Michio Masujima

Applied Mathematical Methods in Theoretical Physics

WILEY-VCH Verlag GmbH & Co. KGaA

Michio Masujima

Applied Mathematical Methods in Theoretical Physics

Michio Masujima

Applied Mathematical Methods in Theoretical Physics

WILEY-VCH Verlag GmbH & Co. KGaA

Cover Picture
K. Schmidt

Library of Congress Card No.: applied for
British Library Cataloging-in-Publication Data:
A catalogue record for this book is available from the British Library

Bibliographic information published by
Die Deutsche Bibliothek
Die Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data is available in the
Internet at <http://dnb.ddb.de>.

Printed in the Federal Republic of Germany
Printed on acid-free paper

Printing Strauss GmbH, Mörlenbach
Bookbinding Litges & Dopf Buchbinderei GmbH, Heppenheim

ISBN-13: 978- 3-527-40534-3
ISBN-10: 3-527-40534-8

Contents

Preface

This book on integral equations and the calculus of variations is intended for use by senior undergraduate students and first-year graduate students in science and engineering. Basic familiarity with theories of linear algebra, calculus, differential equations, and complex analysis on the mathematics side, and classical mechanics, classical electrodynamics, quantum mechanics including the second quantization, and quantum statistical mechanics on the physics side, is assumed. Another prerequisite for this book on the mathematics side is a sound understanding of local and global analysis.

This book grew out of the course notes for the last of the three-semester sequence of *Methods of Applied Mathematics I (Local Analysis)*, *II (Global Analysis)* and *III (Integral Equations and Calculus of Variations)* taught in the Department of Mathematics at MIT. About two-thirds of the course is devoted to integral equations and the remaining one-third to the calculus of variations. Professor Hung Cheng taught the course on integral equations and the calculus of variations every other year from the mid 1960s through the mid 1980s at MIT. Since then, younger faculty have been teaching the course in turn. The course notes evolved in the intervening years. This book is the culmination of these joint efforts.

There will be the obvious question: Why yet another book on integral equations and the calculus of variations? There are already many excellent books on the theory of integral equations. No existing book, however, discusses the singular integral equations in detail; in particular, Wiener–Hopf integral equations and Wiener–Hopf sum equations with the notion of the Wiener–Hopf index. In this book, the notion of the Wiener–Hopf index is discussed in detail.

This book is organized as follows. In Chapter 1 we discuss the notion of function space, the linear operator, the Fredholm alternative and Green's functions, to prepare the reader for the further development of the material. In Chapter 2 we discuss a few examples of integral equations and Green's functions. In Chapter 3 we discuss integral equations of the Volterra type. In Chapter 4 we discuss integral equations of the Fredholm type. In Chapter 5 we discuss the Hilbert–Schmidt theories of the symmetric kernel. In Chapter 6 we discuss singular integral equations of the Cauchy type. In Chapter 7, we discuss the Wiener–Hopf method for the mixed boundary-value problem in classical electrodynamics, Wiener–Hopf integral equations, and Wiener–Hopf sum equations; the latter two topics being discussed in terms of the notion of the index. In Chapter 8 we discuss nonlinear integral equations of the Volterra, Fredholm and Hammerstein type. In Chapter 9 we discuss the calculus of variations, in particular, the second variations, the Legendre test and the Jacobi test, and the relationship between integral equations and applications of the calculus of variations. In Chapter 10 we discuss Feynman's action principle in quantum mechanics and Feynman's variational principle, a system

of the Schwinger–Dyson equations in quantum field theory and quantum statistical mechanics, Weyl's gauge principle and Kibble's gauge principle.

A substantial portion of Chapter 10 is taken from my monograph, "*Path Integral Quantization and Stochastic Quantization*", Vol. 165, Springer Tracts in Modern Physics, Springer, Heidelberg, published in the year 2000.

A reasonable understanding of Chapter 10 requires the reader to have a basic understanding of classical mechanics, classical field theory, classical electrodynamics, quantum mechanics including the second quantization, and quantum statistical mechanics. For this reason, Chapter 10 can be read as a side reference on theoretical physics, independently of Chapters 1 through 9.

The examples are mostly taken from classical mechanics, classical field theory, classical electrodynamics, quantum mechanics, quantum statistical mechanics and quantum field theory. Most of them are worked out in detail to illustrate the methods of the solutions. Those examples which are not worked out in detail are either intended to illustrate the general methods of the solutions or it is left to the reader to complete the solutions.

At the end of each chapter, with the exception of Chapter 1, problem sets are given for sound understanding of the content of the main text. The reader is recommended to solve all the problems at the end of each chapter. Many of the problems were created by Professor Hung Cheng over the past three decades. The problems due to him are designated by the note '(Due to H. C.)'. Some of the problems are those encountered by Professor Hung Cheng in the course of his own research activities.

Most of the problems can be solved by the direct application of the method illustrated in the main text. Difficult problems are accompanied by the citation of the original references. The problems for Chapter 10 are mostly taken from classical mechanics, classical electrodynamics, quantum mechanics, quantum statistical mechanics and quantum field theory.

A bibliography is provided at the end of the book for an in-depth study of the background materials in physics, beside the standard references on the theory of integral equations and the calculus of variations.

The instructor can cover Chapters 1 through 9 in one semester or two quarters with a choice of the topic of his or her own taste from Chapter 10.

I would like to express many heart-felt thanks to Professor Hung Cheng at MIT, who appointed me as his teaching assistant for the course when I was a graduate student in the Department of Mathematics at MIT, for his permission to publish this book under my single authorship and also for his criticism and constant encouragement without which this book would not have materialized.

I would like to thank Professor Francis E. Low and Professor Kerson Huang at MIT, who taught me many of the topics within theoretical physics. I would like to thank Professor Roberto D. Peccei at Stanford University, now at UCLA, who taught me quantum field theory and dispersion theory.

I would like to thank Professor Richard M. Dudley at MIT, who taught me real analysis and theories of probability and stochastic processes. I would like to thank Professor Herman Chernoff, then at MIT, now at Harvard University, who taught me many topics in mathematical statistics starting from multivariate normal analysis, for his supervision of my Ph. D. thesis at MIT.

I would like to thank Dr. Ali Nadim for supplying his version of the course notes and Dr. Dionisios Margetis at MIT for supplying examples and problems of integral equations from his courses at Harvard University and MIT. The problems due to him are designated by the note '(Due to D. M.)'. I would like to thank Dr. George Fikioris at the National Technical University of Athens for supplying the references on the Yagi–Uda semi-infinite arrays.

I would like to thank my parents, Mikio and Hanako Masujima, who made my undergraduate study at MIT possible by their financial support. I also very much appreciate their moral support during my graduate student days at MIT. I would like to thank my wife, Mari, and my son, Masachika, for their strong moral support, patience and encouragement during the period of the writing of this book, when the 'going got tough'.

Lastly, I would like to thank Dr. Alexander Grossmann and Dr. Ron Schulz of Wiley-VCH GmbH & Co. KGaA for their administrative and legal assistance in resolving the copyright problem with Springer.

Michio Masujima

Tokyo, Japan,
June, 2004

Introduction

Many problems within theoretical physics are frequently formulated in terms of ordinary differential equations or partial differential equations. We can often convert them into integral equations with boundary conditions or with initial conditions built in. We can formally develop the perturbation series by iterations. A good example is the Born series for the potential scattering problem in quantum mechanics. In some cases, the resulting equations are nonlinear integro-differential equations. A good example is the Schwinger–Dyson equation in quantum field theory and quantum statistical mechanics. It is the nonlinear integro-differential equation, and is exact and closed. It provides the starting point of Feynman–Dyson type perturbation theory in configuration space and in momentum space. In some singular cases, the resulting equations are Wiener–Hopf integral equations. These originate from research on the radiative equilibrium on the surface of a star. In the two-dimensional Ising model and the analysis of the Yagi–Uda semi-infinite arrays of antennas, among others, we have the Wiener–Hopf sum equation.

The theory of integral equations is best illustrated by the notion of functionals defined on some function space. If the functionals involved are quadratic in the function, the integral equations are said to be linear integral equations, and if they are higher than quadratic in the function, the integral equations are said to be nonlinear integral equations. Depending on the form of the functionals, the resulting integral equations are said to be of the first kind, of the second kind, or of the third kind. If the kernels of the integral equations are square-integrable, the integral equations are said to be nonsingular, and if the kernels of the integral equations are not square-integrable, the integral equations are then said to be singular. Furthermore, depending on whether or not the endpoints of the kernel are fixed constants, the integral equations are said to be of the Fredholm type, Volterra type, Cauchy type, or Wiener–Hopf types, etc. By the discussion of the variational derivative of the quadratic functional, we can also establish the relationship between the theory of integral equations and the calculus of variations. The integro-differential equations can best be formulated in this manner. Analogies of the theory of integral equations with the system of linear algebraic equations are also useful.

The integral equation of Cauchy type has an interesting application to classical electrodynamics, namely, dispersion relations. Dispersion relations were derived by Kramers in 1927 and Kronig in 1926, for X-ray dispersion and optical dispersion, respectively. Kramers-Kronig dispersion relations are of very general validity which only depends on the assumption of the causality. The requirement of the causality alone determines the region of analyticity of dielectric constants. In the mid 1950s, these dispersion relations were also derived from quantum field theory and applied to strong interaction physics. The application of the covariant perturbation theory to strong interaction physics was impossible due to the large coupling

constant. From the mid 1950s to the 1960s, the dispersion-theoretic approach to strong interaction physics was the only realistic approach that provided many sum rules. To cite a few, we have the Goldberger–Treiman relation, the Goldberger–Miyazawa–Oehme formula and the Adler–Weisberger sum rule. In the dispersion-theoretic approach to strong interaction physics, experimentally observed data were directly used in the sum rules. The situation changed dramatically in the early 1970s when quantum chromodynamics, the relativistic quantum field theory of strong interaction physics, was invented by the use of asymptotically-free non-Abelian gauge field theory.

The region of analyticity of the scattering amplitude in the upper-half k-plane in quantum field theory, when expressed in terms of the Fourier transform, is immediate since quantum field theory has microscopic causality. But, the region of analyticity of the scattering amplitude in the upper-half k-plane in quantum mechanics, when expressed in terms of the Fourier transform, is not immediate since quantum mechanics does not have microscopic causality. We shall invoke the generalized triangular inequality to derive the region of analyticity of the scattering amplitude in the upper-half k-plane in quantum mechanics. This region of analyticity of the scattering amplitudes in the upper-half k-plane in quantum mechanics and quantum field theory strongly depends on the fact that the scattering amplitudes are expressed in terms of the Fourier transform. When another expansion basis is chosen, such as the Fourier–Bessel series, the region of analyticity drastically changes its domain.

In the standard application of the calculus of variations to the variety of problems in theoretical physics, we simply write the Euler equation and are rarely concerned with the second variations; the Legendre test and the Jacobi test. Examination of the second variations and the application of the Legendre test and the Jacobi test becomes necessary in some cases of the application of the calculus of variations theoretical physics problems. In order to bring the development of theoretical physics and the calculus of variations much closer, some historical comments are in order here.

Euler formulated Newtonian mechanics by the variational principle; the Euler equation. Lagrange began the whole field of the calculus of variations. He also introduced the notion of generalized coordinates into classical mechanics and completely reduced the problem to that of differential equations, which are presently known as Lagrange equations of motion, with the Lagrangian appropriately written in terms of kinetic energy and potential energy. He successfully converted classical mechanics into analytical mechanics using the variational principle. Legendre constructed the transformation methods for thermodynamics which are presently known as the Legendre transformations. Hamilton succeeded in transforming the Lagrange equations of motion, which are of the second order, into a set of first-order differential equations with twice as many variables. He did this by introducing the canonical momenta which are conjugate to the generalized coordinates. His equations are known as Hamilton's canonical equations of motion. He successfully formulated classical mechanics in terms of the principle of least action. The variational principles formulated by Euler and Lagrange apply only to the conservative system. Hamilton recognized that the principle of least action in classical mechanics and Fermat's principle of shortest time in geometrical optics are strikingly analogous, permitting the interpretation of optical phenomena in mechanical terms and vice versa. Jacobi quickly realized the importance of the work of Hamilton. He noted that Hamilton was using just one particular set of the variables to describe the mechanical system and formulated the canonical transformation theory using the Legendre transformation. He

duly arrived at what is presently known as the Hamilton–Jacobi equation. He formulated his version of the principle of least action for the time-independent case.

Path integral quantization procedure, invented by Feynman in 1942 in the Lagrangian formalism, is usually justified by the Hamiltonian formalism. We deduce the canonical formalism of quantum mechanics from the path integral formalism. As a byproduct of the discussion of the Schwinger–Dyson equation, we deduce the path integral formalism of quantum field theory from the canonical formalism of quantum field theory.

Weyl's gauge principle also attracts considerable attention due to the fact that all forces in nature; the electromagnetic force, the weak force and the strong force, can be unified with Weyl's gauge principle by the appropriate choice of the grand unifying Lie groups as the gauge group. Inclusion of the gravitational force requires the use of superstring theory.

Basic to these are the integral equations and the calculus of variations.

1 Function Spaces, Linear Operators and Green's Functions

1.1 Function Spaces

Consider the set of all complex valued functions of the real variable x, denoted by $f(x), g(x), \ldots$ and defined on the interval (a, b). We shall restrict ourselves to those functions which are *square-integrable*. Define the *inner product* of any two of the latter functions by

$$(f, g) \equiv \int_a^b f^*(x)\, g(x)\; dx, \tag{1.1.1}$$

in which $f^*(x)$ is the complex conjugate of $f(x)$. The following properties of the inner product follow from the definition (1.1.1).

$$\begin{array}{rcl} (f,g)^* & = & (g,f), \\ (f, g+h) & = & (f,g)+(f,h), \\ (f,\alpha g) & = & \alpha(f,g), \\ (\alpha f, g) & = & \alpha^*(f,g), \end{array} \tag{1.1.2}$$

with α a complex scalar.

While the inner product of any two functions is in general a complex number, the inner product of a function with itself is a real number and is non-negative. This prompts us to define the *norm of a function* by

$$\|f\| \equiv \sqrt{(f,f)} = \left[\int_a^b f^*(x) f(x)\; dx\right]^{\frac{1}{2}}, \tag{1.1.3}$$

provided that f is *square-integrable*, i.e., $\|f\| < \infty$. Equation (1.1.3) constitutes a proper definition for a norm since it satisfies the following conditions,

$$\begin{array}{lll} \text{(i)} & \textit{scalar multiplication} & \|\alpha f\| = |\alpha| \cdot \|f\|, \quad \text{for all complex } \alpha, \\ \text{(ii)} & \textit{positivity} & \|f\| > 0, \quad \text{for all } f \neq 0, \\ & & \|f\| = 0, \quad \text{if and only if } f = 0, \\ \text{(iii)} & \textit{triangular inequality} & \|f+g\| \leq \|f\| + \|g\|. \end{array} \tag{1.1.4}$$

A very important inequality satisfied by the *inner product* (1.1.1) is the so-called *Schwarz inequality* which says

$$|(f,g)| \leq \|f\| \cdot \|g\| . \tag{1.1.5}$$

To prove the latter, start with the trivial inequality $\|(f+\alpha g)\|^2 \geq 0$, which holds for any $f(x)$ and $g(x)$ and for any complex number α. With a little algebra, the left-hand side of this inequality may be expanded to yield

$$(f,f) + \alpha^*(g,f) + \alpha(f,g) + \alpha\alpha^*(g,g) \geq 0. \tag{1.1.6}$$

The latter inequality is true for any α, and is thus true for the value of α which minimizes the left-hand side. This value can be found by writing α as $a+ib$ and minimizing the left-hand side of Eq. (1.1.6) with respect to the real variables a and b. A quicker way would be to treat α and α^* as independent variables and requiring $\partial/\partial\alpha$ and $\partial/\partial\alpha^*$ of the left hand side of Eq. (1.1.6) to vanish. This immediately yields $\alpha = -(g,f)/(g,g)$ as the value of α at which the minimum occurs. Evaluating the left-hand side of Eq. (1.1.6) at this minimum then yields

$$\|f\|^2 \geq \frac{|(f,g)|^2}{\|g\|^2}, \tag{1.1.7}$$

which proves the Schwarz inequality (1.1.5).

Once the Schwarz inequality has been established, it is relatively easy to prove the *triangular inequality* (1.1.4)(iii). To do this, we simply begin from the definition

$$\|f+g\|^2 = (f+g, f+g) = (f,f) + (f,g) + (g,f) + (g,g). \tag{1.1.8}$$

Now the right-hand side of Eq. (1.1.8) is a sum of complex numbers. Applying the usual triangular inequality for complex numbers to the right-hand side of Eq. (1.1.8) yields

$$\begin{aligned} |\text{Right-hand side of Eq. (1.1.8)}| &\leq \|f\|^2 + |(f,g)| + |(g,f)| + \|g\|^2 \\ &= (\|f\| + \|g\|)^2. \end{aligned} \tag{1.1.9}$$

Combining Eqs. (1.1.8) and (1.1.9) finally proves the triangular inequality (1.1.4)(iii).

We remark finally that the set of functions $f(x)$, $g(x)$, ... is an example of a *linear vector space*, equipped with an inner product and a norm based on that inner product. A similar set of properties, including the Schwarz and triangular inequalities, can be established for other linear vector spaces. For instance, consider the set of all complex column vectors $\vec{u}$, $\vec{v}$, $\vec{w}$, ... of finite dimension n. If we define the inner product

$$(\vec{u}, \vec{v}) \equiv (\vec{u}^*)^{\mathrm{T}} \vec{v} = \sum_{k=1}^{n} u_k^* v_k, \tag{1.1.10}$$

and the related norm

$$\|\vec{u}\| \equiv \sqrt{(\vec{u}, \vec{u})}, \tag{1.1.11}$$

then the corresponding Schwarz and triangular inequalities can be proven in an identical manner yielding

$$|(\vec{u}, \vec{v})| \leq \|\vec{u}\| \, \|\vec{v}\| \, , \tag{1.1.12}$$

and

$$\|\vec{u} + \vec{v}\| \leq \|\vec{u}\| + \|\vec{v}\| \, . \tag{1.1.13}$$

1.2 Orthonormal System of Functions

Two functions $f(x)$ and $g(x)$ are said to be *orthogonal* if their inner product vanishes, i.e.,

$$(f, g) = \int_a^b f^*(x) g(x) \, dx = 0. \tag{1.2.1}$$

A function is said to be *normalized* if its norm equals to unity, i.e.,

$$\|f\| = \sqrt{(f, f)} = 1. \tag{1.2.2}$$

Consider now a set of normalized functions $\{\phi_1(x), \phi_2(x), \phi_3(x), \ldots\}$ which are mutually orthogonal. Such a set is called an *orthonormal set of functions*, satisfying the orthonormality condition

$$(\phi_i, \phi_j) = \delta_{ij} = \begin{cases} 1, & \text{if } \; i = j, \\ 0, & \text{otherwise}, \end{cases} \tag{1.2.3}$$

where δ_{ij} is the *Kronecker delta symbol* itself defined by Eq. (1.2.3).

An orthonormal set of functions $\{\phi_n(x)\}$is said to form a *basis for a function space*, or to be *complete*, if any function $f(x)$ in that space can be expanded in a series of the form

$$f(x) = \sum_{n=1}^{\infty} a_n \phi_n(x). \tag{1.2.4}$$

(This is not the exact definition of a complete set but it will do for our purposes.) To find the coefficients of the expansion in Eq. (1.2.4), we take the inner product of both sides with $\phi_m(x)$ from the left to obtain

$$\begin{aligned} (\phi_m, f) &= \sum_{n=1}^{\infty} (\phi_m, a_n \phi_n) \\ &= \sum_{n=1}^{\infty} a_n \, (\phi_m, \phi_n) \\ &= \sum_{n=1}^{\infty} a_n \delta_{mn} = a_m. \end{aligned} \tag{1.2.5}$$

In other words, for any n,

$$a_n = (\phi_n, f) = \int_a^b \phi_n^* (x)\, f\,(x)\ \, dx. \tag{1.2.6}$$

An example of an orthonormal system of functions on the interval $(-l, l)$ is the infinite set

$$\phi_n\,(x) = \frac{1}{\sqrt{2l}} \exp\left[\frac{in\pi x}{l}\right], \quad n = 0, \pm 1, \pm 2, \ldots \tag{1.2.7}$$

with which the expansion of a *square-integrable* function $f(x)$ on $(-l, l)$ takes the form

$$f(x) = \sum_{n=-\infty}^{\infty} c_n \exp\left[\frac{in\pi x}{l}\right], \tag{1.2.8a}$$

with

$$c_n = \frac{1}{2l} \int_{-l}^{+l} f(x) \exp\left[-\frac{in\pi x}{l}\right], \tag{1.2.8b}$$

which is the familiar complex form of the *Fourier series* of $f(x)$.

Finally the *Dirac delta function* $\delta\,(x - x')$, defined with x and x' in (a, b), can be expanded in terms of a complete set of orthonormal functions $\phi_n(x)$ in the form

$$\delta\,(x - x') = \sum_n a_n \phi_n(x)$$

with

$$a_n = \int_a^b \phi_n^*(x)\delta(x - x')\ dx = \phi_n^*\,(x')\,.$$

That is,

$$\delta(x - x') = \sum_n \phi_n^*(x')\phi_n(x). \tag{1.2.9}$$

The expression (1.2.9) is sometimes taken as the statement which implies the *completeness of an orthonormal system of functions*.

1.3 Linear Operators

An *operator* can be thought of as a mapping or a transformation which acts on a member of the function space (i.e., a function) to produce another member of that space (i.e., another function). The operator, typically denoted by a symbol such as L, is said to be *linear* if it satisfies

$$L(\alpha f + \beta g) = \alpha L f + \beta L g, \tag{1.3.1}$$

where α and β are complex numbers, and f and g are members of that function space.

Some trivial examples of linear operators L are

(i) multiplication by a constant scalar, i.e.,

$$L\phi = a\phi,$$

(ii) taking the third derivative of a function, i.e.,

$$L\phi = \frac{d^3}{dx^3}\phi \quad \text{or} \quad L = \frac{d^3}{dx^3},$$

which is a differential operator, or,

(iii) multiplying a function by the kernel, $K(x, x')$, and integrating over (a, b) with respect to x', i.e.,

$$L\phi(x) = \int_a^b K(x, x')\phi(x')\, dx',$$

which is an integral operator.

An important concept in the theory of the linear operator is that of the *adjoint* of the operator which is defined as follows. Given the operator L, together with an inner product defined on a vector space, the adjoint L^{adj} of the operator L is that operator for which

$$(\psi, L\phi) = (L^{\text{adj}}\psi, \phi), \tag{1.3.2}$$

is an identity for any two members ϕ and ψ of the vector space. Actually, as we shall see later, in the case of the differential operators, we frequently need to worry to some extent about the boundary conditions associated with the original and the adjoint problems. Indeed, there often arise additional terms on the right-hand side of Eq. (1.3.2) which involve the boundary points, and a prudent choice of the adjoint boundary conditions will need to be made in order to avoid unnecessary difficulties. These issues will be raised in connection with Green's functions for differential equations.

As our first example of the adjoint operator, consider the liner vector space of n-dimensional complex column vectors $\vec{u}$, $\vec{v}$, ... with their associated inner product (1.1.10). In this space, $n\times\ n$ square matrices A, B, ... with complex entries are linear operators when multiplied by the n-dimensional complex column vectors according to the usual rules of matrix multiplication. Consider now the problem of finding the adjoint A^{adj} of the matrix A. According to the definition (1.3.2) of the adjoint operator, we search for the matrix A^{adj} satisfying

$$(\vec{u}, A\vec{v}) = (A^{\text{adj}}\vec{u}, \vec{v}). \tag{1.3.3}$$

Now, from the definition of the inner product (1.1.10), we must have

$$\vec{u}^{*\text{T}}(A^{\text{adj}})^{*\text{T}}\vec{v} = \vec{u}^{*\text{T}}A\vec{v},$$

i.e.,

$$(A^{\text{adj}})^{*\text{T}} = A \quad \text{or} \quad A^{\text{adj}} = A^{*\text{T}}. \tag{1.3.4}$$

That is, the adjoint A^{adj} of a matrix A is equal to the complex conjugate of its transpose, which is also known as its *Hermitian transpose*,

$$A^{\text{adj}} = A^{*\text{T}} \equiv A^{\text{H}}. \tag{1.3.5}$$

As a second example, consider the problem of finding the adjoint of the linear integral operator

$$L = \int_a^b dx' K(x, x'), \tag{1.3.6}$$

on our function space. By definition, the adjoint L^{adj} of L is the operator which satisfies Eq. (1.3.2). Upon expressing the left-hand side of Eq. (1.3.2) explicitly with the operator L given by Eq. (1.3.6), we find

$$(\psi, L\phi) = \int_a^b dx\, \psi^*(x) L\phi(x) = \int_a^b dx' \left[\int_a^b dx K(x, x') \psi^*(x) \right] \phi(x'). \tag{1.3.7}$$

Requiring Eq. (1.3.7) to be equal to

$$(L^{\text{adj}}\psi, \phi) = \int_a^b dx (L^{\text{adj}}\psi(x))^* \phi(x)$$

necessitates defining

$$L^{\text{adj}}\psi(x) = \int_a^b d\xi K^*(\xi, x)\psi(\xi).$$

Hence the adjoint of integral operator (1.3.6) is found to be

$$L^{\text{adj}} = \int_a^b dx' K^*(x', x). \tag{1.3.8}$$

Note that, aside from the complex conjugation of the kernel $K(x, x')$, the integration in Eq. (1.3.6) is carried out with respect to the second argument of $K(x, x')$ while that in Eq. (1.3.8) is carried out with respect to the first argument of $K^*(x', x)$. Also, be careful to note which of the variables throughout the above is the dummy variable of integration.

Before we end this section, let us define what is meant by a *self-adjoint* operator. An operator L is said to be self-adjoint (or *Hermitian*) if it is equal to its own adjoint L^{adj}. Hermitian operators have very nice properties which will be discussed in Section 1.6. Not the least of these is that their eigenvalues are real. (Eigenvalue problems are discussed in the next section.)

Examples of self-adjoint operators are Hermitian matrices, i.e., matrices which satisfies

$$A = A^{\text{H}},$$

and linear integral operators of the type (1.3.6) whose kernel satisfy

$$K(x, x') = K^*(x', x),$$

each on their respective linear spaces and with their respective inner products.

1.4 Eigenvalues and Eigenfunctions

Given a linear operator L on a linear vector space, we can set up the following eigenvalue problem

$$L\phi_n = \lambda_n \phi_n \quad (n = 1, 2, 3, \ldots). \tag{1.4.1}$$

Obviously the trivial solution $\phi(x) = 0$ always satisfies this equation, but it also turns out that for some particular values of λ (called the *eigenvalues* and denoted by λ_n), nontrivial solutions to Eq. (1.4.1) also exist. Note that for the case of the differential operators on bounded domains, we must also specify an appropriate homogeneous boundary condition (such that $\phi = 0$ satisfies those boundary conditions) for the *eigenfunctions* $\phi_n(x)$. We have affixed the subscript n to the eigenvalues and eigenfunctions under the assumption that the eigenvalues are discrete and that they can be counted (i.e., with $n = 1, 2, 3, \ldots$). This is not always the case. The conditions which guarantee the existence of a discrete (and complete) set of eigenfunctions are beyond the scope of this introductory chapter and will not be discussed.

So, for the moment, let us tacitly assume that the eigenvalues λ_n of Eq. (1.4.1) are discrete and that their eigenfunctions ϕ_n form a basis (i.e., a complete set) for their space.

Similarly the adjoint L^{adj} of the operator L would posses a set of eigenvalues and eigenfunctions satisfying

$$L^{\text{adj}}\psi_m = \mu_m \psi_m \quad (m = 1, 2, 3, \ldots). \tag{1.4.2}$$

It can be shown that the eigenvalues μ_m of the adjoint problem are equal to complex conjugates of the eigenvalues λ_n of the original problem. (We will prove this only for matrices but it remains true for general operators.) That is, if λ_n is an eigenvalue of L, λ_n^* is an eigenvalue of L^{adj}. This prompts us to rewrite Eq. (1.4.2) as

$$L^{\text{adj}}\psi_m = \lambda_m^* \psi_m, \quad (m = 1, 2, 3, \ldots). \tag{1.4.3}$$

It is then a trivial matter to show that the eigenfunctions of the adjoint and original operators are all orthogonal, except those corresponding to the same index ($n = m$). To do this, take the inner product of Eq. (1.4.1) with ψ_m from the left, and the inner product of Eq. (1.4.3) with ϕ_n from the right, to find

$$(\psi_m, L\phi_n) = (\psi_m, \lambda_n \phi_n) = \lambda_n (\psi_m, \phi_n) \tag{1.4.4}$$

and

$$(L^{\text{adj}}\psi_m, \phi_n) = (\lambda_m^* \psi_m, \phi_n) = \lambda_m (\psi_m, \phi_n). \tag{1.4.5}$$

Subtract the latter two equations and note that their left-hand sides are equal because of the definition of the adjoint, to get

$$0 = (\lambda_n - \lambda_m)(\psi_m, \phi_n). \tag{1.4.6}$$

This implies

$$(\psi_m, \phi_n) = 0 \quad \text{if} \quad \lambda_n \neq \lambda_m, \tag{1.4.7}$$

which proves the desired result. Also, since each ϕ_n and ψ_m is determined to within a multiplicative constant (e.g., if ϕ_n satisfies Eq. (1.4.1) so does $\alpha\phi_n$), the normalization for the latter can be chosen such that

$$(\psi_m, \phi_n) = \delta_{mn} = \begin{cases} 1, & \text{for } n = m, \\ 0, & \text{otherwise.} \end{cases} \tag{1.4.8}$$

Now, if the set of eigenfunctions ϕ_n $(n = 1, 2, \ldots)$ forms a complete set, any arbitrary function $f(x)$ in the space may be expanded as

$$f(x) = \sum_n a_n \phi_n(x), \tag{1.4.9}$$

and to find the coefficients a_n, we simply take the inner product of both sides with ψ_k to get

$$\begin{aligned}(\psi_k, f) = \sum_n (\psi_k, a_n\phi_n) &= \sum_n a_n(\psi_k, \phi_n) \\ &= \sum_n a_n \delta_{kn} = a_k,\end{aligned}$$

i.e.,

$$a_n = (\psi_n, f), \quad (n = 1, 2, 3, \ldots). \tag{1.4.10}$$

Note the difference between Eqs. (1.4.9) and (1.4.10) and the corresponding formulas (1.2.4) and (1.2.6) for an orthonormal system of functions. In the present case, neither $\{\phi_n\}$ nor $\{\psi_n\}$ form an orthonormal system, but they are orthogonal to one another.

Proof that the eigenvalues of the adjoint matrix are complex conjugates of the eigenvalues of the original matrix.

Above, we claimed without justification that the eigenvalues of the adjoint of an operator are complex conjugates of those of the original operator. Here we show this for the matrix case. The eigenvalues of a matrix A are given by

$$\det(A - \lambda I) = 0. \tag{1.4.11}$$

The latter is the characteristic equation whose n solutions for λ are the desired eigenvalues. On the other hand, the eigenvalues of A^{adj} are determined by setting

$$\det(A^{\text{adj}} - \mu I) = 0. \tag{1.4.12}$$

Since the determinant of a matrix is equal to that of its transpose, we easily conclude that the eigenvalues of A^{adj} are the complex conjugates of λ_n. □

1.5 The Fredholm Alternative

The *Fredholm Alternative*, which may be also called the *Fredholm solvability condition*, is concerned with the existence of the solution $y(x)$ of the inhomogeneous problem

$$Ly(x) = f(x), \tag{1.5.1}$$

where L is a given linear operator and $f(x)$ a known forcing term. As usual, if L is a differential operator, additional boundary or initial conditions must also be specified.

The Fredholm Alternative *states that the unknown function $y(x)$ can be determined uniquely if the corresponding homogeneous problem*

$$L\phi_H(x) = 0 \tag{1.5.2}$$

with homogeneous boundary conditions, has no nontrivial solutions. On the other hand, if the homogeneous problem (1.5.2) *does possess a nontrivial solution, then the inhomogeneous problem* (1.5.1) *has either no solution or infinitely many solutions.*

What determines the latter is the homogeneous solution ψ_H to the adjoint problem

$$L^{\text{adj}}\psi_H = 0. \tag{1.5.3}$$

Taking the inner product of Eq. (1.5.1) with ψ_H from the left,

$$(\psi_H, Ly) = (\psi_H, f).$$

Then, by the definition of the adjoint operator (excluding the case wherein L is a differential operator, to be discussed in Section 1.7.), we have

$$(L^{\text{adj}}\psi_H, y) = (\psi_H, f).$$

The left-hand side of the equation above is zero by the definition of ψ_H, Eq. (1.5.3).

Thus the criteria for the solvability of the inhomogeneous problem Eq. (1.5.1) *is given by*

$$(\psi_H, f) = 0.$$

If these criteria are satisfied, there will be an infinity of solutions to Eq. (1.5.1)*, otherwise Eq.* (1.5.1) *will have no solution.*

To understand the above claims, let us suppose that L and L^{adj} possess complete sets of eigenfunctions satisfying

$$L\phi_n = \lambda_n\phi_n \quad (n = 0, 1, 2, \ldots), \tag{1.5.4a}$$

$$L^{\text{adj}}\psi_n = \lambda_n^*\psi_n \quad (n = 0, 1, 2, \ldots), \tag{1.5.4b}$$

with

$$(\psi_m, \phi_n) = \delta_{mn}. \tag{1.5.5}$$

The existence of a nontrivial homogeneous solution $\phi_H\,(x)$ to Eq. (1.5.2), as well as $\psi_H\,(x)$ to Eq. (1.5.3), is the same as having one of the eigenvalues λ_n in Eqs. (1.5.4a), (1.5.4b) to be zero. If this is the case, i.e., if zero is an eigenvalue of Eq. (1.5.4a) and hence Eq. (1.5.4b), we shall choose the subscript $n = 0$ to signify that eigenvalue ($\lambda_0 = 0$), and in that case

ϕ_0 and ψ_0 are the same as ϕ_H and ψ_H. The two circumstances in the Fredholm Alternative correspond to cases where zero is an eigenvalue of Eqs. (1.5.4a), (1.5.4b) and where it is not.

Let us proceed formally with the problem of solving the inhomogeneous problem Eq. (1.5.1). Since the set of eigenfunctions ϕ_n of Eq. (1.5.4a) is assumed to be complete, both the known function $f(x)$ and the unknown function $y(x)$ in Eq. (1.5.1) can presumably be expanded in terms of $\phi_n(x)$:

$$f(x) = \sum_{n=0}^{\infty} \alpha_n \phi_n(x), \tag{1.5.6}$$

$$y(x) = \sum_{n=0}^{\infty} \beta_n \phi_n(x), \tag{1.5.7}$$

where the α_n are known (since $f(x)$ is known), i.e., according to Eq. (1.4.10)

$$\alpha_n = (\psi_n, f), \tag{1.5.8}$$

while the β_n are unknown. Thus, if all the β_n can be determined, then the solution $y(x)$ to Eq. (1.5.1) is regarded as having been found.

To try to determine the β_n, substitute both Eqs. (1.5.6) and (1.5.7) into Eq. (1.5.1) to find

$$\sum_{n=0}^{\infty} \lambda_n \beta_n \phi_n = \sum_{k=0}^{\infty} \alpha_k \phi_k, \tag{1.5.9}$$

where different summation indices have been used on the two sides to remind the reader that the latter are dummy indices of summation. Next, take the inner product of both sides with ψ_m (with an index which must be different from the two above) to get

$$\sum_{n=0}^{\infty} \lambda_n \beta_n (\psi_m, \phi_n) = \sum_{k=0}^{\infty} \alpha_k (\psi_m, \phi_k),$$

or

$$\sum_{n=0}^{\infty} \lambda_n \beta_n \delta_{mn} = \sum_{k=0}^{\infty} \alpha_k \delta_{mk},$$

i.e.,

$$\lambda_m \beta_m = \alpha_m. \tag{1.5.10}$$

Thus, for any $m = 0, 1, 2, \ldots$, we can solve Eq. (1.5.10) for the unknowns β_m to get

$$\beta_n = \alpha_n / \lambda_n \quad (n = 0, 1, 2, \ldots), \tag{1.5.11}$$

provided that λ_n is not equal to zero. Obviously the only possible difficulty occurs if one of the eigenvalues (which we take to be λ_0) is equal to zero. In that case, equation (1.5.10) with $m = 0$ reads

$$\lambda_0 \beta_0 = \alpha_0 \quad (\lambda_0 = 0). \tag{1.5.12}$$

Now if $\alpha_0 \neq 0$, then we cannot solve for β_0 and thus the problem $Ly = f$ has no solution. On the other hand if $\alpha_0 = 0$, i.e., if

$$(\psi_0, f) = (\psi_H, f) = 0, \tag{1.5.13}$$

implying that f is orthogonal to the homogeneous solution to the adjoint problem, then Eq. (1.5.12) is satisfied by any choice of β_0. All the other β_n ($n = 1, 2, \ldots$) are uniquely determined but there are infinitely many solutions $y(x)$ to Eq. (1.5.1) corresponding to the infinitely many values possible for β_0. The reader must make certain that he or she understands the equivalence of the above with the original statement of the Fredholm Alternative.

1.6 Self-adjoint Operators

Operators which are self-adjoint or Hermitian form a very useful class of operators. They possess a number of special properties, some of which are described in this section.

The first important property of self-adjoint operators is that their *eigenvalues are real*. To prove this, begin with

$$\begin{aligned} L\phi_n &= \lambda_n \phi_n, \\ L\phi_m &= \lambda_m \phi_m, \end{aligned} \tag{1.6.1}$$

and take the inner product of both sides of the former with ϕ_m from the left, and the latter with ϕ_n from the right, to obtain

$$\begin{aligned} (\phi_m, L\phi_n) &= \lambda_n (\phi_m, \phi_n), \\ (L\phi_m, \phi_n) &= \lambda_m^* (\phi_m, \phi_n). \end{aligned} \tag{1.6.2}$$

For a self-adjoint operator $L = L^{\text{adj}}$, the two left-hand sides of Eq. (1.6.2) are equal and hence, upon subtraction of the latter from the former, we find

$$0 = (\lambda_n - \lambda_m^*)(\phi_m, \phi_n). \tag{1.6.3}$$

Now, if $m = n$, the inner product $(\phi_n, \phi_n) = \|\phi_n\|^2$ is nonzero and Eq. (1.6.3) implies

$$\lambda_n = \lambda_n^*, \tag{1.6.4}$$

proving that all the eigenvalues are real. Thus Eq. (1.6.3) can be rewritten as

$$0 = (\lambda_n - \lambda_m)(\phi_m, \phi_n), \tag{1.6.5}$$

indicating that if $\lambda_n \neq \lambda_m$, then the eigenfunctions ϕ_m and ϕ_n are orthogonal. Thus, upon normalizing each ϕ_n, we verify a second important property of self-adjoint operators that (upon normalization) the *eigenfunctions of a self-adjoint operator form an orthonormal set*.

The Fredholm Alternative can also be restated for a self-adjoint operator L in the following form: The inhomogeneous problem $Ly = f$ (with L self-adjoint) is solvable for y, if f is orthogonal to all eigenfunctions ϕ_0 of L with eigenvalue zero (if indeed any exist). If zero is not an eigenvalue of L, the solution is unique. Otherwise, there is no solution if $(\phi_0, f) \neq 0$, and an infinite number of solutions if $(\phi_0, f) = 0$.

Diagonalization of Self-adjoint Operators: Any linear operator can be expanded in some sense in terms of any orthonormal basis set. To elaborate on this, suppose that the orthonormal system $\{e_i(x)\}_i$, with $(e_i, e_j) = \delta_{ij}$ forms a complete set. Any function $f(x)$ can be expanded as

$$f(x) = \sum_{j=1}^{\infty} \alpha_j e_j(x), \quad \alpha_j = (e_j, f). \tag{1.6.6}$$

Thus the function $f(x)$ can be thought of as an infinite dimensional vector with components α_j. Now consider the action of an arbitrary linear operator L on the function $f(x)$. Obviously

$$Lf(x) = \sum_{j=1}^{\infty} \alpha_j L e_j(x). \tag{1.6.7}$$

But L acting on $e_j(x)$ is itself a function of x which can be expanded in the orthonormal basis $\{e_i(x)\}_i$. Thus we write

$$Le_j(x) = \sum_{i=1}^{\infty} l_{ij} e_i(x), \tag{1.6.8}$$

wherein the coefficients l_{ij} of the expansion are found to be $l_{ij} = (e_i, Le_j)$. Substitution of Eq. (1.6.8) into Eq. (1.6.7) then shows

$$Lf(x) = \sum_{i=1}^{\infty} \left(\sum_{j=1}^{\infty} l_{ij} \alpha_j \right) e_i(x). \tag{1.6.9}$$

We discover that just as we can think of $f(x)$ as the infinite dimensional vector with components α_j, we can consider L to be equivalent to an infinite dimensional matrix with components l_{ij}, and we can regard Eq. (1.6.9) as a regular multiplication of the matrix L (components l_{ij}) with the vector f (components α_j). However, this equivalence of the operator L with the matrix whose components are l_{ij}, i.e., $L \Leftrightarrow l_{ij}$, depends on the choice of the orthonormal set.

For a self-adjoint operator $L = L^{\text{adj}}$, the most natural choice of the basis set is the set of eigenfunctions of L. Denoting these by $\{\phi_i(x)\}_i$, the components of the equivalent matrix for L take the form

$$l_{ij} = (\phi_i, L\phi_j) = (\phi_i, \lambda_j \phi_j) = \lambda_j (\phi_i, \phi_j) = \lambda_j \delta_{ij}. \tag{1.6.10}$$

1.7 Green's Functions for Differential Equations

In this section, we describe the conceptual basis of the theory of *Green's functions*. We do this by first outlining the abstract themes involved and then by presenting a simple example. More complicated examples will appear in later chapters.

Prior to discussing Green's functions, recall some of the elementary properties of the so-called Dirac delta function $\delta(x - x')$. In particular, remember that if x' is inside the domain