\begin{pmatrix} 0 & 1 & & & & & \\ & 0 & 1 & & & & \\ & & 0 & & & & \\ & \vdots & & \ddots & & \vdots & \\ & & & & 0 & 1 & \\ & & & \cdots & & 0 & 1 \\ & & & & & & 0 \end{pmatrix}.\end{aligned} \tag{10.122}$$

These $(i, i)$ matrices of (10.122) are called $i$th order Jordan blocks. Notice that the number of those Jordan blocks is $(r_i - r_{i+1})$. Let us expressly define this number as [3]

$$J_i = r_i - r_{i+1}, \tag{10.123}$$

where $J_i$ is the number of the $i$th order Jordan blocks. The total number of Jordan blocks within a whole vector space $V^n = W^{(v)}$ is

$$\sum_{i=1}^{v} J_i = \sum_{i=1}^{v} (r_i - r_{i+1}) = r_1. \tag{10.124}$$

Recalling the dimension theorem mentioned in (9.45), we have

$$\begin{aligned}\dim V^n &= \dim \mathrm{Ker}\ \ N^i + \dim N^i(V^n) \\ &= \dim \mathrm{Ker}\, N^i + \mathrm{rank}\, N^i.\end{aligned} \tag{10.125}$$

Meanwhile, since $W^{(i)} = \mathrm{Ker}\, N^i$, $\dim W^{(i)} = m_i = \dim\ \mathrm{Ker}\, N^i$. From (10.116), $m_0 \equiv 0$. Then (9.45) is now read as

$$\dim V^n = m_i + \mathrm{rank}\, N^i. \tag{10.126}$$

That is

$$n = m_i + \mathrm{rank}\, N^i. \tag{10.127}$$

or

$$m_i = n - \operatorname{rank} N^i. \tag{10.128}$$

Meanwhile, from Table 10.1 we have

$$\dim W^{(i)} = m_i = \sum_{k=1}^{i} r_k, \ \dim W^{(i-1)} = m_{i-1} = \sum_{k=1}^{i-1} r_k. \tag{10.129}$$

Hence, we have

$$r_i = m_i - m_{i-1}. \tag{10.130}$$

Then we get [3].

$$\begin{aligned} J_i &= r_i - r_{i+1} = (m_i - m_{i-1}) - (m_{i+1} - m_i) = 2m_i - m_{i-1} - m_{i+1} \\ &= 2\left(n - \operatorname{rank} N^i\right) - \left(n - \operatorname{rank} N^{i-1}\right) - \left(n - \operatorname{rank} N^{i+1}\right) \\ &= \operatorname{rank} N^{i-1} + \operatorname{rank} N^{i+1} - 2\operatorname{rank} N^i. \end{aligned} \tag{10.131}$$

The number $J_i$ is therefore defined uniquely by $N$. The total number of Jordan blocks $r_1$ is also computed using (10.128) and (10.130) as

$$r_1 = m_1 - m_0 = m_1 = n - \operatorname{rank} N = \dim \operatorname{Ker} N, \tag{10.132}$$

where the last equality arises from the dimension theorem expressed as (9.45).

In Table 10.1, moreover, we have two extreme cases. That is, if $\nu = 1$ in (10.116), i.e., $N = 0$, from (10.132) we have $r_1 = n$ and $r_2 = \cdots = r_n = 0$; see Fig. 10.1a. Also we confirm $n = \sum_{i=1}^{\nu} r_i$ in (10.121). In this case, all the eigenvectors are proper

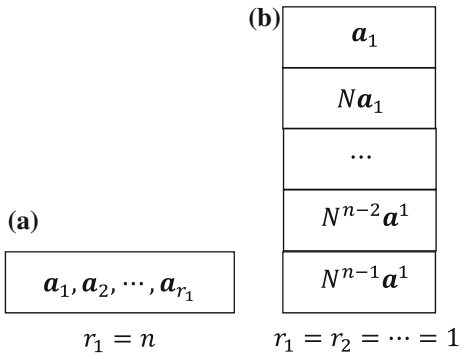


**Fig. 10.1** Examples of the structure of Jordan blocks. **a** $r_1 = n$. **b** $r_1 = r_2 = \cdots = r_n = 1$

eigenvectors with multiplicity of $n$ and we have $n$ first-order Jordan blocks. The other is the case of $v = n$. In that case, we have

$$r_1 = r_2 = \cdots = r_n = 1. \tag{10.133}$$

In the latter case, we also have $n = \sum_{i=1}^{v} r_i$ in (10.121). We have only one proper eigenvector and $(n-1)$ generalized eigenvectors; see Fig. 10.1b. From (10.132), we have this special case where we have only one $n$th order Jordan block, when rank $N = n - 1$.

### 10.6.2 Jordan Blocks

Let us think of (10.81) on the basis of (10.102). Picking up $A^{(i)}$ from (10.102) and considering (10.108), we put

$$N_i = A^{(i)} - \alpha_i E_{n_i} \ (1 \le i \le s), \tag{10.134}$$

where $E_{n_i}$ denotes $(n_i, n_i)$ identity matrix. We express a nilpotent matrix as $N_i$ as before. In (10.134), the number $n_i$ corresponds to $n$ in $V^n$ of Sect. 10.6.1. As $N_i$ is a $(n_i, n_i)$ matrix, we have

$$N_i^{n_i} = 0.$$

Here we are speaking of $v$th order nilpotent matrices $N_i$ such that $N_i^{v-1} \neq 0$ and $N_i^v = 0$ $(1 \le v \le n_i)$. We can deal with $N_i$ in a manner fully consistent with the theory we developed in Sect. 10.6.1. Each $A^{(i)}$ comprises one or more Jordan blocks $A^{(\kappa)}$ that is expressed as

$$A^{(\kappa)} = N_{\kappa_i} + \alpha_i E_{\kappa_i} \ (1 \le \kappa_i \le n_i), \tag{10.135}$$

where $A^{(\kappa)}$ denotes the $\kappa$th Jordan block in $A^{(i)}$. In $A^{(\kappa)}$, $N_{\kappa_i}$ and $E_{\kappa_i}$ are nilpotent $(\kappa_i, \kappa_i)$ matrix and $(\kappa_i, \kappa_i)$identity matrix, respectively. In $N_{\kappa_i}$, $\kappa_i$ zeros are displayed on the principal diagonal and entries of 1 are positioned on the matrix element next above the principal diagonal. All other entries are zero; see, e.g., a matrix of (10.122). As in Sect. 10.6.1, the number $\kappa_i$ is called a dimension of the Jordan block. Thus, $A^{(i)}$ of (10.102) can further be reduced to segmented matrices $A^{(\kappa)}$. Our next task is to find out how many Jordan blocks are contained in individual $A^{(i)}$ and what is the dimension of those Jordan blocks.

Corresponding to (10.122), the matrix representation of the linear transformation by $A^{(\kappa)}$ with respect to the set of $\kappa_i$ vectors is

$$
\begin{aligned}
&([A^{(\kappa)} - \alpha_i E_{\kappa_i}]^{\kappa_i - 1} \boldsymbol{a}_\sigma [A^{(\kappa)} - \alpha_i E_{\kappa_i}]^{\kappa_i - 2} \boldsymbol{a}_\sigma \ldots \boldsymbol{a}_\sigma)[A^{(\kappa)} - \alpha_i E_{\kappa_i}] \\
&= ([A^{(\kappa)} - \alpha_i E_{\kappa_i}]^{\kappa_i - 1} \boldsymbol{a}_\sigma [A^{(\kappa)} - \alpha_i E_{\kappa_i}]^{\kappa_i - 2} \boldsymbol{a}_\sigma \ldots \boldsymbol{a}_\sigma) \\
&\quad \times \begin{pmatrix}
0 & 1 & & & & & \\
& 0 & 1 & & & & \\
& & 0 & & & & \\
\vdots & & & \ddots & & \vdots & \\
& & & & 0 & 1 & \\
& & & \cdots & & 0 & 1 \\
& & & & & & 0
\end{pmatrix}.
\end{aligned} \tag{10.136}
$$

A vector $\boldsymbol{a}_\sigma$ stands for a vector associated with the $\kappa$th Jordan block of $A^{(i)}$. From (10.136), we obtain

$$
\left[A^{(\kappa)} - \alpha_i E_{\kappa_i}\right] \left\{ \left[A^{(\kappa)} - \alpha_i E_{\kappa_i}\right]^{\kappa_i - 1} \boldsymbol{a}_\sigma \right\} = 0. \tag{10.137}
$$

Namely,

$$
A^{(\kappa)} \left\{ \left[A^{(\kappa)} - \alpha_i E_{\kappa_i}\right]^{\kappa_i - 1} \boldsymbol{a}_\sigma \right\} = \alpha_i \left\{ \left[A^{(\kappa)} - \alpha_i E_{\kappa_i}\right]^{\kappa_i - 1} \boldsymbol{a}_\sigma \right\}. \tag{10.138}
$$

This shows that $\left[A^{(\kappa)} - \alpha_i E_{\kappa_i}\right]^{\kappa_i - 1} \boldsymbol{a}_\sigma$ is a proper eigenvector of $A^{(\kappa)}$ that corresponds to an eigenvalue $\alpha_i$. On the other hand, $\boldsymbol{a}_\sigma$ is a generalized eigenvector of rank $\kappa_i$. There are another $(\kappa_i - 2)$ generalized eigenvectors of $\left[A^{(\kappa)} - \alpha_i E_{\kappa_i}\right]^{\mu} \boldsymbol{a}_\sigma$ $(1 \leq \mu \leq \kappa_i - 2)$. In total, there are $\kappa_i$ eigenvectors [a proper eigenvector and $(\kappa_i - 1)$ generalized eigenvectors]. Also we see that the sole proper eigenvector can be found for each Jordan block.

In reference to these $\kappa_i$ eigenvectors as the basis vectors, the $(\kappa_i, \kappa_i)$-matrix $A^{(\kappa)}$ (i.e., a Jordan block) is expressed as

$$
A^{(\kappa)} = \begin{pmatrix}
\alpha_i & 1 & & & & & \\
& \alpha_i & 1 & & & & \\
& & \alpha_i & \ddots & & & \\
\vdots & & & \ddots & \ddots & \vdots & \\
& & & & \alpha_i & 1 & \\
& & & \cdots & & \alpha_i & 1 \\
& & & & & & \alpha_i
\end{pmatrix}. \tag{10.139}
$$

A$(n_i, n_i)$-matrix $A^{(i)}$ of (10.102) pertinent to an eigenvalue $\alpha_i$ contains a direct sum of Jordan blocks whose dimension ranges from 1 to $n_i$. The largest possible number of Jordan blocks of dimension $d$ (that satisfies $[\frac{n_i}{2} + 1] \leq d \leq n_i$, where $[\mu]$ denotes a largest integer that does not exceed $\mu$) is at most one.

An example depicted below is a matrix $A^{(i)}$ that explicitly includes two one-dimensional Jordan blocks, a (3, 3) three-dimensional Jordan block, and a (5, 5) five-dimensional Jordan block:

$$A^{(i)} \sim \begin{pmatrix} \alpha_i & 0 & & & & & & & & \\ & \alpha_i & 0 & & & & & & & \\ & & \alpha_i & 1 & & \cdots & & & & \\ & & & \alpha_i & 1 & & & & & \\ & \vdots & & & \alpha_i & 0 & & \vdots & & \\ & & & & & \alpha_i & 1 & & & \\ & & & & & & \alpha_i & 1 & & \\ & & & & & \cdots & & \alpha_i & 1 & \\ & & & & & & & & \alpha_i & 1 \\ & & & & & & & & & \alpha_i \end{pmatrix},$$

where $A^{(i)}$ is a (10,10) upper triangle matrix in which $\alpha_i$ is displayed on the principal diagonal with entries 0 or 1 on the matrix element next above the principal diagonal with all other entries being zero.

Theorem 10.1 shows that every $(n, n)$ square matrix can be converted to a triangle matrix by suitable similarity transformation. Diagonal elements give eigenvalues. Furthermore, Theorem 10.5 ensures that $A$ can be reduced to generalized eigenspaces $\widetilde{W}_{\alpha_i}$ $(1 \leq i \leq s)$ according to individual eigenvalues. Suppose for example that after a suitable similarity transformation a full matrix $A$ is represented as

$$A \sim \begin{pmatrix} \alpha_1 & & & & & & & & & & & \\ & \alpha_2 & * & * & & & & & & & & \\ & & \alpha_2 & * & & & & & & & & \\ & & & \alpha_2 & & & & & & & & \\ & & & & \ddots & & & & & & & \\ & & & & & \alpha_i & * & * & * & & & \\ & & & & & & \alpha_i & * & * & & & \\ & & & & & & & \alpha_i & * & & & \\ & & & & & & & & \alpha_i & & & \\ & & & & & & & & & \ddots & & \\ & & & & & & & & & & \alpha_s & * \\ & & & & & & & & & & & \alpha_s \end{pmatrix}, \qquad (10.140)$$

where

$$A = A^{(1)} \oplus A^{(2)} \oplus \cdots \oplus A^{(i)} \oplus \cdots \oplus A^{(s)}. \tag{10.141}$$

In (10.141), $A^{(1)}$ is a (1,1) matrix (i.e., simply a number); $A^{(2)}$ is a (3,3) matrix; $\cdots A^{(i)}$ is a (4,4) matrix; $\cdots A^{(s)}$ is a (2, 2) matrix.

The above matrix form allows us to further deal with segmented triangle matrices separately. In the case of (10.140), we may use a following matrix for similarity transformation:

$$\widetilde{P} = \begin{pmatrix} 1 & & & & & & & & & & & \\ & 1 & & & & & & & & & & \\ & & 1 & & & & & & & & & \\ & & & 1 & & & & & & & & \\ & & & & \ddots & & & & & & & \\ & & & & & p_{11} & p_{12} & p_{13} & p_{14} & & & \\ & & & & & p_{21} & p_{22} & p_{23} & p_{24} & & & \\ & & & & & p_{31} & p_{32} & p_{33} & p_{34} & & & \\ & & & & & p_{41} & p_{42} & p_{43} & p_{44} & & & \\ & & & & & & & & & \ddots & & \\ & & & & & & & & & & 1 & \\ & & & & & & & & & & & 1 \end{pmatrix}.$$

where a (4,4) matrix $P$ given by

$$P = \begin{pmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \\ p_{41} & p_{42} & p_{43} & p_{44} \end{pmatrix}.$$

is a non-singular matrix. The matrix $P$ is to be operated on $A^{(i)}$ so that we can separately perform the similarity transformation with respect to a (4, 4) nilpotent matrix $A^{(i)} - \alpha_i E_4$ following the procedures mentioned Sect. 10.6.1.

Thus, only an $\alpha_i$-associated segment can be treated with other segments left unchanged. In a similar fashion, we can consecutively deal with matrix segments related to other eigenvalues. In a practical case, however, it is more convenient to seek different eigenvalues and corresponding (generalized) eigenvectors at once and convert the matrix to Jordan canonical form. To make a guess about the structure of a matrix, however, the following argument will be useful. Let us think of an example after that.

Using (10.140), we have

$$A-\alpha_i E \sim \begin{pmatrix} \alpha_1-\alpha_i & & & & & & & & & & & & \\ & \alpha_2-\alpha_i & * & * & & & & & & & & & \\ & & \alpha_2-\alpha_i & * & & & & & & & & & \\ & & & \alpha_2-\alpha_i & & & & & & & & & \\ & & & & \ddots & & & & & & & & \\ & & & & & 0 & * & * & * & & & & \\ & & & & & & 0 & * & * & & & & \\ & & & & & & & 0 & * & & & & \\ & & & & & & & & 0 & & & & \\ & & & & & & & & & \ddots & & & \\ & & & & & & & & & & \alpha_s-\alpha_i & * & \\ & & & & & & & & & & & \alpha_s-\alpha_i & \end{pmatrix}.$$

Here we are treating the $(n, n)$ matrix $A$ on $V^n$. Note that a matrix $A-\alpha_i E$ is not nilpotent as a whole. Suppose that the multiplicity of $\alpha_i$ is $n_i$; in (10.140) $n_i = 4$. Since eigenvalues $\alpha_1, \alpha_2, \ldots,$ and $\alpha_s$ take different values from one another, $\alpha_1-\alpha_i, \alpha_2-\alpha_i, \ldots,$ and $\alpha_s-\alpha_i \neq 0$. In a triangle matrix diagonal element gives eigenvalues and, hence, $\alpha_1-\alpha_i, \alpha_2-\alpha_i, \ldots,$ and $\alpha_s-\alpha_i$ are nonzero eigenvalues of $A-\alpha_i E$. We rewrite (10.140) as

$$A-\alpha_i E \sim \begin{pmatrix} M^{(1)} & & & & & \\ & M^{(2)} & & & & \\ & & \ddots & & & \\ & & & M^{(i)} & & \\ & & & & \ddots & \\ & & & & & M^{(s)} \end{pmatrix}, \tag{10.142}$$

where $M^{(1)} = (\alpha_1-\alpha_i)$, $M^{(2)} = \begin{pmatrix} \alpha_2-\alpha_i & * & * \\ & \alpha_2-\alpha_i & * \\ & & \alpha_2-\alpha_i \end{pmatrix}, \cdots,$

$$M^{(i)} = \begin{pmatrix} 0 & * & * & * \\ & 0 & * & * \\ & & 0 & * \\ & & & 0 \end{pmatrix}, \ldots, M^{(s)} = \begin{pmatrix} \alpha_s-\alpha_i & * \\ & \alpha_s-\alpha_i \end{pmatrix}.$$

Thus, $M^{(p)}$ $(p \neq i)$ is a non-singular matrix and $M^{(i)}$ is a nilpotent matrix. Note that if we can find $\mu_i$ such that $\left[M^{(i)}\right]^{\mu_i - 1} \neq 0$ and $\left[M^{(i)}\right]^{\mu_i} = 0$, with a minimal polynomial $\varphi_{M^{(i)}}(x)$ for $M^{(i)}$, we have $\varphi_{M^{(i)}}(x) = x^{\mu_i}$. Consequently, we get

$$(A - \alpha_i E)^{\mu_i} \sim \begin{pmatrix} [M^{(1)}]^{\mu_i} & & & & & \\ & [M^{(2)}]^{\mu_i} & & & & \\ & & \ddots & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & [M^{(s)}]^{\mu_i} \end{pmatrix}. \tag{10.143}$$

In (10.143), diagonal elements of non-singular triangle matrices $\left[M^{(p)}\right]^{\mu_i}$ $(p \neq i)$ are $(\alpha_p - \alpha_i)^{\mu_i}$ $(\neq 0)$. Thus, we have a "perforated" matrix $(A - \alpha_i E)^{\mu_i}$ where $\left[M^{(i)}\right]^{\mu_i} = 0$ in (10.143). Putting

$$\Phi_A(x) \equiv \prod_{i=1}^{s} (x - \alpha_i)^{\mu_i},$$

we get

$$\Phi_A(A) \equiv \prod_{i=1}^{s} (A - \alpha_i E)^{\mu_i} = 0.$$

A polynomial $\Phi_A(x)$ gives a minimal polynomial for $f_A(x)$.

From the above argument, we can choose $\mu_i$ for $l_i$ in (10.82). Meanwhile, $M^{(i)}$ in (10.142) is identical to $N_i$ in (10.134). Rewriting (10.134), we get

$$A^{(i)} = M^{(i)} + \alpha_i E_{n_i}. \tag{10.144}$$

Let us think of matrices $\left[A^{(i)} - \alpha_i E_{n_i}\right]^k$ and $(A - \alpha_i E)^k$ $(k \leq \mu_i)$. From (9.45), we find

$$\dim V^n = n = \dim \ \mathrm{Ker}\,(A - \alpha_i E)^k + \mathrm{rank}\ (A - \alpha_i E)^k, \tag{10.145}$$

and

$$\dim\ \widetilde{W}_{\alpha_i} = n_i = \dim\ \mathrm{Ker}\ [A^{(i)} - \alpha_i E_{n_i}]^k + \mathrm{rank}\ [A^{(i)} - \alpha_i E_{n_i}]^k, \tag{10.146}$$

where $\mathrm{rank}\,(A - \alpha_i E)^k = \dim (A - \alpha_i E)^k (V^n)$ and $\mathrm{rank}\ [A^{(i)} - \alpha_i E_{n_i}]^k = \dim [A^{(i)} - \alpha_i E_{n_i}]^k \left(\widetilde{W}_{\alpha_i}\right)$.

Noting that

$$\dim \text{Ker } (A - \alpha_i E)^k = \dim \text{Ker } [A^{(i)} - \alpha_i E_{n_i}]^k, \tag{10.147}$$

we get

$$n - \text{rank } (A - \alpha_i E)^k = n_i - \text{rank } [A^{(i)} - \alpha_i E_{n_i}]^k. \tag{10.148}$$

This notable property comes from the non-singularity of $\left[M^{(p)}\right]^k$ ($p \neq i$, $k$: a positive integer); i.e., all the eigenvalues of $\left[M^{(p)}\right]^k$ are nonzero. In particular, as rank $[A^{(i)} - \alpha_i E_{n_i}]^{\mu_i} = 0$, from (10.148) we have

$$\text{rank}(A - \alpha_i E)^l = n - n_i \ (l \geq \mu_i). \tag{10.149}$$

Meanwhile, putting $k = 1$ in (10.148) and using (10.132) we get

$$\begin{aligned} \dim \text{Ker } [A^{(i)} - \alpha_i E_{n_i}] &= n_i - \text{rank } [A^{(i)} - \alpha_i E_{n_i}] \\ &= n - \text{rank } (A - \alpha_i E) = \dim \ \text{Ker } (A - \alpha_i E) \\ &= r_1^{(i)}. \end{aligned} \tag{10.150}$$

The value $r_1^{(i)}$ gives the number of Jordan blocks with an eigenvalue $\alpha_i$.

Moreover, we must consider a following situation. We know how the matrix $A^{(i)}$ in (10.134) is reduced to Jordan blocks of lower dimension. To get detailed information about it, however, we have to get the information about (generalized) eigenvalues corresponding to eigenvalues other than $\alpha_i$. In this context, (10.148) is useful. Equation (10.131) tells how the number of Jordan blocks in a nilpotent matrix is determined. If we can get this knowledge before we have found out all the (generalized) eigenvectors, it will be easier to address the problem. Let us rewrite (10.131) as

$$\begin{aligned} J_q^{(i)} &= r_q - r_{q+1} \\ &= \text{rank } [A^{(i)} - \alpha_i E_{n_i}]^{q-1} + \text{rank } [A^{(i)} - \alpha_i E_{n_i}]^{q+1} - 2\,\text{rank } [A^{(i)} - \alpha_i E_{n_i}]^q, \end{aligned} \tag{10.151}$$

where we define $J_q^{(i)}$ as the number of the $q$th order Jordan blocks within $A^{(i)}$. Note that these blocks are expressed as $(q, q)$ matrices. Meanwhile, using (10.148) $J_q^{(i)}$ is expressed as

$$J_q^{(i)} = \text{rank } (A - \alpha_i E)^{q-1} + \text{rank } (A - \alpha_i E)^{q+1} - 2\,\text{rank } (A - \alpha_i E)^q. \tag{10.152}$$

This relation can be obtained by replacing $k$ in (10.148) with $q - 1, q + 1$, and $q$, respectively and deleting $n$ and $n_i$ from these three relations. This enables us to gain

access to a whole structure of the linear transformation represented by the $(n, n)$ matrix $A$ without reducing it to subspaces.

To enrich our understanding of Jordan canonical forms, a following tangible example will be beneficial.

### 10.6.3 Example of Jordan Canonical Form

Let us think of a following matrix $A$:

$$A = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 1 & 3 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ -3 & -1 & -2 & 4 \end{pmatrix}. \tag{10.153}$$

The characteristic equation $f_A(x)$ is given by

$$\begin{aligned} f_A(x) &= \begin{vmatrix} x-1 & 1 & 0 & 0 \\ -1 & x-3 & 0 & 0 \\ 0 & 0 & x-2 & 0 \\ 3 & 1 & 2 & x-4 \end{vmatrix} \\ &= (x-4)(x-2)^3. \end{aligned} \tag{10.154}$$

Equating (10.154) to zero, we get an eigenvalue 4 as a simple root and an eigenvalue 2 as a triple root. The vector space $V^4$ is then decomposed to two invariant subspaces. The first is a one-dimensional kernel (or null-space) of the transformation $(A - 4E)$ and the other is a three-dimensional kernel of the transformation $(A - 2E)^3$. We have to seek eigenvectors that span these invariant subspaces.

(i) $x = 4$ :

An eigenvector belonging to the first invariant subspace must satisfy a proper eigenvalue equation since the eigenvalue 4 is simple root. This equation is expressed as

$$(A - 4E)\boldsymbol{x} = \boldsymbol{0}.$$

This reads in a matrix form as

$$\begin{pmatrix} -3 & -1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ -3 & -1 & -2 & 0 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix} = 0. \tag{10.155}$$

This is equivalent to a following set of four equations:

$$\begin{aligned} -3c_1 - c_2 &= 0, \\ c_1 - c_2 &= 0, \\ -2c_3 &= 0, \\ -3c_1 - c_2 - 2c_3 &= 0. \end{aligned}$$

These are equivalent to that $c_3 = 0$ and $c_1 = c_2 = -3c_1$. Therefore, $c_1 = c_2 = c_3 = 0$ with an arbitrarily chosen number of $c_4$, which is chosen as 1 as usual. Hence, designating the proper eigenvector as $\boldsymbol{e}_1^{(4)}$, its column vector representation is

$$\boldsymbol{e}_1^{(4)} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

A (4,4) matrix in (10.155) representing $(A - 4E)$ has a rank 3. The number of Jordan blocks for an eigenvalue 4 is given by (10.150) as

$$r_1^{(4)} = 4 - \text{rank}(A - 4E) = 1. \tag{10.156}$$

In this case, the Jordan block is naturally one-dimensional. In fact, using (10.152) we have

$$\begin{aligned} J_1^{(4)} &= \text{rank}\,(A - 4E)^0 + \text{rank}\,(A - 4E)^2 - 2\text{rank}\,(A - 4E) \\ &= 4 + 3 - 2 \times 3 = 1. \end{aligned} \tag{10.157}$$

In (10.157), $J_1^{(4)}$ gives the number of the first-order Jordan blocks for an eigenvalue 4. We used

$$(A - 4E)^2 = \begin{pmatrix} 8 & 4 & 0 & 0 \\ -4 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 8 & 4 & 4 & 0 \end{pmatrix} \tag{10.158}$$

and confirmed that $\text{rank}(A - 4E)^2 = 3$.

(ii) $x = 2$ :

The eigenvalue 2 has a triple root. Therefore, we must examine how the invariant subspaces can further be decomposed to subspaces of lower dimension. To this end, we first start with a secular equation expressed as

$$(A - 2E)\boldsymbol{x} = \mathbf{0}. \tag{10.159}$$

The matrix representation is

$$\begin{pmatrix} -1 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -3 & -1 & -2 & 2 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix} = 0.$$

This is equivalent to a following set of two equations:

$$\begin{aligned} &c_1 + c_2 = 0, \\ &-3c_1 - c_2 - 2c_3 + 2c_4 = 0. \end{aligned}$$

From the above, we can put $c_1 = c_2 = 0$ and $c_3 = c_4\ (= 1)$. The equations allow the existence of another proper eigenvector. For this we have $c_1 = -c_2 = 1$, $c_3 = 0$, and $c_4 = 1$. Thus, for the two proper eigenvectors corresponding to an eigenvalue 2, we get

$$\boldsymbol{e}_1^{(2)} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \quad \boldsymbol{e}_2^{(2)} = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \end{pmatrix}.$$

A dimension of the invariant subspace corresponding to an eigenvalue 2 is three (due to the triple root) and, hence, there should be one generalized eigenvector. To determine it, we examine a following matrix equation:

$$(A - 2E)^2\boldsymbol{x} = \mathbf{0}. \tag{10.160}$$

The matrix representation is

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -4 & 0 & -4 & 4 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix} = 0.$$

That is,

$$-c_1 - c_3 + c_4 = 0. \tag{10.161}$$

Furthermore, we have

$$(A-2E)^3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -8 & 0 & -8 & 8 \end{pmatrix}. \tag{10.162}$$

Moreover, $\mathrm{rank}(A-2E)^l\ (=1)$ remains unchanged for $l \geq 2$ as expected from (10.149).

It will be convenient to examine a structure of the invariant subspace. For this purpose, we seek the number of Jordan blocks $r_1^{(2)}$ and their order. Using (10.150), we have

$$r_1^{(2)} = 4 - \mathrm{rank}(A-2E) = 4-2 = 2. \tag{10.163}$$

The number of first-order Jordan blocks is

$$\begin{aligned} J_1^{(2)} &= \mathrm{rank}\,(A-2E)^0 + \mathrm{rank}\,(A-2E)^2 - 2\,\mathrm{rank}\,(A-2E) \\ &= 4+1-2\times 2 = 1. \end{aligned} \tag{10.164}$$

In turn, the number of second-order Jordan blocks is

$$\begin{aligned} J_2^{(2)} &= \mathrm{rank}\,(A-2E) + \mathrm{rank}\,(A-2E)^3 - 2\,\mathrm{rank}\,(A-2E)^2 \\ &= 2+1-2\times 1 = 1. \end{aligned} \tag{10.165}$$

In the above, $J_1^{(2)}$ and $J_2^{(2)}$ are obtained from (10.152). Thus, Fig. 10.2 gives a constitution of Jordan blocks for eigenvalues 4 and 2. The overall number of Jordan blocks is three; the number of the first-order and second-order Jordan blocks is two and one, respectively.

The proper eigenvector $\boldsymbol{e}_1^{(2)}$ is related to $J_1^{(2)}$ of (10.164). A set of the proper eigenvector $\boldsymbol{e}_2^{(2)}$ and the corresponding generalized eigenvector $\boldsymbol{g}_2^{(2)}$ is pertinent to $J_2^{(2)}$ of (10.165). We must have the generalized eigenvector $\boldsymbol{g}_2^{(2)}$ in such a way that

$$(A-2E)^2\boldsymbol{g}_2^{(2)} = (A-2E)\left[(A-2E)\boldsymbol{g}_2^{(2)}\right] = (A-2E)\boldsymbol{e}_2^{(2)} = 0. \tag{10.166}$$

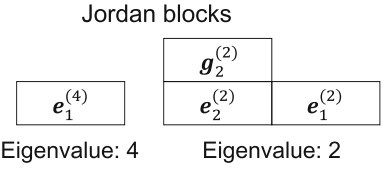


**Fig. 10.2** Structure of Jordan blocks of a matrix shown in (10.170)

From (10.161), we can put $c_1 = c_3 = c_4 = 0$ and $c_2 = -1$. Thus, the matrix representation of the generalized eigenvector $\boldsymbol{g}_2^{(2)}$ is

$$\boldsymbol{g}_2^{(2)} = \begin{pmatrix} 0 \\ -1 \\ 0 \\ 0 \end{pmatrix}. \tag{10.167}$$

We stress here that $\boldsymbol{e}_1^{(2)}$ is not eligible for a proper pair with $\boldsymbol{g}_2^{(2)}$ in $J_2$. It is because from (10.166) we have

$$(A - 2E)\boldsymbol{g}_2^{(2)} = \boldsymbol{e}_2^{(2)},\ (A - 2E)\boldsymbol{g}_2^{(2)} \neq \boldsymbol{e}_1^{(2)}. \tag{10.168}$$

Thus, we have determined a set of (generalized) eigenvectors. The matrix representation $R$ for the basis vectors transformation is given by

$$R = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & -1 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \end{pmatrix} \sim (\boldsymbol{e}_1^{(4)}\ \boldsymbol{e}_1^{(2)}\ \boldsymbol{e}_2^{(2)}\ \boldsymbol{g}_2^{(2)}), \tag{10.169}$$

where the symbol $\sim$ denotes the column vector representation; $\boldsymbol{e}_1^{(4)}, \boldsymbol{e}_1^{(2)}$, and $\boldsymbol{e}_2^{(2)}$ represent proper eigenvectors and $\boldsymbol{g}_2^{(2)}$ is a generalized eigenvector. Performing similarity transformation using this $R$, we get a following Jordan canonical form:

$$\begin{aligned} R^{-1}AR &= \begin{pmatrix} -1 & 0 & -1 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ -1 & -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 & 0 \\ 1 & 3 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ -3 & -1 & -2 & 4 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & -1 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}. \end{aligned} \tag{10.170}$$

The structure of Jordan blocks is shown in Fig. 10.2. Notice here that the trace of $A$ remains unchanged before and after similarity transformation.

Next, we consider column vector representations. According to (9.37), let us view the matrix $A$ as a linear transformation over $V^4$. Then $A$ is given by

$$
\begin{aligned}
A(\boldsymbol{x}) &= (\boldsymbol{e}_1\ \boldsymbol{e}_2\ \boldsymbol{e}_3\ \boldsymbol{e}_4)A\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \\
&= (\boldsymbol{e}_1\ \boldsymbol{e}_2\ \boldsymbol{e}_3\ \boldsymbol{e}_4)\begin{pmatrix} 1 & -1 & 0 & 0 \\ 1 & 3 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ -3 & -1 & -2 & 4 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix},
\end{aligned} \tag{10.171}
$$

where $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$, and $\boldsymbol{e}_4$ are basis vectors and $x_1$, $x_2$, $x_3$, and $x_4$ are corresponding coordinates of a vector $\boldsymbol{x} = \sum_{i=1}^{n} x_i \boldsymbol{e}_i$ $(\in V^4)$. We rewrite (10.171) as

$$
\begin{aligned}
A(\boldsymbol{x}) &= (\boldsymbol{e}_1\ \boldsymbol{e}_2\ \boldsymbol{e}_3\ \boldsymbol{e}_4)RR^{-1}ARR^{-1}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \\
&= (\boldsymbol{e}_1^{(4)}\ \boldsymbol{e}_1^{(2)}\ \boldsymbol{e}_2^{(2)}\ \boldsymbol{g}_2^{(2)})\begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}\begin{pmatrix} x_1' \\ x_2' \\ x_3' \\ x_4' \end{pmatrix},
\end{aligned} \tag{10.172}
$$

where we have

$$
\begin{aligned}
\left(\boldsymbol{e}_1^{(4)}\ \boldsymbol{e}_1^{(2)}\ \boldsymbol{e}_2^{(2)}\ \boldsymbol{g}_2^{(2)}\right) &= (\boldsymbol{e}_1\ \boldsymbol{e}_2\ \boldsymbol{e}_3\ \boldsymbol{e}_4)R \\
\begin{pmatrix} x_1' \\ x_2' \\ x_3' \\ x_4' \end{pmatrix} &= R^{-1}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}.
\end{aligned} \tag{10.173}
$$

After (9.84), we have

$$
\begin{aligned}
\boldsymbol{x} &= (\boldsymbol{e}_1\ \boldsymbol{e}_2\ \boldsymbol{e}_3\ \boldsymbol{e}_4)\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = (\boldsymbol{e}_1\ \boldsymbol{e}_2\ \boldsymbol{e}_3\ \boldsymbol{e}_4)RR^{-1}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \\
&= \left(\boldsymbol{e}_1^{(4)}\ \boldsymbol{e}_1^{(2)}\ \boldsymbol{e}_2^{(2)}\ \boldsymbol{g}_2^{(2)}\right)\begin{pmatrix} x_1' \\ x_2' \\ x_3' \\ x_4' \end{pmatrix}.
\end{aligned} \tag{10.174}
$$

As for $V^n$ in general, let us put $R = (p)_{ij}$ and $R^{-1} = (q)_{ij}$. Also represent $j$th (generalized) eigenvectors by a column vector and denote them by $p^{(j)}$. There we display individual (generalized) eigenvectors in the order of $\left(\boldsymbol{e}^{(1)}\ \boldsymbol{e}^{(2)} \cdots \boldsymbol{e}^{(j)} \cdots \boldsymbol{e}^{(n)}\right)$, where $\boldsymbol{e}^{(j)}$ $(1 \leq j \leq n)$ denotes either a proper eigenvector or a generalized eigenvector according to (10.174). Each $p^{(j)}$ is represented in reference to original basis vectors $(\boldsymbol{e}_1 \ldots \boldsymbol{e}_n)$. Then we have

$$R^{-1}p^{(j)} = \sum_{k=1}^{n} q_{ik}p_k^{(j)} = \delta_i^{(j)}, \tag{10.175}$$

where $\delta_i^{(j)}$ denotes a column vector to which only the $j$th row is 1, otherwise 0. Thus, a column vector $\delta_i^{(j)}$ is an "address" of $\boldsymbol{e}^{(j)}$ in reference to $\left(\boldsymbol{e}^{(1)}\ \boldsymbol{e}^{(2)} \ldots \boldsymbol{e}^{(j)} \ldots \boldsymbol{e}^{(n)}\right)$ taken as basis vectors.

In our present case, in fact, we have

$$\begin{aligned}
R^{-1}p^{(1)} &= \begin{pmatrix} -1 & 0 & -1 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ -1 & -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \\
R^{-1}p^{(2)} &= \begin{pmatrix} -1 & 0 & -1 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ -1 & -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \\
R^{-1}p^{(3)} &= \begin{pmatrix} -1 & 0 & -1 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ -1 & -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \\
R^{-1}p^{(4)} &= \begin{pmatrix} -1 & 0 & -1 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ -1 & -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ -1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.
\end{aligned} \tag{10.176}$$

In (10.176), $p^{(1)}$ is a column vector representation of $\boldsymbol{e}_1^{(4)}$; $p^{(2)}$ is a column vector representation of $\boldsymbol{e}_1^{(2)}$, and so on.

A minimal polynomial $\Phi_A(x)$ is expressed as

$$\Phi_A(x) = (x-4)(x-2)^2.$$

Readers can easily make sure of it.

We remark that a non-singular matrix $R$ pertinent to the similarity transformation is not uniquely determined, but we have arbitrariness. In (10.169), for example, if we adopt $-\boldsymbol{g}_2^{(2)}$ instead of $\boldsymbol{g}_2^{(2)}$, we should adopt $-\boldsymbol{e}_2^{(2)}$ instead of $\boldsymbol{e}_2^{(2)}$ accordingly. Thus, instead of $R$ in (10.169) we may choose

$$R' = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & -1 & 0 \end{pmatrix}. \tag{10.177}$$

In this case, we also get the same Jordan canonical form as before. That is,

$$R'^{-1}AR' = \begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

Suppose that we choose $R''$ such that

$$\begin{aligned}(\boldsymbol{e}_1\ \boldsymbol{e}_2\ \boldsymbol{e}_3\ \boldsymbol{e}_4)R'' &= (\boldsymbol{e}_1\ \boldsymbol{e}_2\ \boldsymbol{e}_3\ \boldsymbol{e}_4)\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & -1 & -1 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 \end{pmatrix} \\ &= (\boldsymbol{e}_1^{(2)}\ \boldsymbol{e}_2^{(2)}\ \boldsymbol{g}_2^{(2)}\ \boldsymbol{e}_1^{(4)}).\end{aligned}$$

In this case, we have

$$R''^{-1}AR'' = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}.$$

Note that we get a different disposition of the matrix elements from that of (10.172).

Next, we decompose $A$ into a semi-simple matrix and a nilpotent matrix. In (10.172), we had

$$R^{-1}AR = \begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

Defining

$$S = \begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} \quad \text{and} \quad N = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

we have

$$R^{-1}AR = S + N \quad \text{i.e.} \quad A = R(S+N)R^{-1} = RSR^{-1} + RNR^{-1}.$$

Performing the above matrix calculations and putting $\widetilde{S} = RSR^{-1}$ and $\widetilde{N} = RNR^{-1}$, we get

$$A = \widetilde{S} + \widetilde{N}$$

with

$$\widetilde{S} = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ -2 & 0 & -2 & 4 \end{pmatrix} \quad \text{and} \quad \widetilde{N} = \begin{pmatrix} -1 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & -1 & 0 & 0 \end{pmatrix}.$$

That is, we have

$$A = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 1 & 3 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ -3 & -1 & -2 & 4 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ -2 & 0 & -2 & 4 \end{pmatrix} + \begin{pmatrix} -1 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & -1 & 0 & 0 \end{pmatrix}. \tag{10.178}$$

Even though matrix forms $S$ and $N$ differ depending on the choice of different matrix forms of similarity transformation $R$ (namely, $R', R''$ represented above), the decomposition (10.178) is unique. That is, $\widetilde{S}$ and $\widetilde{N}$ are uniquely determined, once a matrix $A$ is given. The confirmation is left for readers as an exercise.

We present another simple example. Let $A$ be a matrix described as

$$A = \begin{pmatrix} 0 & 4 \\ -1 & 4 \end{pmatrix}.$$

Eigenvalues of $A$ are 2 as a double root. According to routine, we have an eigenvector $\boldsymbol{e}_1^{(2)}$ as a column vector, e.g.,

$$\boldsymbol{e}_1^{(2)} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

Another eigenvector is a generalized eigenvector $\boldsymbol{g}_1^{(2)}$ of rank 2. This can be decided such that

$$(A - 2E)\boldsymbol{g}_1^{(2)} = \begin{pmatrix} -2 & 4 \\ -1 & 2 \end{pmatrix}\boldsymbol{g}_1^{(2)} = \boldsymbol{e}_1^{(2)}.$$

As an option, we get

$$\boldsymbol{g}_1^{(2)} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Thus, we can choose $R$ for a diagonalizing matrix together with an inverse matrix $R^{-1}$ such that

$$R = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}, \quad \text{and} \quad R^{-1} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}.$$

Therefore, with a Jordan canonical form we have

$$R^{-1}AR = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}. \tag{10.179}$$

As before, putting

$$S = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \quad \text{and} \quad N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$

we have

$$R^{-1}AR = S + N \quad \text{i.e.} \quad A = R(S+N)R^{-1} = RSR^{-1} + RNR^{-1}.$$

Putting $\widetilde{S} = RSR^{-1}$ and $\widetilde{N} = RNR^{-1}$, we get

$$A = \widetilde{S} + \widetilde{N}$$

with

$$\widetilde{S} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \quad \text{and} \quad \widetilde{N} = \begin{pmatrix} -2 & 4 \\ -1 & 2 \end{pmatrix}. \tag{10.180}$$

We may also choose $R'$ for a diagonalizing matrix together with an inverse matrix $R'^{-1}$ instead of $R$ and $R^{-1}$, respectively, such that we have, e.g.,

$$R' = \begin{pmatrix} 2 & -3 \\ 1 & -1 \end{pmatrix}, \quad \text{and} \quad R'^{-1} = \begin{pmatrix} -1 & 3 \\ -1 & 2 \end{pmatrix}.$$

Using these matrices, we get exactly the same Jordan canonical form and the matrix decomposition as (10.179) and (10.180). Thus, again we find that the matrix decomposition is unique.

Another simple example is a lower triangle matrix

$$A = \begin{pmatrix} 2 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Following now familiar procedures, as a diagonalizing matrix we have, e.g.,

$$R = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad R^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Then, we get

$$R^{-1}AR = S = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Therefore, the "decomposition" is

$$A = RSR^{-1} = \begin{pmatrix} 2 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

where the first term is a semi-simple matrix and the second is a nilpotent matrix (i.e., zero matrix). Thus, the decomposition is once again unique.

## 10.7 Diagonalizable Matrices

Among canonical forms of matrices, the simplest form is a diagonalizable matrix. Here we define a diagonalizable matrix as a matrix that can be converted to that whose off-diagonal elements are zero. In Sect. 10.5 we have investigated different properties of the matrices. In this Section we examine basic properties of diagonalizable matrices.

In Sect. 10.6.1 we have shown that $\mathrm{Span}\{N^{i-1}\boldsymbol{a}_\rho,\ N^{i-2}\boldsymbol{a}_\rho, \ldots N\boldsymbol{a}_\rho,\ \boldsymbol{a}_\rho\}$ forms a $N$-invariant subspace of a dimension $i$, where $\boldsymbol{a}_\rho$ satisfies the relations $N^i\boldsymbol{a}_\rho = \boldsymbol{0}$ and $N^{i-1}\boldsymbol{a}_\rho \neq \boldsymbol{0}\ (r_{i+1}+1 \leq j \leq r_i)$ as in (10.122). Of these vectors, only $N^{i-1}\boldsymbol{a}_\rho$ is a sole proper eigenvector that is accompanied by $(i-1)$ generalized eigenvectors. Note that only the proper eigenvector can construct one-dimensional $N$-invariant subspace by itself. This is because regarding other generalized eigenvectors $\boldsymbol{g}$ (here $\boldsymbol{g}$ stands for all of generalized eigenvectors), $N\boldsymbol{g}(\neq 0)$ and $\boldsymbol{g}$ are linearly independent. Note that with a proper eigenvector $\boldsymbol{e}$, we have $N\boldsymbol{e} = \boldsymbol{0}$. A corresponding Jordan block is represented by a matrix as given in (10.122) in reference to the basis vectors comprising these $i$ eigenvectors. Therefore, if a $(n, n)$ matrix $A$ has only proper eigenvectors, all Jordan blocks are one-dimensional. This means that $A$ is diagonalizable. That $A$ has only proper eigenvectors is equivalent to that those eigenvectors form a one-dimensional subspace and that $V^n$ is a direct sum of the subspaces spanned by individual proper eigenvectors. In other words, if $V^n$ is a direct sum of subspaces (i.e., eigenspaces) spanned by individual proper eigenvectors of $A$, $A$ is diagonalizable.

Next, suppose that $A$ is diagonalizable. Then, after an appropriate similarity transformation with a non-singular matrix $P$, $A$ has a following form:

$$P^{-1}AP = \begin{pmatrix} \alpha_1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & \alpha_1 & & & & & & \\ & & & \alpha_2 & & & & & \\ & & & & \ddots & & & & \\ & & & & & \alpha_2 & & & \\ & & & & & & \ddots & & \\ & & & & & & & \alpha_s & \\ & & & & & & & & \ddots \\ & & & & & & & & & \alpha_s \end{pmatrix}. \tag{10.181}$$

In this case, let us examine what form a minimal polynomial $\varphi_A(x)$ for $A$ takes. A characteristic polynomial $f_A(x)$ for $A$ is invariant through similarity transformation, so is $\varphi_A(x)$. That is,

$$\varphi_{P^{-1}AP}(x) = \varphi_A(x). \tag{10.182}$$

From (10.181), we find that $A - \alpha_i E\ (1 \leq i \leq s)$ has a "perforated" form such as (10.143) with the diagonalized form unchanged. Then we have

$$(A - \alpha_1 E)(A - \alpha_2 E)\ldots(A - \alpha_s E) = 0. \tag{10.183}$$

This is because a product of matrices having only diagonal elements is merely a product of individual diagonal elements. Meanwhile, in virtue of Hamilton–Cayley Theorem, we have

$$f_A(A) = \prod_{i=1}^{s} (A - \alpha_i E)^{n_i} = 0.$$

Rewriting this expression, we have

$$(A - \alpha_1 E)^{n_1} (A - \alpha_2 E)^{n_2} \ldots (A - \alpha_s E)^{n_s} = 0. \tag{10.184}$$

In light of (10.183), this implies that a minima polynomial $\varphi_A(x)$ is expressed as

$$\varphi_A(x) = (x - \alpha_1)(x - \alpha_2) \ldots (x - \alpha_s). \tag{10.185}$$

Surely $\varphi_A(x)$ is in (10.185) has a lowest-order polynomial among those satisfying $f(A) = 0$ and a divisor of $f_A(x)$. Also $\varphi_A(x)$ has a highest-order coefficient of 1. Thus, $\varphi_A(x)$ should be a minimal polynomial of $A$ and we conclude that $\varphi_A(x)$ does not have a multiple root.

Then let us think how $V^n$ is characterized in case $\varphi_A(x)$ does not have a multiple root. This is equivalent to that $\varphi_A(x)$ is described by (10.185). To see this, suppose that we have two matrices $A$ and $B$ and let $BV^n = W$. We wish to use the following relation:

$$\begin{aligned} \text{rank}\,(AB) &= \dim ABV^n = \dim AW = \dim W - \dim (A^{-1}\{\mathbf{0}\} \cap W) \\ &\geq \dim W - \dim \left(A^{-1}\{\mathbf{0}\}\right) = \dim BV^n - (n - \dim AV^n) \\ &= \text{rank}\, A + \text{rank}\, B - n. \end{aligned} \tag{10.186}$$

In (10.186), the third equality comes from the fact that the domain of $A$ is restricted to $W$. Concomitantly, $A^{-1}\{\mathbf{0}\}$ is restricted to $A^{-1}\{\mathbf{0}\} \cap W$ as well; notice that $A^{-1}\{\mathbf{0}\} \cap W$ is a subspace. Considering these situations, we use a relation corresponding to that of (9.45). The fourth equality is due to the dimension theorem of (9.45). Applying (10.186) to (10.183) successively, we have

$$\begin{aligned} 0 &= \text{rank}\, [(A - \alpha_1 E)(A - \alpha_2 E) \ldots (A - \alpha_s E)] \\ &\geq \text{rank}\, (A - \alpha_1 E) + \text{rank}\, [(A - \alpha_2 E) \ldots (A - \alpha_s E)] - n \\ &\geq \text{rank}\, (A - \alpha_1 E) + \text{rank}\, (A - \alpha_2 E) + \text{rank}\, [(A - \alpha_3 E) \ldots (A - \alpha_s E)] \\ &\quad - 2n \\ &\geq \ldots \\ &\geq \text{rank}\, (A - \alpha_1 E) + \cdots + \text{rank}\, (A - \alpha_s E) - (s - 1)n \\ &= \sum_{i=1}^{s} \text{rank}\, [(A - \alpha_i E) - n] + n. \end{aligned}$$

Finally we get

$$\sum_{i=1}^{s} \text{rank } [n - (A - \alpha_i E)] \geq n. \tag{10.187}$$

As $\text{rank}[n - (A - \alpha_i E)] = \dim W_{\alpha_i}$, we have

$$\sum_{i-1}^{s} \dim W_{\alpha_i} \geq n. \tag{10.188}$$

Meanwhile, we have

$$\begin{aligned} &V^n \supset W_{\alpha_1} \oplus W_{\alpha_2} \oplus \cdots \oplus W_{\alpha_s} \\ &n \geq \dim (W_{\alpha_1} \oplus W_{\alpha_2} \oplus \cdots \oplus W_{\alpha_s}) = \sum_{i=1}^{s} \dim W_{\alpha_i}. \end{aligned} \tag{10.189}$$

The equality results from the property of a direct sum. From (10.188) and (10.189), we get

$$\sum_{i=1}^{s} \dim W_{\alpha_i} = n. \tag{10.190}$$

Hence,

$$V^n = W_{\alpha_1} \oplus W_{\alpha_2} \oplus \cdots \oplus W_{\alpha_s}. \tag{10.191}$$

Thus, we have proven that if the minimal polynomial does not have a multiple root, $V^n$ is decomposed into direct sum of eigenspaces as in (10.191).

If in turn $V^n$ is decomposed into direct sum of eigenspaces as in (10.191), $A$ can be diagonalized by a similarity transformation. The proof is as follows: Suppose that (10.191) holds. Then, we can take only eigenvectors for the basis vectors of $V^n$. Suppose that $\dim W_{\alpha_i} = n_i$. Then, we can take vectors $\boldsymbol{a}_k$ $\left[\left(\sum_{j=1}^{i-1} n_j\right) + 1 \leq k \leq \sum_{j=1}^{i} n_j\right]$ so that $\boldsymbol{a}_k$ can be the basis vectors of $W_{\alpha_i}$. In reference to this basis set, we describe a vector $\boldsymbol{x} \in V^n$ such that

$$\boldsymbol{x} = (\boldsymbol{a}_1 \boldsymbol{a}_2 \ldots \boldsymbol{a}_n) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

Operating $A$ on $\boldsymbol{x}$, we get

$$
\begin{aligned}
A(\boldsymbol{x}) &= (\boldsymbol{a}_1\ \boldsymbol{a}_2 \ldots \boldsymbol{a}_n)A\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = (\boldsymbol{a}_1 A\ \boldsymbol{a}_2 A \ldots \boldsymbol{a}_n A)\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \\
&= (\alpha_1 \boldsymbol{a}_1\ \alpha_2 \boldsymbol{a}_2 \ldots \alpha_n \boldsymbol{a}_n)\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \\
&= (\boldsymbol{a}_1\ \boldsymbol{a}_2 \ldots \boldsymbol{a}_n)\begin{pmatrix} \alpha_1 & & & \\ & \alpha_2 & & \\ & & \ddots & \\ & & & \alpha_n \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},
\end{aligned} \tag{10.192}
$$

where with the second equality we used the notation (9.40); with the third equality some of $\alpha_i$ $(1 \leq i \leq n)$ may be identical; $\boldsymbol{a}_i$ is an eigenvector that corresponds to an eigenvalue $\alpha_i$. Suppose that $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots,$ and $\boldsymbol{a}_n$ are obtained by transforming an "original" basis set $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots,$ and $\boldsymbol{e}_n$ by $R$. Then, we have

$$
A(\boldsymbol{x}) = (\boldsymbol{e}_1\ \boldsymbol{e}_2 \ldots \boldsymbol{e}_n)R\begin{pmatrix} \alpha_1 & & & \\ & \alpha_2 & & \\ & & \ddots & \\ & & & \alpha_n \end{pmatrix}R^{-1}\begin{pmatrix} x_1^{(0)} \\ x_2^{(0)} \\ \vdots \\ x_n^{(0)} \end{pmatrix}.
$$

We denote the transformation $A$ with respect to a basis set $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots,$ and $\boldsymbol{e}_n$ by $A_0$; see (9.82) with the notation. Then, we have

$$
A(\boldsymbol{x}) = (\boldsymbol{e}_1\ \boldsymbol{e}_2 \ldots \boldsymbol{e}_n)A_0\begin{pmatrix} x_1^{(0)} \\ x_2^{(0)} \\ \vdots \\ x_n^{(0)} \end{pmatrix}.
$$

Therefore, we get

$$
R^{-1}A_0R = \begin{pmatrix} \alpha_1 & & & \\ & \alpha_2 & & \\ & & \ddots & \\ & & & \alpha_n \end{pmatrix}. \tag{10.193}
$$

Thus, $A$ is similar to a diagonal matrix as represented in (10.192) and (10.193).

It is obvious to show a minimal polynomial of a diagonalizable matrix has no multiple root. The proof is left for readers.

Summarizing the above arguments, we have a following theorem:

**Theorem 10.7 [3]** *The following three statements related to A are equivalent*:

(i) *The matrix A is similar to a diagonal matrix.*
(ii) *The minimal polynomial* $\varphi_A(x)$ *does not have a multiple root.*
(iii) *The vector space* $V^n$ *is decomposed into a direct sum of eigenspaces.*

In Example 10.1 we showed the diagonalization of a matrix. There $A$ has two different eigenvalues. Since with a $(n, n)$ matrix having $n$ different eigenvalues its characteristic polynomial does not have a multiple root, the minimal polynomial necessarily has no multiple root. The above theorem therefore ensures that a matrix having no multiple root must be diagonalizable.

Another consequence of this theorem is that an idempotent matrix is diagonalizable. The matrix is characterized by $A^2 = A$. Then $A(A - E) = 0$. Taking its determinant, $(\det A)[\det(A - E)] = 0$. Therefore, we have either $\det A = 0$ or $\det(A - E) = 0$. Hence, eigenvalues of $A$ are zero or 1. Think of $f(x) = x(x - 1)$. As $f(A) = 0$, $f(x)$ should be a minimal polynomial. It has no multiple root, and so the matrix is diagonalizable.

*Example 10.6* Let us revisit Example 10.1, where we dealt with

$$A = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}. \tag{10.32}$$

From (10.33), $f_A(x) = (x - 2)(x - 1)$. Note that $f_A(x) = f_{P^{-1}AP}(x)$. Let us treat a problem according to Theorem 10.5. Also we use the notation of (10.85). Given $f_1(x) = x - 1$ and $f_2(x) = x - 2$, let us decide $M_1(x)$ and $M_2(x)$ such that these can satisfy

$$M_1(x)f_1(x) + M_2(x)f_2(x) = 1. \tag{10.194}$$

We find $M_1(x) = 1$ and $M_2(x) = -1$. Thus, using the notation of Theorem 10.5, Sect. 10.4 we have

$$A_1 = M_1(A)f_1(A) = A - E = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix},$$
$$A_2 = M_2(A)f_2(A) = -A + 2E = \begin{pmatrix} 0 & -1 \\ 0 & 1 \end{pmatrix}.$$

We also have

$$A_1 + A_2 = E,\ A_i A_j = A_j A_i = A_i \delta_{ij}. \tag{10.195}$$

Thus, we find that $A_1$ and $A_2$ are idempotent matrices. As both $A_1$ and $A_2$ are expressed by a polynomial $A$, they are commutative with $A$.

We find that $A$ is represented by

$$A = \alpha_1 A_1 + \alpha_2 A_2, \tag{10.196}$$

where $\alpha_1$ and $\alpha_2$ denote eigenvalues 2 and 1, respectively. Thus choosing proper eigenvectors for basis vectors, we have decomposed a vector space $V^n$ into a direct sum of invariant subspaces comprising the proper eigenvectors. Concomitantly, $A$ is represented as in (10.196). The relevant decomposition is always possible for a diagonalizable matrix.

Thus, idempotent matrices play an important role in the theory of linear vector spaces.

*Example 10.7* Let us think of a following matrix.

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}. \tag{10.197}$$

This is a triangle matrix, and so diagonal elements give eigenvalues. We have an eigenvalue 1 of double root and that 0 of simple root. The matrix can be diagonalized using $P$ such that

$$\widetilde{A} = P^{-1}AP = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{10.198}$$

As can be checked easily, $\widetilde{A}^2 = \widetilde{A}$. We also have

$$E - \widetilde{A} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{10.199}$$

where $(E - \widetilde{A})^2 = E - \widetilde{A}$ holds as well. Moreover, $\widetilde{A}\left(E - \widetilde{A}\right) = \left(E - \widetilde{A}\right)\widetilde{A} = 0$. Thus $\widetilde{A}$ and $E - \widetilde{A}$ behave like $A_1$ and $A_2$ of (10.195).

Next, suppose that $\boldsymbol{x} \in V^n$ is expressed as a linear combination of basis vectors $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots,$ and $\boldsymbol{a}_n$. Then

$$\boldsymbol{x} = c_1 \boldsymbol{a}_1 + c_2 \boldsymbol{a}_2 + \cdots + c_{n-1} \boldsymbol{a}_{n-1} + c_n \boldsymbol{a}_n. \tag{10.200}$$

Here let us define a following linear transformation $P^{(k)}$ such that $P^{(k)}$ "extracts" the $k$th component of $\boldsymbol{x}$. That is,

$$P^{(k)}(\boldsymbol{x}) = P^{(k)}\left(\sum_{j=1}^{n} c_j \boldsymbol{a}_j\right) = \sum_{j=1}^{n}\sum_{i=1}^{n} p_{ij}^{[k]} c_j \boldsymbol{a}_i = c_k \boldsymbol{a}_k \tag{10.201}$$

where $p_{ij}^{[k]}$ is the matrix representation of $P^{(k)}$. In fact, suppose that there is another arbitrarily chosen vector $\boldsymbol{x}$ such that

$$\boldsymbol{y} = d_1\boldsymbol{a}_1 + d_2\boldsymbol{a}_2 + \cdots + d_{n-1}\boldsymbol{a}_{n-1} + d_n\boldsymbol{a}_n. \tag{10.202}$$

Then we have

$$P^{(k)}(a\boldsymbol{x} + b\boldsymbol{y}) = (ac_k + bd_k)\boldsymbol{a}_k = ac_k\boldsymbol{a}_k + bd_k\boldsymbol{a}_k = aP^{(k)}(\boldsymbol{x}) + bP^{(k)}(\boldsymbol{y}). \tag{10.203}$$

Thus $P^{(k)}$ is a linear transformation. In (10.201), for the third equality to hold, we should have

$$(p_{ij}^{[k]}) = \delta_i^{(k)}\delta_{(k)}^{j}, \tag{10.204}$$

where $\delta_i^{(j)}$ has been defined in (10.175). Meanwhile, $\delta_{(k)}^{j}$ denotes a row vector to which only the $k$th column is 1, otherwise 0. Note that $\delta_i^{(k)}$ represents a $(n, 1)$ matrix and that $\delta_{(k)}^{j}$ denotes a $(1, n)$ matrix. Therefore, $\delta_i^{(k)}\delta_{(k)}^{j}$ represents a $(n, n)$ matrix whose $(k, k)$ element is 1, otherwise 0. Thus, $P^{(k)}(\boldsymbol{x})$ is denoted by

$$P^{(k)}(\boldsymbol{x}) = (\boldsymbol{a}_1 \ldots \boldsymbol{a}_n)\begin{pmatrix} 0 & & & & & & \\ & \ddots & & & & & \\ & & 0 & & & & \\ & & & 1 & & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_k\boldsymbol{a}_k, \tag{10.205}$$

where only the $(k,k)$ element is 1, otherwise 0. Then $P^{(k)}\left[P^{(k)}(\boldsymbol{x})\right] = P^{(k)}(\boldsymbol{x})$. That is

$$\left[P^{(k)}\right]^2 = P^{(k)}. \tag{10.206}$$

Also $P^{(k)}[P^{(l)}(\boldsymbol{x})] = 0$ if $k \neq l$. Meanwhile, we have $P^{(1)}(\boldsymbol{x}) + \cdots + P^{(n)}(\boldsymbol{x}) = \boldsymbol{x}$. Hence, $P^{(1)}(\boldsymbol{x}) + \cdots + P^{(n)}(\boldsymbol{x}) = [P^{(1)} + \cdots + P^{(n)}](\boldsymbol{x}) = \boldsymbol{x}$. Since this relation holds with any $\boldsymbol{x} \in V^n$, we get

$$P^{(1)} + \cdots + P^{(n)} = E. \tag{10.207}$$

As shown above, an idempotent matrix such as $P^{(k)}$ always exists. In particular, if the basis vectors comprise only proper eigenvectors, the decomposition as expressed in (10.196) is possible. In that case, it is described as

$$A = \alpha_1 A_1 + \cdots + \alpha_n A_n, \tag{10.208}$$

where $\alpha_1, \ldots,$ and $\alpha_n$ are eigenvalues (some of which may be identical) and $A_1, \ldots,$ and $A_n$ are idempotent matrices such as those represented by (10.205).

Yet, we have to be careful to construct idempotent matrices according to a formalism described in Theorem 10.5. It is because we often encounter a situation where different matrices give an identical characteristic polynomial. We briefly mention this in the next example.

*Example 10.8* Let us think about following two matrices:

$$A = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}, B = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}. \tag{10.209}$$

Then, following Theorem 10.5, Sect. 10.4, we have

$$f_A(x) = f_B(x) = (x-3)(x-2)^2, \tag{10.210}$$

with eigenvalues $\alpha_1 = 3$ and $\alpha_2 = 2$. Also we have $f_1(x) = (x-2)^2$ and $f_2(x) = x - 3$. Following the procedures of (10.85) and (10.86), we obtain

$$M_1(x) = x - 2 \quad \text{and} \quad M_2(x) = -x^2 + 3x - 3.$$

Therefore, we have

$$M_1(x)f_1(x) = (x-2)^3,\ M_2(x)f_2(x) = (x-3)\left(-x^2 + 3x - 3\right). \tag{10.211}$$

Hence, we get

$$\begin{aligned} A_1 &\equiv M_1(A)f_1(A) = (A - 2E)^3, \\ A_2 &\equiv M_2(A)f_2(A) = (A - 3E)\left(-A^2 + 3A - 3E\right). \end{aligned} \tag{10.212}$$

Similarly, we get $B_1$ and $B_2$ by replacing $A$ with $B$ in (10.212). Thus, we have

$$A_1 = B_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},\ A_2 = B_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{10.213}$$

Notice that we get the same idempotent matrix of (10.213), even though the matrix forms of $A$ and $B$ differ. Also we have

$$A_1 + A_2 = B_1 + B_2 = E.$$

Then, we have

$$A = (A_1 + A_2)A = A_1A + A_2A,\ B = (B_1 + B_2)B = B_1B + B_2B.$$

Nonetheless, although $A = 3A_1 + 2A_2$ holds, $B \neq 3B_1 + 2B_2$. That is, the decomposition of the form of (10.208) is not true of $B$. The decomposition of this kind is possible only with diagonalizable matrices.

In summary, a $(n, n)$ matrix with $s$ $(1 \leq s \leq n)$ different eigenvalues has at least $s$ proper eigenvectors. (Note that a diagonalizable matrix has $n$ proper eigenvectors.) In the case of $s < n$, the matrix has multiple root(s) and may have generalized eigenvectors. If the matrix has a generalized eigenvector of rank $v$, the matrix is accompanied by $(v - 1)$ generalized eigenvectors along with a sole proper eigenvector. Those vectors form an invariant subspace along with the proper eigenvector (s). In total, such $n$ (generalized) eigenvectors span a whole vector space $V^n$.

With the eigenvalue equation $A(\boldsymbol{x}) = \alpha\boldsymbol{x}$, we have an indefinite but non-trivial solution $\boldsymbol{x} \neq \boldsymbol{0}$ for only restricted numbers $\alpha$ (i.e., eigenvalues) in a complex plane. However, we have a unique but trivial solution $\boldsymbol{x} = \boldsymbol{0}$ for complex numbers $\alpha$ other than eigenvalues. This is characteristic of the eigenvalue problem.

## References

1. Mirsky L (1990) An Introduction to linear algebra. Dover, New York
2. Hassani S (2006) Mathematical physics. Springer, New York
3. Satake I (1974) Linear algebra. Shokabo, Tokyo (in Japanese)

# Chapter 11
# Inner Product Space

Thus far we have treated the theory of linear vector spaces. The vector spaces, however, were somewhat "structureless," and so it will be desirable to introduce a concept of metric or measure into the linear vector spaces. We call a linear vector space where the inner product is defined an inner product space. In virtue of a concept of the inner product, the linear vector space is given a variety of structures. For instance, introduction of the inner product to the linear vector space immediately leads to the definition of adjoint operators and Gram matrices.

Above all, the concept of inner product can readily be extended to a functional space and facilitate understanding of, e.g., orthogonalization of functions, as was exemplified in Parts I and II. Moreover, definition of the inner product allows us to relate matrix operators and differential operators. In particular, it is a key issue to understand logical structure of quantum mechanics. This can easily be understood from the fact that Paul Dirac, who was known as one of prominent founders of quantum mechanics, invented bra and ket vectors to represent an inner product.

## 11.1 Inner Product and Metric

Inner product relates two vectors to a complex number. To do this, we introduce the notation $|a\rangle$ and $\langle b|$ to represent the vectors. This notation is due to Dirac and widely used in physics and mathematical physics. Usually, $|a\rangle$ and $\langle b|$ are called a "ket" vector and a "bra" vector, respectively, again due to Dirac. Alternatively, we may call $\langle a|$ an adjoint vector of $|a\rangle$. Or we denote $\langle a| \equiv |a\rangle^{\dagger}$. The symbol "†" (dagger) means that for a matrix, its transposed matrix should be taken with complex conjugate matrix elements. That is, $\left(a_{ij}\right)^{\dagger} = (a_{ji}^{*})$. If we represent a full matrix, we have



S. Hotta, *Mathematical Physical Chemistry*,
https://doi.org/10.1007/978-981-10-7671-8_11

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}, \quad A^\dagger = \begin{pmatrix} a_{11}^* & \cdots & a_{n1}^* \\ \vdots & \ddots & \vdots \\ a_{1n}^* & \cdots & a_{nn}^* \end{pmatrix} \tag{11.1}$$

We call $A^\dagger$ an adjoint matrix or adjoint to $A$; see (1.106). The operation of transposition and complex conjugate is commutable. A further remark will be made below after showing the definition of the inner product. The symbols $|a\rangle$ and $\langle b|$ represent vectors, and hence, we do not need to use bold characters to show that those are vectors.

The definition of the inner product is as follows:

$$\langle b|a\rangle = \langle a|b\rangle^*, \tag{11.2}$$

$$\langle a|(\beta|b\rangle + \gamma|c\rangle) = \beta\langle a|b\rangle + \gamma\langle a|c\rangle, \tag{11.3}$$

$$\langle a|a\rangle \geq 0. \tag{11.4}$$

In (11.4), equality holds only if $|a\rangle = 0$. Note here that two vectors are said to be orthogonal to each other if their inner product vanishes, i.e., $\langle b|a\rangle = \langle a|b\rangle = 0$. In particular, if a vector $|a\rangle \in V^n$ is orthogonal to all the vectors in $V^n$, i.e., $\langle x|a\rangle = 0$ for $^\forall x \in V^n$, then $|a\rangle = \mathbf{0}$. This is because if we choose $|a\rangle$ for $|x\rangle$, we have $\langle a|a\rangle = 0$. This means that $|a\rangle = \mathbf{0}$. We call a linear vector space to which the inner product is defined an inner product space.

We can create another structure to a vector space. An example is a metric. Suppose that there is an arbitrary set $Q$. If a real nonnegative number $\rho(a, b)$ is defined as follows with $a, b \in Q$, the set $Q$ is called a metric space [1]:

$$\rho(a, b) = \rho(b, a), \tag{11.5}$$

$$\rho(a, b) \geq 0 \text{ for } ^\forall a, b; \ \rho(a, b) = 0 \quad \text{if and only if } a = b, \tag{11.6}$$

$$\rho(a, b) + \rho(b, c) \geq \rho(a, c). \tag{11.7}$$

In our study, a vector space is chosen for the set $Q$. Here let us define a norm for each vector $a$. The norm is defined as

$$\|a\| = \sqrt{\langle a|a\rangle}. \tag{11.8}$$

If we define $\rho(a, b) \equiv \|a - b\|$, $\|a - b\|$ satisfies the definitions of metric. Equations (11.5) and (11.6) are obvious. For (11.7), let us consider a vector $|c\rangle$ as $|c\rangle = |a\rangle - x\langle b|a\rangle|b\rangle$ with real $x$. Since $\langle c|c\rangle \geq 0$, we have

$$\begin{aligned} &x^2\langle a|b\rangle\langle b|a\rangle\langle b|b\rangle - 2x\langle a|b\rangle\langle b|a\rangle + \langle a|a\rangle \geq 0 \quad \text{or} \\ &x^2|\langle a|b\rangle|^2\langle b|b\rangle - 2x|\langle a|b\rangle|^2 + \langle a|a\rangle \geq 0. \end{aligned} \tag{11.9}$$

The inequality (11.9) related to the quadratic equation in $x$ with real coefficients requires the inequality such that

$$\langle a|a\rangle\langle b|b\rangle \geq \langle a|b\rangle\langle b|a\rangle = |\langle a|b\rangle|^2. \tag{11.10}$$

That is,

$$\sqrt{\langle a|a\rangle} \cdot \sqrt{\langle b|b\rangle} \geq |\langle a|b\rangle|. \tag{11.11}$$

Namely,

$$\|a\| \cdot \|b\| \geq |\langle a|b\rangle| \geq \mathcal{R}e\langle a|b\rangle. \tag{11.12}$$

The relations (11.11) and (11.12) are known as Cauchy–Schwarz inequality. Meanwhile, we have

$$\|a+b\|^2 = \langle a+b|a+b\rangle = \|a\|^2 + \|b\|^2 + 2\mathcal{R}e\langle a|b\rangle, \tag{11.13}$$

$$(\|a\| + \|b\|)^2 = \|a\|^2 + \|b\|^2 + 2\|a\| \cdot \|b\|, \tag{11.14}$$

Comparing (11.13) and (11.14) and using (11.12), we have

$$\|a\| + \|b\| \geq \|a+b\|. \tag{11.15}$$

The inequality (11.15) is known as the triangle inequality. In (11.15), replacing $a \to a-b$ and $b \to b-c$, we get

$$\|a-b\| + \|b-c\| \geq \|a-c\|. \tag{11.16}$$

Thus, (11.16) is equivalent to (11.7). At the same time, the norm defined in (11.8) may be regarded as a "length" of a vector $a$.

As $\beta|b\rangle + \gamma|c\rangle$ represents a vector, we use a shorthand notation for it as

$$|\beta b + \gamma c\rangle \equiv \beta|b\rangle + \gamma|c\rangle. \tag{11.17}$$

According to the definition (11.3)

$$\langle a|\beta b + \gamma c\rangle = \beta\langle a|b\rangle + \gamma\langle a|c\rangle. \tag{11.18}$$

Also from (11.2),

$$\langle \beta b + \gamma c|a\rangle = \langle a|\beta b + \gamma c\rangle^* = [\beta\langle a|b\rangle + \gamma\langle a|c\rangle]^* = \beta^*\langle b|a\rangle + \gamma^*\langle c|a\rangle. \tag{11.19}$$

That is,

$$\langle \beta b + \gamma c| = \beta^*\langle b| + \gamma^*\langle c|. \tag{11.20}$$

Therefore, when we take out a scalar from a bra vector, it should be a complex conjugate. When the scalar is taken out from a ket vector, on the other hand it is unaffected by definition (11.3). To show this, we have $|\alpha a\rangle = \alpha|a\rangle$. Taking its adjoint, $\langle \alpha a| = \alpha^*\langle a|$.

We can view (11.17) as a linear transformation in a vector space. In other words, if we regard $|\cdot\rangle$ as a linear transformation of a vector $a \in V^n$ to $|a\rangle \in \widetilde{V^n}$, $|\cdot\rangle$ is that of $V^n$ to $\widetilde{V^n}$. On the other hand, $\langle\cdot|$ could not be regarded as a linear transformation of $a \in V^n$ to $\langle a| \in \widetilde{V^{n'}}$. Sometimes, the said transformation is referred to as "antilinear" or "sesquilinear." From the point of view of formalism, the inner product can be viewed as an operation: $\widetilde{V^n} \times \widetilde{V^{n'}} \to \mathbb{C}$.

Let us consider $|x\rangle = |y\rangle$ in an inner product space $\widetilde{V^n}$. Then $|x\rangle - |y\rangle = 0$. That is, $|x - y\rangle = 0$. Therefore, we have $x = y$, or $|0\rangle = 0$. This means that the linear transformation $|\cdot\rangle$ converts $0 \in V^n$ to $|0\rangle \in \widetilde{V^n}$. This is a characteristic of a linear transformation represented in (9.44). Similarly, we have $\langle 0| = 0$. However, we do not have to get into further details in this book. Also, if we have to specify a vector space, we simply do so by designating it as $V^n$.

## 11.2 Gram Matrices

Once we have defined an inner product between any pair of vectors $|a\rangle$ and $|b\rangle$ of a vector space $V^n$, we can define and calculate various quantities related to inner products. As an example, let $|a_1\rangle, \ldots, |a_n\rangle$ and $|b_1\rangle, \ldots, |b_n\rangle$ be two sets of vectors in $V^n$. The vectors $|a_1\rangle, \ldots, |a_n\rangle$ may or may not be linearly independent. This is true of $|b_1\rangle, \ldots, |b_n\rangle$. Let us think of a following matrix $M$ defined as below:

$$M = \begin{pmatrix} \langle a_1| \\ \vdots \\ \langle a_n| \end{pmatrix} (|b_1\rangle \cdots |b_n\rangle) = \begin{pmatrix} \langle a_1|b_1\rangle & \cdots & \langle a_1|b_n\rangle \\ \vdots & \ddots & \vdots \\ \langle a_n|b_1\rangle & \cdots & \langle a_n|b_n\rangle \end{pmatrix}. \tag{11.21}$$

We assume the following cases.

(i) Suppose that in (11.21), $|b_1\rangle, \ldots, |b_n\rangle$ are linearly dependent. Then, without loss of generality, we can put

$$|b_1\rangle = c_2|b_2\rangle + c_3|b_3\rangle + \cdots + c_n|b_n\rangle. \tag{11.22}$$

Then we have

$$M = \begin{pmatrix} c_2\langle a_1|b_2\rangle + c_3\langle a_1|b_3\rangle + \cdots + c_n\langle a_1|b_n\rangle & \langle a_1|b_2\rangle & \ldots & \langle a_1|b_n\rangle \\ \vdots & \vdots & \ddots & \vdots \\ c_2\langle a_n|b_2\rangle + c_3\langle a_n|b_3\rangle + \cdots + c_n\langle a_n|b_n\rangle & \langle a_n|b_2\rangle & \ldots & \langle a_n|b_n\rangle \end{pmatrix}. \tag{11.23}$$

Multiplying the second column, ..., and the $n$th column by $(-c_2)$, ..., and $(-c_n)$, respectively, and adding them to the first column to get

$$M = \begin{pmatrix} 0 & \langle a_1|b_2\rangle & \ldots & \langle a_1|b_n\rangle \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \langle a_n|b_2\rangle & \ldots & \langle a_n|b_n\rangle \end{pmatrix}. \tag{11.24}$$

Hence, $\det M = 0$.

(ii) Suppose in turn that in (11.21), $|a_1\rangle, \ldots, |a_n\rangle$ are linearly dependent. In that case, again, without loss of generality, we can put

$$|a_1\rangle = d_2|a_2\rangle + d_3|a_3\rangle + \cdots + d_n|a_n\rangle. \tag{11.25}$$

Focusing attention on individual rows and taking a similar procedure described above, we have

$$M = \begin{pmatrix} 0 & \cdots & 0 \\ \langle a_2|b_1\rangle & \cdots & \langle a_2|b_n\rangle \\ \vdots & \ddots & \vdots \\ \langle a_n|b_1\rangle & \cdots & \langle a_n|b_n\rangle \end{pmatrix} \tag{11.26}$$

Again, $\det M = 0$.

Next, let us examine the case where $\det M = 0$. In that case, $n$ column vectors of $M$ in (11.21) are linearly dependent. Without loss of generality, we suppose that the first column is expressed as a linear combination of the other $(n-1)$ columns such that

$$\begin{pmatrix} \langle a_1|b_1\rangle \\ \vdots \\ \langle a_n|b_1\rangle \end{pmatrix} = c_2 \begin{pmatrix} \langle a_1|b_2\rangle \\ \vdots \\ \langle a_n|b_2\rangle \end{pmatrix} + \cdots + c_n \begin{pmatrix} \langle a_1|b_n\rangle \\ \vdots \\ \langle a_n|b_n\rangle \end{pmatrix}. \tag{11.27}$$

Rewriting this, we have

$$\begin{pmatrix} \langle a_1|b_1 - c_2b_2 - \cdots - c_nb_n\rangle \\ \vdots \\ \langle a_n|b_1 - c_2b_2 - \cdots - c_nb_n\rangle \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{11.28}$$

Multiplying the first row, ..., and the $n$th row of (11.28) by an appropriate complex number $p_1^*$, ..., and $p_n^*$, respectively, we have

$$\begin{aligned} &\langle p_1a_n|b_1 - c_2b_2 - \cdots - c_nb_n\rangle = 0, \\ &\qquad\cdots, \\ &\langle p_na_n|b_1 - c_2b_2 - \cdots - c_nb_n\rangle = 0. \end{aligned} \tag{11.29}$$

Adding all the above, we get

$$\langle p_1a_1 + \cdots + p_na_n|b_1 - c_2b_2 - \cdots - c_nb_n\rangle = 0. \tag{11.30}$$

Now, suppose that $|a_1\rangle, \ldots, |a_n\rangle$ are the basis vectors. Then, $|p_1a_1 + \cdots + p_na_n\rangle$ represent any vectors in a vector space. This implies that $|b_1 - c_2b_2 - \cdots - c_nb_n\rangle = 0$; for this, see remarks after (11.4). That is,

$$|b_1\rangle = c_2|b_2\rangle + \cdots + c_n|b_n\rangle. \tag{11.31}$$

Thus, $|b_1\rangle, \ldots, |b_n\rangle$ are linearly dependent.

Meanwhile, $\det M = 0$ implies that $n$ row vectors of $M$ in (11.21) are linearly dependent. In that case, performing similar calculations to the above, we can readily show that if $|b_1\rangle, \ldots, |b_n\rangle$ are the basis vectors, $|a_1\rangle, \ldots, |a_n\rangle$ are linearly dependent.

We summarize the above discussion by following statement: Suppose that we have two sets of vectors $|a_1\rangle, \ldots, |a_n\rangle$ and $|b_1\rangle, \ldots, |b_n\rangle$.

At least a set of vectors are linearly dependent. $\Leftrightarrow \det M = 0$.

Both the sets of vectors are linearly independent. $\Leftrightarrow \det M \neq 0$.

The latter statement is obtained by considering contraposition of the former statement.

We restate the above in following theorem:

**Theorem 11.1** *Let $|a_1\rangle, \ldots, |a_n\rangle$ and $|b_1\rangle, \ldots, |b_n\rangle$ be two sets of vectors defined in a vector space $V^n$. A necessary and sufficient condition for both these sets of vectors to be linearly independent is that for a matrix $M$ defined below,* $\det M \neq 0$.

$$M = \begin{pmatrix} \langle a_1| \\ \vdots \\ \langle a_n| \end{pmatrix} (|b_1\rangle \cdots |b_n\rangle) = \begin{pmatrix} \langle a_1|b_1\rangle & \cdots & \langle a_1|b_n\rangle \\ \vdots & \ddots & \vdots \\ \langle a_n|b_1\rangle & \cdots & \langle a_n|b_n\rangle \end{pmatrix}.$$

Next, we consider a norm of a vector expressed in reference to a set of basis vectors $|e_1\rangle, \ldots, |e_n\rangle$ of $V^n$. Let us express a vector $|x\rangle$ in an inner product space as follows as in the case of (9.10) and (9.13):

$$\begin{aligned} |x\rangle &= x_1|e_1\rangle + x_2|e_2\rangle + \cdots + x_n|e_n\rangle \\ &= |x_1e_1 + x_2e_2 + \cdots + x_ne_n\rangle \\ &= (|e_1\rangle \cdots |e_n\rangle)\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}. \end{aligned} \tag{11.32}$$

A bra vector $\langle x|$ is then denoted by

$$\langle x| = \left(x_1^* \cdots x_n^*\right)\begin{pmatrix} \langle e_1| \\ \vdots \\ \langle e_n| \end{pmatrix}. \tag{11.33}$$

Thus, we have an inner product described as

$$\begin{aligned} \langle x|x\rangle &= \left(x_1^* \cdots x_n^*\right)\begin{pmatrix} \langle e_1| \\ \vdots \\ \langle e_n| \end{pmatrix}(|e_1\rangle \cdots |e_n\rangle)\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \\ &= \left(x_1^* \cdots x_n^*\right)\begin{pmatrix} \langle e_1|e_1\rangle & \cdots & \langle e_1|e_n\rangle \\ \vdots & \ddots & \vdots \\ \langle e_n|e_1\rangle & \cdots & \langle e_n|e_n\rangle \end{pmatrix}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}. \end{aligned} \tag{11.34}$$

Here the matrix expressed as follows is called a Gram matrix [2, 3]:

$$G = \begin{pmatrix} \langle e_1|e_1\rangle & \cdots & \langle e_1|e_n\rangle \\ \vdots & \ddots & \vdots \\ \langle e_n|e_1\rangle & \cdots & \langle e_n|e_n\rangle \end{pmatrix}. \tag{11.35}$$

As $\langle e_j|e_i\rangle = \langle e_i|e_j\rangle^*$, we have $G = G^\dagger$. From Theorem 11.1, $\det G \neq 0$. With a shorthand notation, we write $(G)_{ij} = (g_{ij}) = \left(\langle e_i|e_j\rangle\right)$. As already mentioned in Sect. 1.1, if for a matrix $H$ we have a relation described by

$$H = H^\dagger, \tag{1.119}$$

it is said to be a Hermitian matrix or a self-adjoint matrix. We often say that such a matrix is Hermitian.

Since the Gram matrices frequently appear in matrix algebra and play a role, their properties are worth examining. Since $G$ is a Hermitian matrix, it can be diagonalized through similarity transformation using a unitary matrix. We will give its proof later (see Sect. 12.3).

Let us deal with (11.34) further. We have

$$\langle x|x\rangle = \left(x_1^* \cdots x_n^*\right) UU^\dagger \begin{pmatrix} \langle e_1|e_1\rangle & \cdots & \langle e_1|e_n\rangle \\ \vdots & \ddots & \vdots \\ \langle e_n|e_1\rangle & \cdots & \langle e_n|e_n\rangle \end{pmatrix} UU^\dagger \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \tag{11.36}$$

where $U$ is defined as $UU^\dagger = U^\dagger U = E$. Such a matrix $U$ is called a unitary matrix. We represent a matrix form of $U$ as

$$U = \begin{pmatrix} u_{11} & \cdots & u_{1n} \\ \vdots & \ddots & \vdots \\ u_{n1} & \cdots & u_{nn} \end{pmatrix}, U^\dagger = \begin{pmatrix} u_{11}^* & \cdots & u_{n1}^* \\ \vdots & \ddots & \vdots \\ u_{1n}^* & \cdots & u_{nn}^* \end{pmatrix}. \tag{11.37}$$

Here, putting

$$U^\dagger \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \text{ or equivalently } \sum_{k=1}^{n} x_k u_{ki}^* \equiv \xi_i \tag{11.38}$$

and taking its adjoint such that

$$\left(x_1^* \cdots x_n^*\right) U = \left(\xi_1^* \cdots \xi_n^*\right) \text{ or equivalently } \sum_{k=1}^{n} x_k^* u_{ki} = \xi_i^*,$$

we have

$$\langle x|x\rangle = \left(\xi_1^* \cdots \xi_n^*\right) U^\dagger \begin{pmatrix} \langle e_1|e_1\rangle & \cdots & \langle e_1|e_n\rangle \\ \vdots & \ddots & \vdots \\ \langle e_n|e_1\rangle & \cdots & \langle e_n|e_n\rangle \end{pmatrix} U \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}. \tag{11.39}$$

We assume that the Gram matrix is diagonalized by a similarity transformation by $U$. After being diagonalized, similarly to (10.192) and (10.193) the Gram matrix has a following form $G'$:

$$U^\dagger G U = G' = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}. \tag{11.40}$$

Thus, we get

$$\langle x|x\rangle = \left(\xi_1^* \cdots \xi_n^*\right) \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix} \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} = \lambda_1|\xi_1|^2 + \cdots + \lambda_n|\xi_n|^2. \quad (11.41)$$

From the relation (11.4), $\langle x|x\rangle \geq 0$. This implies that in (11.41), we have

$$\lambda_i \geq 0 \ (1 \leq i \leq n). \quad (11.42)$$

To show this, suppose that for $^\exists\lambda_i$, $\lambda_i < 0$. Suppose also that for $^\exists\xi_i$, $\xi_i \neq 0$. Then, with $\begin{pmatrix} 0 \\ \vdots \\ \xi_i \\ \vdots \\ 0 \end{pmatrix}$ we have $\langle x|x\rangle = \lambda_i|\xi_i|^2 < 0$, in contradiction.

Since we have $\det G \neq 0$, taking a determinant of (11.40) we have

$$\det G' = \det\left(U^\dagger G U\right) = \det\left(U^\dagger U G\right) = \det E \det G = \det G = \prod_{i=1}^{n} \lambda_i \neq 0. \quad (11.43)$$

Combining (11.42) and (11.43), all the eigenvalues $\lambda_i$ are positive, i.e.,

$$\lambda_i > 0 \ (1 \leq i \leq n). \quad (11.44)$$

The norm $\langle x|x\rangle = 0$, if and only if $\xi_1 = \xi_2 = \cdots = \xi_n = 0$ which corresponds to $x_1 = x_2 = \cdots = x_n = 0$ from (11.38).

For further study, we generalize the aforementioned feature a little further. That is, if $|e_1\rangle, \ldots,$ and $|e_n\rangle$ are linearly dependent, from Theorem 11.1 we get $\det G = 0$. This implies that there is at least one eigenvalue $\lambda_i = 0$ $(1 \leq i \leq n)$ and that for a vector $\begin{pmatrix} 0 \\ \vdots \\ \xi_i \\ \vdots \\ 0 \end{pmatrix}$ with $^\exists\xi_i \neq 0$ we have $\langle x|x\rangle = 0$. (Suppose that in $V^2$ we have linearly dependent vectors $|e_1\rangle$ and $|e_2\rangle = |-e_1\rangle$; see Example 11.2 below.)

Let $H$ be a Hermitian matrix [i.e., $(H)_{ij} = (H^*)_{ji}$]. If we have $\phi$ as a function of complex variables $x_1, \ldots, x_n$ such that

$$\phi(x_1, \ldots, x_n) = \sum_{i,j=1}^{n} x_i^* (H)_{ij} x_j, \quad (11.45)$$

where $H_{ij}$ is a matrix element of a Hermitian matrix; $\phi(x_1, \ldots, x_n)$ is said to be a Hermitian quadratic form. Suppose that $\phi(x_1, \ldots, x_n)$ satisfies $\phi(x_1, \ldots, x_n) = 0$ if and only if $x_1 = x_2 = \cdots = x_n = 0$, and otherwise $\phi(x_1, \ldots, x_n) > 0$ with any other sets of $x_i$ $(1 \leq i \leq n)$. Then, the said Hermitian quadratic form is called positive definite and we write

$$H > 0. \tag{11.46}$$

If $\phi(x_1, \ldots, x_n) \geq 0$ for any $x_n$ $(1 \leq i \leq n)$ and $\phi(x_1, \ldots, x_n) = 0$ for at least a set of $(x_1, \ldots, x_n)$ to which $^{\exists}x_i \neq 0$, $\phi(x_1, \ldots, x_n)$ is said to be positive semi-definite or nonnegative. In that case, we write

$$H \geq 0. \tag{11.47}$$

From the above argument, a Gram matrix comprising linearly independent vectors is positive definite, whereas that comprising linearly dependent vectors is nonnegative. On the basis of the above argument including (11.36) to (11.44), we have

$$\begin{aligned} &H > 0 \Leftrightarrow \ \lambda_i > 0 \ (1 \leq i \leq n), \det H > 0; \\ &H \geq 0 \Leftrightarrow \ \lambda_i \geq 0 \text{ with } {}^{\exists}\lambda_i = 0 \ (1 \leq i \leq n), \det H = 0. \end{aligned} \tag{11.48}$$

Notice here that eigenvalues $\lambda_i$ remain unchanged after (unitary) similarity transformation. Namely, the eigenvalues are inherent to $H$.

We have already encountered several examples of positive definite and nonnegative operators. A typical example of the former case is Hamiltonian of quantum-mechanical harmonic oscillator (see Chap. 2). In this case, energy eigenvalues are all positive (i.e., positive definite). Orbital angular momenta $\boldsymbol{L}^2$ of hydrogen-like atoms, on the other hand, are nonnegative operators, and hence, an eigenvalue of zero is permitted.

Alternatively, the Gram matrix is defined as $B^{\dagger}B$, where $B$ is any $(n, n)$ matrix. If we take an orthonormal basis $|\eta_1\rangle, |\eta_2\rangle, \ldots, |\eta_n\rangle$, $|e_i\rangle$ can be expressed as

$$\begin{aligned} |e_i\rangle &= \sum_{j=1}^{n} b_{ji} |\eta_j\rangle, \\ \langle e_k | e_i \rangle &= \sum_{j=1}^{n} \sum_{l=1}^{n} b_{lk}^* b_{ji} \langle \eta_l | \eta_j \rangle = \sum_{j=1}^{n} \sum_{l=1}^{n} b_{lk}^* b_{ji} \delta_{lj} = \sum_{j=1}^{n} b_{jk}^* b_{ji} \\ &= \sum_{j=1}^{n} \left( B^{\dagger} \right)_{kj} (B)_{ji} = \left( B^{\dagger} B \right)_{ki}. \end{aligned} \tag{11.49}$$

For the second equality of (11.49), we used the orthonormal condition $\langle \eta_i \mid \eta_j \rangle = \delta_{ij}$. Thus, the Gram matrix $G$ defined in (11.35) can be regarded as identical to $B^{\dagger}B$.

In Sect. 9.4, we have dealt with a linear transformation of a set of basis vectors $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots,$ and $\boldsymbol{e}_n$ by a matrix $A$ defined in (9.69) and examined whether the transformed vectors $\boldsymbol{e}_1^{'}, \boldsymbol{e}_2^{'},$ and $\boldsymbol{e}_n^{'}$ are linearly independent. As a result, a necessary and sufficient condition for $\boldsymbol{e}_1^{'}, \boldsymbol{e}_2^{'},$ and $\boldsymbol{e}_n^{'}$ to be linearly independent (i.e., to be a set of basis vectors) is $\det A \neq 0$. Thus, we notice that $B$ plays a same role as $A$ of (9.69) and, hence, $\det B \neq 0$ if and only if the set of vectors $|e_1\rangle, |e_2\rangle,$ and $\cdots |e_n\rangle$ defined in (11.32) are linearly independent. By the same token as the above, we conclude that eigenvalues of $B^\dagger B$ are all positive, only if $\det B \neq 0$ (i.e., $B$ is non-singular).

Alternatively, if $B$ is singular, $\det B^\dagger B = \det B^\dagger \det B = 0$. In that case, at least one of eigenvalues of $B^\dagger B$ must be zero. The Gram matrices appearing in (11.35) are frequently dealt with in the field of mathematical physics in conjunction with quadratic forms. Further topics can be seen in the next chapter.

*Example 11.1* Let us take two vectors $|\varepsilon_1\rangle$ and $|\varepsilon_2\rangle$ that are expressed as

$$|\varepsilon_1\rangle = |e_1\rangle + |e_2\rangle, \; |\varepsilon_2\rangle = |e_1\rangle + i|e_2\rangle. \tag{11.50}$$

Here we have $\langle e_i | e_j \rangle = \delta_{ij}$ $(1 \leq i, j \leq 2)$. Then we have a Gram matrix expressed as

$$G = \begin{pmatrix} \langle \varepsilon_1 | \varepsilon_1 \rangle & \langle \varepsilon_1 | \varepsilon_2 \rangle \\ \langle \varepsilon_2 | \varepsilon_1 \rangle & \langle \varepsilon_2 | \varepsilon_2 \rangle \end{pmatrix} = \begin{pmatrix} 2 & 1+i \\ 1-i & 2 \end{pmatrix}. \tag{11.51}$$

Principal minors of $G$ are $|2| = 2 > 0$ and $\begin{vmatrix} 2 & 1+i \\ 1-i & 2 \end{vmatrix} = 4 - (1+i)(1-i) = 2 > 0$. Therefore, according to Theorem 12.11, $G > 0$ (vide infra).

Let us diagonalize the matrix $G$. To this end, we find roots of the characteristic equation. That is,

$$\det |G - \lambda E| = \begin{vmatrix} 2-\lambda & 1+i \\ 1-i & 2-\lambda \end{vmatrix} = 0, \; \lambda^2 - 4\lambda + 2 = 0. \tag{11.52}$$

We have $\lambda = 2 \pm \sqrt{2}$. Then as a diagonalizing unitary matrix $U$, we get

$$U = \begin{pmatrix} \frac{1+i}{2} & \frac{1+i}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{pmatrix}. \tag{11.53}$$

Thus, we get

$$U^\dagger GU = \begin{pmatrix} \frac{1-i}{2} & \frac{\sqrt{2}}{2} \\ \frac{1-i}{2} & -\frac{\sqrt{2}}{2} \end{pmatrix} \begin{pmatrix} 2 & 1+i \\ 1-i & 2 \end{pmatrix} \begin{pmatrix} \frac{1+i}{2} & \frac{1+i}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{pmatrix} = \begin{pmatrix} 2+\sqrt{2} & 0 \\ 0 & 2-\sqrt{2} \end{pmatrix}. \tag{11.54}$$

The eigenvalues $2+\sqrt{2}$ and $2-\sqrt{2}$ are real positive as expected. That is, the Gram matrix is positive definite.

*Example 11.2* Let $|e_1\rangle$ and $|e_2\rangle (= -|e_1\rangle)$ be two vectors. Then, we have a Gram matrix expressed as

$$G = \begin{pmatrix} \langle e_1|e_1\rangle & \langle e_1|e_2\rangle \\ \langle e_2|e_1\rangle & \langle e_2|e_2\rangle \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}. \tag{11.55}$$

Similarly in the case of Example 11.1, we have as an eigenvalue equation

$$\det|G - \lambda E| = \begin{vmatrix} 1-\lambda & -1 \\ -1 & 1-\lambda \end{vmatrix} = 0,\ \lambda^2 - 2\lambda = 0. \tag{11.56}$$

We have $\lambda = 2$ or 0. As a diagonalizing unitary matrix $U$, we get

$$U = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}. \tag{11.57}$$

Thus, we get

$$U^\dagger GU = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}. \tag{11.58}$$

As expected, we have a diagonal matrix, one of eigenvalues for which is zero. That is, the Gram matrix is nonnegative.

In the present case, let us think of a following Hermitian quadratic form.

$$\langle x|x\rangle = \sum_{i,j=1}^{2} x_i^*(G)_{ij}x_j = \left(x_1^* x_2^*\right) UU^\dagger GUU^\dagger \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$
$$= \left(x_1^* x_2^*\right) U \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} U^\dagger \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \left(\xi_1^* \xi_2^*\right) \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = 2|\xi_1|^2,$$

Then, if we take $\begin{pmatrix} 0 \\ \xi_2 \end{pmatrix}$ with $\xi_2 \neq 0$ as a column vector, $\langle x|x\rangle = 0$. Meanwhile, we have

$$U^\dagger \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}.$$

Thus, for $\begin{pmatrix} 0 \\ \xi_2 \end{pmatrix}$ with $\xi_2 \neq 0$ we get $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. That is, if we had $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, we would have $\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, in contradiction. To be more precise, if we take, e.g., $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ or $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$, we have

$$\langle x|x\rangle = \begin{pmatrix} x_1^* & x_2^* \end{pmatrix} G \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0,$$

$$\langle x|x\rangle = \begin{pmatrix} 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \end{pmatrix} = 1 > 0.$$

Thus, $G$ is nonnegative by definition.

## 11.3 Adjoint Operators

A linear transformation $A$ is similarly defined as before and $A$ transforms $|x\rangle$ of (11.32) such that

$$A(|x\rangle) = (|e_1\rangle \cdots |e_n\rangle) \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \tag{11.59}$$

where $\left(a_{ij}\right)$ is a matrix representation of $A$. Defining $A(|x\rangle) = |A(x)\rangle$ and $\langle (x)A^\dagger| = (\langle x|)A^\dagger = [A(|x\rangle)]^\dagger$ to be consistent with (1.117), we have

$$\langle (x)A^\dagger| = \left(x_1^* \cdots x_n^*\right) \begin{pmatrix} a_{11}^* & \cdots & a_{n1}^* \\ \vdots & \ddots & \vdots \\ a_{1n}^* & \cdots & a_{nn}^* \end{pmatrix} \begin{pmatrix} \langle e_1| \\ \vdots \\ \langle e_n| \end{pmatrix}. \tag{11.60}$$

Therefore, putting $|y\rangle = \sum_{i=1}^{n} y_i |e_i\rangle$, we have

$$\left\langle (x)A^{\dagger} \,\middle|\, y \right\rangle = \left(x_1^* \cdots x_n^*\right) \begin{pmatrix} a_{11}^* & \cdots & a_{n1}^* \\ \vdots & \ddots & \vdots \\ a_{1n}^* & \cdots & a_{nn}^* \end{pmatrix} \begin{pmatrix} \langle e_1| \\ \vdots \\ \langle e_n| \end{pmatrix} \left(|e_1\rangle \cdots |e_n\rangle\right) \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$
$$= \left(x_1^* \cdots x_n^*\right) A^{\dagger} G \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}. \tag{11.61}$$

Meanwhile, we get

$$\langle y \mid A(x)\rangle = \left(y_1^* \cdots y_n^*\right) \begin{pmatrix} \langle e_1| \\ \vdots \\ \langle e_1| \end{pmatrix} \left(|e_1\rangle \cdots |e_1\rangle\right) \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$
$$= \left(y_1^* \cdots y_n^*\right) GA \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}. \tag{11.62}$$

Hence, we have

$$\langle y \mid A(x)\rangle^* = (y_1 \cdots y_n) G^* A^* \begin{pmatrix} x_1^* \\ \vdots \\ x_n^* \end{pmatrix} = (y_1 \cdots y_n) G^{\mathrm{T}} \left(A^{\dagger}\right)^{\mathrm{T}} \begin{pmatrix} x_1^* \\ \vdots \\ x_n^* \end{pmatrix}. \tag{11.63}$$

With the second equality, we used $G^* = \left(G^{\dagger}\right)^{\mathrm{T}} = G^{\mathrm{T}}$ (note that $G$ is Hermitian) and $A^* = \left(A^{\dagger}\right)^{\mathrm{T}}$. A complex conjugate matrix $A^*$ is defined as

$$A^* \equiv \begin{pmatrix} a_{11}^* & \cdots & a_{1n}^* \\ \vdots & \ddots & \vdots \\ a_{n1}^* & \cdots & a_{nn}^* \end{pmatrix}.$$

Comparing (11.61) and (11.63), we find that one is transposed matrix of the other. Also note that $(AB)^{\mathrm{T}} = B^{\mathrm{T}}A^{\mathrm{T}}$, $(ABC)^{\mathrm{T}} = C^{\mathrm{T}}B^{\mathrm{T}}A^{\mathrm{T}}$, etc. Since an inner product can be viewed as a (1,1) matrix, two mutually transposed (1,1) matrices are identical. Hence, we get

$$\left\langle (x)A^{\dagger} \,\middle|\, y \right\rangle = \langle y \mid A(x)\rangle^* = \langle A(x) \mid y\rangle, \tag{11.64}$$

where the second equality is due to (11.2).

The other way around, we may use (11.64) for the definition of an adjoint operator of a linear transformation $A$. In fact, on the basis of (11.64), we have

$$\begin{aligned}\sum_{i,j,k} x_i^* \left(A^\dagger\right)_{ik} (G)_{kj} y_j &= \sum_{i,j,k} y_j (G^*)_{jk} (A^*)_{ki} x_i^* \\ &= \sum_{i,j,k} x_i^* y_j \left[\left(A^\dagger\right)_{ik} (G)_{kj} - (G^*)_{jk} \left(A_{ki}^*\right)\right] \\ &= \sum_{i,j,k} x_i^* y_j \left[\left(A^\dagger\right)_{ik} - (A^*)_{ki} (G)_{kj}\right] = 0.\end{aligned} \tag{11.65}$$

With the third equality of (11.65), we used $(G^*)_{jk} = (G)_{kj}$, i.e., $G^\dagger = G$ (Hermitian matrix). Thanks to the freedom in choice of basis vectors as well as $x_i$ and $y_i$, we must have

$$\left(A^\dagger\right)_{ik} = (A^*)_{ki}. \tag{11.66}$$

Adopting the matrix representation of (11.59) for $A$, we get [1]

$$\left(A^\dagger\right)_{ik} = \left(a_{ki}^*\right). \tag{11.67}$$

Thus, we confirm that the adjoint operator $A^\dagger$ is represented by a complex conjugate transposed matrix of $A$, in accordance with (11.1).

Taking complex conjugate of (11.64), we have

$$\left\langle (x)A^\dagger \,\middle|\, y \right\rangle^* = \left\langle y \,\middle|\, \left(A^\dagger\right)^\dagger (x) \right\rangle = \langle y \mid A(x) \rangle.$$

Comparing both sides of the second equality, we get

$$\left(A^\dagger\right)^\dagger = A. \tag{11.68}$$

In Chap. 9, we have seen how a matrix representation of a linear transformation $A$ is changed from $A_0$ to $A'$ by the basis vectors transformation. We have

$$A' = P^{-1} A_0 P. \tag{9.88}$$

In a similar manner, taking the adjoint of (9.88), we get

$$(A')^\dagger = P^\dagger A^\dagger \left(P^{-1}\right)^\dagger = P^\dagger A^\dagger \left(P^\dagger\right)^{-1}. \tag{11.69}$$

In (11.69), we denote the adjoint operator before the basis vectors transformation simply by $A^\dagger$ to avoid complicated notation. We also have

$$\left(A^\dagger\right)' = (A')^\dagger. \tag{11.70}$$

Meanwhile, suppose that $\left(A^\dagger\right)' = Q^{-1}A^\dagger Q$. Then, from (11.69) to (11.70), we have

$$P^\dagger = Q^{-1}.$$

Next, let us perform a calculation as below:

$$\begin{aligned}\left\langle (\alpha u + \beta v)A^\dagger \,\middle|\, y \right\rangle &= \langle y \mid A(\alpha u + \beta v)\rangle^* = \alpha^* \langle y \mid A(u)\rangle^* + \beta^* \langle y \mid A(v)\rangle^* \\ &= \alpha^* \left\langle (u)A^\dagger \,\middle|\, y \right\rangle + \beta^* \left\langle (v)A^\dagger \,\middle|\, y \right\rangle = \left\langle \alpha(u)A^\dagger + \beta(v)A^\dagger \,\middle|\, y \right\rangle.\end{aligned}$$

As $y$ is an element arbitrarily chosen from a relevant vector space, we have

$$\left\langle (\alpha u + \beta v)A^\dagger \right| = \left\langle \alpha(u)A^\dagger + \beta(v)A^\dagger \right|, \tag{11.71}$$

or

$$(\alpha u + \beta v)A^\dagger = \alpha(u)A^\dagger + \beta(v)A^\dagger. \tag{11.72}$$

Equation (11.71) states the equality of two vectors in an inner product space on both sides, whereas (11.72) states that in a vector space where the inner product is not defined. In either case, both (11.71) and (11.72) show that $A^\dagger$ is indeed a linear transformation. In fact, the matrix representation of (11.66) and (11.67) is independent of the concept of the inner product.

Suppose that there are two (or more) adjoint operators $B$ and $C$ that correspond to $A$. Then, from (11.64), we have

$$\langle (x)B \mid y\rangle = \langle (x)C \mid y\rangle = \langle y \mid A(x)\rangle^*. \tag{11.73}$$

Also we have

$$\langle (x)B - (x)C \mid y\rangle = \langle (x)(B - C) \mid y\rangle = 0. \tag{11.74}$$

As $x$ and $y$ are arbitrarily chosen elements, we get $B = C$, indicating the uniqueness of the adjoint operator.

It is of importance to examine how the norm of a vector is changed by the linear transformation. To this end, let us perform a calculation as below:

$$\begin{aligned}
&\left\langle (x)A^{\dagger} \,\middle|\, A(x) \right\rangle \\
&= \left(x_1^* \cdots x_n^*\right) \begin{pmatrix} a_{11}^* & \cdots & a_{n1}^* \\ \vdots & \ddots & \vdots \\ a_{1n}^* & \cdots & a_{nn}^* \end{pmatrix} \begin{pmatrix} \langle e_1| \\ \vdots \\ \langle e_n| \end{pmatrix} \left(|e_1\rangle \cdots |e_n\rangle\right) \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \\
&= \left(x_1^* \cdots x_n^*\right) A^{\dagger} G A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.
\end{aligned} \tag{11.75}$$

Equation (11.75) gives the norm of vector after its transformation.

We may have a case where the norm is conserved before and after the transformation. Actually, comparing (11.34) and (11.75), we notice that if $A^{\dagger} G A = G$, $\langle x|x\rangle = \left\langle (x)A^{\dagger} \,\middle|\, A(x) \right\rangle$. Let us have a following example for this.

*Example 11.3* Let us take two mutually orthogonal vectors $|e_1\rangle$ and $|e_2\rangle$ as basis vectors in the $xy$-plane (Fig. 11.1).

$$|x\rangle = \left(|e_1\rangle |e_2\rangle\right) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}. \tag{11.76}$$

Then we have

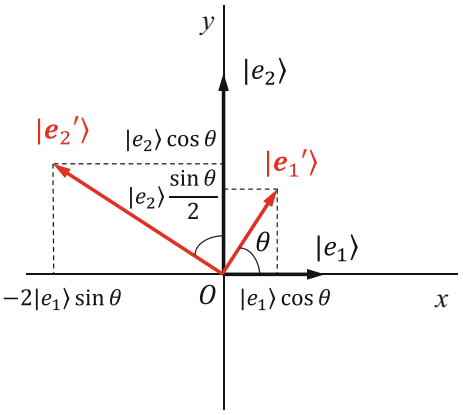


**Fig. 11.1** Basis vectors $|e_1\rangle$ and $|e_2\rangle$ in the $xy$-plane and their linear transformation by $R$

$$\begin{aligned}\langle x|x\rangle &= (x_1\ x_2)\begin{pmatrix}\langle e_1|\\ \langle e_2|\end{pmatrix}(|e_1\rangle|e_2\rangle)\begin{pmatrix}x_1\\ x_2\end{pmatrix}\\ &= (x_1\ x_2)\begin{pmatrix}\langle e_1\mid e_1\rangle & \langle e_1\mid e_2\rangle\\ \langle e_2\mid e_1\rangle & \langle e_2\mid e_2\rangle\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix} = (x_1\ x_2)\begin{pmatrix}1 & 0\\ 0 & 4\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix}\\ &= x_1^2+4x_2^2.\end{aligned} \tag{11.77}$$

In (11.77), we are considering that $|e_1\rangle$ and $|e_2\rangle$ are orthogonal, but we do not assume that $|e_2\rangle$ is normalized. That is, although $||e_1|| = \sqrt{\langle e_1|e_1\rangle} = 1,\ ||e_2|| = \sqrt{\langle e_2|e_2\rangle} = 2$.

Next, let us think of a following linear transformation $R$ whose matrix representation is given by

$$R = \begin{pmatrix}\cos\theta & -2\sin\theta\\ (\sin\theta)/2 & \cos\theta\end{pmatrix}. \tag{11.78}$$

The transformation matrix $R$ is geometrically represented in Fig. 11.1. Following (9.36), we have

$$R(|x\rangle) = (|e_1\rangle|e_2\rangle)\begin{pmatrix}\cos\theta & -2\sin\theta\\ (\sin\theta)/2 & \cos\theta\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix}. \tag{11.79}$$

As a result of the transformation $R$, the basis vectors $|e_1\rangle$ and $|e_2\rangle$ are transformed into $|e_1'\rangle$ and $|e_2'\rangle$, respectively, as in Fig. 11.1 such that

$$(|e_1'\rangle|e_2'\rangle) = (|e_1\rangle|e_2\rangle)\begin{pmatrix}\cos\theta & -2\sin\theta\\ (\sin\theta)/2 & \cos\theta\end{pmatrix}.$$

Taking an inner product of (11.79), we have

$$\begin{aligned}&\Big\langle (x)R^\dagger \Big|\ R(x)\Big\rangle\\ &= (x_1x_2)\begin{pmatrix}\cos\theta & (\sin\theta)/2\\ -2\sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & 4\end{pmatrix}\begin{pmatrix}\cos\theta & -2\sin\theta\\ (\sin\theta)/2 & \cos\theta\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix}\\ &= (x_1x_2)\begin{pmatrix}1 & 0\\ 0 & 4\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix} = x_1^2+4x_2^2.\end{aligned} \tag{11.80}$$

Putting

$$G = \begin{pmatrix}1 & 0\\ 0 & 4\end{pmatrix}, \tag{11.81}$$

we have $R^{\dagger}GR = G$.

Comparing (11.77) and (11.80), we notice that a norm of $|x\rangle$ remains unchanged after the transformation $R$. This means that $R$ is virtually a unitary transformation. A somewhat unfamiliar matrix form of $R$ resulted from the choice of basis vectors other than an orthonormal basis.

## 11.4 Orthonormal Basis

Now we introduce an orthonormal basis, the simplest and most important basis set in an inner product space. If we choose the orthonormal basis so that $\langle e_i|e_j\rangle = \delta_{ij}$, a Gram matrix $G = E$. Thus, $R^{\dagger}GR = G$ reads as $R^{\dagger}R = E$. In that case, a linear transformation is represented by a unitary matrix and it conserves a norm of a vector and an inner product with two arbitrary vectors.

So far we assumed that an adjoint operator $A^{\dagger}$ operates only on a row vector from the right, as is evident from (11.61). At the same time, $A$ operates only on the column vector from the left as in (11.62). To render the notation of (11.61) and (11.62) consistent with the associative law, we have to examine the commutability of $A^{\dagger}$ with $G$. In this context, choice of the orthonormal basis enables us to get through such a troublesome situation and largely eases matrix calculation. Thus,

$$\begin{aligned}
&\left\langle (x)A^{\dagger} \,\middle|\, A(x) \right\rangle \\
&= \left(x_1^* \cdots x_n^*\right) \begin{pmatrix} a_{11}^* & \cdots & a_{n1}^* \\ \vdots & \ddots & \vdots \\ a_{1n}^* & \cdots & a_{nn}^* \end{pmatrix} \begin{pmatrix} \langle e_1| \\ \vdots \\ \langle e_n| \end{pmatrix} \left(|e_1\rangle \cdots |e_n\rangle\right) \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \\
&= \left(x_1^* \cdots x_n^*\right) A^{\dagger} E A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \left(x_1^* \cdots x_n^*\right) A^{\dagger} A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.
\end{aligned} \tag{11.82}$$

At the same time, we adopt a simple notation as below instead of (11.75)

$$\left\langle xA^{\dagger}|Ax \right\rangle = \left(x_1^* \cdots x_n^*\right) A^{\dagger} A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}. \tag{11.83}$$

This notation has become now consistent with the associative law. Note that $A^{\dagger}$ and $A$ operate on either a column or row vector. We can also do without a symbol "|" in (11.83) and express it as

$$\left\langle xA^{\dagger}Ax\right\rangle = \left\langle xA^{\dagger}|Ax\right\rangle. \tag{11.84}$$

Thus, we can freely operate $A^{\dagger}$ and $A$ from both the left and right. By the same token, we rewrite (11.62) as

$$\langle y|Ax\rangle = \langle yAx\rangle = \left(y_1^* \cdots y_n^*\right)\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \left(y_1^* \cdots y_n^*\right)A\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}. \tag{11.85}$$

Here, notice that a vector $|x\rangle$ is represented by a column vector with respect to the orthonormal basis. Using (11.64), we have

$$\begin{aligned} \langle y|Ax\rangle^* &= \left\langle xA^{\dagger}\,\middle|\, y\right\rangle = \left(x_1^* \cdots x_n^*\right)A^{\dagger}\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \\ &= \left(x_1^* \cdots x_n^*\right)\begin{pmatrix} a_{11}^* & \cdots & a_{n1}^* \\ \vdots & \ddots & \vdots \\ a_{1n}^* & \cdots & a_{nn}^* \end{pmatrix}\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}. \end{aligned} \tag{11.86}$$

If in (11.83) we put $|y\rangle = |Ax\rangle$, $\left\langle xA^{\dagger}|Ax\right\rangle = \langle y|y\rangle \geq 0$. Thus, we define a norm of $|Ax\rangle$ as

$$\|Ax\| = \sqrt{\left\langle xA^{\dagger}|Ax\right\rangle}. \tag{11.87}$$

Now we are in a position to construct an orthonormal basis in $V^n$ using $n$ linearly independent vectors $|i\rangle$ $(1 \leq i \leq n)$. The following theorem is well-known as the Gram–Schmidt orthonormalization.

**Theorem 11.2: Gram–Schmidt Orthonormalization Theorem [4]** *Suppose that there are a set of linearly independent vectors $|i\rangle$ $(1 \leq i \leq n)$ in $V^n$. Then one can construct an orthonormal basis $|e_i\rangle$ $(1 \leq i \leq n)$ so that $\langle e_i|e_j\rangle = \delta_{ij}$ $(1 \leq j \leq n)$ and each vector $|e_i\rangle$ can be a linear combination of the vectors $|i\rangle$.*

*Proof* First, let us take $|1\rangle$. This can be normalized such that

$$|e_1\rangle = \frac{|1\rangle}{\sqrt{\langle 1 \mid 1\rangle}}, \quad \langle e_1|e_1\rangle = 1. \tag{11.88}$$

Next, let us take $|2\rangle$ and then make a following vector:

$$|e_2\rangle = \frac{1}{L_2}[|2\rangle - \langle e_1|2\rangle|e_1\rangle], \tag{11.89}$$

where $L_2$ is a normalization constant such that $\langle e_2|e_2\rangle = 1$. Note that $|e_2\rangle$ cannot be a zero vector. This is because if $|e_2\rangle$ were a zero vector, $|2\rangle$ and $|e_1\rangle$ (or $|1\rangle$) would be linearly dependent, in contradiction to the assumption. We have $\langle e_1|e_2\rangle = 0$. Thus, $|e_1\rangle$ and $|e_2\rangle$ are orthonormal.

After this, the proof is based upon mathematical induction. Suppose that the theorem is true of $(n-1)$ vectors. That is, let $|e_i\rangle$ $(1 \le i \le n-1)$ so that $\langle e_i|e_j\rangle = \delta_{ij}$ $(1 \le j \le n-1)$ and each vector $|e_i\rangle$ can be a linear combination of the vectors $|i\rangle$. Meanwhile, let us define

$$\widetilde{|n\rangle} \equiv |n\rangle - \sum_{j=1}^{n-1} \langle e_j|n\rangle|e_j\rangle. \tag{11.90}$$

Again, the vector $\widetilde{|n\rangle}$ cannot be a zero vector as asserted above. We have

$$\langle e_k\widetilde{|n\rangle} = \langle e_k|n\rangle - \sum_{j=1}^{n-1} \langle e_j|n\rangle\langle e_k|e_j\rangle = \langle e_k|n\rangle - \sum_{j=1}^{n-1} \langle e_j|n\rangle\delta_{kj} = 0, \tag{11.91}$$

where $1 \le k \le n-1$. The second equality comes from the assumption of the induction. The vector $\widetilde{|n\rangle}$ can always be normalized such that

$$|e_n\rangle = \frac{\widetilde{|n\rangle}}{\sqrt{\widetilde{\langle n|}\widetilde{|n\rangle}}}, \ \langle e_n|e_n\rangle = 1. \tag{11.92}$$

Thus, the theorem is proven. In (11.92), a phase factor $e^{i\theta}$ ($\theta$: an arbitrarily chosen real number) can be added such that

$$|e_n\rangle = \frac{e^{i\theta}\widetilde{|n\rangle}}{\sqrt{\widetilde{\langle n|}\widetilde{|n\rangle}}}. \tag{11.93}$$

To prove Theorem 11.2, we have used the following simple but important theorem.

**Theorem 11.3** *Let us have any n vectors* $|i\rangle \neq 0$ $(1 \le i \le n)$ *in* $V^n$ *and let these vectors* $|i\rangle$ *be orthogonal to one another. Then the vectors* $|i\rangle$ *are linearly independent.*

*Proof* Let us think of the following equation:

$$c_1|1\rangle + c_2|2\rangle + \cdots + c_n|n\rangle = 0. \tag{11.94}$$

Multiplying (11.94) by $\langle i|$ from the left and considering the orthogonality among the vectors, we have

$$c_i\langle i|i\rangle = 0. \tag{11.95}$$

Since $\langle i|i\rangle \neq 0$, $c_i = 0$. The above is true of any $c_i$ and $|i\rangle$. Then $c_1 = c_2 = \cdots = c_n = 0$. Thus, (11.94) implies that $|1\rangle, |2\rangle, \ldots, |n\rangle$ are linearly independent.

## References

1. Hassani S (2006) Mathematical physics. Springer, New York
2. Satake I (1974) Linear algebra. Shokabo, Tokyo (in Japanese)
3. Riley KF, Hobson MP, Bence SJ (2006) Mathematical methods for physics and engineering, 3rd edn. Cambridge University Press, Cambridge
4. Dennery P, Krzywicki A (1996) Mathematics for physicists. Dover, New York

# Chapter 12
# Hermitian Operators and Unitary Operators

Hermitian operators and unitary operators are quite often encountered in mathematical physics and, in particular, quantum physics. In this chapter, we investigate their basic properties. Both Hermitian operators and unitary operators fall under the category of normal operators. The normal matrices are characterized by an important fact that those matrices can be diagonalized by a unitary matrix. Moreover, Hermitian matrices always possess real eigenvalues. This fact largely eases mathematical treatment of quantum mechanics. In relation to these topics, in this chapter we investigate projection operators systematically. We find their important application to physicochemical problems in Part IV. We further investigate Hermitian quadratic forms and real symmetric quadratic forms as an important branch of matrix algebra. In connection with this topic, positive definiteness and nonnegative property of a matrix are an important concept. This characteristic is readily applicable to theory of differential operators, thus rendering this chapter closely related to basic concepts of quantum physics.

## 12.1 Projection Operators

In Chap. 10 we considered the decomposition of a vector space to direct sum of invariant subspaces. We also mentioned properties of idempotent operators. Moreover, we have shown how an orthonormal basis can be constructed from a set of linearly independent vectors. In this section, an orthonormal basis set is implied as basis vectors in an inner product space $V^n$.

Let us start with a concept of an orthogonal complement. Let $W$ be a subspace in $V^n$. Let us think of a set of vectors $|x\rangle$ such that

$$\left\{|x\rangle; \langle x|y\rangle = 0 \text{ for } {}^{\forall}|y\rangle \in W\right\}.$$



S. Hotta, *Mathematical Physical Chemistry*,
https://doi.org/10.1007/978-981-10-7671-8_12

We name this set $W^{\perp}$ and call it an orthogonal complement of $W$. The set $W^{\perp}$ forms a subspace of $V^n$. In fact, if $|a\rangle, |b\rangle \in W^{\perp}$, $\langle a|y\rangle = 0, \langle b|y\rangle = 0$. Since $(\langle a| + \langle b|)|y\rangle = \langle a|y\rangle + \langle b|y\rangle = 0$. Therefore, $|a\rangle + |b\rangle \in W^{\perp}$ and $\langle \alpha a|y\rangle = \alpha^*$ $\langle a|y\rangle = 0$. Hence, $|\alpha a\rangle = \alpha|a\rangle \in W^{\perp}$. Then, $W^{\perp}$ is a subspace of $V^n$.

**Theorem 12.1** *Let $W$ be a subspace and $W^{\perp}$ be its orthogonal complement in $V^n$. Then,*

$$V^n = W \oplus W^{\perp}. \tag{12.1}$$

*Proof* Suppose that an orthonormal basis comprising $|e_1\rangle, |e_2\rangle,$ and $|e_n\rangle$ spans $V^n$;

$$V^n = \mathrm{Span}\{|e_1\rangle, |e_2\rangle, \ldots, |e_n\rangle\}. \tag{12.2}$$

Of the orthonormal basis let $|e_1\rangle, |e_2\rangle,$ and $|e_r\rangle$ $(r<n)$ span $W$. Let an arbitrarily chosen vector from $V^n$ be $|x\rangle$. Then we have

$$|x\rangle = x_1|e_1\rangle + x_2|e_2\rangle + \cdots + x_n|e_n\rangle = \sum_{i=1}^{n} x_i|e_i\rangle. \tag{12.3}$$

Multiplying $\langle e_j|$ on (12.3) from the left, we have

$$\langle e_j|x\rangle = \sum_{i=1}^{n} x_i\langle e_j|e_i\rangle = \sum_{i=1}^{n} x_i\delta_{ij} = x_j. \tag{12.4}$$

That is,

$$|x\rangle = \sum_{i=1}^{n}\langle e_i|x\rangle|e_i\rangle. \tag{12.5}$$

Meanwhile, put

$$|x'\rangle = \sum_{i=1}^{r}\langle e_i|x\rangle|e_i\rangle. \tag{12.6}$$

Then we have $|x'\rangle \in W$. Also putting $|x''\rangle = |x\rangle - |x'\rangle$ and multiplying $\langle e_i|(1 \le i \le r)$ on it from the left, we get

$$\langle e_i|x''\rangle = \langle e_i|x\rangle - \langle e_i|x'\rangle = \langle e_i|x\rangle - \langle e_i|x\rangle = 0. \tag{12.7}$$

Taking account of $W = \mathrm{Span}\{|e_1\rangle, |e_2\rangle, \ldots, |e_r\rangle\}$, we get $|x''\rangle \in W^{\perp}$. That is, for $^{\forall}|x\rangle \in V^n$

$$|x\rangle = |x'\rangle + |x''\rangle. \tag{12.8}$$

This means that $V^n = W + W^\perp$. Meanwhile, we have $W \cap W^\perp = \{0\}$. In fact, suppose that $|x\rangle \in W \cap W^\perp$. Then $\langle x|x\rangle = 0$ because of the orthogonality. However, this implies that $|x\rangle = 0$. Consequently, we have $V^n = W \oplus W^\perp$. This completes the proof.

The consequence of the Theorem 12.1 is that the dimension of $W^\perp$ is $(n - r)$. In other words, we have

$$\dim V^n = n = \dim W + \dim W^\perp.$$

Moreover, the contents of the Theorem 12.1 can readily be generalized to more subspaces such that

$$V^n = W_1 \oplus W_2 \oplus \cdots \oplus W_r, \tag{12.9}$$

where $W_1, W_2, \ldots,$ and $W_r (r \leq n)$ are mutually orthogonal complements. In this case, $^\forall|x\rangle \in V^n$ can be expressed uniquely as the sum of individual vectors $|w_1\rangle, |w_2\rangle, \ldots,$ and $|w_r\rangle$ of each subspace, i.e.,

$$|x\rangle = |w_1\rangle + |w_2\rangle + \cdots + |w_r\rangle = |w_1 + w_2 + \cdots + w_r\rangle. \tag{12.10}$$

Let us define the following operators similarly to the case of (10.201):

$$P_i(|x\rangle) = |w_i\rangle (1 \leq i \leq r). \tag{12.11}$$

Thus, the operator $P_i$ extracts a vector $|w_i\rangle$ in a subspace $W_i$. Then we have

$$\begin{aligned}(P_1 + P_2 + \cdots + P_r)(|x\rangle) &= P_1|x\rangle + P_2|x\rangle + \cdots + P_r|x\rangle \\ &= |w_1\rangle + |w_2\rangle + \cdots + |w_r\rangle = |x\rangle.\end{aligned} \tag{12.12}$$

Since $|x\rangle$ is an arbitrarily chosen vector, we get

$$P_1 + P_2 + \cdots + P_r = E. \tag{12.13}$$

Moreover,

$$P_i[P_i(|x\rangle)] = P_i(|w_i\rangle) = |w_i\rangle (1 \leq i \leq r). \tag{12.14}$$

Therefore, from (12.11) and (12.14) we have

$$P_i[P_i(|x\rangle)] = P_i(|x\rangle). \tag{12.15}$$

The vector $|x\rangle$ is arbitrarily chosen, and so we get

$$P_i^2 = P_i. \tag{12.16}$$

Choose another arbitrary vector $|y\rangle \in V^n$ such that

$$|y\rangle = |u_1\rangle + |u_2\rangle + \cdots + |u_r\rangle = |u_1 + u_2 + \cdots + u_r\rangle. \tag{12.17}$$

Then, we have

$$\begin{aligned} \langle x|P_i y\rangle &= \langle w_1 + w_2 + \cdots + w_r|P_i|u_1 + u_2 + \cdots + u_r\rangle \\ &= \langle w_1 + w_2 + \cdots + w_r|u_i\rangle = \langle w_i u_i\rangle. \end{aligned} \tag{12.18}$$

With the last equality, we used the mutual orthogonality of the subspaces. Meanwhile, we have

$$\begin{aligned} \langle y|P_i x\rangle^* &= \langle u_1 + u_2 + \cdots + u_r|P_i|w_1 + w_2 + \cdots + w_r\rangle^* \\ &= \langle u_1 + u_2 + \cdots + u_r|w_i\rangle^* = \langle u_i|w_i\rangle^* = \langle w_i|u_i\rangle. \end{aligned} \tag{12.19}$$

Comparing (12.18) and (12.19), we get

$$\langle x|P_i y\rangle = \langle y|P_i x\rangle^* = \left\langle x\middle|P_i^\dagger y\right\rangle, \tag{12.20}$$

where we used (11.64) with the second equality. Since $|x\rangle$ and $|y\rangle$ are arbitrarily chosen, we get

$$P_i^\dagger = P_i. \tag{12.21}$$

Equation (12.21) shows that $P_i$ is Hermitian.

The above discussion parallels that made in Sect. 10.4 with an idempotent operator. We have a following definition about a projection operator.

**Definition 12.1** An operator $P$ is said to be a projection operator if $P^2 = P$ and $P^\dagger = P$. That is, an idempotent and Hermitian operator is a projection operator.

As described above, a projection operator is characterized by (12.16) and (12.21). An idempotent operator does not premise the presence of an inner product space, but we only need a direct sum of subspaces. In contrast, if we deal with the projection operator, we are thinking of orthogonal compliments as subspaces and their direct sum. The projection operator can adequately be defined in an inner product vector space having an orthonormal basis.

From (12.13) we have

$$\begin{aligned}&(P_1+P_2+\cdots+P_r)(P_1+P_2+\cdots+P_r)\\&=\sum_{i=1}^{r}P_i^2+\sum_{i\neq j}P_iP_j=\sum_{i=1}^{r}P_i+\sum_{i\neq j}P_iP_j=E+\sum_{i\neq j}P_iP_j=E.\end{aligned}\tag{12.22}$$

In (12.22), we used (12.13) and (12.16). Therefore, we get

$$\sum_{i\neq j}P_iP_j=0.$$

In particular, we have

$$P_iP_j=0,\quad P_jP_i=0\ (i\neq j).\tag{12.23}$$

In fact, we have

$$P_i\big[P_j(|x\rangle)\big]=P_i\big(|w_j\rangle\big)=0\ (i\neq j,\ 1\leq i,\ j\leq n).\tag{12.24}$$

The second equality comes from $W_i\cap W_j=\{0\}$. Notice that in (12.24), indices $i$ and $j$ are interchangeable. Again, $|x\rangle$ is arbitrarily chosen, and so (12.23) holds. Combining (12.16) and (12.23), we write

$$P_iP_j=\delta_{ij}.\tag{12.25}$$

In virtue of the relation (12.23), $P_i+P_j$ $(i\neq j)$ is a projection operator as well [1]. In fact, we have

$$\big(P_i+P_j\big)^2=P_i^2+P_iP_j+P_jP_i+P_j^2=P_i^2+P_j^2=P_i+P_j,$$

where the second equality comes from (12.23). Also we have

$$\big(P_i+P_j\big)^\dagger=P_i^\dagger+P_j^\dagger=P_i+P_j.$$

The following notation is often used:

$$P_i=\frac{|w_i\rangle\langle w_i|}{\||w_i\|\cdot\||w_i\|}.\tag{12.26}$$

Then we have

$$\begin{aligned} P_i|x\rangle &= \frac{|w_i\rangle\langle w_i|}{\||w_i\| \cdot \||w_i\|}|(|w_1\rangle + |w_2\rangle + \cdots + |w_s\rangle)| \\ &= [|w_i\rangle(\langle w_i|w_1\rangle + \langle w_i|w_2\rangle + \cdots + \langle w_i|w_s\rangle)]/\langle w_i|w_i\rangle \\ &= [|w_i\rangle\langle w_i|w_i\rangle]/\langle w_i|w_i\rangle = |w_i\rangle. \end{aligned}$$

Furthermore, we have

$$P_i^2|x\rangle = P_i|w_i\rangle = |w_i\rangle = P_i|x\rangle. \tag{12.27}$$

Equation (12.16) is recovered accordingly. Meanwhile,

$$(|w_i\rangle\langle w_i|)^\dagger = (\langle w_i|)^\dagger(|w_i\rangle)^\dagger = |w_i\rangle\langle w_i|. \tag{12.28}$$

Hence, we recover

$$P_i^\dagger = P_i. \tag{12.21}$$

In (12.28), we used $(AB)^\dagger = B^\dagger A^\dagger$. In fact, we have

$$\langle x|B^\dagger A^\dagger|y\rangle = \left\langle xB^\dagger \middle| A^\dagger y\right\rangle = \langle yA|Bx\rangle^* = \langle y|AB|x\rangle^* = \langle x|(AB)^\dagger|y\rangle. \tag{12.29}$$

With the second equality of (12.29), we used (11.86) where $A$ is replaced with $B$ and $|y\rangle$ is replaced with $A^\dagger|y\rangle$. Since (12.29) holds for arbitrarily chosen vectors $|x\rangle$ and $|y\rangle$, comparing the first and last sides of (12.29) we have

$$(AB)^\dagger = B^\dagger A^\dagger. \tag{12.30}$$

We can express (12.29) alternatively as follows:

$$\left\langle xB^\dagger \middle| A^\dagger y\right\rangle = \left\langle y\left(A^\dagger\right)^\dagger \middle| \left(B^\dagger\right)^\dagger x\right\rangle^* = \langle yA|Bx\rangle^* = \langle yABx\rangle^*, \tag{12.31}$$

where with the second equality we used (11.68). Also recall the remarks after (11.83) with the expressions of (12.29) and (12.31). Other notations can be adopted.

We can view a projection operator under a more strict condition. Related operators can be defined as well. As in (12.26), let us define an operator such that

$$\widetilde{P_k} = |e_k\rangle\langle e_k|. \tag{12.32}$$

Operating $\widetilde{P}_i$ on $|x\rangle = x_1|e_1\rangle + x_2|e_2\rangle + \cdots + x_n|e_n\rangle$ from the left, we get

$$\widetilde{P_k}|x\rangle = |e_k\rangle\langle e_k|\left(\sum_{j=1}^{n} x_j|e_j\rangle\right) = |e_k\rangle\left(\sum_{j=1}^{n} x_j\langle e_k|e_j\rangle\right) = |e_k\rangle\left(\sum_{j=1}^{n} x_j\delta_{kj}\right)$$
$$= x_k|e_k\rangle.$$

Thus, we find that $\widetilde{P_k}$ plays the same role as $P^{(k)}$ defined in (10.201). Represented by a matrix, $\widetilde{P_k}$ has the same structure as that denoted in (10.205). Evidently,

$$\left(\widetilde{P_k}\right)^2 = \widetilde{P_k},\ \widetilde{P_k}^\dagger = \widetilde{P_k},\ \widetilde{P_1} + \widetilde{P_2} + \cdots + \widetilde{P_n} = E. \tag{12.33}$$

Now let us modify $P^{(k)}$ in (10.201). There $P^{(k)} = \delta_i^{(k)}\delta_{(k)}^j$, where only the $(k,\ k)$ element is 1, otherwise 0, in the $(n,\ n)$ matrix. We define a matrix $P_{(m)}^{(k)} = \delta_i^{(k)}\delta_{(m)}^j$. A full matrix representation for it is

$$P_{(m)}^{(k)} = \begin{pmatrix} 0 & & & & & & \\ & \ddots & & & 1 & & \\ & & 0 & & & & \\ & & & 0 & & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 0 \end{pmatrix}, \tag{12.34}$$

where only the $(k, m)$ element is 1, otherwise 0. In an example of (12.34), $P_{(m)}^{(k)}$ is an upper triangle matrix $(k<m)$. Therefore, its eigenvalues are all zero, and so $P_{(m)}^{(k)}$ is a nilpotent matrix. If $k > m$, the matrix is a lower triangle matrix and nilpotent as well. Such a matrix is not Hermitian (nor a projection operator), as can be immediately seen from the matrix form of (12.34). Because of the properties of nilpotent matrices mentioned in Sect. 10.3, $P_{(m)}^{(k)}(k \neq m)$ is not diagonalizable either.

Various relations can be extracted. As an example, we have

$$P_{(m)}^{(k)}P_{(n)}^{(l)} = \sum_q \delta_i^{(k)}\delta_{(m)}^q\delta_q^{(l)}\delta_{(n)}^j = \delta_i^{(k)}\delta_{ml}\delta_{(n)}^j = \delta_{ml}P_{(n)}^{(k)}. \tag{12.35}$$

Note that $P_{(k)}^{(k)} \equiv P^{(k)}$ defined in (10.201). From (12.35), moreover, we have

$$P_{(m)}^{(k)}P_{(n)}^{(m)} = P_{(n)}^{(k)},\ P_{(m)}^{(n)}P_{(n)}^{(m)} = P_{(n)}^{(n)},\ P_{(n)}^{(m)}P_{(m)}^{(n)} = P_{(m)}^{(m)}$$
$$P_{(m)}^{(k)}P_{(m)}^{(k)} = \delta_{mk}P_{(m)}^{(k)},\ P_{(m)}^{(m)}P_{(m)}^{(m)} = \left[P_{(m)}^{(m)}\right]^2 = P_{(m)}^{(m)}, \text{etc.}$$

These relations remain unchanged after the unitary similarity transformation by $U$. For instance, taking the first equation of the above, we have

$$U^\dagger P^{(k)}_{(m)} P^{(m)}_{(n)} U = \left[U^\dagger P^{(k)}_{(m)} U\right]\left[U^\dagger P^{(m)}_{(n)} U\right] = U^\dagger P^{(k)}_{(n)} U.$$

Among these operators, only $P^{(k)}_{(k)}$ is eligible for a projection operator. We will encounter further examples in Part IV.

Using $P^{(k)}_{(k)}$ in (11.62), we have

$$\left\langle y \middle| P^{(k)}_{(k)}(x) \right\rangle = \left(y_1^* \cdots y_n^*\right) G P^{(k)}_{(k)} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}. \tag{12.36}$$

Within a framework of an orthonormal basis where $G = E$, the representation is largely simplified to be

$$\left\langle y \middle| P^{(k)}_{(k)} x \right\rangle = \left(y_1^* \cdots y_n^*\right) P^{(k)}_{(k)} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = y_k^* x_k. \tag{12.37}$$

## 12.2 Normal Operators

There are a large group of operators called normal operators that play an important role in mathematical physics, especially quantum physics. A normal operator is defined as an operator on an inner product space that commutes with its adjoint operator. That is, let $A$ be a normal operator. Then, we have

$$AA^\dagger = A^\dagger A. \tag{12.38}$$

The normal operators include an Hermitian operator $H$ defined as $H^\dagger = H$ as well as a unitary operator $U$ defined as $UU^\dagger = U^\dagger U = E$.

In this condition, let us estimate the norm of $\left|A^\dagger x\right\rangle$ together with $|Ax\rangle$ defined by (11.87). If $A$ is a normal operator,

$$\left\|A^\dagger x\right\| = \sqrt{\left\langle x\left(A^\dagger\right)^\dagger \middle| A^\dagger x \right\rangle} = \sqrt{\left\langle xAA^\dagger x \right\rangle} = \sqrt{\left\langle xA^\dagger Ax \right\rangle} = \|Ax\|. \tag{12.39}$$

The other way around suppose that $\|Ax\| = \left\|A^\dagger x\right\|$. Then, since $\|Ax\|^2 = \left\|A^\dagger x\right\|^2$, $\left\langle xA^\dagger Ax\right\rangle = \left\langle xAA^\dagger x\right\rangle$. That is, $\langle x|A^\dagger A - AA^\dagger|x\rangle = 0$ for an arbitrarily chosen vector $|x\rangle$. To assert $A^\dagger A - AA^\dagger = 0$, i.e., $A^\dagger A = AA^\dagger$ on the assumption that $\langle x|A^\dagger A - AA^\dagger|x\rangle = 0$, we need the following theorems.

**Theorem 12.2 [2]** *A linear transformation A on an inner product space is the zero transformation if and only if* $\langle y|Ax\rangle = 0$ *for any vectors* $|x\rangle$ and $|y\rangle$.

*Proof* If $A = 0$, then $\langle y|Ax\rangle = \langle y|0\rangle = 0$. This is because in (11.3) putting $\beta = 1 = -\gamma$ and $|b\rangle = |c\rangle$, we get $\langle a|0\rangle = 0$. Conversely, suppose that $\langle y|Ax\rangle = 0$ for any vectors $|x\rangle$ and $|y\rangle$. Then, putting $|y\rangle = |Ax\rangle$, $\left\langle xA^\dagger|Ax\right\rangle = 0$ and $\langle y|y\rangle = 0$. This implies that $|y\rangle = |Ax\rangle = 0$. For $|Ax\rangle = 0$ to hold for any $|x\rangle$ we must have $A = 0$. Note here that if $A$ is a singular matrix, for some vectors $|x\rangle, |Ax\rangle = 0$. However, even though $A$ is singular, for $|Ax\rangle = 0$ to hold for any $|x\rangle$, $A = 0$.

We have another important theorem under a further restricted condition.

**Theorem 12.3 [2]** *A linear transformation A on an inner product space is the zero transformation if and only if* $\langle x|Ax\rangle = 0$ *for any vectors* $|x\rangle$.

*Proof* As in the case of Theorem 12.2, a necessary condition is trivial. To prove a sufficient condition, let us consider the following:

$$\begin{aligned}\langle x+y|A(x+y)\rangle &= \langle x|Ax\rangle + \langle y|Ay\rangle + \langle x|Ay\rangle + \langle y|Ax\rangle,\\ \langle x|Ay\rangle + \langle y|Ax\rangle &= \langle x+y|A(x+y)\rangle - \langle x|Ax\rangle - \langle y|Ay\rangle.\end{aligned} \tag{12.40}$$

From the assumption that $\langle x|Ax\rangle = 0$ with any vectors $|x\rangle$, we have

$$\langle x|Ay\rangle + \langle y|Ax\rangle = 0. \tag{12.41}$$

Meanwhile, replacing $|y\rangle$ by $|iy\rangle$ in (12.41), we get

$$\langle x|Aiy\rangle + \langle iy|Ax\rangle = i[\langle x|Ay\rangle - \langle y|Ax\rangle] = 0. \tag{12.42}$$

That is,

$$\langle x|Ay\rangle - \langle y|Ax\rangle = 0. \tag{12.43}$$

Combining (12.41) and (12.43), we get

$$\langle x|Ay\rangle = 0. \tag{12.44}$$

Theorem 12.2 means that $A = 0$, indicating that the sufficient condition holds. This completes the proof.

Thus returning to the beginning, i.e., remarks made after (12.39), we establish the following theorem.

**Theorem 12.4** *A necessary and sufficient condition for a linear transformation A on an inner product space to be a normal operator is*

$$\left\|A^{\dagger}x\right\| = \|Ax\|. \tag{12.45}$$

## 12.3 Unitary Diagonalization of Matrices

A normal operator has a distinct property. The normal operator can be diagonalized by a similarity transformation by a unitary matrix. The transformation is said to be a unitary similarity transformation. Let us prove the following theorem.

**Theorem 12.5 [3]** *A necessary and sufficient condition for a matrix A to be diagonalized by unitary similarity transformation is that the matrix A is a normal matrix.*

*Proof* To prove the necessary condition, suppose that $A$ can be diagonalized by a unitary matrix $U$. That is,

$$U^{\dagger}AU = D, \quad \text{i.e. } A = UDU^{\dagger} \text{ and } A^{\dagger} = UD^{\dagger}U^{\dagger}, \tag{12.46}$$

where $D$ is a diagonal matrix. Then

$$\begin{aligned} AA^{\dagger} &= \left(UDU^{\dagger}\right)\left(UD^{\dagger}U^{\dagger}\right) = UDD^{\dagger}U^{\dagger} = UD^{\dagger}DU^{\dagger} = \left(UD^{\dagger}U^{\dagger}\right)\left(UDU^{\dagger}\right) \\ &= A^{\dagger}A. \end{aligned} \tag{12.47}$$

For the third equality, we used $DD^{\dagger} = D^{\dagger}D$ (i.e., $D$ and $D^{\dagger}$ are commutable). This shows that $A$ is a normal matrix.

To prove the sufficient condition, let us show that a normal matrix can be diagonalized by unitary similarity transformation. The proof is due to mathematical induction, as is the case with Theorem 10.1.

First we show that Theorem is true of a (2,2) matrix. Suppose that one of eigenvalues of $A_2$ is $\alpha_1$ and that its corresponding eigenvector is $|x_1\rangle$. Following procedures of the proof for Theorem 10.1 and remembering the Gram–Schmidt orthonormalization theorem, we can construct a unitary matrix $U_1$ such that

$$U_1 = (|x_1\rangle|p_1\rangle), \tag{12.48}$$

where $|x_1\rangle$ represents a column vector and $|p_1\rangle$ is another arbitrarily determined column vector. Then we can convert $A_2$ to a triangle matrix such that

$$\widetilde{A_2} \equiv U_1^\dagger A_2 U_1 = \begin{pmatrix} \alpha_1 & x \\ 0 & y \end{pmatrix}. \tag{12.49}$$

Then we have

$$\begin{aligned} \widetilde{A_2}\left[\widetilde{A_2}\right]^\dagger &= \left(U_1^\dagger A_2 U_1\right)\left(U_1^\dagger A_2 U_1\right)^\dagger = \left(U_1^\dagger A_2 U_1\right)\left(U_1^\dagger A_2^\dagger U_1\right) \\ &= U_1^\dagger A_2 A_2^\dagger U_1 = U_1^\dagger A_2^\dagger A_2 U_1 = (U_1^\dagger A_2^\dagger U_1)(U_1^\dagger A_2 U_1) \\ &= [\widetilde{A_2}]^\dagger \widetilde{A_2}. \end{aligned} \tag{12.50}$$

With the fourth equality, we used the supposition that $A_2$ is a normal matrix. Equation (12.50) means that $\widetilde{A_2}$ defined in (12.49) is a normal operator. Via simple matrix calculations, we have

$$\widetilde{A_2}\left[\widetilde{A_2}\right]^\dagger = \begin{pmatrix} |\alpha_1|^2 + |x|^2 & xy^* \\ x^* y & |y|^2 \end{pmatrix}, \quad \left[\widetilde{A_2}\right]^\dagger \widetilde{A_2} = \begin{pmatrix} |\alpha_1|^2 & \alpha_1^* x \\ \alpha_1 x^* & |x|^2 + |y|^2 \end{pmatrix}. \tag{12.51}$$

For (12.50) to hold, we must have $x = 0$ in (12.51). Accordingly, we get

$$\widetilde{A_2} = \begin{pmatrix} \alpha_1 & 0 \\ 0 & y \end{pmatrix}. \tag{12.52}$$

This implies that a normal matrix $A_2$ has been diagonalized by the unitary similarity transformation.

Now let us examine a general case where we consider a $(n,n)$ square normal matrix $A_n$. Let $\alpha_n$ be one of eigenvalues of $A_n$. On the basis of the argument of the (2,2) matrix case, after a suitable similarity transformation by a unitary matrix $\widetilde{U}$ we first have

$$\widetilde{A_n} = \left(\widetilde{U}\right)^\dagger A_n \widetilde{U}, \tag{12.53}$$

where we can put

$$\widetilde{A_n} = \begin{pmatrix} \alpha_n & \boldsymbol{x}^T \\ 0 & B \end{pmatrix}, \tag{12.54}$$

where $\boldsymbol{x}$ is a column vector of order $(n-1)$, 0 is a zero column vector of order $(n-1)$, and $B$ is a $(n-1,\ n-1)$ matrix. Then we have

$$\left[\widetilde{A_n}\right]^\dagger = \begin{pmatrix} \alpha_n^* & 0 \\ \boldsymbol{x}^* & B^\dagger \end{pmatrix}, \tag{12.55}$$

where $\boldsymbol{x}^*$ is a complex column vector. Performing matrix calculations, we have

$$\widetilde{A_n}\left[\widetilde{A_n}\right]^\dagger = \begin{pmatrix} |\alpha_n|^2 + \boldsymbol{x}^T\boldsymbol{x}^* & \boldsymbol{x}^T B^\dagger \\ B\boldsymbol{x}^* & BB^\dagger \end{pmatrix},$$

$$\left[\widetilde{A_n}\right]^\dagger\left[\widetilde{A_n}\right] = \begin{pmatrix} |\alpha_n|^2 & \alpha_n^*\boldsymbol{x}^T \\ \alpha_n\boldsymbol{x}^* & \boldsymbol{x}^*\boldsymbol{x}^T + B^\dagger B \end{pmatrix}. \tag{12.56}$$

For $\widetilde{A_n}\left[\widetilde{A_n}\right]^\dagger = \left[\widetilde{A_n}\right]^\dagger\left[\widetilde{A_n}\right]$ to hold with (12.56), we must have $\boldsymbol{x} = \boldsymbol{0}$. Thus we get

$$\widetilde{A_n} = \begin{pmatrix} \alpha_n & 0 \\ 0 & B \end{pmatrix}. \tag{12.57}$$

Since $\widetilde{A_n}$ is a normal matrix, so is $B$. According to mathematical induction, let us assume that the theorem holds with a $(n-1, n-1)$ matrix, i.e., $B$. Then, also from the assumption there exists a unitary matrix $C$ and a diagonal matrix $D$, both of order $(n-1)$, such that $BC = CD$. Hence,

$$\begin{pmatrix} \alpha_n & 0 \\ 0 & B \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & C \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & C \end{pmatrix}\begin{pmatrix} \alpha_n & 0 \\ 0 & D \end{pmatrix}. \tag{12.58}$$

Here putting

$$\widetilde{C_n} = \begin{pmatrix} 1 & 0 \\ 0 & C \end{pmatrix} \text{ and } \widetilde{D_n} = \begin{pmatrix} \alpha_n & 0 \\ 0 & D \end{pmatrix}, \tag{12.59}$$

we get

$$\widetilde{A_n}\widetilde{C_n} = \widetilde{C_n}\widetilde{D_n}. \tag{12.60}$$

As $\widetilde{C_n}$ is a $(n,n)$ unitary matrix, $\left[\widetilde{C_n}\right]^\dagger\widetilde{C_n} = \widetilde{C_n}\left[\widetilde{C_n}\right]^\dagger = E$. Hence, $\left[\widetilde{C_n}\right]^\dagger\widetilde{A_n}\widetilde{C_n} = \widetilde{D_n}$. Thus, from (12.53) finally we get

$$\left[\widetilde{C_n}\right]^\dagger(\widetilde{U})^\dagger A_n\widetilde{U}\widetilde{C_n} = \widetilde{D_n}. \tag{12.61}$$

Putting $\widetilde{U}\widetilde{C_n} = V$, $V$ being another unitary operator,

$$V^\dagger A_n V = \widetilde{D_n}. \tag{12.62}$$

This completes the proof.

A direct consequence of Theorem 12.5 is that with any normal matrix we can find a set of orthonormal eigenvectors corresponding to individual eigenvalues whether or not those are degenerate.

In Sect. 8.2 we dealt with a decomposition of a linear vector space and relevant reduction of an operator when discussing canonical forms of matrices. In this context, Theorem 12.5 gives a simple and clear criterion for this. Equation (12.57) implies that a $(n-1, n-1)$ submatrix $B$ can further be reduced to matrices having lower dimensions. Considering that a diagonal matrix is a special case of triangle matrices, a normal matrix that has been diagonalized by the unitary similarity transformation gives eigenvalues by its diagonal elements.

From a point of view of the aforementioned aspect, let us consider the characteristics of normal matrices, starting with the discussion about the invariant subspaces. We have a following important theorem.

**Theorem 12.6** *Let $A$ be a normal matrix and let one of its eigenvalues be $\alpha$. Let $W_\alpha$ be an eigenspace corresponding to $\alpha$. Then, $W_\alpha$ is both $A$-invariant and $A^\dagger$-invariant. Also $W_\alpha^\perp$ is both $A$-invariant and $A^\dagger$-invariant.*

*Proof* Theorem 12.5 ensures that a normal matrix is diagonalized by unitary similarity transformation. Therefore, we deal with only "proper" eigenvalues and eigenvectors here. First we show if a subspace $W$ is $A$-invariant, then its orthogonal complements $W^\perp$ is $A^\dagger$-invariant. In fact, suppose that $|x\rangle \in W$ and $|x'\rangle \in W^\perp$. Then, from (11.64) and (11.86), we have

$$\langle x'|Ax\rangle = 0 = \left\langle xA^\dagger|x'\right\rangle^* = \left\langle x\middle|A^\dagger x'\right\rangle^*. \tag{12.63}$$

The first equality comes from the fact that $|x\rangle \in W \Rightarrow A|x\rangle (= |Ax\rangle) \in W$ as $W$ is $A$-invariant. From the last equality of (12.63), we have

$$A^\dagger|x'\rangle \left(= \left|A^\dagger x'\right\rangle\right) \in W^\perp. \tag{12.64}$$

That is, $W^\perp$ is $A^\dagger$-invariant.

Next suppose that $|x\rangle \in W_\alpha$. Then we have

$$AA^\dagger|x\rangle = A^\dagger A|x\rangle = A^\dagger(\alpha|x\rangle) = \alpha A^\dagger|x\rangle. \tag{12.65}$$

Therefore, $A^\dagger|x\rangle \in W_\alpha$. This means that $W_\alpha$ is $A^\dagger$-invariant. From the above remark, $W_\alpha^\perp$ is $\left(A^\dagger\right)^\dagger$-invariant, i.e., $A$-invariant accordingly. This completes the proof.

From Theorem 12.5, we know that the resulting diagonal matrix $\widetilde{D_n}$ in (12.62) has a form with $n$ eigenvalues ($\alpha_n$) some of which may be multiple roots arranged in diagonal elements. After diagonalizing the matrix, those eigenvalues can be sorted

out according to different eigenvalues $\alpha_1, \alpha_2, \ldots,$ and $\alpha_s$. This can also be done by unitary similarity transformation. The relevant unitary matrix $U$ is represented as

$$U = \begin{pmatrix} 1 & & & & & & & & & & \\ & \ddots & & & & & & & & & \\ & & 1 & & & & & & & & \\ & & & 0 & \cdots & 1 & & & & & \\ & & & & 1 & & & & & & \\ & & & \vdots & \ddots & \vdots & & & & & \\ & & & & & & 1 & & & & \\ & & & 1 & \cdots & 0 & & & & & \\ & & & & & & & 1 & & & \\ & & & & & & & & \ddots & & \\ & & & & & & & & & 1 & \end{pmatrix}, \tag{12.66}$$

where except $(i, j)$ and $(j, i)$ elements equal to 1, all the off-diagonal elements are zero. If operated from the left, $U$ exchanges the $i$th and $j$th rows of the matrix. If operated from the right, $U$ exchanges the $i$th and $j$th columns of the matrix. Note that $U$ is at once unitary and Hermitian with eigenvalue 1 or $-1$. Note that $U^2 = E$. This is because exchanging two columns (or two rows) two times produces identity transformation. Thus performing such unitary similarity transformations appropriate times, we get

$$\widetilde{D}_n \sim \begin{pmatrix} \alpha_1 & & & & & & & & & \\ & \ddots & & & & & & & & \\ & & \alpha_1 & & & & & & & \\ & & & \alpha_2 & & & & & & \\ & & & & \ddots & & & & & \\ & & & & & \alpha_2 & & & & \\ & & & & & & \ddots & & & \\ & & & & & & & \alpha_s & & \\ & & & & & & & & \ddots & \\ & & & & & & & & & \alpha_s \end{pmatrix}. \tag{12.67}$$

The matrix is identical to that represented in (10.181).

In parallel, $V^n$ is decomposed to mutually orthogonal subspaces associated with different eigenvalues $\alpha_1, \alpha_2, \ldots,$ and $\alpha_s$ such that

$$V^n = W_{\alpha_1} \oplus W_{\alpha_2} \oplus \cdots \oplus W_{\alpha_s}. \tag{12.68}$$

This expression is formally identical to that represented in (10.191). Note, however, that in (10.181), orthogonal subspaces are not implied. At the same time, $A_n$ is reduced to

$$A_n \sim \begin{pmatrix} A^{(1)} & & \cdots & \\ & A^{(2)} & & \\ \vdots & & \ddots & \vdots \\ & & \cdots & A^{(s)} \end{pmatrix}, \tag{12.69}$$

according to the different eigenvalues.

A normal operator has other distinct properties. Following theorems are good examples.

**Theorem 12.7** *Let $A$ be a normal operator on $V^n$. Then $|x\rangle$ is an eigenvector of $A$ with an eigenvalue $\alpha$, if and only if $|x\rangle$ is an eigenvector of $A^\dagger$ with an eigenvalue $\alpha^*$.*

*Proof* We apply (12.45) for the proof. Both $(A-\alpha E)^\dagger = A^\dagger - \alpha^* E$ and $(A-\alpha E)$ are normal, since $A$ is normal. Consequently, we have $\|(A-\alpha E)x\| = 0$ if and only if $\left\|\left(A^\dagger - \alpha^* E\right)x\right\| = 0$. Since only the zero vector has a zero norm, we get

$$(A-\alpha E)|x\rangle = 0 \text{ if and only if } \left(A^\dagger - \alpha^* E\right)|x\rangle = 0.$$

This completes the proof.

**Theorem 12.8** *Let $A$ be a normal operator on $V^n$. Then, eigenvectors corresponding to different eigenvalues are mutually orthogonal.*

*Proof* Let $A$ be a normal operator on $V^n$. Let $|u\rangle$ be an eigenvector corresponding to an eigenvalue $\alpha$, and $|v\rangle$ be an eigenvector corresponding to an eigenvalue $\beta$ with $\alpha \neq \beta$. Then we have

$$\begin{aligned} \alpha\langle v|u\rangle = \langle v|\alpha u\rangle = \langle v|Au\rangle = \left\langle u\middle|A^\dagger v\right\rangle^* = \left\langle u\middle|\beta^\dagger v\right\rangle^* = \langle \beta^* v|u\rangle \\ = \langle \beta v|u\rangle, \end{aligned} \tag{12.70}$$

where with the fourth equality we used Theorem 12.7. Then we get

$$(\alpha-\beta)\langle v|u\rangle = 0.$$

Since $\alpha - \beta \neq 0$, $\langle v|u\rangle = 0$. Namely, the eigenvectors $|u\rangle$ and $|v\rangle$ are mutually orthogonal.

In (10.208), we mentioned the decomposition of diagonalizable matrices. As for the normal matrices, we have a related matrix decomposition. Let $A$ be a normal operator. Then, according to Theorem 12.5, $A$ can be diagonalized and expressed as (12.67). This is equivalently expressed as a following succinct relation. That is, if we choose $U$ for a diagonalizing unitary matrix, we have

$$U^{\dagger}AU = \alpha_1 P_1 + \alpha_2 P_2 + \cdots + \alpha_s P_s, \tag{12.71}$$

where $\alpha_1, \alpha_2, \ldots,$ and $\alpha_s$ are different eigenvalues of $A$; $P_l (1 \leq l \leq s)$ is described such that, e.g.,

$$P_1 = \begin{pmatrix} E_{n_1} & & \cdots & \\ & 0_{n_2} & & \\ \vdots & & \ddots & \vdots \\ & & \cdots & 0_{n_s} \end{pmatrix}, \tag{12.72}$$

where $E_{n_1}$ stands for a $(n_1, n_1)$ identity matrix with $n_1$ corresponding to multiplicity of $\alpha_1$. A matrix represented by $0_{n_2}$ is a $(n_2, n_2)$ zero matrix, and so forth. This expression is in accordance with (12.69). From a matrix form (12.72), obviously $P_l (1 \leq l \leq s)$ is a projection operator. Thus, operating $U$ and $U^{\dagger}$ on both sides (12.71) from the left and right of (12.71), respectively, we obtain

$$A = \alpha_1 U P_1 U^{\dagger} + \alpha_2 U P_2 U^{\dagger} + \cdots + \alpha_s U P_s U^{\dagger}. \tag{12.73}$$

Defining $\widetilde{P}_l \equiv U P_l U^{\dagger} (1 \leq l \leq s)$, we have

$$A = \alpha_1 \widetilde{P_1} + \alpha_1 \widetilde{P_2} + \cdots + \alpha_s \widetilde{P_s}. \tag{12.74}$$

In (12.74), we can easily check that $\widetilde{P}_l$ is a projection operator with $\alpha_1, \alpha_2, \ldots,$ and $\alpha_s$ being different eigenvalues of $A$. If $\alpha_l (1 \leq l \leq s)$ is degenerate, we express $\widetilde{P}_l$ as $\widetilde{P_l^{\mu}} (1 \leq \mu \leq m_l)$, where $m_l$ is multiplicity of $\alpha_l$. In that case, we may write

$$\widetilde{P}_l = \widetilde{P_l^1} \oplus \cdots \oplus \widetilde{P_l^{m_l}}. \tag{12.75}$$

Also we have

$$\widetilde{P_k}\widetilde{P}_l = U P_k U^{\dagger} U P_l U^{\dagger} = U P_k E P_l U^{\dagger} = U P_k P_l U^{\dagger} = 0 \, (1 \leq k, \; l \leq s).$$

The last equality comes from (12.23). Similarly, we have

$$\widetilde{P}_l \widetilde{P_k} = 0.$$

Thus, we have

$$\widetilde{P_k}\widetilde{P}_l = \delta_{kl}.$$

If the operator is decomposed as in the case of (12.75), we can express

$$\widetilde{P_l^{\mu}}\widetilde{P_l^{\nu}} = \delta_{\mu\nu}(1 \le \mu, \nu \le m_l).$$

Conversely, if an operator $A$ is expressed by (12.74), that operator is normal operator. In fact, we have

$$\begin{aligned} A^{\dagger}A &= \left(\sum_i \alpha_i \widetilde{P}_i\right)^{\dagger}\left(\sum_j \alpha_j \widetilde{P}_j\right) = \sum_{i,j} \alpha_i^* \alpha_j \widetilde{P}_i \widetilde{P}_j = \sum_{i,j} \alpha_i^* \alpha_j \delta_{ij} \widetilde{P}_i = \sum_i |\alpha_i|^2 \widetilde{P}_i, \\ AA^{\dagger} &= \left(\sum_j \alpha_j \widetilde{P}_j\right)\left(\sum_i \alpha_i \widetilde{P}_i\right)^{\dagger} = \sum_{i,j} \alpha_i^* \alpha_j \widetilde{P}_j \widetilde{P}_i = \sum_{i,j} \alpha_i^* \alpha_j \delta_{ji} \widetilde{P}_j = \sum_i |\alpha_i|^2 \widetilde{P}_i. \end{aligned} \tag{12.76}$$

Hence, $A^{\dagger}A = AA^{\dagger}$. If projection operators are further decomposed as in the case of (12.75), we have a related expression to (12.76). Thus, a necessary and sufficient condition for an operator to be a normal operator is that the said operator is expressed as (12.74). The relation (12.74) is well known as a spectral decomposition theorem. Thus, the spectral decomposition theorem is equivalent to Theorem 12.5.

The relations (10.208) and (12.74) are virtually the same, aside from the fact that whereas (12.74) premises an inner product space, (10.208) does not premise it. Correspondingly, whereas the related operators are called projection operators with the case of (12.74), those operators are said to be idempotent operators for (10.208).

*Example 12.1* Let us think of a Gram matrix of Example 11.1, as shown below.

$$G = \begin{pmatrix} 2 & 1+i \\ 1-i & 2 \end{pmatrix}. \tag{11.51}$$

After a unitary similarity transformation, we got

$$U^{\dagger}GU = \begin{pmatrix} 2+\sqrt{2} & 0 \\ 0 & 2-\sqrt{2} \end{pmatrix}. \tag{11.54}$$

Putting $\widetilde{G} = U^{\dagger}GU$ and rewriting (11.54), we have

$$\widetilde{G} = \begin{pmatrix} 2+\sqrt{2} & 0 \\ 0 & 2-\sqrt{2} \end{pmatrix} = \left(2+\sqrt{2}\right)\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \left(2-\sqrt{2}\right)\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

By a back calculation of $U\widetilde{G}U^{\dagger} = G$, we get

$$G = \left(2+\sqrt{2}\right)\begin{pmatrix} \frac{1}{2} & \frac{\sqrt{2}(1+i)}{4} \\ \frac{\sqrt{2}(1-i)}{4} & \frac{1}{2} \end{pmatrix} + \left(2-\sqrt{2}\right)\begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{2}(1+i)}{4} \\ -\frac{\sqrt{2}(1-i)}{4} & \frac{1}{2} \end{pmatrix}. \tag{12.77}$$

Putting eigenvalues $\alpha_1 = 2+\sqrt{2}$ and $\alpha_2 = 2-\sqrt{2}$ along with

$$A_1 = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{2}(1+i)}{4} \\ \frac{\sqrt{2}(1-i)}{4} & \frac{1}{2} \end{pmatrix}, \quad A_2 = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{2}(1+i)}{4} \\ -\frac{\sqrt{2}(1-i)}{4} & \frac{1}{2} \end{pmatrix}, \tag{12.78}$$

we get

$$G = \alpha_1 A_1 + \alpha_2 A_2. \tag{12.79}$$

In the above, $A_1$ and $A_2$ are projection operators. In fact, as anticipated we have

$$A_1^2 = A_1, A_2^2 = A_2, A_1A_2 = A_2A_1 = 0, A_1 + A_2 = E. \tag{12.80}$$

Moreover, (12.78) obviously shows that both $A_1$ and $A_2$ are Hermitian. Thus, Eqs. (12.77) and (12.79) are an example of the spectral decomposition. The decomposition is unique.

The Example 12.1 can be dealt with in parallel to Example 10.5. In Example 10.5, however, an inner product space is not implied, and so we used an idempotent matrix instead of a projection operator. Note that as can be seen in Example 10.5 that idempotent matrix was not Hermitian.

## 12.4 Hermitian Matrices and Unitary Matrices

Of normal matrices, Hermitian matrices and unitary matrices play a crucial role both in fundamental and applied science. Let us think of several topics and examples.

In quantum physics, one frequently treats expectation value of an operator. In general, such an operator is Hermitian, more strictly an observable. Moreover, a vector on an inner product space is interpreted as a state on a Hilbert space. Suppose that there is a linear operator (or observable that represents a physical quantity) $O$ that has discrete (or countable) eigenvalues $\alpha_1, \alpha_2, \ldots$ The number of the eigenvalues may be a finite number or an infinite number, but here we assume the finite number, i.e., let us suppose that we have eigenvalues $\alpha_1, \alpha_2, \ldots,$ and $\alpha_s$, in consistent with our previous discussion.

In quantum physics, we have a well-known Born probability rule. The rule says the following: Suppose that we carry out a physical measurement on $A$ with respect

to a physical state $|u\rangle$. Here we assume that $|u\rangle$ has been normalized. Then, the probability $\wp_l$ that $A$ takes $\alpha_l (1 \leq l \leq s)$ is given by

$$\wp_l = \left\| \widetilde{P}_l u \right\|^2, \tag{12.81}$$

where $\widetilde{P}_l$ is a projection operator that projects $|u\rangle$ to an eigenspace $W_{\alpha_l}$ spanned by $|\alpha_l, k\rangle$. Here $k (1 \leq k \leq n_l)$ reflects the multiplicity $n_l$ of an eigenvalue $\alpha_l$. Hence, we express the $n_l$-dimensional eigenspace $W_{\alpha_l}$ as

$$W_{\alpha_l} = \mathrm{Span}\{|\alpha_l, 1\rangle, |\alpha_l, 2\rangle, \ldots, |\alpha_l, n_l\rangle\}. \tag{12.82}$$

Now, we define an expectation value $\langle A \rangle$ of $A$ such that

$$\langle A \rangle \equiv \sum_{l=1}^{s} \alpha_l \wp_l. \tag{12.83}$$

From (12.81), we have

$$\wp_l = \left\| \widetilde{P}_l u \right\|^2 = \left\langle u \widetilde{P}_l{}^\dagger \,\middle|\, \widetilde{P}_l u \right\rangle = \left\langle u \widetilde{P}_l \,\middle|\, \widetilde{P}_l u \right\rangle = \left\langle u \widetilde{P}_l \widetilde{P}_l u \right\rangle = \left\langle u \widetilde{P}_l u \right\rangle. \tag{12.84}$$

For the third equality, we used the fact that $\widetilde{P}_l$ is Hermitian; for the last equality, we used $\widetilde{P}_l^2 = \widetilde{P}_l$. Summing (12.84) with the index $l$, we have

$$\sum_l \wp_l = \sum_l \langle u \widetilde{P}_l u \rangle = \left\langle u \left( \sum_l \widetilde{P}_l \right) u \right\rangle = \langle uEu \rangle = \langle u | u \rangle = 1,$$

where with the third equality we used (12.33).

Meanwhile, from the spectral decomposition theorem, we have

$$A = \alpha_1 \widetilde{P}_1 + \alpha_2 \widetilde{P}_2 + \cdots + \alpha_s \widetilde{P}_s. \tag{12.74}$$

Operating $\langle u|$ and $|u\rangle$ on both sides of (12.74), we get

$$\begin{aligned} \langle u | A | u \rangle &= \alpha_1 \langle u | \widetilde{P}_1 | u \rangle + \alpha_2 \langle u | \widetilde{P}_2 | u \rangle + \cdots + \alpha_s \langle u | \widetilde{P}_s | u \rangle \\ &= \alpha_1 \wp_1 + \alpha_2 \wp_2 + \cdots + \alpha_s \wp_s. \end{aligned} \tag{12.85}$$

Equating (12.83) and (12.85), we have

$$\langle A \rangle = \langle u | A | u \rangle. \tag{12.86}$$

In quantum physics, a real number is required for an expectation value of an observable (i.e., a physical quantity). To warrant this, we have following theorems.

**Theorem 12.9** *A linear transformation A on an inner product space is Hermitian if and only if* $\langle u|A|u\rangle$ *is real for all* $|u\rangle$ *of that inner product space.*

*Proof* If $A = A^\dagger$, then $\langle u|A|u\rangle^* = \left\langle u\middle|A^\dagger|u\right\rangle = \langle u|A|u\rangle$. Therefore, $\langle u|A|u\rangle$ is real. Conversely, if $\langle u|A|u\rangle$ is real for all $|u\rangle$, we have

$$\langle u|A|u\rangle = \langle u|A|u\rangle^* = \left\langle u\middle|A^\dagger|u\right\rangle.$$

Hence,

$$\langle u|A - A^\dagger|u\rangle = 0. \tag{12.87}$$

From Theorem 12.3, we get $A - A^\dagger = 0$, i.e., $A = A^\dagger$. This completes the proof.

**Theorem 12.10** *The eigenvalues of an Hermitian operator A are real.*

*Proof* Let $\alpha$ be an eigenvalue of $A$ and let $|u\rangle$ be a corresponding eigenvector. Then, $A|u\rangle = \alpha|u\rangle$. Operating $\langle u|$ from the left, $\langle u|A|u\rangle = \alpha\langle u|u\rangle = \alpha||u||$. Thus,

$$\alpha = \langle u|A|u\rangle/||u||. \tag{12.88}$$

Since $A$ is Hermitian, $\langle u|A|u\rangle$ is real. Then, $\alpha$ is real as well.

Unitary matrices have a following conspicuous features: (i) An inner product is held invariant under unitary transformation: Suppose that $|x'\rangle = U|x\rangle$ and $|y'\rangle = U|y\rangle$, where $U$ is a unitary operator. Then $\langle y'|x'\rangle = \left\langle yU^\dagger|Ux\right\rangle = \langle y|x\rangle$. A norm of any vector is held invariant under unitary transformation as well. This is easily checked by replacing $|y\rangle$ with $|x\rangle$ in the above. (ii) Let $U$ be a unitary matrix and suppose that $\lambda$ be an eigenvalue with $|\lambda\rangle$ being its corresponding eigenvector of that matrix. Then we have

$$|\lambda\rangle = U^\dagger U|\lambda\rangle = U^\dagger(\lambda|\lambda\rangle) = \lambda U^\dagger|\lambda\rangle = \lambda\lambda^*|\lambda\rangle, \tag{12.89}$$

where with the last equality we used Theorem 12.7. Thus

$$(1 - \lambda\lambda^*)|\lambda\rangle = 0. \tag{12.90}$$

As $|\lambda\rangle \neq 0$ is assumed, $1 - \lambda\lambda^* = 0$. That is

$$\lambda\lambda^* = |\lambda|^2 = 1. \tag{12.91}$$

Thus, eigenvalues of a unitary matrix have unit absolute value.

*Example 12.2* Let us think of a following unitary matrix $R$:

$$R = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}. \tag{12.92}$$

A characteristic equation is

$$\begin{vmatrix} \cos\theta - \lambda & -\sin\theta \\ \sin\theta & \cos\theta - \lambda \end{vmatrix} = \lambda^2 - 2\lambda\cos\theta + 1. \tag{12.93}$$

Solving (12.93), we have

$$\lambda = \cos\theta \pm \sqrt{\cos^2\theta - 1} = \cos\theta \pm i|\sin\theta|. \tag{12.94}$$

(i) $\theta = 0$: This is a trivial case. The matrix $R$ has automatically been diagonalized to be an identity matrix. Eigenvalues are 1 (double root).
(ii) $\theta = \pi$: This is a trivial case. Eigenvalues are $-1$ (double root).
(iii) $\theta \neq 0,\ \pi$: Let us think of $0 < \theta < \pi$. Then $\lambda = \cos\theta \pm i\sin\theta$. As a diagonalizing unitary matrix $U$, we get

$$U = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{i}{\sqrt{2}} & \frac{i}{\sqrt{2}} \end{pmatrix}, \quad U^\dagger = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{i}{\sqrt{2}} \end{pmatrix}. \tag{12.95}$$

Therefore, we have

$$U^\dagger R U = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{i}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{i}{\sqrt{2}} & \frac{i}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix}. \tag{12.96}$$

A trace of the resulting matrix is $2\cos\theta$. In the case of $\pi < \theta < 2\pi$, we get a diagonal matrix similarly. The conformation is left for readers.

## 12.5 Hermitian Quadratic Forms

The Hermitian quadratic forms appeared in, e.g., (11.34) and (11. 83) in relation to Gram matrices in Sect. 11.2. The Hermitian quadratic forms have wide applications in the field of mathematical physics and materials science.

Let $H$ be an Hermitian operator and $|x\rangle$ on an inner vector space. We define the Hermitian quadratic form in an arbitrary orthonormal basis as follows:

$$\langle x|H|x\rangle = \left(x_1^* \cdots x_n^*\right) H \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \sum_{i,j} x_i^* (H)_{ij} x_j,$$

where $|x\rangle$ is represented as a column vector, as already mentioned in Sect. 11.4. Let us start with unitary diagonalization of (11.40), where a Gram matrix is a kind of

Hermitian matrix. Following similar procedures, as in the case of (11.36) we obtain a diagonal matrix and an inner product such that

$$\langle x|H|x\rangle = \begin{pmatrix}\xi_1^* \cdots \xi_n^*\end{pmatrix}\begin{pmatrix}\lambda_1 & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & \lambda_n\end{pmatrix}\begin{pmatrix}\xi_1\\ \vdots\\ \xi_n\end{pmatrix} = \lambda_1|\xi_1|^2 + \cdots + \lambda_n|\xi_n|^2. \tag{12.97}$$

Notice, however, that the Gram matrix comprising basis vectors (that are linearly independent) is positive definite. Remember that if a Gram matrix is constructed by $A^\dagger A$ (Hermitian as well) according to whether $A$ is non-singular or singular, $A^\dagger A$ is either positive definite or nonnegative. The Hermitian matrix we are dealing with here, in general, does not necessarily possess the positive definiteness or nonnegative feature. Yet, remember that $\langle x|H|x\rangle$ and eigenvalues $\lambda_1, \ldots, \lambda_n$ are real.

Positive definiteness of matrices is an important concept in relation to the Hermitian (and real) quadratic forms (see Sect. 11.2). In particular, in the case where all the matrix elements are real and $|x\rangle$ is defined in a real domain, we are dealing with the quadratic form with respect to a real symmetric matrix. In the case of the real quadratic forms, we sometimes adopt the following notation:

$$A[\boldsymbol{x}] \equiv \boldsymbol{x}^T A\boldsymbol{x} = \sum_{i,j=1}^{n} a_{ij}x_i x_j;\ \boldsymbol{x} = \begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix},$$

where $A = \left(a_{ij}\right)$ is a real symmetric matrix and $x_i (1 \le i \le n)$ are real numbers. The positive definiteness is invariant under a transformation $P^T AP$, where $P$ is non-singular. In fact, if $A > 0$, for $\boldsymbol{x}'^T = \boldsymbol{x}^T P$ and $A' = P^T AP$ we have

$$\boldsymbol{x}^T A\boldsymbol{x} = \boldsymbol{x}^T P\left(P^T AP\right)P^T\boldsymbol{x} = \boldsymbol{x}'^T A'\boldsymbol{x}'.$$

Since $P$ is non-singular, $P^T\boldsymbol{x} = \boldsymbol{x}'$ represents any arbitrary vector. Hence,

$$A' = P^T AP > 0, \tag{12.98}$$

where with the notation $P^T AP > 0$, see (11.46). In particular, using a suitable orthogonal matrix $O$, we obtain

$$O^T AO = \begin{pmatrix}\lambda_1 & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & \lambda_n\end{pmatrix}.$$

From the above argument, $O^T AO > 0$. We have

$$\det O^T AO = \det O^T \det A \det O = (\pm 1) \det A (\pm 1) = \det A.$$

Therefore, from (12.97), we have $\lambda_i > 0\,(1 \leq i \leq n)$. Thus, we get

$$\det A = \prod_{i=1}^{n} \lambda_i > 0.$$

Notice that in the above discussion, $P^T AP$ is said to be an equivalent transformation of $A$ by $P$ and that we distinguish the equivalent transformation from the similarity transformation $P^{-1}AP$. Nonetheless, if we choose an orthogonal matrix $O$ for $P$, the two transformations are the same because $O^T = O^{-1}$.

We often encounter real quadratic forms in the field of electromagnetism. Typical example is a trace of electromagnetic fields observed with an elliptically or circularly polarized light (see Sect. 5.3). A polarizability tensor of an anisotropic media such as crystals (either inorganic or organic) is another example, even though we did not treat it but only assumed isotropic media in Part II.

Regarding the real quadratic forms, we have a following important theorem.

**Theorem 12.11 [4]** *Let $A$ be a $n$-dimensional real symmetric matrix $A = (a_{ij})$. Let $A^{(k)}$ be $k$-dimensional principal submatrices described by*

$$A^{(k)} = \begin{pmatrix} a_{i_1 i_1} & a_{i_1 i_2} & \cdots & a_{i_1 i_k} \\ a_{i_2 i_1} & a_{i_2 i_2} & \cdots & a_{i_2 i_k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i_k i_1} & a_{i_k i_2} & \cdots & a_{i_k i_k} \end{pmatrix} (1 \leq i_1 < i_2 < \cdots < i_k \leq n),$$

where the principal submatrices mean a matrix made by striking out rows and columns on diagonal elements. Alternatively, a principal submatrix of a matrix $A$ can be defined as a matrix whose diagonal elements are a part of the diagonal elements of $A$. As a special case, those include $a_{11}, \ldots,$ or $a_{nn}$ as a (1,1) matrix (i.e., merely a number) as well as $A$ itself. Then, we have

$$A > 0 \Leftrightarrow \det A^{(k)} > 0\ (1 \leq k \leq n). \tag{12.99}$$

*Proof* First, suppose that $A > 0$. Then, in a quadratic form $A[\boldsymbol{x}]$ equating $(n-k)$ variables $x_l = 0 (l \neq i_1, i_2, \ldots, i_k)$, we obtain a quadratic form of

$$\sum_{\mu,\nu=1}^{k} a_{i_\mu i_\nu} x_{i_\mu} x_{i_\nu}.$$

Since $A > 0$, this (partial) quadratic form is positive definite as well, i.e., we have

$$A^{(k)} > 0.$$

Therefore, we get

$$\det A^{(k)} > 0.$$

This is due to (11.48). Notice that $\det A^{(k)}$ is said to be a principal minor. Thus, we have proven $\Rightarrow$ of (12.99).

To prove $\Leftarrow$, in turn, we use mathematical induction. If $n = 1$, we have a trivial case, i.e., $A$ is merely a real positive number. Suppose that $\Leftarrow$ is true of $n - 1$. Then, we have $A^{(n-1)} > 0$ by supposition. Thus, it follows that it will suffice to show $A > 0$ on condition that $A^{(n-1)} > 0$ and $\det A > 0$ in addition. Let $A$ be a $n$-dimensional real symmetric non-singular matrix such that

$$A = \begin{pmatrix} A^{(n-1)} & \boldsymbol{a} \\ \boldsymbol{a}^T & a_n \end{pmatrix},$$

where $A^{(n-1)}$ is a symmetric matrix and non-singular as well. We define $P$ such that

$$P = \begin{pmatrix} E & A^{(n-1)^{-1}}\boldsymbol{a} \\ \boldsymbol{0} & 1 \end{pmatrix},$$

where $E$ is a $(n-1, n-1)$ unit matrix. Notice that $\det P = \det E \cdot 1 = 1 \neq 0$, indicating that $P$ is non-singular. We have

$$P^T = \begin{pmatrix} E & \boldsymbol{0} \\ \boldsymbol{a}^T A^{(n-1)^{-1}} & 1 \end{pmatrix}.$$

For this expression, consider a non-singular matrix $S$. Then, we have $SS^{-1} = E$. Taking its transposition, we have $(S^{-1})^T S^T = E$. Therefore, if $S$ is a symmetric matrix $(S^{-1})^T S = E$, i.e., $(S^{-1})^T = S^{-1}$. Hence, an inverse matrix of a symmetric matrix is symmetric as well. Then, for a symmetric matrix $A^{(n-1)}$ we have

$$\left(\boldsymbol{a}^T A^{(n-1)^{-1}}\right)^T = \left(A^{(n-1)^{-1}}\right)^T \left(\boldsymbol{a}^T\right)^T = A^{(n-1)^{-1}}\boldsymbol{a}.$$

Therefore, $A$ can be expressed as

$$\begin{aligned} A &= P^T \begin{pmatrix} A^{(n-1)} & \boldsymbol{0} \\ \boldsymbol{0} & a_n - A^{(n-1)^{-1}}[\boldsymbol{a}] \end{pmatrix} P. \\ &= \begin{pmatrix} E & \boldsymbol{0} \\ \boldsymbol{a}^T A^{(n-1)^{-1}} & 1 \end{pmatrix} \begin{pmatrix} A^{(n-1)} & \boldsymbol{0} \\ \boldsymbol{0} & a_n - A^{(n-1)^{-1}}[\boldsymbol{a}] \end{pmatrix} \begin{pmatrix} E & A^{(n-1)^{-1}}\boldsymbol{a} \\ \boldsymbol{0} & 1 \end{pmatrix}. \end{aligned} \tag{12.100}$$

Now, taking a determinant of (12.100), we have

$$\begin{aligned}\det A &= \det P^T \left[\det A^{(n-1)}\right] \left\{a_n - A^{(n-1)^{-1}}[\boldsymbol{a}]\right\} \det P \\ &= \left[\det A^{(n-1)}\right] \left\{a_n - A^{(n-1)^{-1}}[\boldsymbol{a}]\right\}.\end{aligned}$$

By supposition, we have $\det A^{(n-1)} > 0$ and $\det A > 0$. Hence, we have

$$a_n - A^{(n-1)^{-1}}[\boldsymbol{a}] > 0.$$

Putting $\widetilde{a_n} \equiv a_n - A^{(n-1)^{-1}}[\boldsymbol{a}]$ and $\boldsymbol{x} \equiv \begin{pmatrix} \boldsymbol{x}^{(n-1)} \\ x_n \end{pmatrix}$, we get

$$\begin{pmatrix} A^{(n-1)} & \boldsymbol{0} \\ \boldsymbol{0} & \widetilde{a_n} \end{pmatrix}[\boldsymbol{x}] = A^{(n-1)}\left[\boldsymbol{x}^{(n-1)}\right] + \widetilde{a_n} x_n^2.$$

Since $A^{(n-1)} > 0$ and $\widetilde{a_n} > 0$, we have

$$\widetilde{A} \equiv \begin{pmatrix} A^{(n-1)} & \boldsymbol{0} \\ \boldsymbol{0} & \widetilde{a_n} \end{pmatrix} > 0.$$

Meanwhile, $A$ is expressed as

$$A = P^T \widetilde{A} P.$$

From (12.98), $A > 0$. These complete the proof.

We also have a related theorem (the following Theorem 12.12) for an Hermitian quadratic form.

**Theorem 12.12** *Let $A = \left(a_{ij}\right)$ be a n-dimensional Hermitian matrix. Let $\widetilde{A^{(k)}}$ be k-dimensional principal submatrices. Then, we have*

$$A > 0 \Leftrightarrow \det \widetilde{A^{(k)}} > 0 \ (1 \leq k \leq n),$$

*where $\widetilde{A^{(k)}}$ is described as*

$$\widetilde{A^{(k)}} = \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{kk} \end{pmatrix}.$$

*The proof is left for readers.*

*Example 12.3* Let us consider a following Hermitian (real symmetric) matrix and corresponding Hermitian quadratic form.

$$H = \begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix}, \quad \langle x|H|x\rangle = (x_1 \quad x_2)\begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}. \tag{12.101}$$

Principal minors of $H$ are $5(>0)$ and $1(>0)$ and $\det H = 5-4 = 1(>0)$. Therefore, from Theorem 12.11 we have $H > 0$. A characteristic equation gives following eigenvalues, i.e.,

$$\lambda_1 = 3+2\sqrt{2}, \quad \lambda_2 = 3-2\sqrt{2}.$$

Both the eigenvalues are positive as anticipated. As a diagonalizing matrix $R$, we get

$$R = \begin{pmatrix} 1+\sqrt{2} & 1-\sqrt{2} \\ 1 & 1 \end{pmatrix}.$$

To obtain a unitary matrix, we have to seek norms of column vectors. Corresponding to $\lambda_1$ and $\lambda_2$, we estimate their norms to be $\sqrt{4+2\sqrt{2}}$ and $\sqrt{4-2\sqrt{2}}$, respectively. Using them, as a unitary matrix $U$ we get

$$U = \begin{pmatrix} \frac{1}{\sqrt{4-2\sqrt{2}}} & -\frac{1}{\sqrt{4+2\sqrt{2}}} \\ \frac{1}{\sqrt{4+2\sqrt{2}}} & \frac{1}{\sqrt{4-2\sqrt{2}}} \end{pmatrix}. \tag{12.102}$$

Thus, performing the matrix diagonalization, we obtain a diagonal matrix $D$ such that

$$D = U^\dagger H U = \begin{pmatrix} 3+2\sqrt{2} & 0 \\ 0 & 3-2\sqrt{2} \end{pmatrix}. \tag{12.103}$$

Let us view the above unitary diagonalization in terms of coordinate transformation. Using the above matrix $U$ and changing (12.101) as in (11.36),

$$\begin{aligned} \langle x|H|x\rangle &= (x_1x_2)UU^\dagger\begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix}UU^\dagger\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ &= (x_1x_2)U\begin{pmatrix} 3+2\sqrt{2} & 0 \\ 0 & 3-2\sqrt{2} \end{pmatrix}U^\dagger\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}. \end{aligned} \tag{12.104}$$

Making an argument analogous to that of Sect. 11.2 and using similar notation, we have

$$\begin{aligned} \langle \tilde{x}|H'|\tilde{x}\rangle &= (\tilde{x_1} \ \tilde{x_2})H'\begin{pmatrix} \tilde{x_1} \\ \tilde{x_2} \end{pmatrix} = (x_1x_2)UU^\dagger HUU^\dagger\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ &= \langle x|H|x\rangle. \end{aligned} \tag{12.105}$$

That is, we have

$$\begin{pmatrix} \widetilde{x_1} \\ \widetilde{x_2} \end{pmatrix} = U^\dagger \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}. \tag{12.106}$$

Or, taking an adjoint of (12.106), we have equivalently $(\,\widetilde{x_1} \quad \widetilde{x_2}\,) = (\,x_1 \quad x_2\,)U$. Notice here that $\begin{pmatrix} \widetilde{x_1} \\ \widetilde{x_2} \end{pmatrix}$ and $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ are real and that $U$ is real (i.e., orthogonal matrix). Likewise,

$$H' = U^\dagger H U. \tag{12.107}$$

Thus, it follows that $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ and $\begin{pmatrix} \widetilde{x_1} \\ \widetilde{x_2} \end{pmatrix}$ are *different* column vector representations of the *identical* vector that is viewed in reference to two different sets of orthonormal bases (i.e., different coordinate systems).

From (12.102), as an approximation we have

$$U \cong \begin{pmatrix} 0.92 & -0.38 \\ 0.38 & 0.92 \end{pmatrix} \cong \begin{pmatrix} \cos 22.5^\circ & -\sin 22.5^\circ \\ \sin 22.5^\circ & \cos 22.5^\circ \end{pmatrix}. \tag{12.108}$$

Equating $\langle x|H|x\rangle$ to a constant, we get a hypersurface in a plane. Choosing 1 for a constant, we get an equation of hypersurface (i.e., an ellipse) as a function of $\widetilde{x_1}$ and $\widetilde{x_2}$ such that

$$\frac{\widetilde{x_1}^2}{\left(\sqrt{\frac{1}{3+2\sqrt{2}}}\right)^2} + \frac{\widetilde{x_2}^2}{\left(\sqrt{\frac{1}{3-2\sqrt{2}}}\right)^2} = 1. \tag{12.109}$$

Figure 12.1 depicts the ellipse.

## 12.6 Simultaneous Eigenstates and Diagonalization

In quantum physics, a concept of simultaneous eigenstate is important and has briefly mentioned in Sect. 3.3. To rephrase this concept, suppose that there are two operators $B$ and $C$ and ask whether the two operators (or more) possess a common set of eigenvectors. The question is boiled down to whether the two operators commute. To address this question, the following theorem is important.

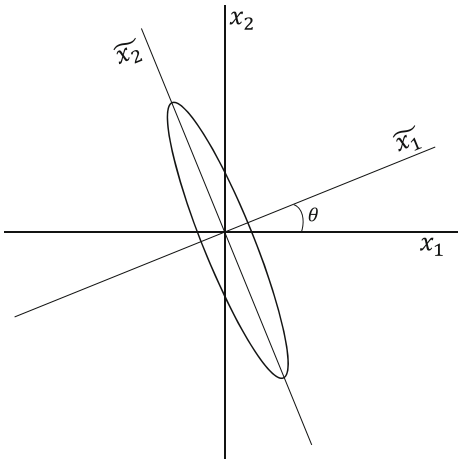


**Fig. 12.1** Ellipse obtained through diagonalization of a Hermitian quadratic form. The angle $\theta$ is about 22.5°

**Theorem 12.13 [2]** *Two Hermitian matrices B and C commute if and only if there exists a complete orthonormal set of common eigenvectors.*

*Proof* Suppose that there exists a complete orthonormal set of common eigenvectors $\{|x_i;\ m_i\rangle\}$ that span the linear vector space, where $i$ and $m_i$ are a positive integer and that $|x_i\rangle$ corresponds to an eigenvalue $b_i$ of $B$ and $c_i$ of $C$. Note that if $m_i$ is 1, we say that the spectrum is non-degenerate and that if $m_i$ is equal to two or more, the spectrum is said to be degenerate. Then we have

$$B(|x_i\rangle) = b_i|x_i\rangle,\ C(|x_i\rangle) = c_i|x_i\rangle. \tag{12.110}$$

Therefore, $BC(|x_i\rangle) = B(c_i|x_i\rangle) = c_iB(|x_i\rangle) = c_ib_i|x_i\rangle$. Similarly, $CB(|x_i\rangle) = c_ib_i|x_i\rangle$. Consequently, $(BC - CB)(|x_i\rangle) = 0$ for any $|x_i\rangle$. As all the set of $|x_i\rangle$ span the vector space, $BC - CB = 0$, namely $BC = CB$.

In turn, assume that $BC = CB$ and that $B(|x_i\rangle) = b_i|x_i\rangle$, where $|x_i\rangle$ are orthonormal. Then, we have

$$CB(|x_i\rangle) = b_iC(|x_i\rangle) \Rightarrow B[C(|x_i\rangle)] = b_iC(|x_i\rangle). \tag{12.111}$$

This implies that $C(|x_i\rangle)$ is an eigenvector of $B$ corresponding to the eigenvalue $b_i$. We have two cases.

(i) The spectrum is non-degenerate: The spectrum is said to be non-degenerate if only one eigenvector belongs to an eigenvalue. In other words, multiplicity of $b_i$ is one. Then, $C(|x_i\rangle)$ must be equal to some constant times $|x_i\rangle$, i.e., $C(|x_i\rangle) = c_i|x_i\rangle$. That is, $|x_i\rangle$ is an eigenvector of $C$ corresponding to an eigenvalue $c_i$. That is, $|x_i\rangle$ is a common eigenvector to $B$ and $C$. This completes the proof.

(ii) The spectrum is degenerate: The spectrum is said to be degenerate if two or more eigenvectors belong to an eigenvalue. Multiplicity of $b_i$ is two or more; here suppose that the multiplicity is $m$ $(m \geq 2)$. Let $\left|x_i^{(1)}\right\rangle, \left|x_i^{(2)}\right\rangle, \ldots,$ and $\left|x_i^{(m)}\right\rangle$ be linearly independent vectors and belong to the eigenvalue $b_i$ of $B$. Then, from the assumption, we have $m$ eigenvectors

$$C\left(\left|x_i^{(1)}\right\rangle\right), C\left(\left|x_i^{(2)}\right\rangle\right), \ldots, C\left(\left|x_i^{(m)}\right\rangle\right)$$

that belong to an eigenvalue $b_i$ of $B$. This means that individual $C\left(\left|x_i^{(\mu)}\right\rangle\right)(1 \leq \mu \leq m)$ are described by linear combination of $\left|x_i^{(1)}\right\rangle, \left|x_i^{(2)}\right\rangle, \ldots,$ and $\left|x_i^{(m)}\right\rangle$. What we want to prove is to show that suitable linear combination of these $m$ vectors constitutes an eigenvector corresponding to an eigenvalue $c_\mu$ of $C$. Here, to avoid complexity, we denote the multiplicity by $m$ instead of the above-mentioned $m_i$.

The vectors $C\left(\left|x_i^{(\mu)}\right\rangle\right)(1 \leq \mu \leq m)$ can be described as

$$C\left(\left|x_i^{(1)}\right\rangle\right) = \sum_{j=1}^{m} \alpha_{j1}\left|x_i^{(j)}\right\rangle, \ldots, C\left(\left|x_i^{(m)}\right\rangle\right) = \sum_{j=1}^{m} \alpha_{jm}\left|x_i^{(j)}\right\rangle. \tag{12.112}$$

Using full matrix representation, we have

$$\begin{aligned} C\left(\sum_{k=1}^{m} \gamma_k \left|x_i^{(k)}\right\rangle\right) &= \left(\left|x_i^{(1)}\right\rangle \cdots \left|x_i^{(m)}\right\rangle\right) C \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_m \end{pmatrix} \\ &= \left(\left|x_i^{(1)}\right\rangle \cdots \left|x_i^{(m)}\right\rangle\right) \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1m} \\ \vdots & \ddots & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mm} \end{pmatrix} \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_m \end{pmatrix}. \end{aligned} \tag{12.113}$$

In (12.113), we adopt the notation of (9.37). Since $(\alpha_{ij})$ is a matrix representation of an Hermitian operator $C$, $(\alpha_{ij})$ is Hermitian as well. More specifically, if we take an inner product of a vector expressed in (12.112) with $\left|x_i^{(l)}\right\rangle$, then we have

$$\begin{aligned} \left\langle x_i^{(l)} \middle| C x_i^{(k)} \right\rangle &= \left\langle x_i^{(l)} \middle| \sum_{j=1}^{m} \alpha_{jk} x_i^{(j)} \right\rangle = \sum_{j=1}^{m} \alpha_{jk} \left\langle x_i^{(l)} \middle| x_i^{(j)} \right\rangle = \sum_{j=1}^{m} \alpha_{jk} \delta_{lj} \\ &= \alpha_{lk} (1 \leq k, l \leq m), \end{aligned} \tag{12.114}$$

where the third equality comes from the orthonormality of the basis vectors. Meanwhile,

$$\left\langle x_i^{(l)} \middle| C x_i^{(k)} \right\rangle = \left\langle x_i^{(k)} \middle| C^\dagger x_i^{(l)} \right\rangle^* = \left\langle x_i^{(k)} \middle| C x_i^{(l)} \right\rangle^* = \alpha_{kl}^*, \tag{12.115}$$

where the second equality comes from the Hermiticity of $C$. From (12.114) and (12.115), we get

$$\alpha_{lk} = \alpha_{kl}^* (1 \le k,\ l \le m). \tag{12.116}$$

This indicates that $(\alpha_{ij})$ is in fact Hermitian.

We are seeking the condition under which linear combinations of the eigenvectors $\left| x_i^{(k)} \right\rangle (1 \le k \le m)$ for $B$ are simultaneously eigenvectors of $C$. If the linear combination $\sum_{k=1}^{m} \gamma_k \left| x_i^{(k)} \right\rangle$ is to be an eigenvector of $C$, we must have

$$C\left( \sum_{k=1}^{m} \gamma_k \left| x_i^{(k)} \right\rangle \right) = c\left( \sum_{j=1}^{m} \gamma_j \left| x_i^{(j)} \right\rangle \right). \tag{12.117}$$

Considering (12.113), we have

$$\left( \left| x_i^{(1)} \right\rangle \cdots \left| x_i^{(m)} \right\rangle \right) \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1m} \\ \vdots & \ddots & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mm} \end{pmatrix} \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_m \end{pmatrix} = \left( \left| x_i^{(1)} \right\rangle \cdots \left| x_i^{(m)} \right\rangle \right) c \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_m \end{pmatrix}. \tag{12.118}$$

The vectors $\left| x_i^{(j)} \right\rangle (1 \le k \le m)$ span an invariant subspace (i.e., an eigenspace corresponding to an eigenvalue of $b_i$). Let us call this subspace $W^m$. Consequently, in (12.118), we can equate the scalar coefficients of individual $\left| x_i^{(j)} \right\rangle (1 \le k \le m)$. Then, we get

$$\begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1m} \\ \vdots & \ddots & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mm} \end{pmatrix} \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_m \end{pmatrix} = c \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_m \end{pmatrix}. \tag{12.119}$$

This is nothing other than an eigenvalue equation. Since $(\alpha_{ij})$ is an Hermitian matrix, there should be $m$ eigenvalues $c_\mu$ some of which may be identical (the degenerate case). Moreover, we can always decide $m$ orthonormal column vectors by solving (12.119). We denote them by $\gamma^{(\mu)} (1 \le \mu \le m)$ that belong to $c_\mu$. Rewriting (12.119), we get

$$\begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1m} \\ \vdots & \ddots & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mm} \end{pmatrix} \begin{pmatrix} \gamma_1^{(\mu)} \\ \vdots \\ \gamma_m^{(\mu)} \end{pmatrix} = c_\mu \begin{pmatrix} \gamma_1^{(\mu)} \\ \vdots \\ \gamma_m^{(\mu)} \end{pmatrix}. \tag{12.120}$$

Equation (12.120) implies that we can construct a $(m,m)$ unitary matrix from the $m$ orthonormal column vectors $\gamma^{(\mu)}$. Using the said unitary matrix, we will be able to diagonalize $(\alpha_{ij})$ according to Theorem 12.5.

Having determined $m$ eigenvectors $\gamma^{(\mu)}(1 \le \mu \le m)$, we can construct a set of eigenvectors such that

$$\left| y_i^{(\mu)} \right\rangle = \sum_{k=1}^{m} \gamma_k^{(\mu)} \left| x_i^{(k)} \right\rangle. \tag{12.121}$$

Finally let us confirm that $\left| y_i^{(\mu)} \right\rangle (1 \le \mu \le m)$ in fact constitute an orthonormal basis. To show this, we have

$$\begin{aligned} \left\langle y_i^{(\nu)} \middle| y_i^{(\mu)} \right\rangle &= \left\langle \sum_{k=1}^{m} \gamma_k^{(\nu)} x_i^{(k)} \middle| \sum_{l=1}^{m} \gamma_l^{(\mu)} x_i^{(l)} \right\rangle \\ &= \sum_{k=1}^{m} \sum_{l=1}^{m} \left[ \gamma_k^{(\nu)} \right]^* \gamma_l^{(\mu)} \left\langle x_i^{(k)} \middle| x_i^{(l)} \right\rangle = \sum_{k=1}^{m} \sum_{l=1}^{m} \left[ \gamma_k^{(\nu)} \right]^* \gamma_l^{(\mu)} \delta_{kl} \\ &= \sum_{k=1}^{m} \left[ \gamma_k^{(\nu)} \right]^* \gamma_k^{(\mu)} = \delta_{\nu\mu}. \end{aligned} \tag{12.122}$$

The last equality comes from the fact that a matrix comprising $m$ orthonormal column vectors $\gamma^{(\mu)}(1 \le \mu \le m)$ forms a unitary matrix. Thus, $\left| y_i^{(\mu)} \right\rangle (1 \le \mu \le m)$ certainly constitute an orthonormal basis.

The above completes the proof.

Theorem 12.13 can be restated as follows: Two Hermitian matrices $B$ and $C$ can be simultaneously diagonalized by a unitary similarity transformation. As mentioned above, we can construct a unitary matrix $U$ such that

$$U = \begin{pmatrix} \gamma_1^{(1)} & \cdots & \gamma_1^{(m)} \\ \vdots & \ddots & \vdots \\ \gamma_m^{(1)} & \cdots & \gamma_m^{(m)} \end{pmatrix}.$$

Then, using $U$, matrices $B$ and $C$ are diagonalized such that

$$U^\dagger BU = \begin{pmatrix} b_i & & & \\ & b_i & & \\ & & \ddots & \\ & & & b_i \end{pmatrix}, \quad U^\dagger CU = \begin{pmatrix} c_1 & & & \\ & c_2 & & \\ & & \ddots & \\ & & & c_m \end{pmatrix}. \tag{12.123}$$

Note that both $B$ and $C$ are represented in an invariant subspace $W^m$.

As we have already seen in Part I that dealt with an eigenvalue problem of a hydrogen-like atom, squared angular momentum ($\boldsymbol{L}^2$) and $z$-component of angular momentum ($L_z$) possess a mutual set of eigenvectors and, hence, their eigenvalues are determined at once. Related matrix representations to (12.123) were given in (3.159). On the other hand, this was not the case with a set of operators $L_x$, and $L_y$, and $L_z$; see (3.30). Yet, we pointed out the exceptional case where these three operators along with $\boldsymbol{L}^2$ take an eigenvalue zero in common which an eigenstate $Y_0^0(\theta, \phi) \equiv \sqrt{1/4\pi}$ corresponds to. Nonetheless, no *complete* orthonormal set of common eigenvectors exists with the set of operators $L_x$, and $L_y$, and $L_z$. This fact is equivalent to that these three operators are non-commutative among them. In contrast, $\boldsymbol{L}^2$ and $L_z$ share a complete orthonormal set of common eigenvectors and, hence, are commutable.

Notice that $C\left(\left|x_i^{(1)}\right\rangle\right), C\left(\left|x_i^{(2)}\right\rangle\right), \ldots$, and $\left(C\left|x_i^{(m)}\right\rangle\right)$ are not necessarily linearly independent (see Sect. 9.4). Suppose that among $m$ eigenvalues $c_\mu$ $(1 \le \mu \le m)$, some $c_\mu = 0$. Then, $\det C = 0$ according to (11.48). This means that $C$ is singular. In that case, $C\left(\left|x_i^{(1)}\right\rangle\right), C\left(\left|x_i^{(2)}\right\rangle\right), \ldots$, and $\left(C\left|x_i^{(m)}\right\rangle\right)$ are linearly dependent. In Sect. 3.3, in fact, we had $\left|L_z Y_0^0(\theta, \phi)\right\rangle = \left|\boldsymbol{L}^2 Y_0^0(\theta, \phi)\right\rangle = 0$. But, this special situation does not affect the proof of Theorem 12.13.

We know that any matrix $A$ can be decomposed such that

$$A = \frac{1}{2}\left(A + A^\dagger\right) + i\left[\frac{1}{2i}\left(A - A^\dagger\right)\right], \tag{12.124}$$

where we put $B \equiv \frac{1}{2}\left(A + A^\dagger\right)$ and $C \equiv \frac{1}{2i}\left(A - A^\dagger\right)$; both $B$ and $C$ are Hermitian. That is any matrix $A$ can be decomposed to two Hermitian matrices in such a way that

$$A = B + iC. \tag{12.125}$$

Note here that $B$ and $C$ commute if and only if $A$ and $A^\dagger$ commute, that is $A$ is a normal matrix. In fact, from (12.125) we get

$$AA^\dagger - A^\dagger A = 2i(CB - BC). \tag{12.126}$$

From (12.126), if and only if $B$ and $C$ commute (i.e., $B$ and $C$ can be diagonalized simultaneously), $AA^\dagger - A^\dagger A = 0$, i.e., $AA^\dagger = A^\dagger A$. This indicates that $A$ is a normal matrix. Thus, the following theorem will follow.

**Theorem 12.14** *A matrix can be diagonalized by a unitary similarity transformation, if and only if it is a normal matrix.*

Thus, Theorem 12.13 is naturally generalized so that it can be stated as follows: Two normal matrices $B$ and $C$ commute if and only if there exists a complete orthonormal set of common eigenvectors.

In Sects. 12.1 and 12.3, we mentioned the spectral decomposition. There, we showed a special case where projection operators commute with one another; see (12.23). Thus, in light of Theorem 12.13, those projection operators can be diagonalized at once to be expressed as, e.g., (12.72). This is a conspicuous feature of the projection operators.

## References

1. Hassani S (2006) Mathematical physics. Springer, New York
2. Byron FW Jr, Fuller RW (1992) Mathematics of classical and quantum physics. Dover, New York
3. Mirsky L (1990) An Introduction to linear algebra. Dover, New York
4. Satake I (1974) Linear algebra. Shokabo, Tokyo (in Japanese)

# Part IV
# Group Theory and Its Chemical Applications

Universe comprises space and matter. These two mutually stipulate their modality of existence. We often comprehend related various aspects as manifestation of symmetry. In this part, we deal with the symmetry from a point of view of group theory. In this last part of the book, we outline and emphasize chemical applications of the methods of mathematical physics. This part supplies us with introductory description of group theory. The group theory forms an important field of both pure and applied mathematics. Starting with the definition of groups, we cover a variety of topics related to group theory. Of these, symmetry groups are familiar to chemists, because they deal with a variety of matter and molecules that are characterized by different types of symmetry. The symmetry group is a kind of finite group and called a point group as well. Meanwhile, we have various infinite groups that include rotation group as a typical example. We also mention an introductory theory of the rotation group of SO(3) that deals with an important topic of, e.g., Euler angles. We also treat successive coordinate transformations.

Next, we describe representation theory of groups. Schur's lemmas and related grand orthogonality theorem underlie the representation theory of groups. In parallel, characters and irreducible representations are important concepts that support the representation theory. We present various representations, e.g., regular representation, direct-product representation, and symmetric and antisymmetric representations. These have wide applications in the field of quantum mechanics and quantum chemistry, and so forth.

In the final chapter, we highlight quantum chemical applications of group theory in relation to a method of molecular orbitals. As tangible examples, we adopt aromatic molecules and methane.

# Chapter 13
# Introductory Group Theory

A group comprises mathematical elements that satisfy four simple definitions. A bunch of groups exists under these simple definitions. This makes the group theory a discriminating field of mathematics. To get familiar with various concepts of groups, we first show several tangible examples. Group elements can be numbers (both real and complex) and matrices. More abstract mathematical elements can be included as well. Examples include transformation, operation as already studied in previous parts. Once those mathematical elements form a group, they share several common notions such as classes, subgroups, and direct-product groups. In this context, readers are encouraged to conceive different kinds of groups close to their heart. Mapping is an important concept as in the case of vector spaces. In particular, isomorphism and homomorphism frequently appear in the group theory. These concepts are closely related to the representation theory that is an important pillar of the group theory.

## 13.1 Definition of Groups

In contrast to a broad range of applications, the definition of the group is simple. Let $\mathcal{G}$ be a set of elements $g_v$, where $v$ is an index either countable (e.g., integers) or uncountable (e.g., real numbers) and the number of elements may be finite or infinite. We denote this by $\mathcal{G} = \{g_v\}$. If a group is a finite group, we express it as

$$\mathcal{G} = \{g_1, g_2, \ldots, g_n\}, \tag{13.1}$$

where $n$ is said to be an order of the group.

Definition of the group comprises the following four axioms with respect to a well-defined "multiplication" rule between any pair of elements. The multiplication is denoted by a symbol $''\diamond''$ below. Note that the symbol $\diamond$ implies an ordinary multiplication, an ordinary addition, etc.



S. Hotta, *Mathematical Physical Chemistry*,
https://doi.org/10.1007/978-981-10-7671-8_13

(A1) If $a$ and $b$ are any elements of the set $\mathcal{g}$, then so is $a \diamond b$. (We sometimes say that the set is "closed" regarding the multiplication.)
(A2) Multiplication is associative; i.e., $a \diamond (b \diamond c) = (a \diamond b) \diamond c$.
(A3) The set $\mathcal{g}$ contains an element of $e$ called the identity element such that we have $a \diamond e = e \diamond a = a$ with any element $a$ of $\mathcal{g}$.
(A4) For any $a$ of $\mathcal{g}$, we have an element $b$ such that $a \diamond b = b \diamond a = e$. The element $b$ is said to be the inverse element of $a$. We denote $b \equiv a^{-1}$.

In the above definitions, we assume that the commutative law does not necessarily hold, that is, $a \diamond b \neq b \diamond a$. In that case, the group $\mathcal{g}$ is said to be a non-commutative group. However, we have a case where the commutative law holds, i.e., $a \diamond b = b \diamond a$. If so, the group $\mathcal{g}$ is called a commutative group or an Abelian group.

Let us think of some examples of groups. Henceforth, we follow the convention and write $ab$ to express $a \diamond b$.

*Example 13.1* We present several examples of groups below. Examples (i) to (iv) are simple, but Example (v) is general.

(i) $\mathcal{g} = \{1, -1\}$. The group $\mathcal{g}$ makes a group with respect to the multiplication. This is an example of a finite group.
(ii) $\mathcal{g} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$. The group $\mathcal{g}$ makes a group with respect to the addition. This is an infinite group. For instance, take $a(>0)$ and make $a+1$ and make $(a+1)+1, [(a+1)+1]+1, \ldots$ again and again. Thus, addition is not closed and, hence, we must have an infinite group.
(iii) Let us start with a matrix $a = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. Then, the inverse $a^{-1} = a^3 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. We have $a^2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$. Its inverse is $a^2$ itself. These four elements make a group. That is, $\mathcal{g} = \{e, a, a^2, a^3\}$. This is an example of cyclic groups.
(iv) $\mathcal{g} = \{1\}$. It is a most trivial case, but sometimes the trivial case is very important as well. We will come back later to this point.
(v) Let us think of a more general case. In Chap. 9, we discussed endomorphism on a vector space and showed the necessary and sufficient condition for the existence of an inverse transformation. In this relation, we consider a set that comprises matrices such that

$$\mathrm{GL}(n, \mathbb{C}) \equiv \{A = (a_{ij}) | i, j = 1, 2, \ldots, n;\ a_{ij} \in \mathbb{C}, \det A \neq 0\}.$$

This may be either a finite group or an infinite group. The former can be a symmetry group and the latter can be a rotation group. This group is characterized by a set of invertible and endomorphic linear transformations over a vector space $V^n$ and called a linear transformation group or a general linear group and denoted by $\mathrm{GL}(n, \mathbb{C})$, $\mathrm{GL}(V^n)$, $\mathrm{GL}(V)$, etc. The relevant transformations are bijective.

We can readily make sure that axioms (A1) to (A4) are satisfied with $\mathrm{GL}(n, \mathbb{C})$. Here, the vector space can be $\mathbb{C}^n$ or a function space.

The structure of a finite group is tabulated in a *multiplication table*. This is made up such that group elements are arranged in a first row and a first column and that an intersection of an element $g_i$ in the row and $g_j$ in the column is designated as a product $g_i \diamond g_j$. Choosing the above (iii) for an example, we make its multiplication table (see Table 13.1). There we define $a^2 = b$ and $a^3 = c$.

Having a look at Table 13.1, we notice that in the individual rows and columns each group element appears once and only once. This is well known as a rearrangement theorem.

**Theorem 13.1: Rearrangement Theorem [1]** *In each row or each column in the group multiplication table, individual group elements appear once and only once. From this, each row and each column list merely rearranged group* elements.

*Proof* Let a set $\mathcal{g} = \{g_1 \equiv e, g_2, \ldots, g_n\}$ be a group. Arbitrarily choosing any element $h$ from $\mathcal{g}$ and multiplying individual elements by $h$, we obtain a set $\mathfrak{H} = \{hg_1, hg_2, \cdots, hg_n\}$. Then, all the group elements of $\mathcal{g}$ appear in $\mathfrak{H}$ once and only once. Choosing any group element $g_i$, let us multiply $g_i$ by $h^{-1}$ to get $h^{-1}g_i$. Since $h^{-1}g_i$ must be a certain element $g_k$ of $\mathcal{g}$, we put $h^{-1}g_i = g_k$. Multiplying both sides by $h$, we have $g_i = hg_k$. Therefore, we are able to find this very element $hg_k$ in $\mathfrak{H}$, i.e., $g_i$ in $\mathfrak{H}$. This implies that the element $g_i$ necessarily appears in $\mathfrak{H}$. Suppose in turn that $g_i$ appears more than once. Then, we must have $g_i = hg_k = hg_l (k \neq l)$. Multiplying the relation by $h^{-1}$, we would get $h^{-1}g_i = g_k = g_l$, in contradiction to the supposition. This means that $g_i$ appears in $\mathfrak{H}$ once and only once. This confirms that the theorem is true of each row of the group multiplication table.

A similar argument applies with a set $\mathfrak{H}' = \{g_1 h, g_2 h, \ldots, g_n h\}$. This confirms in turn that the theorem is true of each column of the group multiplication table. These complete the proof.

## 13.2 Subgroups

As we think of subspaces in a linear vector space, we have subgroups in a group. The definition of the subgroup is that a subset $\mathcal{H}$ of a group makes a group with respect to the multiplication $\diamond$ that is defined for the group $\mathcal{g}$. The identity element

**Table 13.1** Multiplication table of $\{\mathcal{g} = e, a, a^2, a^3\}$

| $\mathcal{g}$ | $e$ | $a$ | $b \equiv a^2$ | $c \equiv a^3$ |
|---|---|---|---|---|
| $e$ | $e$ | $a$ | $b$ | $c$ |
| $a$ | $a$ | $b$ | $c$ | $e$ |
| $b$ | $b$ | $c$ | $e$ | $a$ |
| $c$ | $c$ | $e$ | $a$ | $b$ |

makes a group by itself. Both $\{e\}$ and $\mathcal{G}$ are subgroups as well. We often call subgroups other than $\{e\}$ and $\mathcal{G}$ "proper" subgroups.

A necessary and sufficient condition for the subset $\mathcal{H}$ to be a subgroup is the following:

(1) $h_i, h_j \in \mathcal{H} \Rightarrow h_i \diamond h_j \in \mathcal{H}$.
(2) $h \in \mathcal{H} \Rightarrow h^{-1} \in \mathcal{H}$.

If $\mathcal{H}$ is a subgroup of $\mathcal{G}$, it is obvious that the relations (1) and (2) hold. Conversely, if (1) and (2) hold, $\mathcal{H}$ is a subgroup. In fact, (1) ensures the aforementioned relation (A1). Since $\mathcal{H}$ is a subset of $\mathcal{G}$, this guarantees the associative law (A2). The relation (2) ensures (A4). Finally, in virtue of (1) and (2), $h \diamond h^{-1} = e$ is contained in $\mathcal{H}$; this implies that (A3) is satisfied. Thus, $\mathcal{H}$ is a subgroup, because $\mathcal{H}$ satisfies the axioms (A1) to (A4). Of the above examples, (iii) has a subgroup $\mathcal{H} = \{e, a^2\}$.

It is important to decompose a set into subsets that do not mutually contain an element (except for a special element) among them. We saw this in Part III when we decomposed a linear vector space into subspaces. In that case the said special element was a zero vector. Here let us consider a related question in its similar aspects.

Let $\mathcal{H} = \{h_1 \equiv e, h_2, \ldots, h_s\}$ be a subgroup of $\mathcal{G}$. Also, let us consider $a\mathcal{H}$ where $^{\exists}a \in \mathcal{G}$ and $a \notin \mathcal{H}$. Suppose that $a\mathcal{H}$ is a subset of $\mathcal{G}$ such that $a\mathcal{H} = \{ah_1, ah_2, \ldots, ah_s\}$. Then, we have another subset $\mathcal{H} + a\mathcal{H}$. If $\mathcal{H}$ contains $s$ elements, so does $a\mathcal{H}$. In fact, if it were not the case, namely, if $ah_i = ah_j$, multiplying the both sides by $a^{-1}$ we would have $h_i = h_j$, in contradiction. Next, let us take $b$ such that $b \notin \mathcal{H}$ and $b \notin a\mathcal{H}$ and make up $b\mathcal{H}$ and $\mathcal{H} + a\mathcal{H} + b\mathcal{H}$ successively. Our question is whether these procedures decompose $\mathcal{G}$ into subsets mutually exclusive and collectively exhaustive.

Suppose that we can succeed in such a decomposition and get

$$\mathcal{G} = g_1\mathcal{H} + g_2\mathcal{H} + \cdots + g_k\mathcal{H}, \tag{13.2}$$

where $g_1, g_2, \ldots, g_k$ are mutually different elements with $g_1$ being the identity $e$. In that case (13.2) is said to be the left coset decomposition of $\mathcal{G}$ by $\mathcal{H}$. Similarly, right coset decomposition can be done to give

$$\mathcal{G} = \mathcal{H}g_1 + \mathcal{H}g_2 + \cdots + \mathcal{H}g_k. \tag{13.3}$$

In general, however,

$$g_k\mathcal{H} \neq \mathcal{H}g_k \text{ or } g_k\mathcal{H}g_k^{-1} \neq \mathcal{H}. \tag{13.4}$$

Taking the case of left coset as an example, let us examine whether different cosets mutually contain a common element. Suppose that $g_i\mathcal{H}$ and $g_j\mathcal{H}$ mutually contain a common element. Then, that element would be expressed as $g_ih_p = g_jh_q (1 \leq i, j \leq n; 1 \leq p, q \leq s)$. Thus, we have $g_ih_ph_q^{-1} = g_j$. Since $\mathcal{H}$ is a

subgroup of $\mathcal{G}$, $h_p h_q^{-1} \in \mathcal{H}$. This implies that $g_j \in g_i\mathcal{H}$. It is in contradiction to the definition of left coset. Thus, we conclude that different cosets do not mutually contain a common element.

Suppose that the order of $\mathcal{G}$ and $\mathcal{H}$ is $n$ and $s$, respectively. Different $k$ cosets comprise $s$ elements individually and different cosets do not mutually possess a common element and, hence, we must have

$$n = sk, \tag{13.5}$$

where $k$ is called an index of $\mathcal{H}$. We will have many examples afterward.

## 13.3 Classes

Another method to decompose a group into subsets is called conjugacy classes. A conjugate element is defined as follows: Let $a$ be an element arbitrarily chosen from a group. Then an element $gag^{-1}$ is called a conjugate element or conjugate to $a$. If $c$ is conjugate to $b$ and $b$ is conjugate to $a$, then $c$ is conjugate to $a$. It is because

$$c = g'bg'^{-1}, b = gag^{-1} \Rightarrow c = g'bg'^{-1} = g'gag^{-1}g'^{-1} = g'ga(g'g)^{-1}. \tag{13.6}$$

In the above, a set containing $a$ and all the elements conjugate to $a$ is said to be a (conjugate) class of $a$. Denoting this set by $\mathsf{C}_a$, we have

$$\mathsf{C}_a = \left\{a, g_2 a g_2^{-1}, g_3 a g_3^{-1}, \ldots, g_n a g_n^{-1}\right\}. \tag{13.7}$$

In $\mathsf{C}_a$, a same element may appear repeatedly. It is obvious that in every group the identity element $e$ forms a class by itself. That is,

$$\mathsf{C}_e = \{e\}. \tag{13.8}$$

As in the case of the decomposition of a group into (left or right) cosets, we can decompose a group to classes. If group elements are not exhausted by a set comprising $\mathsf{C}_e$ or $\mathsf{C}_a$, let us take $b$ such that $b \neq e$ and $b \notin \mathsf{C}_a$ and make $\mathsf{C}_b$ similarly to (13.7). Repeating this procedure, we should be able to decompose a group into classes. In fact, if group elements have not yet exhausted after these procedures, take remaining element $z$ and make a class. If the remaining element is only $z$ in this moment, $z$ can make a class by itself (as in the case of $e$). Notice that for an Abelian group every element makes a class by itself. Thus, with a finite group, we have a decomposition such that

$$\mathcal{G} = \mathsf{C}_e + \mathsf{C}_a + \mathsf{C}_b + \cdots \mathsf{C}_z. \tag{13.9}$$

To show that (13.9) is really a decomposition, suppose that for instance a set $\mathsf{C}_a \cap \mathsf{C}_b$ is not an empty set and that $x \in \mathsf{C}_a \cap \mathsf{C}_b$. Then, we must have $\alpha$ and $\beta$ that

satisfy a following relation: $x = \alpha a \alpha^{-1} = \beta b \beta^{-1}$, i.e., $b = \beta^{-1}\alpha a \alpha^{-1}$ $\beta = \beta^{-1}\alpha a(\beta^{-1}\alpha)^{-1}$. This implies that $b$ has already been included in $\mathsf{C}_a$, in contradiction to the supposition. Thus, (13.9) is in fact a decomposition of $\mathscr{g}$ into a finite number of classes.

In the above, we thought of a class conjugate to a single element. This notion can be extended to a class conjugate to a subgroup. Let $\mathcal{H}$ be a subgroup of $\mathscr{g}$. Let $g$ be an element of $\mathscr{g}$. Let us now consider a set $\mathcal{H}' = g\mathcal{H}g^{-1}$. The set $\mathcal{H}'$ is a subgroup of $\mathscr{g}$ and is called a conjugate subgroup.

In fact, let $h_i$ and $h_j$ be any two elements of $\mathcal{H}$, that is, let $gh_ig^{-1}$ and $gh_jg^{-1}$ be any tow elements of $\mathcal{H}'$. Then, we have

$$\left(gh_ig^{-1}\right)\left(gh_jg^{-1}\right) = gh_ih_jg^{-1} = gh_kg^{-1}, \tag{13.10}$$

where $h_k = h_ih_j \in \mathcal{H}$. Hence, $gh_kg^{-1} \in \mathcal{H}'$. Meanwhile, $(gh_ig^{-1})^{-1} = gh_i^{-1}$ $g^{-1} \in \mathcal{H}'$. Thus, conditions (1) and (2) of Sect. 13.2 are satisfied with $\mathcal{H}'$. Therefore, $\mathcal{H}'$ is a subgroup of $\mathscr{g}$. The subgroup $\mathcal{H}'$ has a same order as $\mathcal{H}$. This is because with any two different elements $h_i$ and $h_j$ $gh_ig^{-1} \neq gh_jg^{-1}$.

If for $^{\forall}g \in \mathscr{g}$ and a subgroup $\mathcal{H}$, we have a following equality

$$g^{-1}\mathcal{H}g = \mathcal{H}, \tag{13.11}$$

such a subgroup $\mathcal{H}$ is said to be an invariant subgroup. If (13.11) holds, $\mathcal{H}$ should be a sum of classes (reader, please show this). A set comprising only the identity, i.e., $\{e\}$ forms a class. Therefore, if $\mathcal{H}$ is a proper subgroup, $\mathcal{H}$ must contain two or more classes. The relation (13.11) can be rewritten as

$$g\mathcal{H} = \mathcal{H}g. \tag{13.12}$$

This implies that the left coset is identical to the right coset. Thus, as far as we are dealing with a coset pertinent to an invariant subgroup, we do not have to distinguish left and right cosets.

Now let us anew consider the (left) coset decomposition of $\mathscr{g}$ by an invariant subgroup $\mathcal{H}$

$$\mathscr{g} = g_1\mathcal{H} + g_2\mathcal{H} + \cdots + g_k\mathcal{H}, \tag{13.13}$$

where we have $\mathcal{H} = \{h_1 \equiv e, h_2, \ldots, h_s\}$. Then, multiplication of two elements that belong to the cosets $g_i\mathcal{H}$ and $g_j\mathcal{H}$ is expressed as

$$(g_ih_l)\left(g_jh_m\right) = \left(g_ig_jg_j^{-1}h_l\right)\left(g_jh_m\right) = g_ig_j(g_j^{-1}h_lg_j)h_m = g_ig_jh_ph_m = g_ig_jh_q, \tag{13.14}$$

where the third equality comes from (13.11). That is, we should have $^{\exists}h_p$ such that $g_j^{-1}h_l g_j = h_p$, and $h_p h_m = h_q$. In (13.14), $h_\alpha \in \mathcal{H}$ ($\alpha$ stands for $l$, $m$, $p$, $q$, etc., with $1 \leq \alpha \leq s$). Note that $g_i g_j h_q \in g_i g_j \mathcal{H}$. Accordingly, a product of elements belonging to $g_i\mathcal{H}$ and $g_j\mathcal{H}$ belongs to $g_i g_j \mathcal{H}$. We rewrite (13.14) as a relation between the sets

$$(g_i\mathcal{H})(g_j\mathcal{H}) = g_i g_j \mathcal{H}. \tag{13.15}$$

Viewing LHS of (13.15) as a product of two cosets, we find that the said product is a coset as well. This implies that a collection of the cosets forms a group. Such a group that possesses cosets as elements is said to be a factor group or quotient group. In this context, the multiplication is a product of cosets. We denote the factor group by

$$\mathcal{G}/\mathcal{H}$$

An identity element of this factor group is $\mathcal{H}$. This is because in (13.15) putting $g_i = e$, we get $\mathcal{H}(g_j\mathcal{H}) = g_j\mathcal{H}$. Alternatively, putting $g_j = e$, we have $(g_i\mathcal{H})\mathcal{H} = g_i\mathcal{H}$. In (13.15), moreover, putting $g_j = g_i^{-1}$, we get

$$(g_i\mathcal{H})(g_i^{-1}\mathcal{H}) = g_i g_i^{-1}\mathcal{H} = \mathcal{H}. \tag{13.16}$$

Hence, $(g_i\mathcal{H})^{-1} = g_i^{-1}\mathcal{H}$. That is, the inverse element of $g_i\mathcal{H}$ is $g_i^{-1}\mathcal{H}$.

## 13.4 Isomorphism and Homomorphism

As in the case of the linear vector space, we consider the mapping between group elements. Of these, the notion of isomorphism and homomorphism is important.

**Definition 13.1**
Let $\mathcal{G} = \{x, y, \ldots\}$ and $\mathcal{G}' = \{x', y', \ldots\}$ be groups and let a mapping $\mathcal{G} \to \mathcal{G}'$ exist. Suppose that there is a one-to-one correspondence (i.e., injective mapping)

$$x \leftrightarrow x', y \leftrightarrow y', \cdots$$

between the elements such that $xy = z$ implies that $x'y' = z'$ and vice versa. Meanwhile, any element in $\mathcal{G}'$ must be the image of some element of $\mathcal{G}$. That is, the mapping is surjective as well and, hence, the mapping is bijective. Then, the two groups $\mathcal{G}$ and $\mathcal{G}'$ are said to be *isomorphic*. The relevant mapping is called an *isomorphism*. We symbolically denote this relation by

$$\mathcal{G} \cong \mathcal{G}'.$$

Note that the aforementioned groups can be either a finite group or an infinite group. We did not designate identity elements. Suppose that $x$ is the identity $e$. Then, from the relations $xy = z$ and $x'y' = z'$, we have

$$ey = z = y,\ x'y' = z' = y'. \tag{13.17}$$

Then, we get

$$x' = e', \text{i.e. } e \leftrightarrow e'. \tag{13.18}$$

Also let us put $y = x^{-1}$. Then,

$$xx^{-1} = z = e,\ x'\left(x^{-1}\right)' = e',\ x'y' = z' = e'. \tag{13.19}$$

Comparing the second and third equations of (13.19), we get

$$y' = x'^{-1} = \left(x^{-1}\right)'. \tag{13.20}$$

The bijective character mentioned above can somewhat be loosened in such a way that the one-to-one correspondence is replaced with $n$-to-one correspondence. We have a following definition.

**Definition 13.2**
Let $\mathcal{g} = \{x, y, \ldots\}$ and $\mathcal{g}' = \{x', y', \ldots\}$ be groups and let a mapping $\mathcal{g} \to \mathcal{g}'$ exist. Also let a mapping $\rho$: $\mathcal{g} \to \mathcal{g}'$ exist such that with arbitrarily chosen any two elements, the following relation holds:

$$\rho(x)\rho(y) = \rho(xy). \tag{13.21}$$

Then, the two groups $\mathcal{g}$ and $\mathcal{g}'$ are said to be *homomorphic*. The relevant mapping is called *homomorphism*. We symbolically denote this relation by

$$\mathcal{g} \sim \mathcal{g}'.$$

In this case, we have

$$\rho(e)\rho(e) = \rho(ee) = \rho(e), \text{i.e. } \rho(e) = e',$$

where $e'$ is an identity element of $\mathcal{g}'$. Also, we have

$$\rho(x)\rho\left(x^{-1}\right) = \rho\left(xx^{-1}\right) = \rho(e) = e'.$$

Therefore,

$$[\rho(x)]^{-1} = \rho\left(x^{-1}\right).$$

The two groups can be either a finite group or an infinite group. Note that in the above, the mapping is not injective. The mapping may or may not be surjective.

Regarding the identity and inverse elements, we have the same relations as (13.18) and (13.20). From Definitions 13.1 and 13.2, we say that the bijective homomorphism is the isomorphism.

Let us introduce an important notion of a kernel of a mapping. In this regard, we have a following theorem.

**Definition 13.3**
Let $\mathscr{g} = \{e, x, y, \ldots\}$ and $\mathscr{g}' = \{e', x', y', \ldots\}$ be groups and let $e$ and $e'$ be the identity elements. Suppose that there exists a homomorphic mapping $\rho$: $\mathscr{g} \to \mathscr{g}'$. Also let $\mathcal{F}$ be a subset of $\mathscr{g}$ such that

$$\rho(\mathcal{F}) = e'. \tag{13.22}$$

Then, $\mathcal{F}$ is said to be a kernel of $\rho$.

Regarding the kernel, we have following important theorems.

**Theorem 13.2**
*Let $\mathscr{g} = \{e, x, y, \cdots\}$ and $\mathscr{g}' = \{e', x', y', \cdots\}$ be groups, where e and e′ are identity elements. A necessary and sufficient condition for a surjective and homomorphic mapping $\rho : \mathscr{g} \to \mathscr{g}'$ to be isomorphic is that a kernel $\mathcal{F} = \{e\}$.*

*Proof* We assume that $\mathcal{F} = \{e\}$. Suppose that $\rho(x) = \rho(y)$. Then, we have

$$\rho(x)[\rho(y)]^{-1} = \rho(x)\rho\left(y^{-1}\right) = \rho\left(xy^{-1}\right) = e'. \tag{13.23}$$

The first and second equalities result from the homomorphism of $\rho$. Since $\mathcal{F} = \{e\}$, $xy^{-1} = e$, i.e., $x = y$. Therefore, $\rho$ is injective (i.e., one-to-one correspondence). As $\rho$ is surjective from the assumption, $\rho$ is bijective. The mapping $\rho$ is isomorphic accordingly.

Conversely, suppose that $\rho$ is isomorphic. Also suppose for $^{\exists}x \in \mathscr{g}$ $\rho(x) = e'$. From (13.18), $\rho(e) = e'$. We have $\rho(x) = \rho(e) = e' \Rightarrow x = e$ due to the isomorphism of $\rho$ (i.e., one-to-one correspondence). This implies $\mathcal{F} = \{e\}$. This completes the proof.

We become aware of close relationship between Theorem 13.1 and linear transformation versus kernel already mentioned in Sect. 9.2 of Part III. Figure 13.1 shows this relationship. Figure 13.1a represents homomorphic mapping $\rho$ in a group, whereas Fig. 13.1b shows linear transformation $A$ in a vector space.

**Theorem 13.3** *Suppose that there exists a homomorphic mapping $\rho$: $\mathscr{g} \to \mathscr{g}'$, where $\mathscr{g}$ and $\mathscr{g}'$ are groups. Then, a kernel $\mathcal{F}$ of $\rho$ is an invariant subgroup of $\mathscr{g}$.*

*Proof* Let $k_i$ and $k_j$ be any two arbitrarily chosen elements of $\mathcal{F}$. Then,

$$\rho(k_i) = \rho\left(k_j\right) = e', \tag{13.24}$$

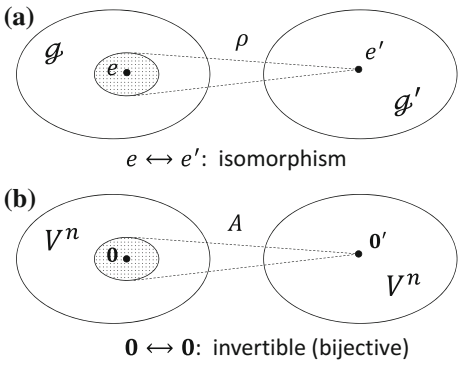


**Fig. 13.1** Mapping in a group and vector space. **a** Homomorphic Mapping $\rho$ in a group. **b** Linear transformation $A$ in a vector space

where $e'$ is the identity element of $\mathcal{G}'$. From (13.21), we have

$$\rho\left(k_i k_j\right) = \rho(k_i)\rho\left(k_j\right) = e'e' = e'. \tag{13.25}$$

Therefore, $k_i k_j \in \mathcal{F}$. Meanwhile, from (13.20), we have

$$\rho\left(k_i^{-1}\right) = [\rho(k_i)]^{-1} = e'^{-1} = e'. \tag{13.26}$$

Then, $k_i^{-1} \in \mathcal{F}$. Thus, $\mathcal{F}$ is a subgroup of $\mathcal{G}$.
Next, for ${}^{\forall}g \in \mathcal{G}$, we have

$$\rho\left(gk_i g^{-1}\right) = \rho(g)\rho(k_i)\rho\left(g^{-1}\right) = \rho(g)e'\rho\left(g^{-1}\right) = e'. \tag{13.27}$$

Accordingly, we have $gk_i g^{-1} \in \mathcal{F}$. Thus, $g\mathcal{F}g^{-1} \subset \mathcal{F}$. Since $g$ is chosen arbitrarily, replacing it with $g^{-1}$ we have $g^{-1}\mathcal{F}g \subset \mathcal{F}$. Multiplying $g$ and $g^{-1}$ on both sides from the left and right, respectively, we get $\mathcal{F} \subset g\mathcal{F}g^{-1}$. Consequently, we get

$$g\mathcal{F}g^{-1} = \mathcal{F}. \tag{13.28}$$

This implies that $\mathcal{F}$ of $\rho$ is an invariant subgroup of $\mathcal{G}$.

**Theorem 13.4 (Homomorphism Theorem)**
*Let $\mathcal{G} = \{x, y, \ldots\}$ and $\mathcal{G}' = \{x', y', \ldots\}$ be groups and let a homomorphic (and surjective) mapping $\rho$: $\mathcal{G} \to \mathcal{G}'$ exist. Also let $\mathcal{F}$ be a kernel of $\mathcal{G}$. Let us define a surjective mapping $\widetilde{\rho}$: $\mathcal{G}/\mathcal{F} \to \mathcal{G}'$ such that*

$$\widetilde{\rho}(g_i\mathcal{F}) = \rho(g_i). \tag{13.29}$$

*Then, $\widetilde{\rho}$ is an isomorphic mapping.*

*Proof* From (13.15) and (13.21), it is obvious that $\widetilde{\rho}$ is homomorphic. The confirmation is left for readers. Let $g_i\mathcal{F}$ and $g_j\mathcal{F}$ be two different cosets. Suppose here that $\rho(g_i) = \rho(g_j)$. Then, we have

$$\rho\left(g_i^{-1}g_j\right) = \rho\left(g_i^{-1}\right)\rho\left(g_j\right) = [\rho(g_i)]^{-1}\rho\left(g_j\right) = [\rho(g_i)]^{-1}\rho(g_i) = e'. \tag{13.30}$$

This implies that $g_i^{-1}g_j \in \mathcal{F}$. That is, we would have $g_j \in g_i\mathcal{F}$. This is in contradiction to the definition of a coset. Thus, we should have $\rho(g_i) \neq \rho\left(g_j\right)$. In other words, the different cosets $g_i\mathcal{F}$ and $g_j\mathcal{F}$ have been mapped into different elements $\rho(g_i)$ and $\rho\left(g_j\right)$ in $\mathcal{g}'$. That is, $\widetilde{\rho}$ is isomorphic; i.e., $\mathcal{g}/\mathcal{F} \cong \mathcal{g}'$.

## 13.5 Direct-Product Groups

So far we have investigated basic properties of groups. In Sect. 13.4, we examined factor groups. The homomorphism theorem shows that the factor group is characterized by division. In the case of a finite group, an order of the group is reduced. In this section, we study the opposite character, i.e., properties of direct product of groups, or direct-product groups.

Let $\mathcal{H} = \{h_1 \equiv e, h_2, \ldots, h_m\}$ and $\mathcal{H}' = \left\{h'_1 \equiv e, h'_2, \ldots, h'_n\right\}$ be groups of the order of $m$ and $n$, respectively. Suppose that (i) ${}^{\forall}h_i (1 \leq i \leq m)$ and ${}^{\forall}h'_j (1 \leq i \leq m)$ commute, i.e., $h_ih'_j = h'_jh_i$ and that (ii) $\mathcal{H} \cap \mathcal{H}' = \{e\}$. Under these conditions let us construct a set $\mathcal{g}$ such that

$$\mathcal{g} = \left\{h_1h'_1 \equiv e, h_ih'_j (1 \leq i \leq m, 1 \leq j \leq n)\right\}. \tag{13.31}$$

In other words, $\mathcal{g}$ is a set comprising $mn$ elements $h_ih'_j$. A product of elements is defined as

$$(h_ih'_j)(h_kh'_l) = h_ih_kh'_jh'_l = h_ph'_q, \tag{13.32}$$

where $h_p = h_ih_k$ and $h'_q = h'_jh'_l$. The identity element is $ee = e; eh_ih_k = h_ih_ke$. The inverse element is $(h_ih'_j)^{-1} = h_j'^{-1}h_i^{-1} = h_i^{-1}h_j'^{-1}$. Associative law is obvious from $h_ih'_j = h'_jh_i$. Thus, $\mathcal{g}$ forms a group. This is said to be a direct product of groups, or a direct-product group. The groups $\mathcal{H}$ and $\mathcal{H}'$ are called direct factors of $\mathcal{g}$. In this case, we succinctly represent

$$\mathcal{g} = \mathcal{H} \otimes \mathcal{H}'.$$

In the above, the condition (ii) is equivalent to that ${}^{\forall}g \in \mathcal{g}$ is uniquely represented as

$$g = hh'; \, h \in \mathcal{H}, \, h' \in \mathcal{H}'. \tag{13.33}$$

In fact, suppose that $\mathcal{H} \cap \mathcal{H}' = \{e\}$ and that $g$ can be represented in two ways such that

$$g = h_1 h_1' = h_2 h_2'; h_1, h_2 \in \mathcal{H}, h_1', h_2' \in \mathcal{H}'. \tag{13.34}$$

Then, we have

$$h_2^{-1} h_1 = h_2' h_1'^{-1}; h_2^{-1} h_1 \in \mathcal{H}, h_2' h_1'^{-1} \in \mathcal{H}'. \tag{13.35}$$

From the supposition, we get

$$h_2^{-1} h_1 = h_2' h_1'^{-1} = e. \tag{13.36}$$

That is, $h_2 = h_1$ and $h_2' = h_1'$. This means that the representation is unique.

Conversely, suppose that the representation is unique and that $x \in \mathcal{H} \cap \mathcal{H}'$. Then, we must have

$$x = xe = ex. \tag{13.37}$$

Thanks to uniqueness of the representation, $x = e$. This implies $\mathcal{H} \cap \mathcal{H}' = \{e\}$.

Now suppose $h \in \mathcal{H}$. Then for $^{\forall}g \in \mathcal{G}$ putting $g = h_\nu h_\mu'$, we have

$$ghg^{-1} = h_\nu h_\mu' h h_\mu'^{-1} h_\nu^{-1} = h_\nu h h_\mu' h_\mu'^{-1} h_\nu^{-1} = h_\nu h h_\nu^{-1} \in \mathcal{H}. \tag{13.38}$$

Then, we have $g\mathcal{H}g^{-1} \subset \mathcal{H}$. Similarly to the proof of Theorem 13.3, we get

$$g\mathcal{H}g^{-1} = \mathcal{H}. \tag{13.39}$$

This shows that $\mathcal{H}$ is an invariant subgroup of $\mathcal{G}$. Similarly, $\mathcal{H}'$ is an invariant subgroup as well.

Regarding the unique representation of the group element of a direct-product group, we become again aware of the close relationship between the direct product and direct sum that was mentioned earlier in Part III.

## Reference

1. Cotton FA (1990) Chemical applications of group theory. Wiley, New York

# Chapter 14
# Symmetry Groups

We have many opportunities to observe symmetry both macroscopic and microscopic in natural world. First, we need to formulate the symmetry appropriately. For this purpose, we must regard various symmetry operations as mathematical elements and classify these operations under several categories. In Part III, we examined various properties of vectors and their transformations. We also showed that the vector transformation can be viewed as the coordinate transformation. On these topics, we focused upon abstract concepts in various ways. On another front, however, we have not paid attention to specific geometric objects, especially molecules. In this chapter, we study the symmetry of these concrete objects. For this, it will be indispensable to correctly understand a variety of symmetry operations. At the same time, we deal with the vector and coordinate transformations as group elements. Among such transformations, rotations occupy a central place in the group theory and related field of mathematics. Regarding the three-dimensional Euclidean space, $SO(3)$ is particularly important. This is characterized by an infinite group in contrast to various symmetry groups (or point groups) we investigate in the former parts of this chapter.

## 14.1 A Variety of Symmetry Operations

To understand various aspects of symmetry operations, it is convenient and essential to consider a general point that is fixed in a three-dimensional Euclidean space and to examine how this point is transformed in the space. In parallel to the description in Part III, we express the coordinate of the general point $P$ as

$$P = \begin{pmatrix} x \\ y \\ z \end{pmatrix}. \tag{14.1}$$



S. Hotta, *Mathematical Physical Chemistry*,
https://doi.org/10.1007/978-981-10-7671-8_14

Note that $P$ may be on or within or outside a geometric object or molecule that we are dealing with. The relevant position vector $\boldsymbol{x}$ for $P$ is expresses as

$$\begin{aligned} \boldsymbol{x} &= x\boldsymbol{e}_1 + y\boldsymbol{e}_2 + z\boldsymbol{e}_3. \\ &= (\boldsymbol{e}_1\, \boldsymbol{e}_2\, \boldsymbol{e}_3) \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \end{aligned} \tag{14.2}$$

where $\boldsymbol{e}_1, \boldsymbol{e}_2, \text{and}, \boldsymbol{e}_3$ denote an orthonormal basis vectors pointing to positive directions of $x$-, $y$-, and $z$-axes, respectively. Similarly, we denote a linear transformation $A$ by

$$A(\boldsymbol{x}) = (\boldsymbol{e}_1\, \boldsymbol{e}_2\, \boldsymbol{e}_3) \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}. \tag{14.3}$$

Among various linear transformations that are represented by matrices, orthogonal transformations are the simplest and most widely used. We use orthogonal matrices to represent the orthogonal transformations accordingly.

Let us think of a movement or translation of a geometric object and an operation that causes such a movement. First suppose that the geometric object is fixed on a coordinate system. Then, the object is moved (or translate) to another place. If before and after such a movement (or translation) one could not tell whether the object has been moved, we say that the object possesses a "symmetry" in a certain sense. Thus, we have to specify this symmetry. In that context, group theory deals with the symmetry and defines it clearly.

To tell whether the object has been moved (to another place), we usually distinguish it by change in (i) positional relationship and (ii) attribute or property. To make the situation simple, let us consider a following example:

*Example. 14.1* We have two round disks; i.e., Disk A and Disk B. Suppose that Disk A is a solid white disk, whereas Disk B is partly painted black (see Fig. 14.1). In Fig. 14.1, we are thinking of a rotation of an object (e.g., round disk) around an axis standing on its center and stretching perpendicularly to the object plane.

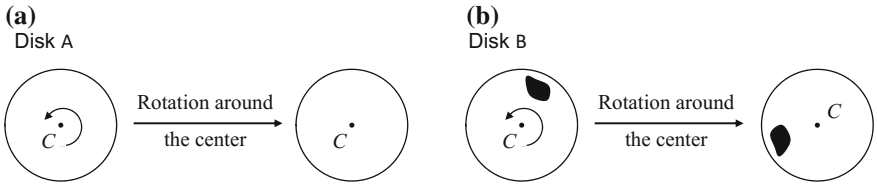


**Fig. 14.1** Rotation of an object. **a** Case where we cannot recognize that the object has been moved. **b** Case where we can recognize that the object has been moved because of its attribute (i.e., because the round disk is partly painted black)

If an arbitrarily chosen positional vector fixed on the object before the rotation is moved to another position that was not originally occupied by the object, then we recognize that the object has certainly been moved. For instance, imagine that a round disk having a through-hole located aside from center is rotating. What about the case where that position was originally occupied by the object, then? We have two possibilities. The first alternative is that we cannot recognize that the object has been moved. The second one is that we can yet recognize that the object has been moved. According to Fig. 14.1a, b, we have the former case and the latter case, respectively. In the latter case, we have recognized the movement of the object by its attribute, i.e., by that the object is partly painted black.

However, we do not have to be rigorous here. We have a clear intuitive criterion for a judgement of whether a geometric object has been moved. From now on, we assume that the geometric character of an object is pertinent to both its positional relationship and attribute. Thus, we define the equivalent (or indistinguishable) disposition of an object and the operation that yields such an equivalent disposition as follows:

**Definition 14.1**

(i) Symmetry operation: A geometric operation that produces an equivalent (or indistinguishable) disposition of an object.
(ii) Equivalent (or indistinguishable) disposition: Suppose that regarding a geometric operation of an object, we cannot recognize that the object has been moved before and after that geometric operation. In that case, the original disposition of the object and the resulting disposition reached after the geometric operation are referred to as an equivalent disposition. The relevant geometric operation is the symmetric operation.

Here we should clearly distinguish *translation* (i.e., parallel displacement) from the above-mentioned symmetry operations. This is because for a geometric object to possess the translation symmetry the object must be infinite in extent, typically an infinite crystal lattice. The relevant discipline is widely studied as space group and has a broad class of applications in physics and chemistry. However, we will not deal with the space group or associated topics, but focus our attention upon symmetry groups in this book.

In the above example, the rotation is a symmetry operation with Fig. 14.1a, but the said geometric operation is not a symmetric operation with Fig. 14.1b.

Let us further inspect properties of the symmetry operation. Let us consider a set $\mathfrak{H}$ consisting of symmetry operations. Let $a$ and $b$ be any two symmetric operations of $\mathfrak{H}$. Then, (i) $a \diamond b$ is a symmetric operation as well. (ii) Multiplication of successive symmetric operations $a, b, c$ is associative; i.e., $a \diamond (b \diamond c) = (a \diamond b) \diamond c$. (iii) The set $\mathfrak{H}$ contains an element of $e$ called the identity element such that we have $a \diamond e = e \diamond a = a$ with any element $a$ of $\mathfrak{H}$. Operating "nothing" should be $e$. If the rotation is relevant, $2\pi$ rotation is thought to be $e$. These are intuitively acceptable. (iv) For any $a$ of $\mathfrak{H}$, we have an element $b$ such that $a \diamond b = b \diamond a = e$. The element $b$ is said to be the inverse element of $a$. We

denote it by $b \equiv a^{-1}$. The inverse element corresponds to an operation that brings the disposition of a geometric object back to the original disposition. Thus, $\mathfrak{H}$ forms a group. We call $\mathfrak{H}$ satisfying the above criteria a *symmetry group*. A symmetry group is called a point group as well. This is because a point group comprises symmetry operations of geometric objects as group elements, and those objects have at least one fixed point after the relevant symmetry operation. The name of a point group comes from this fact.

As mentioned above, the symmetric operation is best characterized by a (3,3) orthogonal matrix. In Example 14.1, e.g., the $\pi$ rotation is represented by an orthogonal matrix $A$ such that

$$A = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{14.4}$$

This operation represents a $\pi$ rotation around the $z$-axis. Let us think of another symmetric operation described by

$$B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \tag{14.5}$$

This produces a mirror symmetry with respect to the $xy$-plane. Then, we have

$$C \equiv AB = BA = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \tag{14.6}$$

The operation $C$ shows an inversion about the origin. Thus, $A$, $B$, and $C$ along with an identity unit $E$ form a group. Here $E$ is expressed as

$$E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{14.7}$$

The above group is represented by four three-dimensional diagonal matrices whose elements are 1 or $-1$. Therefore, it is evident that an inverse element of $A$, $B$, and $C$ is $A$, $B$, and $C$ itself, respectively. The said group is commutative (Abelian) and said to be a *four group* [1].

Meanwhile, we have a number of non-commutative groups. From a point of view of a matrix structure, non-commutativity comes from off-diagonal elements of the matrix. A typical example is a rotation matrix of a rotation angles different from zero or $n\pi$ ($n$: integer). For later use, let us have a matrix form that expresses a $\theta$ rotation around the $z$-axis. Figure 14.2 depicts a graphical illustration for this.

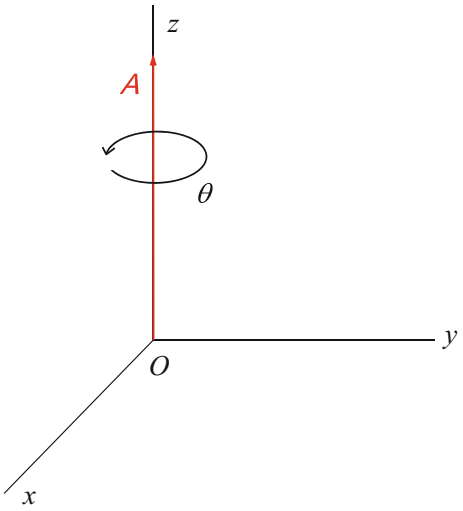


**Fig. 14.2** Rotation by $\theta$ around the $z$-axis

A matrix $R$ has a following form:

$$R = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{14.8}$$

Note that $R$ has been reduced. This implies that the three-dimensional Euclidean space is decomposed into a two-dimensional subspace ($xy$-plane) and a one-dimensional subspace ($z$-axis). The $xy$-coordinates are not mixed with the $z$-component after the rotation $R$. Note, however, that if the rotation axis is oblique against the $xy$-plane, this is not the case. We will come back to this point later.

Taking only the $xy$-coordinates in Fig. 14.3, we make a calculation. Using an addition theorem of trigonometric functions, we get

$$x' = r\cos(\theta+\alpha) = r(\cos\alpha\cos\theta - \sin\alpha\sin\theta) = x\cos\theta - y\sin\theta, \tag{14.9}$$

$$y' = r\sin(\theta+\alpha) = r(\sin\alpha\cos\theta + \cos\alpha\sin\theta) = y\cos\theta + x\sin\theta, \tag{14.10}$$

where we used $x = r\cos\alpha$ and $y = r\sin\alpha$. Combining (14.9) and (14.10), as a matrix form, we get

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \tag{14.11}$$

Equation (14.11) is the same as (9.31) and represents a transformation matrix of a rotation angle $\theta$ within the $xy$-plane. Whereas in Chap. 9, we considered this from the point of view of the transformation of basis vectors, here we deal with the

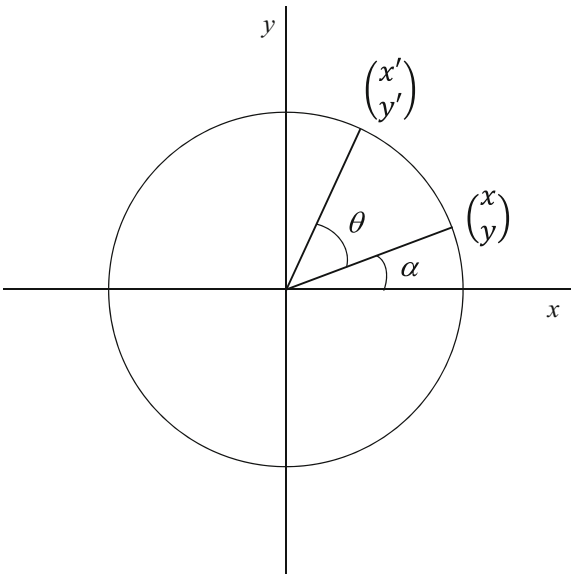


**Fig. 14.3** Transformation of the $xy$-coordinates by a $\theta$ rotation

transformations of coordinates in the fixed coordinate system. If $\theta \neq \pi$, off-diagonal elements do not vanish. Including the $z$-component, we get (14.8).

Let us summarize symmetry operations and their (3,3) matrix representations. The coordinates before and after a symmetry operation are expressed as

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \text{ and } \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}, \text{ respectively.}$$

(i) Identity transformation:
To leave a geometric object or a coordinate system unchanged (or unmoved), by convention, we denote it by a capital letter $E$. It is represented by a (3,3) identity matrix.

(ii) Rotation symmetry around a rotation axis:
Here a "proper" rotation is intended. We denote a rotation by a rotation axis and its magnitude (i.e., rotation angle). Thus, we have

$$R_{z\theta} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad R_{y\phi} = \begin{pmatrix} \cos\phi & 0 & \sin\phi \\ 0 & 1 & 0 \\ -\sin\phi & 0 & \cos\phi \end{pmatrix}, \quad R_{x\varphi} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{pmatrix}. \tag{14.12}$$

With $R_{y\phi}$, we first consider a following coordinate transformation:

$$\begin{pmatrix} z' \\ x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} z \\ x \\ y \end{pmatrix}. \tag{14.13}$$

This can easily be visualized in Fig. 14.4; consider that cyclic permutation of $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ produces $\begin{pmatrix} z \\ x \\ y \end{pmatrix}$. Shuffling the order of coordinates, we get

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cos\phi & 0 & \sin\phi \\ 0 & 1 & 0 \\ -\sin\phi & 0 & \cos\phi \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}. \tag{14.14}$$

By convention of the symmetry groups, the following notation $C_n$ is used to denote a rotation. A subscript $n$ of $C_n$ represents the order of the rotation axis. The order means the largest number of $n$ so that the rotation through $2\pi/n$ gives an equivalent configuration. Successive rotations of $m$ times ($m < n$) are denoted by

$$C_n^m.$$

If $m = n$, the successive rotations produce an equivalent configuration same as the beginning; i.e., $C_n^n = E$. The rotation angles $\theta$, $\varphi$, etc., used above are restricted to $2\pi m/n$ accordingly.

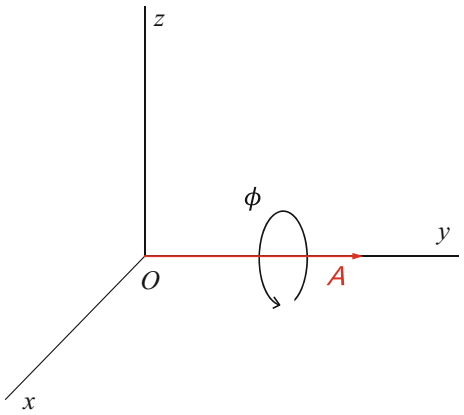


**Fig. 14.4** Rotation by $\phi$ around the $y$-axis

(iii) Mirror symmetry with respect to a plane of mirror symmetry:
We denote a mirror symmetry by a mirror symmetry plane; *xy*-plane, and *yz*-plane, etc. We have

$$M_{xy} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, M_{yz} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, M_{zx} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (14.15)$$

The mirror symmetry is usually denoted by $\sigma_v, \sigma_h$, and $\sigma_d$, whose subscripts stand for "vertical," "horizontal," and "dihedral," respectively. Among these symmetry operations, $\sigma_v$ and $\sigma_d$ include a rotation axis in the symmetry plane, while $\sigma_h$ is perpendicular to the rotation axis if such an axis exists. Notice that a group belonging to $C_s$ symmetry possesses only $E$ and $\sigma_h$. Although $\sigma_h$ can exist by itself as a mirror symmetry, neither $\sigma_v$ nor $\sigma_d$ can exist as a mirror symmetry by itself. We will come back to this point later.

(iv) Inversion symmetry with respect to a center of inversion:
We specify an inversion center if necessary; e.g., an origin of a coordinate system $O$.

$$I_O = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \quad (14.16)$$

Note that as obviously from the matrix form, $I_O$ is commutable with any other symmetry operations. Note also that $I_O$ can be expressed as successive symmetry operations or product of symmetry operations. For instance, we have

$$I_O = R_{z\pi} M_{xy} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \quad (14.17)$$

Note that $R_{z\pi}$ and $M_{xy}$ are commutable; i.e., $R_{z\pi} M_{xy} = M_{xy} R_{z\pi}$.

(iv) Improper rotation:
This is a combination of a proper rotation and a reflection by a mirror symmetry plane. That is, rotation around an axis is performed first and then reflection is carried out by a mirror plane that is perpendicular to the rotation axis. For instance, an improper rotation is expressed as

$$M_{xy}R_{z\theta} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & -1 \end{pmatrix}. \tag{14.18}$$

As mentioned just above, the inversion symmetry $I$ can be viewed as an improper rotation. Note that in this case, the reflection and rotation operations are commutable. However, we will follow a conventional custom that considers the inversion symmetry as an independent symmetry operation. Readers may well wonder why we need to consider the improper rotation. The answer is simple; it solely rests upon the axiom (A1) of the group theory. A group must be closed with respect to the multiplication.

The improper rotations are usually denoted by $S_n$. A subscript $n$ again stands for an order of rotation.

## 14.2 Successive Symmetry Operations

Let us now consider successive reflections in different planes and successive rotations about different axes [2]. Figure 14.5a displays two reflections with respect to the planes $\sigma$ and $\tilde{\sigma}$ both perpendicular to the $xy$-plane. The said planes make a dihedral angle $\theta$ with their intersection line identical to the $z$-axis. Also, the plane $\sigma$ is identical with the $zx$-plane. Suppose that an arrow lies on the $xy$-plane perpendicularly to the $zx$-plane. As in (14.15), an operation $\sigma$ is represented as

$$\sigma = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{14.19}$$

To determine a matrix representation of $\tilde{\sigma}$, we calculate a matrix again as in the above case. As a result, we have

$$\tilde{\sigma} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \cos 2\theta & \sin 2\theta & 0 \\ \sin 2\theta & -\cos 2\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{14.20}$$

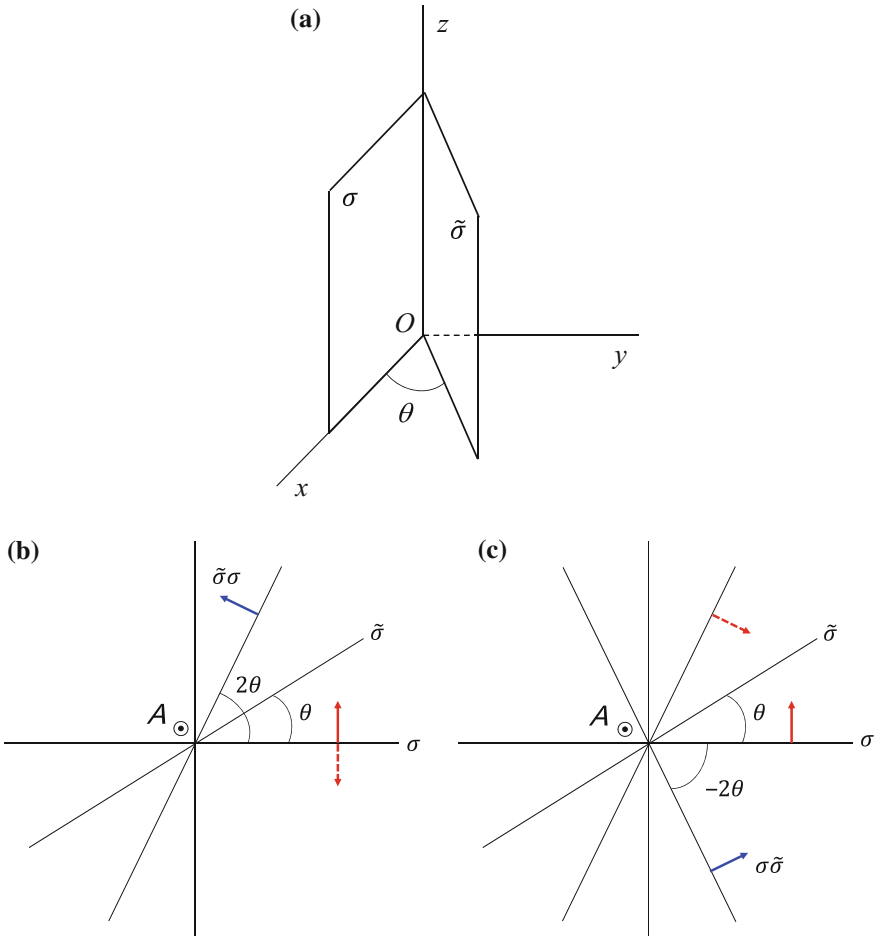


**Fig. 14.5** Successive two reflections about two planes $\sigma$ and $\tilde{\sigma}$ that make an angle $\theta$. **a** Reflections $\sigma$ and $\tilde{\sigma}$ with respect to two planes. **b** Successive operations of $\sigma$ and $\tilde{\sigma}$ in this order. The combined operation is denoted by $\tilde{\sigma}\sigma$. The operations result in a $2\theta$ rotation around the $z$-axis. **c** Successive operations of $\tilde{\sigma}$ and $\sigma$ in this order. The combined operation is denoted by $\sigma\tilde{\sigma}$. The operations result in a $-2\theta$ rotation around the $z$-axis

Notice that this matrix representation is referred to the original $xyz$-coordinate system; see discussion of Sect. 9.4. Hence, we describe the successive transformations $\sigma$ followed by $\tilde{\sigma}$ as

$$\tilde{\sigma}\sigma = \begin{pmatrix} \cos 2\theta & -\sin 2\theta & 0 \\ \sin 2\theta & \cos 2\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{14.21}$$

The expression (14.21) means that the multiplication should be done first by $\sigma$ and then by $\tilde{\sigma}$; see Fig. 14.5b. Note that $\sigma$ and $\tilde{\sigma}$ are conjugate to each other (Sect. 13.3). In this case, the combined operations produce a $2\theta$ rotation around the $z$-axis.

If, on the other hand, the multiplication is made first by $\tilde{\sigma}$ and then by $\sigma$, we have a $-2\theta$ rotation around the $z$-axis; see Fig. 14.5c. As a matrix representation, we have

$$\sigma\tilde{\sigma} = \begin{pmatrix} \cos 2\theta & \sin 2\theta & 0 \\ -\sin 2\theta & \cos 2\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{14.22}$$

Thus, successive operations of reflection by the planes that make a dihedral angle $\theta$ yield a rotation $\pm 2\theta$ around the $z$-axis (i.e., the intersection line of $\sigma$, and, $\tilde{\sigma}$). The operation $\sigma\tilde{\sigma}$ is an inverse to $\tilde{\sigma}\sigma$. That is

$$(\sigma\tilde{\sigma})(\tilde{\sigma}\sigma) = E. \tag{14.23}$$

We have $\det \sigma\tilde{\sigma} = \det \tilde{\sigma}\sigma = 1$.

Meanwhile, putting

$$R_{2\theta} = \begin{pmatrix} \cos 2\theta & -\sin 2\theta & 0 \\ \sin 2\theta & \cos 2\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{14.24}$$

we have

$$\tilde{\sigma}\sigma = R_{2\theta} \quad \text{or} \quad \tilde{\sigma} = R_{2\theta}\sigma. \tag{14.25}$$

This implies the following: Suppose a plane $\sigma$ and a straight line on it. Also, suppose that one first makes a reflection about $\sigma$ and then makes a $2\theta$ rotation around the said straight line. Then, the resulting transformation is equivalent to a reflection about $\tilde{\sigma}$ that makes an angle $\theta$ with $\sigma$. At the same time, the said straight line is an intersection line of $\sigma$ and $\tilde{\sigma}$. Note that a dihedral angle between the two planes $\sigma$ and $\tilde{\sigma}$ is half an angle of the rotation. Thus, any two of the symmetry operations related to (14.25) are mutually dependent; any two of them produce the third symmetry operation.

In the above illustration, we did not take account of the presence of a symmetry axis. If the aforementioned axis is a symmetry axis $C_n$, we must have

$$2\theta = 2\pi/n \quad \text{or} \quad n = \pi/\theta. \tag{14.26}$$

From a symmetry requirement, there should be $n$ planes of mirror symmetry in combination with the $C_n$ axis. Moreover, an intersection line of these $n$ mirror symmetry planes should coincide with that $C_n$ axis. This can be seen as various $C_{nv}$ groups.

Next we consider another successive symmetry operations. Suppose that there are two $C_2$ axes ($C_{x\pi}$ and $C_{a\pi}$ as shown) that intersect at an angle $\theta$ (Fig. 14.6). There $C_{x\pi}$ is identical to the $x$-axis and the other ($C_{a\pi}$) lies on the $xy$-plane making an angle $\theta$ with $C_{x\pi}$. Following procedures similar to the above, we have matrix representations of the successive $C_2$ operations in reference to the $xyz$-system such that

$$
\begin{aligned}
C_{x\pi} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \\
C_{a\pi} &= \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} \cos 2\theta & \sin 2\theta & 0 \\ \sin 2\theta & -\cos 2\theta & 0 \\ 0 & 0 & -1 \end{pmatrix}.
\end{aligned}
\tag{14.27}
$$

where $C_{a\pi}$ can be calculated similarly to (14.20). Again, we get

$$
\begin{aligned}
C_{a\pi}C_{x\pi} &= \begin{pmatrix} \cos 2\theta & -\sin 2\theta & 0 \\ \sin 2\theta & \cos 2\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}, \\
C_{x\pi}C_{a\pi} &= \begin{pmatrix} \cos 2\theta & \sin 2\theta & 0 \\ -\sin 2\theta & \cos 2\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.
\end{aligned}
\tag{14.28}
$$

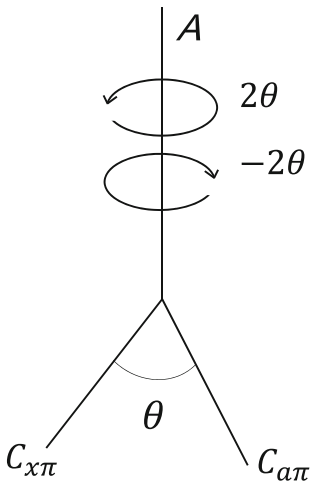


**Fig. 14.6** Successive two $\pi$ rotations around two $C_2$ axes of $C_{x\pi}$, and, $C_{a\pi}$ that intersect at an angle $\theta$. Of these, $C_{x\pi}$ is identical to the $x$-axis (not shown)

Notice that $(C_{x\pi}C_{a\pi})(C_{a\pi}C_{x\pi}) = E$. Once again putting

$$R_{2\theta} = \begin{pmatrix} \cos 2\theta & -\sin 2\theta & 0 \\ \sin 2\theta & \cos 2\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{14.29}$$

we have [1, 2]

$$C_{a\pi}C_{x\pi} = R_{2\theta} \quad \text{or} \quad C_{a\pi} = R_{2\theta}C_{x\pi}. \tag{14.30}$$

Note that in the above two illustrations for the successive symmetry operations, both the relevant operators have been represented in reference to the *original xyz*-system. For this reason, the latter operation was done from the *left* in (14.28).

From a symmetry requirement, once again the aforementioned $C_2$ axes must be present in combination with the $C_n$ axis. Moreover, those $C_2$ axes should be perpendicular to the $C_n$ axis. This can be seen in various $D_n$ groups.

Another illustration of successive symmetry operations is an improper rotation. If the rotation angle is $\pi$, this causes an inversion symmetry. In this illustration, reflection, rotation, and inversion symmetries coexist. A $C_{2h}$ symmetry is a typical example.

Equations (14.21), (14.22), and (14.29) demonstrate the same relation. Namely, two successive mirror symmetry operations about a couple of planes and two successive $\pi$-rotations about a couple of $C_2$ axes cause the same effect with regard to the geometric transformation. In this relation, we emphasize that two successive reflection operations make a determinant of relevant matrices 1. These aspects cause an interesting effect, and we will briefly discuss it in relation to $O$ and $T_d$ groups.

If furthermore the above-mentioned mirror symmetry planes and $C_2$ axes coexist, the symmetry planes coincide with or bisect the $C_2$ axes and vice versa. If these were not the case, another mirror plane or $C_2$ axis would be generated from the symmetry requirement and the newly generated plane or axis would be coupled with the original plane or axis. From the above argument, these processes again produce another $C_n$ axis. That must be prohibited.

Next, suppose that a $C_n$ axis intersects *obliquely* with a plane of mirror symmetry. A rotation of $2\pi/n$ around such an axis produces another mirror symmetry plane. This newly generated plane intersects with the original mirror plane and produces a different $C_n$ axis according to the above discussion. Thus, in this situation, a mirror symmetry plane cannot coexist with a sole rotation axis. In a geometric object with higher symmetry such as $O_h$, however, several mirror symmetry planes can coexist with several rotation axes in such a way that the axes intersect with the mirror planes *obliquely*. In case, the $C_n$ axis intersects *perpendicularly* to a mirror symmetry plane that plane can coexist with a sole rotation axis (see Fig. 14.7). This is actually the case with a geometric object having a $C_{2h}$ symmetry. The mirror symmetry plane is denoted by $\sigma_h$.

Now, let us examine simple examples of molecules and associated symmetry operations.

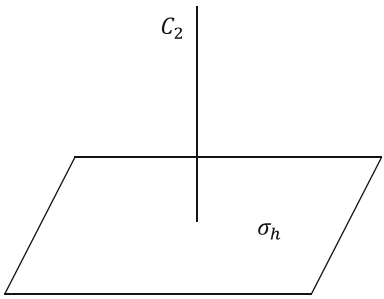


**Fig. 14.7** Mirror symmetry plane $\sigma_h$ and a sole rotation axis perpendicular to it

*Example 14.2* Figure 14.8 shows chemical structural formulae of thiophene, bithiophene, biphenyl, and naphthalene. These molecules belong to $C_{2v}$, $C_{2h}$, $D_2$, and $D_{2h}$, respectively. Note that these symbols are normally used to show specific point groups. Notice also that in biphenyl two benzene rings are twisted relative to the molecular axis. As an example, a multiplication table is shown in Table 14.1 for a $C_{2v}$ group. Table 14.1 clearly demonstrates that the group constitution of $C_{2v}$ differs from that of the group appearing in Example 13.1 (iii), even though the order is four for both the case. Similar tables are given with $C_{2h}$ and $D_2$. This is left for readers as an exercise. We will find that the multiplication tables of these groups have the same structure and that $C_{2v}, C_{2h}$, and, $D_2$ are all isomorphic to one another as a four group. Table 14.2 gives matrix representation of symmetry operations for $C_{2v}$. The representation is defined as the transformation by the symmetry operations of a set of basis vectors $(xyz)$ in $\mathbb{R}^3$.

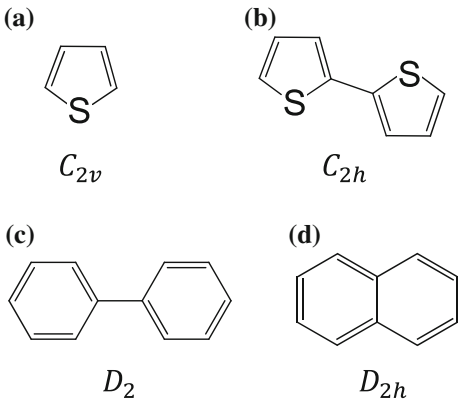


**Fig. 14.8** Chemical structural formulae and point groups of **a** thiophene, **b** bithiophene, **c** biphenyl, and **d** naphthalene

**Table 14.1** Multiplication table of $C_{2v}$

| $C_{2v}$ | $E$ | $C_2(z)$ | $\sigma_v(zx)$ | $\sigma'_v(yz)$ |
|---|---|---|---|---|
| $E$ | $E$ | $C_2$ | $\sigma_v$ | $\sigma'_v$ |
| $C_2(z)$ | $C_2$ | $E$ | $\sigma'_v$ | $\sigma_v$ |
| $\sigma_v(zx)$ | $\sigma_v$ | $\sigma'_v$ | $E$ | $C_2$ |
| $\sigma'_v(yz)$ | $\sigma'_v$ | $\sigma_v$ | $C_2$ | $E$ |

**Table 14.2** Matrix representation of symmetry operations for $C_{2v}$

| | $E$ | $C_2$ (around $z$-axis) | $\sigma_v(zx)$ | $\sigma_v'(yz)$ |
|---|---|---|---|---|
| Matrix | $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ | $\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ | $\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ | $\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ |

Meanwhile, Table 14.3 gives the multiplication table of $D_{2h}$. We recognize that the multiplication table of $C_{2v}$ appears on upper left and lower right blocks. If we suitably rearrange the order of group elements, we can make another multiplication table so that, e.g., $C_{2h}$ may appear on upper left and lower right blocks. As in the case of Table 14.2, Table 14.4 summarizes the matrix representation of symmetry operations for $D_{2h}$. There are eight group elements; i.e., an identity, an inversion, three mutually perpendicular $C_2$ axes and three mutually perpendicular planes of mirror symmetry ($\sigma$). Here, we consider a possibility of constructing subgroups of $D_{2h}$. The order of the subgroups must be a divisor of eight, and so let us list subgroups whose order is four and examine how many subgroups exist. We have ${}_8\mathrm{C}_4 = 70$ combinations, but those allowed should be restricted from the requirement of forming a group. This is because all the groups must contain identity element, and so the number allowed is equal to or no greater than ${}_7\mathrm{C}_3 = 35$.

(i) In light of the aforementioned discussion, two $C_2$ axes mutually intersecting at $\pi/2$ yield another $C_2$ axis around the normal to a plane defined by the intersecting axes. Thus, three $C_2$ axes have been chosen and a $D_2$ symmetry results. In this case, we have only one choice. (ii) In the case of $C_{2v}$, two planes mutually intersecting at $\pi/2$ yield a $C_2$ axis around their line of intersection. There are three possibilities of choosing two axes out of three (i.e., ${}_3\mathrm{C}_2 = 3$). (iii) If we choose the inversion (i) along with, e.g., one of the three $C_2$ axes, a $\sigma$ necessarily results. This is also the case when we first combine a $\sigma$ with $i$ to obtain a $C_2$ axis. We have three possibilities (i.e., ${}_3\mathrm{C}_1 = 3$) as well. Thus, we have only seven choices to construct subgroups of $D_{2h}$ having an order of four. This is summarized in Table 14.5.

**Table 14.3** Multiplication table of $D_{2h}$

| $D_{2h}$ | $E$ | $C_2(z)$ | $\sigma_v(zx)$ | $\sigma_v'(yz)$ | $i$ | $\sigma_v''(xy)$ | $C_2'(y)$ | $C_2''(x)$ |
|---|---|---|---|---|---|---|---|---|
| $E$ | $E$ | $C_2$ | $\sigma_v$ | $\sigma_v'$ | $i$ | $\sigma_v''$ | $C_2'$ | $C_2''$ |
| $C_2(z)$ | $C_2$ | $E$ | $\sigma_v'$ | $\sigma_v$ | $\sigma_v''$ | $i$ | $C_2''$ | $C_2'$ |
| $\sigma_v(zx)$ | $\sigma_v$ | $\sigma_v'$ | $E$ | $C_2$ | $C_2'$ | $C_2''$ | $i$ | $\sigma_v''$ |
| $\sigma_v'(yz)$ | $\sigma_v'$ | $\sigma_v$ | $C_2$ | $E$ | $C_2''$ | $C_2'$ | $\sigma_v''$ | $i$ |
| $i$ | $i$ | $\sigma_v''$ | $C_2'$ | $C_2''$ | $E$ | $C_2$ | $\sigma_v$ | $\sigma_v'$ |
| $\sigma_v''(xy)$ | $\sigma_v''$ | $i$ | $C_2''$ | $C_2'$ | $C_2$ | $E$ | $\sigma_v'$ | $\sigma_v$ |
| $C_2'(y)$ | $C_2'$ | $C_2''$ | $i$ | $\sigma_v''$ | $\sigma_v$ | $\sigma_v'$ | $E$ | $C_2$ |
| $C_2''(x)$ | $C_2''$ | $C_2'$ | $\sigma_v''$ | $i$ | $\sigma_v'$ | $\sigma_v$ | $C_2$ | $E$ |

**Table 14.4** Matrix representation of symmetry operations for $D_{2h}$

| $D_{2h}$ | $E$ | $C_2(z)$ | $C_2(y)$ | $C_2(x)$ | $i$ | $\sigma(xy)$ | $\sigma(zx)$ | $\sigma(yz)$ |
|---|---|---|---|---|---|---|---|---|
| Matrix | $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ | $\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ | $\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$ | $\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$ | $\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$ | $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$ | $\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ | $\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ |

**Table 14.5** Choice of symmetry operations for construction of subgroups of $D_{2h}$

| Subgroup | $E$ | $C_2(z)$ | $\sigma$ | $i$ | Choice |
|---|---|---|---|---|---|
| $D_2$ | 1 | 3 | 0 | 0 | 1 |
| $C_{2v}$ | 1 | 1 | 2 | 0 | 3 |
| $C_{2h}$ | 1 | 1 | 1 | 1 | 3 |

An inverse of any element is that element itself. Therefore, if with any above subgroup one chooses any element out of the remaining four elements and combines it with the identity, one can construct a subgroup $C_s$, $C_2$, or $C_i$ of an order of 2. Since all those subgroups of an order of 4 and 2 are commutative with $D_{2h}$, these subgroups are invariant subgroups. Thus in terms of a direct-product group, $D_{2h}$ can be expressed as various direct factors. Conversely, we can construct factor groups from coset decomposition. For instance, we have

$$D_{2h}/C_{2v} \cong C_s, D_{2h}/C_{2v} \cong C_2, D_{2h}/C_{2v} \cong C_i. \tag{14.31}$$

In turn, we express direct-product groups as, e.g.,

$$D_{2h} = C_{2v} \times C_s, D_{2h} = C_{2v} \times C_2, D_{2h} = C_{2v} \times C_i.$$

*Example 14.3* Figure 14.9 shows an equilateral triangle placed on the $xy$-plane of a three-dimensional Cartesian coordinate. An orthonormal basis $\boldsymbol{e}_1$, and, $\boldsymbol{e}_2$ are designated as shown. As for a chemical species of molecules, we have, e.g., boron trifluoride ($BF_3$). In the molecule, a boron atom is positioned at a molecular center with three fluorine atoms located at vertices of the equilateral triangle. The boron atom and fluorine atoms form a planar molecule (Fig. 14.10).

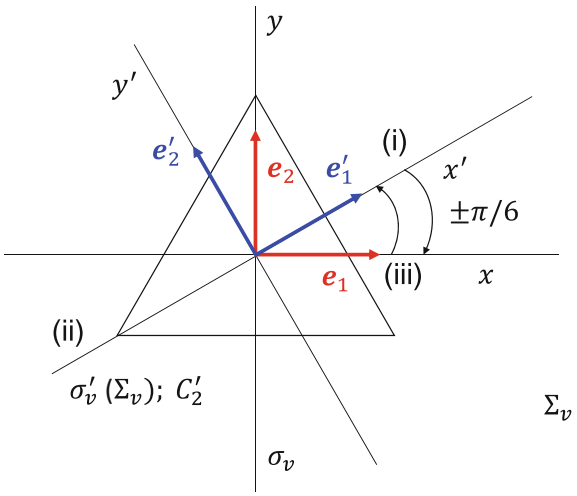


**Fig. 14.9** Equilateral triangle placed on the $xy$-plane. Several symmetry operations of $D_{3h}$ are shown. We consider the successive operations (i), (ii), and (iii) to represent $\sigma_v'$ in reference to the $xy$-coordinate system (see text)

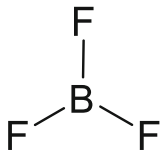


**Fig. 14.10** Boron trifluoride ($BF_3$) belonging to a $D_{3h}$ point group

A symmetry group belonging to $D_{3h}$ comprises twelve symmetry operations such that

$$D_{3h} = \left\{E, C_3, C_3', C_2, C_2', C_2'', \sigma_h, S_3, S_3', \sigma_v, \sigma_v', \sigma_v''\right\}, \tag{14.32}$$

where symmetry operations of the same species but distinct operation are denoted by a "prime" or "double prime." When we represent these operations by a matrix, it is straightforward in most cases. For instance, a matrix for $\sigma_v$ is given by $M_{yz}$ of (14.15). However, we should make some matrix calculations about $C_2', C_2'', \sigma_v', \text{and}, \sigma_v''$.

To determine a matrix representation of, e.g., $\sigma_v'$ in reference to the $xy$-coordinate system with orthonormal basis vectors $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$, we consider the $x'y'$-coordinate system with orthonormal basis vectors $\boldsymbol{e}_1', \text{and}, \boldsymbol{e}_2'$ (see Fig. 14.9). A transformation matrix between the two set of basis vectors is represented by

$$\left(\boldsymbol{e}_1'\boldsymbol{e}_2'\boldsymbol{e}_3'\right) = (\boldsymbol{e}_1\boldsymbol{e}_2\boldsymbol{e}_3)R_{z\frac{\pi}{6}} = (\boldsymbol{e}_1\boldsymbol{e}_2\boldsymbol{e}_3)\begin{pmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ \frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{14.33}$$

This representation corresponds to (9.69). Let $\Sigma_v$ be a reflection with respect to the $z'x'$-plane. This is *the same* operation as $\sigma_v'$. However, a matrix representation is *different*. This is because $\Sigma_v$ is represented in reference to the $x'y'z'$-system, while $\sigma_v'$ is in reference to $xyz$-system. The matrix representation of $\Sigma_v$ is simple and expressed as

$$\Sigma_v = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{14.34}$$

Referring to (9.80), we have

$$R_{z\frac{\pi}{6}}\Sigma_v = \sigma_v' R_{z\frac{\pi}{6}} \quad \text{or} \quad \sigma_v' = R_{z\frac{\pi}{6}}\Sigma_v\left[R_{z\frac{\pi}{6}}\right]^{-1}. \tag{14.35}$$

Thus, we see that in the first equation of (14.35) the order of multiplications are reversed according as the latter operation is expressed in the $xyz$-system or in the $x'y'z'$-system. As a full matrix representation, we get

$$\sigma_v' = \begin{pmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ \frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} & 0 \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{14.36}$$

Notice that this matrix representation is referred to the original *xyz*-coordinate system as before. Graphically, (14.35) corresponds to the multiplication of the symmetry operations done in the order of (i) $-\pi/6$ rotation, (ii) reflection (denoted by $\Sigma_v$), and (iii) $\pi/6$ rotation (Fig. 14.9). The associated matrices are multiplied from the left.

Similarly, with $C_2'$, we have

$$C_2' = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & -1 \end{pmatrix}. \tag{14.37}$$

The matrix form of (14.37) can also be decided in a manner similar to the above according to three successive operations shown in Fig. 14.9. In terms of classes, $\sigma_v, \sigma_v', \text{and}, \sigma_v''$ form a conjugacy class and $C_2, C_2', \text{and}, C_2''$ form another conjugacy class. With regard to the reflection and rotation, we have $\det \sigma_v' = -1$ and $\det C_2 = 1$, respectively.

## 14.3 *O* and $T_d$ Groups

According to a geometric object (or a molecule) which has a higher symmetry, we have to deal with many symmetry operations and relationship between them. As an example, we consider $O$ and $T_d$ groups. Both the groups have 24 symmetry operations and are isomorphic.

Let us think of the group $O$ fist. We start with considering rotations of π/2 around *x*-, *y*-, and *z*-axes. The matrices representing these rotations are obtained from (14.12) to give

$$R_{x\frac{\pi}{2}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, R_{y\frac{\pi}{2}} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix}, R_{z\frac{\pi}{2}} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{14.38}$$

We continue multiplications of these matrices so that the matrices can make up a complete set (i.e., a closure). Counting over those matrices, we have 24 of them and they form a group termed $O$. The group $O$ is a pure rotation group. Here, the pure rotation group is defined as a group whose group elements only comprise proper rotations (with their determinant of 1). An example is shown in Fig. 14.11, where

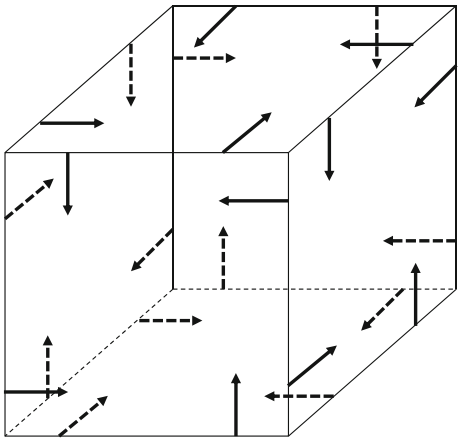

**Fig. 14.11** Cube whose individual vertices have three arrows for the cube not to possess mirror symmetries. This object belongs to a point group $O$ called pure rotation group

individual vertices of the cube have three arrows for the cube not to possess mirror symmetries or inversion symmetry. The group $O$ has five conjugacy classes. Figure 14.12 summarizes them. Geometrical characteristics of individual classes are sketched as well. These classes are categorized by a *trace* of the matrix. This is because the trace is kept unchanged by a similarity transformation. (Remember that

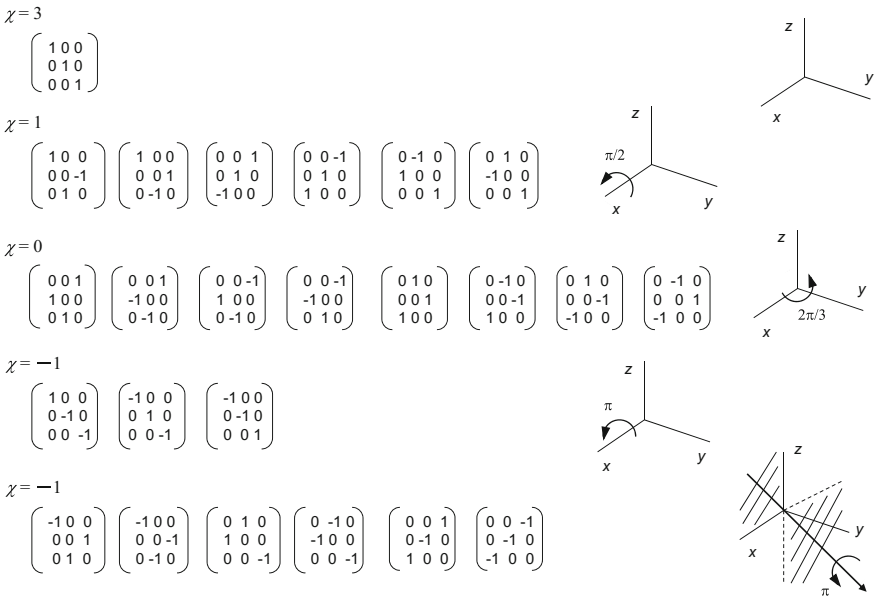


**Fig. 14.12** Point group $O$ and its five conjugacy classes. Geometrical characteristics of individual classes are briefly sketched

elements of the same conjugacy class are connected with a similarity transformation.) Having a look at the sketches, we notice that each operation switches the basis vectors $\boldsymbol{e}_1, \boldsymbol{e}_2, \text{and}, \boldsymbol{e}_3$; i.e., $x$-, $y$-, and $z$-axes. Therefore, the presence of diagonal elements (either 1 or $-1$) implies that the matrix takes the basis vector(s) as an eigenvector with respect to the rotation. Corresponding eigenvalue(s) are either 1 or $-1$ accordingly. This is expected from the fact that the matrix is an orthogonal matrix (i.e., unitary). The trace, namely a summation of diagonal elements, is closely related to the geometrical feature of the operation.

The operations of a $\pi$ rotation around the $x$-, $y$-, and $z$-axes and those of a $\pi$ rotation around an axis bisecting any two of three axes have a trace $-1$. The former operations take all the basis vectors as an eigenvectors; that is, all the diagonal elements are nonvanishing. With the latter operations, however, only one diagonal element is $-1$. This feature comes from that the bisected axes are switched by the rotation, whereas the remaining axis is reversed by the rotation.

Another characteristic is the generation of eight rotation axes that trisect the $x$-, $y$-, and $z$-axes, more specifically, a solid angle $\pi/2$ formed by the $x$-, $y$-, and $z$-axes. Since the rotation switches all the $x$-, $y$-, and $z$-axes, the trace is zero. At the same time, we find that this operation belongs to $C_3$. This operation is generated by successive two $\pi/2$ rotations around two mutually orthogonal axes. To inspect this situation more closely, we consider a conjugacy class of $\pi/2$ rotation that belongs to the $C_4$ symmetry and includes six elements, i.e., $R_{x\frac{\pi}{2}}$, $R_{y\frac{\pi}{2}}$, $R_{z\frac{\pi}{2}}$, $R_{\bar{x}\frac{\pi}{2}}$, $R_{\bar{y}\frac{\pi}{2}}$, and $R_{\bar{z}\frac{\pi}{2}}$. With these notations, e.g., $R_{x\frac{\pi}{2}}$ stands for a $\pi/2$ counterclockwise rotation around the $x$-axis; $R_{\bar{x}\frac{\pi}{2}}$ denotes a $\pi/2$ counterclockwise rotation around the —$x$-axis. Consequently, $R_{\bar{x}\frac{\pi}{2}}$ implies a $\pi/2$ *clockwise* rotation around the $x$-axis and, hence, an inverse element of $R_{x\frac{\pi}{2}}$. Namely, we have

$$R_{\bar{x}\frac{\pi}{2}} = \left(R_{x\frac{\pi}{2}}\right)^{-1}. \tag{14.39}$$

Now let us consider the successive two rotations. This is denoted by the multiplication of matrices that represent the related rotations. For instance, the multiplication of, e.g., $R_{x\frac{\pi}{2}}$ and $R'_{y\frac{\pi}{2}}$ produces the following:

$$R_{xyz\frac{2\pi}{3}} = R_{x\frac{\pi}{2}} R'_{y\frac{\pi}{2}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}. \tag{14.40}$$

In (14.40), we define $R_{xyz\frac{2\pi}{3}}$ as a $2\pi/3$ counterclockwise rotation around an axis that trisects the $x$-, $y$-, and $z$-axes. The prime " $'$ " of $R'_{y\frac{\pi}{2}}$ means that the operation is carried out in reference to the *new* coordinate system reached by the previous operation $R_{x\frac{\pi}{2}}$. For this reason, $R'_{y\frac{\pi}{2}}$ is operated (i.e., multiplied) from the *right* in (14.40). Compare this with the remark made just after (14.30). Changing the order of $R_{x\frac{\pi}{2}}$ and $R'_{y\frac{\pi}{2}}$, we have

$$R_{xyz\frac{2\pi}{3}} = R_{y\frac{\pi}{2}} R'_{x\frac{\pi}{2}} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & -1 \\ -1 & 0 & 0 \end{pmatrix}, \quad (14.41)$$

where $R_{xyz\frac{2\pi}{3}}$ is a $2\pi/3$ counterclockwise rotation around an axis that trisects the $x$-, $y$-, and —$z$-axes. Notice that we used $R'_{x\frac{\pi}{2}}$ this time, because it was performed after $R_{y\frac{\pi}{2}}$. Thus, we notice that there are eight related operations that trisect eight octants of the coordinate system. These operations are further categorized into four sets in which the two elements are an inverse element of each other. For instance, we have

$$R_{\bar{x}\bar{y}\bar{z}\frac{2\pi}{3}} = (R_{xyz\frac{2\pi}{3}})^{-1}. \quad (14.42)$$

Notice that a $2\pi/3$ counterclockwise rotation around an axis that trisects the —$x$-, —$y$-, and —$z$-axes is equivalent to a 2 $\pi/3$ clockwise rotation around an axis that trisects the $x$-, $y$-, and $z$-axes. Also, we have $R_{\bar{x}\bar{y}z\frac{2\pi}{3}} = (R_{xy\bar{z}\frac{2\pi}{3}})^{-1}$, etc. Moreover, we have "cyclic" relations such as

$$R_{x\frac{\pi}{2}} R'_{y\frac{\pi}{2}} = R_{y\frac{\pi}{2}} R'_{z\frac{\pi}{2}} = R_{z\frac{\pi}{2}} R'_{x\frac{\pi}{2}} = R_{xyz\frac{2\pi}{3}}. \quad (14.43)$$

Returning back to Sect. 9.4, we had

$$A[P(\boldsymbol{x})] = [(\boldsymbol{e}_1 \ldots \boldsymbol{e}_n) P A'] \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (\boldsymbol{e}_1 \ldots \boldsymbol{e}_n) \left[ A_O P \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right]. \quad (9.79)$$

Implication of (9.79) is that LHS is related to the transformation of basis vectors while retaining coordinates $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ and that transformation matrices should be operated on the basis vectors from the right. Meanwhile, RHS describes the transformation of coordinates, while retaining basis vectors. In that case, transformation matrices should be operated on the coordinates from the left. Thus, the order of operator multiplication is reversed. Following (9.80), we describe

$$R_{x\frac{\pi}{2}} R'_{y\frac{\pi}{2}} = R_O R_{x\frac{\pi}{2}}, \text{ i.e., } R_O = R_{x\frac{\pi}{2}} R'_{y\frac{\pi}{2}} (R_{x\frac{\pi}{2}})^{-1}, \quad (14.44)$$

where $R_O$ is viewed in reference to the original (or fixed) coordinate system and conjugated to $R'_{y\frac{\pi}{2}}$. Thus, we have

$$R_O = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (14.45)$$

Note that (14.45) is identical to a matrix representation of a $\pi/2$ rotation around the $z$-axis. This is evident from the fact that the $y$-axis is converted to the original $z$-axis by $R_{x\bar{z}}$; readers, imagine it.

We have two conjugacy classes of π rotation (the $C_2$ symmetry). One of them includes six elements, i.e., $R_{xy\pi}$, $R_{yz\pi}$, $R_{zx\pi}$, $R_{\bar{x}y\pi}$, $R_{\bar{y}z\pi}$, and $R_{\bar{z}x\pi}$. For these notations, a subscript, e.g., $xy$ stands for an axis that bisects the angle formed by $x$- and $y$-axes. A subscript $\bar{x}y$ denotes an axis bisecting the angle formed by —$x$- and $y$-axes. Another class includes three elements, i.e., $R_{x\pi}$, $R_{y\pi}$, and $R_{z\pi}$.

As for $R_{x\pi}$, $R_{y\pi}$, and $R_{z\pi}$, a combination of these operations should yield a $C_2$ rotation axis as discussed in Sect. 14.2. Of these three rotation axes, in fact, any two produce a $C_2$ rotation around the remaining axis, as is the case with naphthalene belonging to the $D_{2h}$ symmetry (see Sect. 14.2).

Regarding the class comprising six π rotation elements, a combination of, e.g., $R_{xy\pi}$ and $R_{\bar{x}y\pi}$ crossing each other at a right angle causes a related effect. For the other combinations, the two $C_2$ axes intersect each other at $\pi/3$; see Fig. 14.6 and put $\theta = \pi/3$ there. In this respect, elementary analytic geometry teaches the positional relationship among planes and straight lines. The argument is as follows: A plane determined by three points $\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix}$, $\begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix}$, and $\begin{pmatrix} x_3 \\ y_3 \\ z_3 \end{pmatrix}$ that do not sit on a line is expressed by a following equation:

$$\begin{vmatrix} x & y & z & 1 \\ x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \end{vmatrix} = 0. \tag{14.46}$$

Substituting $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$, and $\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$, we have

$$x - y + z = 0. \tag{14.47}$$

Taking account of direction cosines and using the Hesse's normal form, we get

$$\frac{1}{\sqrt{3}}(x - y + z) = 0, \tag{14.48}$$

where the normal to the plane expressed in (14.48) has direction cosines of $\frac{1}{\sqrt{3}}$, $-\frac{1}{\sqrt{3}}$, and $\frac{1}{\sqrt{3}}$ in relation to the $x$-, $y$-, and $z$-axes, respectively.

Therefore, the normal is given by a straight line connecting the origin and $\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$. In other words, a line connecting the origin and a corner of a cube is the

normal to the plane described by (14.48). That plane is formed by two intersecting lines, i.e., rotation axes of $C_2$ and $C_2'$ (see Fig. 14.13 that depicts a cube of each side of 2). These axes make an angle $\pi/3$; this can easily be checked by taking an inner product between $\begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$. These column vectors are two direction cosines of $C_2$ and $C_2'$. On the basis of the discussion of Sect. 14.2, we must have a rotation axis of $C_3$. That is, this axis trisects a solid angle $\pi/2$ shaped by three intersecting sides.

It is sometimes hard to visualize or envisage the positional relationship among planes and straight lines in three-dimensional space. It will therefore be useful to make a simple kit to help visualize it. Figure 14.14 gives an illustration.

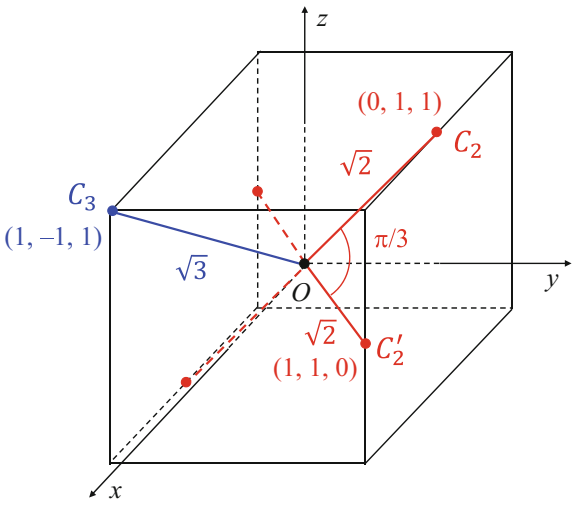


**Fig. 14.13** Rotation axes of $C_2$ and $C_2'$ along with another rotation axis $C_3$ in a point group $O$

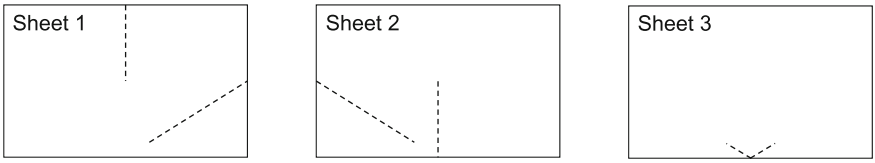


**Fig. 14.14** Simple kit that helps to visualize the positional relationship among planes and straight lines in three-dimensional space. To make it, follow next procedures: **a** Take three thick sheets of paper and make slits (dashed lines) as shown. **b** Insert Sheet 2 into Sheet 1 so that the two sheets can make a right angle. **c** Insert Sheet 3 into combined Sheets 2 and 3

Another typical example having 24 group elements is $T_d$. A molecule of methane belongs to this symmetry. Table 14.6 collects the relevant symmetry operations and their (3,3) matrix representations. As in the case of Fig. 14.12, the matrices show how a set of vectors $(x\,y\,z)$ are transformed according to the symmetry operations. Comparing it with Fig. 14.12, we immediately recognize that the close relationship between $T_d$ and $O$ exists and that these point groups share notable characteristics.

(i) Both $T_d$ and $O$ consist of five conjugacy classes, each of which contains the same number of symmetry species. (ii) Both $T_d$ and $O$ contain a pure rotation group $T$ as a subgroup. The subgroup $T$ consists of 12 group elements $E$, $8C_3$, and $3C_2$. Other remaining twelve group elements of $T_d$ are symmetry species related to reflection; $S_4$ and $\sigma_d$. The elements $6S_4$ and $6\sigma_d$ correspond to $6C_4$ and $6C_2$ of $O$, respectively. That is, successive operations of $S_4$ cause similar effects to those of $C_4$ of $O$. Meanwhile, successive operations of $\sigma_d$ are related to those of $6C_2$ of $O$.

Let us imagine in Fig. 14.15 that a regular tetrahedron is inscribed in a cube. As an example of symmetry operations, suppose that three pairs of planes of $\sigma_d$ (six planes in total) are given by equations of $x = \pm y$ and $y = \pm z$ and $z = \pm x$. Their Hesse's normal forms are represented as

$$\frac{1}{\sqrt{2}}(x \pm y) = 0, \tag{14.49}$$

$$\frac{1}{\sqrt{2}}(y \pm z) = 0, \tag{14.50}$$

$$\frac{1}{\sqrt{2}}(z \pm x) = 0. \tag{14.51}$$

Then, a dihedral angle $\alpha$ of the two planes is given by

$$\cos\alpha = \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} = \frac{1}{2} \text{ or} \tag{14.52}$$

$$\cos\alpha = \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} = 0. \tag{14.53}$$

That is, $\alpha = \pi/3$ or $\alpha = \pi/2$. On the basis of the discussion of Sect. 14.2, the intersection of the two planes must be a rotation axis of $C_3$ or $C_2$. Once again, in the case of $C_3$ the intersection is a straight line connecting the origin and a vertex of the cube. This can readily be verified as follows: For instance, two planes given by $x = y$ and $y = z$ make an angle $\pi/3$ and produce an intersection line $x = y = z$. This line, in turn, connects the origin and a vertex of the cube.

If we choose, e.g., two planes $x = \pm y$ from the above, these planes make a right angle and their intersection must be a $C_2$ axis. The three $C_2$ axes coincide with the $x$-, $y$-, and $z$-axes. In this light, $\sigma_d$ functions similarly to $6C_2$ of $O$ in that their

**Table 14.6** Symmetry operations and their matrix representation of $T_d$

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| $E$ | $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ | | | | | | | |
| $8C_3$ | $\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$ | $\begin{pmatrix} 0 & 0 & 1 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \end{pmatrix}$ | $\begin{pmatrix} 0 & 0 & -1 \\ 1 & 0 & 0 \\ 0 & -1 & 0 \end{pmatrix}$ | $\begin{pmatrix} 0 & 0 & -1 \\ -1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$ | $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$ | $\begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 1 & 0 & 0 \end{pmatrix}$ | $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & -1 \\ -1 & 0 & 0 \end{pmatrix}$ | $\begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & 1 \\ -1 & 0 & 0 \end{pmatrix}$ |
| $3C_2$ | $\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ | $\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$ | $\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$ | | | | | |
| $6S_4$ | $\begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$ | $\begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}$ | $\begin{pmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ -1 & 0 & 0 \end{pmatrix}$ | $\begin{pmatrix} 0 & 0 & -1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$ | $\begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$ | $\begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$ | | |
| $6\sigma_{\mathrm{d}}$ | $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$ | $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix}$ | $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$ | $\begin{pmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix}$ | $\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ | $\begin{pmatrix} 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ | | |

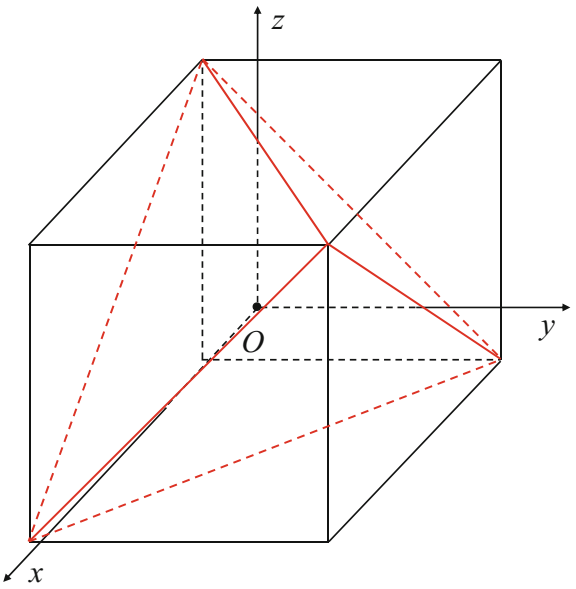


**Fig. 14.15** Regular tetrahedron inscribed in a cube. As an example of symmetry operations, we can choose three pairs of planes of $\sigma_d$ (six planes in total) given by equations of $x = \pm y$ and $y = \pm z$ and $z = \pm x$

combinations produce $8C_3$ or $3C_2$. Thus, constitution and operation of $T_d$ and $O$ are related.

Let us more closely inspect the structure and constitution of $O$ and $T_d$. First we construct mapping $\rho$ between group elements of $O$ and $T_d$ such that

$$\rho : g \in O \leftrightarrow g' \in T_d, \text{ if}, g \in T,$$

$$\rho : g \in O \leftrightarrow -(g')^{-1} \in T_d, \text{ if}, g \notin T.$$

In the above relation, the minus sign indicates that with an inverse representation matrix $R$ must be replaced with $-R$. Then, $\rho(g) = g'$ is an isomorphic mapping. In fact, comparing Fig. 14.12 and Table 14.6, $\rho$ gives identical matrix representation for $O$ and $T_d$. For example, taking the first matrix of $S_4$ of $T_d$, we have

$$-\begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}^{-1} = -\begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}.$$

The resulting matrix is identical to the first matrix of $C_4$ of $O$. Thus, we find $T_d$ and $O$ are isomorphic to each other.

Both $O$ and $T_d$ consist of 24 group elements and isomorphic to a *symmetric group* $S_4$; do not confuse it with the same symbol $S_4$ as a group element of $T_d$. The subgroup $T$ consists of three conjugacy classes $E$, $8C_3$, and $3C_2$. Since $T$ is constructed only by entire classes, it is an invariant subgroup; in this respect see the discussion of Sect. 13.3. The groups $O$, $T_d$, and $T$ along with $T_h$ and $O_h$ form cubic

groups [2]. In Table 14.7, we list these cubic groups together with their name and order. Of these, $O$ is a pure rotation subgroup of $O_h$ and $T$ is a pure rotation subgroup of $T_d$ and $T_h$.

Symmetry groups are related to permutations of $n$ elements (or objects or numbers). The permutation has already appeared in (9.56) when we defined a determinant of a matrix. That was defined as

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{pmatrix},$$

where $\sigma$ means the permutation of the numbers $1, 2, \cdots, n$. Therefore, the above symmetry group has $n!$ group elements (i.e., different ways of rearrangements). Although we do not dwell on symmetric groups much, we describe a following important theorem related to finite groups without proof. Interested readers are referred to literature [1].

**Theorem 14.1** (Cayley's Theorem [1]) *Every finite group $\mathcal{G}$ of order $n$ is isomorphic to a subgroup (containing a whole group) of the symmetric group $S_n$.*

## 14.4 Special Orthogonal Group $SO(3)$

In Part III and Part IV thus far, we have dealt with a broad class of linear transformations. Related groups are finite groups. Here we will describe characteristics of special orthogonal group $SO(3)$, a kind of infinite groups. The $SO(3)$ represents rotations in three-dimensional Euclidean space $\mathbb{R}^3$. Rotations are made around an axis (a line through the origin) with the origin fixed. The rotation is defined by an azimuth of direction and a magnitude (angle) of rotation. The azimuth of direction is defined by two parameters, and the magnitude is defined by one parameter, a rotation angle. Hence, the rotation is defined by three independent parameters. Since those parameters are continuously variable, $SO(3)$ is one of continuous groups.

Two rotations result in another rotation with the origin again fixed. A reverse rotation is unambiguously defined. An identity transformation is naturally defined. An associative law holds as well. Thus, the relevant rotations form a group, i.e., $SO(3)$. The rotation is represented by a real (3,3) matrix whose determinant is 1. A matrix

**Table 14.7** Several cubic groups and their characteristics

| Notation | Group name | Order | Remark |
|---|---|---|---|
| $T$ | Tetrahedral rotation | 12 | subgroup of $T_h$, $T_d$ |
| $T_h$, $T_d$ | Tetrahedral | 24 | |
| $O$ | Octahedral rotation | 24 | subgroup of $O_h$ |
| $O_h$ | Octahedral | 48 | |

representation is uniquely determined once an orthonormal basis is set in $\mathbb{R}^3$. Any rotation is represented by a rotation matrix accordingly. The rotation matrix $R$ is defined by

$$R^T R = RR^T = E, \tag{14.54}$$

where $R$ is a real matrix with $\det R = 1$. Notice that we exclude the case where $\det R = -1$. Matrices that satisfy (14.54) with $\det R = \pm 1$ are referred to as orthogonal matrices that cause orthogonal transformation. Correspondingly, orthogonal groups (represented by orthogonal matrices) contain rotation groups as a special case. In other words, the orthogonal groups contain the rotation groups as a subgroup. An orthogonal group in $\mathbb{R}^3$ denoted $O(3)$ contains $SO(3)$ as a subgroup. By the same token, orthogonal matrices contain rotation matrices as a special case.

In Sect. 14.2, we treated reflection and improper rotation with a determinant of their matrices being –1. In this section, these transformations are excluded and only rotations are dealt with. We focus on geometric characteristics of the rotation groups. Readers are referred to more detailed representation theory of $SO(3)$ in appropriate literature [1].

### 14.4.1 Rotation Axis and Rotation Matrix

In this section, we represent a vector such as $|x\rangle$. We start with showing that any rotation has a unique presence of a rotation axis. The rotation axis is defined by the following: Suppose that there is a rigid body with some point within the body fixed. Here the said rigid body can be that with infinite extent. Then the rigid body exerts rotation. The rotation axis is a line on which every point is unmoved during the rotation. As a matter of course, identity matrix $E$ has the linearly independent rotation axes. (Practically, this represents *no* rotation.)

**Theorem 14.2** *Any rotation matrix R is accompanied by at least one rotation axis. Unless the rotation matrix is identity, the rotation matrix should be accompanied by one and only one rotation axis.*

*Proof* As $R$ is an orthogonal matrix of a determinant 1, so are $R^T$ and $R^{-1}$. Then, we have

$$(R - E)^T = R^T - E = R^{-1} - E. \tag{14.55}$$

Hence, we get

$$\begin{aligned} &\det(R - E) = \det(R^T - E) = \det(R^{-1} - E) = \det[R^{-1}(E - R)] \\ &= \det(R^{-1})\det(E - R) = \det(E - R) = -\det(R - E). \end{aligned} \tag{14.56}$$

Note here that with any (3,3) matrix $A$ of $\mathbb{R}^3$,

$$\det A = -(-1)^9 \det A = -\det(-A).$$

This equality holds with $\mathbb{R}^n\,(n:\text{odd})$, but in $\mathbb{R}^n\,(n:\text{even})$, we have $\det A = \det(-A)$. Then, (14.56) results in a trivial equation 0 = 0 accordingly. Therefore, the discussion made below only applies to $\mathbb{R}^n\,(n:\text{odd})$. Thus, from (14.56), we have

$$\det(R-E) = 0. \tag{14.57}$$

This implies that for $^{\exists}|x_0\rangle \neq 0$

$$(R-E)|x_0\rangle = 0. \tag{14.58}$$

Therefore, we get

$$R(a|x_0\rangle) = a|x_0\rangle, \tag{14.59}$$

where $a$ is an arbitrarily chosen real number. In this case, an eigenvalue of $R$ is 1, which an eigenvector $a|x_0\rangle$ corresponds to. Thus, as a rotation axis, we have a straight line expressed as

$$l = \text{Span}\{a|x_0\rangle;\, a \in \mathbb{R}\}. \tag{14.60}$$

This proves the presence of a rotation axis.

Next suppose that there are two (or more) rotation axes. The presence of two rotation axes naturally implies that there are two linearly independent vectors (i.e., two straight lines that mutually intersect at the fixed point). Suppose that such vectors are $|u\rangle$ and $|v\rangle$. Then, we have

$$(R-E)|u\rangle = 0, \tag{14.61}$$

$$(R-E)|v\rangle = 0. \tag{14.62}$$

Let us consider a vector $^{\forall}|y\rangle$ that is chosen from $\text{Span}\{|u\rangle,|v\rangle\}$, to which we assume that

$$\text{Span}\{|u\rangle,|v\rangle\} \equiv \text{Span}\{a|u\rangle, b|v\rangle; a,b \in \mathbb{R}\}. \tag{14.63}$$

That is, $\text{Span}\{|u\rangle,|v\rangle\}$ represents a plane $P$ formed by two mutually intersecting straight lines. Then, we have

$$y = s|u\rangle + t|v\rangle, \tag{14.64}$$

where $s$ and $t$ are some real numbers. Operating $R - E$ on (14.64), we have

$$(R-E)|y\rangle = (R-E)(s|u\rangle + t|v\rangle) = s(R-E)|u\rangle + t(R-E)|v\rangle = 0. \quad (14.65)$$

This indicates that any vectors in $P$ can be an eigenvector of $R$, implying that an infinite number of rotation axes exist.

Now, take another vector $|w\rangle$ that is perpendicular to the plane $P$ (see Fig. 14.16). Let us consider an inner product $\langle y|Rw\rangle$. Since

$$R|u\rangle = |u\rangle, \langle u|R^\dagger = \langle u|R^T = \langle u|. \quad (14.66)$$

Similarly, we have

$$\langle v|R^T = \langle v|. \quad (14.67)$$

Therefore, using the relation (14.64), we get

$$\langle y|R^T = \langle y|. \quad (14.68)$$

Here we are dealing with real numbers, and hence, we have

$$R^\dagger = R^T. \quad (14.69)$$

Now we have

$$\langle y|Rw\rangle = \langle yR^T|Rw\rangle = \langle y|R^TRw\rangle = \langle y|Ew\rangle = \langle y|w\rangle = 0. \quad (14.70)$$

In (14.70), the second equality comes from the associative law; the third is due to (14.54). The last equality comes from that $|w\rangle$ is perpendicular to $P$.

From (14.70), we have

$$\langle y|(R-E)w\rangle = 0. \quad (14.71)$$

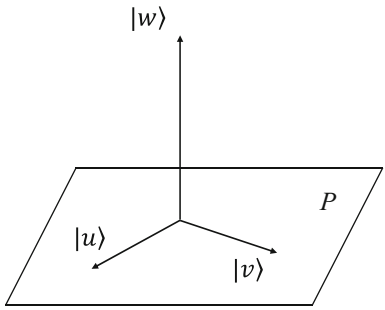


**Fig. 14.16** Plane $P$ formed by two mutually intersecting straight lines represented by $|u\rangle$ and $|v\rangle$. Another vector $|w\rangle$ is perpendicular to the plane $P$

However, we should be careful not to conclude immediately from (14.71) that $(R-E)|w\rangle = 0$; i.e., $R|w = |w\rangle$. This is because in (14.70) $|y\rangle$ does not represent all vectors in $\mathbb{R}^3$ but merely represent all the vectors in Span$\{|u\rangle, |v\rangle\}$. Nonetheless, both $|w\rangle$ and $|Rw\rangle$ are perpendicular to $P$, and so

$$|Rw\rangle = a|w\rangle, \tag{14.72}$$

where $a$ is an arbitrarily chosen real number. From (14.72), we have

$$\left\langle wR^{\dagger}|Rw\right\rangle = \langle wR^{T}|Rw\rangle = \langle w|w\rangle = |a|^2\langle w|w\rangle, \quad \text{i.e.} \quad a = \pm 1,$$

where the first equality comes from that $R$ is an orthogonal matrix. Since det $R$ = 1, $a$ = 1. Thus, from (14.72) this time around, we have $|Rw\rangle = |w\rangle$; that is,

$$(R-E)|w\rangle = 0. \tag{14.73}$$

Equations (14.65) and (14.73) imply that for any vector $|x\rangle$ arbitrarily chosen from $\mathbb{R}^3$, we have

$$(R-E)|x\rangle = 0. \tag{14.74}$$

Consequently, we get

$$R - E = 0 \quad \text{or} \quad R = E. \tag{14.75}$$

The above procedures represented by (14.61)–(14.75) indicate that the presence of two rotation axes necessarily requires a transformation matrix to be identity. This implies that all the vectors in $\mathbb{R}^3$ are an eigenvector of a rotation matrix.

Taking contraposition of the above, unless the rotation matrix is identity, the relevant rotation cannot have two rotation axes. Meanwhile, the proof of the former half ensures the presence at least one rotation axis. Consequently, any rotation is characterized by a unique rotation axis except for the identity transformation. This completes the proof.

An immediate consequence of Theorem 14.2 is that the rotation matrix should have an eigenvalue 1 which an eigenvector representing the rotation axis corresponds to. This statement includes a trivial fact that all the eigenvalues of the identity matrix are 1. In Sect. 12.4, we calculated eigenvalues of a two-dimensional rotation matrix. The eigenvalues were $e^{i\theta}$ or $e^{-i\theta}$, where $\theta$ is a rotation angle.

Let us consider rotation matrices that we dealt with in Sect. 14.1. The matrix representing the rotation around the $z$-axis by a rotation angle $\theta$ is expressed by

$$R = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{14.8}$$

In Sect. 12.4, we treated diagonalization of a rotation matrix. As $R$ is reduced, the diagonalization can be performed in a manner essentially the same as (12.92). That is, as a diagonalizing unitary matrix, we have

$$U = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{i}{\sqrt{2}} & \frac{i}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad U^\dagger = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & -\frac{i}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{14.76}$$

As a result of the unitary similarity transformation, we get

$$\begin{aligned} U^\dagger R U &= \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & -\frac{i}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{i}{\sqrt{2}} & \frac{i}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} e^{i\theta} & 0 & 0 \\ 0 & e^{-i\theta} & 0 \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned} \tag{14.77}$$

Thus, eigenvalues are 1, $e^{i\theta}$, and $e^{-i\theta}$. The eigenvalue 1 results from the existence of the unique rotation axis. When $\theta = 0$, (14.77) gives an identity matrix with all the eigenvalues 1 as expected. When $\theta = \pi$, eigenvalues are $-1$, $-1$, and 1. The eigenvalue 1 is again associated with the unique rotation axis. The (unitary) similarity transformation keeps a trace unchanged, that is, the trace $\chi$ is

$$\chi = 1 + 2\cos\theta. \tag{14.78}$$

As $R$ is a normal operator, spectral decomposition can be done as in the case of Example 12.1. Here we only show the result below.

$$R = e^{i\theta} \begin{pmatrix} \frac{1}{2} & \frac{i}{2} & 0 \\ -\frac{i}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{pmatrix} + e^{-i\theta} \begin{pmatrix} \frac{1}{2} & -\frac{i}{2} & 0 \\ \frac{i}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Three matrices of the above equation are projection operators.

### 14.4.2 Euler Angles and Related Topics

Euler angles are well known and have been being used in various fields of science. We wish to connect the above discussion with Euler angles.

In Part III, we dealt with successive linear transformations. This can be extended to the case of three or more successive transformations. Suppose that we have three successive transformation $R_1$, $R_2$, and $R_3$ and that the coordinate system

(a three-dimensional orthonormal basis) is transformed from O → I → II → III accordingly. The symbol "O" stands for the original coordinate system, and I, II, III represent successively transformed systems.

With the discussion that follows, let us denote the transformation by $R_2'$, $R_3'$, $R_3''$, etc., in reference to the coordinate system I. For example, $R_3'$ means that the third transformation is viewed from the system I. The transformation $R_3''$ indicates the third transformation which is viewed from the system II. That is, the number of primes " $'$ " denotes the number of the coordinate system to distinguish the systems I and II. Let $R_2$ (without prime) stand for the second transformation viewed from the system O.

Meanwhile, we have

$$R_1 R_2' = R_2 R_1. \tag{14.79}$$

This notation is in parallel to (9.80). Similarly, we have

$$R_2' R_3'' = R_3' R_2' \quad \text{and} \quad R_1 R_3' = R_3 R_1. \tag{14.80}$$

Therefore, we get [3]

$$R_1 R_2' R_3'' = R_1 R_3' R_2' = R_3 R_1 R_2' = R_3 R_2 R_1. \tag{14.81}$$

Also combining (14.79) and (14.80), we have

$$R_3'' = (R_2 R_1)^{-1} R_3 (R_2 R_1). \tag{14.82}$$

Let us call $R_2'$, $R_3''$, etc., a transformation on a "moving" coordinate system (i.e., the system I, II, III, $\cdots$). On the other hand, we call $R_1$, $R_2$, etc., a transformation on a "fixed" system (i.e., original coordinate system O). Thus, (14.81) shows that the multiplication order is reversed with respect to the moving system and fixed system [3].

For a practical purpose, it would be enough to consider three successive transformations. Let us think of, however, a general case where $n$ successive transformations are involved ($n$ denotes a positive integer). For the purpose of succinct notation, let us define the linear transformations and relevant coordinate systems as those in Fig. 14.17. Also, we define a following orthogonal transformation $R_j^{(i)}$:

$$R_j^{(i)} (0 \le i < j \le n) \quad \text{and} \quad R_i \equiv R_i^{(0)}, \tag{14.83}$$

where $R_j^{(i)}$ is defined as a transformation $R_j$ described in reference to the coordinate system $i$; $R_i^{(0)}$ means that $R_i$ is referred to the original coordinate system (i.e., the fixed coordinate). Then, we have

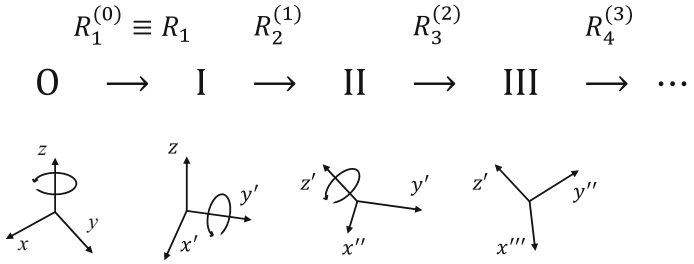


**Fig. 14.17** Successive orthogonal transformations and relevant coordinate systems

$$R_{i-1}^{(i-2)} R_k^{(i-1)} = R_k^{(i-2)} R_{i-1}^{(i-2)} \ (k > i-1). \tag{14.84}$$

Particularly, when $i = 3$

$$R_2^{(1)} R_k^{(2)} = R_k^{(1)} R_2^{(1)} \ (k > 2). \tag{14.85}$$

For $i = 2$ , we have

$$R_1 R_k^{(1)} = R_k R_1 \ (k > 1). \tag{14.86}$$

We define $n$ time successive transformations on a moving coordinate system as $\widetilde{R_n}$ such that

$$\widetilde{R_n} \equiv R_1 R_2^{(1)} R_3^{(2)} \cdots R_{n-2}^{(n-3)} R_{n-1}^{(n-2)} R_n^{(n-1)}. \tag{14.87}$$

Applying (14.84) on $R_{n-1}^{(n-2)} R_n^{(n-1)}$ and rewriting (14.87), we have

$$\widetilde{R_n} = R_1 R_2^{(1)} R_3^{(2)} \cdots R_{n-2}^{(n-3)} R_n^{(n-2)} R_{n-1}^{(n-2)}. \tag{14.88}$$

Applying (14.84) again on $R_{n-2}^{(n-3)} R_n^{(n-2)}$, we get

$$\widetilde{R_n} = R_1 R_2^{(1)} R_3^{(2)} \cdots R_n^{(n-3)} R_{n-2}^{(n-3)} R_{n-1}^{(n-2)}. \tag{14.89}$$

Proceeding similarly, we have

$$\widetilde{R_n} = R_1 R_n^{(1)} R_2^{(1)} R_3^{(2)} \cdots R_{n-2}^{(n-3)} R_{n-1}^{(n-2)} = R_n R_1 R_2^{(1)} R_3^{(2)} \cdots R_{n-2}^{(n-3)} R_{n-1}^{(n-2)}, \tag{14.90}$$

where with the last equality, we used (14.86). In this case, we have

$$R_1 R_n^{(1)} = R_n R_1. \tag{14.91}$$

To reach RHS of (14.90), we applied (14.84) $(n-1)$ times in total. Then we repeat the above procedures with respect to $R_{n-1}^{(n-2)}$ another $(n-2)$ times to get

$$\widetilde{R_n} = R_n R_{n-1} R_1 R_2^{(1)} R_3^{(2)} \cdots R_{n-2}^{(n-3)}. \tag{14.92}$$

Further proceeding similarly, we finally get

$$\widetilde{R_n} = R_n R_{n-1} R_{n-2} \cdots R_3 R_2 R_1. \tag{14.93}$$

In total, we have applied the permutation of (14.84) $n(n-1)/2$ times. When $n$ is 2, $n(n-1)/2 = 1$. This is the case with (14.79). If $n$ is 3, $n(n-1)/2 = 3$. This is the case with (14.81). Thus, (14.93) once again confirms that the multiplication order is reversed with respect to the moving system and fixed system.

Meanwhile, we define $\tilde{P}$ as

$$\tilde{P} \equiv R_1 R_2^{(1)} R_3^{(2)} \cdots R_{n-2}^{(n-3)} R_{n-1}^{(n-2)}. \tag{14.94}$$

Alternately, we describe

$$\tilde{P} = R_{n-1} R_{n-2} \cdots R_3 R_2 R_1. \tag{14.95}$$

Then, from (14.87) and (14.90), we get

$$\widetilde{R_n} = \tilde{P} R_n^{(n-1)} = R_n \tilde{P}. \tag{14.96}$$

Equivalently, we have

$$R_n = \tilde{P} R_n^{(n-1)} \tilde{P}^{-1} \quad \text{or} \tag{14.97}$$

$$R_n^{(n-1)} = \tilde{P}^{-1} R_n \tilde{P}. \tag{14.98}$$

Moreover, we have

$$\begin{aligned} \tilde{P}^T &= \left[R_1 R_2^{(1)} R_3^{(2)} \cdots R_{n-2}^{(n-3)} R_{n-1}^{(n-2)}\right]^T = \left[R_{n-1}^{(n-2)}\right]^T \left[R_{n-2}^{(n-3)}\right]^T \cdots \left[R_2^{(1)}\right]^T [R_1]^T \\ &= \left[R_{n-1}^{(n-2)}\right]^{-1} R \left[{}^{(n-3)}_{n-2}\right]^{-1} \cdots \left[R_2^{(1)}\right]^{-1} \left[R_1^{-1}\right] \\ &= \left[R_1 R_2^{(1)} R_3^{(2)} \cdots R_{n-2}^{(n-3)} R_{n-1}^{(n-2)}\right]^{-1} = \tilde{P}^{-1}. \end{aligned} \tag{14.99}$$

The third equality of (14.99) comes from the fact that matrices $R_{n-1}^{(n-2)}$, $R_{n-2}^{(n-3)}$, $R_2^{(1)}$, and $R_1$ are orthogonal matrices. Then, we get

$$\tilde{P}^T\tilde{P} = \tilde{P}\tilde{P}^T = E. \tag{14.100}$$

Thus, $\tilde{P}$ is an orthogonal matrix.

From a point of view of practical application, (14.97) and (14.98) are very useful. This is because $R_n$ and $R_n^{(n-1)}$ are conjugate to each other. Consequently, $R_n$ has the same eigenvalues and trace as $R_n^{(n-1)}$. In light of (9.81), we see that (14.97) and (14.98) relate $R_n$ (i.e., viewed in reference to the original coordinate system) to $R_n^{(n-1)}$ [i.e., the same transformation viewed in reference to the coordinate system reached after the $(n-1)$ transformations]. Since the transformation $R_n^{(n-1)}$ is usually described in a simple form, matrix calculations to compute $R_n$ can readily be done. Now let us consider an example.

*Example 14.4: Successive Rotations* A typical illustration of three successive transformation in moving coordinate systems is well known as Euler angles. This contains the following three steps:

(i) Rotation by $\alpha$ around the $z$-axis in the original coordinate system (O).
(ii) Rotation by $\beta$ around the $y'$-axis in the transferred coordinate system (I).
(iii) Rotation by $\gamma$ around the $z''$-axis (the same as $z'$-axis) in the transferred coordinate system (II).

Three steps are represented by matrices of $R_{z\alpha}$, $R'_{y'\beta}$, and $R''_{z''\gamma}$ in (14.12). That is, as a total transformation $\widetilde{R_3}$, we have

$$\begin{aligned}\widetilde{R_3} &= \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{pmatrix}\begin{pmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \cos\alpha\cos\beta\cos\gamma - \sin\alpha\sin\gamma & -\cos\alpha\cos\beta\sin\gamma - \sin\alpha\cos\gamma & \cos\alpha\sin\beta \\ \sin\alpha\cos\beta\cos\gamma + \cos\alpha\sin\gamma & -\sin\alpha\cos\beta\sin\gamma + \cos\alpha\cos\gamma & \sin\alpha\sin\beta \\ -\sin\beta\cos\gamma & \sin\beta\sin\gamma & \cos\beta \end{pmatrix}.\end{aligned} \tag{14.101}$$

This matrix corresponds to (14.87), where $n = 3$. The angles $\alpha$, $\beta$, and $\gamma$ in (14.101) are well known as Euler angles, and the associated matrix is widely used in quantum mechanics and related fields of natural science. The matrix notation, however, differs from literature to literature, and so care should be taken [3–5]. Using the notation of (14.87), we have

$$R_1 = \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{14.102}$$

$$R_2^{(1)} = \begin{pmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{pmatrix}, \tag{14.103}$$

$$R_3^{(2)} = \begin{pmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{14.104}$$

From (14.94) we also get

$$\tilde{P} \equiv R_1 R_2^{(1)} = \begin{pmatrix} \cos\alpha\cos\beta & -\sin\alpha & \cos\alpha\sin\beta \\ \sin\alpha\cos\beta & \cos\alpha & \sin\alpha\sin\beta \\ -\sin\beta & 0 & \cos\beta \end{pmatrix}. \tag{14.105}$$

Corresponding to (14.97), we have

$$R_3 = \tilde{P} R_3^{(2)} \tilde{P}^{-1} = R_1 R_2^{(1)} R_3^{(2)} \left[R_2^{(1)}\right]^{-1} R_1^{-1}. \tag{14.106}$$

Now matrix calculations are readily performed such that

$$R_3 = \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{pmatrix} \begin{pmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\times \begin{pmatrix} \cos\beta & 0 & -\sin\beta \\ 0 & 1 & 0 \\ \sin\beta & 0 & \cos\beta \end{pmatrix} \begin{pmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} (\cos^2\alpha\cos^2\beta + \sin^2\alpha)\cos\gamma & \cos\alpha\sin\alpha\sin^2\beta(1-\cos\gamma) & \cos\alpha\cos\beta\sin\beta(1-\cos\gamma) \\ \quad + \cos^2\alpha\sin^2\beta & \quad - \cos\beta\sin\gamma & \quad + \sin\alpha\sin\beta\sin\gamma \\ \cos\alpha\sin\alpha\sin^2\beta(1-\cos\gamma) & (\sin^2\alpha\cos^2\beta + \cos^2\alpha)\cos\gamma & \sin\alpha\cos\beta\sin\beta(1-\cos\gamma) \\ \quad + \cos\beta\sin\gamma & \quad + \sin^2\alpha\sin^2\beta & \quad - \cos\alpha\sin\beta\sin\gamma \\ \cos\beta\sin\beta\cos\alpha(1-\cos\gamma) & \sin\alpha\cos\beta\sin\beta(1-\cos\gamma) & \sin^2\beta\cos\gamma \\ \quad - \sin\alpha\sin\beta\sin\gamma & \quad + \cos\alpha\sin\beta\sin\gamma & \quad + \cos^2\beta \end{pmatrix} \tag{14.107}$$

Notice that in (14.107), we have a trace $\chi$ described as

$$\chi = 1 + 2\cos\gamma. \tag{14.108}$$

The trace is same as that of $R_3^{(2)}$, as expected from (14.106) and (14.97).

Equation (14.107) apparently seems complicated but has simple and well-defined meaning. The rotation represented by $R_3^{(2)}$ is characterized by a rotation

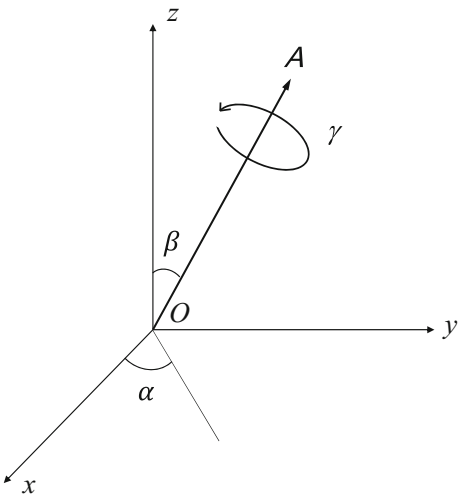


**Fig. 14.18** Rotation $\gamma$ around the rotation axis $A$. The orientation of $A$ is defined by angles $\alpha$ and $\beta$ as shown

by $\gamma$ around the $z''$-axis. Figure 14.18 represents the orientation of the $z''$-axis viewed in reference to the original $xyz$-system. That is, the $z''$-axis (identical to the rotation axis $A$) is designated by an azimuthal angle $\alpha$ and a zenithal angle $\beta$ as shown. The operation $R_3$ is represented by a rotation by $\gamma$ around the axis $A$ in the $xyz$-system. The angles $\alpha$, $\beta$, and $\gamma$ coincide with the Euler angles designated with the same independent parameters $\alpha$, $\beta$, and $\gamma$.

From (14.77) and (14.107), a diagonalizing matrix for $R_3$ is $\tilde{P}U$. That is,

$$U^\dagger \tilde{P}^{-1} R_3 \tilde{P} U = (\tilde{P}U)^\dagger R_3 \tilde{P} U = U^\dagger R_3^{(2)} U = \begin{pmatrix} e^{i\gamma} & 0 & 0 \\ 0 & e^{-i\gamma} & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{14.109}$$

Note that as $\tilde{P}$ is a real matrix, we have

$$\tilde{P}^\dagger = \tilde{P}^T = \tilde{P}^{-1} \tag{14.110}$$

and

$$\begin{aligned} \tilde{P}U &= \begin{pmatrix} \cos\alpha\cos\beta & -\sin\alpha & \cos\alpha\sin\beta \\ \sin\alpha\cos\beta & \cos\alpha & \sin\alpha\sin\beta \\ -\sin\beta & 0 & \cos\beta \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{i}{\sqrt{2}} & \frac{i}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{\sqrt{2}}\cos\alpha\cos\beta + \frac{i}{\sqrt{2}}\sin\alpha & \frac{1}{\sqrt{2}}\cos\alpha\cos\beta - \frac{i}{\sqrt{2}}\sin\alpha & \cos\alpha\sin\beta \\ \frac{1}{\sqrt{2}}\sin\alpha\cos\beta - \frac{i}{\sqrt{2}}\cos\alpha & \frac{1}{\sqrt{2}}\sin\alpha\cos\beta + \frac{i}{\sqrt{2}}\cos\alpha & \sin\alpha\sin\beta \\ -\frac{1}{\sqrt{2}}\sin\beta & -\frac{1}{\sqrt{2}}\sin\beta & \cos\beta \end{pmatrix}. \end{aligned} \tag{14.111}$$

A vector representing the rotation axis corresponds to an eigenvalue 1. The direction cosines of $x$-, $y$-, and $z$-components for the rotation axis $A$ are $\cos\alpha\sin\beta$, $\sin\alpha\sin\beta$, and $\cos\beta$ (see Fig. 3.1), respectively, when viewed in reference to the original $xyz$-coordinate system. This can directly be shown as follows: The characteristic equation of $R_3$ is expressed as

$$|R_3 - \lambda E| = 0. \tag{14.112}$$

Using (14.107), we have

$$|R_3 - \lambda E| = \begin{vmatrix} (\cos^2\alpha\cos^2\beta + \sin^2\alpha)\cos\gamma + \cos^2\alpha\sin^2\beta - \lambda & \cos\alpha\sin\alpha\sin^2\beta(1-\cos\gamma) - \cos\beta\sin\gamma & \cos\alpha\cos\beta\sin\beta(1-\cos\gamma) + \sin\alpha\sin\beta\sin\gamma \\ \cos\alpha\sin\alpha\sin^2\beta(1-\cos\gamma) + \cos\beta\sin\gamma & (\sin^2\alpha\cos^2\beta + \cos^2\alpha)\cos\gamma + \sin^2\alpha\sin^2\beta - \lambda & \sin\alpha\cos\beta\sin\beta(1-\cos\gamma) - \cos\alpha\sin\beta\sin\gamma \\ \cos\alpha\cos\beta\sin\beta(1-\cos\gamma) - \sin\alpha\sin\beta\sin\gamma & \sin\alpha\cos\beta\sin\beta(1-\cos\gamma) + \cos\alpha\sin\beta\sin\gamma & \sin^2\beta\cos\gamma + \cos^2\beta - \lambda \end{vmatrix} \tag{14.113}$$

When $\lambda = 1$, we must have the direction cosines of the rotation axis as an eigenvector. That is,

$$\begin{vmatrix} (\cos^2\beta\cos^2\alpha + \sin^2\alpha)\cos\gamma + \sin^2\beta\cos^2\alpha - 1 & \sin^2\beta\cos\alpha\sin\alpha(1-\cos\gamma) - \cos\beta\sin\gamma & \cos\beta\sin\beta\cos\alpha(1-\cos\gamma) + \sin\beta\sin\alpha\sin\gamma \\ \sin^2\beta\cos\alpha\sin\alpha(1-\cos\gamma) + \cos\beta\sin\gamma & (\cos^2\beta\sin^2\alpha + \cos^2\alpha)\cos\gamma + \sin^2\alpha\sin^2\beta - 1 & \sin\alpha\cos\beta\sin\beta(1-\cos\gamma) - \cos\alpha\sin\beta\sin\gamma \\ \cos\alpha\cos\beta\sin\beta(1-\cos\gamma) - \sin\alpha\sin\beta\sin\gamma & \sin\alpha\cos\beta\sin\beta(1-\cos\gamma) + \cos\alpha\sin\beta\sin\gamma & \sin^2\beta\cos\gamma + \cos^2\beta - 1 \end{vmatrix} \times \begin{pmatrix} \cos\alpha\sin\beta \\ \sin\alpha\sin\beta \\ \cos\beta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \tag{14.114}$$

The above matrix calculations certainly verify that (14.114) holds; readers, check it.

As an application of (14.107) to an illustrative example, let us consider a $2\pi/3$ rotation around an axis trisecting a solid angle $\pi/2$ formed by the $x$-, $y$-, and $z$-axes (see Fig. 14.19 and Sect. 14.3). Then we have

$$\begin{aligned} &\cos\alpha = \sin\alpha = 1/\sqrt{2}, \cos\beta = 1/\sqrt{3}, \sin\beta = 2/\sqrt{3}, \\ &\cos\gamma = -1/2, \sin\gamma = \sqrt{3}/2. \end{aligned} \tag{14.115}$$

Substituting (14.115) for (14.107), we get

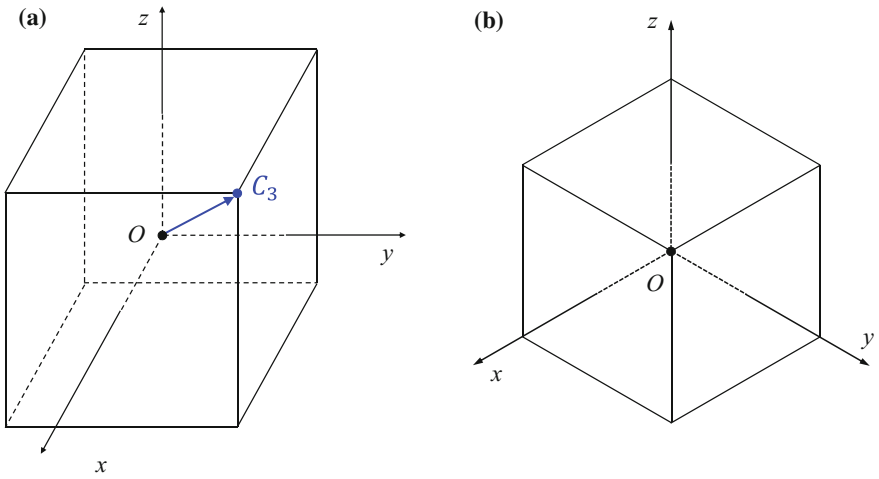


**Fig. 14.19** Rotation axis of $C_3$ that permutates basis vectors. **a** The $C_3$ axis is a straight line that connects the origin and each vertex of a cube. **b** Cube viewed along the $C_3$ axis that connects the origin and vertex

$$R_3 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}. \tag{14.116}$$

If we write a linear transformation $R_3$ following (9.37), we get

$$R_3(|x\rangle) = (|e_1\rangle\,|e_2\rangle\,|e_3\rangle) \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \tag{14.117}$$

where $|e_1\rangle, |e_2\rangle$, and $|e_3\rangle$ represent unit basis vectors in the direction of $x$-, $y$-, and $z$-axes, respectively. We have $|x\rangle = x_1|e_1\rangle + x_2|e_2\rangle + x_3|e_3\rangle$. Thus, we get

$$R_3(|x\rangle) = (|e_2\rangle|e_3\rangle|e_1\rangle) \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}. \tag{14.118}$$

This implies a cyclic permutation of the basis vectors. This is well-characterized by Fig. 14.19. Alternatively, (14.117) can be expressed in terms of the column vectors (i.e., coordinate) transformation as

$$R_3(|x\rangle) = (|e_1\rangle\,|e_2\rangle\,|e_3\rangle) \begin{pmatrix} x_3 \\ x_1 \\ x_2 \end{pmatrix}. \tag{14.119}$$

Care should be taken on which linear transformation is intended out of the basis vectors transformation or coordinate transformation.

As mentioned above, we have compared geometric features on the moving coordinate system and fixed coordinate system. The features apparently seem to differ at a glance but are essentially the same.

## References

1. Hamermesh M (1989) Group theory and its application to physical problems. Dover, New York
2. Cotton FA (1990) Chemical applications of group theory. Wiley, New York
3. Edmonds AR (1957) Angular momentum in quantum mechanics. Princeton University Press, Princeton
4. Chen JQ, Ping J, Wang F (2002) Group representation theory for physicists, 2nd edn. World Scientific, Singapore
5. Arfken GB, Weber HJ, Harris FE (2013) Mathematical methods for physicists, 7th edn. Academic Press, Waltham

# Chapter 15
# Representation Theory of Groups

Representation theory is an important pillar of the group theory. As we shall see soon, a word "representation" and its definition sound a bit daunting. To be short, however, we need "numbers" or "matrices" to do a mathematical calculation. Therefore, we may think of the representation as merely numbers and matrices. Individual representations have their dimension. If that dimension is one, we treat a representation as a number (real or complex). If the dimension is two or more, we are going to deal with matrices; in the case of the $n$-dimension, it is a $(n,n)$ square matrix. In this chapter, we focus on the representation theory of finite groups. In this case, we have an important theorem stating that a representation of any finite group can be converted to a unitary representation by a similarity transformation. That is, group elements of a finite group are represented by a unitary matrix. According to the dimension of the representation, we have the same number of basis vectors. Bearing these things firmly in mind, we can pretty easily understand this important notion of representation.

## 15.1 Definition of Representation

In Sect. 13.4, we dealt with various aspects of the mapping between group elements. Of these, we studied fundamental properties of isomorphism and homomorphism. In this section, we introduce the notion of representation of groups and study it. If we deal with a finite group consisting of $n$ elements, we describe it as $\mathscr{g} = \{g_1 \equiv e, g_2, \cdots, g_n\}$ as in the case of Sect. 13.1.

**Definition 15.1** Let $\mathscr{g} = \{g_v\}$ be a group comprising elements $g_v$. Suppose that a $(d,d)$ matrix $D(g_v)$ is given for each group element $g_v$. Suppose also that in correspondence with



S. Hotta, *Mathematical Physical Chemistry*,
https://doi.org/10.1007/978-981-10-7671-8_15

$$g_\mu g_\nu = g_\rho, \tag{15.1}$$

$$D(g_\mu)D(g_\nu) = D(g_\rho) \tag{15.2}$$

holds. Then a set $\mathcal{L}$ consisting of $D(g_\nu)$, that is,

$$\mathcal{L} = \{D(g_\nu)\}$$

is said to be a representation.

Although the definition seems somewhat daunting, the representation is, as already seen, merely a homomorphism.

We call individual matrices $D(g_\nu)$ a representation matrix. A dimension $d$ of the matrix is said to be a dimension of representation as well. In correspondence with $g_\nu e = g_\nu$, we have

$$D(g_\mu)D(e) = D(g_\mu). \tag{15.3}$$

Therefore, we get

$$D(e) = E, \tag{15.4}$$

where $E$ is an identity matrix of dimension $d$. Also in correspondence with $g_\nu g_\nu^{-1} = g_\nu^{-1} g_\nu = E$, we have

$$D(g_\nu)D(g_\nu^{-1}) = D(g_\nu^{-1})D(g_\nu) = D(e) = E. \tag{15.5}$$

That is,

$$D(g_\nu^{-1}) = [D(g_\nu)]^{-1}. \tag{15.6}$$

Namely, an inverse matrix corresponds to an inverse element. If the representation has one-to-one correspondence, the representation is said to be faithful. In the case of $n$-to-one correspondence, the representation is homomorphic. In particular, representing all group elements by 1 [as (1,1) matrix] is called an "identity representation."

Illustrative examples of the representation have already appeared in Sects. 14.2 and 14.3. In that case, the representation was faithful. For instance, in Tables 14.2, 14.4, and 14.6 as well as Fig. 14.12, the number of representation matrices is the same as that of group elements. In Sects. 14.2 and 14.3, in most cases we have used real orthogonal matrices. Such matrices are included among unitary matrices. Representation using unitary matrices is said to be a "unitary representation." In fact, a representation of a finite group can be converted to a unitary representation.

**Theorem 15.1** [1] *A representation of any finite group can be converted to a unitary representation by a similarity transformation.*

*Proof* Let $\mathcal{G} = \{g_1, g_2, \cdots, g_n\}$ be a finite group of order $n$. Let $D(g_v)$ be a representation matrix of $g_v$ of dimension $d$. Here we suppose that $D(g_v)$ is not unitary, but non-singular. Using $D(g_v)$, we construct the following matrix $H$ such that

$$H = \sum_{i=1}^{n} D(g_i)^\dagger D(g_i). \tag{15.7}$$

Then, for an arbitrarily chosen group element $g_j$, we have

$$\begin{aligned} D(g_j)^\dagger H D(g_j) &= \sum_{i=1}^{n} D(g_j)^\dagger D(g_i)^\dagger D(g_i) D(g_j) \\ &= \sum_{i=1}^{n} \left[D(g_i) D(g_j)\right]^\dagger D(g_i) D(g_j) \\ &= \sum_{i=1}^{n} \left[D(g_i g_j)\right]^\dagger D(g_i g_j) = \sum_{k=1}^{n} D(g_k)^\dagger D(g_k) \\ &= H, \end{aligned} \tag{15.8}$$

where the third equality comes from homomorphism of the representation and the second last equality is due to rearrangement theorem (see Sect. 13.1).

Note that each matrix $D(g_i)^\dagger D(g_i)$ is a Hermitian Gram matrix constructed by a non-singular matrix and that $H$ is a summation of such Gram matrices. Consequently, on the basis of the argument of Sect. 11.2, $H$ is positive definite and all the eigenvalues of $H$ are positive. Then, using an appropriate unitary matrix $U$, we can get a diagonal matrix $\Lambda$ such that

$$U^\dagger H U = \Lambda. \tag{15.9}$$

Here, the diagonal matrix $\Lambda$ is given by

$$\Lambda = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_d \end{pmatrix}, \tag{15.10}$$

with $\lambda_i > 0$ $(1 \leq i \leq d)$. We define $\Lambda^{1/2}$ such that

$$\Lambda^{1/2} = \begin{pmatrix} \sqrt{\lambda_1} & & & \\ & \sqrt{\lambda_2} & & \\ & & \ddots & \\ & & & \sqrt{\lambda_d} \end{pmatrix}, \tag{15.11}$$

$$(\Lambda^{1/2})^{-1} = \begin{pmatrix} \sqrt{\lambda_1^{-1}} & & & \\ & \sqrt{\lambda_2^{-1}} & & \\ & & \ddots & \\ & & & \sqrt{\lambda_d^{-1}} \end{pmatrix}. \tag{15.12}$$

Notice that both $\Lambda^{1/2}$ and $(\Lambda^{1/2})^{-1}$ are non-singular. Furthermore, we define a matrix $V$ such that

$$V = U(\Lambda^{1/2})^{-1}. \tag{15.13}$$

Then, multiplying both sides of (15.8) by $V^{-1}$ from the left and by $V$ from the right and inserting $VV^{-1} = E$ in between, we get

$$V^{-1}D(g_j)^{\dagger}VV^{-1}HVV^{-1}D(g_j)V = V^{-1}HV. \tag{15.14}$$

Meanwhile, we have

$$V^{-1}HV = \Lambda^{1/2}U^{\dagger}HU\left(\Lambda^{1/2}\right)^{-1} = \Lambda^{1/2}\Lambda(\Lambda^{1/2})^{-1} = \Lambda. \tag{15.15}$$

With the second equality of (15.15), we used (15.9). Inserting (15.15) into (15.14), we get

$$V^{-1}D(g_j)^{\dagger}V\Lambda V^{-1}D(g_j)V = \Lambda. \tag{15.16}$$

Multiplying both sides of (15.16) by $\Lambda^{-1}$ from the left, we get

$$\Lambda^{-1}V^{-1}D(g_j)^{\dagger}(V\Lambda)\cdot V^{-1}D(g_j)V = E. \tag{15.17}$$

Using (15.13), we have

$$\Lambda^{-1}V^{-1} = (V\Lambda)^{-1} = \left[U\Lambda^{1/2}\right]^{-1} = \left(\Lambda^{1/2}\right)^{-1}U^{-1} = \left(\Lambda^{1/2}\right)^{-1}U^{\dagger}.$$

Meanwhile, taking adjoint of (15.13) and noting (15.12), we get

$$V^{\dagger} = \left[(\Lambda^{1/2})^{-1}\right]^{\dagger}U^{\dagger} = (\Lambda^{1/2})^{-1}U^{\dagger}.$$

Using (5.13) once again, we have

$$V\Lambda = U\Lambda^{1/2}.$$

Also using (5.13), we have

$$\left(V^{-1}\right)^{\dagger}=\left\{\left[U\left(\Lambda^{1/2}\right)^{-1}\right]^{-1}\right\}^{\dagger}=\left[\Lambda^{1/2}U^{\dagger}\right]^{\dagger}=\left[U^{\dagger}\right]^{\dagger}\left(\Lambda^{1/2}\right)^{\dagger}=U\Lambda^{1/2},$$

where we used Hermiticity of $U$ and (15.11).

Using the above relations and rewriting (15.17), finally we get

$$V^{\dagger}D\left(g_j\right)^{\dagger}\left(V^{-1}\right)^{\dagger}\cdot V^{-1}D\left(g_j\right)V=E. \tag{15.18}$$

Defining $\widetilde{D}\left(g_j\right)$ as follows

$$\widetilde{D}\left(g_j\right)\equiv V^{-1}D\left(g_j\right)V, \tag{15.19}$$

and taking adjoint of both sides of (15.19), we get

$$V^{\dagger}D\left(g_j\right)^{\dagger}\left(V^{-1}\right)^{\dagger}=\widetilde{D}\left(g_j\right)^{\dagger}. \tag{15.20}$$

Then, from (15.18), we have

$$\widetilde{D}\left(g_j\right)^{\dagger}\widetilde{D}\left(g_j\right)=E. \tag{15.21}$$

Equation (15.21) implies that a representation of any finite group can be converted to a unitary representation by a similarity transformation of (15.19). This completes the proof.

## 15.2 Basis Functions of Representation

In Part III, we considered a linear transformation of a vector. In that case, we have defined vectors as abstract elements with which operation laws of (9.1)–(9.8) hold. We assumed that the operation is addition. In this part, so far we have dealt with vectors mostly in $\mathbb{R}^3$. Therefore, vectors naturally possess geometric features. In this section, we extend a notion of vectors so that they can be treated under a wider scope. More specifically, we include mathematical functions treated in analysis as vectors.

To this end, let us think of a basis of a representation. According to the dimension $d$ of the representation, we have $d$ basis vectors. Here we adopt $d$ linearly independent basis vectors for the representation.

Let $\psi_1,\psi_2,\cdots,$ and $\psi_d$ be linearly independent vectors in a vector space $V^d$. Let $\mathcal{G}=\{g_1,g_2,\cdots,g_n\}$ be a finite group of order $n$. Here we assume that $g_i\ (1\leq i\leq n)$

is a linear transformation (or operator) such as a symmetry operation dealt with in Chap. 14. Suppose that the following relation holds with $g_i \in \mathcal{G}$ $(1 \le i \le n)$:

$$g_i(\psi_v) = \sum_{\mu=1}^{d} \psi_\mu D_{\mu v}(g_i) \quad (1 \le v \le d). \tag{15.22}$$

Here we followed the notation of (9.37) that represented a linear transformation of a vector. Rewriting (15.22) more explicitly, we have

$$g_i(\psi_v) = (\psi_1 \psi_2 \cdots \psi_d) \begin{pmatrix} D_{11}(g_i) & \cdots & D_{1d}(g_i) \\ \vdots & \ddots & \vdots \\ D_{d1}(g_i) & \cdots & D_{dd}(g_i) \end{pmatrix}. \tag{15.23}$$

Comparing (15.23) with (9.37), we notice that $\psi_1, \psi_2, \cdots,$ and $\psi_d$ act as vectors. The corresponding coordinates of the vectors (or a column vector), i.e., $x_1, x_2, \cdots,$ and $x_d$, have been omitted.

Now let us make sure that a set $\mathcal{L}$ consisting of $\{D(g_1), D(g_1), \cdots, D(g_n)\}$ form a representation. Operating $g_j$ on (15.22), we have

$$\begin{aligned} g_j[g_i(\psi_v)] &= [g_j g_i](\psi_v) = g_j \sum_{\mu=1}^{d} \psi_\mu D_{\mu v}(g_i) = \sum_{\mu=1}^{d} g_j[\psi_\mu D_{\mu v}(g_i)] \\ &= \sum_{\mu=1}^{d} (g_j \psi_\mu) D_{\mu v}(g_i) = \sum_{\mu=1}^{d} \sum_{\lambda=1}^{d} \psi_\lambda D_{\lambda\mu}(g_j) D_{\mu v}(g_i) \\ &= \sum_{\lambda=1}^{d} \psi_\lambda [D(g_j) D(g_i)]_{\lambda v}. \end{aligned} \tag{15.24}$$

Putting $g_j g_i = g_k$ according to multiplication of group elements, we get

$$g_k(\psi_v) = \sum_{\lambda=1}^{d} \psi_\lambda [D(g_j) D(g_i)]_{\lambda v}. \tag{15.25}$$

Meanwhile, replacing $g_i$ in (15.22) with $g_k$, we have

$$g_k(\psi_v) = \sum_{\mu=1}^{d} \psi_\lambda D_{\lambda v}(g_k). \tag{15.26}$$

Comparing (15.25) and (15.26) and considering the uniqueness of vector representation based on linear independence of the basis vectors (see Sect. 9.1), we get

$$[D(g_j) D(g_i)]_{\lambda v} = D_{\lambda v}(g_k) \equiv [D(g_k)]_{\lambda v}, \tag{15.27}$$

where the last identity follows the notation (9.38) of Sect. 9.2. Hence, we have

$$D(g_j)D(g_i) = D(g_k). \tag{15.28}$$

Thus, the set $\mathcal{L}$ consisting of $\{D(g_1), D(g_2), \cdots, D(g_n)\}$ is certainly a representation of the group $\mathcal{G} = \{g_1, g_2, \cdots, g_n\}$. In such a case, the set $\mathcal{B}$ consisting of linearly independent $d$ functions (i.e., vectors)

$$\mathcal{B} = \{\psi_1, \psi_2, \cdots, \psi_d\} \tag{15.29}$$

is said to be basis functions of the representation $D$. The number $d$ equals the dimension of representation. Correspondingly, the representation matrix is a $(d,d)$ square matrix. As remarked above, the correspondence between the elements of $\mathcal{L}$ and $\mathcal{G}$ is not necessarily one-to-one (isomorphic), but may be $n$-to-one (homomorphic).

**Definition 15.2** Let $D$ and $D'$ be two representations of $\mathcal{G} = \{g_1, g_2, \cdots, g_n\}$. Suppose that these representations are related to each other by similarity transformation such that

$$D'(g_i) = T^{-1}D(g_i)T \ (1 \leq i \leq n), \tag{15.30}$$

where $T$ is a non-singular matrix. Then, $D$ and $D'$ are said to be equivalent representations, or simply equivalent. If the representations are not equivalent, they are called inequivalent.

Suppose that $\mathcal{B} = \{\psi_1, \psi_2, \cdots, \psi_d\}$ is a basis of a representation $D$ of $\mathcal{G} = \{g_1, g_2, \cdots, g_n\}$. Then we have (15.22). Let $T$ be a non-singular matrix. Using $T$, we want to transform a basis of the representation $D$ from $\mathcal{B}$ to a new set $\mathcal{B}' = \{\psi'_1, \psi'_2, \cdots, \psi'_d\}$. Individual elements of the new basis are expressed as

$$\psi'_\nu = \sum_{\lambda=1}^{d} \psi_\lambda T_{\lambda\nu}. \tag{15.31}$$

Since $T$ is non-singular, this ensures that $\psi'_1, \psi'_2, \cdots,$ and $\psi'_d$ are linearly independent and, hence, that $\mathcal{B}'$ forms another basis set of $D$ (see discussion of Sect. 9.4). Thus, we can describe $\psi_\mu$ $(1 \leq \mu \leq n)$ in terms of $\psi'_\nu$. That is,

$$\psi_\mu = \sum_{\lambda=1}^{d} \psi'_\lambda \left(T^{-1}\right)_{\lambda\mu}. \tag{15.32}$$

Operating $g_i$ on both sides of (15.31), we have

$$
\begin{aligned}
g_i(\psi'_v) &= \sum_{\lambda=1}^{d} g_i(\psi_\lambda) T_{\lambda v} = \sum_{\lambda=1}^{d} \sum_{\mu=1}^{d} \psi_\mu D_{\mu\lambda}(g_i) T_{\lambda v} \\
&= \sum_{\lambda=1}^{d} \sum_{\mu=1}^{d} \sum_{\kappa=1}^{d} \psi'_\kappa \left(T^{-1}\right)_{\kappa\mu} D_{\mu\lambda}(g_i) T_{\lambda v} = \sum_{\kappa=1}^{d} \psi'_\kappa \left[T^{-1} D(g_i) T\right]_{\kappa v}.
\end{aligned}
\tag{15.33}
$$

Let $D'$ be a representation of $\mathcal{G}$ in reference to $\mathcal{B}' = \{\psi'_1, \psi'_2, \cdots, \psi'_d\}$. Then we have

$$
g_i(\psi'_v) = \sum_{\kappa=1}^{d} \psi'_\kappa D'_{\kappa v}. \tag{15.34}
$$

Hence, (15.30) follows in virtue of the linear independence of $\psi'_1, \psi'_2, \cdots,$ and $\psi'_d$. Thus, we see that the transformation of basis vectors via (15.31) causes similarity transformation between representation matrices. This is in parallel with (9.81) and (9.88).

Below we show several important notions as a definition. We assume that a group is a finite group.

**Definition 15.3** Let $D$ and $\widetilde{D}$ be two representations of $\mathcal{G} = \{g_1, g_2, \cdots, g_n\}$. Let $D(g_i)$ and $\widetilde{D}(g_i)$ be two representation matrices of $g_i$ $(1 \leq i \leq n)$. If we construct $\Delta(g_i)$ such that

$$
\Delta(g_i) = \begin{pmatrix} D(g_i) & \\ & \widetilde{D}(g_i) \end{pmatrix}, \tag{15.35}
$$

then $\Delta(g_i)$ is a representation as well. The representation $\Delta(g_i)$ is said to be a direct sum of $D(g_i)$ and $\widetilde{D}(g_i)$. We denote it by

$$
\Delta(g_i) = D(g_i) \oplus \widetilde{D}(g_i). \tag{15.36}
$$

A dimension $\Delta(g_i)$ is a sum of that of $D(g_i)$ and that of $\widetilde{D}(g_i)$.

**Definition 15.4** Let $D$ be a representation of $\mathcal{G}$ and let $D(g_i)$ be a representation matrix of a group element $g_i$. If we can convert $D(g_i)$ to a block matrix such as (15.35) by its similarity transformation, $D$ is said to be a reducible representation, or simply reducible. If the representation is not reducible, it is irreducible.

A reducible representation can be decomposed (or reduced) to a direct sum of plural irreducible representations. Such an operation (or procedure) is called reduction.

**Definition 15.5** Let $\mathcal{G} = \{g_1, g_2, \cdots, g_n\}$ be a group and let $V$ be a vector space. Then, we denote a representation $D$ of $\mathcal{G}$ operating on $V$ by

$$
D : \mathcal{G} \to \mathrm{GL}(V).
$$

We call $V$ a representation space (or carrier space) $\mathcal{L}_S$ of $D$ [2, 3].

From Definition 15.5, the dimension of representation is identical with the dimension of $V$. Suppose that there is a subspace $W$ in $V$. If $W$ is $D(g_i)$-invariant (see Sect. 10.2) with all ${}^{\forall}g_i \in \mathcal{g}$, $W$ is said to be an invariant subspace of $V$. Here, we say that $W$ is $D(g_i)$-invariant if we have $|x\rangle \in W \Rightarrow D(g_i)|x\rangle \in W$; see Sect. 10.2. Notice that $|x\rangle$ may well represent a function $\psi_\nu$ of (15.22). In this context, we have a following important theorem.

**Theorem 15.2** [2] *Let $D$: $\mathcal{g} \to \mathrm{GL}(V)$ be a unitary representation over $V$. Let $W$ be a $D(g_i)$-invariant subspace of $V$, where $g_i$ $(1 \le i \le n) \in \mathcal{g} = \{g_1, g_2, \cdots, g_n\}$. Then, $W^\perp$ is a $D(g_i)$-invariant subspace of $V$ as well.*

*Proof* Suppose that $|a\rangle \in W^\perp$ and let $|b\rangle \in W$. Then, from (11.86), we have

$$\langle b|D(g_i)|a\rangle = \left\langle a\middle|D(g_i)^\dagger\middle|b\right\rangle^* = \left\langle a\middle|[D(g_i)]^{-1}\middle|b\right\rangle^* = \left\langle a\middle|D\left(g_i^{-1}\right)\middle|b\right\rangle^*,$$

where we used the fact that $D(g_i)$ is a unitary matrix; see (15.6). But, since $W$ is $D(g_i)$-invariant from the supposition, $|D\left(g_i^{-1}\right)|b\rangle \in W$. Here notice that $g_i^{-1}$ must be ${}^{\exists}g_i$ $(1 \le i \le n) \in \mathcal{g}$. Therefore, we should have

$$\langle |bD(g_i)a|\rangle = \left\langle a\middle|D\left(g_i^{-1}\right)\middle|b\right\rangle^* = 0.$$

This implies that $|D(g_i)|a\rangle \in W^\perp$. This means that $W^\perp$ is a $D(g_i)$-invariant subspace of $V$. This completes the proof.

From Theorem 12.1, we have

$$V = W \oplus W^\perp.$$

Correspondingly, as in the case of (10.102), $D(g_i)$ is reduced as follows:

$$D(g_i) = \begin{pmatrix} D^{(W)}(g_i) & \\ & D^{(W^\perp)}(g_i) \end{pmatrix},$$

where $D^{(W)}(g_i)$ and $D^{(W^\perp)}(g_i)$ are representation matrices associated with subspaces $W$ and $W^\perp$, respectively. This is the converse of the additive representations that appeared in Definition 15.3. From (9.17) to (9.18), we have

$$\dim V = \dim W + \dim W^\perp. \tag{15.37}$$

If $V$ is a $d$-dimensional vector space, $V$ is spanned by $d$ linearly independent basis vectors (or functions) $\psi_\mu$ $(1 \le \mu \le d)$. Suppose that the dimension $W$ and $W^\perp$ is $d^{(W)}$ and $d^{(W^\perp)}$, respectively. Then, $W$ and $W^\perp$ are spanned by $d^{(W)}$ linearly independent vectors and $d^{(W^\perp)}$ linearly independent vectors, respectively. The subspaces $W$ and $W^\perp$ may well further be decomposed into orthogonal complements [i.e., $D(g_i)$-invariant subspaces of $V$]. In this situation, in general, we write

$$D(g_i) = D^{(1)}(g_i) \oplus D^{(2)}(g_i) \oplus \cdots \oplus D^{(\omega)}(g_i). \tag{15.38}$$

We will develop further detailed discussion in Sect. 15.4.

If the aforementioned decomposition cannot be done, $D$ is irreducible. Therefore, it will be of great importance to examine a dimension of irreducible representations contained in (15.38). This is closely related to the choice of basis vectors and properties of the invariant subspaces. Notice that the same irreducible representation may repeatedly occur in (15.38).

Before advancing to the next section, however, let us think of an example to get used to abstract concepts. This is at the same time for the purpose of taking the contents of Chap. 16 in advance.

*Example 15.1* Figure 15.1 shows structural formulae including resonance structures for allyl radical. Thanks to the resonance structures, the allyl radical belongs to $C_{2v}$; see the multiplication table in Table 14.1. The molecule lies on the $yz$-plane and a line connecting $C_1$ and a bonded H is the $C_2$ axis (see Fig. 15.1). The $C_2$ axis is identical to the $z$-axis. The molecule has mirror symmetries with respect to the $yz$- and $zx$-planes. We denote $\pi$-orbitals of $C_1$, $C_2$, and $C_3$ by $\phi_1$, $\phi_2$, and $\phi_3$, respectively. We suppose that these orbitals extend toward the $x$-direction with a positive sign in the upper side and a negative sign in the lower side relative to the plane of paper. Notice that we follow custom of the group theory notation with the coordinate setting in Fig. 15.1.

We consider an inner vector space $V^3$ spanned by $\phi_1$, $\phi_2$, and $\phi_3$. To explicitly show that these are vectors of the inner vector space, we express them as $|\phi_1\rangle$, $|\phi_2\rangle$, and $|\phi_3\rangle$ in this example. Then, according to (9.19) of Sect. 9.1, we write

$$V^3 = \text{Span } \{|\phi_1\rangle, |\phi_2\rangle, |\phi_3\rangle\}.$$

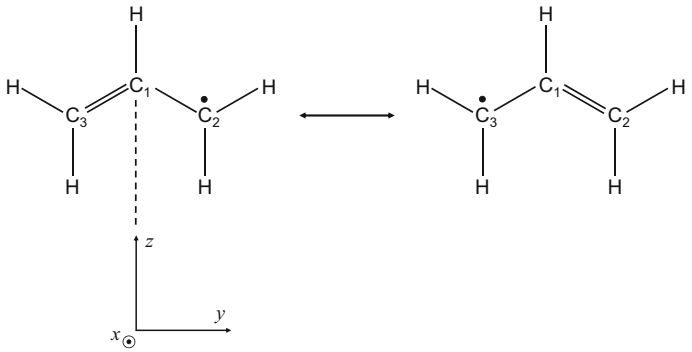


**Fig. 15.1** Allyl radical and its resonance structure. The molecule is placed on the $yz$-plane. The $z$-axis is identical with a straight line connecting $C_1$ and H (i.e., $C_2$ axis)

The vector space $V^3$ is a representation space pertinent to a representation $D$ of the present example. Also, in parallel to (11.32) of Sect. 11.2, we express an arbitrary vector $|\psi\rangle \in V^3$ as

$$|\psi\rangle = c_1|\phi_1\rangle + c_2|\phi_2\rangle + c_3|\phi_3\rangle = |c_1\phi_1 + c_2\phi_2 + c_3\phi_3\rangle$$
$$= (|\phi_1\rangle|\phi_2\rangle|\phi_3\rangle)\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}.$$

Let us now operate a group element of $C_{2v}$. For example, choosing $C_2(z)$, we have

$$C_2(z)(|\psi\rangle) = (|\phi_1\rangle|\phi_2\rangle|\phi_3\rangle)\begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}. \tag{15.39}$$

Thus, $C_2(z)$ is represented by a (3,3) matrix. Other group elements are represented similarly. These results are collected in Table 15.1, where each matrix is given with respect to $|\phi_1\rangle$, $|\phi_2\rangle$, and $|\phi_3\rangle$ as basis vectors. Notice that the matrix representations differ from those of Table 14.2, where we chose $|x\rangle, |y\rangle$, and $|z\rangle$ for the basis vectors. From Table 15.1, we immediately see that the representation matrices are reduced to an upper (1,1) diagonal matrix (i.e., just a number) and a lower (2,2) diagonal matrix.

In Table 15.1, we find that all the matrices are Hermitian (as well as unitary). Since $C_{2v}$ is an Abelian group (i.e., commutative), in light of Theorem 12.13, we should be able to diagonalize these matrices by a single unitary similarity transformation all at once. In fact, $E$ and $\sigma_v'(yz)$ are invariant with respect to unitary similarity transformation, and so we only have to diagonalize $C_2(z)$ and $\sigma_v(zx)$ at once. As a characteristic equation of the above (3,3) matrix of (15.39), we have

$$\begin{vmatrix} -\lambda - 1 & 0 & 0 \\ 0 & -\lambda & -1 \\ 0 & -1 & -\lambda \end{vmatrix} = 0.$$

Solving the above equation, we get $\lambda = 1$ or $\lambda = -1$ (as a double root). Also as a diagonalizing unitary matrix $U$, we get

**Table 15.1** Matrix representation of symmetry operations given with respect to $|\phi_1\rangle$, $|\phi_2\rangle$, and $|\phi_3\rangle$

| $E$ | $C_2(z)$ | $\sigma_v(zx)$ | $\sigma_v'(yz)$ |
|---|---|---|---|
| $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ | $\begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix}$ | $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$ | $\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$ |

$$U = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} = U^{\dagger}. \tag{15.40}$$

Thus, (15.39) can be rewritten as

$$\begin{aligned} C_2(z)(|\psi\rangle) &= (|\phi_1\rangle|\phi_2\rangle|\phi_3\rangle)UU^{\dagger}\begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix}UU^{\dagger}\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} \\ &= \left(|\phi_1\rangle \frac{1}{\sqrt{2}}(|\phi_2+\phi_3\rangle)\frac{1}{\sqrt{2}}(|\phi_2-\phi_3\rangle)\right)\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} c_1 \\ \frac{1}{\sqrt{2}}(c_2+c_3) \\ \frac{1}{\sqrt{2}}(c_2-c_3) \end{pmatrix}. \end{aligned}$$

The diagonalization of the representation matrix $\sigma_v(zx)$ using the same $U$ as the above is left for the readers as an exercise. Table 15.2 shows the results of the diagonalized representations with regard to the vectors $|\phi_1\rangle$, $\frac{1}{\sqrt{2}}(|\phi_2+\phi_3\rangle)$ and $\frac{1}{\sqrt{2}}(|\phi_2-\phi_3\rangle)$. Notice that traces of the matrices remain unchanged in Tables 15.1 and 15.2, i.e., before and after the unitary similarity transformation. From Table 15.2, we find that the vectors are eigenfunctions of the individual symmetry operations. Each diagonal element is a corresponding eigenvalue of those operations. Of the three vectors, $|\phi_1\rangle$ and $\frac{1}{\sqrt{2}}(|\phi_2+\phi_3\rangle)$ have the same eigenvalues with respect to individual symmetry operations. With symmetry operations $C_2$ and $\sigma_v(zx)$, another vector $\frac{1}{\sqrt{2}}(|\phi_2-\phi_3\rangle)$ has an eigenvalue of a sign opposite to that of the former two vectors. Thus, we find that we have arrived at "symmetry-adapted" vectors by taking linear combination of original vectors. Returning back to Table 14.2, we see that the representation matrices have already been diagonalized. In terms of the representation space, we constructed the representation matrices with respect to $x$, $y$, and $z$ as symmetry-adapted basis vectors. In the next chapter, we make the most of such vectors constructed by the symmetry-adapted linear combination.

Meanwhile, allocating appropriate numbers to coefficients $c_1$, $c_2$, and $c_3$, we represent any vector in $V^3$ spanned by $|\phi_1\rangle$, $|\phi_2\rangle$, and $|\phi_3\rangle$. In the next chapter, we deal with molecular orbital (MO) calculations. Including the present case of the allyl radical, we solve the energy eigenvalue problem by appropriately determining those

**Table 15.2** Matrix representation of symmetry operations given with respect to $|\phi_1\rangle$, $\frac{1}{\sqrt{2}}(|\phi_2+\phi_3\rangle)$, and $\frac{1}{\sqrt{2}}(|\phi_2-\phi_3\rangle)$

| $E$ | $C_2(z)$ | $\sigma_v(zx)$ | $\sigma_v'(yz)$ |
|---|---|---|---|
| $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ | $\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ | $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$ | $\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$ |

coefficients (i.e., eigenvectors) and corresponding energy eigenvalues in a representation space. A dimension of the representation space depends on the number of molecular orbitals. The representation space is decomposed into several or more (but finite) orthogonal complements according to (15.38). We will come back to the present example in Sect. 16.3.4 and further investigate the problems there.

As can be seen in the above example, the representation space accommodates various types of vectors, e.g., mathematical functions in the present case. If the representation space is decomposed into invariant subspaces, we can choose appropriate basis vectors for each subspace the number of which is equal to the dimensionality of each irreducible representation [3]. In this context, the situation is related to that of Part III where we have studied how a linear vector space is decomposed into direct sum of invariant subspaces that are spanned by associated basis vectors. In particular, it will be of great importance to construct mutually orthogonal symmetry-adapted vectors through linear combination of original basis vectors in each subspace associated with the irreducible representation. We further study these important subjects in the following several sections.

## 15.3 Schur's Lemmas and Grand Orthogonality Theorem (GOT)

Pivotal notions of the representation theory of finite groups rest upon Schur's lemmas (first lemma and second lemma).

**Schur's First Lemma** [1] *Let $D$ and $\widetilde{D}$ be two irreducible representations of $\mathcal{G}$. Let dimensions of representations of $D$ and $\widetilde{D}$ be $m$ and $n$, respectively. Suppose that with $^{\forall}g \in \mathcal{G}$ the following relation holds*:

$$D(g)M = M\widetilde{D}(g), \tag{15.41}$$

*where $M$ is a $(m,n)$ matrix. Then we must have*

Case (i): $M = 0$ *or*
Case (ii): *$M$ is a square matrix* (*i.e.*, $m = n$) *with* $\det M \neq 0$.

*In* Case (i), *the representations of $D$ and $\widetilde{D}$ are inequivalent. In* Case (ii), *on the other hand, $D$ and $\widetilde{D}$ are equivalent.*

*Proof*

(a) First, suppose that $m > n$. Let $\mathcal{B} = \{\psi_1, \psi_2, \cdots, \psi_m\}$ be a basis set of the representation space related to $D$. Then we have

$$g(\psi_\nu) = \sum_{\mu=1}^{m} \psi_\mu D_{\mu\nu}(g) \ (1 \leq \nu \leq m).$$

Next, we form a linear combination of $\psi_1, \psi_2, \cdots,$ and $\psi_m$ such that

$$\phi_\nu = \sum_{\mu=1}^{m} \psi_\mu M_{\mu\nu} \ (1 \le \nu \le n). \tag{15.42}$$

Operating $g$ on both sides of (15.42), we have

$$\begin{aligned} g(\phi_\nu) &= \sum_{\mu=1}^{m} g(\psi_\mu) M_{\mu\nu} = \sum_{\mu=1}^{m} \left[ \sum_{\lambda=1}^{m} \psi_\lambda D_{\lambda\mu}(g) \right] M_{\mu\nu} \\ &= \sum_{\lambda=1}^{m} \psi_\lambda \left[ \sum_{\mu=1}^{m} D_{\lambda\mu}(g) M_{\mu\nu} \right] = \sum_{\lambda=1}^{m} \psi_\lambda \left[ \sum_{\mu=1}^{n} M_{\lambda\mu} \widetilde{D}_{\mu\nu}(g) \right] \\ &= \sum_{\mu=1}^{n} \left[ \sum_{\lambda=1}^{m} \psi_\lambda M_{\lambda\mu} \right] \widetilde{D}_{\mu\nu}(g) = \sum_{\mu=1}^{n} \phi_\mu \widetilde{D}_{\mu\nu}(g). \end{aligned} \tag{15.43}$$

With the fourth equality of (15.43), we used (15.41). Therefore, $\tilde{\mathcal{B}} = \{\phi_1, \phi_2, \cdots, \phi_n\}$ is a representation of $\widetilde{D}$.

If $m > n$, we would have been able to construct a basis of the representation $D$ using $n$ ($< m$) functions of $\phi_\nu$ ($1 \le \nu \le n$) that are a linear combination of $\psi_\mu$ ($1 \le \mu \le m$). In that case, we would have obtained a representation of a smaller dimension $n$ for $\widetilde{D}$ by taking a linear combination of $\psi_1, \psi_2, \cdots,$ and $\psi_m$. As $n < m$ by supposition, this implies that as in (15.35), $D$ would be decomposed into representations whose dimension is smaller than $m$. It is in contradiction to the supposition that $D$ is *irreducible*. To avoid this contradiction, we must have $M = 0$. This makes (15.41) trivially hold.

(b) Next, we consider a case of $m < n$. Taking a complex conjugate of (15.41) and exchanging both sides, we get

$$\widetilde{D}(g)^\dagger M^\dagger = M^\dagger D(g)^\dagger. \tag{15.44}$$

Then, from (15.21) to (15.28), we have

$$D(g)^\dagger D(g) = E \quad \text{and} \quad D(g) D\left(g^{-1}\right) = D(e). \tag{15.45}$$

Therefore, we get

$$D\left(g^{-1}\right) = D(g)^{-1} = D(g)^\dagger.$$

Similarly, we have

$$\widetilde{D}\left(g^{-1}\right) = \widetilde{D}(g)^{-1} = \widetilde{D}(g)^\dagger.$$

Thus, (15.44) is rewritten as

$$\widetilde{D}\left(g^{-1}\right)M^{\dagger} = M^{\dagger}D\left(g^{-1}\right). \tag{15.46}$$

The number of rows of $M^{\dagger}$ is larger than that of columns. This time, dimension $n$ of $\widetilde{D}(g^{-1})$ is larger than $m$ of $D(g^{-1})$. A group element $g$ designates an arbitrary element of $\mathcal{G}$, so does $g^{-1}$. Thus, (15.46) can be treated in parallel with (15.41) where $m > n$ with a $(m,n)$ matrix $M$. Figure 15.2 graphically shows the magnitude relationship in representation dimensions between representation matrices and $M$. Thus, in parallel with the above argument of (a), we exclude the case where $m < n$ as well.

(c) We consider the third case of $m = n$. Similarly as before, we make a linear combination of vectors contained in the basis set $\mathcal{B} = \{\psi_1, \psi_2, \cdots, \psi_m\}$ such that

$$\phi_v = \sum_{\mu=1}^{m} \psi_{\mu} M_{\mu v} \quad (1 \leq v \leq m), \tag{15.47}$$

where $M$ is a $(m,m)$ square matrix. If $\det M = 0$, then $\phi_1, \phi_2, \cdots,$ and $\phi_m$ are linearly dependent (see Sect. 9.4). With the number of linearly independent vectors $p$, we have $p < m$ accordingly. As in (15.43) again, this implies that we would have obtained a representation of a smaller dimension $p$ for $\widetilde{D}$, in contradiction to the supposition that $\widetilde{D}$ is *irreducible*. To avoid this contradiction, we must have $\det M \neq 0$. These complete the proof.

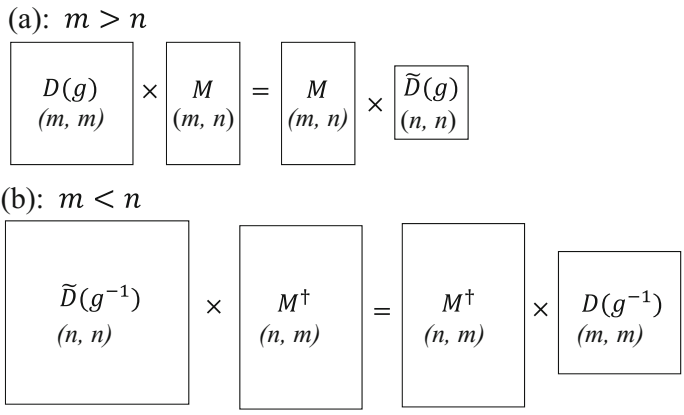


**Fig. 15.2** Magnitude relationship between dimensions of representation matrices and $M$. **a** $m > n$. **b** $m < n$. The diagram is based on (15.41) and (15.46) of the text

**Schur's Second Lemma** [1] *Let $D$ be a representation of $\mathcal{G}$. Suppose that with $^{\forall}g \in \mathcal{G}$, we have*

$$D(g)M = MD(g). \tag{15.48}$$

*Then, if $D$ is irreducible, $M = cE$ where $c$ being an arbitrary complex number and $E$ is an identity matrix.*

*Proof* Let $c$ be an arbitrarily chosen complex number. From (15.48), we have

$$D(g)(M - cE) = (M - cE)D(g). \tag{15.49}$$

If $D$ is irreducible, Schur's First Lemma implies that we must have either (i) $M - cE = 0$ or (ii) $\det(M - cE) \neq 0$. A matrix $M$ has at least one proper eigenvalue $\lambda$ (Sect. 10.1), and so choosing $\lambda$ for $c$ in (15.49), we have $\det(M - \lambda E) = 0$. Consequently, only former case is allowed. That is, we have

$$M = cE. \tag{15.50}$$

Schur's lemmas lead to important orthogonality theorem that plays a fundamental role in many scientific fields. The orthogonality theorem includes that of matrices and their traces (or characters).

**Theorem 15.3** (Grand Orthogonality Theorem (GOT) [4]) *Let $D^{(1)}, D^{(2)}, \cdots$ be all inequivalent irreducible representations of a group $\mathcal{G} = \{g_1, g_2, \cdots, g_n\}$ of order $n$. Let $D^{(\alpha)}$* and *$D^{(\beta)}$ be two irreducible representations chosen from among $D^{(1)}, D^{(2)}, \cdots$. Then, regarding their* matrix *representations, we have the following relationship*:

$$\sum_g D_{ij}^{(\alpha)}(g)^* D_{kl}^{(\beta)}(g) = \frac{n}{d_\alpha}\delta_{\alpha\beta}\delta_{ik}\delta_{jl}, \tag{15.51}$$

*where $\sum_g$ means that the summation should be taken over all $n$ group elements; $d_\alpha$ denotes a dimension of the representation $D^{(\alpha)}$. The symbol $\delta_{\alpha\beta}$ means that $\delta_{\alpha\beta} = 1$ when $D^{(\alpha)}$* and *$D^{(\beta)}$ are equivalent* and *that $\delta_{\alpha\beta} = 0$* when *$D^{(\alpha)}$* and *$D^{(\beta)}$ are inequivalent*.

*Proof* First we prove the case where $D^{(\alpha)} = D^{(\beta)}$. For the sake of simple expression, we omit a superscript and denote $D^{(\alpha)}$ simply by $D$. Let us construct a matrix $A$ such that

$$A = \sum_g D(g)XD(g^{-1}), \tag{15.52}$$

where $X$ is an arbitrary matrix. Hence,

$$\begin{aligned} D(g')A &= \sum_g D(g')D(g)XD\left(g^{-1}\right) = \sum_g D(g')D(g)XD\left(g^{-1}\right)D\left(g'^{-1}\right)D(g') \\ &= \sum_g D(g')D(g)XD\left(g^{-1}g'^{-1}\right)D(g') \\ &= \sum_g D(g'g)XD\left[(g'g)^{-1}\right]D(g'). \end{aligned} \tag{15.53}$$

Thanks to the rearrangement theorem, for fixed $g'$, the element $g'g$ runs through all the group elements as $g$ does so. Therefore, we have

$$\sum_g D(g'g)XD[(g'g)^{-1}] = \sum_g D(g)XD(g^{-1}) = A. \tag{15.54}$$

Thus,

$$D(g)A = AD(g). \tag{15.55}$$

According to Schur's Second Lemma, we have

$$A = \lambda E. \tag{15.56}$$

The value of a constant $\lambda$ depends upon the choice of $X$. Let $X$ be $\delta_i^{(l)}\delta^j_{(m)}$ where all the matrix elements are zero except for the $(l,m)$-component that takes 1 (Sects. 10.5 and 10.6). Thus from (15.53), we have

$$\sum_{g,p,q} D_{ip}(g)\delta_p^{(l)}\delta^q_{(m)}D_{qj}\left(g^{-1}\right) = \sum_g D_{il}(g)D_{mj}\left(g^{-1}\right) = \lambda_{lm}\delta_{ij}, \tag{15.57}$$

where $\lambda_{lm}$ is a constant to be determined. Using the unitary representation, we have

$$\sum_g D_{il}(g)D_{jm}(g)^* = \lambda_{lm}\delta_{ij}. \tag{15.58}$$

Next, we wish to determine coefficients $\lambda_{lm}$. To this end, setting $i = j$ and summing over $i$ in (15.57), we get for LHS

$$\begin{aligned} \sum_g \sum_i D_{il}(g)D_{mi}(g^{-1}) &= \sum_g \left[D\left(g^{-1}\right)D(g)\right]_{ml} = \sum_g \left[D\left(g^{-1}g\right)\right]_{ml} \\ &= \sum_g [D(e)]_{ml} = \sum_g \delta_{ml} = n\delta_{ml}, \end{aligned} \tag{15.59}$$

where $n$ is equal to the order of group. As for RHS, we have

$$\sum_i \lambda_{lm}\delta_{ii} = \lambda_{lm}d, \tag{15.60}$$

where $d$ is equal to a dimension of $D$. From (15.59) to (15.60), we get

$$\lambda_{lm}d = n\delta_{lm} \quad \text{or} \quad \lambda_{lm} = \frac{n}{d}\delta_{lm}. \tag{15.61}$$

Therefore, from (15.58)

$$\sum_g D_{il}(g)D_{jm}(g)^* = \frac{n}{d}\delta_{lm}\delta_{ij}. \tag{15.62}$$

Specifying a species of the *irreducible* representation, we get

$$\sum_g D^{(\alpha)}_{il}(g)D^{(\alpha)}_{jm}(g)^* = \frac{n}{d_\alpha}\delta_{lm}\delta_{ij}, \tag{15.63}$$

where $d_\alpha$ is a dimension of $D^{(\alpha)}$.

Next, we examine the relationship between two inequivalent irreducible representations. Let $D^{(\alpha)}$ and $D^{(\beta)}$ be such representations with dimensions $d_\alpha$ and $d_\beta$, respectively. Let us construct a matrix $B$ such that

$$B = \sum_g D^{(\alpha)}(g)XD^{(\beta)}\left(g^{-1}\right), \tag{15.64}$$

where $X$ is again an arbitrary matrix. Hence,

$$\begin{aligned} D^{(\alpha)}(g')B &= \sum_g D^{(\alpha)}(g')D^{(\alpha)}(g)XD^{(\beta)}(g^{-1}) \\ &= \sum_g D^{(\alpha)}(g')D^{(\alpha)}(g)XD^{(\beta)}(g^{-1})D^{(\beta)}\left(g'^{-1}\right)D^{(\beta)}(g') \\ &= \sum_g D^{(\alpha)}(g'g)XD^{(\beta)}\left[(g'g)^{-1}\right]D^{(\beta)}(g') = BD^{(\beta)}(g'). \end{aligned} \tag{15.65}$$

According to Schur's First Lemma, we have

$$B = 0. \tag{15.66}$$

Putting $X = \delta^{(l)}_i \delta^j_{(m)}$ as before and rewriting (15.64), we get

$$\sum_g D^{(\alpha)}_{il}(g)D^{(\beta)}_{jm}(g)^* = 0. \tag{15.67}$$

Combining (15.63) and (15.67), we get (15.51). These procedures complete the proof.

## 15.4 Characters

Representation matrices of a group are square matrices. In Part III, we examined properties of a trace, i.e., a sum of diagonal elements of a square matrix. In group theory, the trace is called a character.

**Definition 15.6** Let $D$ be a (matrix) representation of a group $\mathscr{g} = \{g_1, g_2, \cdots, g_n\}$. The sum of diagonal elements $\chi(g)$ is defined as follows:

$$\chi(g) \equiv \mathrm{Tr}D(g) = \sum_{i=1}^{d} D_{ii}(g), \tag{15.68}$$

where $g$ stands for group elements $g_1, g_2, \cdots,$ and $g_n$; Tr stands for "trace"; $d$ is a dimension of the representation $D$. Let $\mathsf{c}$ be a set defined as

$$\mathsf{c} = \{\chi(g_1), \chi(g_2), \cdots, \chi(g_n)\}. \tag{15.69}$$

Then, the set $\mathsf{c}$ is called a character of $D$. A character of an irreducible representation is said to be an irreducible character.

Let us describe several properties of the character or trace.

(i) A character of the identity element $\chi(e)$ is equal to a dimension $d$ of a representation. This is because the identity is given by a unit matrix.

(ii) Let $P$ and $Q$ be two square matrices. Then, we have

$$\mathrm{Tr}(PQ) = \mathrm{Tr}(QP). \tag{15.70}$$

This is because

$$\sum_i (PQ)_{ii} = \sum_i \sum_j P_{ij} Q_{ji} = \sum_j \sum_i Q_{ji} P_{ij} = \sum_j (QP)_{jj} \tag{15.71}$$

Putting $Q = SP^{-1}$ in (15.70), we get

$$\mathrm{Tr}\left(PSP^{-1}\right) = \mathrm{Tr}\left(SP^{-1}P\right) = \mathrm{Tr}(S). \tag{15.72}$$

Therefore, we have the following property:

(iii) Characters of group elements that are conjugate to each other are equal. If $g_i$ and $g_j$ are conjugate, these elements are connected by means of a suitable element $g$ such that

$$g g_i g^{-1} = g_j. \tag{15.73}$$

Accordingly, a representation matrix is expressed as

$$D(g)D(g_i)D\left(g^{-1}\right) = D(g)D(g_i)[D(g)]^{-1} = D\left(g_j\right). \tag{15.74}$$

Taking a trace of both sides of (15.74), we have

$$\chi(g_i) = \chi(g_j). \tag{15.75}$$

(iv) Any two equivalent representations have the same trace. This immediately follows from (15.30).

There are several orthogonality theorems about a trace. Among them, a following theorem is well-known.

**Theorem 15.4** *A trace of irreducible representations satisfies the following orthogonality relation*:

$$\sum_g \chi^{(\alpha)}(g)^* \chi^{(\beta)}(g) = n\delta_{\alpha\beta}, \tag{15.76}$$

*where* $\chi^{(\alpha)}$ *and* $\chi^{(\beta)}$ *are traces of irreducible representations* $D^{(\alpha)}$ *and* $D^{(\beta)}$, *respectively*.

*Proof* In (15.51), putting $i = j$ and $k = l$ in both sides and summing over all $i$ and $k$, we have

$$\sum_g \sum_{i,k} D_{ii}^{(\alpha)}(g)^* D_{kk}^{(\beta)}(g) = \sum_{i,k} \frac{n}{d_\alpha} \delta_{\alpha\beta}\delta_{ik}\delta_{ik} = \sum_{i,k} \frac{n}{d_\alpha} \delta_{\alpha\beta}\delta_{ik} = \frac{n}{d_\alpha} \delta_{\alpha\beta} d_\alpha = n\delta_{\alpha\beta}. \tag{15.77}$$

From (15.68) to (15.77), we get (15.76). This completes the proof.

Since a character is identical with group elements belonging to the same conjugacy class $K_l$, we may write it as $\chi(K_l)$ and rewrite a summation of (15.76) as a summation of the conjugacy classes. Thus, we have

$$\sum_{l=1}^{n_c} \chi^{(\alpha)}(K_l)^* \chi^{(\beta)}(K_l) k_l = n\delta_{\alpha\beta}, \tag{15.78}$$

where $n_c$ denotes the number of conjugacy classes in a group and $k_l$ indicates the number of group elements contained in a class $K_l$.

We have seen a case where a representation matrix can be reduced to two (or more) block matrices as in (15.35). Also as already seen in Part III, the block matrices' decomposition takes place with normal matrices (including unitary matrices). The character is often used to examine a constitution of a reducible representation or reducible matrix.

Alternatively, if a unitary matrix (a normal matrix, more widely) is decomposed into block matrices, we say that the unitary matrix comprises a direct sum of those block matrices. In physics and chemistry dealing with atoms, molecules, crystals, etc., we very often encounter such a situation. Extending (15.36), the relation can generally be summarized as

$$D(g_i) = D^{(1)}(g_i) \oplus D^{(2)}(g_i) \oplus \cdots \oplus D^{(\omega)}(g_i), \tag{15.79}$$

where $D(g_i)$ is a reducible representation for a group element $g_i$; $D^{(1)}(g_i)$, $D^{(2)}(g_i)$, $\cdots$, and $D^{(\omega)}(g_i)$ are irreducible representations in a group. The notation $D^{(\omega)}(g_i)$ means that $D^{(\omega)}(g_i)$ may be equivalent (or identical) to $D^{(1)}(g_i)$, $D^{(2)}(g_i)$, etc., or may be inequivalent to them. More specifically, the same irreducible representations may well appear several times.

To make the above situation clear, we usually use the following equation instead:

$$D(g_i) = \sum_{\alpha} q_\alpha D^{(\alpha)}(g_i), \tag{15.80}$$

where $q_\alpha$ is zero or a positive integer and $D^{(\alpha)}$ is different types of irreducible representations. If the same $D^{(\alpha)}$ repeatedly appears in the direct sum, then $q_\alpha$ specifies how many times $D^{(\alpha)}$ appears in the direct sum. Unless $D^{(\alpha)}$ appears, $q_\alpha$ is zero.

Bearing the above in mind, we take a trace of (15.80). Then we have

$$\chi(g) = \sum_{\alpha} q_\alpha \chi^{(\alpha)}(g), \tag{15.81}$$

where we omitted a subscript $i$ indicating an element. To find $q_\alpha$, let us multiply both sides of (15.81) by $\chi^{(\alpha)}(g)^*$ and take summation over group elements. That is,

$$\sum_{g} \chi^{(\alpha)}(g)^* \chi(g) = \sum_{\beta} q_\beta \sum_{g} \chi^{(\alpha)}(g)^* \chi^{(\beta)}(g) = \sum_{\beta} q_\beta n \delta_{\alpha\beta} = q_\alpha n, \tag{15.82}$$

where we used (15.76) with the second equality. Thus, we get

$$q_\alpha = \frac{1}{n} \sum_{g} \chi^{(\alpha)}(g)^* \chi(g). \tag{15.83}$$

The integer $q_\alpha$ explicitly gives the number of appearance of $D^{(\alpha)}(g)$ that appears in a reducible representation $D(g)$. The expression pertinent to the classes is

$$q_\alpha = \frac{1}{n} \sum_{i} \chi^{(\alpha)}(K_i)^* \chi(K_i) k_i, \tag{15.84}$$

where $K_i$ and $k_i$ denote the $i$th class of the group and the number of elements belonging to $K_i$, respectively.

## 15.5 Regular Representation and Group Algebra

Now, the readers may wonder how many different irreducible representations exist for a group. To answer this question, let us introduce a special representation of the *regular representation*.

**Definition 15.7** Let $\mathcal{G} = \{g_1, g_2, \cdots, g_n\}$ be a group. Let us define a $(n,n)$ square matrix $D^{(\mathrm{R})}(g_v)$ for an arbitrary group element $g_v$ $(1 \le v \le n)$ such that

$$\left[D^{(\mathrm{R})}(g_v)\right]_{ij} = \delta\left(g_i^{-1} g_v g_j\right) \ (1 \le i,j \le n), \tag{15.85}$$

where

$$\delta(g_v) = \begin{cases} 1 & \text{for } g_v = e \text{ (i.e. identity)}, \\ 0 & \text{for } g_v \neq e. \end{cases} \tag{15.86}$$

Let us consider a set

$$\mathcal{R} = \left\{D^{(\mathrm{R})}(g_1), D^{(\mathrm{R})}(g_2), \cdots, D^{(\mathrm{R})}(g_n)\right\}. \tag{15.87}$$

Then, the set $\mathcal{R}$ is said to be a regular representation of the group $\mathcal{G}$.

In fact, $\mathcal{R}$ is a representation. This is confirmed as follows: In (15.85), if $g_i^{-1} g_v g_j = e$, $\delta\left(g_i^{-1} g_v g_j\right) = 1$. This occurs when $g_v g_j = g_i$ (A). Meanwhile, let us consider a situation where $g_j^{-1} g_\mu g_k = e$. This occurs when $g_\mu g_k = g_j$ (B). Replacing $g_j$ in (A) with that in (B), we have

$$g_v g_\mu g_k = g_i. \tag{15.88}$$

That is,

$$g_i^{-1} g_v g_\mu g_k = e. \tag{15.89}$$

If we choose $g_i$ and $g_v$, $g_j$ is uniquely decided from (A). If $g_\mu$ is separately chosen, then $g_k$ is uniquely decided from (B) as well, because $g_j$ has already been uniquely decided. Thus, performing the following matrix calculations, we get

$$\begin{aligned} \sum_j \left[D^{(\mathrm{R})}(g_v)\right]_{ij} \left[D^{(\mathrm{R})}(g_\mu)\right]_{jk} &= \sum_j \delta\left(g_i^{-1} g_v g_j\right) \delta\left(g_j^{-1} g_\mu g_k\right) = \delta\left(g_i^{-1} g_v g_\mu g_k\right) \\ &= \left[D^{(\mathrm{R})}\left(g_v g_\mu\right)\right]_{ik}. \end{aligned} \tag{15.90}$$

Rewriting (15.90) in a matrix product form, we get

$$D^{(\mathrm{R})}(g_\nu)D^{(\mathrm{R})}\left(g_\mu\right) = D^{(\mathrm{R})}\left(g_\nu g_\mu\right). \tag{15.91}$$

Thus, $D^{(\mathrm{R})}$ is certainly a representation.

To further confirm this, let us think of an example.

*Example 15.2* Let us consider a thiophene molecule that we have already examined in Sect. 14.2. In a multiplication table, we arrange $E$, $C_2$, $\sigma_v, \sigma_{v'}$ in a first column and their inverse element $E$, $C_2$, $\sigma_v, \sigma_{v'}$ in a first row. In this case, the inverse element is the same as original element itself. Paying attention, e.g., to $C_2$, we allocate the number 1 on the place where $C_2$ appears and the number 0 otherwise. Then, that matrix is a regular representation of $C_2$; see Table 15.3. Thus as $D^{(\mathrm{R})}(C_2)$, we get

$$D^{(\mathrm{R})}(C_2) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \tag{15.92}$$

As evidenced in (15.92), the rearrangement theorem ensures that the number 1 appears once and only once in each column and each row in such a way that individual column and row vectors become linearly independent. Thus, at the same time, we confirm that the matrix is unitary.

Another characteristic of the regular representation is that the identity is represented by an identity matrix. In this example, we have

$$D^{(\mathrm{R})}(E) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{15.93}$$

For the other symmetry operations, we have

$$D^{(\mathrm{R})}(\sigma_v) = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \quad D^{(\mathrm{R})}(\sigma_{v'}) = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

**Table 15.3** How to make a regular representation of $C_{2v}$

| $C_{2v}$ | $E^{-1}$ | $C_2(z)^{-1}$ | $\sigma_v(zx)^{-1}$ | $\sigma'_v(yz)^{-1}$ |
|---|---|---|---|---|
| $E$ | $E$ | $C_2$ | $\sigma_v$ | $\sigma'_v$ |
| $C_2(z)$ | $C_2$ | $E$ | $\sigma'_v$ | $\sigma_v$ |
| $\sigma_v(zx)$ | $\sigma_v$ | $\sigma'_v$ | $E$ | $C_2$ |
| $\sigma'_v(yz)$ | $\sigma'_v$ | $\sigma_v$ | $C_2$ | $E$ |

Let $\chi^{(\mathrm{R})}(g_v)$ be a character of the regular representation. Then, according to the definition of (15.85),

$$\chi^{(\mathrm{R})}(g_v) = \sum_{i=1}^{n} \delta\left(g_i^{-1} g_v g_i\right) = \begin{cases} n & \text{for } g_v = e, \\ 0 & \text{for } g_v \neq e. \end{cases} \tag{15.94}$$

As can be seen from (15.92), the regular representation is reducible because the matrix is decomposed into block matrices. Therefore, the representation can be reduced to a direct sum of irreducible representations such that

$$D^{(\mathrm{R})} = \sum_{\alpha} q_\alpha D^{(\alpha)}, \tag{15.95}$$

where $q_\alpha$ is a positive integer or zero and $D^{(\alpha)}$ is an irreducible representation. Then, from (15.81), we have

$$\chi^{(\mathrm{R})}(g_v) = \sum_{\alpha} q_\alpha \chi^{(\alpha)}(g_v), \tag{15.96}$$

where $\chi^{(\alpha)}$ is a trace of the irreducible representation $D^{(\alpha)}$. Using (15.83) and (15.94),

$$\begin{aligned} q_\alpha &= \frac{1}{n} \sum_{g} \chi^{(\alpha)}(g)^* \chi^{(\mathrm{R})}(g) = \frac{1}{n} \chi^{(\alpha)}(e)^* \chi^{(\mathrm{R})}(e) = \frac{1}{n} \chi^{(\alpha)}(e)^* n = \chi^{(\alpha)}(e)^* \\ &= d_\alpha. \end{aligned} \tag{15.97}$$

Note that a dimension $d_\alpha$ of the representation $D^{(\alpha)}$ is equal to a trace of its identity matrix. Also, notice from (15.95) to (15.97) that $D^{(\mathrm{R})}$ contains every irreducible representation $D^{(\alpha)}$ $d_\alpha$ times. To show it more clearly, Table 15.4 gives a character table of $C_{2v}$. The regular representation matrices are given in Example 15.2. For this, we have

$$D^{(\mathrm{R})}(C_{2v}) = A_1 + A_2 + B_1 + B_2. \tag{15.98}$$

**Table 15.4** Character table of $C_{2v}$

| $C_{2v}$ | $E$ | $C_2(z)$ | $\sigma_v(zx)$ | $\sigma_v'(yz)$ | |
|---|---|---|---|---|---|
| $A_1$ | 1 | 1 | 1 | 1 | $z; x^2, y^2, z^2$ |
| $A_2$ | 1 | 1 | −1 | −1 | $xy$ |
| $B_1$ | 1 | −1 | 1 | −1 | $x; zx$ |
| $B_2$ | 1 | −1 | −1 | 1 | $y; yz$ |

This relation obviously indicates that all the irreducible representations of $C_{2v}$ are contained one time (that is equal to the dimension of representation of $C_{2v}$).

Returning to (15.94) and (15.96) and replacing $q_\alpha$ with $d_\alpha$ there, we get

$$\sum_\alpha d_\alpha \chi^{(\alpha)}(g_v) = \begin{cases} n & \text{for } g_v = e, \\ 0 & \text{for } g_v \neq e. \end{cases} \tag{15.99}$$

In particular, when $g_v = e$, again we have $\chi^{(\alpha)}(e) = d_\alpha$ ($d_\alpha$ is a real number!). That is,

$$\sum_\alpha d_\alpha^2 = n. \tag{15.100}$$

This is a very important relation in the representation theory of a finite group in that (15.100) sets an upper limit to the number of irreducible representations and their dimensions. That number cannot exceed the order of a group.

In (15.76) and (15.78), we have shown the orthogonality relationship between traces. We have another important orthogonality relationship between them. To prove this theorem, we need a notion of group algebra [4]. The argument is as follows:

(a) Let us think of a set comprising group elements expressed as

$$\aleph = \sum_g a_g g, \tag{15.101}$$

where $g$ is a group element of a group $\mathcal{g} = \{g_1, g_2, \cdots, g_n\}$; $a_g$ is an arbitrarily chosen complex number; $\Sigma_g$ means that summation should be taken over group elements. Let $\aleph'$ be another set similarly defined as (15.101). That is,

$$\aleph' = \sum_{g'} a'_{g'} g'. \tag{15.102}$$

Then we can define a following sum:

$$\aleph + \aleph' = \sum_g a_g g + \sum_{g'} a'_{g'} g' = \sum_g \left( a_g g + a'_g g' \right) = \sum_g \left( a_g + a'_g \right) g. \tag{15.103}$$

Also, we get

$$\begin{aligned}\aleph \cdot \aleph' &= \left(\sum_g a_g g\right)\left(\sum_{g'} a'_{g'} g'\right) = \sum_{g'}\left(\sum_g a_g g g'\right) a'_{g'} \\ &= \sum_{g'^{-1}}\left(\sum_g a_g g g'^{-1}\right) a'_{g'^{-1}} = \sum_{g'^{-1}}\left(\sum_{gg'} a_{gg'} g g' g'^{-1}\right) a'_{g'^{-1}} \\ &= \sum_{g'^{-1}}\left(\sum_{gg'} a_{gg'} g\right) a'_{g'^{-1}} = \sum_{gg'}\left(\sum_{g'^{-1}} a_{gg'} a'_{g'^{-1}}\right) g \\ &= \sum_g \left(\sum_{g'} a_{gg'} a'_{g'^{-1}}\right) g,\end{aligned} \tag{15.104}$$

where we used the rearrangement theorem and suitable exchange of group elements. Thus, we see that the above-defined set $\aleph$ is closed under summations (i.e., linear combinations) and multiplications. A set closed under summations and multiplications is said to be an algebra. If the set forms a group, the said set is called a *group algebra*.

So far we have treated calculations as a multiplication between two elements $g$ and $g'$, i.e., $g \diamond g'$ (see Sect. 13.1). Now we start regarding the calculations as summation as well. In that case, group elements act as basis vectors in a vector space. Bearing this in mind, let us further define specific group algebra.

(b) Let $K_i$ be the $i$th conjugacy class of a group $\mathscr{g} = \{g_1, g_2, \cdots, g_n\}$. Also, let $K_i$ be such that

$$K_i = \left\{A_1^{(i)}, A_2^{(i)}, \cdots, A_{k_i}^{(i)}\right\}, \tag{15.105}$$

where $k_i$ is the number of elements belonging to $K_i$. Now think of a set $gK_ig^{-1}\left(^\forall g \in \mathscr{g}\right)$. Then, due to the definition of a class, we have

$$gK_ig^{-1} \subset K_i. \tag{15.106}$$

Multiplying $g^{-1}$ from the left and $g$ from the right of both sides, we get $K_i \subset g^{-1}K_ig$. Since $g$ is arbitrarily chosen, replacing $g$ with $g^{-1}$, we have

$$K_i \subset gK_ig^{-1}. \tag{15.107}$$

Therefore, we get

$$gK_ig^{-1} = K_i. \tag{15.108}$$

Meanwhile, for $A_\alpha^{(i)}, A_\beta^{(i)} \in K_i; A_\alpha^{(i)} \neq A_\beta^{(i)}$ $(1 \leq \alpha, \beta \leq k_i)$, we have

$$gA_{\alpha}^{(i)}g^{-1} \neq gA_{\beta}^{(i)}g^{-1} \ \left({}^{\forall}g \in \mathcal{G}\right). \tag{15.109}$$

This is because if the equality holds with (15.109), we have $A_{\alpha}^{(i)} = A_{\beta}^{(i)}$, in contradiction.

(c) Let $K$ be a set collecting several classes and described as

$$K = \sum_i a_i K_i, \tag{15.110}$$

where $a_i$ is a positive integer or zero. Thanks to (15.105) and (15.110), we have

$$gKg^{-1} = K. \tag{15.111}$$

Conversely, if a group algebra $K$ satisfies (15.111), $K$ can be expressed as a sum of classes such as (15.110). Here suppose that $K$ is not expressed by (15.110), but described by

$$K = \sum_i a_i K_i + Q, \tag{15.112}$$

where $Q$ is an "incomplete" set that does not form a class. Then,

$$\begin{aligned} gKg^{-1} &= g\left(\sum_i a_i K_i + Q\right)g^{-1} = g\left(\sum_i a_i K_i\right)g^{-1} + gQg^{-1} = \sum_i a_i K_i + gQg^{-1} \\ &= K = \sum_i a_i K_i + Q, \end{aligned} \tag{15.113}$$

where the equality before the last comes from (15.111). Thus, from (15.113), we get

$$gQg^{-1} = Q \ \left({}^{\forall}g \in \mathcal{G}\right). \tag{15.114}$$

By definition of the classes, this implies that $Q$ must form a "complete" class, in contradiction to the supposition. Thus, (15.110) holds.

In Sect. 13.3, we described several characteristics of an invariant subgroup. As readily seen from (13.11) to (15.111), any invariant subgroup consists of two or more entire classes. Conversely, if a group comprises entire classes, it must be an invariant subgroup.

(d) Let us think of a product of classes. Let $K_i$ and $K_j$ be sets described as (15.105). We define a product $K_iK_j$ as a set containing products $A_{\alpha}^{(i)}A_{\beta}^{(j)}$ $\left(1 \leq \alpha \leq K_i, 1 \leq \beta \leq K_j\right)$. That is,

$$K_i K_j = \sum_{l=1}^{g_i} \sum_{m=1}^{g_j} A_l^{(i)} A_m^{(j)}. \tag{15.115}$$

Multiplying (15.115) by $g\left({}^{\forall}g \in \mathscr{g}\right)$ from the left and by $g^{-1}$from the right, we get

$$gK_iK_jg^{-1} = (gK_ig)\left(g^{-1}K_jg^{-1}\right) = K_iK_j. \tag{15.116}$$

From the above discussion, we get

$$K_iK_j = \sum_l c_{ijl}K_l, \tag{15.117}$$

where $c_{ijl}$ is a positive integer or zero. In fact, when we take $gK_iK_jg^{-1}$, we merely permute the terms in (15.117).

(e) If two group elements of the group $\mathscr{g} = \{g_1, g_2, \cdots, g_n\}$ are conjugate to each other, their inverse elements are conjugate to each other as well. In fact, suppose that for $g_\mu, g_\nu \in K_i$, we have

$$g_\mu = gg_\nu g^{-1}. \tag{15.118}$$

Then, taking the inverse of (15.118), we get

$$g_\mu^{-1} = gg_\nu^{-1}g^{-1}. \tag{15.119}$$

Thus, given a class $K_i$, there exists another class $K_{i'}$ that consists of the inverses of the elements of $K_i$. If $g_\mu \neq g_\nu$, $g_\mu^{-1} \neq g_\nu^{-1}$. Therefore, $K_i$ and $K_{i'}$ are of the same order. Suppose that the number of elements contained in $K_i$ is $k_i$ and that in $K_{i'}$ is $k_{i'}$. Then, we have

$$k_i = k_{i'}. \tag{15.120}$$

If $K_j \neq K_{i'}$ (i.e., $K_j$ is not identical with $K_{i'}$ as a set), $K_iK_j$ does not contain $e$. In fact, suppose that for $g_\rho \in K_i$, there was a group element $b \in K_j$ such that $bg_\rho = e$. Then, we would have

$$b = g_\rho^{-1} \quad \text{and} \quad b \in K_{i'}. \tag{15.121}$$

This would mean that $b \in K_j \cap K_{i'}$, implying that $K_j = K_{i'}$ by definition of class. It is in contradiction to $K_j \neq K_{i'}$. Consequently, $K_iK_j$ does not contain $e$. Taking the product of the classes $K_i$ and $K_{i'}$, we obtain the identity $e$ precisely $k_i$ times. Rewriting (15.117), we have

$$K_i K_j = c_{ij1} K_1 + \sum_{l \neq 1} c_{ijl} K_l, \tag{15.122}$$

where $K_1 = \{e\}$. As mentioned below, we are most interested in the first term in (15.122). In (15.122), if $K_j = K_{i'}$, we have

$$c_{ij1} = k_i. \tag{15.123}$$

On the other hand, if $K_j \neq K_{i'}$, $c_{ij1} = 0$. Summarizing the above arguments, we get

$$c_{ij1} = \begin{cases} k_i & \text{for } K_j = K_{i'}, \\ 0 & \text{for } K_j \neq K_{i'}. \end{cases} \tag{15.124}$$

Or, symbolically we write

$$c_{ij1} = k_i \delta_{ji'}. \tag{15.125}$$

## 15.6 Classes and Irreducible Representations

After the aforementioned considerations, we have the following theorem:

**Theorem 15.5** *Let $\mathcal{g} = \{g_1, g_2, \cdots, g_n\}$ be a group. A trace of irreducible representations satisfies the following orthogonality relation*:

$$\sum_{\alpha=1}^{n_r} \chi^{(\alpha)}(K_i) \chi^{(\alpha)}(K_j)^* = \frac{n}{k_i} \delta_{ij}, \tag{15.126}$$

where summation $\alpha$ is taken over all inequivalent $n_r$ irreducible representations; $K_i$ and $K_j$ indicate conjugacy classes; $k_i$ denotes the number of elements contained in the $i$th class $K_i$.

*Proof* Rewriting (15.108), we have

$$gK_i = K_i g \left( {}^{\forall} g \in \mathcal{g} \right). \tag{15.127}$$

Since a homomorphic correspondence holds between a group element and its representation matrix, a similar correspondence holds as well with (15.127). Let $\widehat{K}_i$ be a sum of $K_i$ matrices of the $a$th irreducible representation $D^{(\alpha)}$ and let $\widehat{K}_i$ be expressed as

$$\widehat{K}_i = \sum_{g \in K_i} D^{(\alpha)}. \tag{15.128}$$

Note that in (15.128), $D^{(\alpha)}$ functions as a linear transformation with respect to a group algebra. From (15.127), we have

$$D^{(\alpha)}\widehat{K}_i = \widehat{K}_i D^{(\alpha)} \left({}^{\forall} g \in \mathcal{G}\right). \tag{15.129}$$

Since $D^{(\alpha)}$ is an irreducible representation, $\widehat{K}_i$ must be expressed on the basis of Schur's Second Lemma as

$$\widehat{K}_i = \lambda E. \tag{15.130}$$

To determine $\lambda$, we take a trace of both sides of (15.130). Then, from (15.128) to (15.130), we get

$$k_i \chi^{(\alpha)}(K_i) = \lambda d_\alpha. \tag{15.131}$$

Thus, we get

$$\widehat{K}_i = \frac{k_i}{d_\alpha} \chi^{(\alpha)}(K_i) E. \tag{15.132}$$

Next, corresponding to (15.114), we have

$$\widehat{K}_i \widehat{K}_j = \sum_l c_{ijl} \widehat{K}_l. \tag{15.133}$$

Replacing $\widehat{K}_l$ in (15.133) with that of (15.132), we get

$$k_i k_j \chi^{(\alpha)}(K_i) \chi^{(\alpha)}\left(K_j\right) = d_\alpha \sum_l c_{ijl} k_l \chi^{(\alpha)}(K_l). \tag{15.134}$$

Returning to (15.99) and rewriting it, we have

$$\sum_\alpha d_\alpha \chi^{(\alpha)}(K_i) = n\delta_{i1}, \tag{15.135}$$

where again we have $K_1 = \{e\}$. With respect to (15.134), we sum over all the irreducible representations $\alpha$. Then we have

$$\sum_\alpha k_i k_j \chi^{(\alpha)}(K_i) \chi^{(\alpha)}\left(K_j\right) = \sum_l \left[ c_{ijl} k_l \sum_\alpha d_\alpha \chi^{(\alpha)}(K_l) \right] = \sum_l c_{ijl} k_l n \delta_{l1} = c_{ij1} n. \tag{15.136}$$

In (15.136), we remark that $k_1 = 1$, meaning that the number of group elements that $K_1 = \{e\}$ contains is 1. Rewriting (15.136), we get

$$k_i k_j \sum_{\alpha} \chi^{(\alpha)}(K_i)\chi^{(\alpha)}\left(K_j\right) = c_{ij1} n = k_i n \delta_{ji'}, \tag{15.137}$$

where we used (15.125) with the last equality. Moreover, using

$$\chi^{(\alpha)}(K_{i'}) = \chi^{(\alpha)}(K_i)^*, \tag{15.138}$$

we get

$$k_i k_j \sum_{\alpha} \chi^{(\alpha)}(K_i)\chi^{(\alpha)}(K_j)^* = c_{ij1} n = k_i n \delta_{ji}. \tag{15.139}$$

Rewriting (15.139), we finally get

$$\sum_{\alpha=1}^{n_r} \chi^{(\alpha)}(K_i)\chi^{(\alpha)}(K_j)^* = \frac{n}{k_i}\delta_{ij}. \tag{15.126}$$

Equations (15.78) and (15.126) are well-known as orthogonality relations. So far, we have no idea about the mutual relationship between the two numbers $n_c$ and $n_r$ in magnitude. In (15.78), let us consider a following set $S$

$$S = \left\{\sqrt{K_1}\chi^{(\alpha)}(K_1), \sqrt{K_2}\chi^{(\alpha)}(K_2), \cdots, \sqrt{K_{n_c}}\chi^{(\alpha)}(K_{n_c})\right\}. \tag{15.140}$$

Viewing individual components in $S$ as coordinates of a $n_c$-dimensional vector, (15.78) can be considered as an inner product as expressed using (complex) coordinates of the two vectors. At the same time, (15.78) represents an orthogonal relationship between the vectors. Since we can obtain at most $n_c$ mutually orthogonal (i.e., linearly independent) vectors in a $n_c$-dimensional space, for the number ($n_r$) of such vectors, we have

$$n_r \leq n_c. \tag{15.141}$$

Here $n_r$ is equal to the number of different $\alpha$, i.e., the number of irreducible representations.

Meanwhile, in (15.126), we consider a following set $S'$

$$S' = \left\{\chi^{(1)}(K_i), \chi^{(2)}(K_i), \cdots, \chi^{(n_r)}(K_i)\right\}. \tag{15.142}$$

Similarly, individual components in $S'$ can be considered as coordinates of a $n_r$-dimensional vector. Again, (15.126) implies the orthogonality relation among vectors. Therefore, as for the number ($n_c$) of mutually orthogonal vectors, we have

$$n_c \leq n_r. \tag{15.143}$$

Thus, from (15.141) to (15.143), we finally reach a simple but very important conclusion about the relationship between $n_c$ and $n_r$ such that

$$n_r = n_c. \tag{15.144}$$

That is, the number ($n_r$) of inequivalent irreducible representations is equal to that ($n_c$) of conjugacy classes of the group. An immediate and important consequence of (15.144) together with (15.100) is that the representation of an Abelian group is one-dimensional. This is because in the Abelian group individual group elements constitute a conjugacy class. We have $n_c = n$ accordingly.

## 15.7 Projection Operators

In Sect. 15.2, we have described how basis vectors (or functions) are transformed by a symmetry operation. In that case, the symmetry operation is performed by a group element that belongs to a transformation group. More specifically, given a group $\mathcal{G} = \{g_1, g_2, \cdots, g_n\}$ and a set of basis vectors $\psi_1, \psi_2, \cdots,$ and $\psi_d$, the basis vectors are transformed by $g_i \in \mathcal{G} (1 \leq i \leq n)$ such that

$$g_i(\psi_v) = \sum_{\mu=1}^{d} \psi_\mu D_{\mu v}(g_i). \tag{15.22}$$

We may well ask then how we can construct such basis vectors. In this section, we address this question. A central concept about this is a projection operator. We have already studied the definition and basic properties of the projection operators. In this section, we deal with them bearing in mind that we apply the group theory to molecular science, especially to quantum chemical calculations.

In Sect. 15.6, we examined the permissible number of irreducible representations. We have reached a conclusion that the number ($n_r$) of inequivalent irreducible representations is equal to that ($n_c$) of conjugacy classes of the group. According to this conclusion, we modify the above Eq. (15.22) such that

$$g\left(\psi_i^{(\alpha)}\right) = \sum_{j=1}^{d_\alpha} \psi_j^{(\alpha)} D_{ji}^{(\alpha)}(g), \tag{15.145}$$

where $\alpha$ and $d_\alpha$ denote the $\alpha$th irreducible representation and its dimension, respectively; the subscript $i$ is omitted from $g_i$ for simplicity. This naturally leads to the next question of how the basis vectors $\psi_j^{(\alpha)}$ are related to $\psi_j^{(\beta)}$ that are basis vectors as well, but belong to a different irreducible representation $\beta$.

Suppose that we have an arbitrarily chosen function $f$. Then, $f$ is assumed to contain various components of different irreducible representations. Thus, let us assume that $f$ is decomposed into the component such that

$$f = \sum_{\alpha} \sum_{m} c_m^{(\alpha)} \psi_m^{(\alpha)}, \tag{15.146}$$

where $c_m^{(\alpha)}$ is a coefficient of the expansion and $\psi_m^{(\alpha)}$ is the $m$th component of $\alpha$th irreducible representation.

Now, let us define the following operator:

$$P_{l(m)}^{(\alpha)} = \frac{d_\alpha}{n} \sum_{g} D_{lm}^{(\alpha)}(g)^* g, \tag{15.147}$$

where $\sum_g$ means that the summation should be taken over all $n$ group elements $g$; $d_\alpha$ denotes a dimension of the representation $D^{(\alpha)}$. Operating $P_{l(m)}^{(\alpha)}$ on $f$, we have

$$\begin{aligned}
P_{l(m)}^{(\alpha)} f &= \frac{d_\alpha}{n} \sum_{g} D_{lm}^{(\alpha)}(g)^* g f = \frac{d_\alpha}{n} \sum_{g} D_{lm}^{(\alpha)}(g)^* g \sum_{v} \sum_{k} c_k^{(v)} \psi_k^{(v)} \\
&= \frac{d_\alpha}{n} \sum_{g} D_{lm}^{(\alpha)}(g)^* \sum_{v} \sum_{k} c_k^{(v)} g \psi_k^{(v)} \\
&= \frac{d_\alpha}{n} \sum_{g} D_{lm}^{(\alpha)}(g)^* \sum_{v} \sum_{k} c_k^{(v)} \sum_{j=1}^{d_v} \psi_j^{(v)} D_{jk}^{(v)}(g) \\
&= \frac{d_\alpha}{n} \sum_{v} \sum_{k} c_k^{(v)} \sum_{j=1}^{d_v} \psi_j^{(v)} \left[ \sum_{g} D_{lm}^{(\alpha)}(g)^* D_{jk}^{(v)}(g) \right] \\
&= \frac{d_\alpha}{n} \sum_{v} \sum_{k} c_k^{(v)} \sum_{j=1}^{d_v} \psi_j^{(v)} \frac{n}{d_\alpha} \delta_{\alpha v} \delta_{lj} \delta_{mk} = c_m^{(\alpha)} \psi_l^{(\alpha)},
\end{aligned} \tag{15.148}$$

where we used (15.51) for the equality before the last.

Thus, an implication of the operator $P_{l(m)}^{(\alpha)}$ is that if $f$ contains the component $\psi_m^{(\alpha)}$, i.e., $c_m^{(\alpha)} \neq 0$, $P_{l(m)}^{(\alpha)}$ plays a role in *extracting* the component $c_m^{(\alpha)} \psi_m^{(\alpha)}$ from $f$ and then *converting* it to $c_m^{(\alpha)} \psi_l^{(\alpha)}$. If the $\psi_m^{(\alpha)}$ component is not contained in $f$, that means from (15.148) $P_{l(m)}^{(\alpha)} f = 0$. In that case, we choose a more suitable function for $f$. In this context, we will investigate an example of a quantum chemical calculation later.

In the above case, let us call $P_{l(m)}^{(\alpha)}$ a projection operator *sensu lato*. In Sect. 12.1, we dealt with several aspects of the projection operators. There we have mentioned that a projection operator should be idempotent and Hermitian in a rigorous sense

(Definition 12.1). To address another important aspect of the projection operator, let us prove an important relation in the following theorem.

**Theorem 15.6** [1] *Let* $P^{(\alpha)}_{l(m)}$ and $P^{(\beta)}_{s(t)}$ *be projection operators defined in* (15.147). *Then, the following equation holds*:

$$P^{(\alpha)}_{l(m)} P^{(\beta)}_{s(t)} = \delta_{\alpha\beta}\delta_{ms} P^{(\alpha)}_{l(t)}. \tag{15.149}$$

*Proof* We have

$$\begin{aligned}
P^{(\alpha)}_{l(m)} P^{(\beta)}_{s(t)} &= \frac{d_\alpha}{n}\sum_g D^{(\alpha)}_{lm}(g)^* g \frac{d_\beta}{n}\sum_{g'} D^{(\beta)}_{st}(g')^* g' \\
&= \frac{d_\alpha}{n}\sum_g D^{(\alpha)}_{lm}(g)^* g \frac{d_\beta}{n}\sum_{g'} D^{(\beta)}_{st}(g^{-1}g')^* g^{-1}g' \\
&= \frac{d_\alpha}{n}\frac{d_\beta}{n}\sum_g\sum_{g'} D^{(\alpha)}_{lm}(g)^* D^{(\beta)}_{st}(g^{-1}g')^* g g^{-1} g' \\
&= \frac{d_\alpha}{n}\frac{d_\beta}{n}\sum_g\sum_{g'} D^{(\alpha)}_{lm}(g)^* \left\{ \left[D^{(\beta)}(g)\right]^\dagger D^{(\beta)}(g') \right\}^*_{st} e g' \\
&= \frac{d_\alpha}{n}\frac{d_\beta}{n}\sum_g\sum_{g'} D^{(\alpha)}_{lm}(g)^* \left\{ \sum_k D^{(\beta)}(g)^*_{ks} D^{(\beta)}(g')_{kt} \right\}^* g' \\
&= \frac{d_\alpha}{n}\frac{d_\beta}{n}\sum_k\sum_{g'} \left[ \sum_g D^{(\alpha)}_{lm}(g)^* D^{(\beta)}(g)_{ks} \right] D^{(\beta)}(g')^*_{kt} g' \\
&= \frac{d_\alpha}{n}\frac{d_\beta}{n}\sum_k\sum_{g'} \frac{n}{d_\alpha}\delta_{\alpha\beta}\delta_{lk}\delta_{ms} D^{(\beta)}(g')^*_{kt} g' = \frac{d_\beta}{n}\sum_{g'} \delta_{\alpha\beta}\delta_{ms} D^{(\beta)}(g')^*_{lt} g' \\
&= \delta_{\alpha\beta}\delta_{ms}\frac{d_\beta}{n}\sum_{g'} D^{(\beta)}(g')^*_{lt} g' = \delta_{\alpha\beta}\delta_{ms} P^{(\alpha)}_{l(t)}.
\end{aligned} \tag{15.150}$$

This completes the proof.

In the above proof, we used the rearrangement theorem and grand orthogonality theorem (GOT) as well as the homomorphism and unitarity of the representation matrices.

Comparing (15.146) and (15.148), we notice that the term $c^{(\alpha)}_m \psi^{(\alpha)}_m$ is not extracted entirely, but $c^{(\alpha)}_m \psi^{(\alpha)}_l$ is given instead. This is due to the linearity of $P^{(\alpha)}_{l(m)}$. Nonetheless, this is somewhat inconvenient for a practical purpose. To overcome this inconvenience, we modify (15.149). In (15.149), putting $s = m$, we have

$$P^{(\alpha)}_{l(m)} P^{(\beta)}_{m(t)} = \delta_{\alpha\beta} P^{(\alpha)}_{l(t)}. \tag{15.151}$$

We further modify the relation. Putting $\beta = \alpha$, we have

$$P^{(\alpha)}_{l(m)} P^{(\alpha)}_{m(t)} = P^{(\alpha)}_{l(t)}. \tag{15.152}$$

Putting $t = l$ furthermore,

$$P^{(\alpha)}_{l(m)} P^{(\alpha)}_{m(l)} = P^{(\alpha)}_{l(l)}. \tag{15.153}$$

In particular, putting $m = l$ moreover, we get

$$P^{(\alpha)}_{l(l)} P^{(\alpha)}_{l(l)} = \left[P^{(\alpha)}_{l(l)}\right]^2 = P^{(\alpha)}_{l(l)}. \tag{15.154}$$

In fact, putting $m = l$ in (15.148), we get

$$P^{(\alpha)}_{l(l)} f = c^{(\alpha)}_l \psi^{(\alpha)}_l. \tag{15.155}$$

This means that the term $c^{(\alpha)}_l \psi^{(\alpha)}_l$ has been extracted entirely.

Moreover, in (15.149), putting $\beta = \alpha$, $s = l$, and $t = m$, we have

$$P^{(\alpha)}_{l(m)} P^{(\alpha)}_{l(m)} = \left[P^{(\alpha)}_{l(m)}\right]^2 = \delta_{ml} P^{(\alpha)}_{l(m)}.$$

Therefore, for $P^{(\alpha)}_{l(m)}$ to be an idempotent operator, we must have $m = l$. That is, of various operators $P^{(\alpha)}_{l(m)}$, only $P^{(\alpha)}_{l(l)}$ is eligible for an idempotent operator.

Meanwhile, fully describing $P^{(\alpha)}_{l(l)}$, we have

$$P^{(\alpha)}_{l(l)} = \frac{d_\alpha}{n} \sum_g D^{(\alpha)}_{ll}(g)^* g. \tag{15.156}$$

Taking complex conjugate transposition (i.e., adjoint) of (15.156), we have

$$\begin{aligned} \left[P^{(\alpha)}_{l(l)}\right]^\dagger &= \frac{d_\alpha}{n} \sum_g D^{(\alpha)}_{ll}(g) g^\dagger = \frac{d_\alpha}{n} \sum_{g^{-1}} D^{(\alpha)}_{ll}\left(g^{-1}\right)\left(g^{-1}\right)^\dagger \\ &= \frac{d_\alpha}{n} \sum_g D^{(\alpha)}_{ll}(g)^* g = P^{(\alpha)}_{l(l)}, \end{aligned} \tag{15.157}$$

where we used unitarity of $g$ (with the third equality) and equivalence of summation over $g$ and $g^{-1}$. Notice that the notation of $g^\dagger$ is less common. This should be interpreted as meaning that $g^\dagger$ operates on a vector constituting a representation space. Thus, notation $g^\dagger$ implies that $g^\dagger$ is equivalent to its unitary representation

matrix $D^{(\alpha)}(g)$. Note also that $D^{(\alpha)}_{ll}$ is not a matrix but a (complex) *number*. Namely, in (15.156), $D^{(\alpha)}_{ll}(g)^*$ is a coefficient of an operator $g$.

Equations (15.154) and (15.157) establish that $P^{(\alpha)}_{l(l)}$ is a projection operator in a rigorous sense. Let us call $P^{(\alpha)}_{l(l)}$ a projection operator sensu *stricto* accordingly. Also, we notice that $c^{(\alpha)}_m \psi^{(\alpha)}_m$ is entirely extracted from an arbitrary function $f$ including a coefficient. This situation resembles that of (10.205).

Regarding $P^{(\alpha)}_{l(m)}$, on the other hand, we have

$$\begin{aligned}\left[P^{(\alpha)}_{l(m)}\right]^\dagger &= \frac{d_\alpha}{n}\sum_g D^{(\alpha)}_{lm}(g)g^\dagger = \frac{d_\alpha}{n}\sum_{g^{-1}} D^{(\alpha)}_{lm}\left(g^{-1}\right)\left(g^{-1}\right)^\dagger \\ &= \frac{d_\alpha}{n}\sum_g D^{(\alpha)}_{ml}(g)^* g = P^{(\alpha)}_{m(l)}.\end{aligned} \tag{15.158}$$

Hence, $P^{(\alpha)}_{l(m)}$ is not Hermitian. We have many other related equations. For instance,

$$\left[P^{(\alpha)}_{l(m)}P^{(\alpha)}_{m(l)}\right]^\dagger = \left[P^{(\alpha)}_{m(l)}\right]^\dagger\left[P^{(\alpha)}_{l(m)}\right]^\dagger = P^{(\alpha)}_{l(m)}P^{(\alpha)}_{m(l)}. \tag{15.159}$$

Therefore, $P^{(\alpha)}_{l(m)}P^{(\alpha)}_{m(l)}$ is Hermitian, recovering the relation (15.153).

In (15.149), putting $m = l$ and $t = s$, we get

$$P^{(\alpha)}_{l(l)}P^{(\beta)}_{s(s)} = \delta_{\alpha\beta}\delta_{ls}P^{(\alpha)}_{l(s)}. \tag{15.160}$$

As in the case of (15.146), we assume that $h$ is described as

$$h = \sum_\alpha \sum_m d^{(\alpha)}_m \phi^{(\alpha)}_m, \tag{15.161}$$

where $\phi^{(\alpha)}_m$ is transformed in the same manner as that for $\psi^{(\alpha)}_m$ in (15.146), but linearly independent of $\psi^{(\alpha)}_m$. Namely, we have

$$g\left(\phi^{(\alpha)}_i\right) = \sum_{j=1}^{d_\alpha} \phi^{(\alpha)}_j D^{(\alpha)}_{ji}(g). \tag{15.162}$$

Tangible examples can be seen in Chap. 16.

Operating $P^{(\alpha)}_{l(l)}$ on both sides of (15.155), we have

$$\left[P^{(\alpha)}_{l(l)}\right]^2 f = P^{(\alpha)}_{l(l)}\left[c^{(\alpha)}_l \psi^{(\alpha)}_l\right] = P^{(\alpha)}_{l(l)} f = c^{(\alpha)}_l \psi^{(\alpha)}_l, \tag{15.163}$$

where with the second equality we used $[P^{(\alpha)}_{l(l)}]^2 = P^{(\alpha)}_{l(l)}$. That is, we have

$$P^{(\alpha)}_{l(l)}\left[c^{(\alpha)}_l \psi^{(\alpha)}_l\right] = c^{(\alpha)}_l \psi^{(\alpha)}_l. \tag{15.164}$$

This equation means that $c^{(\alpha)}_l \psi^{(\alpha)}_l$ is an eigenfunction corresponding to an eigenvalue 1 of $P^{(\alpha)}_{l(l)}$. In other words, once $c^{(\alpha)}_l \psi^{(\alpha)}_l$ is extracted from $f$, it belongs to the "position" $l$ of an irreducible representation $\alpha$. Furthermore, for some constants $c$ and $d$ as well as the functions $f$ and $h$ that appeared in (15.146) and (15.161), we consider a following equation:

$$\begin{aligned} P^{(\alpha)}_{l(l)}\left[cc^{(\alpha)}_l \psi^{(\alpha)}_l + dd^{(\alpha)}_l \phi^{(\alpha)}_l\right] &= cP^{(\alpha)}_{l(l)} c^{(\alpha)}_l \psi^{(\alpha)}_l + dP^{(\alpha)}_{l(l)} d^{(\alpha)}_l \phi^{(\alpha)}_l \\ &= cc^{(\alpha)}_l \psi^{(\alpha)}_l + dd^{(\alpha)}_l \phi^{(\alpha)}_l, \end{aligned} \tag{15.165}$$

where with the last equality we used (15.164). This means that an arbitrary linear combination of $c^{(\alpha)}_l \psi^{(\alpha)}_l$ and $d^{(\alpha)}_l \phi^{(\alpha)}_l$ again belongs to the position $l$ of an irreducible representation $\alpha$. If $\psi^{(\alpha)}_l$ and $\phi^{(\alpha)}_l$ are linearly independent, we can construct two orthonormal basis vectors following Theorem 11.2 (Gram–Schmidt orthonormalization Theorem). If there are other linearly independent vectors $p^{(\alpha)}_l \xi^{(\alpha)}_l$, $q^{(\alpha)}_l \varphi^{(\alpha)}_l$, $\cdots$, where $\xi^{(\alpha)}_l$, $\varphi^{(\alpha)}_l$, etc., belong to the position $l$ of an irreducible representation $\alpha$, then $cc^{(\alpha)}_l \psi^{(\alpha)}_l + dd^{(\alpha)}_l \phi^{(\alpha)}_l + pp^{(\alpha)}_l \xi^{(\alpha)}_l + qq^{(\alpha)}_l \varphi^{(\alpha)}_l + \cdots$ again belongs to the position $l$ of an irreducible representation $\alpha$. Thus, we can construct orthonormal basis vectors of the representation space according to Theorem 11.2.

Regarding arbitrary functions $f$ and $h$ as vectors and using (15.160), we make an inner product of (15.160) such that

$$\begin{aligned} \left\langle h \middle| P^{(\alpha)}_{l(l)} P^{(\beta)}_{s(s)} \middle| f \right\rangle &= \left\langle h \middle| P^{(\alpha)}_{l(l)} P^{(\alpha)}_{l(l)} P^{(\beta)}_{s(s)} \middle| f \right\rangle = \left\langle h \middle| P^{(\alpha)}_{l(l)} \delta_{\alpha\beta}\delta_{ls} P^{(\alpha)}_{l(s)} \middle| f \right\rangle \\ &= \delta_{\alpha\beta}\delta_{ls} \left\langle hP^{(\alpha)}_{l(l)} \middle| P^{(\alpha)}_{l(s)} \middle| f \right\rangle, \end{aligned} \tag{15.166}$$

where with the first equality we used $P^{(\alpha)}_{l(l)} = P^{(\alpha)}_{l(l)} P^{(\alpha)}_{l(l)}$ (15.154). Meanwhile, from (15.148), we have

$$P^{(\alpha)}_{l(s)} \middle| f \right\rangle = c^{(\alpha)}_s \psi^{(\alpha)}_l \tag{15.167}$$

Also using (15.155), we get

$$P^{(\alpha)}_{l(l)} \middle| h \right\rangle = d^{(\alpha)}_l \middle| \phi^{(\alpha)}_l \right\rangle. \tag{15.168}$$

Taking adjoint of (15.168), we have

$$\langle h|\left[P^{(\alpha)}_{l(l)}\right]^{\dagger} = \langle h|P^{(\alpha)}_{l(l)} = \left[d^{(\alpha)}_{l}\right]^{*}\left\langle \phi^{(\alpha)}_{l}\right|, \tag{15.169}$$

where we used $\left[P^{(\alpha)}_{l(l)}\right]^{\dagger} = P^{(\alpha)}_{l(l)}$ (15.157). The relation (15.169) is due to the notation of Sect. 11.3. Substituting (15.167) and (15.168) for (15.166),

$$\left\langle h\middle|P^{(\alpha)}_{l(l)}P^{(\beta)}_{s(s)}\middle|f\right\rangle = \delta_{\alpha\beta}\delta_{ls}\left[d^{(\alpha)}_{l}\right]^{*}c^{(\alpha)}_{s}\left\langle \phi^{(\alpha)}_{l}|\psi^{(\alpha)}_{l}\right\rangle.$$

Meanwhile, (15.166) can also be described as

$$\left\langle h\middle|P^{(\alpha)}_{l(l)}P^{(\beta)}_{s(s)}\middle|f\right\rangle = \left\langle d^{(\alpha)}_{l}\phi^{(\alpha)}_{l}|c^{(\beta)}_{s}\psi^{(\beta)}_{s}\right\rangle = \left[d^{(\alpha)}_{l}\right]^{*}c^{(\beta)}_{s}\left\langle \phi^{(\alpha)}_{l}|\psi^{(\beta)}_{s}\right\rangle. \tag{15.170}$$

For (15.166) and (15.170) to be identical, we must have

$$\delta_{\alpha\beta}\delta_{ls}\left[d^{(\alpha)}_{l}\right]^{*}c^{(\alpha)}_{s}\left\langle \phi^{(\alpha)}_{l}|\psi^{(\alpha)}_{l}\right\rangle = \left[d^{(\alpha)}_{l}\right]^{*}c^{(\beta)}_{s}\left\langle \phi^{(\alpha)}_{l}|\psi^{(\beta)}_{s}\right\rangle.$$

Deleting coefficients, we get

$$\left\langle \phi^{(\alpha)}_{l}|\psi^{(\beta)}_{s}\right\rangle = \delta_{\alpha\beta}\delta_{ls}\left\langle \phi^{(\alpha)}_{l}|\psi^{(\alpha)}_{l}\right\rangle. \tag{15.171}$$

The relation (15.171) is frequently used to estimate whether definite integrals vanish. Functional forms depend upon actual problems we encounter in various situations. We will deal with this problem in Chap. 16 in relation to, e.g., discussion on optical transition and evaluation of overlap integrals.

To evaluate (15.170), if $\alpha \neq \beta$ or $l \neq s$, we get simply

$$\left\langle h\middle|P^{(\alpha)}_{l(l)}P^{(\beta)}_{s(s)}\middle|f\right\rangle = 0 \ (\alpha \neq \beta \text{ or } l \neq s). \tag{15.172}$$

That is, under a given condition $\alpha \neq \beta$ or $l \neq s$, $P^{(\beta)}_{s(s)}|f\rangle$ and $P^{(\alpha)}_{l(l)}|h\rangle$ are orthogonal (see Theorem 11.3 of Sect. 11.4). The relation clearly indicates that functions belonging to different irreducible representations ($\alpha \neq \beta$) are mutually orthogonal. Even though the functions belong to the same irreducible representation, the functions are orthogonal if they are allocated to the different "place" as a basis vector designated by $l$, $s$, etc. Here the place means the index $j$ of $\psi^{(\alpha)}_{j}$ in (15.145) or $\phi^{(\alpha)}_{j}$ in (15.162) that designates the "order" of $\psi^{(\alpha)}_{j}$ within $\psi^{(\alpha)}_{1}, \psi^{(\alpha)}_{2}, \cdots, \psi^{(\alpha)}_{d_\alpha}$ or $\phi^{(\alpha)}_{j}$ within $\phi^{(\alpha)}_{1}, \phi^{(\alpha)}_{2}, \cdots, \phi^{(\alpha)}_{d_\alpha}$. This takes place if the representation is multidimensional (i.e., dimensionality: $d_\alpha$).

In (15.160), putting $\alpha = \beta$, we get

$$P^{(\alpha)}_{l(l)} P^{(\alpha)}_{s(s)} = \delta_{ls} P^{(\alpha)}_{l(s)}.$$

In the above relation, furthermore, unless $l = s$, we have

$$P^{(\alpha)}_{l(l)} P^{(\alpha)}_{s(s)} = 0.$$

Therefore, on the basis of the discussion of Sect. 12.1, $P^{(\alpha)}_{l(l)} + P^{(\alpha)}_{s(s)}$ is a projection operator as well in the case of $l \neq s$. Notice, however, that if $l = s$, $P^{(\alpha)}_{l(l)} + P^{(\alpha)}_{s(s)}$ is *not* a projection operator. Readers can readily show it. Moreover, let us define $P^{(\alpha)}$ as below

$$P^{(\alpha)} \equiv \sum_{l=1}^{d_\alpha} P^{(\alpha)}_{l(l)}. \tag{15.173}$$

Similarly, $P^{(\alpha)}$ is again a projection operator as well.

From (15.156), $P^{(\alpha)}$ in (15.173) can be rewritten as

$$P^{(\alpha)} = \frac{d_\alpha}{n} \sum_g \sum_{l=1}^{d_\alpha} D^{(\alpha)}_{ll}(g)^* g = \frac{d_\alpha}{n} \sum_g \left[\chi^{(\alpha)}(g)\right]^* g. \tag{15.174}$$

Returning to (15.155) and taking summation over $l$ there, we have

$$\sum_{l=1}^{d_\alpha} P^{(\alpha)}_{l(l)} f = \sum_{l=1}^{d_\alpha} c^{(\alpha)}_l \psi^{(\alpha)}_l.$$

Using (15.173), we get

$$P^{(\alpha)} f = \sum_{l=1}^{d_\alpha} c^{(\alpha)}_l \psi^{(\alpha)}_l. \tag{15.175}$$

Thus, an operator $P^{(\alpha)}$ has a clear meaning. That is, $P^{(\alpha)}$ plays a role in extracting all the vectors (or functions) including their coefficients. Defining those functions as

$$\psi^{(\alpha)} \equiv \sum_{l=1}^{d_\alpha} c^{(\alpha)}_l \psi^{(\alpha)}_l, \tag{15.176}$$

we succinctly rewrite (15.175) as

$$P^{(\alpha)} f = \psi^{(\alpha)}. \tag{15.177}$$

In turn, let us calculate $P^{(\beta)}P^{(\alpha)}$. This can be done with (15.174) as follows:

$$P^{(\beta)}P^{(\alpha)} = \frac{d_\beta}{n}\sum_g \left[\chi^{(\beta)}(g)\right]^* g \frac{d_\alpha}{n}\sum_{g'}\left[\chi^{(\alpha)}(g')\right]^* g'. \tag{15.178}$$

To carry out the calculation, (i) first we replace $g'$ with $g^{-1}g'$ and rewrite the summation over $g'$ as that over $g^{-1}g'$. (ii) Using the homomorphism and unitarity of the representation, we rewrite $\left[\chi^{(\alpha)}(g')\right]^*$ as

$$\left[\chi^{(\alpha)}(g')\right]^* = \sum_{i,j}\left[D^{(\alpha)}(g)\right]_{ji}\left[D^{(\alpha)}(g')\right]^*_{ji}. \tag{15.179}$$

Rewriting (15.178), we have

$$\begin{aligned} P^{(\beta)}P^{(\alpha)} &= \frac{d_\beta}{n}\frac{d_\alpha}{n}\sum_g\sum_{k,i,j}\left[D^{(\beta)}(g)\right]^*_{kk}\left[D^{(\alpha)}(g)\right]_{ji}\sum_{g'}\left[D^{(\alpha)}(g')\right]^*_{ji} g' \\ &= \frac{d_\beta}{n}\frac{d_\alpha}{n}\sum_{k,i,j}\frac{n}{d_\beta}\delta_{\alpha\beta}\delta_{kj}\delta_{ki}\sum_{g'}\left[D^{(\alpha)}(g')\right]^*_{ji} g' = \frac{d_\alpha}{n}\sum_k\sum_{g'}\left[D^{(\alpha)}(g')\right]^*_{kk} g' \\ &= \delta_{\alpha\beta}\frac{d_\alpha}{n}\sum_{g'}\left[\chi^{(\alpha)}(g')\right]^* g' = \delta_{\alpha\beta}P^{(\alpha)}. \end{aligned} \tag{15.180}$$

These relationships are anticipated from the fact that both $P^{(\alpha)}$ and $P^{(\beta)}$ are projection operators.

As in the case of (15.160), (15.180) is useful to evaluate an inner product of functions. Again taking an inner product of (15.180) with arbitrary functions $f$ and $g$, we have

$$\left\langle g\middle|P^{(\beta)}P^{(\alpha)}\middle|f\right\rangle = \left\langle g\middle|\delta_{\alpha\beta}P^{(\alpha)}\middle|f\right\rangle = \delta_{\alpha\beta}\left\langle g\middle|P^{(\alpha)}\middle|f\right\rangle. \tag{15.181}$$

If we define $P$ such that

$$P \equiv \sum_{\alpha=1}^{n_r} P^{(\alpha)}, \tag{15.182}$$

then $P$ is once again a projection operator (see Sect. 12.1). Now taking summation in (15.177) over all the irreducible representations $\alpha$, we have

$$\sum_{\alpha=1}^{n_r} P^{(\alpha)}f = \sum_{\alpha=1}^{n_r}\psi^{(\alpha)} = f. \tag{15.183}$$

The function $f$ has been arbitrarily taken and, hence, we get

$$P = E. \tag{15.184}$$

As mentioned in Example 15.1, the concept of the representation space is not only important but also very useful for addressing various problems of physics and chemistry. For instance, molecular orbital methods to be dealt with in Chap. 16 consider a representation space whose dimension is equal to the number of electrons of a molecule about which we wish to know, e.g., energy eigenvalues of those electrons. In that case, the dimension of representation space is equal to the number of molecular orbitals. According to a symmetry species of the molecule, the representation matrix is decomposed into a direct sum of invariant eigenspaces relevant to individual irreducible representations. If the basis vectors belong to different irreducible representation, such vectors are orthogonal to each other in virtue of (15.171). Even though those vectors belong to the same irreducible representation, they are orthogonal if they are allocated to a different place. However, it is often the case that the vectors belong to the same place of the same irreducible representation. Then, it is always possible according to Theorem 11.2 to make them mutually orthogonal by taking their linear combination. In such a way, we can construct an orthonormal basis set throughout the representation space. In fact, the method is a powerful tool for solving an energy eigenvalue problem and for determining associated eigenfunctions (or molecular orbitals).

## 15.8 Direct-Product Representation

In Sect. 11.5, we have studied basic properties of direct-product groups. Correspondingly, we examine in this section the properties of direct-product representation. This notion is very useful to investigate optical transitions in molecular systems and selection rules relevant to those transitions.

Let $D^{(\alpha)}$ and $D^{(\beta)}$ be two different irreducible representations whose dimensions are $d_\alpha$ and $d_\beta$, respectively. Then, operating a group element $g$ on the basis functions, we have

$$g(\psi_i) = \sum_{k=1}^{d_\alpha} \psi_k D_{ki}^{(\alpha)}(g) \ (1 \le i \le d_\alpha), \tag{15.185}$$

$$g(\phi_j) = \sum_{l=1}^{d_\beta} \phi_l D_{lj}^{(\beta)}(g) \ (1 \le j \le d_\beta), \tag{15.186}$$

where $\psi_k \ (1 \le i \le d_\alpha)$ and $\phi_l \ (1 \le l \le d_\beta)$ are basis functions of $D^{(\alpha)}$ and $D^{(\beta)}$, respectively. We can construct $d_\alpha d_\beta$ new basis vectors using $\psi_k \phi_l$. These functions are transformed according to $g$ such that

$$g(\psi_i\phi_j) = g(\psi_i)g(\phi_j) = \left[\sum_k \psi_k D^{(\alpha)}_{ki}(g)\right]\left[\sum_l \phi_l D^{(\beta)}_{lj}(g)\right] = \sum_k\sum_l \psi_k\phi_l D^{(\alpha)}_{ki}(g)D^{(\beta)}_{lj}(g). \tag{15.187}$$

Here let us define the following matrix $\left[D^{(\alpha\times\beta)}(g)\right]_{kl,ij}$ such that

$$\left[D^{(\alpha\times\beta)}(g)\right]_{kl,ij} \equiv D^{(\alpha)}_{ki}(g)D^{(\beta)}_{lj}(g). \tag{15.188}$$

Then we have

$$g(\psi_i\phi_j) = \sum_k\sum_l \psi_k\phi_l\left[D^{(\alpha\times\beta)}(g)\right]_{kl,ij}. \tag{15.189}$$

The notation using double scripts is somewhat complicated. We notice, however, that in (15.188), the order of subscript of *kilj* is converted to *kl*, *ij*, i.e., the subscripts $i$ and $l$ have been interchanged.

We write $D^{(\alpha)}$ and $D^{(\beta)}$ in explicit forms as follows:

$$D^{(\alpha)}(g) = \begin{pmatrix} d^{(\alpha)}_{1,1} & \cdots & d^{(\alpha)}_{1,d_\alpha} \\ \vdots & \ddots & \vdots \\ d^{(\alpha)}_{d_\alpha,1} & \cdots & d^{(\alpha)}_{d_\alpha,d_\alpha} \end{pmatrix}, \qquad D^{(\beta)}(g) = \begin{pmatrix} d^{(\beta)}_{1,1} & \cdots & d^{(\beta)}_{1,d_\beta} \\ \vdots & \ddots & \vdots \\ d^{(\beta)}_{d_\beta,1} & \cdots & d^{(\beta)}_{d_\beta,d_\beta} \end{pmatrix}. \tag{15.190}$$

Thus,

$$D^{(\alpha\times\beta)}(g) = D^{(\alpha)}(g)\otimes D^{(\beta)}(g) = \begin{pmatrix} d^{(\alpha)}_{1,1}D^{(\beta)}(g) & \cdots & d^{(\alpha)}_{1,d_\alpha}D^{(\beta)}(g) \\ \vdots & \ddots & \vdots \\ d^{(\alpha)}_{d_\alpha,1}D^{(\beta)}(g) & \cdots & d^{(\alpha)}_{d_\alpha,d_\alpha}D^{(\beta)}(g) \end{pmatrix}. \tag{15.191}$$

To get familiar with the double-scripted notation, we describe a case of (2,2) matrices. Denoting

$$D^{(\alpha)}(g) = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad \text{and} \quad D^{(\beta)}(g) = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}, \tag{15.192}$$

we get

$$D^{(\alpha\times\beta)}(g) = D^{(\alpha)}(g)\otimes D^{(\beta)}(g) = \begin{pmatrix} a_{11}b_{11} & a_{11}b_{12} & a_{12}b_{11} & a_{12}b_{12} \\ a_{11}b_{21} & a_{11}b_{22} & a_{12}b_{21} & a_{12}b_{22} \\ a_{21}b_{11} & a_{21}b_{12} & a_{22}b_{11} & a_{22}b_{12} \\ a_{21}b_{21} & a_{21}b_{22} & a_{22}b_{21} & a_{22}b_{22} \end{pmatrix}. \tag{15.193}$$

Corresponding to (15.22), (15.189) describes the transformation of $\psi_i\phi_j$ regarding the double script. At the same time, a set $\{\psi_i\phi_j; (1\le i\le d_\alpha, 1\le j\le d_\beta)\}$ is a basis of $D^{(\alpha\times\beta)}$. In fact,

$$\begin{aligned}\left[D^{(\alpha\times\beta)}(gg')\right]_{kl,ij} &= D^{(\alpha)}_{ki}(gg')D^{(\beta)}_{lj}(gg') \\ &= \left[\sum_\mu D^{(\alpha)}_{k\mu}(g)D^{(\alpha)}_{\mu i}(g')\right]\left[\sum_\nu D^{(\beta)}_{l\nu}(g)D^{(\beta)}_{\nu j}(g')\right] \\ &= \sum_\mu\sum_\nu D^{(\alpha)}_{k\mu}(g)D^{(\beta)}_{l\nu}(g)D^{(\alpha)}_{\mu i}(g')D^{(\beta)}_{\nu j}(g') \\ &= \sum_\mu\sum_\nu \left[D^{(\alpha\times\beta)}(g)\right]_{kl,\mu\nu}\left[D^{(\alpha\times\beta)}(g')\right]_{\mu\nu,ij}.\end{aligned} \tag{15.194}$$

Equation (15.194) shows that the rule of matrix calculation with respect to the double subscript is satisfied. Consequently, we get

$$D^{(\alpha\times\beta)}(gg') = D^{(\alpha\times\beta)}(g)D^{(\alpha\times\beta)}(g'). \tag{15.195}$$

The relation (15.195) indicates that $D^{(\alpha\times\beta)}$ is certainly a representation of a group. This representation is said to be a direct-product representation.

A character of the direct-product representation is given by putting $k = i$ and $l = j$ in (15.188). That is,

$$\left[D^{(\alpha\times\beta)}(g)\right]_{ij,ij} = D^{(\alpha)}_{ii}(g)D^{(\beta)}_{jj}(g). \tag{15.196}$$

Denoting

$$\chi^{(\alpha\times\beta)}(g) \equiv \left[D^{(\alpha\times\beta)}(g)\right]_{ij,ij}, \tag{15.197}$$

we have

$$\chi^{(\alpha\times\beta)}(g) = \chi^{(\alpha)}(g)\chi^{(\beta)}(g). \tag{15.198}$$

Even though $D^{(\alpha)}$ and $D^{(\beta)}$ are both irreducible, $D^{(\alpha x\beta)}$ is not necessarily irreducible. Suppose that

$$D^{(\alpha)}(g) \otimes D^{(\beta)}(g) = \sum_{\omega} q_{\omega} D^{(\omega)}(g), \tag{15.199}$$

where $q_\gamma$ is given by (15.83). Then, we have

$$q_\gamma = \frac{1}{n}\sum_{g} \chi^{(\gamma)}(g)^* \chi^{(\alpha\times\beta)}(g) = \frac{1}{n}\sum_{g} \chi^{(\gamma)}(g)^* \chi^{(\alpha)}(g)\chi^{(\beta)}(g), \tag{15.200}$$

where $n$ is an order of the group. This relation is often used to perform quantum mechanical or chemical calculations, especially to evaluate optical transitions of matter including atoms and molecular systems. This is also useful to examine whether a definite integral of a product of functions (or an inner product of vectors) vanishes.

In Sect. 13.5, we investigated definition and properties of direct-product groups. Similarly to the case of the direct-product representation, we consider a representation of the direct-product groups. Let us consider two groups $\mathcal{G}$ and $\mathcal{H}$ and assume that a direct-product group $\mathcal{G} \otimes \mathcal{H}$ is defined (Sect. 13.5). Also, let $D^{(\alpha)}$ and $D^{(\beta)}$ be $d_\alpha$- and $d_\beta$-dimensional representations of groups $\mathcal{G}$ and $\mathcal{H}$, respectively. Furthermore, let us define a matrix $D^{(\alpha\times\beta)}(ab)$ as in (15.188) such that

$$\left[D^{(\alpha\times\beta)}(ab)\right]_{kl,ij} \equiv D^{(\alpha)}_{ki}(a) D^{(\beta)}_{lj}(b), \tag{15.201}$$

where $a$ and $b$ are arbitrarily chosen from $\mathcal{G}$ and $\mathcal{H}$, respectively, and $ab \in \mathcal{G} \otimes \mathcal{H}$. Then a set comprising $D^{(\alpha\times\beta)}(ab)$ forms a representation of $\mathcal{G} \otimes \mathcal{H}$. (Readers, please verify it.) A dimension of $D^{(\alpha\times\beta)}$ is $d_\alpha d_\beta$. The character is given in (15.69), and so in the present case by putting $i = k$ and $j = l$ in (15.201), we get

$$\begin{aligned}\chi^{(\alpha\times\beta)}(ab) &= \sum_k \sum_l \left[D^{(\alpha\times\beta)}(ab)\right]_{kl,kl} = \sum_k \sum_l D^{(\alpha)}_{kk}(a) D^{(\beta)}_{ll}(b) \\ &= \chi^{(\alpha)}(a)\chi^{(\beta)}(b).\end{aligned} \tag{15.202}$$

Equation (15.202) resembles (15.198). Hence, we should be careful not to confuse them. In (15.198), we were thinking of a direct-product representation within a *sole* group $\mathcal{G}$. In (15.202), however, we are considering a representation of the direct-product group comprising two *different* groups. In fact, even though a character is computed with respect to a sole group element $g$ in (15.198), in (15.202) we evaluate a character regarding two group elements $a$ and $b$ chosen from different groups $\mathcal{G}$ and $\mathcal{H}$, respectively.

## 15.9 Symmetric Representation and Antisymmetric Representation

As mentioned in Sect. 15.2, we have viewed a group element $g$ as a linear transformation over a vector space $V$. There we dealt with widely chosen functions as vectors. In this chapter, we introduce other useful ideas so that we can apply them to molecular science and atomic physics.

In the previous section, we defined direct-product representation. In $D^{(\alpha\times\beta)}(g) = D^{(\alpha)}(g) \otimes D^{(\beta)}(g)$, we can freely consider a case where $D^{(\alpha)}(g) = D^{(\beta)}(g)$. Then we have

$$g(\psi_i\phi_j) = g(\psi_i)g(\phi_j) = \sum_k \sum_l \psi_k \phi_l D^{(\alpha)}_{ki}(g) D^{(\alpha)}_{lj}(g). \tag{15.203}$$

Regarding a product function $\psi_j\phi_i$, we can get a similar equation such that

$$g(\psi_j\phi_i) = g(\psi_j)g(\phi_i) = \sum_k \sum_l \psi_k \phi_l D^{(\alpha)}_{kj}(g) D^{(\alpha)}_{li}(g). \tag{15.204}$$

On the basis of the linearity of the relations (15.203) and (15.204), let us construct a linear combination of the product functions. That is,

$$\begin{aligned} g(\psi_i\phi_j \pm \psi_j\phi_i) &= g(\psi_i\phi_j) \pm g(\psi_j\phi_i) \\ &= \sum_k \sum_l \psi_k \phi_l D^{(\alpha)}_{ki}(g) D^{(\alpha)}_{lj}(g) \pm \sum_k \sum_l \psi_k \phi_l D^{(\alpha)}_{kj}(g) D^{(\alpha)}_{li}(g) \\ &= \sum_k \sum_l (\psi_k \phi_l \pm \psi_l \phi_k) D^{(\alpha)}_{ki}(g) D^{(\alpha)}_{lj}(g). \end{aligned} \tag{15.205}$$

Here, defining $\Psi^{\pm}_{ij}$ as

$$\Psi^{\pm}_{ij} = \psi_i\phi_j \pm \psi_j\phi_i, \tag{15.206}$$

we rewrite (15.205) as

$$g\Psi^{\pm}_{ij} = \sum_k \sum_l \Psi^{\pm}_{kl} \left\{ \frac{1}{2} \left[ D^{(\alpha)}_{ki}(g) D^{(\alpha)}_{lj}(g) \pm D^{(\alpha)}_{li}(g) D^{(\alpha)}_{kj}(g) \right] \right\}. \tag{15.207}$$

Notice that $\Psi^{+}_{ij}$ and $\Psi^{-}_{ij}$ in (15.206) are symmetric and antisymmetric with respect to the interchange of subscript $i$ and $j$, respectively. That is,

$$\Psi_{ij}^{\pm} = \pm\Psi_{ji}^{\pm}. \tag{15.208}$$

We may naturally ask how we can constitute representation (matrices) with (15.207). To answer this question, we have to carry out calculations by replacing $g$ with $gg'$ in (15.207) and following the procedures of Sect. 15.2. Thus, we have

$$\begin{aligned}
gg'\Psi_{ij}^{\pm} &= \sum_k \sum_l \Psi_{kl}^{\pm} \left\{ \frac{1}{2} \left[ D_{ki}^{(\alpha)}(gg') D_{lj}^{(\alpha)}(gg') \pm D_{li}^{(\alpha)}(gg') D_{kj}^{(\alpha)}(gg') \right] \right\} \\
&= \sum_k \sum_l \Psi_{kl}^{\pm} \left\{ \frac{1}{2} \left[ \sum_\mu D_{k\mu}^{(\alpha)}(g) D_{\mu i}^{(\alpha)}(g') \sum_\nu D_{l\nu}^{(\alpha)}(g) D_{\nu j}^{(\alpha)}(g') \right.\right. \\
&\quad \left.\left. \pm \sum_\mu D_{l\mu}^{(\alpha)}(g) D_{\mu i}^{(\alpha)}(g') \sum_\nu D_{k\nu}^{(\alpha)}(g) D_{\nu j}^{(\alpha)}(g') \right] \right\} \\
&= \sum_k \sum_l \Psi_{kl}^{\pm} \left\{ \frac{1}{2} \sum_\mu \sum_\nu \left[ D_{k\mu}^{(\alpha)}(g) D_{l\nu}^{(\alpha)}(g) \pm D_{l\mu}^{(\alpha)}(g) D_{k\nu}^{(\alpha)}(g) \right] D_{\mu i}^{(\alpha)}(g') D_{\nu j}^{(\alpha)}(g') \right\} \\
&= \sum_k \sum_l \Psi_{kl}^{\pm} \left\{ \frac{1}{2} \sum_\mu \sum_\nu \left\{ \left[ D^{(\alpha\times\alpha)}(g) \right]_{kl,\mu\nu} \pm \left[ D^{(\alpha\times\alpha)}(g) \right]_{lk,\mu\nu} \right\} D_{\mu i}^{(\alpha)}(g') D_{\nu j}^{(\alpha)}(g'),
\end{aligned} \tag{15.209}$$

where the last equality follows from the definition of a direct-product representation (15.188). We notice that the terms of $\left[D^{(\alpha\times\alpha)}(gg')\right]_{kl,\mu\nu} \pm \left[D^{(\alpha\times\alpha)}(gg')\right]_{lk,\mu\nu}$ in (15.209) are symmetric and antisymmetric with respect to the interchange of subscripts $k$ and $l$ together with subscripts $\mu$ and $\nu$, respectively. Comparing both sides of (15.209), we see that this must be the case with $i$ and $j$ as well. Then, the last factor of (15.209) should be rewritten as:

$$D_{\mu i}^{(\alpha)}(g') D_{\nu j}^{(\alpha)}(g') = \frac{1}{2} \left[ D_{\mu i}^{(\alpha)}(g') D_{\nu j}^{(\alpha)}(g') \pm D_{\nu i}^{(\alpha)}(g') D_{\mu j}^{(\alpha)}(g') \right].$$

Now, we define the following notations accordingly:

$$\begin{aligned}
\left\{ D^{[\alpha\times\alpha]}(g) \right\}_{kl,\mu\nu} &\equiv \frac{1}{2} \left\{ \left[ D^{(\alpha\times\alpha)}(g) \right]_{kl,\mu\nu} + \left[ D^{(\alpha\times\alpha)}(g) \right]_{lk,\mu\nu} \right\} \\
&= \frac{1}{2} \left[ D_{k\mu}^{(\alpha)}(g) D_{l\nu}^{(\alpha)}(g) + D_{l\mu}^{(\alpha)}(g) D_{k\nu}^{(\alpha)}(g) \right].
\end{aligned} \tag{15.210}$$

$$\begin{aligned}
\left\{ D^{\{\alpha\times\alpha\}}(g) \right\}_{kl,\mu\nu} &\equiv \frac{1}{2} \left\{ \left[ D^{(\alpha\times\alpha)}(g) \right]_{kl,\mu\nu} - \left[ D^{(\alpha\times\alpha)}(g) \right]_{lk,\mu\nu} \right\} \\
&= \frac{1}{2} \left[ D_{k\mu}^{(\alpha)}(g) D_{l\nu}^{(\alpha)}(g) - D_{l\mu}^{(\alpha)}(g) D_{k\nu}^{(\alpha)}(g) \right].
\end{aligned} \tag{15.211}$$

Meanwhile, using $D^{(\alpha)}_{\mu i}(g')D^{(\alpha)}_{\nu j}(g') = \left[D^{(\alpha\times\alpha)}(g')\right]_{\mu\nu,ij}$ and considering the exchange of summation with respect to the subscripts $\mu$ and $\nu$, we can also define $D^{[\alpha\times\alpha]}(g')$ and $D^{\{\alpha\times\alpha\}}(g)$ according to the symmetric and antisymmetric cases, respectively. Thus for the symmetric case, we have

$$\begin{aligned} gg'\Psi^{+}_{ij} &= \sum_k \sum_l \sum_\mu \sum_\nu \Psi^{+}_{kl} \left\{D^{[\alpha\times\alpha]}(g)\right\}_{kl,\mu\nu} \left\{D^{[\alpha\times\alpha]}(g')\right\}_{\mu\nu,ij} \\ &= \sum_k \sum_l \Psi^{+}_{kl} \left\{D^{[\alpha\times\alpha]}(g)D^{[\alpha\times\alpha]}(g')\right\}_{kl,ij}. \end{aligned} \tag{15.212}$$

Using (15.210), (15.207) can be rewritten as

$$g\Psi^{+}_{ij} = \sum_k \sum_l \Psi^{+}_{kl} \left\{D^{[\alpha\times\alpha]}(g)\right\}_{kl,ij}. \tag{15.213}$$

Then we have

$$gg'\Psi^{+}_{ij} = \sum_k \sum_l \Psi^{+}_{kl} \left\{D^{[\alpha\times\alpha]}(gg')\right\}_{kl,ij}. \tag{15.214}$$

Comparing (15.212) and (15.214), we finally get

$$D^{[\alpha\times\alpha]}(gg') = D^{[\alpha\times\alpha]}(g)D^{[\alpha\times\alpha]}(g'). \tag{15.215}$$

Similarly, for the antisymmetric case, we have

$$gg'\Psi^{-}_{ij} = \sum_k \sum_l \Psi^{-}_{kl} \left\{D^{\{\alpha\times\alpha\}}(g)D^{\{\alpha\times\alpha\}}(g')\right\}_{kl,ij}, \tag{15.216}$$

$$g\Psi^{-}_{ij} = \sum_k \sum_l \Psi^{-}_{kl} \left\{D^{\{\alpha\times\alpha\}}(g)\right\}_{kl,ij}, \tag{15.217}$$

$$D^{\{\alpha\times\alpha\}}(gg') = D^{\{\alpha\times\alpha\}}(g)D^{\{\alpha\times\alpha\}}(g'). \tag{15.218}$$

Thus, both $D^{[\alpha\times\alpha]}(g)$ and $D^{\{\alpha\times\alpha\}}(g)$ produce well-defined representations.

Letting dimension of the representation $\alpha$ be $d_\alpha$, we have $d_\alpha(d_\alpha+1)/2$ functions belonging to the symmetric representation and $d_\alpha(d_\alpha-1)/2$ functions belonging to the antisymmetric representation. With the two-dimensional representation, for instance, functions belonging to the symmetric representation are

$$\psi_1\phi_1, \psi_1\phi_2 + \psi_2\phi_1, \text{ and } \psi_2\phi_2.$$

A function belonging to the antisymmetric representation is

$$\psi_1\phi_2 - \psi_2\phi_1.$$

Note that these vectors have not yet been normalized.

From (15.210) to (15.211), we can readily get useful expressions with characters of symmetric and antisymmetric representations. In (15.210), putting $\mu = k$ and $\nu = l$ and summing over $k$ and $l$,

$$\begin{aligned}\chi^{[\alpha\times\alpha]}(g) &= \sum_k \sum_l \left\{D^{[\alpha\times\alpha]}(g)\right\}_{kl,kl} \\ &= \frac{1}{2}\sum_k \sum_l \left[D^{(\alpha)}_{kk}(g)D^{(\alpha)}_{ll}(g) + D^{(\alpha)}_{lk}(g)D^{(\alpha)}_{kl}(g)\right] \\ &= \frac{1}{2}\left\{\left[\chi^{(\alpha)}(g)\right]^2 + \left[\chi^{(\alpha)}\left(g^2\right)\right]\right\}.\end{aligned} \tag{15.219}$$

Similarly, we have

$$\chi^{\{\alpha\times\alpha\}}(g) = \frac{1}{2}\left\{\left[\chi^{(\alpha)}(g)\right]^2 - \left[\chi^{(\alpha)}\left(g^2\right)\right]\right\}. \tag{15.220}$$

## References

1. Inui T, Tanabe Y, Onodera Y (1980) Group theory and its applications in physics, expanded edn. Shokabo, Tokyo (in Japanese)
2. Hassani S (2006) Mathematical physics. Springer, New York
3. Chen JQ, Ping J, Wang F (2002) Group representation theory for physicists, 2nd edn. World Scientific, Singapore
4. Hamermesh M (1989) Group theory and its application to physical problems. Dover, New York

# Chapter 16
# Applications of Group Theory to Physical Chemistry

On the basis of studies of group theory, now in this last chapter we apply the knowledge to the molecular orbital (MO) calculations (or quantum chemical calculations). As tangible examples, we adopt aromatic hydrocarbons (ethylene, cyclopropenyl radical, benzene, and ally radical) and methane. The approach is based upon a method of linear combination of atomic orbitals (LCAO). To seek an appropriate LCAO MO, we make the most of a method based on a symmetry-adapted linear combination (SALC). To use projection operators is a powerful tool for this purpose. For the sake of correct understanding, it is desired to consider transformation of functions. To this end, we first show several examples.

In the process of carrying out MO calculations, we encounter a secular equation as an eigenvalue equation. Using a SALC eases the calculations of the secular equation. Molecular science relies largely on spectroscopic measurements, and researchers need to assign individual spectral lines to a specific transition between the relevant molecular states. Representation theory works well in this situation. Thus, the group theory finds a perfect fit with its applications in the molecular science.

## 16.1 Transformation of Functions

Before showing individual examples, let us consider the transformation of functions (or vectors) by the symmetry operation. Here, we consider scalar functions.

For example, let us suppose an arbitrary function $f(x, y)$ on a Cartesian $xy$-coordinate. Figure 16.1 shows a contour map of $f(x, y) =$ constant. Then, suppose that the map is rotated around the origin. More specifically, the position vector $\boldsymbol{r}_0$ fixed on a "summit" [i.e., a point that gives a maximal value of $f(x, y)$] undergoes a symmetry operation, namely rotation around the origin. As a result, $\boldsymbol{r}_0$ is transformed to $\boldsymbol{r}_0'$. Here, we assume that a "mountain" represented by $f(x, y)$ is a rigid



S. Hotta, *Mathematical Physical Chemistry*,
https://doi.org/10.1007/978-981-10-7671-8_16

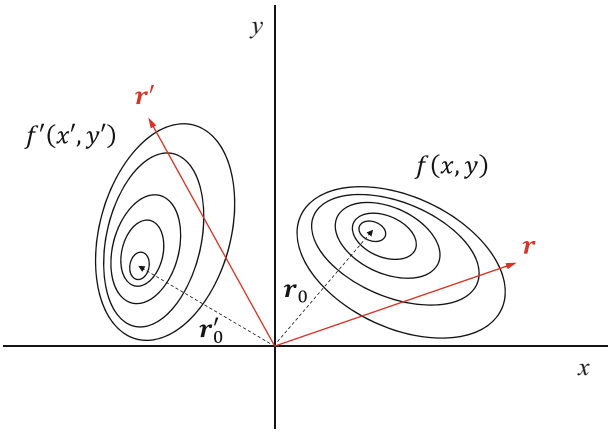


**Fig. 16.1** Contour map of $f(x, y)$ and $f'(x', y')$. The function $f'(x', y')$ is obtained by rotating a map of $f(x, y)$ around the $z$-axis

body before and after the transformation. Concomitantly, a general point $\boldsymbol{r}$ (see Sect. 14.1) is transformed to $\boldsymbol{r}'$ in exactly the same way as $\boldsymbol{r}_0$

A new function $f'$ gives a new contour map after the rotation. Let us describe $f'$ as

$$f' \equiv O_R f \tag{16.1}$$

The RHS of (16.1) means that $f'$ is produced as a result of operating the rotation $R$ on $f$. We describe the position vector $\boldsymbol{r}'_0$. Following the notation of Sect. 9.1, $\boldsymbol{r}'_0$ and $\boldsymbol{r}'$ are expressed as

$$\boldsymbol{r}'_0 = R(\boldsymbol{r}_0) \text{ and } \boldsymbol{r}' = R(\boldsymbol{r}) \tag{16.2}$$

The matrix representation for $R$ is given by, e.g., (9.35). In (16.2), we have

$$\boldsymbol{r}_0 = (\boldsymbol{e}_1\ \boldsymbol{e}_2)\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} \text{ and } \boldsymbol{r}'_0 = (\boldsymbol{e}_1\ \boldsymbol{e}_2)\begin{pmatrix} x'_0 \\ y'_0 \end{pmatrix} \tag{16.3}$$

where $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ are orthonormal basis vectors in the $xy$-plane. Also we have

$$\boldsymbol{r} = (\boldsymbol{e}_1\ \boldsymbol{e}_2)\begin{pmatrix} x \\ y \end{pmatrix} \text{ and } \boldsymbol{r}' = (\boldsymbol{e}_1\ \boldsymbol{e}_2)\begin{pmatrix} x' \\ y' \end{pmatrix} \tag{16.4}$$

Meanwhile, the following equation must hold:

$$f'(x', y') = f(x, y) \tag{16.5}$$

Or, using (16.1), we have

$$O_R f(x', y') = f(x, y) \tag{16.6}$$

The above argument can be extended to a three-dimensional (or higher dimensional) space. In that case, similarly we have

$$O_R f(x', y', z') = f(x, y, z),\ O_R f(\boldsymbol{r}') = f(\boldsymbol{r}),\ \text{or}\ f'(\boldsymbol{r}') = f(\boldsymbol{r}) \tag{16.7}$$

where

$$\boldsymbol{r} = (\boldsymbol{e}_1 \boldsymbol{e}_2 \boldsymbol{e}_3) \begin{pmatrix} x \\ y \\ z \end{pmatrix} \text{and}\, \boldsymbol{r}' = (\boldsymbol{e}_1 \boldsymbol{e}_2 \boldsymbol{e}_3) \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} \tag{16.8}$$

The last relation of (16.7) comes from (16.1). Using (16.2) and (16.3), we rewrite (16.7) as

$$O_R f[R(\boldsymbol{r})] = f(\boldsymbol{r}) \tag{16.9}$$

Replacing $\boldsymbol{r}$ with $R^{-1}(\boldsymbol{r})$, we get

$$O_R f\left\{R\left[R^{-1}(\boldsymbol{r})\right]\right\} = f\left[R^{-1}(\boldsymbol{r})\right] \tag{16.10}$$

That is,

$$O_R f(\boldsymbol{r}) = f[R^{(-1)}(\boldsymbol{r})] \tag{16.11}$$

More succinctly, (16.11) may be rewritten as

$$Rf(\boldsymbol{r}) = f\left[R^{-1}(\boldsymbol{r})\right] \tag{16.12}$$

Comparing (16.11) with (16.12) and considering (16.1), we have

$$f' \equiv O_R f \equiv Rf \tag{16.13}$$

To gain a good understanding of the function transformation, let us think of some examples.

*Example 16.1* Let $f(x, y)$ be a function described by

$$f(x, y) = (x - a)^2 + (y - b)^2; a, b > 0 \tag{16.14}$$

A contour is shown in Fig. 16.2. We consider a $\pi/2$ rotation around the $z$-axis. Then, $f(x, y)$ is transformed into $Rf(x, y) = f'(x, y)$ such that

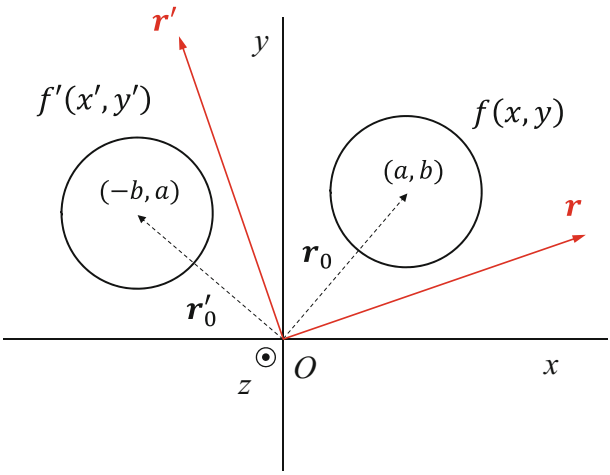


**Fig. 16.2** Contour map of $f(x,y) = (x-a)^2 + (y-b)^2$ and $f'(x',y') = (x'+b)^2 + (y'-a)^2$ before and after a $\pi/2$ rotation around the $z$-axis

$$Rf(x,y) = f'(x,y) = (x+b)^2 + (y-a)^2 \tag{16.15}$$

We also have

$$\boldsymbol{r}_0 = (\boldsymbol{e}_1\boldsymbol{e}_2)\begin{pmatrix} a \\ b \end{pmatrix},$$

$$\boldsymbol{r}_0' = (\boldsymbol{e}_1\boldsymbol{e}_2)\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = (\boldsymbol{e}_1\boldsymbol{e}_2)\begin{pmatrix} -b \\ a \end{pmatrix},$$

where we define $R$ as

$$R = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

Similarly,

$$\boldsymbol{r} = (\boldsymbol{e}_1\boldsymbol{e}_2)\begin{pmatrix} x \\ y \end{pmatrix} \text{ and } \boldsymbol{r}' = (\boldsymbol{e}_1\boldsymbol{e}_2)\begin{pmatrix} -y \\ x \end{pmatrix}$$

From (16.15), we have

$$f'(x',y') = (x'+b)^2 + (y'-a)^2 = (-y+b)^2 + (x-a)^2 = f(x,y) \tag{16.16}$$

This ensures that (16.5) holds. The implication of (16.16) combined with (16.5) is that a view of $f'(x', y')$ from $(-b, a)$ is the same as that of $f(x, y)$ from $(a, b)$. Imagine that if we are standing at $(-b, a)$ of $f'(x', y')$, we are in the bottom of the "valley" of $f'(x', y')$. Exactly in the same manner, if we are standing at $(a, b)$ of $f(x, y)$, we are in the bottom of the valley of $f(x, y)$ as well. Notice that for both $f(x, y)$ and $f'(x', y')$, $(a, b)$ and $(-b, a)$ are the lowest point, respectively.

Meanwhile, we have

$$R^{-1} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \; R^{-1}(\boldsymbol{r}) = (\boldsymbol{e}_1 \boldsymbol{e}_2)\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = (\boldsymbol{e}_1 \boldsymbol{e}_2)\begin{pmatrix} y \\ -x \end{pmatrix}$$

Then, we have

$$f\left[R^{-1}(\boldsymbol{r})\right] = \left[(y) - a\right]^2 + \left[(-x) - b\right]^2 = (x+b)^2 + (y-a)^2 = Rf(\boldsymbol{r}) \tag{16.17}$$

Thus, (16.12) certainly holds.

*Example 16.2* Let $f(\boldsymbol{r})$ and $g(\boldsymbol{r})$ be functions described by

$$f(\boldsymbol{r}) = \mathrm{e}^{-2\left[(x-a)^2 + y^2 + z^2\right]} + \mathrm{e}^{-2\left[(x+a)^2 + y^2 + z^2\right]} (a > 0), \tag{16.18}$$

$$g(\boldsymbol{r}) = \mathrm{e}^{-2\left[(x-a)^2 + y^2 + z^2\right]} - \mathrm{e}^{-2\left[(x+a)^2 + y^2 + z^2\right]} (a > 0). \tag{16.19}$$

Figure 16.3 shows an outline of the contour of $f(\boldsymbol{r})$. We consider a following symmetry operation in a three-dimensional coordinate system:

$$R = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ and } R^{-1} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{16.20}$$

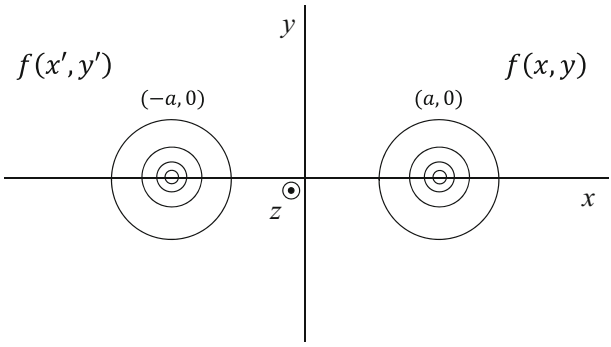


**Fig. 16.3** Contour map of $f(\boldsymbol{r}) = e^{-2\left[(x-a)^2 + y^2 + z^2\right]} + e^{-2\left[(x+a)^2 + y^2 + z^2\right]} (a > 0)$. The function form remains unchanged by a reflection with respect to the $yz$-plane

This represents a reflection with respect to the $yz$-plane. Then, we have

$$f\left[R^{-1}(\boldsymbol{r})\right] = Rf(\boldsymbol{r}) = f(\boldsymbol{r}), \tag{16.21}$$

$$g\left[R^{-1}(\boldsymbol{r})\right] = Rg(\boldsymbol{r}) = -g(\boldsymbol{r}), \tag{16.22}$$

Plotting $f(\boldsymbol{r})$ and $g(\boldsymbol{r})$ as a function of $x$ on the $x$-axis, we depict results in Fig. 16.4.

Looking at (16.21) and (16.22), we find that $f(\boldsymbol{r})$ and $g(\boldsymbol{r})$ are solutions of an eigenvalue equation for an operator $R$. Corresponding eigenvalues are 1 and $-1$ for $f(\boldsymbol{r})$ and $g(\boldsymbol{r})$, respectively. In particular, $f(\boldsymbol{r})$ holds the functional form after the transformation $R$. In this case, $f(\boldsymbol{r})$ is said to be invariant with the transformation $R$. Moreover, $f(\boldsymbol{r})$ is invariant with the following eight transformations:

$$R = \begin{pmatrix} \pm 1 & 0 & 0 \\ 0 & \pm 1 & 0 \\ 0 & 0 & \pm 1 \end{pmatrix} \tag{16.23}$$

These transformations form a group that is isomorphic to $D_{2h}$. Therefore, $f(\boldsymbol{r})$ is eligible for a basis function of the totally symmetric representation $A_g$ of $D_{2h}$. Notice that $f(\boldsymbol{r})$ is invariant as well with a rotation of an arbitrary angle around the $x$-axis. On the other hand, $g(\boldsymbol{r})$ belongs to $B_{3u}$.

## 16.2 Method of Molecular Orbitals (MOs)

Bearing in mind these arguments, we examine several examples of quantum chemical calculations. Our approach is based upon the molecular orbital theory. The theory assumes the existence of molecular orbitals (MOs) in a molecule, as the notion of atomic orbitals has been well-established in atomic physics. Furthermore, we assume that the molecular orbitals comprise a linear combination of atomic orbitals (LCAO).

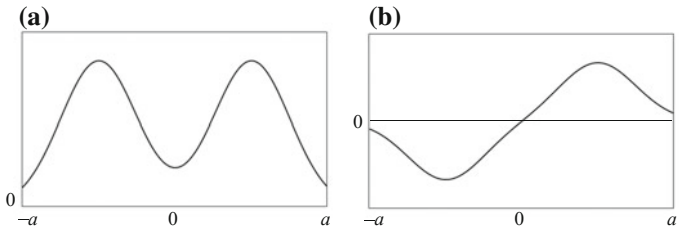


**Fig. 16.4** Plots of $f(\boldsymbol{r}) = e^{-2\left[(x-a)^2+y^2+z^2\right]} + e^{-2\left[(x+a)^2+y^2+z^2\right]}$ and $g(\boldsymbol{r}) = e^{-2\left[(x-a)^2+y^2+z^2\right]} - e^{-2\left[(x+a)^2+y^2+z^2\right]} (a > 0)$ as a function of $x$ on the $x$-axis. **a** $f(\boldsymbol{r})$. **b** $g(\boldsymbol{r})$

This notion is equivalent to that individual electrons in a molecule are independently moving in a potential field produced by nuclei and other electrons. In other words, we assume that each electron is moving along an MO that is extended over the whole molecule. Electronic state in the molecule is formed by different MOs of various energies that are occupied by electrons. As in the case of an atom, an MO $\psi_i(\boldsymbol{r})$ occupied by an electron is described as

$$H\psi_i(\boldsymbol{r}) \equiv \left[-\frac{\hbar^2}{2m}\nabla^2 + V(\boldsymbol{r})\right]\psi_i(\boldsymbol{r}) = E_i\psi_i(\boldsymbol{r}), \tag{16.24}$$

where $H$ is Hamiltonian of a molecule; $m$ is a mass of an electron (note that we do not use a reduced mass $\mu$ here); $\boldsymbol{r}$ is a position vector of the electron; $\nabla^2$ is the Laplacian (Laplace operator); $V(\boldsymbol{r})$ is a potential of the molecule at $\boldsymbol{r}$; $E_i$ is an energy of the electron occupying $\psi_i$ and said to be a molecular orbital energy. We assume that $V(\boldsymbol{r})$ possesses a symmetry the same as that of the molecule.

Let $\mathscr{g}$ be a symmetry group, and let a group element arbitrarily chosen from among $\mathscr{g}$ be $R$. Suppose that an arbitrary position vector $\boldsymbol{r}$ is moved to another position $\boldsymbol{r}'$. This transformation is expressed as (16.2). Since $V(\boldsymbol{r})$ has the same symmetry as the molecule, an electron "feels" the same potential field at $\boldsymbol{r}'$ as that at $\boldsymbol{r}$. That is,

$$V(\boldsymbol{r}) = V'(\boldsymbol{r}') = V(\boldsymbol{r}') \tag{16.25}$$

Or we have

$$V(\boldsymbol{r})\psi(\boldsymbol{r}) = V(\boldsymbol{r}')\psi'(\boldsymbol{r}') \tag{16.26}$$

where $\psi$ is an arbitrary function. Defining

$$V(\boldsymbol{r})\psi(\boldsymbol{r}) \equiv [V\psi](\boldsymbol{r}) = V\psi(\boldsymbol{r}), \tag{16.27}$$

and recalling (16.1) and (16.7), we get

$$\begin{aligned}[RV]\psi(\boldsymbol{r}') &= R[V\psi(\boldsymbol{r}')] = V'\psi'(\boldsymbol{r}') = V'(\boldsymbol{r}')\psi'(\boldsymbol{r}') = V(\boldsymbol{r}')\psi'(\boldsymbol{r}') \\ &= V(\boldsymbol{r}')R\psi(\boldsymbol{r}') = VR\psi(\boldsymbol{r}') = [VR]\psi(\boldsymbol{r}')\end{aligned} \tag{16.28}$$

Comparing the first and last sides and remembering that $\psi$ is an arbitrary function, we get

$$RV = VR \tag{16.29}$$

Next, let us examine the symmetry of the Laplacian $\nabla^2$. The Laplacian is defined in Sect. 1.2 as

$$\nabla^2 \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}, \tag{1.24}$$

where $x$, $y$, and $z$ denote the Cartesian coordinates. Let $S$ be an orthogonal matrix that transforms the $xyz$-coordinate system to $x'y'z'$-coordinate system. We suppose that $S$ is expressed as

$$S = \begin{pmatrix} s_{11} & s_{12} & s_{13} \\ s_{21} & s_{22} & s_{23} \\ s_{31} & s_{32} & s_{33} \end{pmatrix} \text{ and } \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} s_{11} & s_{12} & s_{13} \\ s_{21} & s_{22} & s_{23} \\ s_{31} & s_{32} & s_{33} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \tag{16.30}$$

Since $S$ is an orthonormal matrix, we have

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} s_{11} & s_{21} & s_{31} \\ s_{12} & s_{22} & s_{32} \\ s_{13} & s_{23} & s_{33} \end{pmatrix} \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} \tag{16.31}$$

The equation is due to

$$SS^T = S^T S = E, \tag{16.32}$$

where $E$ is a unit matrix. Then, we have

$$\frac{\partial}{\partial x'} = \frac{\partial x}{\partial x'}\frac{\partial}{\partial x} + \frac{\partial y}{\partial x'}\frac{\partial}{\partial y} + \frac{\partial z}{\partial x'}\frac{\partial}{\partial z} = s_{11}\frac{\partial}{\partial x} + s_{12}\frac{\partial}{\partial y} + s_{13}\frac{\partial}{\partial z} \tag{16.33}$$

Partially differentiating (16.33), we have

$$\begin{aligned} \frac{\partial}{\partial x'^2} &= s_{11}\frac{\partial}{\partial x'}\frac{\partial}{\partial x} + s_{12}\frac{\partial}{\partial x'}\frac{\partial}{\partial y} + s_{13}\frac{\partial}{\partial x'}\frac{\partial}{\partial z} \\ &= s_{11}\left(s_{11}\frac{\partial}{\partial x} + s_{12}\frac{\partial}{\partial y} + s_{13}\frac{\partial}{\partial z}\right)\frac{\partial}{\partial x} + s_{12}\left(s_{11}\frac{\partial}{\partial x} + s_{12}\frac{\partial}{\partial y} + s_{13}\frac{\partial}{\partial z}\right)\frac{\partial}{\partial y} \\ &\quad + s_{13}\left(s_{11}\frac{\partial}{\partial x} + s_{12}\frac{\partial}{\partial y} + s_{13}\frac{\partial}{\partial z}\right)\frac{\partial}{\partial z} \\ &= s_{11}^2\frac{\partial^2}{\partial x^2} + s_{11}s_{12}\frac{\partial^2}{\partial y \partial x} + s_{11}s_{13}\frac{\partial^2}{\partial z \partial x} + s_{12}s_{11}\frac{\partial^2}{\partial x \partial y} + s_{12}^2\frac{\partial^2}{\partial y^2} + s_{12}s_{13}\frac{\partial^2}{\partial z \partial y} \\ &\quad + s_{13}s_{11}\frac{\partial^2}{\partial x \partial z} + s_{13}s_{12}\frac{\partial^2}{\partial y \partial z} + s_{13}^2\frac{\partial^2}{\partial z^2} \end{aligned} \tag{16.34}$$

Calculating terms of $\frac{\partial}{\partial y'^2}$ and $\frac{\partial}{\partial z'^2}$, we have similar results. Then, summarizing all those 27 terms, we get

$$\left(s_{11}^2 + s_{21}^2 + s_{31}^2\right)\frac{\partial^2}{\partial x^2} = \frac{\partial^2}{\partial x^2}, \tag{16.35}$$

where we used an orthogonal relationship of $S$. Cross terms of $\frac{\partial^2}{\partial x \partial y}, \frac{\partial^2}{\partial y \partial x}$, etc., all vanish. Consequently, we get

$$\frac{\partial}{\partial x'^2} + \frac{\partial}{\partial y'^2} + \frac{\partial}{\partial z'^2} = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \tag{16.36}$$

Defining $\nabla'^2$ as

$$\nabla'^2 \equiv \frac{\partial}{\partial x'^2} + \frac{\partial}{\partial y'^2} + \frac{\partial}{\partial z'^2}, \tag{16.37}$$

we have

$$\nabla'^2 = \nabla^2 \tag{16.38}$$

Notice that (16.38) holds with not only the symmetry operation of the molecule, but also any rotation operation in $\mathbb{R}^3$ (see Sect. 14.4).

As in the case of (16.28), we have

$$\begin{aligned}\left[R\nabla^2\right]\psi(\boldsymbol{r}') &= R\left[\nabla^2\psi(\boldsymbol{r}')\right] = \nabla'^2\psi'(\boldsymbol{r}') = \nabla^2\psi'(\boldsymbol{r}') \\ &= \nabla^2 R\psi(\boldsymbol{r}') = \left[\nabla^2 R\right]\psi(\boldsymbol{r}')\end{aligned} \tag{16.39}$$

Consequently, we get

$$R\nabla^2 = \nabla^2 R \tag{16.40}$$

Adding both sides of (16.29) and (16.40), we get

$$R\left(\nabla^2 + V\right) = \left(\nabla^2 + V\right)R \tag{16.41}$$

From (16.24), we have

$$RH = HR \tag{16.42}$$

Thus, we confirm that the Hamiltonian $H$ commutes with any symmetry operation $R$. In other words, $H$ is invariant with the coordinate transformation relevant to the symmetry operation.

Now, we consider matrix representation of (16.42). Also, let $D$ be an irreducible representation of the symmetry group $\mathscr{g}$. Then, we have

$$D(g)H = HD(g)\big(g \in \mathscr{g}\big)$$

Let $\{\psi_1, \psi_2, \ldots, \psi_d\}$ be a set of basis vectors that span a representation space $\mathcal{L}_S$ associated with $D$. Then, on the basis of Schur's Second Lemma of Sect. 15.3, (16.42) immediately leads to an important conclusion that if $H$ is represented by a matrix, we must have

$$H = \lambda E, \tag{16.43}$$

where $\lambda$ is an arbitrary complex constant. That is, $H$ is represented such that

$$H = \begin{pmatrix} \lambda & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda \end{pmatrix} \tag{16.44}$$

where the above matrix is $(d, d)$ square diagonal matrix and. This implies that $H$ is reduced to

$$H = \lambda D^{(0)} \oplus \cdots \oplus \lambda D^{(0)},$$

where $D^{(0)}$ denotes the totally symmetric representation; notice that it is given by 1 for any symmetry operation. Thus, the commutability of an operator with all symmetry operation is equivalent to that the operator belonging to the totally symmetric representation.

Since $H$ is Hermitian, $\lambda$ should be real. Operating both sides of (16.42) on $\{\psi_1, \psi_2, \ldots, \psi_d\}$ from the right, we have

$$\begin{aligned}(\psi_1\psi_2\ldots\psi_d)RH &= (\psi_1\psi_2\ldots\psi_d)HR \\ &= (\psi_1\psi_2\ldots\psi_d)\begin{pmatrix} \lambda & & \cdots & \\ & \lambda & & \\ \vdots & & \ddots & \vdots \\ & & \cdots & \lambda \end{pmatrix} R = (\lambda\psi_1\lambda\psi_2\ldots\lambda\psi_d)R \\ &= \lambda(\psi_1\psi_2\ldots\psi_d)R.\end{aligned} \tag{16.45}$$

In particular, putting $R = E$, we get

$$(\psi_1\psi_2\ldots\psi_d)H = \lambda(\psi_1\psi_2\ldots\psi_d)$$

Simplifying this equation, we have

$$\psi_i H = \lambda \psi_i (1 \le i \le d) \tag{16.46}$$

Thus, $\lambda$ is found to be an energy eigenvalue of $H$, and $\psi_i (1 \le i \le d)$ are eigenfunctions belonging to the eigenvalue $\lambda$.

Meanwhile, using

$$R(\psi_i) = \sum_{k=1}^{d} \psi_k D_{ki}(R),$$

we rewrite (16.45) as

$$\begin{aligned}(\psi_1 \psi_2 \ldots \psi_d) RH &= (\psi_1 R \psi_2 R \ldots \psi_d R) H \\ &= \left( \sum_{k=1}^{d} \psi_k D_{k1}(R) \sum_{k=1}^{d} \psi_k D_{k2}(R) \ldots \sum_{k=1}^{d} \psi_k D_{kd}(R) \right) H \\ &= \lambda \left( \sum_{k=1}^{d} \psi_k D_{k1}(R) \sum_{k=1}^{d} \psi_k D_{k2}(R) \ldots \sum_{k=1}^{d} \psi_k D_{kd}(R) \right),\end{aligned}$$

where with the second equality we used the relation (9.42), i.e.,

$$\psi_i R = R(\psi_i)$$

Thus, we get

$$\sum_{k=1}^{d} \psi_k D_{ki}(R) H = \lambda \sum_{k=1}^{d} \psi_k D_{ki}(R) \ (1 \le i \le d). \tag{16.47}$$

The relations (16.45)–(16.47) imply that $\psi_1, \psi_2, \ldots,$ and $\psi_d$ as well as their linear combinations using a representation matrix $D(R)$ are eigenfunctions that belong to the same eigenvalue $\lambda$. That is, $\psi_1, \psi_2, \ldots,$ and $\psi_d$ are said to be degenerate with a multiplication $d$.

After the remarks of Theorem 12.5, we can construct an orthonormal basis set $\left\{ \widetilde{\psi}_1, \widetilde{\psi}_2, \ldots, \widetilde{\psi}_d \right\}$ as eigenfunctions. These vectors can be constructed via linear combinations of $\psi_1, \psi_2, \ldots,$ and $\psi_d$. After transforming the vectors $\widetilde{\psi}_i (1 \le i \le d)$ by $R$, we get

$$\begin{aligned}\left\langle \sum_{k=1}^{d} \widetilde{\psi}_k D_{ki}(R) \middle| \sum_{l=1}^{d} \widetilde{\psi}_l D_{lj}(R) \right\rangle &= \sum_{k=1}^{d} \sum_{l=1}^{d} D_{ki}(R)^* D_{lj}(R) \left\langle \widetilde{\psi}_k \middle| \widetilde{\psi}_l \right\rangle \\ &= \sum_{k=1}^{d} D_{ki}(R)^* D_{kj}(R) = \sum_{k=1}^{d} \left[ D^\dagger(R) \right]_{ik} [D(R)]_{kj} \\ &= \left[ D^\dagger(R) D(R) \right]_{ij} = \delta_{ij}.\end{aligned}$$

With the last equality, we used unitarity of $D(R)$. Thus, the orthonormal basis set is retained after unitary transformation of $\left\{\widetilde{\psi}_1, \widetilde{\psi}_2, \ldots, \widetilde{\psi}_d\right\}$.

In the above discussion, we have assumed that $\psi_1, \psi_2, \ldots,$ and $\psi_d$ belong to a certain irreducible representation $D^{(\nu)}$. Since $\widetilde{\psi}_1, \widetilde{\psi}_2, \ldots,$ and $\widetilde{\psi}_d$ consist of their linear combinations, $\widetilde{\psi}_1, \widetilde{\psi}_2, \ldots,$ and $\widetilde{\psi}_d$ belong to $D^{(\nu)}$ as well. Also, we have assumed that $\widetilde{\psi}_1, \widetilde{\psi}_2, \ldots,$ and $\widetilde{\psi}_d$ form an orthonormal basis, and hence, according to Theorem 11.3, these vectors constitute basis vectors that belong to $D^{(\nu)}$. In particular, functions derived using projection operators share these characteristics. In fact, the above principles underlie molecular orbital calculations dealt with in Sects. 16.4 and 16.5. We will go into more detail in subsequent sections.

Bearing the aforementioned argument in mind, we make the most of the relation expressed by (15.171) to evaluate inner products of functions (or vectors). In this connection, we often need to calculate matrix elements of an operator. One of the most typical examples is matrix elements of Hamiltonian. In the field of molecular science, we have to estimate an overlap integral, Coulomb integral, resonance integral, etc. Other examples include transition matrix elements pertinent to, e.g., electric dipole transition. To this end, we deal with direct-product representation (see Sects. 15.7 and 15.8). Let $O^{(\gamma)}$ be an Hermitian operator belonging to the $\gamma$th irreducible representation and. Let us think of a following inner product:

$$\left\langle \phi_l^{(\beta)} | O^{(\gamma)} \psi_s^{(\alpha)} \right\rangle, \tag{16.48}$$

where $\psi_s^{(\alpha)}$ and $\phi_l^{(\beta)}$ are the $s$th component of $\alpha$th irreducible representation and the $l$th component of $\beta$th irreducible representation, respectively; see (15.146) and (15.161).

(i) Let us think of first the case where $O^{(\gamma)}$ is Hamiltonian $H$. Suppose that $\psi_s^{(\alpha)}$ belongs to an eigenvalue $\lambda$ of $H$. Then, we have

$$H\left|\psi_l^{(\alpha)}\right\rangle = \lambda\left|\psi_l^{(\alpha)}\right\rangle \tag{16.49}$$

In that case, with (16.48) we get

$$\left\langle \phi_l^{(\beta)} \middle| H\psi_s^{(\alpha)} \right\rangle = \lambda\left\langle \phi_l^{(\beta)} \middle| \psi_s^{(\alpha)} \right\rangle. \tag{16.50}$$

This equation is essentially the same as (15.171). That is,

$$\left\langle \phi_s^{(\beta)} \middle| \psi_l^{(\alpha)} \right\rangle = \delta_{\alpha\beta}\delta_{ls}\left\langle \phi_s^{(\beta)} \middle| \psi_l^{(\alpha)} \right\rangle \tag{15.171}$$

At the first glance, this equation seems trivial, but the equation gives us a powerful tool to save us a lot of troublesome calculations. In fact, if we

encounter a series of inner product calculations (i.e., definite integrals), we ignore many of them. It is because matrix elements of Hamiltonian do not vanish only if $\alpha = \beta$ and $l = s$. That is, we only have to estimate $\left\langle \phi_s^{(\alpha)} \middle| \psi_s^{(\alpha)} \right\rangle$. Otherwise, the inner products vanish. The functional forms of $\phi_s^{(\alpha)}$ and $\psi_s^{(\alpha)}$ are determined depending upon individual practical problems. We will show tangible examples later (Sect. 16.4).

(ii) Next, let us consider matrix elements of the optical transition. In this case, we are thinking of transition probability between quantum states. Assuming the dipole approximation, we use $\boldsymbol{\varepsilon}_e \cdot \boldsymbol{P}^{(\gamma)}$ for $O^{(\gamma)}$ in $\left\langle \phi_l^{(\beta)} \middle| O^{(\gamma)} \psi_s^{(\alpha)} \right\rangle$ of (16.48). The quantity $\boldsymbol{P}^{(\gamma)}$ represents an electric dipole moment associated with the position vector. Normally, we can readily find it in a character table, in which we examine which irreducible representation $\gamma$ the position vector components $x$, $y$, and $z$ indicated in a rightmost column correspond to. Table 15.4 is an example. If we take a unit polarization vector $\boldsymbol{\varepsilon}_e$ in parallel with the position vector component that the character table designates, $\boldsymbol{\varepsilon}_e \cdot \boldsymbol{P}^{(\gamma)}$ is nonvanishing. Suppose that $\psi_s^{(\alpha)}$ and $\phi_l^{(\beta)}$ are an initial state and final state, respectively. At the first glance, we would wish to use (15.200) and count how many times a representation $\beta$ occurs for a direct-product representation $D^{(\gamma \times \alpha)}$. It is often the case, however, where $\beta$ is not an irreducible representation.

Even in such a case, if either $\alpha$ or $\beta$ belongs to a totally symmetric representation, the handling will be easier. This situation corresponds to that an initial electronic configuration or a final electronic configuration forms a closed shell an electronic state of which is totally symmetric [1]. The former occurs when we consider the optical absorption that takes place, e.g., in a molecule of a ground electronic state. The latter corresponds to an optical emission that ends up with a ground state. Let us consider the former case. Since $O_j^{(\gamma)}$ is Hermitian, we rewrite (16.48) as

$$\left\langle \phi_l^{(\beta)} \middle| O_j^{(\gamma)} \psi_s^{(\alpha)} \right\rangle = \left\langle O_j^{(\gamma)} \phi_l^{(\beta)} \middle| \psi_s^{(\alpha)} \right\rangle = \left\langle \psi_s^{(\alpha)} \middle| O_j^{(\gamma)} \phi_l^{(\beta)} \right\rangle^* \tag{16.51}$$

where we assume that $\psi_s^{(\alpha)}$ is a ground state having a closed-shell electronic configuration. Therefore, $\psi_s^{(\alpha)}$ belongs to a totally symmetric irreducible representation and. For (16.48) not to vanish, therefore, we may alternatively state that it is necessary for

$$D^{(\gamma \times \beta)} = D^{(\gamma)} \otimes D^{(\beta)}$$

to contain $D^{(\alpha)}$ belonging to a totally symmetric representation. Note that in group theory we usually write $D^{(\gamma)} \times D^{(\beta)}$ instead of $D^{(\gamma)} \otimes D^{(\beta)}$; see (15.191).

If $\phi_l^{(\beta)}$ belongs to a reducible representation and, from (15.80) we have

$$D^{(\beta)} = \sum_{\omega} q_\omega D^{(\omega)},$$

where $D^{(\omega)}$ belongs to an irreducible representation and $\omega$. Then, we get

$$D^{(\gamma)} \times D^{(\beta)} = \sum_{\omega} q_\omega D^{(\gamma)} \times D^{(\omega)}.$$

Thus, applying (15.199) and (15.200), we can obtain a direct sum of irreducible representations. After that, we examine whether an irreducible representation $\alpha$ is contained in $D^{(\gamma)} \times D^{(\beta)}$. Since the totally symmetric representation is one-dimensional, $s = 1$ in (16.61).

## 16.3 Calculation Procedures of Molecular Orbitals (MOs)

We describe a brief outline of the MO method based on LCAO (LCAOMO). Suppose that a molecule consists of $n$ atomic orbitals and that each MO $\psi_i (1 \leq i \leq n)$ comprises a linear combination of those $n$ atomic orbitals $\phi_k (1 \leq k \leq n)$. That is, we have

$$\psi_i = \sum_{k=1}^{n} c_{ki} \phi_k (1 \leq i \leq n), \tag{16.52}$$

where $c_{ki}$ are complex coefficients. We assume that $\phi_k$ are normalized. That is,

$$\langle \phi_k | \phi_k \rangle \equiv \int \phi_k^* \phi_k \mathrm{d}\tau = 1, \tag{16.53}$$

where $\mathrm{d}\tau$ implies that an integration should be taken over $\mathbb{R}^3$. The notation $\langle g|f \rangle$ means an inner product defined in Sect. 11.1. The inner product is usually defined by a definite integral of $g^* f$ whose integration range covers a part or all of $\mathbb{R}^3$ depending on a constitution of a physical system.

First, we try to solve Schrödinger equation given as an eigenvalue equation. The said equation is described as

$$H\psi = \lambda\psi \text{ or } (H - \lambda\psi) = 0 \tag{16.54}$$

Replacing $\psi$ with (16.65), we have

$$\sum_{k=1}^{n} c_k (H - \lambda)\phi_k = 0 (1 \leq k \leq n), \tag{16.55}$$

where the subscript $i$ in (16.52) has been omitted for simplicity. Multiplying $\phi_j^*$ from the left and integrating over whole $\mathbb{R}^3$, we have

$$\sum_{k=1}^{n} c_k \int \phi_j^*(H-\lambda)\phi_k \mathrm{d}\tau = 0 \tag{16.56}$$

Rewriting (16.56), we get

$$\sum_{k=1}^{n} c_k \int \left(\phi_j^* H \phi_k - \lambda \phi_j \phi_k\right) \mathrm{d}\tau = 0 \tag{16.57}$$

Here, let us define following quantities:

$$H_{jk} = \int \phi_j^* H \phi_k \mathrm{d}\tau \text{ and } S_{jk} = \int \phi_j^* \phi_k \mathrm{d}\tau, \tag{16.58}$$

where $S_{ii} = 1 (1 \leq i \leq n)$ due to a normalized function of $\phi_i$. Then, we get

$$\sum_{k=1}^{n} c_k (H_{jk} - \lambda S_{jk}) = 0. \tag{16.59}$$

Rewriting (16.59) in a matrix form, we get

$$\begin{pmatrix} H_{11}-\lambda & \cdots & H_{1n}-\lambda S_{1n} \\ \vdots & \ddots & \vdots \\ H_{n1}-\lambda S_{n1} & \cdots & H_{nn}-\lambda \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = 0, \tag{16.60}$$

where note that $S_{ii} = 1 (1 \leq i \leq n)$ because of normalization of $\phi_i$. Suppose that by solving (16.59) or (16.60), we get $\lambda_i (1 \leq i \leq n)$, some of which may be identical (i.e., the degenerate case), and obtain $n$ different column eigenvectors corresponding to $n$ eigenvalues $\lambda_i$. In light of (10.4) of Sect. 7.3, the following condition must be satisfied for this to get eigenvectors for which not all $c_k$ is zero:

$$\det\left(H_{jk} - \lambda S_{jk}\right) = 0 \text{ or } \begin{vmatrix} H_{11}-\lambda & \cdots & H_{1n}-\lambda S_{1n} \\ \vdots & \ddots & \vdots \\ H_{n1}-\lambda S_{n1} & \cdots & H_{nn}-\lambda \end{vmatrix} = 0 \tag{16.61}$$

Equation (16.61) is called a secular equation. This equation is pertinent to a determinant of an order $n$, and so we are expected to get $n$ roots for this, some of which are identical (i.e., a degenerate case).

In the above discussion, it is useful to introduce the following notation:

$$\langle \phi_j | H \phi_k \rangle \equiv H_{jk} = \int \phi_j^* H \phi_k \mathrm{d}\tau \text{ and } \langle \phi_j | \phi_k \rangle \equiv S_{jk} = \int \phi_j^* \phi_k \mathrm{d}\tau. \tag{16.62}$$

The above notation has already been introduced in Sect. 1.4. Equation (16.62) certainly satisfies the definition of the inner product described in (11.2) to (11.4); readers, please check it. On the basis of (11.64), we have

$$\langle y | Hx \rangle^* = \left\langle x H^\dagger | y \right\rangle \tag{16.63}$$

Since $H$ is Hermitian, $H^\dagger = H$. That is,

$$\langle y | Hx \rangle^* = \langle xH | y \rangle. \tag{16.64}$$

To solve (16.61) with an enough large number $n$ is usually formidable. However, if we can find appropriate conditions, (16.61) can pretty easily be solved. An essential point rests upon how we deal with off-diagonal elements $H_{jk}$ and $S_{jk}$. If we are able to appropriately choose basis vectors so that we can get

$$H_{jk} = 0 \text{ and } S_{jk} = 0 \text{ for } j \neq k, \tag{16.65}$$

the secular equation is completely reduced to a simple form

$$\begin{vmatrix} \widetilde{H}_{11} - \lambda \widetilde{S}_{11} & & & \\ & \widetilde{H}_{22} - \lambda \widetilde{S}_{22} & & \\ & & \ddots & \\ & & & \widetilde{H}_{nn} - \lambda \widetilde{S}_{nn} \end{vmatrix} = 0, \tag{16.66}$$

where all the off-diagonal elements are zero, and an eigenvalue $\lambda$ is given by

$$\lambda_i = \widetilde{H}_{ii} / \widetilde{S}_{ii} \tag{16.67}$$

This means that the eigenvalue equation has automatically been solved.

The best way to achieve this is to choose basis functions (vectors) such that the functions conform to the symmetry which the molecule belongs to. That is, we "shuffle" the atomic orbitals as follows:

$$\xi_i = \sum_{k=1}^{n} d_{ki} \phi_k (1 \leq i \leq n), \tag{16.68}$$

where $\xi_i$ are new functions chosen instead of $\psi_i$ of (16.52). That is, we construct a linear combination of the atomic orbitals that belongs to individual irreducible representation. The said linear combination is called symmetry-adapted linear

combination (SALC). We remark that all atomic orbitals are not necessarily included in the SALC. Then, we have

$$\widetilde{H}_{jk} = \int \xi_j^* H \xi_k d\tau \text{ and } \widetilde{S}_{jk} = \int \xi_j^* \xi_k d\tau \tag{16.69}$$

Thus, instead of (16.61), we have a new secular equation of

$$\det\left(\widetilde{H}_{jk} - \lambda \widetilde{S}_{jk}\right) = 0 \tag{16.70}$$

Since the Hamiltonian $H$ is totally symmetric, in terms of the direct-product representation $H\xi_k$ in (16.69) belongs to an irreducible representation in which $\xi_k$ belongs to. This can intuitively be understood. But, to assert this, use (15.76) and (15.200) also the fact that characters of the totally symmetric representation are 1.

In light of (15.171) and (15.181), if $\xi_k$ and $\xi_j$ belong to different irreducible representations, $\widetilde{H}_{jk}$ and $\widetilde{S}_{jk}$ both vanish at once. From (15.172), at the same time, $\xi_k$ and $\xi_j$ are orthogonal to each other. The most ideal situation is to get (16.66) with all the off-diagonal elements vanishing. Regarding the diagonal elements of (16.70), we always have the direct product of the same representation, and hence, the integrals (16.69) do not vanish. Thus, we get a powerful guideline for the evaluation of (16.70) in an as simplest as possible form.

If the SALC orbitals $\xi_j$ and $\xi_k$ belong to the same irreducible representation, the relevant matrix elements are generally nonvanishing. Even in that case, however, according to Theorem 11.2 we can construct a set of orthonormal vectors by taking appropriate linear combination of $\xi_j$ and $\xi_k$. The resulting vectors naturally belong to the same irreducible representation. This can be done by solving the secular equation as described below. On top of it, Hermiticity of the Hamiltonian ensures that those vectors can be rendered orthogonal (see Theorem 12.5). If $n$ electrons are present in a molecule, we are to deal with $n$ SALC orbitals which we view as vectors accordingly. In terms of representation theory, these vectors span a representation space where the vectors undergo symmetry operations. Thus, we can construct orthonormal basis vectors belonging to various irreducible representations throughout the representation space.

To address our problems, we take the following procedures: (i) First, we have to determine a symmetry species (i.e., point group) of a molecule. (ii) Next, we pick up atomic orbitals contained in a molecule and examine how those orbitals are transformed by symmetry operations. (iii) We examine how those atomic orbitals are transformed according to symmetry operations. Since the symmetry operations are represented by a $(n, n)$ unitary matrix, we can readily decide a trace of the matrix. Generally, that matrix representation is reducible, and so we reduce the representation according to the procedures of (15.81)–(15.83). Thus, we are able to determine how many irreducible representations are contained in the original reducible representation. (iv) After having determined the irreducible representations, we construct SALCs and constitute a secular equation using them. (v) Solving the

secular equation, we determine molecular orbital energies and decide functional forms of the corresponding MOs. (vi) We examine physicochemical properties such as optical transition within a molecule.

In the procedure (iii) of the above, if the same irreducible representations appear more than once, we have to solve a secular equation of an order of two or more. Even in that case, we can render a set of resulting eigenvectors orthogonal to one another during the process of solving a problem.

To construct the above-mentioned SALC, we make the most of projection operators that are defined in Sect. 12.1. In (15.147), putting $m = l$, we have

$$P_{l(l)}^{(\alpha)} = \frac{d_\alpha}{n} \sum_g D_{ll}^{(\alpha)}(g)^* g \tag{16.71}$$

Or we can choose a projection operator $P^{(\alpha)}$ described as

$$P^{(\alpha)} = \frac{d_\alpha}{n} \sum_g \sum_{l=1}^{d_\alpha} D_{ll}^{(\alpha)}(g)^* g = \frac{d_\alpha}{n} \sum_g [\chi^{(\alpha)}(g)]^* g \tag{15.174}$$

In the one-dimensional representation and, $P_{l(l)}^{(\alpha)}$ and $P^{(\alpha)}$ are identical. As expressed in (15.155) and (15.175), these projection operators act on an arbitrary function and extract specific component(s) pertinent to a specific irreducible representation and of the point group which the molecule belongs to.

At first glance, the definition of $P_{l(l)}^{(\alpha)}$ and $P^{(\alpha)}$ looks daunting, but use of character tables relieves a calculation task. In particular, all the irreducible representations are one-dimensional (i.e., just a number!) with Abelian groups as mentioned in Sect. 15.6. For this, notice that individual group elements form a class by themselves. Therefore, utilization of character tables becomes easier for Abelian groups. Even though we encounter a case where a dimension of representation is more than one (i.e., the case of non-commutative groups), $D_{ll}^{(\alpha)}$ can be determined without much difficulty.

## 16.4 MO Calculations Based on π-Electron Approximation

On the basis of the general argument developed in Sects. 16.2 and 16.3, we perform molecular orbital calculations of individual molecules. First, we apply group theory to the molecular orbital calculations about aromatic hydrocarbons such as ethylene, cyclopropenyl radical, and cation as well as benzene and allyl radical. In these cases, in addition to adoption of the molecular orbital theory, we adopt so-called π-electron approximation. With the first three molecules, we will not have to "solve" a secular equation. But, for allyl radical, we deal with two SALC orbitals belonging

to the same irreducible representations, and hence, the final molecular orbitals must be obtained by solving the secular equation.

### 16.4.1 Ethylene

We start with one of the simplest examples, ethylene. Ethylene is a planar molecule and belongs to $D_{2h}$ symmetry (see Sect. 14.2). In the molecule, two π-electrons extend vertically to the molecular plane toward upper and lower directions (Fig. 16.5). The molecular plane forms a node to atomic $2p_z$ orbitals; that is, those atomic orbitals change a sign relative to the molecular plane (i.e., the *xy*-plane).

Figure 16.5, two $p_z$ atomic orbitals of carbon are denoted by $\phi_1$ and $\phi_2$. We should be able to construct basis vectors using $\phi_1$ and $\phi_2$. Corresponding to the two atomic orbitals, we are dealing with a two-dimensional vector space. Let us consider how $\phi_1$ and $\phi_2$ are transformed by a symmetry operation. First, we examine an operation $C_2(z)$. This operation exchanges $\phi_1$ and $\phi_2$. That is, we have

$$C_2(z)(\phi_1) = \phi_2 \tag{16.72}$$

and

$$C_2(z)(\phi_2) = \phi_1 \tag{16.73}$$

Equation (16.73) can be combined into a following equation:

$$(\phi_1 \phi_2) C_2(z) = (\phi_2 \phi_1) \tag{16.74}$$

Thus, using a matrix representation, we have

$$C_2(z) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{16.75}$$

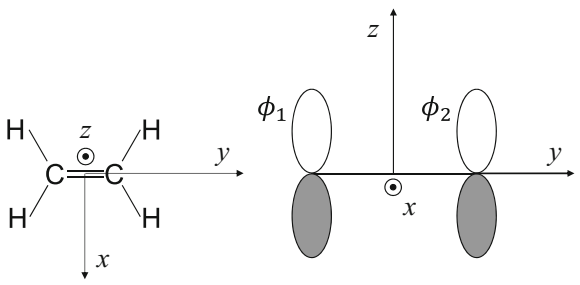


**Fig. 16.5** Ethylene molecule placed on the *xyz*-coordinate system. Two $p_z$ atomic orbitals of carbon are denoted by $\phi_1$ and $\phi_2$. The atomic orbitals change a sign relative to the molecular plane (i.e., the *xy*-plane)

In Sect. 14.2, we had

$$R_{z\theta} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{16.76}$$

or

$$C_2(z) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{16.77}$$

Whereas (16.76) and (16.77) show the transformation of a position vector in $\mathbb{R}^3$, (16.75) represents the transformation of functions in a two-dimensional vector space. A vector space composed of functions may be finite-dimensional or infinite-dimensional. We have already encountered the latter case in Part I where we dealt with the quantum mechanics of a harmonic oscillator. Such a function space is often referred to as a Hilbert space. An essence of (16.75) is characterized by that a trace (or character) of the matrix is zero.

Let us consider another transformation $C_2(y)$. In this case, the situation is different from the above case in that $\phi_1$ is converted to $-\phi_1$ by $C_2(y)$ and that $\phi_2$ is converted to $-\phi_2$. Notice again that the molecular plane forms a node to atomic $p$-orbitals. Thus, $C_2(y)$ is represented by

$$C_2(y) = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \tag{16.78}$$

The trace of the matrix $C_2(y)$ is $-2$. In this way, choosing $\phi_1$ and $\phi_2$ for the basis functions for the representation of $D_{2h}$ we can determine the characters for individual symmetry transformations of atomic $2p_z$ orbitals in ethylene belonging to $D_{2h}$. We collect the results in Table 16.1.

Next, we examine what kind of irreducible representations is contained in our present representation of ethylene. To do this, we need a character table of $D_{2h}$ (see Table 16.2). If a specific kind of irreducible representation is contained, then we want to examine how many times that specific representation takes place. Equation (15.83) is very useful for this purpose. In the present case, $n = 8$ in (15.83). Also taking into account (15.79) and (15.80), we get

$$\Gamma = B_{1u} \oplus B_{3g}, \tag{16.79}$$

**Table 16.1** Characters for individual symmetry transformations of atomic $2p_z$ orbitals in ethylene

| $D_{2h}$ | $E$ | $C_2(z)$ | $C_2(y)$ | $C_2(x)$ | $i$ | $\sigma(xy)$ | $\sigma(zx)$ | $\sigma(yz)$ |
|---|---|---|---|---|---|---|---|---|
| $\Gamma$ | 2 | 0 | −2 | 0 | 0 | −2 | 0 | 2 |

**Table 16.2** Character table of $D_{2h}$

| $D_{2h}$ | $E$ | $C_2(z)$ | $C_2(y)$ | $C_2(x)$ | $i$ | $\sigma(xy)$ | $\sigma(zx)$ | $\sigma(yz)$ | |
|---|---|---|---|---|---|---|---|---|---|
| $A_g$ | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | $x^2, y^2, z^2$ |
| $B_{1g}$ | 1 | 1 | −1 | −1 | 1 | 1 | −1 | −1 | $xy$ |
| $B_{2g}$ | 1 | −1 | 1 | −1 | 1 | −1 | 1 | −1 | $zx$ |
| $B_{3g}$ | 1 | −1 | −1 | 1 | 1 | −1 | −1 | 1 | $yz$ |
| $A_u$ | 1 | 1 | 1 | 1 | −1 | −1 | −1 | −1 | |
| $B_{1u}$ | 1 | 1 | −1 | −1 | −1 | −1 | 1 | 1 | $z$ |
| $B_{2u}$ | 1 | −1 | 1 | −1 | −1 | 1 | −1 | 1 | $y$ |
| $B_{3u}$ | 1 | −1 | −1 | 1 | −1 | 1 | 1 | −1 | $x$ |

where $\Gamma$ is a reducible representation for a set consisting of two $p_z$ atomic orbitals of ethylene. Equation (16.79) clearly shows that $\Gamma$ is a direct sum of two irreducible representations $B_{1u}$ and $B_{3g}$ that belong to $D_{2h}$. Instead of (16.79), we usually simply express the direct product in quantum chemical notation as

$$\Gamma = B_{1u} + B_{3g}. \tag{16.80}$$

As a next step, we are going to find an appropriate basis function belonging to two irreducible representations of (16.80). For this purpose, we use projection operators expressed as

$$P^{(\alpha)} = \frac{d_\alpha}{n} \sum\nolimits_g \left[\chi^{(\alpha)}(g)\right]^* g \tag{15.174}$$

Taking $\phi_1$ for instance, we apply (15.174) to $\phi_1$. That is,

$$\begin{aligned} P^{(B_{1u})}\phi_1 &= \frac{1}{8}\sum_g \left[\chi^{(B_{1u})}(g)\right]^* g\phi_1 \\ &= \frac{1}{8}[1\cdot\phi_1 + 1\cdot\phi_2 + (-1)(-\phi_1) + (-1)(-\phi_2) + (-1)(-\phi_2) \\ &\quad + (-1)(-\phi_1) + 1\cdot\phi_2 + 1\cdot\phi_1] = \frac{1}{2}(\phi_1 + \phi_2) \end{aligned} \tag{16.81}$$

Also with the $B_{3g}$, we apply (15.174) to $\phi_1$ and get

$$\begin{aligned} P^{(B_{3g})}\phi_1 &= \frac{1}{8}\sum_g \left[\chi^{(B_{3g})}(g)\right]^* g\phi_1 \\ &= \frac{1}{8}[1\cdot\phi_1 + (-1)\phi_2 + (-1)(-\phi_1) + 1(-\phi_2) + 1(-\phi_2) + (-1)(-\phi_1) \\ &\quad + (-1)\phi_2 + 1\cdot\phi_1] = \frac{1}{2}(\phi_1 - \phi_2) \end{aligned} \tag{16.82}$$

Thus, after going through routine but sure procedures, we have reached appropriate basis functions that belong to each irreducible representation.

As mentioned earlier (see Table 14.5), $D_{2h}$ can be expressed as a direct-product group and $C_2$ group is contained in $D_{2h}$ as a subgroup. We can use it for the present analysis. Suppose now that we have a group $\mathcal{G}$ and that $\mathcal{H}$ is a subgroup of $\mathcal{G}$. Let $g$ be an arbitrary element of $\mathcal{G}$ and let $D(g)$ be a representation of $g$. Meanwhile, let $h$ be an arbitrary element of $\mathcal{H}$. Then, with $^{\forall}h \in \mathcal{H}$, a collection of $D(h)$ is a representation of $\mathcal{H}$. We write this relation as

$$D \downarrow \mathcal{H} \tag{16.83}$$

This representation is called a subduced representation of $D$ to $\mathcal{H}$. Table 16.3 shows a character table of irreducible representations of $C_2$. In the present case, we are thinking of $C_2(z)$ as $C_2$; see Table 14.5 and Fig. 16.5. Then, we have

$$B_{1u} \downarrow C_2 = A \text{ and } B_{3g} \downarrow C_2 = B. \tag{16.84}$$

The expression of (16.84) is called a compatibility relation. Note that in (16.84) $C_2$ is not a symmetry operation, but means a subgroup of $D_{2h}$. Thus, (16.81) is reduced to

$$P^{(A)}\phi_1 = \frac{1}{2}\sum_g \left[\chi^{(A)}(g)\right]^* g\phi_1 = \frac{1}{2}[1\cdot\phi_1 + 1\cdot\phi_2] = \frac{1}{2}(\phi_1 + \phi_2) \tag{16.85}$$

Also, we have

$$P^{(B)}\phi_1 = \frac{1}{2}\sum_g \left[\chi^{(B)}(g)\right]^* g\phi_1 = \frac{1}{2}[1\cdot\phi_1 + (-1)\phi_2] = \frac{1}{2}(\phi_1 - \phi_2) \tag{16.86}$$

The relations of (16.85) and (16.86) are essentially the same as (16.81) and (16.82), respectively.

We can easily construct a character table (see Table 16.3). There should be two irreducible representations. Regarding the totally symmetric representation, we allocate 1 to each symmetry operation. For another representation, we allocate 1 to an identity element $E$ and $-1$ to an element $C_2$ so that the row and column of the character table are orthogonal to each other.

Readers might well ask why we bother to make circuitous approaches to reaching predictable results such as (16.81) and (16.82) or (16.85) and (16.86). This question seems natural when we are dealing with a case where the number of basis vectors (i.e., a dimension of the vector space) is small, typically two as in the

**Table 16.3** Character table of $C_2$

| $C_2$ | $E$ | $C_2$ | |
|---|---|---|---|
| $A$ | 1 | 1 | $z; x^2, y^2, z^2, xy$ |
| $B$ | 1 | $-1$ | $x, y; yz, zx$ |

present case. With increasing dimension of the vector space, however, to seek and determine appropriate SALCs become increasingly complicated and difficult. Under such circumstances, a projection operator is an indispensable tool to address the problems.

We have a two-dimensional secular equation to be solved such that

$$\begin{vmatrix} \widetilde{H}_{11} - \lambda \widetilde{S}_{11} & \widetilde{H}_{12} - \lambda \widetilde{S}_{12} \\ \widetilde{H}_{21} - \lambda \widetilde{S}_{21} & \widetilde{H}_{22} - \lambda \widetilde{S}_{22} \end{vmatrix} = 0 \tag{16.87}$$

In the above argument, $\frac{1}{2}(\phi_1 + \phi_2)$ belongs to $B_{1u}$ and $\frac{1}{2}(\phi_1 - \phi_2)$ belongs to $B_{3g}$. Since they belong to different irreducible representations, we have

$$\widetilde{H}_{12} = \widetilde{S}_{12} = 0 \tag{16.88}$$

Thus, the secular equation (16.84) is reduced to

$$\begin{vmatrix} \widetilde{H}_{11} - \lambda \widetilde{S}_{11} & 0 \\ 0 & \widetilde{H}_{22} - \lambda \widetilde{S}_{22} \end{vmatrix} = 0 \tag{16.89}$$

As expected, (16.89) has automatically been solved to give a solution

$$\lambda_1 = \widetilde{H}_{11}/\widetilde{S}_{11} \text{ and } \lambda_2 = \widetilde{H}_{22}/\widetilde{S}_{22} \tag{16.90}$$

The next step is to determine the energy eigenvalue of the molecule. Note here that a role of SALCs is to determine a suitable irreducible representation that corresponds to a "direction" of a vector. As the coefficient keeps the direction of a vector unaltered, it would be of secondary importance. The final form of normalized MOs can be decided last. That procedure includes the normalization of a vector. Thus, we tentatively choose the following functions for SALCs, i.e.,

$$\xi_1 = \phi_1 + \phi_2 \text{ and } \xi_2 = \phi_1 - \phi_2$$

Then, we have

$$\begin{aligned} \widetilde{H}_{11} &= \int \xi_1^* H \xi_1 \mathrm{d}\tau = \int (\phi_1 + \phi_2)^* H (\phi_1 + \phi_2) \mathrm{d}\tau \\ &= \int \phi_1^* H \phi_1 \mathrm{d}\tau + \int \phi_1^* H \phi_2 \mathrm{d}\tau + \int \phi_2^* H \phi_1 \mathrm{d}\tau + \int \phi_2^* H \phi_2 \mathrm{d}\tau \\ &= H_{11} + H_{12} + H_{21} + H_{22} = H_{11} + 2H_{12} + H_{22} \end{aligned} \tag{16.91}$$

Similarly, we have

$$\widetilde{H}_{22} = \int \xi_2^* H \xi_2 \mathrm{d}\tau = H_{11} - 2H_{12} + H_{22} \tag{16.92}$$

The last equality comes from the fact that we have chosen real functions for $\phi_1$ and $\phi_2$ as studied in Part I. Moreover, we have

$$\begin{aligned} H_{11} &\equiv \int \phi_1^* H \phi_1 \mathrm{d}\tau = \int \phi_1 H \phi_1 \mathrm{d}\tau = \int \phi_2 H \phi_2 \mathrm{d}\tau = H_{22}, \\ H_{12} &\equiv \int \phi_1^* H \phi_2 \mathrm{d}\tau = \int \phi_1 H \phi_2 \mathrm{d}\tau = \int \phi_2^* H \phi_1 \mathrm{d}\tau = \int \phi_2 H \phi_1 \mathrm{d}\tau = H_{21} \end{aligned} \tag{16.93}$$

The first equation comes from the fact that both $H_{11}$ and $H_{22}$ are calculated using the same $2p_z$ atomic orbital of carbon. The second equation results from the fact that $H$ is Hermitian. Notice that both $\phi_1$ and $\phi_2$ are real functions. Following the convention, we denote

$$\alpha \equiv H_{11} = H_{22} \text{ and } \beta \equiv H_{12} \tag{16.94}$$

where $\alpha$ is called Coulomb integral and $\beta$ is said to be resonance integral. Then, we have

$$\widetilde{H}_{11} = 2(\alpha + \beta) \tag{16.95}$$

In a similar manner, we get

$$\widetilde{H}_{22} = 2(\alpha - \beta) \tag{16.96}$$

Meanwhile, we have

$$\begin{aligned} \widetilde{S}_{11} = \langle \xi_1 | \xi_1 \rangle &= \int (\phi_1 + \phi_2)^* (\phi_1 + \phi_2) \mathrm{d}\tau = \int (\phi_1 + \phi_2)^2 \mathrm{d}\tau \\ &= \int \phi_1^2 \mathrm{d}\tau + \int \phi_2^2 \mathrm{d}\tau + 2 \int \phi_1 \phi_2 \mathrm{d}\tau = 2 + 2 \int \phi_1 \phi_2 \mathrm{d}\tau, \end{aligned} \tag{16.97}$$

where we used the fact that $\phi_1$ and $\phi_2$ have been normalized. Also following the convention, we denote

$$S \equiv \int \phi_1 \phi_2 \mathrm{d}\tau = S_{12}, \tag{16.98}$$

where $S$ is called overlap integral. Thus, we have

$$\widetilde{S}_{11} = 2(1 + S) \tag{16.99}$$

Similarly, we get

$$\widetilde{S}_{22} = 2(1 - S) \tag{16.100}$$

Substituting (16.95) and (16.96) along with (16.99) and (16.100) for (16.90), we get as the energy eigenvalue

$$\lambda_1 = \frac{\alpha + \beta}{1 + S} \text{ and } \lambda_2 = \frac{\alpha - \beta}{1 - S} \tag{16.101}$$

From (16.97), we get

$$||\xi_1|| = \sqrt{\langle \xi_1 | \xi_1 \rangle} = \sqrt{2(1+S)} \tag{16.102}$$

Thus, for one of MOs corresponding to an energy eigenvalue $\lambda_1$, we get

$$\Psi_1 = \frac{|\xi_1\rangle}{||\xi_1||} = \frac{\phi_1 + \phi_2}{\sqrt{2(1+S)}} \tag{16.103}$$

Also, we have

$$||\xi_2|| = \sqrt{\langle \xi_2 | \xi_2 \rangle} = \sqrt{2(1-S)} \tag{16.104}$$

For another MO corresponding to an energy eigenvalue $\lambda_2$, we get

$$\Psi_2 = \frac{|\xi_2\rangle}{||\xi_2||} = \frac{\phi_1 - \phi_2}{\sqrt{2(1-S)}} \tag{16.105}$$

Note that both normalized MOs and energy eigenvalues depend upon whether we ignore an overlap integral, as being the case with the simplest Hückel approximation. Nonetheless, MO functional forms (in this case, either $\phi_1 + \phi_2$ or $\phi_1 - \phi_2$) remain the same regardless of the approximation levels about the overlap integral. Regarding quantitative evaluation of $\alpha$, $\beta$, and $S$, we will briefly mention it later.

Once we have decided symmetry of MO (or an irreducible representation which the orbital belongs to) and its energy eigenvalue, we will be in a position to examine various physicochemical properties of the molecule. One of them is an optical transition within a molecule, particularly electric dipole transition. In most cases, the most important transition is that occurring among the highest occupied molecular orbital (HOMO) and lowest unoccupied molecular orbital (LUMO). In the case of ethylene, those levels are depicted in Fig. 16.6. In a ground state

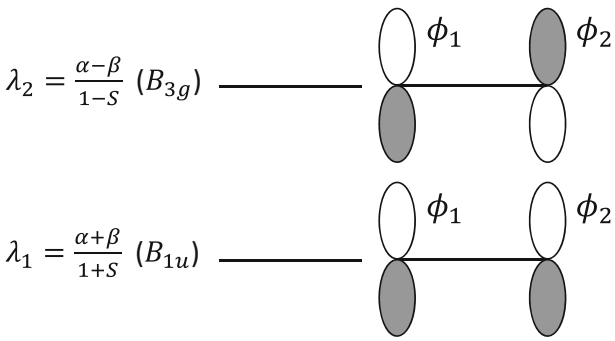


**Fig. 16.6** HOMO and LUMO energy levels and their assignments of ethylene

(i.e., the most stable state), two electrons are positioned in a $B_{1u}$ state (HOMO). An excited state is assigned to $B_{3g}$ (LUMO). The ground state that consists only of fully occupied MOs belongs to a totally symmetric representation. In the case of ethylene, the ground state belongs to $A_g$ accordingly.

If a photon is absorbed by a molecule, that molecule is excited by an energy of the photon. In ethylene, this process takes place by exciting an electron from $B_{1u}$ to $B_{3g}$. The resulting electronic state ends up as an electron remaining in $B_{1u}$ and another excited to $B_{3g}$ (a final state). Thus, the representation of the final state electronic configuration (denoted by $\Gamma_f$) is described as

$$\Gamma_f = B_{1u} \times B_{3g} \tag{16.106}$$

That is, the final excited state is expressed as a direct product of the states associated with the optical transition. To determine the symmetry of the final sate, we use (15.198) and (15.200). If $\Gamma$ in (16.106) is reducible, using (15.200) we can determine the number of times that individual representations take place. Calculating $\chi^{(B_{1u})}(g)\chi^{(B_{3g})}(g)$ for each group element, we can readily get the result. Using a character table, we have

$$\Gamma_f = B_{2u} \tag{16.107}$$

Thus, we find that the transition is $A_g \rightarrow B_{2u}$, where $A_g$ is called an initial state electronic configuration and $B_{2u}$ is called a final state electronic configuration. This transition is characterized by an electric dipole transition moment operator $\boldsymbol{P}$. Here, we have an important physical quantity of transition matrix element. This quantity $T_{fi}$ is approximated by

$$T_{fi} \equiv \left\langle \Theta_f | \boldsymbol{\varepsilon}_e \cdot \boldsymbol{P} | \Theta_i \right\rangle, \tag{16.108}$$

where $\boldsymbol{\varepsilon}_e$ is a unit polarization vector of the electric field; $\Theta_i$ and $\Theta_f$ are the initial state and final sate electronic configuration, respectively. The description (16.108) is in parallel with (4.5). Notice that in (4.5) we dealt with a single-electron system such as a particle confined in a square-well potential, a sole one-dimensional harmonic oscillator, and an electron in a hydrogen atom.

In the present case, however, we are dealing with a two-electron system, ethylene. Consequently, we cannot describe $\Theta_f$ by a simple wave function, but must use a more elaborated function. Nonetheless, when we discuss the optical transition of a molecule, it is often the case that when we study optical absorption or optical emission we first wish to know whether such a phenomenon truly takes place. In such a case, qualitative prediction for this is of great importance. This can be done by judging whether the integral (16.108) vanishes. If the integral does not vanish, the relevant transition is said to be *allowed*. If on the other hand the integral vanishes, the transition is called *forbidden*. In this context, a systematic approach based on group theory is a powerful tool for this.

Let us consider optical absorption of ethylene. With the transition matrix element for this, we have

$$T_{fi} = \left\langle \Theta_f^{(B_{2u})} | \boldsymbol{\varepsilon}_e \cdot \boldsymbol{P} | \Theta_i^{(A_g)} \right\rangle \tag{16.109}$$

Again, note that a closed-shell electronic configuration belongs to a totally symmetric representation [1]. Suppose that the position vector $\boldsymbol{x}$ belongs to an irreducible representation $\eta$. A necessary condition for (16.109) not to vanish is that $D^{(\eta)} \times D^{(A_g)}$ contains the irreducible representation $D^{(B_{2u})}$. In the present case, all the representations are one-dimensional, and so we can use $\chi^{(\omega)}$ instead of $D^{(\omega)}$, where $\omega$ shows an arbitrary irreducible representation of $D_{2h}$. This procedure is straightforward as shown in Sect. 16.2.

However, if the character is real (it is the case with many symmetry groups and with the point group $D_{2h}$ as well), the situation will be easier. Suppose in general that we are examining whether a following matrix element vanishes:

$$M_{fi} = \left\langle \Phi_f^{(\beta)} \middle| O^{(\gamma)} \middle| \Phi_i^{(\alpha)} \right\rangle, \tag{16.110}$$

where $\alpha$, $\beta$, and $\gamma$ stand for irreducible representations, and $O$ is an appropriate operator. In this case, (15.200) can be rewritten as

$$\begin{aligned} q_\alpha &= \frac{1}{n}\sum_g \chi^{(\alpha)}(g)^* \chi^{(\gamma\times\beta)}(g) = \frac{1}{n}\sum_g \chi^{(\alpha)}(g)\chi^{(\gamma\times\beta)}(g) \\ &= \frac{1}{n}\sum_g \chi^{(\alpha)}(g)\chi^{(\gamma)}(g)\chi^{(\beta)}(g) = \frac{1}{n}\sum_g \chi^{(\gamma)}(g)\chi^{(\alpha)}(g)\chi^{(\beta)}(g) \\ &= \frac{1}{n}\sum_g \chi^{(\gamma)}(g)\chi^{(\alpha\times\beta)}(g) = \frac{1}{n}\sum_g \chi^{(\gamma)}(g)^* \chi^{(\alpha\times\beta)}(g) = q_\gamma \end{aligned} \tag{16.111}$$

Consequently, the number of times that $D^{(\gamma)}$ appears in $D^{(\alpha\times\beta)}$ is identical to the number of times that $D^{(\alpha)}$ appears in $D^{(\gamma\times\beta)}$. Thus, it suffices to examine whether $D^{(\gamma\times\beta)}$ contains $D^{(\alpha)}$. In other words, we only have to examine whether $q_\alpha \neq 0$ in (16.111).

Thus, applying (16.111)–(16.109), we examine whether $D^{(B_{2u}\times A_g)}$ contains the irreducible representation $D^{(\eta)}$ that is related to $\boldsymbol{x}$. We easily get

$$B_{2u} = B_{2u} \times A_g \tag{16.112}$$

Therefore, if $\boldsymbol{\varepsilon}_e \cdot \boldsymbol{P}$ (or $\boldsymbol{x}$) belongs to $B_{2u}$, the transition is allowed. Consulting the character table, we find that $y$ belongs to $B_{2u}$. In this case, in fact (16.111) reads as

$$q_{B_{2u}} = \frac{1}{8}[1 \times 1 + (-1) \times (-1) + 1 \times 1 + (-1) \times (-1) + (-1) \times (-1) + 1 \times 1 + (-1) \times (-1) + 1 \times 1] = 1$$

Equivalently, we simply write

$$B_{2u} \times B_{2u} = A_g$$

This means that if a light polarized along the $y$-axis is incident, i.e., $\boldsymbol{\varepsilon}_e$ is parallel to the $y$-axis, the transition is allowed. In that situation, ethylene is said to be polarized along the $y$-axis or polarized in the direction of the $y$-axis. As a molecular axis is parallel to the $y$-axis, ethylene is polarized in the direction of the molecular axis. This is often the case with aromatic molecules having a well-defined molecular long axis such as ethylene.

We would examine whether ethylene is polarized along, e.g., the $x$-axis. From a character table of $D_{2h}$ (see Table 16.2), $x$ belongs to $B_{3u}$. In that case, using (16.111) we have

$$q_{B_{3u}} = \frac{1}{8}[1 \times 1 + (-1) \times (-1) + 1 \times (-1) + (-1) \times 1 + (-1) \times (-1) + 1 \times 1 + (-1) \times 1 + 1 \times (-1)] = 0.$$

This implies that $B_{3u}$ is not contained in $B_{1u} \times B_{3g} (= B_{2u})$.

The above results on the optical transitions are quite obvious. Once we get used to using a character table, quick estimation will be done.

### 16.4.2 Cyclopropenyl Radical [1]

Let us think of another example, cyclopropenyl radical that has three resonant structures (Fig. 16.7). It is a planar molecule, and three carbon atoms form an equilateral triangle. Hence, the molecule belongs to $D_{3h}$ symmetry (see Sect. 14.2). In the molecule, three π-electrons extend vertically to the molecular plane toward upper and lower directions. Suppose that cyclopropenyl radical is placed on the $xy$-plane. Three $p_z$ atomic orbitals of carbons located at vertices of an equilateral triangle are denoted by $\phi_1$, $\phi_2$, and $\phi_3$ in Fig. 16.8. The orbitals are numbered *clockwise* so that the calculations can be consistent with the conventional notation

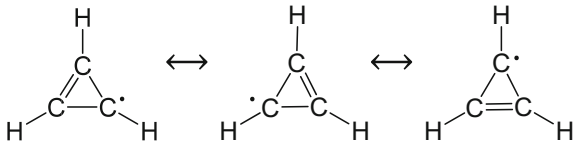

**Fig. 16.7** Three resonant structures of cyclopropenyl radical

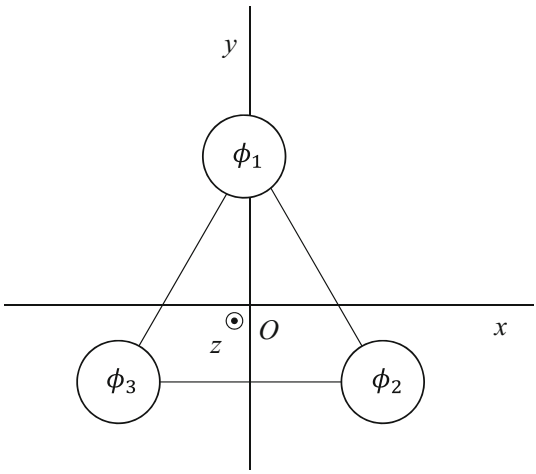


**Fig. 16.8** Three $p_z$ atomic orbitals of carbons for cyclopropenyl radical that is placed on the *xy*-plane. The carbon atoms are located at vertices of an equilateral triangle. The atomic orbitals are denoted by $\phi_1$, $\phi_2$, and $\phi_3$

of a character table (vide infra). We assume that these π-orbitals take positive and negative signs on the upper and lower sides of the plane of paper, respectively, with a nodal plane lying on the *xy*-plane. The situation is similar to that of ethylene, and the problem can be treated in parallel with the case of ethylene.

As in the case of ethylene, we can choose $\phi_1$, $\phi_2$, and $\phi_3$ as real functions. We construct basis vectors using these vectors. What we want to do to address the problem is as follows: (i) We examine how $\phi_1$, $\phi_2$, and $\phi_3$ are transformed by the symmetry operations of $D_{3h}$. According to the analysis, we can determine what irreducible representations SALCs should be assigned. (ii) On the basis of knowledge obtained in (i), we construct proper MOs.

In Table 16.4, we list a character table of $D_{3h}$ along with symmetry species. First, we examine traces (characters) of representation matrices. Similarly, in the case of ethylene, a subgroup $C_3$ of $D_{3h}$ plays an essential role (vide infra). This subgroup contains three group elements such that

$$C_3 = \left\{E, C_3, C_3^2\right\}$$

**Table 16.4** Character table of $D_{3h}$

| $D_{3h}$ | $E$ | $2C_3$ | $3C_2$ | $\sigma_h$ | $2S_3$ | $3\sigma_v$ | |
|---|---|---|---|---|---|---|---|
| $A_1'$ | 1 | 1 | 1 | 1 | 1 | 1 | $x^2+y^2, z^2$ |
| $A_2'$ | 1 | 1 | −1 | 1 | 1 | −1 | |
| $E'$ | 2 | −1 | 0 | 2 | −1 | 0 | $(x, y); (x^2-y^2, xy)$ |
| $A_1''$ | 1 | 1 | 1 | −1 | −1 | −1 | |
| $A_2''$ | 1 | 1 | −1 | −1 | −1 | 1 | $z$ |
| $E''$ | 2 | −1 | 0 | −2 | 1 | 0 | $(yz, zx)$ |

In the above, we use the same notation for the group name and group element, and so we should be careful not to confuse them.

They are transformed as follows:

$$C_3(z)(\phi_1) = \phi_3, C_3(z)(\phi_2) = \phi_1, \text{and } C_3(z)(\phi_3) = \phi_2; \tag{16.113}$$

$$C_3^2(z)(\phi_1) = \phi_2, C_3^2(z)(\phi_2) = \phi_3, \text{and } C_3^2(z)(\phi_3) = \phi_1. \tag{16.114}$$

Equation (16.113) can be combined into a following form:

$$(\phi_1\phi_2\phi_3)C_3(z) = (\phi_3\phi_1\phi_2) \tag{16.115}$$

Using a matrix representation, we have

$$C_3(z) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}. \tag{16.116}$$

In turn, (16.114) is expressed as

$$C_3^2(z) = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \tag{16.117}$$

Both traces of (16.116) and (16.117) are zero.

Similarly, let us check the representation matrices of other symmetry species. Of these, e.g., for $C_2$ related to the $y$-axis (see Fig. 16.8) we have

$$(\phi_1\phi_2\phi_3)C_2 = (-\phi_1 - \phi_3 - \phi_2) \tag{16.118}$$

Therefore,

$$C_2 = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix} \tag{16.119}$$

We have a trace $-1$ accordingly. Regarding $\sigma_h$, we have

$$(\phi_1\phi_2\phi_3)\sigma_h = (-\phi_1 - \phi_2 - \phi_3) \text{ and } \sigma_h = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \tag{16.120}$$

In this way, we can determine the trace for individual symmetry transformations of basis functions $\phi_1$, $\phi_2$, and $\phi_3$. We collect the results of characters of a reducible representation $\Gamma$ in Table 16.5. It can be reduced to a summation of irreducible representations according to the procedures given in (15.81) to (15.83) and using a character table of $D_{3h}$ (Table 16.4). As a result, we get

**Table 16.5** Characters for individual symmetry transformations of $2p_z$ orbitals in cyclopropenyl radical

| $D_{3h}$ | $E$ | $2C_3$ | $3C_2$ | $\sigma_h$ | $2S_3$ | $3\sigma_v$ |
|---|---|---|---|---|---|---|
| $\Gamma$ | 3 | 0 | −1 | −3 | 0 | 1 |

$$\Gamma = A_2'' + E'' \tag{16.121}$$

As in the case of ethylene, we make the best use of the information of a subgroup $C_3$ of $D_{3h}$. Let us consider a subduced representation of $D$ of $D_{3h}$ to $C_3$. For this, in Table 16.6 we show a character table of irreducible representations of $C_3$. We can readily construct this character table. There should be three irreducible representations. Regarding the totally symmetric representation, we allocate 1 to each symmetry operation. Hence, for other representations, we allocate 1 to an identity element $E$ and two other triple roots of 1, i.e., $\varepsilon$ and $\varepsilon^*$ [where $\varepsilon = \exp(i2\pi/3)$] to an element $C_3$ and $C_3^2$ as shown so that the row and column vectors of the character table are orthogonal to each other.

Returning to the construction of the subduced representation, we have

$$A_2'' \downarrow C_3 = A \text{ and } E'' \downarrow C_3 = E \tag{16.122}$$

Then, corresponding to (15.147), we have

$$\begin{aligned} P^{(A)}\phi_1 &= \frac{1}{3}\sum_g \left[\chi^{(A)}(g)\right]^* g\phi_1 = \frac{1}{3}[1\cdot\phi_1 + 1\cdot\phi_2 + 1\cdot\phi_3] \\ &= \frac{1}{3}(\phi_1 + \phi_2 + \phi_3) \end{aligned} \tag{16.123}$$

Also, we have

$$\begin{aligned} P^{[E^{(1)}]}\phi_1 &= \frac{1}{3}\sum_g \left[\chi^{[E^{(1)}]}(g)\right]^* g\phi_1 = \frac{1}{3}[1\cdot\phi_1 + \varepsilon^*\phi_3 + (\varepsilon^*)^*\phi_2] \\ &= \frac{1}{3}(\phi_1 + \varepsilon\phi_2 + \varepsilon^*\phi_3) \end{aligned} \tag{16.124}$$

**Table 16.6** Character table of $C_3$

| $C_3$ | $E$ | $C_3$ | $C_3^2$ | $\varepsilon = \exp(i2\pi/3)$ |
|---|---|---|---|---|
| $A$ | 1 | 1 | 1 | $z; x^2+y^2, z^2$ |
| $E$ | $\left\{\begin{matrix}1\\1\end{matrix}\right.$ | $\begin{matrix}\varepsilon\\\varepsilon^*\end{matrix}$ | $\left.\begin{matrix}\varepsilon^*\\\varepsilon\end{matrix}\right\}$ | $(x,y); (x^2-y^2, xy), (yz, zx)$ |

Also, we get

$$P^{[E^{(2)}]}\phi_1 = \frac{1}{3}\sum_g [\chi^{[E^{(2)}]}(g)]^* g\phi_1 = \frac{1}{3}[1\cdot\phi_1 + (\varepsilon^*)^*\phi_3 + \varepsilon^*\phi_2] = \frac{1}{3}(\phi_1 + \varepsilon^*\phi_2 + \varepsilon\phi_3) \tag{16.125}$$

Here is the best place to mention an eigenvalue of a symmetry operator. Let us designate SALCs as follows:

$$\xi_1 = \phi_1 + \phi_2 + \phi_3, \xi_2 = \phi_1 + \varepsilon\phi_2 + \varepsilon^*\phi_3, \text{and } \xi_3 = \phi_1 + \varepsilon^*\phi_2 + \varepsilon\phi_3 \tag{16.126}$$

Let us choose $C_3$ for a symmetry operator. Then, we have

$$C_3(\xi_1) = C_3(\phi_1 + \phi_2 + \phi_3) = C_3\phi_1 + C_3\phi_2 + C_3\phi_3 = \phi_3 + \phi_1 + \phi_2 = \xi_1, \tag{16.127}$$

where for the second equality we used the fact that $C_3$ is a linear operator. That is, regarding a SALC $\xi_1$, an eigenvalue of $C_3$ is 1.

Similarly, we have

$$C_3(\xi_2) = C_3(\phi_1 + \varepsilon\phi_2 + \varepsilon^*\phi_3) = C_3\phi_1 + \varepsilon C_3\phi_2 + \varepsilon^* C_3\phi_3 = \phi_3 + \varepsilon\phi_1 + \varepsilon^*\phi_2 = \varepsilon(\phi_1 + \varepsilon\phi_2 + \varepsilon^*\phi_3) = \varepsilon\xi_2 \tag{16.128}$$

Furthermore, we get

$$C_3(\xi_3) = \varepsilon^*\xi_3 \tag{16.129}$$

Thus, we find that regarding SALCs $\xi_2$ and $\xi_3$, eigenvalues of $C_3$ are $\varepsilon$ and $\varepsilon^*$, respectively. These pieces of information imply that if we appropriately choose proper functions for basis vectors, a character of a symmetry operation for a one-dimensional representation is identical to an eigenvalue of the said symmetry operation (see Table 16.6).

Regarding the last parts of the calculations, we follow the procedures described in the case of ethylene. Using the above functions $\xi_1$, $\xi_2$, and $\xi_3$, we construct the secular equation such that

$$\begin{vmatrix} \widetilde{H}_{11} - \lambda\widetilde{S}_{11} & & \\ & \widetilde{H}_{22} - \lambda\widetilde{S}_{22} & \\ & & \widetilde{H}_{33} - \lambda\widetilde{S}_{33} \end{vmatrix} = 0 \tag{16.130}$$

Since we have obtained three SALCs that are assigned to individual irreducible representations $A$, $E^{(1)}$, and $E^{(2)}$, these SALCs span the representation space $V^3$.

This makes off-diagonal elements of the secular equation vanish, and it is simplified as in (16.130). Here, we have

$$\begin{aligned}\widetilde{H}_{11} &= \int \xi_1^* H \xi_1 \mathrm{d}\tau = \int (\phi_1 + \phi_2 + \phi_3)^* H (\phi_1 + \phi_2 + \phi_3)\mathrm{d}\tau \\ &= 3(\alpha + 2\beta)\end{aligned} \tag{16.131}$$

where we used the same $\alpha$ and $\beta$ as defined in (16.94). Strictly speaking, $\alpha$ and $\beta$ appearing in (16.131) should be slightly different from those of (16.94), because a Hamiltonian is different. This approximation, however, would be enough for the present studies. In a similar manner, we get

$$\widetilde{H}_{22} = \int \xi_2^* H \xi_2 \mathrm{d}\tau = 3(\alpha - \beta) \text{ and } \widetilde{H}_{33} = \int \xi_3^* H \xi_3 \mathrm{d}\tau = 3(\alpha - \beta). \tag{16.132}$$

Meanwhile, we have

$$\begin{aligned}\widetilde{S}_{11} = \langle \xi_1 | \xi_1 \rangle &= \int (\phi_1 + \phi_2 + \phi_3)^* (\phi_1 + \phi_2 + \phi_3)\mathrm{d}\tau = \int (\phi_1 + \phi_2 + \phi_3)^2 \mathrm{d}\tau \\ &= 3(1 + 2S)\end{aligned} \tag{16.133}$$

Similarly, we get

$$\widetilde{S}_{22} = \widetilde{S}_{33} = 3(1 - S) \tag{16.134}$$

Readers are urged to verify (16.133) and (16.134).

Substituting (16.131) through (16.134) for (16.130), we get as the energy eigenvalue

$$\lambda_1 = \frac{\alpha + 2\beta}{1 + 2S}, \lambda_2 = \frac{\alpha - \beta}{1 - S}, \text{and } \lambda_3 = \frac{\alpha - \beta}{1 - S} \tag{16.135}$$

Notice that two MOs belonging to $E''$ have the same energy. These MOs are said to be energetically *degenerate*. This situation is characteristic of a two-dimensional representation. Actually, even though the group $C_3$ has only one-dimensional representations (because it is an Abelian group), the two complex conjugate representations labeled $E$ behave as if they were a two-dimensional representation [2]. We will again encounter the same situation in a next example, benzene.

From (16.126), we get

$$||\xi_1|| = \sqrt{\langle \xi_1 | \xi_1 \rangle} = \sqrt{3(1 + 2S)} \tag{16.136}$$

Thus, as one of MOs corresponding to an energy eigenvalue $\lambda_1$, i.e., $\Psi_1$, we get

$$\Psi_1 = \frac{|\xi_1\rangle}{||\xi_1||} = \frac{\phi_1 + \phi_2 + \phi_3}{\sqrt{3(1 + 2S)}} \tag{16.137}$$

Also, we have

$$||\xi_2|| = \sqrt{\langle \xi_2 | \xi_2 \rangle} = \sqrt{3(1-S)} \text{ and } ||\xi_3|| = \sqrt{\langle \xi_3 | \xi_3 \rangle} = \sqrt{3(1-S)} \quad (16.138)$$

Thus, for another MO corresponding to an energy eigenvalue $\lambda_2$, we get

$$\Psi_2 = \frac{|\xi_2\rangle}{||\xi_2||} = \frac{\phi_1 + \varepsilon\phi_2 + \varepsilon^*\phi_3}{\sqrt{3(1-S)}} \quad (16.139)$$

Also, with a MO corresponding to $\lambda_3 (= \lambda_2)$, we have

$$\Psi_3 = \frac{|\xi_3\rangle}{||\xi_3||} = \frac{\phi_1 + \varepsilon^*\phi_2 + \varepsilon\phi_3}{\sqrt{3(1-S)}} \quad (16.140)$$

Equations (16.139) and (16.140) include complex numbers, and so it is inconvenient to computer analysis. In that case, we can convert it to real numbers. In Part III, we examined properties of unitary transformations. Since the unitary transformation keeps a norm of a vector unchanged, this is suited to our present purpose. This can be done using a following unitary matrix $U$:

$$U = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{i}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} \end{pmatrix} \quad (16.141)$$

Then, we have

$$(\Psi_2 \Psi_3) U = \left( \frac{1}{\sqrt{2}}(\Psi_2 + \Psi_3) \frac{i}{\sqrt{2}}(-\Psi_2 + \Psi_3) \right) \quad (16.142)$$

Thus, defining $\widetilde{\Psi_2}$ and $\widetilde{\Psi_3}$ as

$$\widetilde{\Psi_2} = \frac{1}{\sqrt{2}}(\Psi_2 + \Psi_3) \text{ and } \widetilde{\Psi_3} = \frac{i}{\sqrt{2}}(-\Psi_2 + \Psi_3) \quad (16.143)$$

we get

$$\begin{aligned} \widetilde{\Psi_2} &= \frac{1}{\sqrt{6(1-S)}}[2\phi_1 + (\varepsilon + \varepsilon^*)\phi_2 + (\varepsilon^* + \varepsilon)\phi_3] \\ &= \frac{1}{\sqrt{6(1-S)}}(2\phi_1 - \phi_2 - \phi_3) \end{aligned} \quad (16.144)$$

Also, we have

$$\widetilde{\Psi_3} = \frac{i}{\sqrt{6(1-S)}}[(\varepsilon^* - \varepsilon)\phi_2 + (\varepsilon - \varepsilon^*)\phi_3] = \frac{i}{\sqrt{6(1-S)}}\left[-i\sqrt{3}\phi_2 + i\sqrt{3}\phi_3\right]$$
$$= \frac{1}{\sqrt{2(1-S)}}(\phi_2 - \phi_3) \tag{16.145}$$

Thus, we have successfully converted complex functions to real functions. In the above unitary transformation, notice that a norm of the vectors remains unchanged before and after the unitary transformation.

As cyclopropenyl radical has three π electrons, two occupy the lowest energy level of $A''$. Another electron occupies a level $E''$. Since this level possesses an energy higher than $\alpha$, the electron occupying this level is anticipated to be unstable. Under such a circumstance, a molecule tends to lose the said electron so as to be a cation. Following the argument given in the previous case of ethylene, it is easy to make sure that the allowed transition of the cyclopropenyl radical takes place when the light is polarized parallel to the molecular plane (i.e., the *xy*-plane in Fig. 16.8). The proof is left for readers as an exercise. This polarizing feature is typical of planar molecules with high molecular symmetry.

### *16.4.3 Benzene*

Benzene has structural formula which is shown in Fig. 16.9. It is a planar molecule, and six carbon atoms form a regular hexagon. Hence, the molecule belongs to $D_{6h}$ symmetry. In the molecule, six π-electrons extend vertically to the molecular plane toward upper and lower directions as in the case of ethylene and cyclopropenyl radical. This is a standard illustration of quantum chemistry and dealt with in many textbooks. As in the case of ethylene and cyclopropenyl radical, the problem can be treated similarly.

As before, six equivalent $p_z$ atomic orbitals of carbon are denoted by $\phi_1$ to $\phi_6$ in Fig. 16.10. These vectors or their linear combinations span a six-dimensional representation space. We construct basis vectors using these vectors. Following the previous procedures, we construct proper SALC orbitals along with MOs. Similarly

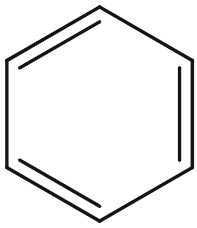

**Fig. 16.9** Structural formula of benzene. It belongs to $D_{6h}$ symmetry

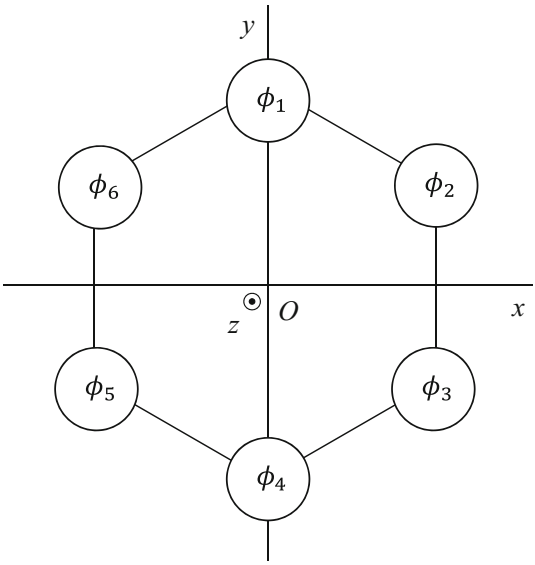


**Fig. 16.10** Six equivalent $p_z$ atomic orbitals of carbon of benzene

as before, a subgroup $C_6$ of $D_{6h}$ plays an essential role. This subgroup contains six group elements such that

$$C_6 = \left\{E, C_6, C_3, C_2, C_3^2, C_6^5\right\} \tag{16.146}$$

Taking $C_6(z)$ as an example, we have

$$(\phi_1\phi_2\phi_3\phi_4\phi_5\phi_6)C_6(z) = (\phi_6\phi_1\phi_2\phi_3\phi_4\phi_5) \tag{16.147}$$

Using a matrix representation, we have

$$C_6(z) = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{16.148}$$

Once again, we can determine the trace for individual symmetry transformations belonging to $D_{6h}$. We collect the results in Table 16.7. The representation is reducible, and this is reduced as follows using a character table of $D_{6h}$ (Table 16.8). As a result, we get

**Table 16.7** Characters for individual symmetry transformations of $2p_z$ orbitals in benzene

| $D_{6h}$ | $E$ | $2C_6$ | $2C_3$ | $C_2$ | $3C_2'$ | $3C_2''$ | $i$ | $2S_3$ | $2S_6$ | $\sigma_h$ | $3\sigma_d$ | $3\sigma_v$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\Gamma$ | 6 | 0 | 0 | 0 | −2 | 0 | 0 | 0 | 0 | −6 | 0 | 2 |

**Table 16.8** Character table of $D_{6h}$

| $D_{6h}$ | $E$ | $2C_6$ | $2C_3$ | $C_2$ | $3C_2'$ | $3C_2''$ | $i$ | $2S_3$ | $2S_6$ | $\sigma_h$ | $3\sigma_d$ | $3\sigma_v$ | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $A_{1g}$ | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | $x^2+y^2, z^2$ |
| $A_{2g}$ | 1 | 1 | 1 | 1 | −1 | −1 | 1 | 1 | 1 | 1 | −1 | −1 | |
| $B_{1g}$ | 1 | −1 | 1 | −1 | 1 | −1 | 1 | −1 | 1 | −1 | 1 | −1 | |
| $B_{2g}$ | 1 | −1 | 1 | −1 | −1 | 1 | 1 | −1 | 1 | −1 | −1 | 1 | |
| $E_{1g}$ | 2 | 1 | −1 | −2 | 0 | 0 | 2 | 1 | −1 | −2 | 0 | 0 | $(yz, zx)$ |
| $E_{2g}$ | 2 | −1 | −1 | 2 | 0 | 0 | 2 | −1 | −1 | 2 | 0 | 0 | $(x^2-y^2, xy)$ |
| $A_{1u}$ | 1 | 1 | 1 | 1 | 1 | 1 | −1 | −1 | −1 | −1 | −1 | −1 | |
| $A_{2u}$ | 1 | 1 | 1 | 1 | −1 | −1 | −1 | −1 | −1 | −1 | 1 | 1 | $z$ |
| $B_{1u}$ | 1 | −1 | 1 | −1 | 1 | −1 | −1 | 1 | −1 | 1 | −1 | 1 | |
| $B_{2u}$ | 1 | −1 | 1 | −1 | −1 | 1 | −1 | 1 | −1 | 1 | 1 | −1 | |
| $E_{1u}$ | 2 | 1 | −1 | −2 | 0 | 0 | −2 | −1 | 1 | 2 | 0 | 0 | $(x, y)$ |
| $E_{2u}$ | 2 | −1 | −1 | 2 | 0 | 0 | −2 | 1 | 1 | −2 | 0 | 0 | |

$$\Gamma = A_{2u} + B_{2g} + E_{1g} + E_{2u} \tag{16.149}$$

As a subduced representation of $D$ of $D_{6h}$ to $C_6$, we have

$$A_{2u} \downarrow C_6 = A,\ B_{2g} \downarrow C_6 = B,\ E_{1g} \downarrow C_6 = 2E_1,\ E_{2u} \downarrow C_6 = 2E_2. \tag{16.150}$$

Here, we used Table 16.9 that shows a character table of irreducible representations of $C_6$. Following the previous procedures, as SALCs we have

$$\begin{aligned}
6P^{(A)}\phi_1 &\equiv \xi_1 = \phi_1 + \phi_2 + \phi_3 + \phi_4 + \phi_5 + \phi_6,\\
6P^{(B)}\phi_1 &\equiv \xi_2 = \phi_1 - \phi_2 + \phi_3 - \phi_4 + \phi_5 - \phi_6,\\
P^{[E_1^{(1)}]}\phi_1 &\equiv \xi_3 = \phi_1 + \varepsilon\phi_2 - \varepsilon^*\phi_3 - \phi_4 - \varepsilon\phi_5 + \varepsilon^*\phi_6,\\
P^{[E_1^{(2)}]}\phi_1 &\equiv \xi_4 = \phi_1 + \varepsilon^*\phi_2 - \varepsilon\phi_3 - \phi_4 - \varepsilon^*\phi_5 + \varepsilon\phi_6,\\
P^{[E_2^{(1)}]}\phi_1 &\equiv \xi_5 = \phi_1 - \varepsilon^*\phi_2 - \varepsilon\phi_3 + \phi_4 - \varepsilon^*\phi_5 - \varepsilon\phi_6,\\
P^{[E_2^{(2)}]}\phi_1 &\equiv \xi_6 = \phi_1 - \varepsilon\phi_2 - \varepsilon^*\phi_3 + \phi_4 - \varepsilon\phi_5 - \varepsilon^*\phi_6,
\end{aligned} \tag{16.151}$$

where $\varepsilon = \exp(i\pi/3)$.

**Table 16.9** Character table of $C_6$

| $C_6$ | $E$ | $C_6$ | $C_3$ | $C_2$ | $C_3^2$ | $C_6^5$ | |
|---|---|---|---|---|---|---|---|
| $A$ | 1 | 1 | 1 | 1 | 1 | 1 | $z; x^2+y^2, z^2$ |
| $B$ | 1 | −1 | 1 | −1 | 1 | −1 | |
| $E_1$ | $\left\{\begin{matrix}1\\1\end{matrix}\right.$ | $\begin{matrix}\varepsilon\\ \varepsilon^*\end{matrix}$ | $\begin{matrix}-\varepsilon^*\\ -\varepsilon\end{matrix}$ | $\begin{matrix}-1\\ -1\end{matrix}$ | $\begin{matrix}-\varepsilon\\ -\varepsilon^*\end{matrix}$ | $\left.\begin{matrix}\varepsilon^*\\ \varepsilon\end{matrix}\right\}$ | $(x, y); (yz, zx)$ |
| $E_2$ | $\left\{\begin{matrix}1\\1\end{matrix}\right.$ | $\begin{matrix}-\varepsilon^*\\ -\varepsilon\end{matrix}$ | $\begin{matrix}-\varepsilon\\ -\varepsilon^*\end{matrix}$ | $\begin{matrix}1\\ 1\end{matrix}$ | $\begin{matrix}-\varepsilon^*\\ -\varepsilon\end{matrix}$ | $\left.\begin{matrix}-\varepsilon\\ -\varepsilon^*\end{matrix}\right\}$ | $(x^2-y^2, xy)$ |

Correspondingly, we have a diagonal secular equation of a sixth order such that

$$\begin{vmatrix} \widetilde{H}_{11}-\lambda\widetilde{S}_{11} & & & & & \\ & \widetilde{H}_{22}-\lambda\widetilde{S}_{22} & & & & \\ & & \widetilde{H}_{33}-\lambda\widetilde{S}_{33} & & & \\ & & & \widetilde{H}_{44}-\lambda\widetilde{S}_{44} & & \\ & & & & \widetilde{H}_{55}-\lambda\widetilde{S}_{55} & \\ & & & & & \widetilde{H}_{66}-\lambda\widetilde{S}_{66} \end{vmatrix} = 0 \tag{16.152}$$

Here, we have for example

$$\widetilde{H}_{11} = \int \xi_1^* H \xi_1 \mathrm{d}\tau = \langle \xi_1 | H \xi_1 \rangle = 6(\alpha + 2\beta + 2\beta' + \beta'') \tag{16.153}$$

In (16.153), we used the same $\alpha$ and $\beta$ defined in (16.94) as in the case of cyclopropenyl radical. That is, $\alpha$ is a Coulomb integral and $\beta$ is a resonance integral between two adjacent $2p_z$ orbitals of carbon. Meanwhile, $\beta'$ is a resonance integral between orbitals of "meta" positions such as $\phi_1$ and $\phi_3$. A quantity $\beta''$ is a resonance integral between orbitals of "para" positions such as $\phi_1$ and $\phi_4$. It is unfamiliar to include such kind of resonance integrals of $\beta'$ and $\beta''$ at a simple $\pi$-electron approximation level. To ignore such resonance integrals is because of a practical purpose to simplify the calculations. However, we have no reason to exclude them. Or rather, the use of appropriate SALCs makes it feasible to include $\beta'$ and $\beta''$.

In a similar manner, we get

$$\begin{aligned} \widetilde{H}_{22} = \int \xi_2^* H \xi_2 \mathrm{d}\tau = \langle \xi_2 | H \xi_2 \rangle &= 6(\alpha - 2\beta + 2\beta' - \beta''), \\ \widetilde{H}_{33} = \widetilde{H}_{44} &= 6(\alpha + \beta - \beta' - \beta''), \\ \widetilde{H}_{55} = \widetilde{H}_{66} &= 6(\alpha - \beta - \beta' + \beta''). \end{aligned} \tag{16.154}$$

Meanwhile, we have

$$\begin{aligned} \widetilde{S}_{11} &= \langle \xi_1 | \xi_1 \rangle = 6(1 + 2S + 2S' + S''), \\ \widetilde{S}_{22} &= \langle \xi_2 | \xi_2 \rangle = 6(1 - 2S + 2S' - S''), \\ \widetilde{S}_{33} &= \widetilde{S}_{44} = 6(1 + S - S' - S''), \\ \widetilde{S}_{55} &= \widetilde{S}_{66} = 6(1 - S - S' + S''), \end{aligned} \tag{16.155}$$

where $S$, $S'$, and $S''$ are overlap integrals between the ortho, meta, and para positions, respectively. Substituting (16.153) through (16.155) for (16.152), the energy eigenvalues are readily obtained as

$$\lambda_1 = \frac{\alpha + 2\beta + 2\beta' + \beta''}{1 + 2S + 2S' + S''}, \lambda_2 = \frac{\alpha - 2\beta + 2\beta' - \beta''}{1 - 2S + 2S' - S''},$$
$$\lambda_3 = \lambda_4 = \frac{\alpha + \beta - \beta' - \beta''}{1 + S - S' - S''}, \lambda_5 = \lambda_6 = \frac{\alpha - \beta - \beta' + \beta''}{1 - S - S' + S''} \tag{16.156}$$

Notice that two MOs $\xi_3$ and $\xi_4$ as well as $\xi_5$ and $\xi_6$ are degenerate. As can be seen, $\lambda_3$ and $\lambda_4$ are doubly degenerate. So are $\lambda_5$ and $\lambda_6$.

From (16.151), we get for instance

$$||\xi_1|| = \sqrt{\langle \xi_1 | \xi_1 \rangle} = \sqrt{6(1 + 2S + 2S' + S'')} \tag{16.157}$$

Thus, for one of normalized MOs corresponding to an energy eigenvalue $\lambda_1$, we get

$$\Psi_1 = \frac{|\xi_1\rangle}{||\xi_1||} = \frac{\phi_1 + \phi_2 + \phi_3 + \phi_4 + \phi_5 + \phi_6}{\sqrt{6(1 + 2S + 2S' + S'')}} \tag{16.158}$$

Following the previous examples, we have other normalized MOs. That is, we have

$$\Psi_2 = \frac{|\xi_2\rangle}{||\xi_2||} = \frac{\phi_1 - \phi_2 + \phi_3 - \phi_4 + \phi_5 - \phi_6}{\sqrt{6(1 - 2S + 2S' - S'')}},$$
$$\Psi_3 = \frac{|\xi_3\rangle}{||\xi_3||} = \frac{\phi_1 + \varepsilon\phi_2 - \varepsilon^*\phi_3 - \phi_4 - \varepsilon\phi_5 + \varepsilon^*\phi_6}{\sqrt{6(1 + S - S' - S'')}},$$
$$\Psi_4 = \frac{|\xi_4\rangle}{||\xi_4||} = \frac{\phi_1 + \varepsilon^*\phi_2 - \varepsilon\phi_3 - \phi_4 - \varepsilon^*\phi_5 + \varepsilon\phi_6}{\sqrt{6(1 + S - S' - S'')}},$$
$$\Psi_5 = \frac{|\xi_5\rangle}{||\xi_5||} = \frac{\phi_1 - \varepsilon^*\phi_2 - \varepsilon\phi_3 + \phi_4 - \varepsilon^*\phi_5 - \varepsilon\phi_6}{\sqrt{6(1 - S - S' + S'')}},$$
$$\Psi_6 = \frac{|\xi_6\rangle}{||\xi_6||} = \frac{\phi_1 - \varepsilon\phi_2 - \varepsilon^*\phi_3 + \phi_4 - \varepsilon\phi_5 - \varepsilon^*\phi_6}{\sqrt{6(1 - S - S' + S'')}} \tag{16.159}$$

The eigenfunction $\Psi_i (1 \le i \le 6)$ corresponds to the eigenvalue $\lambda_i$. As in the case of cyclopropenyl radical, $\Psi_3$ and $\Psi_4$ can be transformed to $\widetilde{\Psi}_3$ and $\widetilde{\Psi}_4$, respectively, through a unitary matrix of (16.141). Thus, we get

$$\widetilde{\Psi}_3 = \frac{2\phi_1 + \phi_2 - \phi_3 - 2\phi_4 - \phi_5 + \phi_6}{\sqrt{12(1 + S - S' - S'')}},$$
$$\widetilde{\Psi}_4 = \frac{\phi_2 + \phi_3 - \phi_5 - \phi_6}{2\sqrt{1 + S - S' - S''}} \tag{16.160}$$

Similarly, transforming $\Psi_5$ and $\Psi_6$ to $\widetilde{\Psi}_5$ and $\widetilde{\Psi}_6$, respectively, we have

$$\begin{aligned}\widetilde{\Psi}_5 &= \frac{2\phi_1 - \phi_2 - \phi_3 + 2\phi_4 - \phi_5 - \phi_6}{\sqrt{12(1 - S - S' + S'')}},\\ \widetilde{\Psi}_6 &= \frac{\phi_2 - \phi_3 + \phi_5 - \phi_6}{2\sqrt{1 - S - S' + S''}}\end{aligned} \tag{16.161}$$

Figure 16.11 shows an energy diagram and MO assignments of benzene.

A major optical transition takes place among HOMO $(E_{1g})$ and LUMO $(E_{2u})$ levels. In the case of optical absorption, an initial electronic configuration is assigned to the totally symmetric representation $A_{1g}$ and the symmetry of the final electronic configuration is described as

$$\Gamma = E_{1g} \times E_{2u}$$

Therefore, a transition matrix element is expressed as

$$\Phi^{(E_{1g}\times E_{2u})} | \boldsymbol{\varepsilon}_e \cdot \boldsymbol{P} | \Phi^{(A_{1g})}, \tag{16.162}$$

where $\Phi^{(A_{1g})}$ stands for the totally symmetric ground state electronic configuration; $\Phi^{(E_{1g}\times E_{2u})}$ denotes an excited-state electronic configuration represented by a direct-product representation. This representation is reducible and expressed as a direct sum of irreducible representations such that

$$\Gamma = B_{1u} + B_{2u} + E_{1u} \tag{16.163}$$

Notice that unlike ethylene a direct-product representation associated with the final state is reducible.

To examine whether (16.162) is nonvanishing, as in the case of (16.109) we estimate whether a direct-product representation $A_{1g} \times E_{1g} \times E_{2u} = E_{1g} \times E_{2u} =$

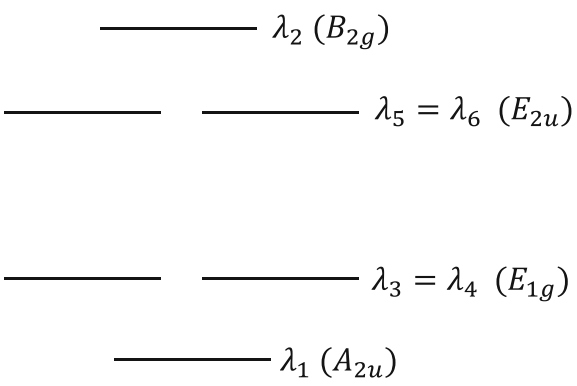


**Fig. 16.11** Energy diagram and MO assignments of benzene. Energy eigenvalues $\lambda_1$ to $\lambda_6$ are given in (16.156)

$B_{1u} + B_{2u} + E_{1u}$ contains an irreducible representation which $\boldsymbol{\varepsilon}_e \cdot \boldsymbol{P}$ belongs to. Consulting a character table of $D_{6h}$, we find that $x$ and $y$ belong to an irreducible representation $E_{1u}$ and that $z$ belongs to $A_{2u}$. Since the direct sum $\Gamma$ in (16.163) contains $E_{1u}$, benzene is expected to be polarized along both $x$- and $y$-axes (see Fig. 16.10). Since (16.163) does not contain $A_{2u}$, the transition along the $z$-axis is forbidden. Accordingly, the transition takes place when the light is polarized parallel to the molecular plane (i.e., the $xy$-plane). This is a common feature among planar aromatic molecules including benzene and cyclopropenyl radical. On the other hand, $A_{2u}$ is not contained in $\Gamma$, and so we do not expect the optical transition to occur in the direction of the $z$-axis.

### 16.4.4 Allyl Radical [1]

We revisit the allyl radical and perform its MO calculations. As already noted, Tables 15.1 and 15.2 of Example 15.1 collected representation matrices of individual symmetry operations in reference to the basis vectors comprising three atomic orbitals of allyl radical. As usual, we examine traces (or characters) of those matrices. Table 16.10 collects them. The representation is readily reduced according to the character table of $C_{2v}$ (see Table 15.4) so that we have

$$\Gamma = A_2 + 2B_1 \tag{16.164}$$

We have two SALC orbitals that belong to the same irreducible representation of $B_1$. As noted in Sect. 16.3, the orbitals obtained by a linear combination of these two SALCs belong to $B_1$ as well. Such a linear combination is given by a unitary transformation. In the present case, it is convenient to transform the basis vectors two times. The first transformation is carried out to get SALCs, and the second one will be done in the process of solving a secular equation using the SALCs. Schematically showing the procedures, we have

$$\begin{Bmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \end{Bmatrix} \rightarrow \begin{Bmatrix} \Psi_1 \\ \Psi_2 \\ \Psi_3 \end{Bmatrix} \rightarrow \begin{Bmatrix} \Phi_1 \\ \Phi_2 \\ \Phi_3 \end{Bmatrix},$$

where $\phi_1, \phi_2, \phi_3$ show the original atomic orbitals; $\Psi_1, \Psi_2, \Psi_3$ the SALCs; $\Phi_1, \Phi_2, \Phi_3$ the final MOs. Thus, the three sets of vectors are connected through unitary transformations.

**Table 16.10** Characters for individual symmetry transformations of $2p_z$ orbitals in allyl radical

| $C_{2v}$ | $E$ | $C_2(z)$ | $\sigma_v(zx)$ | $\sigma_v'(yz)$ |
|---|---|---|---|---|
| $\Gamma$ | 3 | −1 | 1 | −3 |

Starting with $\phi_1$ and following the previous cases, we have, e.g.,

$$P^{(B_1)}\phi_1 = \frac{1}{4}\sum_g [\chi^{(B_1)}(g)]^* g\phi_1 = \phi_1.$$

Also starting with $\phi_2$, we have

$$P^{(B_1)}\phi_2 = \frac{1}{4}\sum_g \left[\chi^{(B_1)}(g)\right]^* g\phi_2 = \frac{1}{2}(\phi_2 + \phi_3)$$

Meanwhile, we get

$$P^{(A_2)}\phi_1 = 0,$$

$$P^{(A_2)}\phi_2 = \frac{1}{4}\sum_g \left[\chi^{(A_2)}(g)\right]^* g\phi_2 = \frac{1}{2}(\phi_2 - \phi_3)$$

Thus, we recovered the results of Example 15.1. Notice that $\phi_1$ does not participate in $A_2$, but take part in $B_1$ by itself.

Normalized SALCs are given as follows:

$$\Psi_1 = \phi_1, \Psi_2 = (\phi_2 + \phi_3)/\sqrt{2(1+S')}, \Psi_3 = (\phi_2 - \phi_3)/\sqrt{2(1-S')},$$

where we define $S' \equiv \int \phi_2\phi_3 \mathrm{d}\tau$. If $\Psi_1$, $\Psi_2$, and $\Psi_3$ belonged to different irreducible representations (as in the cases of previous three examples of ethylene, cyclopropenyl radical, and benzene), the secular equation would be fully reduced to a form of (16.66). In the present case, however, $\Psi_1$ and $\Psi_2$ belong to the same irreducible representation $B_1$, making the situation a bit complicated. Nonetheless, we can use (16.61) and the secular equation is "partially" reduced.

Defining

$$H_{jk} = \int \Psi_j^* H \Psi_k \mathrm{d}\tau \text{ and } S_{jk} = \int \Psi_j^* \Psi_k \mathrm{d}\tau,$$

we have a following secular equation, the same form as (16.61):

$$\det\left(H_{jk} - \lambda S_{jk}\right) = 0.$$

More specifically, we have

$$\begin{vmatrix} \alpha - \lambda & \sqrt{\frac{2}{1+S'}}(\beta - S\lambda) & 0 \\ \sqrt{\frac{2}{1+S'}}(\beta - S\lambda) & \frac{\alpha+\beta'}{1+S'} - \lambda & 0 \\ 0 & 0 & \frac{\alpha-\beta'}{1-S'} - \lambda \end{vmatrix} = 0,$$

where $\alpha$, $\beta$, and $S$ are similarly defined as (16.94) and (16.98). The quantities $\beta'$ are defined as

$$\beta' \equiv \int \phi_2^* H \phi_3 \mathrm{d}\tau$$

Thus, the secular equation is separated into the following two:

$$\begin{vmatrix} \alpha - \lambda & \sqrt{\frac{2}{1+S'}}(\beta - S\lambda) \\ \sqrt{\frac{2}{1+S'}}(\beta - S\lambda) & \frac{\alpha+\beta'}{1+S'} - \lambda \end{vmatrix} = 0 \text{ and } \frac{\alpha - \beta'}{1 - S'} - \lambda = 0 \tag{16.165}$$

The second equation immediately gives

$$\lambda = \frac{\alpha - \beta'}{1 - S'}$$

The first equation of (16.165) is somewhat complicated, and so we adopt the next approximation. That is,

$$S' = \beta' = 0 \tag{16.166}$$

This approximation is justified, because two carbon atoms $C_2$ and $C_3$ are pretty remote, and so the interaction between them is likely to be weak enough. Thus, we rewrite the first equation of (16.165) as

$$\begin{vmatrix} \alpha - \lambda & \sqrt{2}(\beta - S\lambda) \\ \sqrt{2}(\beta - S\lambda) & \alpha - \lambda \end{vmatrix} = 0 \tag{16.167}$$

Moreover, we assume that since $S$ is a small quantity compared to 1, a square term of $S^2 \ll 1$. Hence, we ignore $S^2$.

Using the approximation of (16.166) and assuming $S^2 \approx 0$, from (16.167) we have a following quadratic equation:

$$\lambda^2 - 2(\alpha - 2\beta S)\lambda + \alpha^2 - 2\beta^2 = 0$$

Solving this equation, we have

$$\lambda = \alpha - 2\beta S \pm \sqrt{2}\beta\sqrt{1 - \frac{2\alpha S}{\beta}} \approx \alpha - 2\beta S \pm \sqrt{2}\beta\left(1 - \frac{\alpha S}{\beta}\right),$$

where the last approximation is based on

$$\sqrt{1 - x} \approx 1 - \frac{1}{2}x$$

with a small quantity $x$. Thus, we get

$$\lambda_L \approx \left(\alpha+\sqrt{2}\beta\right)\left(1-\sqrt{2}S\right) \text{ and } \lambda_H \approx \left(\alpha-\sqrt{2}\beta\right)\left(1+\sqrt{2}S\right), \tag{16.168}$$

where $\lambda_L < \lambda_H$. To determine the corresponding eigenfunctions $\Phi$ (i.e., MOs), we use a linear combination of two SALCs $\Psi_1$ and $\Psi_2$. That is, putting

$$\Phi = c_1\Psi_1 + c_2\Psi_2 \tag{16.169}$$

and from (16.167), we obtain

$$(\alpha-\lambda)c_1 + \sqrt{2}(\beta - S\lambda)c_2 = 0.$$

Thus, with $\lambda = \lambda_L$, we get $c_1 = c_2$, and for $\lambda = \lambda_H$, we get $c_1 = -c_2$. Consequently, as a normalized eigenfunction $\Phi_L$ corresponding to $\lambda_L$, we get

$$\Phi_L = \frac{1}{\sqrt{2}}(\Psi_1+\Psi_2) = \frac{1}{2}(1+\sqrt{2}S)^{-1/2}\left(\sqrt{2}\phi_1+\phi_2+\phi_3\right) \tag{16.170}$$

Likewise, as a normalized eigenfunction $\Phi_H$ corresponding to $\lambda_H$, we get

$$\Phi_H = \frac{1}{\sqrt{2}}(-\Psi_1+\Psi_2) = \frac{1}{2}(1-\sqrt{2}S)^{-1/2}\left(-\sqrt{2}\phi_1+\phi_2+\phi_3\right) \tag{16.171}$$

As another eigenvalue $\lambda_0$ and corresponding eigenfunction $\Phi_0$, we have

$$\lambda_0 \approx \alpha \text{ and } \Phi_0 = \frac{1}{\sqrt{2}}(\phi_2-\phi_3), \tag{16.172}$$

where we have $\lambda_L < \lambda_0 < \lambda_H$. The eigenfunction $\Phi_0$ does not participate in chemical bonding and, hence, is said to be a non-bonding orbital.

It is worth noting that eigenfunctions $\Phi_L$, $\Phi_0$, and $\Phi_H$ have the same function forms as those obtained by the simple Hückel theory that ignores the overlap integrals $S$. It is because the interaction between $\phi_2$ and $\phi_3$ that construct SALCs of $\Psi_2$ and $\Psi_3$ is weak.

Notice that two sets of basis vectors $|\phi_1\rangle, \frac{1}{\sqrt{2}}|\phi_2+\phi_3\rangle$, and $\frac{1}{\sqrt{2}}|\phi_2-\phi_3\rangle$ and $|\Phi_L\rangle$, $|\Phi_H\rangle$, and $|\Phi_0\rangle$ are connected by a following unitary matrix $V$:

$$\begin{aligned}(|\Phi_L\rangle|\Phi_H\rangle|\Phi_0\rangle) &= \left(|\phi_1\rangle\frac{1}{\sqrt{2}}|\phi_2+\phi_3\rangle\frac{1}{\sqrt{2}}|\phi_2-\phi_3\rangle\right)V \\ &= \left(|\phi_1\rangle\frac{1}{\sqrt{2}}|\phi_2+\phi_3\rangle\frac{1}{\sqrt{2}}|\phi_2-\phi_3\rangle\right)\begin{pmatrix}\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0\\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0\\ 0 & 0 & 1\end{pmatrix}\end{aligned} \tag{16.173}$$

Although both SALCs $\Psi_1$ and $\Psi_2$ belong to the same irreducible representation $B_1$, they are not orthogonal. As shown in (16.170) and (16.171), however, we find that $\Phi_L$ and $\Phi_H$ that are sought by solving the secular equation (16.167) have been mutually orthogonal. Starting from $|\phi_1\rangle$, $|\phi_2\rangle$, and $|\phi_3\rangle$ of Example 15.1, we reached $|\Phi_L\rangle$, $|\Phi_H\rangle$, and $|\Phi_0\rangle$ via two-step unitary transformations (15.40) and (16.173). The combined unitary transformations $W = UV$ are unitary again. That is, we have

$$\begin{aligned}(|\phi_1\rangle|\phi_2\rangle|\phi_3\rangle)UV &= (|\phi_1\rangle|\phi_2\rangle|\phi_3\rangle)\begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{\sqrt{2}} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{\sqrt{2}} \end{pmatrix} \\ &= (|\Phi_L\rangle|\Phi_H\rangle|\Phi_0\rangle)\end{aligned} \tag{16.174}$$

The optical transition of allyl radical represents general features of the optical transitions of molecules. To make a story simple, let us consider a case of allyl cation. Figure 16.12 shows electronic configurations together with symmetry species of individual eigenstates and their corresponding energy eigenvalues for the allyl cation. In Fig. 16.13, we redraw its geometry where the origin is located at the center of a line segment connecting $C_2$ and $C_3$ ($\boldsymbol{r}_2 + \boldsymbol{r}_3 = 0$). Major optical transitions (or optical absorption) are $\Phi_L \to \Phi_0$ and $\Phi_L \to \Phi_H$.

(i) $\Phi_L \to \Phi_0$: In this case, following (16.109) the transition matrix element $T_{fi}$ is described by

$$T_{fi} = \left\langle \Theta_f^{(B_1 \times A_2)} | \boldsymbol{\varepsilon}_e \cdot \boldsymbol{P} | \Theta_i^{(A_1)} \right\rangle \tag{16.175}$$

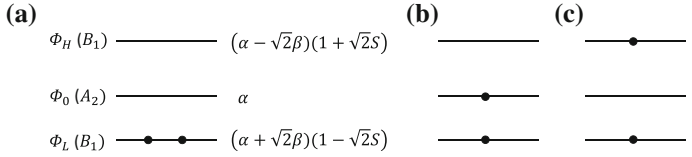


**Fig. 16.12** Electronic configurations and symmetry species of individual eigenstates along with their corresponding energy eigenvalues for the allyl cation. **a** Ground state. **b** First excited state. **c** Second excited state

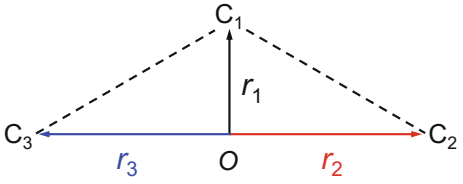


**Fig. 16.13** Geometry and position vectors of carbon atoms of the allyl cation. The origin is located at the center of a line segment connecting $C_2$ and $C_3$ ($\boldsymbol{r}_2 + \boldsymbol{r}_3 = 0$)

In the above equation, we designate the irreducible representation of eigenstates according to (16.164). Therefore, the symmetry of the final electronic configuration is described as

$$\Gamma = B_1 \times A_2 = B_2$$

Since a direct product of the initial electronic configuration $\left[\Theta_i^{(A_1)}\right]$ and the final configuration $\Theta_f^{(B_1 \times A_2)}$ is $B_1 \times A_2 \times A_1 = B_2$. Then, if $\boldsymbol{\varepsilon}_e \cdot \boldsymbol{P}$ belongs to the same irreducible representation $B_2$, the associated optical transition should be allowed. Consulting a character table of $C_{2v}$, we find that $y$ belongs to $B_2$. Thus, the allowed optical transition is polarized along the $y$-direction.

(ii) $\Phi_L \rightarrow \Phi_H$: In parallel with the above case, $T_{fi}$ is described by

$$T_{fi} = \left\langle \Theta_f^{(B_1 \times B_1)} | \boldsymbol{\varepsilon}_e \cdot \boldsymbol{P} | \Theta_i^{(A_1)} \right\rangle \tag{16.176}$$

Thus, the transition is characterized by

$$A_1 \rightarrow A_1,$$

where the former $A_1$ indicates the electronic ground state and the latter $A_1$ indicates the excited state given by $B_1 \times B_1 = A_1$. The direct product of them is simply described as $B_1 \times B_1 \times A_1 = A_1$. Consulting a character table again, we find that $z$ belongs to $A_1$. This implies that the allowed transition is polarized along the $z$-direction.

Next, we investigate the above optical transition in a semiquantitative manner. The transition matrix element $T_{fi}$ should be estimated from (16.175) and (16.176) that use electronic configurations of two-electron system. Nonetheless, a formulation of (4.7) of Sect. 4.1 based upon one-electron states well serves our present purpose. For $\Phi_L \rightarrow \Phi_0$ transition, we have

$$\begin{aligned}
T_{fi}(\Phi_L \rightarrow \Phi_0) &= \int \Phi_0^* \boldsymbol{\varepsilon}_e \cdot \boldsymbol{P} \Phi_L \mathrm{d}\tau = e\boldsymbol{\varepsilon}_e \cdot \int \Phi_0^* \boldsymbol{r} \Phi_L \mathrm{d}\tau \\
&= e\boldsymbol{\varepsilon}_e \cdot \int \frac{1}{2}\left(\sqrt{2}\phi_1 + \phi_2 + \phi_3\right)\boldsymbol{r}\frac{1}{\sqrt{2}}(\phi_2 - \phi_3)\mathrm{d}\tau \\
&= \frac{e\boldsymbol{\varepsilon}_e}{2\sqrt{2}} \cdot \int \left(\sqrt{2}\phi_1 \boldsymbol{r}\phi_2 - \sqrt{2}\phi_1 \boldsymbol{r}\phi_3 + \phi_2 \boldsymbol{r}\phi_2 + \phi_3 \boldsymbol{r}\phi_2 - \phi_3 \boldsymbol{r}\phi_3\right)\mathrm{d}\tau \\
&\approx \frac{e\boldsymbol{\varepsilon}_e}{2\sqrt{2}} \cdot \int (\phi_2 \boldsymbol{r}\phi_2 - \phi_3 \boldsymbol{r}\phi_3)\mathrm{d}\tau \\
&\approx \frac{e\boldsymbol{\varepsilon}_e}{2\sqrt{2}} \cdot \left[\boldsymbol{r}_2 \int |\phi_2|^2 d\tau - \boldsymbol{r}_3 \int |\phi_3|^2 \mathrm{d}\tau\right] = \frac{e\boldsymbol{\varepsilon}_e}{2\sqrt{2}} \cdot (\boldsymbol{r}_2 - \boldsymbol{r}_3) = \frac{e\boldsymbol{\varepsilon}_e}{\sqrt{2}} \cdot \boldsymbol{r}_2,
\end{aligned}$$

where with the first near equality we ignored integrals $\int \phi_i \boldsymbol{r} \phi_j \mathrm{d}\tau (i \neq j)$ and with the last near equality $\boldsymbol{r} \approx \boldsymbol{r}_2$ or $\boldsymbol{r}_3$. For these approximations, we assumed that an electron density is very high near $C_2$ or $C_3$ with ignorable density at a place remote from them. Choosing $\boldsymbol{\varepsilon}_e$ for the direction of $\boldsymbol{r}_2$, we get

$$T_{fi}(\Phi_L \rightarrow \Phi_0) \approx \frac{e}{\sqrt{2}} |\boldsymbol{r}_2|$$

With $\Phi_L \rightarrow \Phi_H$ transition, similarly we have

$$T_{fi}(\Phi_L \rightarrow \Phi_H) = \int \Phi_H^* \boldsymbol{\varepsilon}_e \cdot \boldsymbol{P} \Phi_L \mathrm{d}\tau \approx \frac{e\boldsymbol{\varepsilon}_e}{4} \cdot (\boldsymbol{r}_3 - 2\boldsymbol{r}_1 + \boldsymbol{r}_2) = -\frac{e\boldsymbol{\varepsilon}_e}{2} \cdot \boldsymbol{r}_1$$

Choosing $\boldsymbol{\varepsilon}_e$ for the direction of $\boldsymbol{r}_1$, we get

$$T_{fi}(\Phi_L \rightarrow \Phi_H) \approx \frac{e}{2} |\boldsymbol{r}_1|$$

Transition probability is proportional to a square of an absolute value of $T_{fi}$. Using $|\boldsymbol{r}_2| \approx \sqrt{3} |\boldsymbol{r}_1|$, we have

$$\left| T_{fi}(\Phi_L \rightarrow \Phi_0) \right|^2 \approx \frac{e^2}{2} \left| \boldsymbol{r}_2 \right|^2 = \frac{3e^2}{2} |\boldsymbol{r}_1|^2,$$

$$\left| T_{fi}(\Phi_L \rightarrow \Phi_H) \right|^2 \approx \frac{e^2}{4} \left| \boldsymbol{r}_1 \right|^2$$

Thus, we obtain

$$\left| T_{fi}(\Phi_L \rightarrow \Phi_0) \right|^2 \approx 6 \left| T_{fi}(\Phi_L \rightarrow \Phi_H) \right|^2 \tag{16.177}$$

Thus, the transition probability of $\Phi_L \rightarrow \Phi_0$ is about six times that for $\Phi_L \rightarrow \Phi_H$. Note that in the above simple estimation we ignore an overlap integral $S$.

From the above discussion, we conclude that (i) the $\Phi_L \rightarrow \Phi_0$ transition is polarized along the $\boldsymbol{r}_2$ direction (i.e., the molecular long axis) and that the $\Phi_L \rightarrow \Phi_H$ transition is polarized along the $\boldsymbol{r}_1$ direction (i.e., the molecular short axis). (ii) Transition probability of $\Phi_L \rightarrow \Phi_0$ is about six times that of $\Phi_L \rightarrow \Phi_H$. Note that the polarized characteristics are consistent with those obtained from the discussion based on the group theory. The conclusion reached by the semiquantitative estimation of a simple molecule of allyl cation well typifies the general optical features of more complicated molecules having a well-defined molecular long axis such as polyenes.

## 16.5 MO Calculations of Methane

So far, we investigated MO calculations of aromatic molecules based upon $\pi$-electron approximation. These are a homogeneous system that has the same quality of electrons. Here, we deal with methane that includes a carbon and surrounding four hydrogens. These hydrogen atoms form a regular tetrahedron with the carbon atom positioned at a center of the tetrahedron. It is therefore considered as a heterogeneous system. The calculation principle, however, is consistent; namely, we make the most of projection operators and construct appropriate SALCs of methane.

We deal with four 1s electrons of hydrogen along with two 2s electrons and two $2p$ electrons of carbon. Regarding basis functions of carbon, however, we consider $2s$ atomic orbital and three $2p$ orbitals (i.e., $2p_x, 2p_y, 2p_z$ orbitals). That is, we deal with eight atomic orbitals all together. These are depicted in Fig. 16.14. The dimension of the vector space (i.e., representation space) is eight accordingly.

As before, we wish to determine irreducible representations which individual MOs belong to. As already mentioned in Sect. 14.3, there are 24 symmetry operations in a point group $T_d$ which methane belongs to (see Table 14.6). According to the symmetry operations, we decide transformation matrices related to each operation. For example, $C_3^{xyz}$ transforms basis functions as follows:

$$\begin{aligned}&\left(H_1H_2H_3H_4C2sC2p_xC2p_yC2p_z\right)C_3^{xyz}\\&=\left(H_1H_3H_4H_2C2sC2p_yC2p_zC2p_x\right),\end{aligned}$$

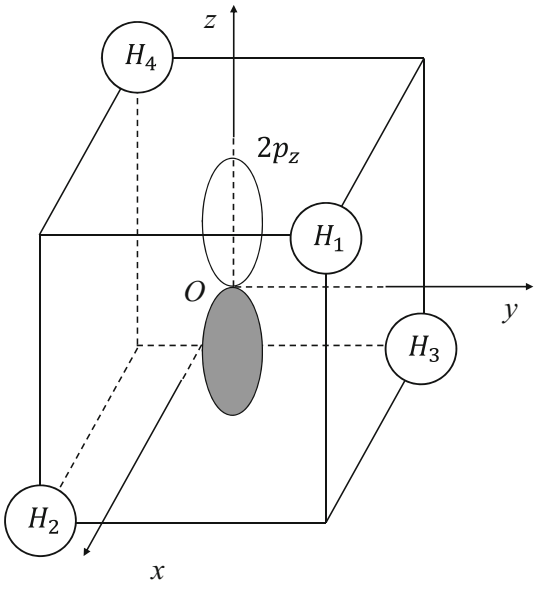


**Fig. 16.14** Four 1s atomic orbitals of hydrogen and a $2p_z$ orbital of carbon. The former orbitals are represented by $H_1$ to $H_4$. $2p_x$ and $2p_y$ orbitals of carbon are omitted for simplicity

where by the above notations we denoted atomic species and molecular orbitals. Hence, as a matrix representation, we have

$$C_3^{xyz} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \tag{16.178}$$

where $C_3^{xyz}$ is the same operation as $R_{xyz\frac{2\pi}{3}}$ appeared in Sect. 14.3. Therefore,

$$\chi\left(C_3^{xyz}\right) = 2.$$

As another example $\sigma_d^{yz}$, we have

$$\begin{aligned} &\left(H_1H_2H_3H_4C2sC2p_xC2p_yC2p_z\right)\sigma_d^{yz} \\ &= \left(H_1H_2H_4H_3C2sC2p_xC2p_zC2p_y\right), \end{aligned}$$

where $\sigma_d^{yz}$ represents a mirror symmetry with respect to the plane that includes the $x$-axis and bisects the angle formed by the $y$- and $z$-axes. Thus, we have

$$\sigma_d^{yz} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \tag{16.179}$$

Then, we have

$$\chi\left(\sigma_d^{yz}\right) = 4.$$

Taking some more examples, for $C_2^z$ we get

$$\begin{aligned} &\left(H_1H_2H_3H_4C2sC2p_xC2p_yC2p_z\right)C_2^z \\ &= \left(H_4H_3H_2H_1C2s - C2p_x - C2p_yC2p_z\right), \end{aligned}$$

where $C_2^z$ means a rotation by $\pi$ around the $z$-axis. Also, we have

$$C_2^z = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{16.180}$$

$$\chi\left(C_2^z\right) = 0.$$

With $S_4^{z\frac{\pi}{2}}$ (i.e., an improper rotation by $\frac{\pi}{2}$ around the $z$-axis), we get

$$\begin{aligned} &\left(H_1H_2H_3H_4C2sC2p_xC2p_yC2p_z\right)S_4^{z\frac{\pi}{2}} \\ &= \left(H_3H_1H_4H_2C2sC2p_y - C2p_x - C2p_z\right), \end{aligned}$$

$$S_4^{z\frac{\pi}{2}} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix} \tag{16.181}$$

$$\chi\left(S_4^{z\frac{\pi}{2}}\right) = 0$$

As for the identity matrix $E$, we have

$$\chi(E) = 8$$

Thus, Table 16.11 collects characters of individual symmetry transformations with respect to hydrogen $1s$ and carbon $2s$ and $2p$ orbitals in methane. From the above examples, we notice that all the symmetry operators $R$ reduce an eight-dimensional representation space $V^8$ to subspaces of $\text{Span}\{H_1H_2H_3H_4\}$ and $\text{Span}\left\{C2sC2p_xC2p_yC2p_z\right\}$. In terms of a notation of Sect. 10.2, we have

$$V^8 = \text{Span}\{H_1H_2H_3H_4\} \oplus \text{Span}\left\{C2sC2p_xC2p_yC2p_z\right\} \tag{16.182}$$

In other words, $V^8$ is decomposed into the above two $R$-invariant subspaces (see Part III): One is the hydrogen-related subspace, and the other is the carbon-related subspace.

**Table 16.11** Characters for individual symmetry transformations of hydrogen 1*s* and carbon 2*p* orbitals in methane

| $T_d$ | $E$ | $8C_3$ | $3C_2$ | $6S_4$ | $6\sigma_d$ |
|---|---|---|---|---|---|
| $\Gamma$ | 8 | 2 | 0 | 0 | 4 |

Correspondingly, a representation $D$ comprising the above representation matrices should be reduced to a direct sum of irreducible representations $D^{(\alpha)}$. That is, we should have

$$D = \sum_{\alpha} q_{\alpha} D^{(\alpha)},$$

where $q_\alpha$ is a positive integer or zero. Here, we are thinking of decomposition of (8,8) matrices such as (16.178) into submatrices. We estimate $q_\alpha$ using (15.83). That is,

$$q_{\alpha} = \frac{1}{n} \sum_{g} \chi^{(\alpha)}(g)^{*} \chi(g). \tag{15.83}$$

With an irreducible representation $A_1$ of $T_d$ for instance, we have

$$q_{A_1} = \frac{1}{24} \sum_{g} \chi^{(A_1)}(g)^{*} \chi(g) = \frac{1}{24} (1 \cdot 8 + 1 \cdot 8 \cdot 2 + 1 \cdot 6 \cdot 4) = 2$$

As for $A_2$, we have

$$q_{A_2} = \frac{1}{24} [1 \cdot 8 + 1 \cdot 8 \cdot 2 + (-1) \cdot 6 \cdot 4] = 0$$

Regarding $T_2$, we get

$$q_{T_2} = \frac{1}{24} (3 \cdot 8 + 1 \cdot 6 \cdot 4) = 2$$

For other irreducible representations of $T_d$, we get $q_\alpha = 0$. Consequently, we have

$$D = 2D^{(A_1)} + 2D^{(T_2)} \tag{16.183}$$

Evidently, both for the hydrogen-related representation $D^{(H)}$ and for the carbon-related representation $D^{(C)}$, we have individually

$$D^{(H)} = D^{(A_1)} + D^{(T_2)} \text{ and } D^{(C)} = D^{(A_1)} + D^{(T_2)}, \tag{16.184}$$

where $D = D^{(H)} + D^{(C)}$.

In fact, (16.180), for example, in a subspace Span$\{H_1H_2H_3H_4\}$, $C_2^z$ is expressed as

$$C_2^z = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

Following routine procedures based on a characteristic polynomials, we get eigenvalues $+1$ (as a double root) and $-1$ (a double root as well). Unitary similarity transformation using a unitary matrix $P$ expressed as

$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

yields a diagonal matrix such that

$$P^{-1}C_2^zP = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Note that this diagonal matrix is identical to submatrix of (16.180) for Span$\{C2sC2p_xC2p_yC2p_z\}$. Similarly, $S_4^{z\frac{\pi}{2}}$ of (16.181) gives eigenvalues $\pm 1$ and $\pm i$ for both the subspaces. Notice that $S_4^{z\frac{\pi}{2}}$ is unitary, and its eigenvalues take a complex number with an absolute value of 1. Since these symmetry operation matrices are unitary, these matrices must be diagonalized according to Theorem 12.5. Using unitary matrices whose column vectors are chosen from eigenvectors of the matrices, diagonal elements are identical to eigenvalues including their multiplicity. Writing representation matrices of symmetry operations for the hydrogen-associated subspace and carbon-associated subspace as $H$ and $C$, we find that $H$ and $C$ have the same eigenvalues in common. Notice here that different types of transformation matrices (e.g., $C_2^z$, $S_4^{z\frac{\pi}{2}}$) give a different set of eigenvalues. Via unitary similarity transformation using unitary matrices $P$ and $Q$, we get

$$P^{-1}HP = Q^{-1}CQ \quad \text{or} \quad (PQ^{-1})^{-1}H(PQ^{-1}) = C.$$

Namely, $H$ and $C$ are similar; i.e., the representation is equivalent. Thus, recalling Schur's First Lemma, Eq. (16.184) results.

Our next task is to construct SALCs, i.e., proper basis vectors using projection operators. From the above, we anticipate that the MOs comprise a linear combination of the hydrogen-associated SALC and carbon-associated SALC that belongs

to the same irreducible representation (i.e., $A_1$ or $T_2$). For this purpose, we first find proper SALCs using projection operators described as

$$P^{(\alpha)}_{l(l)} = \frac{d_\alpha}{n}\sum_g D^{(\alpha)}_{ll}(g)^* g. \tag{15.156}$$

We apply this operator to Span$\{H_1H_2H_3H_4\}$. Taking, e.g., $H_1$ and operating both sides of (15.156) on $H_1$, as a basis vector corresponding to a one-dimensional representation of $A_1$, we have

$$\begin{aligned}
P^{(A_1)}_{1(1)}H_1 &= \frac{d_{A_1}}{n}\sum_g D^{(A_1)}_{11}(g)^* g H_1 \\
&= \frac{1}{24}[(1\cdot H_1) + (1\cdot H_1 + 1\cdot H_1 + 1\cdot H_3 \\
&\quad + 1\cdot H_4 + 1\cdot H_3 + 1\cdot H_2 + 1\cdot H_4 + 1\cdot H_2) + (1\cdot H_4 + 1\cdot H_3 + 1\cdot H_2) \\
&\quad + (1\cdot H_3 + 1\cdot H_2 + 1\cdot H_2 + 1\cdot H_4 + 1\cdot H_4 + 1\cdot H_3) \\
&\quad + (1\cdot H_1 + 1\cdot H_4 + 1\cdot H_1 + 1\cdot H_2 + 1\cdot H_1 + 1\cdot H_3)] \\
&= \frac{1}{24}(6\cdot H_1 + 6\cdot H_2 + 6\cdot H_3 + 6\cdot H_4) \\
&= \frac{1}{4}(H_1 + H_2 + H_3 + H_4)
\end{aligned} \tag{16.185}$$

The case of $C2s$ is simple, because all the symmetry operations convert $C2s$ to itself. That is, we have

$$P^{(A_1)}_{1(1)}C2s = \frac{1}{24}(24\cdot C2s) = C2s \tag{16.186}$$

Regarding $C2p$, taking $C2p_x$ for instance, we have

$$\begin{aligned}
P^{(A_1)}_{1(1)}C2p_x &= \frac{1}{24}\{(1\cdot C2p_x) + [1\cdot C2p_y + 1\cdot C2p_z + 1\cdot(-C2p_y) \\
&\quad + 1\cdot(-C2p_z) + 1\cdot(-C2p_y) + 1\cdot C2p_z + 1\cdot C2p_y + 1\cdot(-C2p_z)] \\
&\quad + [(1\cdot C2p_x + 1\cdot(-C2p_x) + 1\cdot(-C2p_x)] + [1\cdot(-C2p_x) \\
&\quad + 1\cdot(-C2p_x) + 1\cdot(-C2p_z) + 1\cdot C2p_z + 1\cdot C2p_y + 1\cdot(-C2p_y)] \\
&\quad + (1\cdot C2p_y + 1\cdot C2p_x + 1\cdot C2p_z + 1\cdot(-C2p_y) + 1\cdot C2p_x + 1\cdot(-C2p_z)\} \\
&= 0.
\end{aligned}$$

The calculation is somewhat tedious, but it is natural that since $C2s$ is spherically symmetric, it belongs to the totally symmetric representation. Conversely, it is natural to think that $C2p_x$ is totally unlikely to contain a totally symmetric representation. This is also the case with $C2p_y$ and $C2p_z$.

Table 16.12 shows the character table of $T_d$ in which the three-dimensional irreducible representation $T_2$ is spanned by basis vectors $(xyz)$. Since in Table 14.6 each (3,3) matrix is given in reference to the vectors $(xyz)$, it can directly be utilized to represent $T_2$. More specifically, we can directly choose the diagonal elements (1,1), (2,2), and (3,3) of the individual (3,3) matrices for $D_{11}^{(T_2)}$, $D_{22}^{(T_2)}$, and $D_{33}^{(T_2)}$ elements of the projection operators, respectively. Thus, we can construct SALCs using projection operators explained in Sect. 15.7. For example, using $H_1$ we obtain SALCs that belong to $T_2$ such that

$$\begin{aligned} P_{1(1)}^{(T_2)} H_1 &= \frac{3}{24} \sum_g D_{11}^{(T_2)}(g)^* g H_1 = \frac{3}{24} \{(1 \cdot H_1) \\ &\quad + [(-1) \cdot H_4 + (-1) \cdot H_3 + 1 \cdot H_2] \\ &\quad + [0 \cdot (-H_3) + 0 \cdot (-H_2) + 0 \cdot (-H_2) + 0 \cdot (-H_4) + (-1) \cdot H_4 + (-1) \cdot H_3] \\ &\quad + (0 \cdot H_1 + 0 \cdot H_4 + 1 \cdot H_1 + 1 \cdot H_2 + 0 \cdot H_1 + 0 \cdot H_3)\} \\ &= \frac{3}{24} [2 \cdot H_1 + 2 \cdot H_2 + (-2) \cdot H_3 + (-2) \cdot H_4] \\ &= \frac{1}{4} (H_1 + H_2 - H_3 - H_4) \end{aligned} \tag{16.187}$$

Similarly, we have

$$P_{2(2)}^{(T_2)} H_1 = \frac{1}{4} (H_1 - H_2 + H_3 - H_4), \tag{16.188}$$

$$P_{3(3)}^{(T_2)} H_1 = \frac{1}{4} (H_1 - H_2 - H_3 + H_4) \tag{16.189}$$

Now, we can easily guess that $P_{1(1)}^{(T_2)} C2p_x$ solely contains $C2p_x$. Likewise, $P_{2(2)}^{(T_2)} C2p_y$ and $P_{3(3)}^{(T_2)} C2p_z$ contain only $C2p_y$ and $C2p_z$, respectively. In fact, we obtain what we anticipate. That is,

$$P_{1(1)}^{(T_2)} C2p_x = C2p_x,\ P_{2(2)}^{(T_2)} C2p_y = C2p_y,\ P_{3(3)}^{(T_2)} C2p_z = C2p_z. \tag{16.190}$$

**Table 16.12** Character table of $T_d$

| $T_d$ | $E$ | $8C_3$ | $3C_2$ | $6S_4$ | $6\sigma_d$ | |
|---|---|---|---|---|---|---|
| $A_1$ | 1 | 1 | 1 | 1 | 1 | $x^2 + y^2 + z^2$ |
| $A_2$ | 1 | 1 | 1 | −1 | −1 | |
| $E$ | 2 | −1 | 2 | 0 | 0 | $(2z^2 - x^2 - y^2, x^2 - y^2)$ |
| $T_1$ | 3 | 0 | −1 | 1 | −1 | |
| $T_2$ | 3 | 0 | −1 | −1 | 1 | $(x, y, z); (xy, yz, zx)$ |

This gives a good example to illustrate the general concept of projection operators and related calculations of inner products discussed in Sect. 15.7. For example, we have a nonvanishing inner product of $\langle P^{(T_2)}_{1(1)} H_1 | P^{(T_2)}_{1(1)} C2p_x \rangle$ and an inner product of, e.g., $\langle P^{(T_2)}_{2(2)} H_1 | P^{(T_2)}_{1(1)} C2p_x \rangle$ is zero. This significantly reduces efforts to solve a secular equation (vide infra). Notice that functions $P^{(T_2)}_{1(1)} H_1$, $P^{(T_2)}_{1(1)} C2p_x$, etc., are linearly independent of one another. Recall also that $P^{(T_2)}_{1(1)} H_1$ and $P^{(T_2)}_{1(1)} C2p_x$ correspond to $\phi^{(\alpha)}_l$ and $\psi^{(\alpha)}_l$ in (15.171), respectively. That is, the former functions are linearly independent, while they belong to the same place "1" of the same three-dimensional irreducible representation $T_2$.

Equations (16.187)–(16.190) seem intuitively obvious. This is because if we draw a molecular geometry of methane (see Fig. 16.14), we can immediately recognize the relationship between the directionality in $\mathbb{R}^3$ and "directionality" of SALCs represented by sings of hydrogen atomic orbitals (or $C2p_x$, $C2p_y$, and $C2p_z$ of carbon).

As stated above, we have successfully obtained SALCs relevant to methane. Therefore, our next task is to solve an eigenvalue problem and construct appropriate MOs. To do this, let us first normalize the SALCs. We assume that carbon-based SALCs have already been normalized as well-studied atomic orbitals. We suppose that all the functions are real. For the hydrogen-based SALCs

$$\begin{aligned} \langle H_1 + H_2 + H_3 + H_4 | H_1 + H_2 + H_3 + H_4 \rangle &= 4\langle H_1 | H_1 \rangle + 12 \langle H_1 | H_2 \rangle \\ &= 4 + 12 S^{HH} \\ &= 4\left(1 + 3S^{HH}\right), \end{aligned} \tag{16.191}$$

where the second equality comes from the fact that $|H_1\rangle$, i.e., a 1 s atomic orbital is normalized. We define $\langle H_1 | H_2 \rangle = S^{HH}$, i.e., an overlap integral between two adjacent hydrogen atoms. Note also that $\langle H_i | H_j \rangle (1 \leq i, j \leq 4; i \neq j)$ is the same because of the symmetry requirement.

Thus, as a normalized SALC, we have

$$H^{(A_1)} \equiv \frac{|H_1 + H_2 + H_3 + H_4\rangle}{2\sqrt{1 + 3S^{HH}}} \tag{16.192}$$

Defining a denominator as $c\left(= 2\sqrt{1 + 3S^{HH}}\right)$, we have

$$H^{(A_1)} = |H_1 + H_2 + H_3 + H_4\rangle / c \tag{16.193}$$

Also, we have

$$\begin{aligned}\langle H_1+H_2-H_3-H_4|H_1+H_2-H_3-H_4\rangle &= 4\langle H_1|H_1\rangle - 4\langle H_1|H_2\rangle \\ &= 4-4S^{HH} = 4\left(1-S^{HH}\right)\end{aligned} \tag{16.194}$$

Thus, we have

$$H_1^{(T_2)} \equiv \frac{|H_1+H_2-H_3-H_4\rangle}{2\sqrt{1-S^{HH}}} \tag{16.195}$$

Also defining a denominator as $d\left(=2\sqrt{1-S^{HH}}\right)$, we have

$$H_1^{(T_2)} \equiv |H_1+H_2-H_3-H_4\rangle/d$$

Similarly, we define other hydrogen-based SALCs as

$$\begin{aligned}H_2^{(T_2)} &\equiv |H_1-H_2+H_3-H_4\rangle/d, \\ H_3^{(T_2)} &\equiv |H_1-H_2-H_3+H_4\rangle/d.\end{aligned}$$

The next step is to construct MOs using the above SALCs. To this end, we make a linear combination using SALCs belonging to the same irreducible representations. In the case of $A_1$, we choose $H^{(A_1)}$ and $C2s$. Naturally, we anticipate two linear combinations of

$$a_1 H^{(A_1)} + b_1 C2s, \tag{16.196}$$

where $a_1$ and $b_1$ are arbitrary constants. On the basis of the discussions of projection operators in Sect. 15.7, both the above two linear combinations belong to $A_1$ as well. Similarly, according to the projection operators $P_{1(1)}^{(T_2)}$, $P_{2(2)}^{(T_2)}$, and $P_{3(3)}^{(T_2)}$, we make three sets of linear combinations

$$q_1 P_{1(1)}^{(T_2)} H_1 + r_1 C2p_x;\ q_2 P_{2(2)}^{(T_2)} H_1 + r_2 C2p_y;\ q_3 P_{3(3)}^{(T_2)} H_1 + r_3 C2p_z, \tag{16.197}$$

where $q_1$, $r_1$, etc., are arbitrary constants. These three sets of linear combinations belong to individual "addresses" 1, 2, and 3 of $T_2$. What we have to do is to determine coefficients of the above MOs and to normalize them by solving the secular equations. With two different energy eigenvalues, we get two orthogonal (i.e., linearly independent) MOs for the individual four sets of linear combinations of (16.196) and (16.197). Thus, total eight linear combinations constitute MOs of methane.

In light of (16.183), the secular equation can be reduced as follows according to the representations $A_1$ and $T_2$. There we have changed the order of entries in the equation so that we can deal with the equation easily. Then, we have

$$\begin{vmatrix} H_{11}-\lambda H_{12}-\lambda S_{12} & & & & \\ H_{21}-\lambda S_{21} H_{22}-\lambda & & & & \\ & G_{11}-\lambda G_{12}-\lambda T_{12} & & & \\ & G_{21}-\lambda T_{21} G_{22}-\lambda & & & \\ & & F_{11}-\lambda F_{12}-\lambda V_{12} & & \\ & & F_{21}-\lambda V_{21} F_{22}-\lambda & & \\ & & & K_{11}-\lambda K_{12}-\lambda W_{12} & \\ & & & K_{21}-\lambda W_{21} K_{22}-\lambda & \end{vmatrix} = 0 \tag{16.198}$$

where blanc off-diagonal elements are all zero. This is because of the symmetry requirement (see Sect. 16.2). Thus, in (16.198) the secular equation is decomposed into four (2,2) blocks. The first block is pertinent to $A_1$ of a hydrogen-based component and carbon-based component from the left, respectively. Lower three blocks are pertinent to $T_2$ of hydrogen-based and carbon-based components from the left, respectively, in order of $P^{(T_2)}_{1(1)}$, $P^{(T_2)}_{2(2)}$, and $P^{(T_2)}_{3(3)}$ SALCs from the top. The notations follow those of (16.59).

We compute these equations. The calculations are equivalent to solving the following four two-dimensional secular equations.

$$\begin{aligned} \begin{vmatrix} H_{11}-\lambda & H_{12}-\lambda S_{12} \\ H_{21}-\lambda S_{21} & H_{22}-\lambda \end{vmatrix} &= 0, \\ \begin{vmatrix} G_{11}-\lambda & G_{12}-\lambda T_{12} \\ G_{21}-\lambda T_{21} & G_{22}-\lambda \end{vmatrix} &= 0, \\ \begin{vmatrix} F_{11}-\lambda & F_{12}-\lambda V_{12} \\ F_{21}-\lambda V_{21} & F_{22}-\lambda \end{vmatrix} &= 0, \\ \begin{vmatrix} K_{11}-\lambda & K_{12}-\lambda W_{12} \\ K_{21}-\lambda W_{21} & K_{22}-\lambda \end{vmatrix} &= 0. \end{aligned} \tag{16.199}$$

Notice that these four secular equations are the same as a (2,2) determinant of (16.59) in the form of a secular equation. Note, at the same time, that while (16.59) did not assume SALCs, (16.199) takes account of SALCs. That is, (16.199) expresses a secular equation with respect to two SALCs that belong to the same irreducible representation.

The first equation of (16.199) reads as

$$\left(1-S_{12}^2\right)\lambda^2 - (H_{11}+H_{22}-2H_{12}S_{12})\lambda + H_{11}H_{22} - H_{12}^2 = 0 \tag{16.200}$$

In (16.200), we define quantities as follows:

$$S_{12} \equiv \int H^{(A_1)} C2s d\tau \equiv \left\langle H^{(A_1)} | C2s \right\rangle = \frac{\langle H_1 + H_2 + H_3 + H_4 | C2s \rangle}{2\sqrt{1+3S^{HH}}} = \frac{2\langle H_1 | C2s \rangle}{\sqrt{1+3S^{HH}}} = \frac{2S_{A_1}^{CH}}{\sqrt{1+3S^{HH}}} \tag{16.201}$$

In (16.201), an overlap integral between hydrogen atomic orbitals and $C2s$ is identical from the symmetry requirement and it is defined as

$$S_{A_1}^{CH} \equiv \langle H_1 | C2s \rangle \tag{16.202}$$

Also in (16.200), other quantities are defined as follows:

$$H_{11} \equiv \left\langle H^{(A_1)} | \mathfrak{H} H^{(A_1)} \right\rangle = \frac{\alpha^H + 3\beta^{HH}}{1+3S^{HH}} \tag{16.203}$$

$$H_{22} \equiv \langle C2s | \mathfrak{H} C2s \rangle \tag{16.204}$$

$$H_{12} \equiv \left\langle H^{(A_1)} | \mathfrak{H} C2s \right\rangle = \frac{\langle H_1 + H_2 + H_3 + H_4 | \mathfrak{H} C2s \rangle}{2\sqrt{1+3S^{HH}}} = \frac{2\langle H_1 | \mathfrak{H} C2s \rangle}{\sqrt{1+3S^{HH}}} = \frac{2\beta_{A_1}^{CH}}{\sqrt{1+3S^{HH}}} \tag{16.205}$$

where $\mathfrak{H}$ is a Hamiltonian of a methane molecule. In (16.203) and (16.205), moreover, we define the quantities as

$$\alpha^H \equiv \langle H_1 | \mathfrak{H} H_1 \rangle,\ \beta^{HH} \equiv \langle H_1 | \mathfrak{H} H_2 \rangle,\ \beta_{A_1}^{CH} \equiv \langle H_1 | \mathfrak{H} C2s \rangle. \tag{16.206}$$

The quantity of $H_{11}$ is a "Coulomb" integral of the hydrogen-based SALC that involves four hydrogen atoms. Solving the first equation of (16.199), we get

$$\lambda = \frac{H_{11} + H_{22} - 2H_{12}S_{12} \pm \sqrt{(H_{11} - H_{22})^2 + 4\left[H_{12}^2 + H_{11}H_{22}S_{12}^2 - H_{12}S_{12}(H_{11} + H_{22})\right]}}{2\left(1 - S_{12}^2\right)} \tag{16.207}$$

Similarly, we obtain related solutions for the latter three eigenvalue equations of (16.199). With the second equation of (16.199), for instance, we have

$$\left(1 - T_{12}^2\right)\lambda^2 - (G_{11} + G_{22} - 2G_{12}T_{12})\lambda + G_{11}G_{22} - G_{12}^2 = 0 \tag{16.208}$$

Solving this, we get

$$\lambda = \frac{G_{11} + G_{22} - 2G_{12}T_{12} \pm \sqrt{(G_{11} - G_{22})^2 + 4\left[G_{12}^2 + G_{11}G_{22}T_{12}^2 - G_{12}T_{12}(G_{11} + G_{22})\right]}}{2\left(1 - T_{12}^2\right)} \tag{16.209}$$

In (16.208) and (16.209), we define these quantities as follows:

$$\begin{aligned} T_{12} \equiv \int H_1^{(T_2)} C2p_x \mathrm{d}\tau \equiv \left\langle H_1^{(T_2)} | 2Cp_x \right\rangle &= \frac{\langle H_1 + H_2 - H_3 - H_4 | C2p_x \rangle}{2\sqrt{1 - S^{HH}}} \\ &= \frac{2\langle H_1 | C2p_x \rangle}{\sqrt{1 - S^{HH}}} \\ &= \frac{2S_{T_2}^{CH}}{\sqrt{1 - S^{HH}}} \end{aligned} \tag{16.210}$$

$$G_{11} \equiv \left\langle H_1^{(T_2)} | \mathfrak{H} H_1^{(T_2)} \right\rangle = \frac{\alpha^H - \beta^{HH}}{1 - S^{HH}} \tag{16.211}$$

$$G_{22} \equiv \langle C2p_x | \mathfrak{H} C2p_x \rangle \tag{16.212}$$

$$\begin{aligned} G_{12} \equiv \left\langle H_1^{(T_2)} | \mathfrak{H} C2p_x \right\rangle &= \frac{\langle H_1 + H_2 - H_3 - H_4 | \mathfrak{H} C2p_x \rangle}{2\sqrt{1 - S^{HH}}} \\ &= \frac{2\langle H_1 | \mathfrak{H} C2p_x \rangle}{\sqrt{1 - S^{HH}}} = \frac{2\beta_{T_2}^{CH}}{\sqrt{1 - S^{HH}}} \end{aligned} \tag{16.213}$$

In the above equations, we further define integrals such that

$$S_{T_2}^{CH} \equiv \langle H_1 | C2p_x \rangle,\ \beta_{T_2}^{CH} \equiv \langle H_1 | \mathfrak{H} C2p_x \rangle. \tag{16.214}$$

In (16.210), an overlap integral $T_{12}$ between four hydrogen atomic orbitals and $C2p_x$ is identical again from the symmetry requirement. That is, the integrals of (16.213) are *additive* regarding a product of plus components of a hydrogen atomic orbital of (16.187) and $C2p_x$ as well as another product of minus components of a hydrogen atomic orbital of (16.187) and $C2p_x$; see Fig. 16.14. Notice that $C2p_x$ has a node on $yz$-plane.

The third and fourth equations of (16.199) give exactly the same eigenvalues as that given in (16.209). This is obvious from the fact that all the latter three equations of (16.199) are associated with a irreducible representation $T_2$. The corresponding three MOs are triply degenerate.

In (16.207) and (16.209), the plus sign gives a higher orbital energy and the minus sign gives a lower energy. Equations (16.207) and (16.209), however, look somewhat complicated. To simplify the situation, (i) in, e.g., (16.207), let us consider a case where $|H_{11}| \gg |H_{22}|$ or $|H_{11}| \ll |H_{22}|$. In that case, $(H_{11} - H_{22})^2$

dominates inside a square root, and so ignoring $-2H_{12}S_{12}$ and $-S_{12}^2$, we have $\lambda \approx H_{11}$ or $\lambda \approx H_{22}$. Inserting these values into (16.199), we have either $\Psi \approx H^{(A_1)}$ or $\Psi \approx C2s$, where $\Psi$ is a resulting MO. This implies that no interaction would arise between $H^{(A_1)}$ and $C2s$. (ii) If, on the other hand, $H_{11} = H_{22}$, we would get

$$\lambda = \frac{H_{11} - H_{12}S_{12} \pm |H_{12} - H_{11}S_{12}|}{1 - S_{12}^2} \tag{16.215}$$

As $H_{12} - H_{11}S_{12}$ is positive or negative, we have a following alternative:

Case I: $H_{12} - H_{11}S_{12} > 0$. We have $\lambda_H = \frac{H_{11}+H_{12}}{1+S_{12}}$ and $\lambda_L = \frac{H_{11}-H_{12}}{1-S_{12}}$, where $\lambda_H > \lambda_L$.
Case II: $H_{12} - H_{11}S_{12} < 0$. We have $\lambda_H = \frac{H_{11}-H_{12}}{1-S_{12}}$ and $\lambda_L = \frac{H_{11}+H_{12}}{1+S_{12}}$, where $\lambda_H > \lambda_L$.

In this case, we would anticipate maximum orbital mixing between $H^{(A_1)}$ and $C2s$. (iii) If $H_{11}$ and $H_{22}$ are moderately different in between the above (i) and (ii), the orbital mixing is likely to be moderate. This is an actual situation. With (16.209), we have eigenvalues expressed similarly to those of (16.207). Therefore, classifications related to the above Cases I and II hold with $G_{12} - G_{11}T_{12}$, $F_{12} - F_{11}V_{12}$, and $K_{12} - K_{11}W_{12}$ as well.

In spite of simplicity of (16.215), the quantities $H_{11}$, $H_{12}$, and $S_{12}$ are hard to calculate. In general cases including the present example (i.e., methane), the difficulty in calculating these quantities results essentially from the fact that we are dealing with a many-particle interactions that include electron repulsion. Nonetheless, for a simplest case of a hydrogen molecular ion $H_2^+$, the estimation is feasible [3]. Let us estimate $H_{12} - H_{11}S_{12}$ quantitatively according to Atkins and Friedman [3].

The Hamiltonian of $H_2^+$ is described as

$$H = -\frac{\hbar^2}{2m}\nabla^2 - \frac{e^2}{4\pi\varepsilon_0 r_A} - \frac{e^2}{4\pi\varepsilon_0 r_B} + \frac{e^2}{4\pi\varepsilon_0 R}, \tag{16.216}$$

where $m$ is an electron rest mass, and other symbols are defined in Fig. 16.15; the last term represents the repulsive interaction between the two hydrogen nuclei. To estimate the quantities in (16.215), it is convenient to use dimensionless ellipsoidal coordinates $\begin{pmatrix} \mu \\ \nu \\ \phi \end{pmatrix}$ such that [3]

$$\mu = \frac{r_A + r_B}{R} \text{ and } \nu = \frac{r_A - r_B}{R}. \tag{16.217}$$

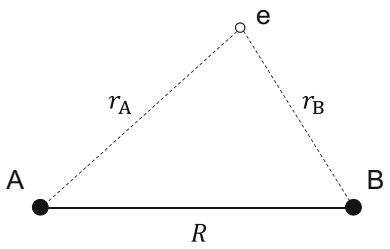


**Fig. 16.15** Configuration of electron (e) and hydrogen nuclei (A and B). $r_A$ and $r_B$ denote a separation between the electron and A and that between the electron and B, respectively. $R$ denotes a separation between A and B

The quantity $\phi$ is an azimuthal angle around the molecular axis (i.e., the straight line connecting the two nuclei). Then, we have

$$\begin{aligned} H_{11} &= \langle A|H|A\rangle = \left\langle A\left|-\frac{\hbar^2}{2m}\nabla^2 - \frac{e^2}{4\pi\varepsilon_0 r_A}\right|A\right\rangle + \left\langle A - \left|\frac{e^2}{4\pi\varepsilon_0 r_B}\right|A\right\rangle + \left\langle A\left|\frac{e^2}{4\pi\varepsilon_0 R}\right|A\right\rangle \\ &= E_{1s} - \frac{e^2}{4\pi\varepsilon_0}\left\langle A\left|\frac{1}{r_B}\right|A\right\rangle + \frac{e^2}{4\pi\varepsilon_0 R}, \end{aligned} \tag{16.218}$$

where $E_{1s}$ is the same as that given in (3.258) with $Z = n = 1$ and $\mu$ replaced with $m$ (i.e., $E_{1s} = -\frac{\hbar^2}{2ma^2}$). Using a coordinate representation and of (3.301), we have

$$|A\rangle = \sqrt{\frac{1}{\pi}}a^{-3/2}e^{-r_A/a} \tag{16.219}$$

Moreover considering (16.217), we have

$$\left\langle A\left|\frac{1}{r_B}\right|A\right\rangle = \frac{1}{\pi a^3}\int d\tau e^{-2r_A/a}\frac{1}{r_B}$$

Converting Cartesian coordinates to ellipsoidal coordinates [3] such that

$$\int \mathrm{d}\tau = \left(\frac{R}{2}\right)^3 \int_0^{2\pi} \mathrm{d}\phi \int_1^{\infty} \mathrm{d}\mu \int_{-1}^{1} \mathrm{d}\nu\left(\mu^2 - \nu^2\right),$$

we have

$$\begin{aligned} \left\langle A\left|\frac{1}{r_B}\right|A\right\rangle &= \frac{R^3}{8\pi a^3}\cdot 2\pi \int_1^{\infty}\mathrm{d}\mu \int_{-1}^{1}\mathrm{d}\nu\left(\mu^2-\nu^2\right)\frac{e^{-(\mu+\nu)R/a}}{\frac{R}{2}(\mu-\nu)} \\ &= \frac{R^2}{2a^3}\int_1^{\infty}\mathrm{d}\mu\int_{-1}^{1}\mathrm{d}\nu(\mu+\nu)e^{-(\mu+\nu)R/a} \end{aligned} \tag{16.220}$$

Putting $I = \int_1^{\infty}\mathrm{d}\mu\int_{-1}^{1}\mathrm{d}\nu(\mu+\nu)e^{-(\mu+\nu)R/a}$, we obtain

$$I = \int_1^\infty \mu e^{-\mu R/a} \mathrm{d}\mu \int_{-1}^1 e^{-\nu R/a} \mathrm{d}\nu + \int_1^\infty e^{-\mu R/a} \mathrm{d}\mu \int_{-1}^1 \nu e^{-\nu R/a} \mathrm{d}\nu \tag{16.221}$$

The above definite integrals can readily be calculated using the methods described in Sect. 3.7.2; see, e.g., (3.262) and (3.263). For instance, we have

$$\int_{-1}^1 e^{-c\nu} \mathrm{d}\nu = \left[\frac{e^{-cx}}{-c}\right]_{-1}^1 = \frac{1}{c}(e^c - e^{-c}) \tag{16.222}$$

Differentiating (16.222) with respect to the parameter $c$, we get

$$\int_{-1}^1 \nu e^{-c\nu} \mathrm{d}\nu = \frac{1}{c^2}[(1-c)e^c - (1+c)e^{-c}] \tag{16.223}$$

In the present case, $c$ is to be replaced with $R/a$. Other calculations of the definite integrals are left for readers. Thus, we obtain

$$I = \frac{a^3}{R^3}\left[2 - 2\left(1 + \frac{R}{a}\right)e^{-2R/a}\right]$$

In turn, we have

$$\left\langle A \middle| \frac{1}{r_\mathrm{B}} \middle| A \right\rangle = \frac{R^2}{2a^3} \cdot I = \frac{1}{R}\left[1 - \left(1 + \frac{R}{a}\right)e^{-2R/a}\right]$$

Introducing a symbol according to Atkins and Friedman [3]

$$j_0 \equiv \frac{e^2}{4\pi\varepsilon_0},$$

we finally obtain

$$H_{11} = E_{1s} - \frac{j_0}{R}\left[1 - \left(1 + \frac{R}{a}\right)e^{-2R/a}\right] + \frac{j_0}{R} \tag{16.224}$$

The quantity $S_{12}$ can be obtained as follows:

$$S_{12} = \langle B|A\rangle = \frac{R^3}{8\pi a^3} \int_0^{2\pi} \mathrm{d}\phi \int_1^\infty \mathrm{d}\mu \int_{-1}^1 \mathrm{d}\nu (\mu^2 - \nu^2) e^{-\mu R/a}, \tag{16.225}$$

where we used $|B\rangle = \sqrt{\frac{1}{\pi}} a^{-3/2} e^{-r_\mathrm{B}/a}$ and (16.217). Noting that

$$\int_1^\infty \mathrm{d}\mu \int_{-1}^{1} \mathrm{d}\nu\left(\mu^2-\nu^2\right)e^{-\mu R/a} = \int_1^\infty \mathrm{d}\mu\left(2\mu^2-\frac{2}{3}\right)e^{-\mu R/a}$$

and following procedures similar to those described above, we get

$$S_{12} = \left[1+\frac{R}{a}+\frac{1}{3}\left(\frac{R}{a}\right)^2\right]e^{-R/a} \tag{16.226}$$

In turn, for $H_{12}$, we have

$$\begin{aligned} H_{12} = \langle A|H|B\rangle &= \left\langle A\left|-\frac{\hbar^2}{2m}\nabla^2-\frac{e^2}{4\pi\varepsilon_0 r_\mathrm{B}}\right|B\right\rangle + \left\langle A\left|-\frac{e^2}{4\pi\varepsilon_0 r_\mathrm{A}}\right|B\right\rangle + \left\langle A\left|\frac{e^2}{4\pi\varepsilon_0 R}\right|B\right\rangle \\ &= E_{1s}\langle A|B\rangle - \frac{e^2}{4\pi\varepsilon_0}\left\langle A\left|\frac{1}{r_\mathrm{A}}\right|B\right\rangle + \frac{e^2}{4\pi\varepsilon_0 R}\langle A|B\rangle \end{aligned} \tag{16.227}$$

Note that in (16.227) $|B\rangle$ is an eigenfunction belonging to $E_{1s} = -\frac{\hbar^2}{2ma^2}$ that is an eigenvalue of an operator $-\frac{\hbar^2}{2m}\nabla^2 - \frac{e^2}{4\pi\varepsilon_0 r_\mathrm{B}}$. From (3.258), we estimate $E_{1s}$ to be 13.61 eV. Using (16.217) and converting Cartesian coordinates to ellipsoidal coordinates once again, we get

$$\left\langle A\left|\frac{1}{r_\mathrm{A}}\right|B\right\rangle = \frac{1}{a}\left(1+\frac{R}{a}\right)e^{-R/a}.$$

Thus, we have

$$H_{12} = \left(E_{1s}+\frac{j_0}{R}\right)S_{12} - \frac{j_0}{a}\left(1+\frac{R}{a}\right)e^{-R/a} \tag{16.228}$$

We are now in the position to evaluate $H_{12} - H_{11}S_{12}$ in (16.213). According to Atkins and Friedman [3], we define the following notations:

$$j' \equiv j_0\left\langle A\left|\frac{1}{r_\mathrm{B}}\right|A\right\rangle \text{ and } k' \equiv j_0\left\langle A\left|\frac{1}{r_\mathrm{A}}\right|B\right\rangle$$

Then, we have

$$H_{12} - H_{11}S_{12} = -k' + j'S_{12} \tag{16.229}$$

The calculation of this quantity is straightforward. The result is

$$H_{12} - H_{11}S_{12} = j_0\left(\frac{1}{R} - \frac{2R}{3a^2}\right)e^{-R/a} - \frac{j_0}{R}\left(1 + \frac{R}{a}\right)\left[1 + \frac{R}{a} + \frac{1}{3}\left(\frac{R}{a}\right)^2\right]e^{-3R/a}$$
$$= \frac{j_0}{a}\left(\frac{a}{R} - \frac{2R}{3a}\right)e^{-R/a} - \frac{j_0}{a}\left(1 + \frac{a}{R}\right)\left[1 + \frac{R}{a} + \frac{1}{3}\left(\frac{R}{a}\right)^2\right]e^{-3R/a} \tag{16.230}$$

In (16.230), we notice that whereas the second term is always negative, the first term may be negative or positive depending upon $R$. We could not tell a priori whether $H_{12} - H_{11}S_{12}$ is negative accordingly. If we had $R \ll a$, (16.230) would become positive.

Let us then make a quantitative estimation. The Bohr radius $a$ is about 52.9 pm (using an electron rest mass). As an experimental result, $R$ is approximately 106 pm [3]. Hence, for a $H_2^+$ ion, we have

$$R/a \approx 2.0$$

Using this number, we get

$$H_{12} - H_{11}S_{12} \approx -0.13j_0/a < 0$$

We estimate $H_{12} - H_{11}\ S_{12}$ to be $\sim -3.5\,\text{eV}$. Therefore, from (16.215), we get

$$\lambda_L = \frac{H_{11} + H_{12}}{1 + S_{12}} \text{ and } \lambda_H = \frac{H_{11} - H_{12}}{1 - S_{12}} \tag{16.231}$$

where $\lambda_L$ and $\lambda_H$ indicate lower and higher energy eigenvalues, respectively. Namely, Case II in the above is more likely. Correspondingly, for MOs, we have

$$\Psi_L = \frac{|A\rangle + |B\rangle}{\sqrt{2(1 + S_{12})}} \text{ and } \Psi_H = \frac{|A\rangle - |B\rangle}{\sqrt{2(1 - S_{12})}} \tag{16.232}$$

where $\Psi_L$ and $\Psi_H$ belong to $\lambda_L$ and $\lambda_H$, respectively. The results are virtually the same as those given in (16.101)–(16.105) of Sect. 16.4.1. In (16.101) and Fig. 16.6, however, we merely assumed that $\lambda_1 = \frac{\alpha+\beta}{1+S}$ is lower than $\lambda_2 = \frac{\alpha-\beta}{1-S}$. Here, we have confirmed that this is truly the case. A chemical bond is formed in such a way that an electron is distributed as much as possible along the molecular axis (see Fig. 16.4 for a schematic) and that in this configuration a minimized orbital energy is achieved.

In our present case, $H_{11}$ in (16.203) and $H_{22}$ in (16.204) should differ. This is also the case with $G_{11}$ in (16.211) and $G_{22}$ in (16.212). Here, we return back to (16.199). Suppose that we get a MO by solving, e.g., the first secular equation of (16.199) such that

$$\Psi = a_1 H^{(A_1)} + b_1 C2\mathrm{s}$$

From the secular equation, we get

$$a_1 = -\frac{H_{12} - \lambda S_{12}}{H_{11} - \lambda} b_1, \tag{16.233}$$

where $S_{12}$ and $H_{12}$ were defined as in (16.201) and (16.205), respectively. A normalized MO $\widetilde{\psi}$ is described by

$$\widetilde{\Psi} = \frac{a_1 H^{(A_1)} + b_1 C2\mathrm{s}}{\sqrt{a_1^2 + b_1^2 + 2a_1 b_1 S_{12}}} \tag{16.234}$$

Thus, according to two different energy eigenvalues $\lambda$, we will get two linearly independent MOs. Other three secular equations are dealt with similarly.

From the energetical consideration of $H_2^+$, we infer that $a_1 b_1 > 0$ with a bonding MO and $a_1 b_1 < 0$ with an anti-bonding MO. Meanwhile, since both $H^{(A_1)}$ and $C2\mathrm{s}$ belong to the irreducible representation $A_1$, so does $\widetilde{\Psi}$ on the basis of the discussion on the projection operators of Sect. 15.7. Thus, we should be able to construct proper MOs that belong to $A_1$. Similarly, we get proper MOs belonging to $T_2$ by solving other three secular equations of (16.199). In this case, three bonding MOs are triply degenerate, so are three anti-bonding MOs. All these six MOs belong to the irreducible representation $T_2$. Thus, we can get a complete set of MOs for methane. These eight MOs span the representation space $V^8$.

To precisely determine the energy levels, we need to take more elaborate approaches to approximate and calculate various parameters that appear in (16.199). At the same time, we need to perform detailed experiments including spectroscopic measurements and interpret those results carefully [4]. Taking account of these situations, Fig. 16.16 [5] displays as an example of MO calculations that give a probable energy diagram and MO symmetry species of methane. The diagram comprises a ground state bonding $a_1$ state and its corresponding anti-bonding state $a_1^*$ along with triply degenerate bonding $t_2$ states and their corresponding anti-bonding state $t_2^*$.

We emphasize that the said elaborate approaches ensue truly from the "paper-and-pencil" methods based upon group theory. The group theory thus supplies us with a powerful tool and clear guideline for addressing various quantum chemical problems, a few of which we are introducing as an example in this book.

Finally, let us examine the optical transition of methane. In this case, we have to consider electronic configurations of the initial and final states. If we are dealing with optical absorption, the initial state is the ground state $A_1$ that is described by the totally symmetric representation. The final state, on the other hand, will be an excited state, which is described by a direct-product representation related to the

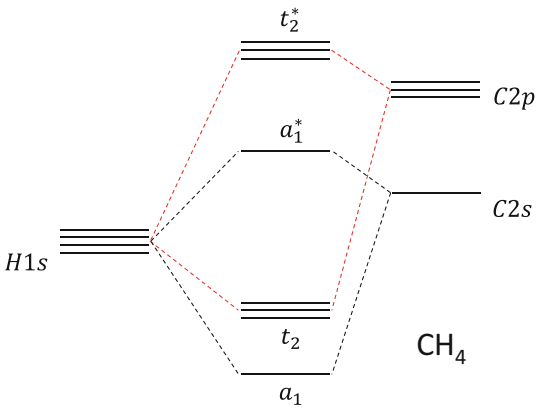


**Fig. 16.16** Probable energy diagram and MO symmetry species of methane. Adapted from http://www.science.oregonstate.edu/%7Egablek/CH334/Chapter1/methane_MOs.htm with kind permission of Professor Kevin P. Gable

two states that are associated with the optical transition. The matrix element is expressed as

$$\left\langle \Theta^{(\alpha\times\beta)} \middle| \boldsymbol{\varepsilon}_e \cdot \boldsymbol{P} \middle| \Theta^{(A_1)} \right\rangle, \tag{16.235}$$

where $\Theta^{(A_1)}$ stands for an electronic configuration of the totally symmetric ground state; $\Theta^{(\alpha\times\beta)}$ denotes an electronic configuration of an excited state represented by a direct-product representation pertinent to irreducible representations $\alpha$ and $\beta$; $\boldsymbol{P}$ is an electric dipole operator; and $\boldsymbol{\varepsilon}_e$ is a unit polarization vector of the electric field. From a character table for $T_d$, we find that $\boldsymbol{\varepsilon}_e \cdot \boldsymbol{P}$ belongs to the irreducible representation $T_2$ (see Table 16.12).

The ground state electronic configuration is $A_1$ (totally symmetric). It is denoted by

$$a_1^2 t_2^2 t_2'^2 t_2''^2,$$

where three $T_2$ states are distinguished by a prime and double prime. For possible configuration of excited states, we have

$$\begin{aligned} &a_1^2 t_2 t_2'^2 t_2''^2 t_2^* \,(A_1 \to T_2 \times T_2),\ a_1^2 t_2 t_2'^2 t_2''^2 a_1^* \,(A_1 \to T_2 \times A_1 = T_2), \\ &a_1 t_2^2 t_2'^2 t_2''^2 t_2^* \,(A_1 \to A_1 \times T_2 = T_2),\ a_1 t_2^2 t_2'^2 t_2''^2 a_1^* \,(A_1 \to A_1 \times A_1 = A_1). \end{aligned} \tag{16.236}$$

In (16.236), we have optical excitations of $t_2 \to t_2^*$, $t_2 \to a_1^*$, $a_1 \to t_2^*$, and $a_1 \to a_1^*$, respectively. Consequently, the excited states are denoted by $T_2 \times T_2$, $T_2$, $T_2$, and $A_1$, respectively. Since $\boldsymbol{\varepsilon}_e \cdot \boldsymbol{P}$ is Hermitian as shown in (16.51) of Sect. 16.2, we have

$$\left\langle \Theta^{(\alpha\times\beta)} \middle| \varepsilon_e \cdot \boldsymbol{P} \middle| \Theta^{(A_1)} \right\rangle = \left\langle \Theta^{(A_1)} \middle| \varepsilon_e \cdot \boldsymbol{P} \middle| \Theta^{(\alpha\times\beta)} \right\rangle^*. \tag{16.237}$$

Therefore, according to the general theory of Sect. 16.2, we need to examine whether $T_2 \times D^{(\alpha)} \times D^{(\beta)}$ contains $A_1$ to judge whether the optical transition is allowed. As mentioned above, $D^{(\alpha)} \times D^{(\beta)}$ is chosen from among $T_2 \times T_2$, $T_2$, $T_2$, and $A_1$. The results are given as below:

$$T_2 \times T_2 \times T_2 = 3T_1 + 4T_2 + 2E + A_2 + A_1, \tag{16.238}$$

$$T_2 \times T_2 = T_1 + T_2 + E + A_1, \tag{16.239}$$

$$T_2 \times A_1 = T_2. \tag{16.240}$$

Admittedly, (16.238) and (16.239) contain $A_1$, and so the transition is allowed. As for (16.240), however, the transition is forbidden because it does not contain $A_1$. In light of the character table of $T_d$ (Table 16.12), we find that the allowed transitions (i.e., $t_2 \to t_2^*$, $t_2 \to a_1^*$, and $a_1 \to t_2^*$) equally take place in the direction polarized along the $x$-, $y$-, $z$-axes. This is often the case with molecules having higher symmetries such as methane. The transition $a_1 \to a_1^*$, on the other hand, is forbidden.

In the above argument including the energetic consideration, we could not tell magnitude relationship of MO energies or photon energies associated with the optical transition. Once again, this requires more accurate calculations and experiments. Yet, the discussion we have developed gives us a strong guiding principle in the investigation of molecular science.

## References

1. Cotton FA (1990) Chemical applications of group theory. Wiley, New York
2. Hamermesh M (1989) Group theory and its application to physical problems. Dover, New York
3. Atkins P, Friedman R (2005) Molecular quantum mechanics, 4th edn. Oxford University Press, Oxford
4. Anslyn EV, Dougherty DA (2006) Modern physical organic chemistry. University Science Books, Sausalito
5. Gable KP (2014) Molecular orbitals: Methane, http://www.science.oregonstate.edu/~gablek/CH334/Chapter1/methane_MOs.htm

# Index




S. Hotta, *Mathematical Physical Chemistry*,
https://doi.org/10.1007/978-981-10-7671-8