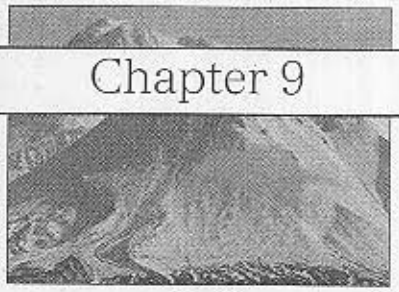
Chapter 9



Sturm–Liouville Theory—Orthogonal Functions

In the preceding chapter, we developed two linearly independent solutions of the second-order linear homogeneous differential equation and proved that no third linearly independent solution existed. In this chapter, the emphasis shifts from solving the differential equation to developing and understanding general properties of the set of solutions. There is a close analogy between the concepts in this chapter and those of linear algebra in Chapter 3. Functions here play the role of vectors there, and linear operators play that of matrices in Chapter 3. The diagonalization of a real symmetric matrix in Chapter 3 corresponds here to the solution of an ordinary differential equation (ODE), defined by a **self-adjoint** operator $\mathcal{L}$, in terms of its eigenfunctions, which are the "continuous" analog of the eigenvectors in Chapter 3, and real eigenvalues that correspond to physically observable quantities in the laboratory. Just as a column eigenvector vector $\mathbf{a}$ is written as $|a\rangle$ in the Dirac notation of Chapter 3, we now write an eigenfunction as $|\varphi\rangle$. In the Cartesian component $a_i = \hat{\mathbf{x}}_i \cdot \mathbf{a}$, the discrete index i of the coordinate unit vectors is now replaced by the continuous variable x in $\varphi(x)$.

In Section 9.1, the concepts of self-adjoint operator, eigenfunction, eigenvalue, and Hermitian operator are presented. The concept of adjoint operator, given first in terms of differential equations, is then redefined in accordance with usage in quantum mechanics. The vital properties of reality of eigenvalues and orthogonality of eigenfunctions are derived in Section 9.2. In Section 9.3, we discuss the Gram–Schmidt procedure for systematically constructing sets of orthogonal functions. Finally, the general property of the completeness of a set of eigenfunctions is explored in Section 9.4.

9.1 Self-Adjoint ODEs

In Chapter 8, we studied, classified, and solved linear, second-order ODEs corresponding to linear, second-order differential operators of the general form

$$\mathcal{L} = p_0(x)\frac{d^2}{dx^2} + p_1(x)\frac{d}{dx} + p_2(x), \tag{9.1}$$

defined over the region $a \le x \le b$. A number of restrictions apply. The coefficients $p_0(x)$, $p_1(x)$, and $p_2(x)$ are real functions of x, and the first $2 - i$ derivatives of $p_i(x)$ are continuous. Reference to Eqs. (8.42) and (8.44) shows that $P(x) = p_1(x)/p_0(x)$ and $Q(x) = p_2(x)/p_0(x)$. Hence, $p_0(x)$ must not vanish for $a < x < b$. The zeros of $p_0(x)$ are singular points (Section 8.4), and the preceding statement means that our interval $[a, b]$ must be such that there are no singular points in the interior of the interval. There may be and often are singular points on the boundaries. Moreover, $b \to \infty$ and/or $a \to -\infty$ are possible in certain problems.

For a linear operator $\mathcal{L}$, the analog of a quadratic form for a matrix in Chapter 3 is the integral in Dirac's notation of Section 3.2:

$$\begin{aligned}\langle u|\mathcal{L}|u\rangle \equiv \langle u|\mathcal{L}u\rangle &\equiv \int_a^b u^*(x)\mathcal{L}u(x)\,dx \\ &= \int_a^b u(p_0u'' + p_1u' + p_2u)\,dx,\end{aligned} \tag{9.2}$$

taking $u = u^*$ to be real. This integral is the continuous analog of the **inner product** of vectors in Chapter 1 and here of u and $\mathcal{L}u$. Two vectors $u(x)$, $v(x)$ are **orthogonal** if their inner product vanishes,

$$\langle v|u\rangle \equiv \int_a^b v^*(x)u(x)\,dx = 0.$$

If we shift the derivatives to the first factor u in Eq. (9.2) by integrating by parts once or twice, we are led to the equivalent expression

$$\begin{aligned}\langle u|\mathcal{L}u\rangle &= [u(x)(p_1 - p_0')u(x)]_{x=a}^b \\ &\quad + \int_a^b \left[\frac{d^2}{dx^2}(p_0u) - \frac{d}{dx}(p_1u) + p_2u\right]u\,dx \\ &= [u(x)(p_1 - p_0')u(x)]_{x=a}^b + \left\langle\left[\frac{d^2}{dx^2}(p_0u) - \frac{d}{dx}(p_1u) + p_2u\right]\middle| u\right\rangle.\end{aligned} \tag{9.3}$$

For Eqs. (9.2) and (9.3) to agree for all u, the integrands have to be equal. The comparison yields

$$u(p_0'' - p_1')u + 2u(p_0' - p_1)u' = 0$$

or

$$p_0'(x) = p_1(x). \tag{9.4}$$

The terms at the boundary $x = a$ and $x = b$ in Eq. (9.3) then also vanish.

Because of the analogy with the transpose matrix in Chapter 3, it is convenient to define the linear operator in Eq. (9.3)

$$\begin{aligned}\tilde{\mathcal{L}}u &= \frac{d^2}{dx^2}(p_0 u) - \frac{d}{dx}(p_1 u) + p_2 u \\ &= p_0\frac{d^2u}{dx^2} + (2p_0' - p_1)\frac{du}{dx} + (p_0'' - p_1' + p_2)u\end{aligned} \tag{9.5}$$

as the **adjoint**[1] **operator** $\tilde{\mathcal{L}}$ so that $\langle u|\mathcal{L}u\rangle = \langle\tilde{\mathcal{L}}u|u\rangle$ for all u. The necessary and sufficient condition that $\tilde{\mathcal{L}} = \mathcal{L}$, or $\langle u|\mathcal{L}u\rangle = \langle\mathcal{L}u|u\rangle$, is that Eq. (9.4) be satisfied for all u. When this condition is satisfied,

$$\tilde{\mathcal{L}}u = \mathcal{L}u = \frac{d}{dx}\left[p(x)\frac{du(x)}{dx}\right] + q(x)u(x), \tag{9.6}$$

where $p(x) = p_0(x)$ and $q(x) = p_2(x)$, and the operator $\mathcal{L}$ is said to be **self-adjoint**. The importance of the form of Eq. (9.6) is that we will be able to carry out two integrations by parts in Eq. (9.3).[2]

The ODEs introduced in Section 8.3, Legendre's equation, and the linear oscillator equation are self-adjoint, but others, such as the Laguerre and Hermite equations, are not. However, the theory of linear, second-order, self-adjoint differential equations is perfectly general because we can **always** transform the non-self-adjoint operator into the required self-adjoint form. Consider Eq. (9.1) with $p_0' \neq p_1$. If we multiply $\mathcal{L}$ by[3]

$$\frac{1}{p_0(x)}\exp\left[\int^x \frac{p_1(t)}{p_0(t)}dt\right],$$

[1]The **adjoint** operator bears a somewhat forced relationship to the **transpose** matrix. A better justification for the nomenclature is found in a comparison of the **self-adjoint** operator (plus appropriate boundary conditions) with the **symmetric** matrix. These significant properties are developed in Section 9.2. Because of these properties, we are interested in **self-adjoint** operators. When adjoint or self-adjoint operators are discussed in the context of a Hilbert space, all functions of that space will satisfy the boundary conditions.

[2]The importance of the self-adjoint form (plus boundary conditions) will become apparent in Section 9.2, Eq. (9.22) and after.

[3]If we multiply $\mathcal{L}$ by $f(x)/p_0(x)$ and then demand that

$$f'(x) = \frac{fp_1}{p_0}$$

so that the new operator will be self-adjoint, we obtain

$$f(x) = \exp\left[\int^x \frac{p_1(t)}{p_0(t)}dt\right].$$

we obtain

$$\frac{1}{p_0(x)}\exp\left[\int^x \frac{p_1(t)}{p_0(t)}dt\right]\mathcal{L}u(u) = \frac{d}{dx}\left\{\exp\left[\int^x \frac{p_1(t)}{p_0(t)}dt\right]\frac{du(x)}{dx}\right\} + \frac{p_2(x)}{p_0(x)}\cdot\exp\left[\int^x \frac{p_1(t)}{p_0(t)}dt\right]u, \quad (9.7)$$

which is clearly self-adjoint [see Eq. (9.6)]. Notice the $p_0(x)$ in the denominator. This is why we require $p_0(x) \neq 0$, $a < x < b$. In the following development, we assume that $\mathcal{L}$ has been put into self-adjoint form.

Eigenfunctions and Eigenvalues

Schrödinger's time-independent wave equation for a single-particle system,

$$H\psi(x) = E\psi(x),$$

is a major example of an eigenvalue equation in physics; here, the differential operator $\mathcal{L}$ is defined by the Hamiltonian H and the eigenvalue is the total energy E of the system. The eigenfunction $\psi(x)$ is usually called a wave function. A variational derivation of this Schrödinger equation appears in Section 18.5. Based on spherical, cylindrical, or some other symmetry properties, a three- or four-dimensional partial differential equation (PDE) or eigenvalue equation, such as the Schrödinger equation, often separates into three (or more) eigenvalue equations in a single variable. In this context, an eigenvalue equation sometimes takes the more general self-adjoint form,

$$\mathcal{L}u(x) + \lambda w(x)u(x) = 0, \quad \text{or} \quad \mathcal{L}|u\rangle + \lambda w|u\rangle = 0, \quad (9.8)$$

where the constant λ is the eigenvalue, $\mathcal{L}$ is self-adjoint, and $w(x)$ is a known **weight** or density function; $w(x) > 0$ except possibly at isolated points at which $w(x) = 0$. The Schrödinger equation for the simple harmonic oscillator is a particular case of Eq. (9.8) with $w = 1$. The analysis of Eq. (9.8) with $\mathcal{L}$ as defined in Eq. (9.6) and its solutions is called **Sturm–Liouville theory**. For a given choice of the parameter λ, a function $u_\lambda(x)$, which satisfies Eq. (9.8) **and the imposed boundary conditions** discussed later, is called an **eigenfunction** corresponding to λ. The constant λ is then called an eigenvalue. There is no guarantee that an eigenfunction $u_\lambda(x)$ will exist for an arbitrary choice of the parameter λ. Indeed, the requirement that there be an eigenfunction often restricts the acceptable values of λ to a discrete set.

The **inner product** (or overlap integral) of two functions

$$\langle v|u\rangle \equiv \int v^*(x)u(x)w(x)dx$$

depends on the weight function and generalizes our previous definition for $w = 1$ to $w \neq 1$. The latter also modifies the definition of **orthogonality** of two eigenfunctions: They are orthogonal if their overlap integral vanishes, $\langle u_{\lambda'}|u_\lambda\rangle = 0$. Examples of eigenvalues for the Legendre and Hermite equations appear in the exercises of Section 8.5. Here, we have the mathematical approach to the process of quantization in quantum mechanics.

Table 9.1

Equation	$p(x)$	$q(x)$	λ	$w(x)$
Legendre	$1-x^2$	0	$l(l+1)$	1
Shifted Legendre	$x(1-x)$	0	$l(l+1)$	1
Associated Legendre	$1-x^2$	$-m^2/(1-x^2)$	$l(l+1)$	1
Bessel[a]	x	$-\frac{n^2}{x}$	a^2	x
Laguerre	xe^{-x}	0	α	e^{-x}
Associated Laguerre	$x^{k+1}e^{-x}$	0	$\alpha-k$	$x^k e^{-x}$
Hermite	e^{-x^2}	0	2α	e^{-x^2}
Simple harmonic oscillator[b]	1	0	n^2	1

[a] Orthogonality of Bessel functions is special. Compare Section 12.1 for details.
[b] This will form the basis for Chapter 14, Fourier series.

The extra weight function $w(x)$ sometimes appears as an asymptotic wave function ψ_∞ that is a common factor in all solutions of a PDE, such as the Schrödinger equation, for example, when the potential $V(x) \to 0$ as $x \to \infty$ in $H = T + V$. We find ψ_∞ when we set $V = 0$ in the Schrödinger equation. Another source for $w(x)$ may be a nonzero angular momentum barrier of the form $l(l+1)/x^2$ in a PDE or separated ODE that has a regular singularity and dominates at $x \to 0$. In such a case the indicial equation, such as Eqs. (8.65) or (8.88), shows that the wave function has x^l as an overall factor. Since the wave function enters twice in matrix elements $\langle v|Hu\rangle$ and orthogonality relations $\langle v|u\rangle$, the weight functions in Table 9.1 come from these common factors in both radial wave functions. This is the physical reason why the $\exp(-x)$ for Laguerre polynomials and $x^k \exp(-x)$ for associated Laguerre polynomials in Table 9.1 arise and is explained in more detail in the next example. The mathematical reason is that the weight function is needed to make the ODE self-adjoint.

EXAMPLE 9.1.1

Asymptotic Behavior, Weight Function Let us look at the asymptotic forms for small and large r of the radial Schrödinger equation for a particle of mass m moving in a spherically symmetric potential $V(r)$,

$$\left(-\frac{\hbar^2}{2m}\frac{d^2}{dr^2} + \frac{\hbar^2}{2m}\frac{l(l+1)}{r^2} + V(r) - E\right) u_l(r) = 0,$$

where $R_l(r) = u_l(r)/r$ is the radial wave function, E is the energy eigenvalue, and l is the orbital angular momentum (see Exercises 2.5.12 and 2.5.14). From the asymptotic ODEs we derive the asymptotic forms of the radial wave function. The boundary conditions are that $u_l(0) = 0$ and $u_l(r) \to 0$ for large r. Let us explain these boundary conditions for $l = 0$. We ask that $u_0 \to 0$ as $r \to 0$ because if $u_0(0)$ diverged the wave function could not be normalized; that is, $\langle u_0|u_0\rangle = 1$ cannot be enforced. If $u_0(0)$ is a finite nonzero constant, then $R_0(r) \sim 1/r$ for $r \to 0$. In that case, the kinetic energy, $\sim \nabla^2 \frac{1}{r} \sim \delta(\mathbf{r})$, would generate a singularity at the origin in the Schrödinger equation.

First, we explore how the angular momentum barrier ($\sim\frac{l(l+1)}{r^2}$) affects the solution for r close to zero, when $l \neq 0$. We assume that the potential is no more singular than a Coulomb potential, that is, $r^2V(r) \to 0$ as $r \to 0$, so that the potential and the energy eigenvalue are negligible compared to the angular momentum barrier for small r. Then the regular solution of the (asymptotic, meaning approximate for small r) radial ODE

$$\left(-\frac{\hbar^2}{2m}\frac{d^2}{dr^2}+\frac{\hbar^2}{2m}\frac{l(l+1)}{r^2}\right)u_l(r)=0$$

is $u_l(r) \sim r^{l+1}$ for small r, the regular solution of this radial ODE, whereas the irregular solution $u_l(r) \sim r^{-l}$ is not finite at small r and is not an acceptable solution. Thus, for $r \to 0$, the radial wave function $R_l(r) \sim r^l$ due to the barrier.

Now we turn our attention to large $r \to \infty$, assuming $V(r) \to 0$, so that we have to solve the asymptotic ODE

$$-\frac{\hbar^2}{2m}\frac{d^2u_l(r)}{dr^2}=Eu_l(r)$$

because the angular momentum barrier is also negligible at large r. For bound states $E < 0$, the solution is $u_l(r) \sim e^{-\kappa r}$ for large r with $E = -\hbar^2\kappa^2/2m$, whereas the other independent solution $e^{\kappa r} \to \infty$ for large r.

This settles the asymptotic behavior of the wave functions, which must have these limiting forms, r^l at small r and $e^{-\kappa r}$ at large r. In Chapter 13, the complete solution will be given in terms of associated Laguerre polynomials $L(r)$ so that $R_l(r) \sim r^l e^{-\kappa r}L(r)$. Therefore, orthonormality integrals $\langle\psi|\psi\rangle$ will contain the weight function $r^{2l+2}e^{-2\kappa r}$ along with a product of Laguerre polynomials, as shown in Table 9.1, except for scaling $2\kappa r \to x$ and renaming $2l+2 \to k$, and the corresponding ODEs are self-adjoint. ■

EXAMPLE 9.1.2

Legendre's Equation Legendre's equation is given by

$$(1-x^2)y''-2xy'+n(n+1)y=0, \tag{9.9}$$

over the interval $-1 \le x \le 1$ and with boundary condition that $y(\pm1)$ is finite. From Eqs. (9.1), (9.8), and (9.9),

$$\begin{aligned} &p_0(x)=1-x^2=p, \quad w(x)=1,\\ &p_1(x)=-2x=p', \quad \lambda=n(n+1),\\ &p_2(x)=0=q. \end{aligned}$$

Recall that our series solutions of Legendre's equation (Exercise 8.5.5)[4] diverged, unless n was restricted to an integer. This also represents a quantization of the eigenvalue λ. ■

[4]Compare also Exercise 5.2.11.

When the equations of Chapter 8 are transformed into the self-adjoint form, we find the values of the coefficients and parameters (Table 9.1). The coefficient $p(x)$ is the coefficient of the second derivative of the eigen-function. The eigenvalue λ is the parameter that is available in a term of the form $\lambda w(x)y(x)$; any x dependence apart from the eigenfunction becomes the weighting function $w(x)$. If there is another term containing the eigenfunction (not the derivatives), the coefficient of the eigenfunction in this additional term is identified as $q(x)$. If no such term is present, $q(x)$ is simply zero.

EXAMPLE 9.1.3

Deuteron Further insight into the concepts of eigenfunction and eigenvalue may be provided by an extremely simple model of the deuteron. The neutron–proton nuclear interaction is represented by a square well potential: $V = V_0 < 0$ for $0 \le r < a$, $V = 0$ for $r \ge a$. The Schrödinger wave equation is

$$-\frac{\hbar^2}{2M}\nabla^2\psi + V\psi = E\psi, \tag{9.10}$$

where $\psi = \psi(r)$ is the probability amplitude for finding a neutron–proton pair at relative distance r. The boundary conditions are $\psi(0)$ finite and $\psi(r) \to 0$ for large r.

We may write $u(r) = r\psi(r)$, and using Exercises 2.5.12 and 2.5.14 the radial wave equation becomes

$$\frac{d^2u}{dr^2} + k_1^2 u = 0, \tag{9.11}$$

with

$$k_1^2 = \frac{2M}{\hbar^2}(E - V_0) > 0 \tag{9.12}$$

for the interior range, $0 \le r < a$. Here, M is the reduced mass of the neutron–proton system. Note that $V_0 < E < 0$ for a bound state, leading to the sign of k_1^2 in Eq. (9.11). For $a \le r < \infty$, we have

$$\frac{d^2u}{dr^2} - k_2^2 u = 0, \tag{9.13}$$

with

$$k_2^2 = -\frac{2ME}{\hbar^2} > 0 \tag{9.14}$$

because $E < 0$ for a bound state with $V \to 0$ as $r \to \infty$. From the boundary condition that ψ remain finite, $u(0) = 0$, and

$$u_{\text{in}}(r) = \sin k_1 r, \quad 0 \le r < a. \tag{9.15}$$

In the range outside the potential well, we have a linear combination of the two exponentials,

$$u_{\text{ex}}(r) = A\exp k_2 r + B\exp(-k_2 r), \quad a \le r < \infty. \tag{9.16}$$

At $r = a$ the solution to the ODE must be continuous, with a continuous derivative, demanding that $u_{\text{in}}(a) = u_{\text{ex}}(a)$ and that $u'_{\text{in}}(a) = u'_{\text{ex}}(a)$. These **joining or matching conditions** give

$$\begin{aligned} \sin k_1 a &= A \exp k_2 a + B \exp(-k_2 a), \\ k_1 \cos k_1 a &= k_2 A \exp k_2 a - k_2 B \exp(-k_2 a). \end{aligned} \tag{9.17}$$

The boundary condition for large r means that $A = 0$. Dividing the preceding pair of equations (to cancel B), we obtain

$$\tan k_1 a = -\frac{k_1}{k_2} = -\sqrt{\frac{E - V_0}{-E}}, \tag{9.18}$$

a transcendental equation for the energy E with only certain discrete solutions. If E is such that Eq. (9.18) can be satisfied, our solutions $u_{\text{in}}(r)$ and $u_{\text{ex}}(r)$ can satisfy the boundary conditions. If Eq. (9.18) is not satisfied, **no acceptable solution exists.** The values of E, for which Eq. (9.18) is satisfied, are the eigenvalues; the corresponding function $\psi(r) = u_{in}/r$ for $r < a$ and $\psi(r) = u_{\text{ex}}/r$ for $r > a$ is the eigenfunction. For the actual deuteron problem, there is one (and only one) negative value of E satisfying Eq. (9.18); that is, the deuteron has one and only one bound state.

Now, what happens if E does not satisfy Eq. (9.18) (i.e., $E \neq E_0$ is not an eigenvalue)? In graphical form, imagine that E and therefore k_1 are varied slightly. For $E = E_1 < E_0$, k_1 is reduced, and $\sin k_1 a$ has not turned down enough to match $\exp(-k_2 a)$. The joining conditions, Eq. (9.17), require $A > 0$ and the wave function goes to $+\infty$, exponentially. For $E = E_2 > E_0$, k_1 is larger, $\sin k_1 a$ peaks sooner and has descended more rapidly at $r = a$. The joining conditions demand $A < 0$, and the wave function goes to $-\infty$ exponentially. Only for $E = E_0$, an eigenvalue, will the wave function have the required negative exponential asymptotic behavior (Fig. 9.1). ■

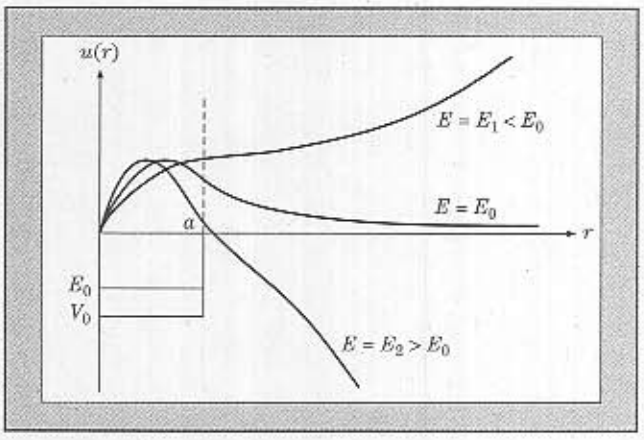


Figure 9.1
Deuteron Wave Functions; Eigenfunction for $E = E$

Boundary Conditions

In the foregoing definition of eigenfunction, it was noted that the eigenfunction $u_\lambda(x)$ was required to satisfy certain imposed boundary conditions. The term boundary conditions includes as a special case the concept of initial conditions, for instance, specifying the initial position x_0 and the initial velocity v_0 in some dynamical problem. The only difference in the present usage of boundary conditions in these one-dimensional problems is that we are going to apply the conditions on **both** ends of the allowed range of the variable to ensure a self-adjoint ODE.

Usually, the form of the differential equation or the boundary conditions on the eigenfunctions will guarantee that, at the ends of our interval (i.e., at the boundary), the following products will vanish:

$$p(x)v^*(x)\frac{du(x)}{dx}\bigg|_{x=a} = 0 \qquad \text{and} \qquad p(x)v^*(x)\frac{du(x)}{dx}\bigg|_{x=b} = 0. \tag{9.19}$$

Here, $u(x)$ and $v(x)$ are solutions of the particular ODE [Eq. (9.8)] being considered. A reason for the particular form of Eq. (9.19) is suggested later. If we recall the radial wave function u of the hydrogen atom in Example 9.1.1 with $u(0) = 0$ and $du/dr \sim e^{-kr} \to 0$ as $r \to \infty$, then both boundary conditions are satisfied. Similarly, in the deuteron Example 9.1.3, $\sin k_1 r \to 0$ as $r \to 0$ and $d(e^{-k_2 r})/dr \to 0$ as $r \to \infty$, both boundary conditions are obeyed.

We can, however, work with a less restrictive set of boundary conditions,

$$v^* pu' \,|_{x=a} = v^* pu' \,|_{x=b}, \tag{9.20}$$

in which $u(x)$ and $v(x)$ are solutions of the differential equation corresponding to the same or to different eigenvalues. Equation (9.20) might well be satisfied if we were dealing with a periodic physical system, such as a crystal lattice.

Equations (9.19) and (9.20) are written in terms of v^*, complex conjugate. When the solutions are real, $v = v^*$ and the asterisk may be ignored. However, in Fourier exponential expansions and in quantum mechanics the functions will be complex and the complex conjugate will be needed.

Hermitian Operators

We now prove an important property of the combination self-adjoint, second-order differential operator [Eq. (9.6)], with functions $u(x)$ and $v(x)$ that satisfy boundary conditions given by Eq. (9.20) and explain the special form of the latter.

By integrating v^* (complex conjugate) times the second-order self-adjoint differential operator $\mathcal{L}$ (operating on u) over the range $a \le x \le b$, we obtain

$$\int_a^b v^* \mathcal{L}u \, dx = \int_a^b v^*(pu')' dx + \int_a^b v^* q u \, dx \tag{9.21}$$

using Eq. (9.6), or in Dirac notation,

$$\langle v|\mathcal{L}u\rangle = \langle v|\left(\frac{d}{dx}p\frac{d}{dx}+q\right)u\rangle.$$

Integrating by parts, we have

$$\int_a^b v^*(pu')'dx = v^*pu'\Big|_a^b - \int_a^b v'^*pu'dx. \tag{9.22}$$

The integrated part vanishes on application of the boundary conditions [Eq. (9.20)]. Integrating the remaining integral by parts a second time, we have

$$-\int_a^b v'^*pu'dx = -v'^*pu\Big|_a^b + \int_a^b u(pv'^*)'dx. \tag{9.23}$$

Again, the integrated part vanishes in an application of Eq. (9.20). A combination of Eqs. (9.21)–(9.23) gives us

$$\langle v|\mathcal{L}u\rangle = \int_a^b v^*\mathcal{L}u\,dx = \int_a^b u\mathcal{L}v^*dx = \langle \mathcal{L}v|u\rangle. \tag{9.24}$$

This property, given by Eq. (9.24), is expressed by stating that the operator $\mathcal{L}$ is Hermitian with respect to the functions $u(x)$ and $v(x)$, which satisfy the boundary conditions specified by Eq. (9.20). Note that if this Hermitian property follows from self-adjointness in a Hilbert space, then it includes that boundary conditions are imposed on all functions of that space. The integral in Eq. (9.24) may also be recognized as **inner product**, $\langle v|\mathcal{L}u\rangle$, of $|v\rangle$ and $|\mathcal{L}u\rangle$.

These properties [Eqs. (9.19) or (9.20)] are so important for the concept of Hermitian operator (discussed next) and the consequences (Section 9.2) that the interval (a, b) must be so as to ensure that either Eq. (9.19) or Eq. (9.20) is **satisfied**. The boundary conditions of the problem determine the range of integration. If our solutions are polynomials, the coefficient $p(x)$ may restrict the range of integration. Note that $p(x)$ also determines the singular points of the differential equation (Section 8.4). For nonpolynomial solutions, for example, $\sin nx$, $\cos nx$; $(p = 1)$, the range of integration is determined by the boundary conditions of each problem, as explained in the next example.

EXAMPLE 9.1.4

Integration Interval, $[a, b]$ For $\mathcal{L} = d^2/dx^2$ a possible eigenvalue equation is

$$\frac{d^2}{dx^2}y(x) + n^2y(x) = 0, \tag{9.25}$$

with eigenfunctions

$$u_n = \cos nx, \qquad v_m = \sin mx.$$

Equation (9.20) becomes

$$-n\sin mx\sin nx\,|_a^b = 0, \qquad \text{or} \quad m\cos mx\cos nx\,|_a^b = 0,$$

interchanging u_n and v_m. Since $\sin mx$ and $\cos nx$ are periodic with period 2π (for n and m integral), Eq. (9.20) is clearly satisfied if $a = x_0$ and $b = x_0 + 2\pi$. If a problem prescribes a different interval, the eigenfunctions and eigenvalues will change along with the boundary conditions. The functions must always be so that the boundary conditions [Eq. (9.20), etc.] are satisfied. For this case (Fourier series), the usual cases are $x_0 = 0$, leading to $(0, 2\pi)$, and $x_0 = -\pi$ leading to $(-\pi, \pi)$. Here and throughout the following several chapters, the relevant functions will satisfy the boundary conditions prescribed by the integration interval [Eq. (9.20)]. The interval $[a, b]$ and the weighting factor $w(x)$ for the most commonly encountered second-order ODEs are listed in Table 9.2. ■

Table 9.2

Equation	a	b	$w(x)$
Legendre	-1	1	1
Shifted Legendre	0	1	1
Associated Legendre	-1	1	1
Laguerre	0	∞	e^{-x}
Associated Laguerre	0	∞	$x^k e^{-x}$
Hermite	$-\infty$	∞	e^{-x^2}
Simple harmonic oscillator	0	2π	1
	$-\pi$	π	1

[1]The orthogonality interval $[a, b]$ is determined by the boundary conditions of Section 9.1. $p(x)$, $q(x)$ are given in Table 9.1.
[2]The weighting function is established by putting the ODE in self-adjoint form.

Hermitian Operators in Quantum Mechanics

The preceding discussion focused on the classical second-order differential operators of mathematical physics. Generalizing our Hermitian operator theory, as required in quantum mechanics, we have an extension: The operators need be neither second-order differential operators nor real. For example, the linear momentum operator $p_x = -i\hbar(\partial/\partial x)$ represents a real physical observable and will be an Hermitian operator. We simply assume (as is customary in quantum mechanics) that the wave functions satisfy appropriate boundary conditions in one or three (or other number of) dimensions, vanishing sufficiently strongly at infinity or having periodic behavior (as in a crystal lattice or unit intensity for waves). In practice, this means that the wave functions are in a given Hilbert space. The operator $\mathcal{L}$ is called **Hermitian** if

$$\langle \psi_1 | \mathcal{L}\psi_2 \rangle = \int \psi_1^* \mathcal{L}\psi_2 d\tau = \int (\mathcal{L}\psi_1)^* \psi_2 d\tau = \langle \mathcal{L}\psi_1 | \psi_2 \rangle \tag{9.26}$$

for all ψ_1, ψ_2 of a given Hilbert space. Apart from the simple extension to complex quantities, this definition is identical to Eq. (9.24).

The **adjoint** $A^\dagger$ of an operator A is defined by

$$\langle\psi_1|A^\dagger\psi_2\rangle = \int \psi_1^* A^\dagger \psi_2 d\tau \equiv \int (A\psi_1)^* \psi_2 d\tau = \langle A\psi_1|\psi_2\rangle. \tag{9.27}$$

Comparing with our classical, second derivative operator-oriented Eq. (9.5) defining $\tilde{\mathcal{L}}$, we see that $\mathcal{L}^\dagger = \tilde{\mathcal{L}}^*$ so that we have generalized the adjoint operator to the complex domain (of quantum mechanics). Here, the adjoint is defined in terms of the resultant integral, with the $A^\dagger$ as part of the integrand. Clearly, if $A = A^\dagger$ (**self-adjoint**) and the (space of) functions, on which it acts, satisfy the previously mentioned boundary conditions, then A is Hermitian.

The **expectation value** of an operator $\mathcal{L}$ is defined as

$$\langle\mathcal{L}\rangle = \int \psi^* \mathcal{L}\psi \, d\tau = \langle\psi|\mathcal{L}\psi\rangle. \tag{9.28a}$$

In the framework of quantum mechanics $\langle\mathcal{L}\rangle$ corresponds to the theoretical value of the physical observable represented by $\mathcal{L}$, if the physical system is in a state described by the wave function ψ. When this property is measured experimentally, $\langle\mathcal{L}\rangle$ may be obtained as the mean or average of many measurements of the observable $\mathcal{L}$ of the physical system in the state ψ.

If we require $\mathcal{L}$ to be Hermitian, it is easy to show that $\langle\mathcal{L}\rangle$ is real (as would be expected from a measurement in a physical theory). Taking the complex conjugate of Eq. (9.28a), we obtain

$$\langle\mathcal{L}\rangle^* = \left[\int \psi^* \mathcal{L}\psi \, d\tau\right]^* = \int \psi \mathcal{L}^* \psi^* d\tau.$$

Rearranging the factors in the integrand, we have

$$\langle\mathcal{L}\rangle^* = \int (\mathcal{L}\psi)^* \psi \, d\tau = \langle\mathcal{L}\psi|\psi\rangle.$$

Then, applying our definition of Hermitian operator [Eq. (9.26)], we get

$$\langle\mathcal{L}\rangle^* = \int \psi^* \mathcal{L}\psi d\tau = \langle\mathcal{L}\rangle, \tag{9.28b}$$

or $\langle\mathcal{L}\rangle$ is real. It is worth noting that ψ is not necessarily an eigenfunction of $\mathcal{L}$.

EXERCISES

9.1.1 Show that Laguerre's equation may be put into self-adjoint form by multiplying by e^{-x} and that $w(x) = e^{-x}$ is the weighting function.

9.1.2 Show that the Hermite equation may be put into self-adjoint form by multiplying by e^{-x^2} and that this gives $w(x) = e^{-x^2}$ as the appropriate density function.

9.1.3 Show the following when the linear second-order differential equation is expressed in self-adjoint form:

(a) The Wronskian is equal to a constant divided by the initial coefficient p.

$$W[y_1, y_2] = \frac{C}{p(x)}.$$

(b) A second solution is given by

$$y_2(x) = Cy_1(x) \int^x \frac{dt}{p[y_1(t)]^2}.$$

9.1.4 For the very special case $\lambda = 0$ and $q(x) = 0$, the self-adjoint eigenvalue equation becomes

$$\frac{d}{dx}\left[p(x)\frac{du(x)}{dx}\right] = 0,$$

satisfied by

$$\frac{du}{dx} = \frac{1}{p(x)}.$$

Use this to obtain a "second" solution of the following:
(a) Legendre's equation;
(b) Laguerre's equation; and
(c) Hermite's equation.

ANS. (a) $u_2(x) = \frac{1}{2}\ln\frac{1+x}{1-x}$,

(b) $u_2(x) - u_2(x_0) = \int_{x_0}^{x} e^t \frac{dt}{t}$,

(c) $u_2(x) = \int_0^x e^{t^2} dt$.

These second solutions illustrate the divergent behavior usually found in a second solution.
Note. In all three cases $u_1(x) = 1$.

9.1.5 Given that $\mathcal{L}u = 0$ and $g\mathcal{L}u$ is self-adjoint, show that for the adjoint operator $\tilde{\mathcal{L}}$, $\tilde{\mathcal{L}}(gu) = 0$.

9.1.6 For a second-order differential operator $\mathcal{L}$ that is self-adjoint show that

$$\langle y_2|\mathcal{L}y_1\rangle - \langle y_1|\mathcal{L}y_2\rangle = \int_a^b [y_2\mathcal{L}y_1 - y_1\mathcal{L}y_2]dx = p(y_1'y_2 - y_1y_2')\Big|_a^b.$$

9.1.7 Show that if a function ψ is required to satisfy Laplace's equation in a finite region of space and to satisfy Dirichlet boundary conditions over the entire closed bounding surface, then ψ is unique.
Hint. One of the forms of Green's theorem (Section 1.10) will be helpful.

9.1.8 Consider the solutions of the Legendre, Hermite, and Laguerre equations to be polynomials. Show that the ranges of integration that guarantee that the Hermitian operator boundary conditions will be satisfied are
(a) Legendre $[-1, 1]$, (b) Hermite $(-\infty, \infty)$, (c) Laguerre $[0, \infty)$.

9.1.9 Within the framework of quantum mechanics [Eq. (9.26) and following], show that the following are Hermitian operators:

(a) momentum $\mathbf{p} = -i\hbar\nabla \equiv -i\dfrac{h}{2\pi}\nabla$; and

(b) angular momentum $\mathbf{L} = -i\hbar\mathbf{r}\times\nabla \equiv -i\dfrac{h}{2\pi}\mathbf{r}\times\nabla$.

Hint. In Cartesian form $\mathbf{L}$ is a linear combination of noncommuting Hermitian operators.

9.1.10 (a) A is a non-Hermitian operator. In the sense of Eqs. (9.26) and (9.27), show that

$$A + A^\dagger \qquad \text{and} \qquad i(A - A^\dagger)$$

are Hermitian operators.

(b) Using the preceding result, show that every non-Hermitian operator may be written as a linear combination of two Hermitian operators.

9.1.11 U and V are two arbitrary operators, not necessarily Hermitian. In the sense of Eq. (9.27), show that

$$(UV)^\dagger = V^\dagger U^\dagger.$$

Note the resemblance to Hermitian adjoint matrices.
Hint. Apply the definition of adjoint operator [Eq. (9.27)].

9.1.12 Prove that the product of two Hermitian operators is Hermitian [Eq. (9.26)] if and only if the two operators commute.

9.1.13 A and B are noncommuting quantum mechanical operators:

$$AB - BA = iC.$$

Show that C is Hermitian. Assume that appropriate boundary conditions are satisfied.

9.1.14 The operator $\mathcal{L}$ is Hermitian. Show that $\langle\mathcal{L}^2\rangle \geq 0$.

9.1.15 A quantum mechanical expectation value is defined by

$$\langle A\rangle = \int \psi^*(x)A\psi(x)dx = \langle\psi|A\psi\rangle,$$

where A is a linear operator. Show that demanding that $\langle A\rangle$ be real means that A must be Hermitian with respect to $\psi(x)$.

9.1.16 From the definition of adjoint [Eq. (9.27)], show that $A^{\dagger\dagger} = A$ in the sense that $\int \psi_1^* A^{\dagger\dagger}\psi_2 d\tau = \int \psi_1^* A\psi_2 d\tau$. The adjoint of the adjoint is the original operator.
Hint. The functions ψ_1 and ψ_2 of Eq. (9.27) represent a class of functions. The subscripts 1 and 2 may be interchanged or replaced by other subscripts.

9.1.17 For a quantum particle moving in a potential well, $V(x) = \frac{1}{2}m\omega^2x^2$, the Schrödinger wave equation is

$$-\frac{\hbar^2}{2m}\frac{d^2\psi(x)}{dx^2} + \frac{1}{2}m\omega^2x^2\psi(x) = E\psi(x)$$

or

$$\frac{d^2\psi(z)}{dz^2} - z^2\psi(z) = -\frac{2E}{\hbar\omega}\psi(z),$$

where $z = (m\omega/\hbar)^{1/2}x$. Since this operator is even, we expect solutions of definite parity. For the initial conditions that follow, integrate out from the origin and determine the minimum constant $2E/\hbar\omega$ that will lead to $\psi(\infty) = 0$ in each case. (You may take $z = 6$ as an approximation of infinity.)

(a) For an even eigenfunction,

$$\psi(0) = 1, \qquad \psi'(0) = 0.$$

(b) For an odd eigenfunction,

$$\psi(0) = 0, \qquad \psi'(0) = 1.$$

Note. Analytical solutions appear in Section 13.1.

9.2 Hermitian Operators

Hermitian or self-adjoint operators with appropriate boundary conditions have the following properties that are of extreme importance in classical and quantum physics:

1. The eigenvalues of an Hermitian operator are real.
2. The eigenfunctions of an Hermitian operator are orthogonal.
3. The eigenfunctions of an Hermitian operator form a complete set, meaning that under suitable conditions a function can be expanded in a series of eigenfunctions.[5]

Real Eigenvalues

We proceed to prove the first two of these three properties. Let

$$\mathcal{L}u_i + \lambda_i w u_i = 0, \quad \text{or} \quad \mathcal{L}|u_i\rangle + \lambda_i w|u_i\rangle = 0. \tag{9.29}$$

Assuming the existence of a second eigenvalue and eigenfunction

$$\mathcal{L}u_j + \lambda_j w u_j = 0, \quad \text{or} \quad \mathcal{L}|u_j\rangle + \lambda_j w|u_j\rangle = 0. \tag{9.30}$$

[5]This property is not universal. It **does** hold for our linear, second-order differential operators in Sturm–Liouville (self-adjoint) form. Completeness is defined and discussed in more detail in Section 9.4. A proof that the eigenfunctions of our linear, second-order, self-adjoint, differential equations form a complete set may be developed from the calculus of variations of Section 18.6.

Then, taking the Hermitian adjoint, using $\mathcal{L}^\dagger = \mathcal{L}$ we obtain

$$\mathcal{L}u_j^* + \lambda_j^* w u_j^* = 0, \quad \text{or} \quad \langle u_j|\mathcal{L} + \langle u_j|\lambda_j^* w = 0, \tag{9.31}$$

where p and q are real functions of x, and $w(x)$ is a real function. However, we permit λ_k, the eigenvalues, and u_k, the eigenfunctions, to be complex. Multiplying Eq. (9.29) by u_j^* (or $\langle u_j|$) and Eq. (9.31) by u_i (or $|u_i\rangle$) and then subtracting, we have

$$u_j^*\mathcal{L}u_i - u_i\mathcal{L}u_j^* = (\lambda_j^* - \lambda_i)wu_iu_j^*. \tag{9.32}$$

We integrate over the range $a \le x \le b$,

$$\int_a^b u_j^*\mathcal{L}u_i dx - \int_a^b u_i\mathcal{L}u_j^* dx = (\lambda_j^* - \lambda_i)\int_a^b u_iu_j^* w\, dx, \tag{9.33}$$

or in Dirac notation,

$$\langle u_j|\mathcal{L}u_i\rangle - \langle \mathcal{L}u_j|u_i\rangle = (\lambda_j^* - \lambda_i)\langle u_j|u_i\rangle.$$

Since $\mathcal{L}$ is Hermitian, the left-hand side vanishes by Eq. (9.27) and

$$(\lambda_j^* - \lambda_i)\int_a^b u_iu_j^* w\, dx = (\lambda_j^* - \lambda_i)\langle u_j|u_i\rangle = 0. \tag{9.34}$$

If $i = j$, the integral cannot vanish [$w(x) > 0$, apart from isolated points], except in the trivial case $u_i = 0$. Hence, the coefficient $(\lambda_i^* - \lambda_i)$ must be zero,

$$\lambda_i^* = \lambda_i, \tag{9.35}$$

which states that the eigenvalue is real. Since λ_i can represent any one of the eigenvalues, this proves the first property. This is an exact analog of the nature of the eigenvalues of real symmetric (and Hermitian) matrices (compare Sections 3.3 and 3.4).

Real eigenvalues of Hermitian operators have a fundamental significance in quantum mechanics. In quantum mechanics the eigenvalues correspond to observable (precisely measurable or sharp) quantities, such as energy and angular momentum. When a single measurement of an observable $\mathcal{L}$ is made, the result must be one of its eigenvalues. With the theory formulated in terms of Hermitian operators, this proof of real eigenvalues guarantees that the theory will predict real numbers for these measurable physical quantities. In Section 18.6, it will be seen that for some operators, such as Hamiltonians, the set of real eigenvalues has a lower bound. Physically important Hermitian operators are real potentials $V^* = V$ and the momentum operator $-id/dx$. The latter is Hermitian because upon using integration by parts and discarding the integrated term, we have

$$\left\langle \psi_1 \middle| -i\frac{d\psi_2}{dx}\right\rangle = \int_{-\infty}^{\infty} \psi_1^*\left(-i\frac{d\psi_2}{dx}\right)dx = -i\psi_2\psi_1^*|_{-\infty}^{\infty} + i\int_{-\infty}^{\infty}\frac{d\psi_1^*}{dx}\psi_2 dx$$
$$= \int_{-\infty}^{\infty}\left(-i\frac{d\psi_1}{dx}\right)^*\psi_2 dx = \left\langle -i\frac{d\psi_1}{dx}\middle|\psi_2\right\rangle.$$

Orthogonal Eigenfunctions

If we take $i \neq j$ and if $\lambda_i \neq \lambda_j$ in Eq. (9.34), the integral of the product of the two different eigenfunctions must vanish:

$$\langle u_j | u_i \rangle = \int_a^b u_j^* u_i w \, dx = 0. \tag{9.36}$$

This condition, called **orthogonality, is the continuum analog of the vanishing of a scalar (or inner) product of two vectors.**[6] We say that the eigenfunctions $u_i(x)$ and $u_j(x)$ are orthogonal with respect to the weighting function $w(x)$ over the interval $[a, b]$. Equation (9.36) constitutes a partial proof of the second property of our Hermitian operators. Again, the precise analogy with matrix analysis should be noted. Indeed, we can establish a one-to-one correspondence between this Sturm–Liouville theory of differential equations and the treatment of Hermitian matrices. Historically, this correspondence has been significant in establishing the mathematical equivalence of matrix mechanics developed by Heisenberg and wave mechanics developed by Schrödinger. Today, the two diverse approaches are merged into the theory of quantum mechanics, and the mathematical formulation that is more convenient for a particular problem is used for that problem. Actually, the mathematical alternatives do not end here. Integral equations form a third equivalent and sometimes more convenient or more powerful approach. Similarly, any two functions u, v, not necessarily eigenfunctions, are orthogonal if $\langle v | u \rangle = \int_a^b v^* u w dx = 0$.

This proof of orthogonality is not quite complete. There is a loophole because we may have $u_i \neq u_j$ but still have $\lambda_i = \lambda_j$. Such eigenvalues are labeled **degenerate**. Illustrations of degeneracy are given at the end of this section. If $\lambda_i = \lambda_j$, the integral in Eq. (9.34) need not vanish. This means that linearly independent eigenfunctions corresponding to the same eigenvalue are not automatically orthogonal and that some other method must be sought to obtain an orthogonal set. Although the eigenfunctions in this degenerate case may not be orthogonal, they can always be made orthogonal. One method is developed in the next section. See also the discussion after Eq. (4.13) for degeneracy due to symmetry.

We shall see in succeeding chapters that it is just as desirable to have a given set of functions orthogonal as it is to have an orthogonal coordinate system. We can work with nonorthogonal functions, but they are likely to prove as messy as an oblique coordinate system.

---

[6] From the definition of Riemann integral,

$$\int_a^b f(x)g(x)dx = \lim_{N\to\infty} \left(\sum_{i=1}^{N} f(x_i) g(x_i) \Delta x \right),$$

where $x_0 = a$, $x_N = b$, and $x_i - x_{i-1} = \Delta x$. If we interpret $f(x_i)$ and $g(x_i)$ as the ith components of an N component vector, then this sum (and therefore this integral) corresponds directly to a scalar product of vectors, Eq. (1.11). The vanishing of the scalar product is the condition for **orthogonality** of the vectors—or functions.

EXAMPLE 9.2.1 **Fourier Series—Orthogonality** To continue Example 9.1.4 with the interval $-\pi \le x \le \pi$, the eigenvalue equation [Eq. (9.25)],

$$\frac{d^2}{dx^2}y(x) + n^2 y(x) = 0,$$

subject to Eq. (9.20), may describe a vibrating violin string with eigenfunctions—$\sin nx$ subject to the boundary conditions $\sin(\pm n\pi) = 0$ so that n is an integer, and the orthogonality integrals become

$$\text{(a)} \int_{x_0-\pi}^{x_0+\pi} \sin mx \sin nx\, dx = C_n \delta_{nm}, \quad x_0 = 0.$$

For an interval of length 2π the preceding analysis guarantees the Kronecker delta in (a). Our Sturm–Liouville theory says nothing about the values of C_n.

Similarly, a quantum mechanical particle in a box may have eigenfunctions $\cos nx$ subject to the boundary conditions $\frac{d\cos nx}{dx}|_{x=\pm\pi} = 0$ giving integer n again. Then

$$\text{(b)} \int_{x_0-\pi}^{x_0+\pi} \cos mx \cos nx\, dx = D_n \delta_{nm}, \quad x_0 = 0,$$

where D_n remains undetermined. Actual calculation yields

$$C_n = \begin{cases} \pi, & n \neq 0, \\ 0, & n = 0, \end{cases} \qquad D_n = \begin{cases} \pi, & n \neq 0, \\ 2\pi, & n = 0. \end{cases}$$

Finally, inspection shows that

$$\text{(c)} \int_{x_0-\pi}^{x_0+\pi} \sin mx \cos nx\, dx = 0$$

always vanishes for all integral m and n. ■

Expansion in Orthogonal Eigenfunctions

Starting from some Hamiltonian and its eigenvalue equation $H|\psi\rangle = E|\psi\rangle$, we determine the set of eigenvalues E_j and eigenfunctions $|\varphi_j\rangle$ taking the latter to be orthonormal; that is,

$$\langle \varphi_k | \varphi_j \rangle = \delta_{jk}.$$

The property of completeness of the set $|\varphi_j\rangle$ means that certain classes of (e.g., sectionally or piecewise continuous) functions may be represented by a series of orthogonal eigenfunctions to any desired degree of accuracy. We now assume $|\psi\rangle$ is in that class and expand it as

$$|\psi\rangle = \sum_j a_j |\varphi_j\rangle.$$

We determine the coefficient a_k by projection

$$\left\langle \varphi_k \middle| H \sum_j a_j \varphi_j \right\rangle = E \left\langle \varphi_k \middle| \sum_j a_j \varphi_j \right\rangle = E \sum_j a_j \langle \varphi_k | \varphi_j \rangle = E \sum_j a_j \delta_{kj} = E a_k.$$

Calling $\langle\varphi_k|H\varphi_j\rangle \equiv H_{k,j}$ the **matrix elements of the Hamiltonian**, we have the eigenvalue equations

$$\sum_j H_{k,j}a_j = Ea_k, \tag{9.37}$$

from which the column vector of admixture coefficients a_j may be determined, along with the eigenvalue E. This usually infinite set of linear equations is truncated in practice.

The choice of eigenfunction is made on the basis of convenience. To illustrate the expansion technique, let us choose the eigenfunctions of Example 9.2.1, $\cos nx$ and $\sin nx$. The eigenfunction series is conveniently (and conventionally) written as the Fourier series

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty}(a_n \cos nx + b_n \sin nx).$$

From the orthogonality integrals of Example 9.2.1 the coefficients are given by

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\cos nt\, dt, \quad b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\sin nt\, dt, \qquad n = 0, 1, 2 \ldots .$$

EXAMPLE 9.2.2

Square Wave Now consider the square wave

$$f(x) = \begin{cases} \dfrac{h}{2}, & 0 < x < \pi, \\ -\dfrac{h}{2}, & -\pi < x < 0. \end{cases}$$

Direct substitution of $\pm h/2$ for $f(t)$ yields

$$a_n = 0,$$

which is expected because of the antisymmetry, $f(-x) = -f(x)$, and

$$b_n = \frac{h}{n\pi}(1 - \cos n\pi) = \begin{cases} 0, & n \text{ even}, \\ \dfrac{2h}{n\pi}, & n \text{ odd}. \end{cases}$$

Hence, the eigenfunction (Fourier) expansion of the square wave is

$$f(x) = \frac{2h}{\pi}\sum_{n=0}^{\infty}\frac{\sin(2n+1)x}{2n+1}. \tag{9.38}$$

Additional examples, using other eigenfunctions, appear in Chapters 11–13. ■

Degeneracy

The concept of degeneracy was introduced earlier. If N linearly independent eigenfunctions correspond to the same eigenvalue, the eigenvalue is said to be N-fold degenerate. A particularly simple illustration is provided by the eigenvalues and eigenfunctions of the linear oscillator equation of classical mechanics (Example 9.2.1). For each value of the eigenvalue n, there are two possible solutions: $\sin nx$ and $\cos nx$ (and any linear combination). We may say the eigenfunctions are degenerate or the eigenvalue is degenerate.

When an underlying symmetry, such as rotational invariance, is causing the degeneracies, states belonging to the same energy eigenvalue will then form a multiplet or representation of the symmetry group. The powerful group–theoretical methods are treated in Chapter 4 in detail.

In Section 9.3, we show an alternative method of how such functions may be made orthogonal.

Biographical Data

Sturm, Jacques Charles. Sturm, who was born in 1803 in Geneva, Switzerland and died in 1855, was Poisson's successor at the Sorbonne and worked with his friend Liouville on heat flow, from which the eigenvalue problems arose now named after both.

Liouville, Joseph. Liouville (1809–1882), a professor at the Collège de France, made contributions to elliptic functions, analytic functions, and quadratic forms.

EXERCISES

9.2.1 The functions $u_1(x)$ and $u_2(x)$ are eigenfunctions of the same Hermitian operator but for distinct eigenvalues λ_1 and λ_2. Show that $u_1(x)$ and $u_2(x)$ are linearly independent.

9.2.2 (a) The vectors $\mathbf{e}_n$ are orthogonal to each other: $\mathbf{e}_n \cdot \mathbf{e}_m = 0$ for $n \neq m$. Show that they are linearly independent.

(b) The functions $\psi_n(x)$ are orthogonal to each other over the interval $[a, b]$ and with respect to the weighting function $w(x)$. Show that the $\psi_n(x)$ are linearly independent.

9.2.3 Given that

$$P_1(x) = x \qquad \text{and} \qquad Q_0(x) = \tfrac{1}{2}\ln\left(\frac{1+x}{1-x}\right)$$

are solutions of Legendre's differential equation corresponding to different eigenvalues:

(a) Evaluate their orthogonality integral

$$\int_{-1}^{1} \frac{x}{2}\ln\left(\frac{1+x}{1-x}\right) dx.$$

(b) Explain why these two functions are not orthogonal, that is, why the proof of orthogonality does not apply.

9.2.4 $T_0(x) = 1$ and $V_1(x) = (1 - x^2)^{1/2}$ are solutions of the Chebyshev differential equation corresponding to different eigenvalues. Explain, in terms of the boundary conditions, why these two functions are not orthogonal.

9.2.5 (a) Show that the first derivatives of the Legendre polynomials satisfy a self-adjoint differential equation with eigenvalue $\lambda = n(n+1) - 2$.

(b) Show that these Legendre polynomial derivatives satisfy an orthogonality relation

$$\int_{-1}^{1} P_m'(x) P_n'(x)(1 - x^2)\,dx = 0, \qquad m \neq n.$$

Note. In Section 11.5, $(1 - x^2)^{1/2} P_n'(x)$ will be labeled an associated Legendre polynomial, $P_n^1(x)$.

9.2.6 A set of functions $u_n(x)$ satisfies the Sturm–Liouville equation

$$\frac{d}{dx}\left[p(x)\frac{d}{dx}u_n(x)\right] + \lambda_n w(x) u_n(x) = 0.$$

The functions $u_m(x)$ and $u_n(x)$ satisfy boundary conditions that lead to orthogonality. The corresponding eigenvalues λ_m and λ_n are distinct. Prove that for appropriate boundary conditions $u_m'(x)$ and $u_n'(x)$ are orthogonal with $p(x)$ as a weighting function.

9.2.7 A linear operator A has n distinct eigenvalues and n corresponding eigenfunctions: $A\psi_i = \lambda_i \psi_i$. Show that the n eigenfunctions are linearly independent. A is not necessarily Hermitian.
Hint. Assume linear dependence—that $\psi_n = \sum_{i=1}^{n-1} a_i \psi_i$. Use this relation and the operator–eigenfunction equation first in one order and then in the reverse order. Show that a contradiction results.

9.2.8 With $\mathcal{L}$ **not** self-adjoint, $\tilde{\mathcal{L}} \neq \mathcal{L}$,

$$\mathcal{L}u_i + \lambda_i w u_i = 0$$

and

$$\tilde{\mathcal{L}}v_j + \lambda_j w v_j = 0.$$

(a) Show that

$$\int_a^b v_j \mathcal{L} u_i\,dx = \int_a^b u_i \tilde{\mathcal{L}} v_j\,dx,$$

provided

$$u_i p_0 v_j' \Big|_a^b = v_j p_0 u_i' \Big|_a^b$$

and

$$u_i(p_1 - p_0')v_j\Big|_a^b = 0.$$

(b) Show that the orthogonality integral for the eigenfunctions u_i and v_j becomes

$$\int_a^b u_i v_j w\, dx = 0 \qquad (\lambda_i \neq \lambda_j).$$

9.3 Gram–Schmidt Orthogonalization

The Gram–Schmidt orthogonalization is a method that takes a nonorthogonal set of linearly independent functions[7] and constructs an orthogonal set over an arbitrary interval and with respect to an arbitrary weight or density factor w that may or may not originate from our basic Eq. (9.8). The choice of these weights gives a particular set of **orthonormal** functions in the end (orthogonal plus unit normalization). In the language of linear algebra, the process is equivalent to a matrix transformation relating an orthogonal set of basis vectors (functions) to a nonorthogonal set. The functions involved may be real or complex. Here, for convenience they are assumed to be real. The generalization to more than one dimension or to complex cases offers no difficulty.

Before taking up orthogonalization, we should consider normalization of functions. So far, no normalization has been specified. This means that

$$\langle \varphi_i | \varphi_i \rangle = \int_a^b \varphi_i^2 w\, dx = N_i^2,$$

but no attention has been paid to the value of N_i. We now demand that each function $\varphi_i(x)$ be multiplied by N_i^{-1} so that the new (normalized) φ_i will satisfy

$$\langle \varphi_i | \varphi_i \rangle = \int_a^b \varphi_i^2(x) w(x) dx = 1 \tag{9.39}$$

and

$$\langle \varphi_j | \varphi_i \rangle = \int_a^b \varphi_i(x) \varphi_j(x) w(x) dx = \delta_{ij}. \tag{9.40}$$

Equation (9.39) states that we have normalized to unity. Including the property of orthogonality, we have Eq. (9.40). Functions satisfying this equation are said to be **orthonormal** (orthogonal plus unit normalization). Other normalizations are certainly possible, and indeed, by historical convention, each of the special functions of mathematical physics treated in Chapters 12 and 13 will be normalized differently.

[7] Such a set of functions might well arise from the solutions of a PDE, in which the eigenvalue was independent of one or more of the constants of separation. Note, however, that the origin of the set of functions is irrelevant to the Gram–Schmidt orthogonalization procedure.

We consider three sets of functions: an original, linearly independent given set $u_n(x)$, $n = 0, 1, 2, \ldots$; an orthogonalized set $\psi_n(x)$ to be constructed; and a final set of functions $\varphi_n(x)$ that are the normalized ψ_n. The original u_n may be degenerate eigenfunctions, but this is not necessary. We shall have

$u_n(x)$	$\psi_n(x)$	$\varphi_n(x)$
Linearly independent	*Linearly independent*	*Linearly independent*
Nonorthogonal	*Orthogonal*	*Orthogonal*
Unnormalized	Unnormalized	*Normalized* (Orthonormal)

The Gram–Schmidt procedure takes the nth ψ function, ψ_n, to be $u_n(x)$ plus an unknown linear combination of the previous φ. The presence of the new $u_n(x)$ will guarantee linear independence. The requirement that $\psi_n(x)$ be orthogonal to each of the previous φ yields just enough constraints to determine each of the unknown coefficients. Then the fully determined ψ_n will be normalized to unity, yielding $\varphi_n(x)$. Then the sequence of steps is repeated for $\psi_{n+1}(x)$.

We start with $n = 0$, letting

$$\psi_0(x) = u_0(x) \tag{9.41}$$

with no "previous" φ to worry about. Then normalize

$$\varphi_0(x) = \frac{\psi_0(x)}{[\int_a^b \psi_0^2 w\, dx]^{1/2}}, \quad \text{or} \quad |\varphi_0\rangle = \frac{|\psi_0\rangle}{[\langle\psi_0|\psi_0\rangle]^{1/2}}. \tag{9.42}$$

For $n = 1$, let

$$\psi_1(x) = u_1(x) + a_{10}\varphi_0(x). \tag{9.43}$$

We demand that $\psi_1(x)$ be orthogonal to $\varphi_0(x)$. [At this stage the normalization of $\psi_1(x)$ is irrelevant.] This demand of orthogonality leads to

$$\int_a^b \psi_1\varphi_0 w\, dx = \int_a^b u_1\varphi_0 w\, dx + a_{10}\int \varphi_0^2 w\, dx = 0. \tag{9.44}$$

Since φ_0 is normalized to unity [Eq. (9.42)], we have

$$a_{10} = -\int_a^b u_1\varphi_0 w\, dx = -\langle\varphi_0|u_1\rangle, \tag{9.45}$$

fixing the value of a_{10}. In Dirac notation we write Eq. (9.43) as

$$|\psi_1\rangle = |u_1\rangle - \langle\varphi_0|u_1\rangle|\varphi_0\rangle \tag{9.43a}$$

and Eq. (9.44) as

$$0 = \langle\varphi_0|\psi_1\rangle = \langle\varphi_0|u_1\rangle + a_{10}\langle\varphi_0|\varphi_0\rangle. \tag{9.44a}$$

In this form we recognize that the coefficient a_{10} is determined by projection, similar to expanding a vector $\mathbf{u}$ in terms of Cartesian coordinate or basis vectors $\hat{\mathbf{x}}_i$ as

$$\mathbf{u} = (\mathbf{u}\cdot\hat{\mathbf{x}}_1)\hat{\mathbf{x}}_1 + (\mathbf{u}\cdot\hat{\mathbf{x}}_2)\hat{\mathbf{x}}_2 + (\mathbf{u}\cdot\hat{\mathbf{x}}_3)\hat{\mathbf{x}}_3 = \sum_i u_i\hat{\mathbf{x}}_i. \tag{9.45a}$$

Normalizing, we define

$$\varphi_1(x) = \frac{\psi_1(x)}{\left(\int \psi_1^2 w\,dx\right)^{1/2}}, \quad \text{or} \quad |\varphi_1\rangle = \frac{|\psi_1\rangle}{[\langle\psi_1|\psi_1\rangle]^{1/2}}. \tag{9.46}$$

Finally, we generalize so that

$$\varphi_i(x) = \frac{\psi_i(x)}{\left(\int \psi_i^2(x)w(x)dx\right)^{1/2}}, \tag{9.47}$$

where

$$\psi_i(x) = u_i + a_{i0}\varphi_0 + a_{i1}\varphi_1 + \cdots + a_{ii-1}\varphi_{i-1}. \tag{9.48}$$

The coefficients a_{ij} are again given by projection (using orthogonality)

$$a_{ij} = -\int u_i\varphi_j w\,dx = -\langle\varphi_j|u_i\rangle. \tag{9.49}$$

Equation (9.49) holds for unit normalization. If some other normalization is selected,

$$\int_a^b [\varphi_j(x)]^2 w(x)dx = N_j^2,$$

then Eq. (9.47) is replaced by

$$\varphi_i(x) = N_i\frac{\psi_i(x)}{\left(\int \psi_i^2 w\,dx\right)^{1/2}} \tag{9.47a}$$

and a_{ij} becomes

$$a_{ij} = -\frac{\int u_i\varphi_j w\,dx}{N_j^2}. \tag{9.49a}$$

Equations (9.48) and (9.49) may be rewritten in terms of **projection operators**, P_j. If we consider the $\varphi_n(x)$ to form a linear vector space, then the integral in Eq. (9.49) may be interpreted as the projection of u_i into the φ_j "coordinate" or the jth component of u_i. With

$$P_j u_i(x) = \left\{\int u_i(t)\varphi_j(t)w(t)dt\right\}\varphi_j(x) = |\varphi_j\rangle\langle\varphi_j|u_i\rangle,$$

that is, $P_j = |\varphi_j\rangle\langle\varphi_j|$. Equation (9.48) becomes

$$\psi_i(x) = \left\{1 - \sum_{j=1}^{i-1} P_j\right\}u_i(x). \tag{9.48a}$$

Subtracting off the jth components, $j = 1$ to $i - 1$, leaves $\psi_i(x)$ orthogonal to all the $\varphi_j(x)$.

Note that although this Gram–Schmidt procedure is one possible way of constructing an orthogonal or orthonormal set, the functions $\varphi_i(x)$ are not unique. There is an infinite number of possible orthonormal sets for a given interval and a given density function. As an illustration of the freedom involved, consider two (nonparallel) vectors $\mathbf{A}$ and $\mathbf{B}$ in the xy-plane. We may normalize $\mathbf{A}$ to unit magnitude and then form $\mathbf{B}' = a\mathbf{A} + \mathbf{B}$ so that $\mathbf{B}'$ is perpendicular to $\mathbf{A}$. By normalizing $\mathbf{B}'$ we have completed the Gram–Schmidt orthogonalization for two vectors. However, any two perpendicular unit vectors, such as $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$, could have been chosen as our orthonormal set. Again, with an infinite number of possible rotations of $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ about the z-axis, we have an infinite number of possible orthonormal sets.

EXAMPLE 9.3.1

Legendre Polynomials by Gram–Schmidt Orthogonalization Let us form an orthonormal set from the set of functions $u_n(x) = x^n$, $n = 0, 1, 2, \ldots$. The interval is $-1 \le x \le 1$ and the density function is $w(x) = 1$.

In accordance with the Gram–Schmidt orthogonalization process described,

$$u_0 = 1, \qquad \text{hence} \qquad \varphi_0 = \frac{1}{\sqrt{2}}. \tag{9.50}$$

Then

$$\psi_1(x) = x + a_{10}\frac{1}{\sqrt{2}} \tag{9.51}$$

and

$$a_{10} = -\int_{-1}^{1} \frac{x}{\sqrt{2}}dx = 0 \tag{9.52}$$

by symmetry. We normalize ψ_1 to obtain

$$\varphi_1(x) = \sqrt{\frac{3}{2}}x. \tag{9.53}$$

Then continue the Gram–Schmidt process with

$$\psi_2(x) = x^2 + a_{20}\frac{1}{\sqrt{2}} + a_{21}\sqrt{\frac{3}{2}}x, \tag{9.54}$$

where

$$a_{20} = -\int_{-1}^{1} \frac{x^2}{\sqrt{2}}dx = -\frac{\sqrt{2}}{3}, \tag{9.55}$$

$$a_{21} = -\int_{-1}^{1} \sqrt{\frac{3}{2}}x^3 dx = 0, \tag{9.56}$$

again by symmetry. Therefore,

$$\psi_2(x) = x^2 - \frac{1}{3}, \tag{9.57}$$

and, on normalizing to unity, we have

$$\varphi_2(x) = \sqrt{\frac{5}{2}} \cdot \frac{1}{2}(3x^2 - 1). \tag{9.58}$$

The next function $\varphi_3(x)$ becomes

$$\varphi_3(x) = \sqrt{\frac{7}{2}} \cdot \frac{1}{2}(5x^3 - 3x). \tag{9.59}$$

Reference to Chapter 11 will show that

$$\varphi_n(x) = \sqrt{\frac{2n+1}{2}} P_n(x), \tag{9.60}$$

where $P_n(x)$ is the nth-order Legendre polynomial. Our Gram–Schmidt process provides a possible, but very cumbersome, method of generating the Legendre polynomials. It illustrates how a power series expansion in $u_n(x) = x^n$, which is not orthogonal, can be converted into an orthogonal series over the finite interval $[-1, 1]$.

The equations for Gram–Schmidt orthogonalization tend to be ill conditioned because of the subtractions. A technique for avoiding this difficulty using the polynomial recurrence relation is discussed by Hamming.[8]

In Example 9.3.1, we have specified an orthogonality interval $[-1, 1]$, a unit weighting function, and a set of functions, x^n, to be taken one at a time in increasing order. Given all these specifications the Gram–Schmidt procedure is unique (to within a normalization factor and an overall sign as discussed subsequently). Our resulting orthogonal set, the Legendre polynomials, P_0 through P_n, forms a complete set for the description of polynomials of order $\leq n$ over $[-1, 1]$. This concept of completeness is taken up in detail in Section 9.4. Expansions of functions in series of Legendre polynomials are discussed in Section 11.3. ■

Orthogonal Polynomials

The previous example was chosen strictly to illustrate the Gram–Schmidt procedure. Although it has the advantage of introducing the Legendre polynomials, the initial functions $u_n = x^n$ are not degenerate eigenfunctions and are not solutions of Legendre's equation. They are simply a set of functions that we have rearranged to create an orthonormal set for the given interval and

[8] Hamming, R. W. (1973). *Numerical Methods for Scientists and Engineers*, 2nd ed. McGraw-Hill, New York. See Section 27.2 and references given there.

Table 9.3

Orthogonal Polynomials Generated by Gram–Schmidt Orthogonalization of $u_n(x) = x^n, n = 0, 1, 2, \ldots$

Polynomials	Interval	Weighting Function $w(x)$	Standard Normalization
Legendre	$-1 \le x \le 1$	1	$\int_{-1}^{1} [P_n(x)]^2 dx = \frac{2}{2n+1}$
Shifted Legendre	$0 \le x \le 1$	1	$\int_{0}^{1} [P_n^*(x)]^2 dx = \frac{1}{2n+1}$
Laguerre	$0 \le x < \infty$	e^{-x}	$\int_{0}^{\infty} [L_n(x)]^2 e^{-x} dx = 1$
Associated Laguerre	$0 \le x < \infty$	$x^k e^{-x}$	$\int_{0}^{\infty} [L_n^k(x)]^2 x^k e^{-x} dx = \frac{(n+k)!}{n!}$
Hermite	$-\infty < x < \infty$	e^{-x^2}	$\int_{-\infty}^{\infty} [H_n(x)]^2 e^{-x^2} dx = 2^n \pi^{1/2} n!$

given weighting function. The fact that we obtained the Legendre polynomials is not quite magic but a direct consequence of the choice of interval and weighting function. The use of $u_n(x) = x^n$, but with other choices of interval and weighting function or a different ordering of the functions, leads to other sets of orthogonal polynomials as shown in Table 9.3. We consider these polynomials in detail in Chapters 11 and 13 as solutions of particular differential equations.

An examination of this orthogonalization process reveals two arbitrary features. First, as emphasized previously, it is not necessary to normalize the functions to unity. In the example just given, we could have required

$$\int_{-1}^{1} \varphi_n(x)\varphi_m(x)dx = \frac{2}{2n+1}\delta_{nm}, \tag{9.61}$$

and the resulting set would have been the actual Legendre polynomials. Second, the sign of φ_n is always indeterminate. In the example, we chose the sign by requiring the coefficient of the highest power of x in the polynomial to be positive. For the Laguerre polynomials, on the other hand, we would require the coefficient of the highest power to be $(-1)^n/n!$.

EXERCISES

9.3.1 Rework Example 9.3.1 by replacing $\varphi_n(x)$ by the conventional Legendre polynomial, $P_n(x)$:

$$\int_{-1}^{1} [P_n(x)]^2 dx = \frac{2}{2n+1}.$$

Using Eqs. (9.47a) and (9.49a), construct P_0, $P_1(x)$, and $P_2(x)$.

ANS. $P_0 = 1, \quad P_1 = x, \quad P_2 = \frac{3}{2}x^2 - \frac{1}{2}.$

9.3.2 Following the Gram–Schmidt procedure, construct a set of polynomials $P_n^*(x)$ orthogonal (unit weighting factor) over the range [0, 1] from the set $[1, x]$. Normalize so that $P_n^*(1) = 1$.

ANS. $P_n^*(x) = 1,$
$P_1^*(x) = 2x - 1,$
$P_2^*(x) = 6x^2 - 6x + 1,$
$P_3^*(x) = 20x^3 - 30x^2 + 12x - 1.$

These are the first four **shifted** Legendre polynomials.
Note. The asterisk is the standard notation for "shifted": [0, 1] instead of [−1, 1]. It does **not** mean complex conjugate.

9.3.3 Apply the Gram–Schmidt procedure to form the first three Laguerre polynomials, using

$$u_n(x) = x^n, \qquad n = 0, 1, 2, \ldots, \qquad 0 \le x < \infty, \qquad w(x) = e^{-x}.$$

The conventional normalization is

$$\langle L_m | L_n \rangle = \int_0^\infty L_m(x) L_n(x) e^{-x} dx = \delta_{mn}.$$

ANS. $L_0 = 1, \quad L_1 = (1 - x), \quad L_2 = \dfrac{2 - 4x + x^2}{2}.$

9.3.4 You are given
(a) a set of functions $u_n(x) = x^n$, $n = 0, 1, 2, \ldots$,
(b) an interval $(0, \infty)$, and
(c) a weighting function $w(x) = xe^{-x}$. Use the Gram–Schmidt procedure to construct the first **three orthonormal** functions from the set $u_n(x)$ for this interval and this weighting function.

ANS. $\varphi_0(x) = 1, \quad \varphi_1(x) = (x - 2)/\sqrt{2}, \quad \varphi_2(x) = (x^2 - 6x + 6)/2\sqrt{3}.$

9.3.5 Using the Gram–Schmidt orthogonalization procedure, construct the lowest three Hermite polynomials, using

$$u_n(x) = x^n, \qquad n = 0, 1, 2, \ldots, \qquad -\infty < x < \infty, \qquad w(x) = e^{-x^2}.$$

For this set of polynomials the usual normalization is

$$\langle H_m | H_n \rangle = \int_{-\infty}^\infty H_m(x) H_n(x) w(x) dx = \delta_{mn} 2^m m! \pi^{1/2}.$$

ANS. $H_0 = 1, \quad H_1 = 2x, \quad H_2 = 4x^2 - 2.$

9.3.6 As a modification of Exercise 9.3.5, apply the Gram–Schmidt orthogonalization procedure to the set $u_n(x) = x^n$, $n = 0, 1, 2, \ldots, 0 \le x < \infty$. Take $w(x)$ to be $\exp[-x^2]$. Find the first two nonvanishing polynomials. Normalize so that the coefficient of the highest power of x is unity. In Exercise 9.3.5, the interval $(-\infty, \infty)$ led to the Hermite polynomials. These are certainly not the Hermite polynomials.

ANS. $\varphi_0 = 1, \quad \varphi_1 = x - \pi^{-1/2}.$

9.3.7 Form an orthogonal set over the interval $0 \le x < \infty$, using $u_n(x) = e^{-nx}$, $n = 1, 2, 3, \ldots$. Take the weighting factor, $w(x)$, to be unity. These functions are solutions of $u_n'' - n^2 u_n = 0$, which is clearly already in Sturm–Liouville (self-adjoint) form. Why doesn't the Sturm–Liouville theory guarantee the orthogonality of these functions?

9.4 Completeness of Eigenfunctions

The third important property of an Hermitian operator is that its eigenfunctions form a complete set. This completeness means that any well-behaved (at least piecewise continuous) function $F(x)$ can be approximated by a series

$$F(x) = \sum_{n=0}^{\infty} a_n \varphi_n(x) \tag{9.62}$$

to any desired degree of accuracy.[9] More precisely, the set $\varphi_n(x)$ is called complete[10] if the limit of the mean square error vanishes:

$$\lim_{m\to\infty} \int_a^b \left[F(x) - \sum_{n=0}^{m} a_n \varphi_n(x) \right]^2 w(x)\,dx = 0. \tag{9.63}$$

We have not required that the error vanish identically in $[a, b]$ but only that the integral of the error squared go to zero. This convergence in the mean [Eq. (9.63)] should be compared with uniform convergence [Section 5.5, Eq. (5.43)]. Clearly, uniform convergence implies convergence in the mean, but the converse does not hold; convergence in the mean is less restrictive. Specifically, Eq. (9.63) is not upset by piecewise continuous functions with only a finite number of finite discontinuities. Equation (9.63) is perfectly adequate for our purposes and is far more convenient than Eq. (5.43). Indeed, since we frequently use expansions in eigenfunctions to describe discontinuous functions, convergence in the mean is all we can expect.

In the language of linear algebra, we have a linear space, a function vector space. The linearly independent, orthonormal functions $\varphi_n(x)$ form the basis for this (infinite-dimensional) space. Equation (9.62) is a statement that the functions $\varphi_n(x)$ span this linear space. With an inner product defined by Eq. (9.36), our linear space is a **Hilbert space** spanned by the complete set of basis states $\varphi_n(x)$; it contains all square-integrable functions F that can be expanded in the sense of Eq. (9.63).

The question of completeness of a set of functions is often determined by comparison with a Laurent series (Section 6.5). In Section 14.1, this is done for Fourier series, thus establishing the completeness of Fourier series. For all orthogonal polynomials mentioned in Section 9.3, it is possible to find a

[9] If we have a finite set, as with vectors, the summation is over the number of linearly independent members of the set.

[10] Many authors use the term *closed* here.

polynomial expansion of each power of z,

$$z^n = \sum_{i=0}^{n} a_i P_i(z), \tag{9.64}$$

where $P_i(z)$ is the ith polynomial. Exercises 11.4.6, 13.1.8, and 13.2.5 are specific examples of Eq. (9.64). Using Eq. (9.64), we may reexpress the Laurent expansion of $f(z)$ in terms of the polynomials, showing that the polynomial expansion exists (when it exists, it is unique; Exercise 9.4.1). The limitation of this Laurent series development is that it requires the function to be analytic. Equations (9.62) and (9.63) are more general. $F(x)$ may be only piecewise continuous. Numerous examples of the representation of such piecewise continuous functions appear in Chapter 14 (Fourier series). A proof that our Sturm–Liouville eigenfunctions form complete sets appears in Courant and Hilbert.[11]

In Eq. (9.62) the expansion coefficients a_m may be determined by

$$a_m = \int_a^b F(x)\varphi_m(x)w(x)\,dx = \langle \varphi_m | F \rangle. \tag{9.65}$$

This follows from multiplying Eq. (9.62) by $\varphi_m(x)w(x)$ and integrating. In Dirac notation,

$$|F\rangle = \sum_n a_n |\varphi_n\rangle \quad \text{implies} \quad \langle \varphi_m | F \rangle = \sum_n a_n \langle \varphi_m | \varphi_n \rangle = \sum_n a_n \delta_{mn} = a_m$$

provided the $|\varphi_n\rangle$ are normalized to unity. From the orthogonality of the eigenfunctions, $\varphi_n(x)$, only the mth term survives. Here, we see the value of orthogonality. Equation (9.65) may be compared with the dot or inner product of vectors (Section 1.2) and a_m interpreted as the mth projection of the function $F(x)$. Often, the coefficient a_m is called a generalized Fourier coefficient.

For a known function $F(x)$, Eq. (9.65) gives a_m as a **definite** integral that can always be evaluated, by computer if not analytically.

For examples of particular eigenfunction expansions, see the following: Fourier series, Chapter 14; Legendre series, Section 11.3; Laplace series, Section 11.6; Bessel and Fourier–Bessel expansions, Section 12.1; Hermite series, Section 13.1; and Laguerre series, Section 13.2. An explicit case of a Fourier expansion is the square wave (Example 9.2.2). The corresponding Hilbert space contains only periodic functions that can be expanded in a series of $\sin nx$, $\cos nx$, the eigenfunctions of one-dimensional square-well potentials in quantum mechanics under suitable boundary conditions.

It may also happen that the eigenfunction expansion [Eq. (9.62)] is the expansion of an unknown $F(x)$ in a series of known eigenfunctions $\varphi_n(x)$ with unknown coefficients a_n. An example is the quantum chemist's attempt to describe an (unknown) molecular wave function as a linear combination of known atomic wave functions. The unknown coefficients a_n would be determined by a variational technique, Rayleigh–Ritz (Section 18.6).

[11] Courant, R., and Hilbert, D. (1953). *Methods of Mathematical Physics* (English translation), Vol. 1. Interscience, New York. Reprinted, Wiley (1989), Chap. 6, Section 3.

Bessel's Inequality

If the set of functions $\varphi_n(x)$ does not form a complete set, possibly because we have not included the required infinite number of members of an infinite set, we are led to an inequality. First, consider the case of a finite sum of components. Let A be an n component vector,

$$\mathbf{A} = \mathbf{e}_1 a_1 + \mathbf{e}_2 a_2 + \cdots + \mathbf{e}_n a_n, \tag{9.66}$$

in which $\mathbf{e}_i$ form a set of orthonormal unit vectors and a_i is the corresponding component (projection) of $\mathbf{A}$; that is,

$$a_i = \mathbf{A} \cdot \mathbf{e}_i. \tag{9.67}$$

Then

$$\left(\mathbf{A} - \sum_{i=1}^{n} \mathbf{e}_i a_i\right)^2 \geq 0. \tag{9.68}$$

If we sum over all n components, clearly the summation equals $\mathbf{A}$ by Eq. (9.66) and the equality holds. If, however, the summation does not include all n components, the inequality results. By expanding Eq. (9.68) and remembering that the orthogonal unit vectors satisfy orthogonality relations

$$\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}, \tag{9.69}$$

we have

$$A^2 \geq \sum_i a_i^2. \tag{9.70}$$

This is **Bessel's inequality**.

For functions we consider the integral

$$\int_a^b \left[f(x) - \sum_{i=1}^{n} a_i \varphi_i(x)\right]^2 w(x)\,dx \geq 0. \tag{9.71}$$

This is the continuum analog of Eq. (9.68), letting $n \to \infty$ and replacing the summation by an integration. Again, with the weighting factor $w(x) > 0$, the integrand is nonnegative. The integral vanishes by Eq. (9.62) if we have a complete set. Otherwise it is positive. Expanding the squared term, we obtain

$$\int_a^b [f(x)]^2 w(x)dx - 2\sum_{i=1}^{n} a_i \int_a^b f(x)\varphi_i(x)w(x)dx + \sum_{i=1}^{n} a_i^2 \geq 0. \tag{9.72}$$

Applying Eq. (9.65), we have

$$\int_a^b [f(x)]^2 w(x)dx \geq \sum_{i=1}^{n} a_i^2. \tag{9.73}$$

Hence, the sum of the squares of the expansion coefficients a_i is less than or equal to the weighted integral of $[f(x)]^2$, the equality holding if and only if the expansion is exact—that is, if the set of functions $\varphi_n(x)$ is a complete set and $n \to \infty$.

In later chapters, when we consider eigenfunctions that form complete sets (such as Legendre polynomials), Eq. (9.73) with the equal sign holding is called a Parseval relation.

Bessel's inequality has a variety of uses, including proof of convergence of the Fourier series.

Schwarz Inequality

The frequently used Schwarz inequality is similar to the Bessel inequality. Consider the quadratic equation with unknown x:

$$\sum_{i=1}^{n}(a_i x + b_i)^2 = \sum_{i=1}^{n} a_i^2 \left(x + \frac{b_i}{a_i}\right)^2 = 0. \tag{9.74}$$

If $b_i/a_i =$ constant c, then the solution is $x = -c$. If b_i/a_i is not a constant, all terms cannot vanish simultaneously for real x. Therefore, the solution must be complex. Expanding, we find that

$$x^2 \sum_{i=1}^{n} a_i^2 + 2x \sum_{i=1}^{n} a_i b_i + \sum_{i=1}^{n} b_i^2 = 0, \tag{9.75}$$

and since x is complex (or $= -b_i/a_i$), the quadratic formula[12] for x leads to

$$\left(\sum_{i=1}^{n} a_i b_i\right)^2 \le \left(\sum_{i=1}^{n} a_i^2\right)\left(\sum_{i=1}^{n} b_i^2\right), \tag{9.76}$$

the equality holding when b_i/a_i equals a constant.

Once more, in terms of vectors, we have

$$(\mathbf{a} \cdot \mathbf{b})^2 = a^2 b^2 \cos^2\theta \le a^2 b^2, \tag{9.77}$$

where θ is the angle included between $\mathbf{a}$ and $\mathbf{b}$.

The analogous Schwarz inequality for functions has the form

$$\left|\int_a^b f^*(x) g(x) w(x)\, dx\right|^2 \le \int_a^b f^*(x) f(x) w(x)\, dx \int_a^b g^*(x) g(x) w(x)\, dx, \tag{9.78}$$

the equality holding if and only if $g(x) = \alpha f(x)$, with α being a constant. To prove this function form of the Schwarz inequality,[13] consider a complex function $\psi(x) = f(x) + \lambda g(x)$ with λ a complex constant. The functions $f(x)$ and $g(x)$ are any two functions (for which the integrals exist). Multiplying by

[12] With discriminant $b^2 - 4ac$ negative (or zero).

[13] An alternate derivation is provided by the inequality $\int\int [f(x)g(y) - f(y)g(x)]^*[f(x)g(y) - f(y)g(x)]w(x)w(y)dxdy \ge 0$.

the complex conjugate and integrating, we obtain

$$\int_a^b \psi^*\psi w(x)dx \equiv \int_a^b f^* f w(x)dx + \lambda \int_a^b f^* g w(x)dx + \lambda^* \int_a^b g^* f w(x)dx + \lambda\lambda^* \int_a^b g^* g w(x)dx \geq 0. \tag{9.79}$$

The ≥ 0 appears since $\psi^*\psi$ is nonnegative, the equal ($=$) sign holding only if $\psi(x)$ is identically zero. Noting that λ and λ^* are linearly independent, we differentiate with respect to one of them and set the derivative equal to zero to minimize $\int_a^b \psi^*\psi dx$:

$$\frac{\partial}{\partial \lambda^*} \int_a^b \psi^*\psi w(x)dx = \int_a^b g^* f w(x)dx + \lambda \int_a^b g^* g w(x)dx = 0.$$

This yields

$$\lambda = -\frac{\int_a^b g^* f w(x)dx}{\int_a^b g^* g w(x)dx}. \tag{9.80a}$$

Taking the complex conjugate, we obtain

$$\lambda^* = -\frac{\int_a^b f^* g w(x)dx}{\int_a^b g^* g w(x)dx}. \tag{9.80b}$$

Substituting these values of λ and λ^* back into Eq. (9.79), we obtain Eq. (9.78), the Schwarz inequality.

In quantum mechanics $f(x)$ and $g(x)$ might each represent a state or configuration of a physical system. Then the Schwarz inequality guarantees that the inner product $\int_a^b f^*(x)g(x)w(x)dx$ exists. In some texts, the Schwarz inequality is a key step in the derivation of the Heisenberg uncertainty principle.

The function notation of Eqs. (9.78) and (9.79) is relatively cumbersome. In advanced mathematical physics, and especially in quantum mechanics, it is common to use the Dirac bra-ket notation:

$$\langle f|g\rangle \equiv \int_a^b f^*(x)g(x)w(x)dx.$$

Using this new notation, we simply understand the range of integration, (a, b), and any weighting function. In this notation the Schwarz inequality becomes

$$|\langle f|g\rangle|^2 \leq \langle f|f\rangle\langle g|g\rangle. \tag{9.78a}$$

If $g(x)$ is a normalized eigenfunction, $\varphi_i(x)$, Eq. (9.78) yields

$$a_i^* a_i \leq \int_a^b f^*(x)f(x)w(x)dx, \tag{9.81}$$

a result that also follows from Eq. (9.73).

Summary of Vector Spaces—Completeness

Here we summarize some properties of vector space, first with the vectors taken to be the familiar real vectors of Chapter 1 and then with the vectors taken to be ordinary functions—polynomials. The concept of completeness is developed for finite vector spaces and carried over into infinite vector spaces.

1v. We shall describe our vector space with a set of n linearly independent vectors $\mathbf{e}_i$, $i = 1, 2, \ldots, n$. If $n = 3$, $\mathbf{e}_1 = \hat{\mathbf{x}}$, $\mathbf{e}_2 = \hat{\mathbf{y}}$, and $\mathbf{e}_3 = \hat{\mathbf{z}}$. The n $\mathbf{e}_i$ **span** the linear vector space and are defined to be a **basis**.

1f. We shall describe our vector (function) space with a set of n linearly independent functions, $\varphi_i(x)$, $i = 0, 1, \ldots, n-1$. The index i starts with 0 to agree with the labeling of the classical polynomials. Here, $\varphi_i(x)$ is assumed to be a polynomial of degree i. The n $\varphi_i(x)$ span the linear vector (function) space forming a **basis**.

2v. The vectors in our vector space satisfy the following relations (Section 1.2; the vector components are numbers):

a. Vector addition is commutative	$\mathbf{u}+\mathbf{v}=\mathbf{v}+\mathbf{u}$
b. Vector addition is associative	$[\mathbf{u}+\mathbf{v}]+\mathbf{w}=\mathbf{u}+[\mathbf{v}+\mathbf{w}]$
c. There is a null vector	$\mathbf{0}+\mathbf{v}=\mathbf{v}$
d. Multiplication by a scalar	
Distributive	$a[\mathbf{u}+\mathbf{v}]=a\mathbf{u}+a\mathbf{v}$
Distributive	$(a+b)\mathbf{u}=a\mathbf{u}+b\mathbf{u}$
Associative	$a[b\mathbf{u}]=(ab)\mathbf{u}$
e. Multiplication	
By unit scalar	$1\mathbf{u}=\mathbf{u}$
By zero	$0\mathbf{u}=0$
f. Negative vector	$(-1)\mathbf{u}=-\mathbf{u}$.

2f. The functions in our linear function space satisfy the properties listed for vectors (substitute "function" for "vector"):

a. $$f(x)+g(x)=g(x)+f(x)$$

b. $$[f(x)+g(x)]+h(x)=f(x)+[g(x)+h(x)]$$

c. $$0+f(x)=f(x)$$

d. $$a[f(x)+g(x)]=af(x)+ag(x)$$
$$(a+b)f(x)=af(x)+bf(x)$$
$$a[bf(x)]=(ab)f(x)$$

e. $$1\cdot f(x)=f(x)$$
$$0\cdot f(x)=0$$

f. $$(-1)\cdot f(x)=-f(x).$$

3v. In n-dimensional vector space an arbitrary vector $\mathbf{c}$ is described by its n components $(c_1, c_2, \ldots, c_n)$ or

$$\mathbf{c} = \sum_{i=1}^{n} c_i \mathbf{e}_i, \quad c_i = \hat{\mathbf{e}}_i \cdot \mathbf{c}.$$

When (i) n $\mathbf{e}_i$ are linearly independent and (ii) span the n-dimensional vector space, then the $\mathbf{e}_i$ form a basis and constitute a **complete** set.

3f. In n-dimensional function space a polynomial of degree $m \leq n-1$ is described by

$$f(x) = \sum_{i=0}^{n-1} c_i \varphi_i(x), \quad c_i = \frac{\langle \varphi_i | f \rangle}{\langle \varphi_i | \varphi_i \rangle}.$$

When (i) the $n\varphi_i(x)$ are linearly independent and (ii) span the n-dimensional function space, then the $\varphi_i(x)$ form a basis and constitute a **complete** set (for describing polynomials of degree $m \leq n-1$).

4v. An inner product (scalar, dot product) is defined by

$$\mathbf{c} \cdot \mathbf{d} = \sum_{i=1}^{n} c_i d_i.$$

If $\mathbf{c}$ and $\mathbf{d}$ have complex components, the inner product is defined as $\sum_{i=1}^{n} c_i^* d_i$. The inner product has the properties of

a. Distributive law of addition $\quad \mathbf{c} \cdot (\mathbf{d} + \mathbf{e}) = \mathbf{c} \cdot \mathbf{d} + \mathbf{c} \cdot \mathbf{e}$

b. Scalar multiplication $\quad \mathbf{c} \cdot a\mathbf{d} = a\mathbf{c} \cdot \mathbf{d}$

c. Complex conjugation $\quad \mathbf{c} \cdot \mathbf{d} = (\mathbf{d} \cdot \mathbf{c})^*$.

4f. An inner product is defined by

$$\langle f | g \rangle = \int_a^b f^*(x) g(x) w(x) dx.$$

The choice of the weighting function $w(x)$ and the interval (a, b) follows from the differential equation satisfied by $\varphi_i(x)$ and the boundary conditions (Section 9.1). In matrix terminology (Section 3.2), $|g\rangle$ is a column vector and $\langle f|$ is a row vector, the adjoint of $|f\rangle$.

The inner product has the properties listed for vectors:

a. $\langle f | g + h \rangle = \langle f | g \rangle + \langle f | h \rangle$

b. $\langle f | ag \rangle = a \langle f | g \rangle$

c. $\langle f | g \rangle = \langle g | f \rangle^*$.

5v. Orthogonality:

$$\mathbf{e}_j \cdot \mathbf{e}_j = 0, \quad i \neq j.$$

If the n $\mathbf{e}_i$ are not already orthogonal, the Gram–Schmidt process may be used to create an orthogonal set.

5f. Orthogonality:

$$\langle \varphi_i | \varphi_j \rangle = \int_a^b \varphi_i^*(x) \varphi_j(x) w(x) dx = 0, \quad i \neq j.$$

If the n $\varphi_i(x)$ are not already orthogonal, the Gram–Schmidt process (Section 9.3) may be used to create an orthogonal set.

6v. Definition of norm:

$$|\mathbf{c}| = (\mathbf{c}\cdot\mathbf{c})^{1/2} = \left(\sum_{i=1}^{n} c_i^2\right)^{1/2}.$$

The basis vectors $\mathbf{e}_i$ are taken to have unit norm (length) $\mathbf{e}_i \cdot \mathbf{e}_i = 1$. The components of $\mathbf{c}$ are given by

$$c_i = \mathbf{e}_i \cdot \mathbf{c}, \quad i = 1, 2, \ldots, n.$$

6f. Definition of norm:

$$\|f\| = \langle f|f\rangle^{1/2} = \left[\int_a^b |f(x)|^2 w(x)\,dx\right]^{1/2} = \left[\sum_{i=0}^{n-1} |c_i|^2\right]^{1/2},$$

Parseval's identity: $\|f\| > 0$ unless $f(x)$ is identically zero. The basis functions $\varphi_i(x)$ may be taken to have unit norm (unit normalization)

$$\|\varphi_i\| = 1.$$

Note that Legendre polynomials are not normalized to unity.

The expansion coefficients of our polynomial $f(x)$ are given by

$$c_i = \langle \varphi_i | f\rangle, \quad i = 0, 1, \ldots, n-1.$$

7v. Bessel's inequality:

$$\mathbf{c}\cdot\mathbf{c} \geq \sum_i c_i^2.$$

If the equal sign holds for all $\mathbf{c}$, it indicates that the $\mathbf{e}_i$ span the vector space; that is, they are complete.

7f. Bessel's inequality:

$$\langle f|f\rangle = \int_a^b |f(x)|^2 w(x)\,dx \geq \sum_i |c_i|^2.$$

If the equal sign holds for all allowable f, it indicates that the $\varphi_i(x)$ span the function space; that is, they are complete.

8v. Schwarz inequality:

$$\mathbf{c}\cdot\mathbf{d} \leq |\mathbf{c}|\cdot|\mathbf{d}|.$$

The equal sign holds when $\mathbf{c}$ is a multiple of $\mathbf{d}$. If the angle included between $\mathbf{c}$ and $\mathbf{d}$ is θ, then $|\cos\theta| \leq 1$.

8f. Schwarz inequality:

$$|\langle f|g\rangle| \leq \langle f|f\rangle^{1/2}\langle g|g\rangle^{1/2} = \|f\|\cdot\|g\|.$$

The equals sign holds when $f(x)$ and $g(x)$ are linearly dependent; that is, when $f(x)$ is a multiple of $g(x)$.

9v. Now, let $n \to \infty$, forming an infinite-dimensional linear vector space, l^2.

In an infinite-dimensional space our vector $\mathbf{c}$ is

$$\mathbf{c} = \sum_{i=1}^{\infty} c_i \mathbf{e}_i.$$

We require that

$$\sum_{i=1}^{\infty} c_i^2 < \infty.$$

The components of $\mathbf{c}$ are given by

$$c_i = \mathbf{e}_i \cdot \mathbf{c}, \quad i = 1, 2, \ldots, \infty,$$

exactly as in a finite-dimensional vector space.

9f. Then let $n \to \infty$, forming an infinite-dimensional vector (function) space, L^2. Then the superscript 2 stands for the quadratic norm [i.e., the 2 in $|f(x)|^2$]. Our functions need no longer be polynomials, but we do require that $f(x)$ be at least piecewise continuous (Dirichlet conditions for Fourier series) and that $\langle f|f\rangle = \int_a^b |f(x)|^2 w(x)dx$ exist. This latter condition is often stated as a requirement that $f(x)$ be square integrable.

Cauchy sequence: Let

$$f_n(x) = \sum_{i=0}^{n} c_i \varphi_i(x).$$

If

$$\|f(x) - f_n(x)\| \to 0 \quad \text{as} \quad n \to \infty$$

or

$$\lim_{n\to\infty} \int \left| f(x) - \sum_{i=0}^{n} c_i \varphi_i x \right|^2 w(x)dx = 0,$$

then we have convergence in the mean. This is analogous to the partial sum–Cauchy sequence criterion for the convergence of an infinite series (Section 5.1).

If every Cauchy sequence of allowable vectors (square integrable, piecewise continuous functions) converges to a limit vector in our linear space, the space is said to be complete. Then

$$f(x) = \sum_{i=0}^{\infty} c_i \varphi_i(x) \quad \text{(almost everywhere)}$$

in the sense of convergence in the mean. As noted previously, this is a weaker requirement than pointwise convergence (fixed value of x) or uniform convergence.

Expansion (Fourier) Coefficients

For a function f its Fourier coefficients are defined as

$$c_i = \langle \varphi_i | f \rangle, \quad i = 0, 1, \ldots, \infty,$$

exactly as in a finite-dimensional vector space. Hence,

$$f(x) = \sum_i \langle \varphi_i | f \rangle \varphi_i(x).$$

A linear space (finite- or infinite-dimensional) that (i) has an inner product defined, $(f \mid g)$, and (ii) is complete is a **Hilbert space**.

Infinite-dimensional Hilbert space provides a natural mathematical framework for modern quantum mechanics because bound-state wave functions are normalized (square-integrable) and usually are eigenfunctions of some Hamiltonian that provides a basis of the Hilbert space. A physical state may be expanded in a set of basis vectors, which are eigenstates of some observable. The expansion coefficients squared give the probabilities of the different eigenvalues of the observable in the given state. Apart from quantum mechanics, Hilbert space retains its abstract mathematical power and beauty, but the necessity for its use is reduced.

SUMMARY

The Sturm–Liouville theory of second-order ODEs with boundary conditions leads to eigenvalue problems whose solutions are eigenfunctions with orthogonality properties. Special functions, such as Legendre polynomials, Bessel functions, and Laguerre polynomials, arise in this context. Eigenfunction expansions are important in quantum mechanics and many other areas of physics and engineering.

Biographical Data

Hilbert, David. Hilbert, a German mathematician, was born in 1862 in Königsberg, Prussia (now Russia), and died in 1943 in Göttingen, Germany. Son of a judge, he obtained his Ph.D. in mathematics at the University of Königsberg in 1885 and became a professor at Göttingen in 1895. In 1899, in his *Foundations of Geometry* he established the first consistent set of geometric axioms, which helped the axiomatic method for the foundation of mathematics to gain general recognition. He contributed to most active branches of mathematics and, with Poincaré, is considered one of the greatest mathematicians of the 20th century. He solved Waring's problem in number theory, developed solutions for integral equations, and is famous for an influential list of unsolved mathematics problems presented in 1900 at the International Congress of Mathematicians in Paris, which deeply influenced the development of mathematics in the 20th century.

EXERCISES

9.4.1 A function $f(x)$ is expanded in a series of orthonormal eigenfunctions

$$f(x) = \sum_{n=0}^{\infty} a_n \varphi_n(x).$$

Show that the series expansion is unique for a given set of $\varphi_n(x)$. The functions $\varphi_n(x)$ are being taken here as the **basis** vectors in an infinite-dimensional Hilbert space.

9.4.2 A function $f(x)$ is represented by a finite set of basis functions $\varphi_i(x)$,

$$f(x) = \sum_{i=1}^{N} c_i \varphi_i(x).$$

Show that the components c_i are unique; that no different set c_i' exists.
Note. Your basis functions are automatically linearly independent. They are not necessarily orthogonal.

9.4.3 A function $f(x)$ is approximated by a power series $\sum_{i=0}^{n-1} c_i x^i$ over the interval [0, 1]. Show that minimizing the mean square error leads to a set of linear equations

$$\mathsf{A}\mathbf{c} = \mathbf{b},$$

where

$$A_{ij} = \int_0^1 x^{i+j} dx = \frac{1}{i+j+1}, \quad i, j = 0, 1, 2, \ldots, n-1$$

and

$$b_i = \int_0^1 x^i f(x) dx, \quad i = 0, 1, 2, \ldots, n-1.$$

Note. The A_{ij} are the elements of the Hilbert matrix of order n. The determinant of this Hilbert matrix is a rapidly decreasing function of n. For $n = 5$, $\det \mathsf{A} = 3.7 \times 10^{-12}$ and the set of equations $\mathsf{A}\mathbf{c} = \mathbf{b}$ is becoming ill conditioned and unstable.

9.4.4 In place of the expansion of a function $F(x)$ given by

$$F(x) = \sum_{n=0}^{\infty} a_n \varphi_n(x),$$

with

$$a_n = \int_a^b F(x) \varphi_n(x) w(x) dx,$$

take the **finite** series approximation

$$F(x) \approx \sum_{n=0}^{m} c_n \varphi_n(x).$$

Show that the mean square error

$$\int_a^b \left[F(x) - \sum_{n=0}^{m} c_n \varphi_n(x) \right]^2 w(x) dx$$

is minimized by taking $c_n = a_n$.
Note. The values of the coefficients are independent of the number of terms in the finite series. This independence is a consequence of orthogonality and would not hold for a least-squares fit using powers of x.

9.4.5 From Example 9.2.2,

$$f(x) = \left\{ \begin{array}{ll} h/2, & 0 < x < \pi \\ -h/2, & -\pi < x < 0 \end{array} \right\} = \frac{2h}{\pi} \sum_{n=0}^{\infty} \frac{\sin(2n+1)x}{2n+1}.$$

(a) Show that

$$\int_{-\pi}^{\pi} [f(x)]^2 dx = \frac{\pi}{2} h^2 = \frac{4h^2}{\pi} \sum_{n=0}^{\infty} (2n+1)^{-2}.$$

For a finite upper limit, this would be Bessel's inequality. For the upper limit, ∞, this is Parseval's identity.

(b) Verify that

$$\frac{\pi}{2} h^2 = \frac{4h^2}{\pi} \sum_{n=0}^{\infty} (2n+1)^{-2}$$

by evaluating the series.

Hint. The series can be expressed as the Riemann zeta function.

9.4.6 Differentiate Eq. (9.79),

$$\langle \psi | \psi \rangle = \langle f | f \rangle + \lambda \langle f | g \rangle + \lambda^* \langle g | f \rangle + \lambda \lambda^* \langle g | g \rangle,$$

with respect to λ^* and show that you get the Schwarz inequality [Eq. (9.78)].

9.4.7 Derive the Schwarz inequality from the identity

$$\left[\int_a^b f(x) g(x)\, dx \right]^2 = \int_a^b [f(x)]^2 dx \int_a^b [g(x)]^2 dx - \frac{1}{2} \int_a^b \int_a^b [f(x)g(y) - f(y)g(x)]^2 dx\, dy.$$

9.4.8 If the functions $f(x)$ and $g(x)$ of the Schwarz inequality [Eq. (9.78)] may be expanded in a series of eigenfunctions $\varphi_i(x)$, show that Eq. (9.78) reduces to Eq. (9.76) (with n possibly infinite).

Note the description of $f(x)$ as a vector in a function space in which $\varphi_i(x)$ corresponds to the unit vector $\mathbf{e}_1$.

9.4.9 The operator H is Hermitian and positive definite; that is,

$$\int_a^b f^* H f\, dx > 0.$$

Prove the generalized Schwarz inequality:

$$\left| \int_a^b f^* H g\, dx \right|^2 \le \int_a^b f^* H f\, dx \int_a^b g^* H g\, dx.$$

9.4.10 A normalized wave function $\psi(x) = \sum_{n=0}^{\infty} a_n \varphi_n(x)$. The expansion coefficients a_n are known as probability amplitudes. We may define a density matrix ρ with elements $\rho_{ij} = a_i a_j^*$. Show that

$$(\rho^2)_{ij} = \rho_{ij}$$

or

$$\rho^2 = \rho.$$

This result, by definition, makes ρ a projection operator.
Hint. Use

$$\int \psi^* \psi dx = 1.$$

9.4.11 Show that
(a) the operator

$$|\varphi_i\rangle\langle\varphi_i|$$

operating on

$$f(t) = \sum_j c_j \langle t|\varphi_j\rangle$$

yields $c_i|\varphi_i(t)\rangle, \quad \langle t|\varphi_j\rangle \equiv \varphi_j(t).$

(b) $\sum_i |\varphi_i\rangle\langle\varphi_i| = 1.$

(c) For $\langle x|\varphi_p\rangle = e^{ipx}/\sqrt{2\pi}$ derive from (a) the Fourier integral of the function $f(t)$ and from (b) the Fourier integral representation of Dirac's δ function (see Chapter 1). Note that p is a continuous variable (momentum) replacing the discrete index i.

This operator is a **projection operator** projecting $f(x)$ onto the ith coordinate, selectively picking out the ith component $c_i|\varphi_i\rangle$ of $f(x)$.

Hint. The operator operates via the well-defined inner product. In the coordinate representation you actually have to work with $\langle x|\varphi_i\rangle\langle\varphi_i|t\rangle = \varphi_i(x)^\dagger \varphi_i(t)$.

Additional Reading

Byron, F. W., Jr., and Fuller, R. W. (1969). *Mathematics of Classical and Quantum Physics*. Addison-Wesley, Reading, MA.

Dennery, P., and Krzywicki, A. (1996). *Mathematics for Physicists* (reprinted). Dover, New York.

Hirsch, M. (1974). *Differential Equations, Dynamical Systems, and Linear Algebra*. Academic Press, San Diego.

Miller, K. S. (1963). *Linear Differential Equations in the Real Domain*. Norton, New York.

Titchmarsh, E. C. (1962). *Eigenfunction Expansions Associated with Second Order Differential Equations*, 2nd ed., Vol. 1. Oxford Univ. Press, London