17

Eigenfunction methods for differential equations

In the previous three chapters we dealt with the solution of differential equations of order n by two methods. In one method, we found n independent solutions of the equation and then combined them, weighted with coefficients determined by the boundary conditions; in the other we found solutions in terms of series whose coefficients were related by (in general) an n-term recurrence relation and thence fixed by the boundary conditions. For both approaches the linearity of the equation was an important or essential factor in the utility of the method, and in this chapter our aim will be to exploit the superposition properties of linear differential equations even further.

We will be concerned with the solution of equations of the inhomogeneous form

$$\mathcal{L}y(x) = f(x), \tag{17.1}$$

where $f(x)$ is a prescribed or general function *and* the boundary conditions to be satisfied by the solution $y = y(x)$, for example at the limits $x = a$ and $x = b$, are given. The expression $\mathcal{L}y(x)$ stands for a linear differential operator $\mathcal{L}$ acting upon the function $y(x)$.

In general, unless $f(x)$ is both known and simple, it will not be possible to find particular integrals of (17.1), even if complementary functions can be found that satisfy $\mathcal{L}y = 0$. The idea is therefore to exploit the linearity of $\mathcal{L}$ by building up the required solution $y(x)$ as a *superposition*, generally containing an infinite number of terms, of some set of functions $\{y_i(x)\}$ that each individually satisfy the boundary conditions. Clearly this brings in a quite considerable complication but since, within reason, we may select the set of functions to suit ourselves, we can obtain sizeable compensation for this complication. Indeed, if the set chosen is one containing functions that, when acted upon by $\mathcal{L}$, produce particularly simple results then we can 'show a profit' on the operation. In particular, if the set consists of those functions y_i for which

$$\mathcal{L}y_i(x) = \lambda_i y_i(x), \tag{17.2}$$

where λ_i is a constant (and which satisfy the boundary conditions), then a distinct advantage may be obtained from the manoeuvre because all the differentiation will have disappeared from (17.1).

Equation (17.2) is clearly reminiscent of the equation satisfied by the *eigenvectors* $\mathbf{x}^i$ of a linear operator $\mathcal{A}$, namely

$$\mathcal{A}\mathbf{x}^i = \lambda_i \mathbf{x}^i, \tag{17.3}$$

where λ_i is a constant and is called the *eigenvalue* associated with $\mathbf{x}^i$. By analogy, in the context of differential equations a function $y_i(x)$ satisfying (17.2) is called an *eigenfunction* of the operator $\mathcal{L}$ (under the imposed boundary conditions) and λ_i is then called the eigenvalue associated with the eigenfunction $y_i(x)$. Clearly, the eigenfunctions $y_i(x)$ of $\mathcal{L}$ are only determined up to an arbitrary scale factor by (17.2).

Probably the most familiar equation of the form (17.2) is that which describes a simple harmonic oscillator, i.e.

$$\mathcal{L}y \equiv -\frac{d^2y}{dt^2} = \omega^2 y, \qquad \text{where } \mathcal{L} \equiv -d^2/dt^2. \tag{17.4}$$

Imposing the boundary condition that the solution is periodic with period T, the eigenfunctions in this case are given by $y_n(t) = A_n e^{i\omega_n t}$, where $\omega_n = 2\pi n/T$, $n = 0, \pm1, \pm2, \ldots$ and the A_n are constants. The eigenvalues are $\omega_n^2 = n^2\omega_1^2 = n^2(2\pi/T)^2$. (Sometimes ω_n is referred to as the eigenvalue of this equation, but we will avoid such confusing terminology here.)

We may discuss a somewhat wider class of differential equations by considering a slightly more general form of (17.2), namely

$$\mathcal{L}y_i(x) = \lambda_i \rho(x) y_i(x), \tag{17.5}$$

where $\rho(x)$ is a *weight function*. In many applications $\rho(x)$ is unity for all x, in which case (17.2) is recovered; in general, though, it is a function determined by the choice of coordinate system used in describing a particular physical situation. The only requirement on $\rho(x)$ is that it is real and does not change sign in the range $a \le x \le b$, so that it can, without loss of generality, be taken to be non-negative throughout; of course, $\rho(x)$ must be the same function for all values of λ_i. A function $y_i(x)$ that satisfies (17.5) is called an eigenfunction of the operator $\mathcal{L}$ with respect to the weight function $\rho(x)$.

This chapter will not cover methods used to determine the eigenfunctions of (17.2) or (17.5), since we have discussed those in previous chapters, but, rather, will use the properties of the eigenfunctions to solve inhomogeneous equations of the form (17.1). We shall see later that the sets of eigenfunctions $y_i(x)$ of a particular class of operators called *Hermitian operators* (the operator in the simple harmonic oscillator equation is an example) have particularly useful properties and these will be studied in detail. It turns out that many of the interesting differential operators met within the physical sciences are Hermitian. Before continuing our discussion of the eigenfunctions of Hermitian operators, however, we will consider some properties of general sets of functions.

17.1 Sets of functions

In chapter 8 we discussed the definition of a vector space but concentrated on spaces of finite dimensionality. We consider now the *infinite*-dimensional space of all reasonably well-behaved functions $f(x)$, $g(x)$, $h(x)$, $\ldots$ on the interval $a \leq x \leq b$. That these functions form a linear vector space is shown by noting the following properties. The set is closed under

(i) addition, which is commutative and associative, i.e.

$$f(x) + g(x) = g(x) + f(x),$$
$$[f(x) + g(x)] + h(x) = f(x) + [g(x) + h(x)]\,,$$

(ii) multiplication by a scalar, which is distributive and associative, i.e.

$$\lambda\,[f(x) + g(x)] = \lambda f(x) + \lambda g(x),$$
$$\lambda\,[\mu f(x)] = (\lambda\mu) f(x),$$
$$(\lambda + \mu) f(x) = \lambda f(x) + \mu f(x).$$

Furthermore, in such a space

(iii) there exists a 'null vector' 0 such that $f(x) + 0 = f(x)$,
(iv) multiplication by unity leaves any function unchanged, i.e. $1 \times f(x) = f(x)$,
(v) each function has an associated negative function $-f(x)$ that is such that $f(x) + [-f(x)] = 0$.

By analogy with finite-dimensional vector spaces we now introduce a set of linearly independent basis functions $y_n(x)$, $n = 0, 1, \ldots, \infty$, such that *any* 'reasonable' function in the interval $a \leq x \leq b$ (i.e. it obeys the Dirichlet conditions discussed in chapter 12) can be expressed as the linear sum of these functions:

$$f(x) = \sum_{n=0}^{\infty} c_n y_n(x).$$

Clearly if a different set of linearly independent basis functions $u_n(x)$ is chosen then the function can be expressed in terms of the new basis,

$$f(x) = \sum_{n=0}^{\infty} d_n u_n(x),$$

where the d_n are a different set of coefficients. In each case, provided the basis functions are linearly independent, the coefficients are unique.

We may also define an *inner product* on our function space by

$$\langle f|g\rangle = \int_a^b f^*(x)g(x)\rho(x)\,dx, \tag{17.6}$$

where $\rho(x)$ is the weight function, which we require to be real and non-negative in the interval $a \le x \le b$. As mentioned above, $\rho(x)$ is often unity for all x. Two functions are said to be *orthogonal* (with respect to the weight function $\rho(x)$) on the interval $[a, b]$ if

$$\langle f|g\rangle = \int_a^b f^*(x)g(x)\rho(x)\,dx = 0, \tag{17.7}$$

and the *norm* of a function is defined as

$$\|f\| = \langle f|f\rangle^{1/2} = \left[\int_a^b f^*(x)f(x)\rho(x)\,dx\right]^{1/2} = \left[\int_a^b |f(x)|^2\rho(x)\,dx\right]^{1/2}. \tag{17.8}$$

It is also common practice to define a *normalised* function by $\hat{f} = f/\|f\|$, which has unit norm.

An infinite-dimensional vector space of functions, for which an inner product is defined, is called a *Hilbert space*. Using the concept of the inner product, we can choose a basis of linearly independent functions $\hat{\phi}_n(x)$, $n = 0, 1, 2, \ldots$ that are orthonormal, i.e. such that

$$\langle\hat{\phi}_i|\hat{\phi}_j\rangle = \int_a^b \hat{\phi}_i^*(x)\hat{\phi}_j(x)\rho(x)\,dx = \delta_{ij}. \tag{17.9}$$

If $y_n(x)$, $n = 0, 1, 2, \ldots$, are a linearly independent, but not orthonormal, basis for the Hilbert space then an orthonormal set of basis functions $\hat{\phi}_n$ may be produced (in a similar manner to that used in the construction of a set of orthogonal eigenvectors of an Hermitian matrix; see chapter 8) by the following procedure:

$$\begin{aligned}
\phi_0 &= y_0,\\
\phi_1 &= y_1 - \hat{\phi}_0\langle\hat{\phi}_0|y_1\rangle,\\
\phi_2 &= y_2 - \hat{\phi}_1\langle\hat{\phi}_1|y_2\rangle - \hat{\phi}_0\langle\hat{\phi}_0|y_2\rangle,\\
&\;\vdots\\
\phi_n &= y_n - \hat{\phi}_{n-1}\langle\hat{\phi}_{n-1}|y_n\rangle - \cdots - \hat{\phi}_0\langle\hat{\phi}_0|y_n\rangle,\\
&\;\vdots
\end{aligned}$$

It is straightforward to check that each ϕ_n is orthogonal to all its predecessors ϕ_i, $i = 0, 1, 2, \ldots, n-1$. This method is called *Gram–Schmidt orthogonalisation*. Clearly the functions ϕ_n form an orthogonal set, but in general they do not have unit norms.

►*Starting from the linearly independent functions $y_n(x) = x^n$, $n = 0, 1, \ldots$, construct three orthonormal functions over the range $-1 < x < 1$, assuming a weight function of unity.*

The first unnormalised function ϕ_0 is simply equal to the first of the original functions, i.e.

$$\phi_0 = 1.$$

The normalisation is carried out by dividing by

$$\langle \phi_0 | \phi_0 \rangle^{1/2} = \left(\int_{-1}^{1} 1 \times 1 \, du \right)^{1/2} = \sqrt{2},$$

with the result that the first normalised function $\hat{\phi}_0$ is given by

$$\hat{\phi}_0 = \frac{\phi_0}{\sqrt{2}} = \sqrt{\tfrac{1}{2}}.$$

The second unnormalised function is found by applying the above Gram–Schmidt orthogonalisation procedure, i.e.

$$\phi_1 = y_1 - \hat{\phi}_0 \langle \hat{\phi}_0 | y_1 \rangle.$$

It can easily be shown that $\langle \hat{\phi}_0 | y_1 \rangle = 0$, and so $\phi_1 = x$. Normalising then gives

$$\hat{\phi}_1 = \phi_1 \left(\int_{-1}^{1} u \times u \, du \right)^{-1/2} = \sqrt{\tfrac{3}{2}}\, x.$$

The third unnormalised function is similarly given by

$$\begin{aligned} \phi_2 &= y_2 - \hat{\phi}_1 \langle \hat{\phi}_1 | y_2 \rangle - \hat{\phi}_0 \langle \hat{\phi}_0 | y_2 \rangle \\ &= x^2 - 0 - \tfrac{1}{3}, \end{aligned}$$

which, on normalising, gives

$$\hat{\phi}_2 = \phi_2 \left(\int_{-1}^{1} \left(u^2 - \tfrac{1}{3} \right)^2 du \right)^{-1/2} = \tfrac{1}{2} \sqrt{\tfrac{5}{2}} (3x^2 - 1).$$

By comparing the functions $\hat{\phi}_0$, $\hat{\phi}_1$ and $\hat{\phi}_2$ with the list in subsection 18.1.1, we see that this procedure has generated (multiples of) the first three Legendre polynomials. ◄

If a function is expressed in terms of an *orthonormal* basis $\hat{\phi}_n(x)$ as

$$f(x) = \sum_{n=0}^{\infty} c_n \hat{\phi}_n(x) \tag{17.10}$$

then the coefficients c_n are given by

$$c_n = \langle \hat{\phi}_n | f \rangle = \int_a^b \hat{\phi}_n^*(x) f(x) \rho(x) \, dx. \tag{17.11}$$

Note that this is true only if the basis is orthonormal.

17.1.1 Some useful inequalities

Since for a Hilbert space $\langle f|f\rangle \geq 0$, the inequalities discussed in subsection 8.1.3 hold. The proofs are not repeated here, but the relationships are listed for completeness.

(i) The Schwarz inequality states that

$$|\langle f|g\rangle| \leq \langle f|f\rangle^{1/2}\langle g|g\rangle^{1/2}, \tag{17.12}$$

where the equality holds when $f(x)$ is a scalar multiple of $g(x)$, i.e. when they are linearly dependent.

(ii) The triangle inequality states that

$$\|f+g\| \leq \|f\| + \|g\|, \tag{17.13}$$

where again equality holds when $f(x)$ is a scalar multiple of $g(x)$.

(iii) Bessel's inequality requires the introduction of an *orthonormal* basis $\hat{\phi}_n(x)$ so that any function $f(x)$ can be written as

$$f(x) = \sum_{n=0}^{\infty} c_n \hat{\phi}_n(x),$$

where $c_n = \langle \hat{\phi}_n|f\rangle$. Bessel's inequality then states that

$$\langle f|f\rangle \geq \sum_n |c_n|^2. \tag{17.14}$$

The equality holds if the summation is over all the basis functions. If some values of n are omitted from the sum then the inequality results (unless, of course, the c_n happen to be zero for all values of n omitted, in which case the equality remains).

17.2 Adjoint, self-adjoint and Hermitian operators

Having discussed general sets of functions, we now return to the discussion of eigenfunctions of linear operators. We begin by introducing the *adjoint* of an operator $\mathcal{L}$, denoted by $\mathcal{L}^\dagger$, which is defined by

$A = \frac{1}{\rho}\mathcal{L}$

$$* \quad \int_a^b f^*(x)\,[\underbrace{\mathcal{L}}_{\uparrow 1/\rho}\underbrace{g(x)}_{\uparrow \rho}]\,dx = \int_a^b [\underbrace{\mathcal{L}^\dagger}_{\uparrow 1/\rho} f(x)]^*\underbrace{g(x)}_{\uparrow \rho}\,dx + \text{boundary terms}, \tag{17.15}$$

$\frac{1}{\rho}\rho = 1$

where the boundary terms are evaluated at the end-points of the interval $[a, b]$. Thus, for any given linear differential operator $\mathcal{L}$, the adjoint operator $\mathcal{L}^\dagger$ can be found by repeated integration by parts.

An operator is said to be *self-adjoint* if $\mathcal{L}^\dagger = \mathcal{L}$. If, in addition, certain boundary conditions are met by the functions f and g on which a self-adjoint operator acts,

$* \ \langle f|Ag\rangle = \langle A^\dagger f|g\rangle$

operador $A = \frac{1}{\rho}\mathcal{L}$

* $\langle f | \frac{1}{\rho} \mathcal{L} g \rangle = \langle \frac{1}{\rho} \mathcal{L} f | g \rangle$

or by the operator itself, such that the boundary terms in (17.15) vanish, then the operator is said to be *Hermitian* over the interval $a \le x \le b$. Thus, in this case,

*
$$\int_a^b f^*(x)\,[\mathcal{L}g(x)]\;dx = \int_a^b [\mathcal{L}f(x)]^* g(x)\,dx. \qquad (17.16)$$

A little careful study will reveal the similarity between the definition of an Hermitian operator and the definition of an Hermitian matrix given in chapter 8.

►*Show that the linear operator $\mathcal{L} = d^2/dt^2$ is self-adjoint, and determine the required boundary conditions for the operator to be Hermitian over the interval t_0 to $t_0 + T$.*

Substituting into the LHS of the definition of the adjoint operator (17.15) and integrating by parts gives

$$\int_{t_0}^{t_0+T} f^* \frac{d^2 g}{dt^2}\,dt = \left[f^* \frac{dg}{dt} \right]_{t_0}^{t_0+T} - \int_{t_0}^{t_0+T} \frac{df^*}{dt}\frac{dg}{dt}\,dt.$$

Integrating the second term on the RHS by parts once more yields

$$\int_{t_0}^{t_0+T} f^* \frac{d^2 g}{dt^2}\,dt = \left[f^* \frac{dg}{dt} \right]_{t_0}^{t_0+T} + \left[-\frac{df^*}{dt} g \right]_{t_0}^{t_0+T} + \int_{t_0}^{t_0+T} g \frac{d^2 f^*}{dt^2}\,dt,$$

which, by comparison with (17.15), proves that $\mathcal{L}$ is a self-adjoint operator. Moreover, from (17.16), we see that $\mathcal{L}$ is an Hermitian operator over the required interval provided

$$\left[f^* \frac{dg}{dt} \right]_{t_0}^{t_0+T} = \left[\frac{df^*}{dt} g \right]_{t_0}^{t_0+T}. \blacktriangleleft$$

We showed in chapter 8 that the eigenvalues of Hermitian matrices are real and that their eigenvectors can be chosen to be orthogonal. Similarly, the eigenvalues of Hermitian operators are real and their eigenfunctions can be chosen to be orthogonal (we will prove these properties in the following section). Hermitian operators (or matrices) are often used in the formulation of quantum mechanics. The eigenvalues then give the possible measured values of an observable quantity such as energy or angular momentum, and the physical requirement that such quantities must be real is ensured by the reality of these eigenvalues. Furthermore, the infinite set of eigenfunctions of an Hermitian operator form a complete basis set over the relevant interval, so that it is possible to expand any function $y(x)$ obeying the appropriate conditions in an eigenfunction series over this interval:

$$y(x) = \sum_{n=0}^{\infty} c_n y_n(x), \qquad (17.17)$$

where the choice of suitable values for the c_n will make the sum arbitrarily close to $y(x)$.[§] These useful properties provide the motivation for a detailed study of Hermitian operators.

§ The proof of the completeness of the eigenfunctions of an Hermitian operator is beyond the scope of this book. The reader should refer, for example, to R. Courant and D. Hilbert, *Methods of Mathematical Physics* (New York: Interscience, 1953).

17.3 Properties of Hermitian operators

We now provide proofs of some of the useful properties of Hermitian operators. Again much of the analysis is similar to that for Hermitian matrices in chapter 8, although the present section stands alone. (Here, and throughout the remainder of this chapter, we will write out inner products in full. We note, however, that the inner product notation often provides a neat form in which to express results.)

17.3.1 *Reality of the eigenvalues*

Consider an Hermitian operator for which (17.5) is satisfied by at least two eigenfunctions $y_i(x)$ and $y_j(x)$, which have corresponding eigenvalues λ_i and λ_j, so that

$$\frac{1}{\rho}\mathcal{L}y_i = \lambda_i y_i \qquad \mathcal{L}y_i = \lambda_i \rho(x) y_i, \tag{17.18}$$

$$\mathcal{L}y_j = \lambda_j \rho(x) y_j, \tag{17.19}$$

where we have allowed for the presence of a weight function $\rho(x)$. Multiplying (17.18) by y_j^* and (17.19) by y_i^* and then integrating gives

$y_j^* \times \rho$

$$\langle y_j | \tfrac{1}{\rho}\mathcal{L} y_i \rangle = \int_a^b y_j^* \mathcal{L} y_i \, dx = \lambda_i \int_a^b y_j^* y_i \rho \, dx, = \lambda_i \langle y_j | y_i \rangle \tag{17.20}$$

$$\langle y_i | \tfrac{1}{\rho}\mathcal{L} y_j \rangle = \int_a^b y_i^* \mathcal{L} y_j \, dx = \lambda_j \int_a^b y_i^* y_j \rho \, dx. = \lambda_j \langle y_i | y_j \rangle \tag{17.21}$$

Remembering that we have required $\rho(x)$ to be real, the complex conjugate of (17.20) becomes

$$\int_a^b y_j (\mathcal{L} y_i)^* \, dx = \lambda_i^* \int_a^b y_i^* y_j \rho \, dx, = \lambda_i^* \langle y_i | y_j \rangle \tag{17.22}$$

and using the definition of an Hermitian operator (17.16) it follows that the LHS of (17.22) is equal to the LHS of (17.21). Thus

$$(\lambda_i^* - \lambda_j)\langle y_i | y_j \rangle = (\lambda_i^* - \lambda_j) \int_a^b y_i^* y_j \rho \, dx = 0. \tag{17.23}$$

If $i = j$ then $\lambda_i = \lambda_i^*$ (since $\int_a^b y_i^* y_i \rho \, dx \neq 0$), which is a statement that the eigenvalue λ_i is real.

17.3.2 *Orthogonality and normalisation of the eigenfunctions*

From (17.23), it is immediately apparent that two eigenfunctions y_i and y_j that correspond to different eigenvalues, i.e. such that $\underline{\lambda_i \neq \lambda_j}$, satisfy

$$\langle y_i | y_j \rangle = 0 \qquad \int_a^b y_i^* y_j \rho \, dx = 0, \tag{17.24}$$

which is a statement of the orthogonality of y_i and y_j.

If one (or more) of the eigenvalues is degenerate, however, we have different eigenfunctions corresponding to the same eigenvalue, and the proof of orthogonality is not so straightforward. Nevertheless, an orthogonal set of eigenfunctions may be constructed using the *Gram–Schmidt orthogonalisation* method mentioned earlier in this chapter and used in chapter 8 to construct a set of orthogonal eigenvectors of an Hermitian matrix. We repeat the analysis here for completeness.

Suppose, for the sake of our proof, that λ_0 is k-fold degenerate, i.e.

$$\mathcal{L}y_i = \lambda_0 \rho y_i \quad \text{for } i = 0, 1, \ldots, k-1, \tag{17.25}$$

but that λ_0 is different from any of λ_k, λ_{k+1}, etc. Then any linear combination of these y_i is also an eigenfunction with eigenvalue λ_0 since

$$\mathcal{L}z \equiv \mathcal{L}\sum_{i=0}^{k-1} c_i y_i = \sum_{i=0}^{k-1} c_i \mathcal{L} y_i = \sum_{i=0}^{k-1} c_i \lambda_0 \rho y_i = \lambda_0 \rho z. \tag{17.26}$$

If the y_i defined in (17.25) are not already mutually orthogonal then consider the new eigenfunctions z_i constructed by the following procedure, in which each of the new functions z_i is to be normalised, to give $\hat{z}_i$, before proceeding to the construction of the next one (the normalisation can be carried out by dividing the eigenfunction z_i by $(\int_a^b z_i^* z_i \rho\, dx)^{1/2}$):

$$\begin{aligned}
z_0 &= y_0, \\
z_1 &= y_1 - \left(\hat{z}_0 \int_a^b \hat{z}_0^* y_1 \rho\, dx\right), \\
z_2 &= y_2 - \left(\hat{z}_1 \int_a^b \hat{z}_1^* y_2 \rho\, dx\right) - \left(\hat{z}_0 \int_a^b \hat{z}_0^* y_2 \rho\, dx\right), \\
&\ \vdots \\
z_{k-1} &= y_{k-1} - \left(\hat{z}_{k-2} \int_a^b \hat{z}_{k-2}^* y_{k-1} \rho\, dx\right) - \cdots - \left(\hat{z}_0 \int_a^b \hat{z}_0^* y_{k-1} \rho\, dx\right).
\end{aligned}$$

Each of the integrals is just a number and thus each new function z_i is, as can be shown from (17.26), an eigenvector of $\mathcal{L}$ with eigenvalue λ_0. It is straightforward to check that each z_i is orthogonal to all its predecessors. Thus, by this explicit construction we have shown that an orthogonal set of eigenfunctions of an Hermitian operator $\mathcal{L}$ can be obtained. Clearly the orthogonal set obtained, z_i, is not unique.

In general, since $\mathcal{L}$ is linear, the normalisation of its eigenfunctions $y_i(x)$ is arbitrary. It is often convenient, however, to work in terms of the normalised eigenfunctions $\hat{y}_i(x)$, so that $\int_a^b \hat{y}_i^* \hat{y}_i \rho\, dx = 1$. These therefore form an orthonormal set and we can write

$$\int_a^b \hat{y}_i^* \hat{y}_j \rho \, dx = \delta_{ij}, \tag{17.27}$$

which is valid for all pairs of values i, j.

17.3.3 Completeness of the eigenfunctions

As noted earlier, the eigenfunctions of an Hermitian operator may be shown to form a complete basis set over the relevant interval. One may thus expand any (reasonable) function $y(x)$ obeying appropriate boundary conditions in an eigenfunction series over the interval, as in (17.17). Working in terms of the normalised eigenfunctions $\hat{y}_n(x)$, we may thus write

$$\begin{aligned} f(x) &= \sum_n \hat{y}_n(x) \int_a^b \hat{y}_n^*(z) f(z) \rho(z) \, dz \\ &= \int_a^b f(z)\rho(z) \sum_n \hat{y}_n(x)\hat{y}_n^*(z) \, dz. \end{aligned}$$

Since this is true for any $f(x)$, we must have that

$$\rho(z) \sum_n \hat{y}_n(x)\hat{y}_n^*(z) = \delta(x - z). \tag{17.28}$$

This is called the *completeness* or *closure* property of the eigenfunctions. It defines a complete set. If the spectrum of eigenvalues of $\mathcal{L}$ is anywhere continuous then the eigenfunction $y_n(x)$ must be treated as $y(n, x)$ and an integration carried out over n.

We also note that the RHS of (17.28) is a δ-function and so is only non-zero when $z = x$; thus $\rho(z)$ on the LHS can be replaced by $\rho(x)$ if required, i.e.

$$\rho(z) \sum_n \hat{y}_n(x)\hat{y}_n^*(z) = \rho(x) \sum_n \hat{y}_n(x)\hat{y}_n^*(z). \tag{17.29}$$

17.3.4 Construction of real eigenfunctions

Recall that the eigenfunction y_i satisfies

$$\mathcal{L}y_i = \lambda_i \rho y_i \tag{17.30}$$

and that the complex conjugate of this gives

$$\mathcal{L}y_i^* = \lambda_i^* \rho y_i^* = \lambda_i \rho y_i^*, \tag{17.31}$$

where the last equality follows because the eigenvalues are real, i.e. $\lambda_i = \lambda_i^*$. Thus, y_i and y_i^* are eigenfunctions corresponding to the same eigenvalue and hence, because of the linearity of $\mathcal{L}$, at least one of $y_i^* + y_i$ and $i(y_i^* - y_i)$, which are both real, is a non-zero eigenfunction corresponding to that eigenvalue. It follows that the eigenfunctions can always be made real by taking suitable linear combinations, though taking such linear combinations will only be necessary in cases where a particular λ is degenerate, i.e. corresponds to more than one linearly independent eigenfunction.

17.4 Sturm–Liouville equations

One of the most important applications of our discussion of Hermitian operators is to the study of *Sturm–Liouville equations*, which take the general form

$$p(x)\frac{d^2y}{dx^2} + r(x)\frac{dy}{dx} + q(x)y + \lambda\rho(x)y = 0, \qquad \text{where } r(x) = \frac{dp(x)}{dx} \tag{17.32}$$

and p, q and r are real functions of x.§ A variational approach to the Sturm–Liouville equation, which is useful in estimating the eigenvalues λ for a given set of boundary conditions on y, is discussed in chapter 22. For now, however, we concentrate on demonstrating that solutions of the Sturm–Liouville equation that satisfy appropriate boundary conditions are the eigenfunctions of an Hermitian operator.

It is clear that (17.32) can be written

$$\mathcal{L}y = \lambda\rho(x)y, \qquad \text{where } \mathcal{L} \equiv -\left[p(x)\frac{d^2}{dx^2} + r(x)\frac{d}{dx} + q(x)\right]. \tag{17.33}$$

Using the condition that $r(x) = p'(x)$, it will be seen that the general Sturm–Liouville equation (17.32) can also be rewritten as

$$(py')' + qy + \lambda\rho y = 0, \tag{17.34}$$

where primes denote differentiation with respect to x. Using (17.33) this may also be written $\mathcal{L}y \equiv -(py')' - qy = \lambda\rho y$, which defines a more useful form for the Sturm–Liouville linear operator, namely

$$\mathcal{L} \equiv -\left[\frac{d}{dx}\left(p(x)\frac{d}{dx}\right) + q(x)\right]. \tag{17.35}$$

17.4.1 Hermitian nature of the Sturm–Liouville operator

As we now show, the linear operator of the Sturm–Liouville equation (17.35) is self-adjoint. Moreover, the operator is Hermitian over the range $[a, b]$ provided

§ We note that sign conventions vary in this expression for the general Sturm–Liouville equation; some authors use $-\lambda\rho(x)y$ on the LHS of (17.32).

certain boundary conditions are met, namely that any two eigenfunctions y_i and y_j of (17.33) must satisfy

$$\left[y_i^* p y_j'\right]_{x=a} = \left[y_i^* p y_j'\right]_{x=b} \qquad \text{for all } i, j. \tag{17.36}$$

Rearranging (17.36), we can write

$$\left[y_i^* p y_j'\right]_{x=a}^{x=b} = 0 \tag{17.37}$$

as an equivalent statement of the required boundary conditions. These boundary conditions are in fact not too restrictive and are met, for instance, by the sets $y(a) = y(b) = 0$; $y(a) = y'(b) = 0$; $p(a) = p(b) = 0$ and by many other sets. It is important to note that in order to satisfy (17.36) and (17.37) one boundary condition must be specified at each end of the range.

►*Prove that the Sturm–Liouville operator is Hermitian over the range* $[a, b]$ *and under the boundary conditions (17.37).*

Putting the Sturm–Liouville form $\mathcal{L}y = -(py')' - qy$ into the definition (17.16) of an Hermitian operator, the LHS may be written as a sum of two terms, i.e.

$$-\int_a^b \left[y_i^*(py_j')' + y_i^* q y_j\right] dx = -\int_a^b y_i^*(py_j')'\, dx - \int_a^b y_i^* q y_j\, dx.$$

The first term may be integrated by parts to give

$$-\left[y_i^* p y_j'\right]_a^b + \int_a^b (y_i^*)' p y_j'\, dx.$$

The boundary-value term in this is zero because of the boundary conditions, and so integrating by parts again yields

$$\left[(y_i^*)' p y_j\right]_a^b - \int_a^b ((y_i^*)' p)' y_j\, dx.$$

Again, the boundary-value term is zero, leaving us with

$$-\int_a^b \left[y_i^*(py_j')' + y_i^* q y_j\right] dx = -\int_a^b \left[y_j(p(y_i^*)')' + y_j q y_i^*\right] dx,$$

which proves that the Sturm–Liouville operator is Hermitian over the prescribed interval. ◄

It is also worth noting that, since $p(a) = p(b) = 0$ is a valid set of boundary conditions, many Sturm–Liouville equations possess a 'natural' interval $[a, b]$ over which the corresponding differential operator $\mathcal{L}$ is Hermitian *irrespective* of the boundary conditions satisfied by its eigenfunctions at $x = a$ and $x = b$ (the only requirement being that they are regular at these end-points).

17.4.2 Transforming an equation into Sturm–Liouville form

Many of the second-order differential equations encountered in physical problems are examples of the Sturm–Liouville equation (17.34). Moreover, *any* second-order

Equation	$p(x)$	$q(x)$	λ	$\rho(x)$
Hypergeometric	$x^c(1-x)^{a+b-c+1}$	0	$-ab$	$x^{c-1}(1-x)^{a+b-c}$
Legendre	$1-x^2$	0	$\ell(\ell+1)$	1
Associated Legendre	$1-x^2$	$-m^2/(1-x^2)$	$\ell(\ell+1)$	1
Chebyshev	$(1-x^2)^{1/2}$	0	ν^2	$(1-x^2)^{-1/2}$
Confluent hypergeometric	$x^c e^{-x}$	0	$-a$	$x^{c-1}e^{-x}$
Bessel*	x	$-\nu^2/x$	α^2	x
Laguerre	xe^{-x}	0	ν	e^{-x}
Associated Laguerre	$x^{m+1}e^{-x}$	0	ν	$x^m e^{-x}$
Hermite	e^{-x^2}	0	2ν	e^{-x^2}
Simple harmonic	1	0	ω^2	1

Table 17.1 The Sturm–Liouville form (17.34) for important ODEs in the physical sciences and engineering. The asterisk denotes that, for Bessel's equation, a change of variable $x \to x/a$ is required to give the conventional normalisation used here, but is not needed for the transformation into Sturm–Liouville form.

differential equation of the form

$$p(x)y'' + r(x)y' + q(x)y + \lambda\rho(x)y = 0 \tag{17.38}$$

can be converted into Sturm–Liouville form by multiplying through by a suitable integrating factor, which is given by

$$F(x) = \exp\left\{\int^x \frac{r(u) - p'(u)}{p(u)}\,du\right\}. \tag{17.39}$$

It is easily verified that (17.38) then takes the Sturm–Liouville form,

$$[F(x)p(x)y']' + F(x)q(x)y + \lambda F(x)\rho(x)y = 0, \tag{17.40}$$

with a different, but still non-negative, weight function $F(x)\rho(x)$. Table 17.1 summarises the Sturm–Liouville form (17.34) for several of the equations listed in table 16.1. These forms can be determined using (17.39), as illustrated in the following example.

►*Put the following equations into Sturm–Liouville (SL) form:*

(i) $(1-x^2)y'' - xy' + \nu^2 y = 0$ *(Chebyshev equation)*;

(ii) $xy'' + (1-x)y' + \nu y = 0$ *(Laguerre equation)*;

(iii) $y'' - 2xy' + 2\nu y = 0$ *(Hermite equation)*.

(i) From (17.39), the required integrating factor is

$$F(x) = \exp\left(\int^x \frac{u}{1-u^2}\,du\right) = \exp\left[-\tfrac{1}{2}\ln(1-x^2)\right] = (1-x^2)^{-1/2}.$$

Thus, the Chebyshev equation becomes

$$(1-x^2)^{1/2}y'' - x(1-x^2)^{-1/2}y' + \nu^2(1-x^2)^{-1/2}y = \left[(1-x^2)^{1/2}y'\right]' + \nu^2(1-x^2)^{-1/2}y = 0,$$

which is in SL form with $p(x) = (1-x^2)^{1/2}$, $q(x) = 0$, $\rho(x) = (1-x^2)^{-1/2}$ and $\lambda = \nu^2$.

(ii) From (17.39), the required integrating factor is

$$F(x) = \exp\left(\int^x -1\,du\right) = \exp(-x).$$

Thus, the Laguerre equation becomes

$$xe^{-x}y'' + (1-x)e^{-x}y' + \nu e^{-x}y = (xe^{-x}y')' + \nu e^{-x}y = 0,$$

which is in SL form with $p(x) = xe^{-x}$, $q(x) = 0$, $\rho(x) = e^{-x}$ and $\lambda = \nu$.

(iii) From (17.39), the required integrating factor is

$$F(x) = \exp\left(\int^x -2u\,du\right) = \exp(-x^2).$$

Thus, the Hermite equation becomes

$$e^{-x^2}y'' - 2xe^{-x^2}y' + 2\nu e^{-x^2}y = (e^{-x^2}y')' + 2\nu e^{-x^2}y = 0,$$

which is in SL form with $p(x) = e^{-x^2}$, $q(x) = 0$, $\rho(x) = e^{-x^2}$ and $\lambda = 2\nu$. ◄

From the $p(x)$ entries in table 17.1, we may read off the natural interval over which the corresponding Sturm–Liouville operator (17.35) is Hermitian; in each case this is given by $[a, b]$, where $p(a) = p(b) = 0$. Thus, the natural interval for the Legendre equation, the associated Legendre equation and the Chebyshev equation is $[-1, 1]$; for the Laguerre and associated Laguerre equations the interval is $[0, \infty]$; and for the Hermite equation it is $[-\infty, \infty]$. In addition, from (17.37), one sees that for the simple harmonic equation one requires only that $[a, b] = [x_0, x_0 + 2\pi]$. We also note that, as required, the weight function in each case is finite and non-negative over the natural interval. Occasionally, a little more care is required when determining the conditions for a Sturm–Liouville operator of the form (17.35) to be Hermitian over some natural interval, as is illustrated in the following example.

►*Express the hypergeometric equation,*

$$x(1-x)y'' + [\,c - (a+b+1)x\,]y' - aby = 0,$$

in Sturm–Liouville form. Hence determine the natural interval over which the resulting Sturm–Liouville operator is Hermitian and the corresponding conditions that one must impose on the parameters a, b and c.

As usual for an equation not already in SL form, we first determine the appropriate integrating factor. This is given, as in equation (17.39), by

$$\begin{aligned}
F(x) &= \exp\left[\int^x \frac{c-(a+b+1)u-1+2u}{u(1-u)}\,du\right] \\
&= \exp\left[\int^x \frac{c-1-(a+b-1)u}{u(1-u)}\,du\right] \\
&= \exp\left[\int^x \left(\frac{c-1}{1-u}+\frac{c-1}{u}-\frac{a+b-1}{1-u}\,du\right)\right] \\
&= \exp\left[\,(a+b-c)\ln(1-x)+(c-1)\ln x\,\right] \\
&= x^{c-1}(1-x)^{a+b-c}.
\end{aligned}$$

When the equation is multiplied through by $F(x)$ it takes the form

$$\left[\,x^c(1-x)^{a+b-c+1}y'\,\right]' - abx^{c-1}(1-x)^{a+b-c}y = 0.$$

Now, for the corresponding Sturm–Liouville operator to be Hermitian, the conditions to be imposed are as follows.

(i) The boundary condition (17.37); if $c > 0$ and $a+b-c+1 > 0$, this is satisfied automatically for $0 \le x \le 1$, which is thus the natural interval in this case.
(ii) The weight function $x^{c-1}(1-x)^{a+b-c}$ must be finite and not change sign in the interval $0 \le x \le 1$. This means that both exponents in it must be positive, i.e. $c-1 > 0$ and $a+b-c > 0$.

Putting together the conditions on the parameters gives the double inequality $a+b > c > 1$. ◄

Finally, we consider Bessel's equation,

$$x^2y'' + xy' + (x^2 - \nu^2)y = 0,$$

which may be converted into Sturm–Liouville form, but only in a somewhat unorthodox fashion. It is conventional first to divide the Bessel equation by x and then to change variables to $\bar{x} = x/\alpha$. In this case, it becomes

$$\bar{x}y''(\alpha\bar{x}) + y'(\alpha\bar{x}) - \frac{\nu^2}{\bar{x}}y(\alpha\bar{x}) + \alpha^2\bar{x}y(\alpha\bar{x}) = 0, \qquad (17.41)$$

where a prime now indicates differentiation with respect to $\bar{x}$. Dropping the bars on the independent variable, we thus have

$$[xy'(\alpha x)]' - \frac{\nu^2}{x}y(\alpha x) + \alpha^2 xy(\alpha x) = 0, \qquad (17.42)$$

which is in SL form with $p(x) = x$, $q(x) = -\nu^2/x$, $\rho(x) = x$ and $\lambda = \alpha^2$. It should be noted, however, that in this case the eigenvalue (actually its square root) appears in the argument of the dependent variable.

17.5 Superposition of eigenfunctions: Green's functions

We have already seen that if

$$\mathcal{L}y_n(x) = \lambda_n \rho(x) y_n(x), \tag{17.43}$$

where $\mathcal{L}$ is an Hermitian operator, then the eigenvalues λ_n are real and the eigenfunctions $y_n(x)$ are orthogonal (or can be made so). Let us assume that we know the eigenfunctions $y_n(x)$ of $\mathcal{L}$ that individually satisfy (17.43) and some imposed boundary conditions (for which $\mathcal{L}$ is Hermitian).

Now let us suppose we wish to solve the inhomogeneous differential equation

$$\mathcal{L}y(x) = f(x), \tag{17.44}$$

subject to the same boundary conditions. Since the eigenfunctions of $\mathcal{L}$ form a complete set, the full solution, $y(x)$, to (17.44) may be written as a superposition of eigenfunctions, i.e.

$$y(x) = \sum_{n=0}^{\infty} c_n y_n(x), \tag{17.45}$$

for some choice of the constants c_n. Making full use of the linearity of $\mathcal{L}$, we have

$$f(x) = \mathcal{L}y(x) = \mathcal{L}\left(\sum_{n=0}^{\infty} c_n y_n(x)\right) = \sum_{n=0}^{\infty} c_n \mathcal{L} y_n(x) = \sum_{n=0}^{\infty} c_n \lambda_n \rho(x) y_n(x). \tag{17.46}$$

Multiplying the first and last terms of (17.46) by y_j^* and integrating, we obtain

$$\int_a^b y_j^*(z) f(z)\, dz = \sum_{n=0}^{\infty} \int_a^b c_n \lambda_n y_j^*(z) y_n(z) \rho(z)\, dz, \tag{17.47}$$

where we have used z as the integration variable for later convenience. Finally, using the orthogonality condition (17.27), we see that the integrals on the RHS are zero unless $n = j$, and so obtain

$$c_n = \frac{1}{\lambda_n} \frac{\int_a^b y_n^*(z) f(z)\, dz}{\int_a^b y_n^*(z) y_n(z) \rho(z)\, dz}. \tag{17.48}$$

Thus, if we can find all the eigenfunctions of a differential operator then (17.48) can be used to find the weighting coefficients for the superposition, to give as the full solution

$$y(x) = \sum_{n=0}^{\infty} \frac{1}{\lambda_n} \frac{\int_a^b y_n^*(z) f(z)\, dz}{\int_a^b y_n^*(z) y_n(z) \rho(z)\, dz}\, y_n(x). \tag{17.49}$$

If we work with normalised eigenfunctions $\hat{y}_n(x)$, so that

$$\int_a^b \hat{y}_n^*(z) \hat{y}_n(z) \rho(z)\, dz = 1 \qquad \text{for all } n,$$

and we assume that we may interchange the order of summation and integration, then (17.49) can be written as

$$y(x) = \int_a^b \left\{ \sum_{n=0}^{\infty} \left[\frac{1}{\lambda_n} \hat{y}_n(x) \hat{y}_n^*(z) \right] \right\} f(z)\, dz.$$

The quantity in braces, which is a function of x and z only, is usually written $G(x,z)$, and is the *Green's function* for the problem. With this notation,

$$y(x) = \int_a^b G(x,z) f(z)\, dz, \tag{17.50}$$

where

$$G(x,z) = \sum_{n=0}^{\infty} \frac{1}{\lambda_n} \hat{y}_n(x) \hat{y}_n^*(z). \tag{17.51}$$

We note that $G(x,z)$ is determined entirely by the boundary conditions and the eigenfunctions $\hat{y}_n$, and hence by $\mathcal{L}$ itself, and that $f(z)$ depends purely on the RHS of the inhomogeneous equation (17.44). Thus, for a given $\mathcal{L}$ and boundary conditions we can establish, once and for all, a function $G(x,z)$ that will enable us to solve the inhomogeneous equation for *any* RHS. From (17.51) we also note that

$$G(x,z) = G^*(z,x). \tag{17.52}$$

We have already met the Green's function in the solution of second-order differential equations in chapter 15, as the function that satisfies the equation $\mathcal{L}[G(x,z)] = \delta(x-z)$ (and the boundary conditions). The formulation given above is an alternative, though equivalent, one.

►*Find an appropriate Green's function for the equation*

$$y'' + \tfrac{1}{4} y = f(x),$$

with boundary conditions $y(0) = y(\pi) = 0$. *Hence, solve for* (i) $f(x) = \sin 2x$ *and* (ii) $f(x) = x/2$.

One approach to solving this problem is to use the methods of chapter 15 and find a complementary function and particular integral. However, in order to illustrate the techniques developed in the present chapter we will use the superposition of eigenfunctions, which, as may easily be checked, produces the same solution.

The operator on the LHS of this equation is already Hermitian under the given boundary conditions, and so we seek its eigenfunctions. These satisfy the equation

$$y'' + \tfrac{1}{4} y = \lambda y.$$

This equation has the familiar solution

$$y(x) = A \sin\left(\sqrt{\tfrac{1}{4} - \lambda}\right) x + B \cos\left(\sqrt{\tfrac{1}{4} - \lambda}\right) x.$$

Now, the boundary conditions require that $B = 0$ and $\sin\left(\sqrt{\tfrac{1}{4} - \lambda}\right)\pi = 0$, and so

$$\sqrt{\tfrac{1}{4} - \lambda} = n, \qquad \text{where } n = 0, \pm 1, \pm 2, \ldots.$$

Therefore, the independent eigenfunctions that satisfy the boundary conditions are

$$y_n(x) = A_n \sin nx,$$

where n is any non-negative integer, and the corresponding eigenvalues are $\lambda_n = \frac{1}{4} - n^2$. The normalisation condition further requires

$$\int_0^\pi A_n^2 \sin^2 nx \, dx = 1 \qquad \Rightarrow \qquad A_n = \left(\frac{2}{\pi}\right)^{1/2}.$$

Comparison with (17.51) shows that the appropriate Green's function is therefore given by

$$G(x, z) = \frac{2}{\pi} \sum_{n=0}^{\infty} \frac{\sin nx \sin nz}{\frac{1}{4} - n^2}.$$

Case (i). Using (17.50), the solution with $f(x) = \sin 2x$ is given by

$$y(x) = \frac{2}{\pi} \int_0^\pi \left(\sum_{n=0}^{\infty} \frac{\sin nx \sin nz}{\frac{1}{4} - n^2} \right) \sin 2z \, dz = \frac{2}{\pi} \sum_{n=0}^{\infty} \frac{\sin nx}{\frac{1}{4} - n^2} \int_0^\pi \sin nz \sin 2z \, dz.$$

Now the integral is zero unless $n = 2$, in which case it is

$$\int_0^\pi \sin^2 2z \, dz = \frac{\pi}{2}.$$

Thus

$$y(x) = -\frac{2}{\pi} \frac{\sin 2x}{15/4} \frac{\pi}{2} = -\frac{4}{15} \sin 2x$$

is the full solution for $f(x) = \sin 2x$. This is, of course, exactly the solution found by using the methods of chapter 15.

Case (ii). The solution with $f(x) = x/2$ is given by

$$y(x) = \int_0^\pi \left(\frac{2}{\pi} \sum_{n=0}^{\infty} \frac{\sin nx \sin nz}{\frac{1}{4} - n^2} \right) \frac{z}{2} \, dz = \frac{1}{\pi} \sum_{n=0}^{\infty} \frac{\sin nx}{\frac{1}{4} - n^2} \int_0^\pi z \sin nz \, dz.$$

The integral may be evaluated by integrating by parts. For $n \neq 0$,

$$\begin{aligned}
\int_0^\pi z \sin nz \, dz &= \left[-\frac{z \cos nz}{n} \right]_0^\pi + \int_0^\pi \frac{\cos nz}{n} \, dz \\
&= \frac{-\pi \cos n\pi}{n} + \left[\frac{\sin nz}{n^2} \right]_0^\pi \\
&= -\frac{\pi(-1)^n}{n}.
\end{aligned}$$

For $n = 0$ the integral is zero, and thus

$$y(x) = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{\sin nx}{n\left(\frac{1}{4} - n^2\right)},$$

is the full solution for $f(x) = x/2$. Using the methods of subsection 15.1.2, the solution is found to be $y(x) = 2x - 2\pi \sin(x/2)$, which may be shown to be equal to the above solution by expanding $2x - 2\pi \sin(x/2)$ as a Fourier sine series. ◀

17.6 A useful generalisation

Sometimes we encounter inhomogeneous equations of a form slightly more general than (17.1), given by

$$\mathcal{L}y(x) - \mu\rho(x)y(x) = f(x) \tag{17.53}$$

for some Hermitian operator $\mathcal{L}$, with y subject to the appropriate boundary conditions and μ a given (i.e. *fixed*) constant. To solve this equation we expand $y(x)$ and $f(x)$ in terms of the eigenfunctions $y_n(x)$ of the operator $\mathcal{L}$, which satisfy

$$\mathcal{L}y_n(x) = \lambda_n\rho(x)y_n(x).$$

Working in terms of the normalised eigenfunctions $\hat{y}_n(x)$, we first expand $f(x)$ as follows:

$$\begin{aligned} f(x) &= \sum_{n=0}^{\infty} \hat{y}_n(x)\int_a^b \hat{y}_n^*(z)f(z)\rho(z)\,dz \\ &= \int_a^b \rho(z)\sum_{n=0}^{\infty} \hat{y}_n(x)\hat{y}_n^*(z)f(z)\,dz. \end{aligned} \tag{17.54}$$

Using (17.29) this becomes

$$\begin{aligned} f(x) &= \int_a^b \rho(x)\sum_{n=0}^{\infty} \hat{y}_n(x)\hat{y}_n^*(z)f(z)\,dz \\ &= \rho(x)\sum_{n=0}^{\infty} \hat{y}_n(x)\int_a^b \hat{y}_n^*(z)f(z)\,dz. \end{aligned} \tag{17.55}$$

Next, we expand $y(x)$ as $y = \sum_{n=0}^{\infty} c_n\hat{y}_n(x)$ and seek the coefficients c_n. Substituting this and (17.55) into (17.53) we have

$$\rho(x)\sum_{n=0}^{\infty}(\lambda_n - \mu)c_n\hat{y}_n(x) = \rho(x)\sum_{n=0}^{\infty} \hat{y}_n(x)\int_a^b \hat{y}_n^*(z)f(z)\,dz,$$

from which we find that

$$c_n = \sum_{n=0}^{\infty} \frac{\int_a^b \hat{y}_n^*(z)f(z)\,dz}{\lambda_n - \mu}.$$

Hence the solution of (17.53) is given by

$$y = \sum_{n=0}^{\infty} c_n\hat{y}_n(x) = \sum_{n=0}^{\infty} \frac{\hat{y}_n(x)}{\lambda_n - \mu}\int_a^b \hat{y}_n^*(z)f(z)\,dz = \int_a^b \sum_{n=0}^{\infty} \frac{\hat{y}_n(x)\hat{y}_n^*(z)}{\lambda_n - \mu}f(z)\,dz.$$

From this we may identify the Green's function

$$G(x,z) = \sum_{n=0}^{\infty} \frac{\hat{y}_n(x)\hat{y}_n^*(z)}{\lambda_n - \mu}.$$

We note that if $\mu = \lambda_n$, i.e. if μ equals one of the eigenvalues of $\mathcal{L}$, then $G(x, z)$ becomes infinite and this method runs into difficulty. No solution then exists unless the RHS of (17.53) satisfies the relation

$$\int_a^b \hat{y}_n^*(x) f(x)\, dx = 0.$$

If the spectrum of eigenvalues of the operator $\mathcal{L}$ is anywhere continuous, the orthonormality and closure relationships of the normalised eigenfunctions become

$$\int_a^b \hat{y}_n^*(x)\hat{y}_m(x)\rho(x)\, dx = \delta(n-m),$$
$$\int_0^\infty \hat{y}_n^*(z)\hat{y}_n(x)\rho(x)\, dn = \delta(x-z).$$

Repeating the above analysis we then find that the Green's function is given by

$$G(x, z) = \int_0^\infty \frac{\hat{y}_n(x)\hat{y}_n^*(z)}{\lambda_n - \mu}\, dn.$$

17.7 Exercises

17.1 By considering $\langle h|h\rangle$, where $h = f + \lambda g$ with λ real, prove that, for two functions f and g,

$$\langle f|f\rangle\langle g|g\rangle \geq \tfrac{1}{4}[\langle f|g\rangle + \langle g|f\rangle]^2.$$

The function $y(x)$ is real and positive for all x. Its Fourier cosine transform $\tilde{y}_c(k)$ is defined by

$$\tilde{y}_c(k) = \int_{-\infty}^{\infty} y(x)\cos(kx)\, dx,$$

and it is given that $\tilde{y}_c(0) = 1$. Prove that

$$\tilde{y}_c(2k) \geq 2[\tilde{y}_c(k)]^2 - 1.$$

17.2 Write the homogeneous Sturm-Liouville eigenvalue equation for which $y(a) = y(b) = 0$ as

$$\mathscr{L}(y;\lambda) \equiv (py')' + qy + \lambda\rho y = 0,$$

where $p(x), q(x)$ and $\rho(x)$ are continuously differentiable functions. Show that if $z(x)$ and $F(x)$ satisfy $\mathcal{L}(z;\lambda) = F(x)$, with $z(a) = z(b) = 0$, then

$$\int_a^b y(x)F(x)\, dx = 0.$$

Demonstrate the validity of this general result by direct calculation for the specific case in which $p(x) = \rho(x) = 1$, $q(x) = 0$, $a = -1$, $b = 1$ and $z(x) = 1 - x^2$.

17.3 Consider the real eigenfunctions $y_n(x)$ of a Sturm–Liouville equation,

$$(py')' + qy + \lambda\rho y = 0, \qquad a \leq x \leq b,$$

in which $p(x)$, $q(x)$ and $\rho(x)$ are continuously differentiable real functions and $p(x)$ does not change sign in $a \leq x \leq b$. Take $p(x)$ as positive throughout the interval, if necessary by changing the signs of all eigenvalues. For $a \le x_1 \le x_2 \le b$, establish the identity

$$(\lambda_n - \lambda_m) \int_{x_1}^{x_2} \rho y_n y_m \, dx = \left[y_n \, p \, y_m' - y_m \, p \, y_n' \right]_{x_1}^{x_2}.$$

Deduce that if $\lambda_n > \lambda_m$ then $y_n(x)$ must change sign between two successive zeros of $y_m(x)$.
[The reader may find it helpful to illustrate this result by sketching the first few eigenfunctions of the system $y'' + \lambda y = 0$, with $y(0) = y(\pi) = 0$, and the Legendre polynomials $P_n(z)$ for $n = 2, 3, 4, 5$.]

17.4 Show that the equation

$$y'' + a\delta(x)y + \lambda y = 0,$$

with $y(\pm\pi) = 0$ and a real, has a set of eigenvalues λ satisfying

$$\tan(\pi\sqrt{\lambda}) = \frac{2\sqrt{\lambda}}{a}.$$

Investigate the conditions under which negative eigenvalues, $\lambda = -\mu^2$, with μ real, are possible.

17.5 Use the properties of Legendre polynomials to carry out the following exercises.

(a) Find the solution of $(1 - x^2)y'' - 2xy' + by = f(x)$, valid in the range $-1 \le x \le 1$ and finite at $x = 0$, in terms of Legendre polynomials.
(b) If $b = 14$ and $f(x) = 5x^3$, find the explicit solution and verify it by direct substitution.

[The first six Legendre polynomials are listed in Subsection 18.1.1.]

17.6 Starting from the linearly independent functions $1, x, x^2, x^3, \ldots$, in the range $0 \le x < \infty$, find the first three orthogonal functions ϕ_0, ϕ_1 and ϕ_2, with respect to the weight function $\rho(x) = e^{-x}$. By comparing your answers with the Laguerre polynomials generated by the recurrence relation (18.115), deduce the form of $\phi_3(x)$.

17.7 Consider the set of functions, $\{f(x)\}$, of the real variable x, defined in the interval $-\infty < x < \infty$, that $\to 0$ at least as quickly as x^{-1} as $x \to \pm\infty$. For unit weight function, determine whether each of the following linear operators is Hermitian when acting upon $\{f(x)\}$:

$$\text{(a)}\ \frac{d}{dx} + x; \quad \text{(b)}\ -i\frac{d}{dx} + x^2; \quad \text{(c)}\ ix\frac{d}{dx}; \quad \text{(d)}\ i\frac{d^3}{dx^3}.$$

17.8 A particle moves in a parabolic potential in which its natural angular frequency of oscillation is $\frac{1}{2}$. At time $t = 0$ it passes through the origin with velocity v. It is then suddenly subjected to an additional acceleration, of $+1$ for $0 \le t \le \pi/2$, followed by -1 for $\pi/2 < t \le \pi$. At the end of this period it is again at the origin. Apply the results of the worked example in section 17.5 to show that

$$v = -\frac{8}{\pi} \sum_{m=0}^{\infty} \frac{1}{(4m+2)^2 - \frac{1}{4}} \approx -0.81.$$

17.9 Find an eigenfunction expansion for the solution, with boundary conditions $y(0) = y(\pi) = 0$, of the inhomogeneous equation

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

where κ is a constant and

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

17.10 Consider the following two approaches to constructing a Green's function.

(a) Find those eigenfunctions $y_n(x)$ of the self-adjoint linear differential operator d^2/dx^2 that satisfy the boundary conditions $y_n(0) = y_n(\pi) = 0$, and hence construct its Green's function $G(x, z)$.

(b) Construct the same Green's function using a method based on the complementary function of the appropriate differential equation and the boundary conditions to be satisfied at the position of the δ-function, showing that it is

$$G(x, z) = \begin{cases} x(z - \pi)/\pi & 0 \le x \le z, \\ z(x - \pi)/\pi & z \le x \le \pi. \end{cases}$$

(c) By expanding the function given in (b) in terms of the eigenfunctions $y_n(x)$, verify that it is the same function as that derived in (a).

17.11 The differential operator $\mathcal{L}$ is defined by

$$\mathcal{L}y = -\frac{d}{dx}\left(e^x \frac{dy}{dx}\right) - \tfrac{1}{4}e^x y.$$

Determine the eigenvalues λ_n of the problem

$$\mathcal{L}y_n = \lambda_n e^x y_n \qquad 0 < x < 1,$$

with boundary conditions

$$y(0) = 0, \qquad \frac{dy}{dx} + \tfrac{1}{2}y = 0 \quad \text{at} \quad x = 1.$$

(a) Find the corresponding unnormalised y_n, and also a weight function $\rho(x)$ with respect to which the y_n are orthogonal. Hence, select a suitable normalisation for the y_n.

(b) By making an eigenfunction expansion, solve the equation

$$\mathcal{L}y = -e^{x/2}, \qquad 0 < x < 1,$$

subject to the same boundary conditions as previously.

17.12 Show that the linear operator

$$\mathcal{L} \equiv \tfrac{1}{4}(1 + x^2)^2 \frac{d^2}{dx^2} + \tfrac{1}{2}x(1 + x^2)\frac{d}{dx} + a,$$

acting upon functions defined in $-1 \le x \le 1$ and vanishing at the end-points of the interval, is Hermitian with respect to the weight function $(1 + x^2)^{-1}$.

By making the change of variable $x = \tan(\theta/2)$, find two even eigenfunctions, $f_1(x)$ and $f_2(x)$, of the differential equation

$$\mathcal{L}u = \lambda u.$$

17.13 By substituting $x = \exp t$, find the normalised eigenfunctions $y_n(x)$ and the eigenvalues λ_n of the operator $\mathcal{L}$ defined by

$$\mathcal{L}y = x^2 y'' + 2xy' + \tfrac{1}{4}y, \qquad 1 \le x \le e,$$

with $y(1) = y(e) = 0$. Find, as a series $\sum a_n y_n(x)$, the solution of $\mathcal{L}y = x^{-1/2}$.

17.14 Express the solution of Poisson's equation in electrostatics,

$$\nabla^2\phi(\mathbf{r}) = -\rho(\mathbf{r})/\epsilon_0,$$

where ρ is the non-zero charge density over a finite part of space, in the form of an integral and hence identify the Green's function for the ∇^2 operator.

17.15 In the quantum-mechanical study of the scattering of a particle by a potential, a Born-approximation solution can be obtained in terms of a function $y(\mathbf{r})$ that satisfies an equation of the form

$$(-\nabla^2 - K^2)y(\mathbf{r}) = F(\mathbf{r}).$$

Assuming that $y_{\mathbf{k}}(\mathbf{r}) = (2\pi)^{-3/2}\exp(i\mathbf{k}\cdot\mathbf{r})$ is a suitably normalised eigenfunction of $-\nabla^2$ corresponding to eigenvalue k^2, find a suitable Green's function $G_K(\mathbf{r},\mathbf{r}')$. By taking the direction of the vector $\mathbf{r}-\mathbf{r}'$ as the polar axis for a $\mathbf{k}$-space integration, show that $G_K(\mathbf{r},\mathbf{r}')$ can be reduced to

$$\frac{1}{4\pi^2|\mathbf{r}-\mathbf{r}'|}\int_{-\infty}^{\infty}\frac{w\sin w}{w^2-w_0^2}\,dw,$$

where $w_0 = K|\mathbf{r}-\mathbf{r}'|$.

[This integral can be evaluated using a contour integration (chapter 24) to give $(4\pi|\mathbf{r}-\mathbf{r}'|)^{-1}\exp(iK|\mathbf{r}-\mathbf{r}'|)$.]

17.8 Hints and answers

17.1 Express the condition $\langle h|h\rangle \geq 0$ as a quadratic equation in λ and then apply the condition for no real roots, noting that $\langle f|g\rangle + \langle g|f\rangle$ is real. To put a limit on $\int y\cos^2 kx\,dx$, set $f = y^{1/2}\cos kx$ and $g = y^{1/2}$ in the inequality.

17.3 Follow an argument similar to that used for proving the reality of the eigenvalues, but integrate from x_1 to x_2, rather than from a to b. Take x_1 and x_2 as two successive zeros of $y_m(x)$ and note that, if the sign of y_m is α then the sign of $y_m'(x_1)$ is α whilst that of $y_m'(x_2)$ is $-\alpha$. Now assume that $y_n(x)$ does not change sign in the interval and has a constant sign β; show that this leads to a contradiction between the signs of the two sides of the identity.

17.5 (a) $y = \sum a_n P_n(x)$ with

$$a_n = \frac{n+1/2}{b-n(n+1)}\int_{-1}^{1} f(z)P_n(z)\,dz;$$

(b) $5x^3 = 2P_3(x)+3P_1(x)$, giving $a_1 = 1/4$ and $a_3 = 1$, leading to $y = 5(2x^3-x)/4$.

17.7 (a) No, $\int gf^{*\prime}dx \neq 0$; (b) yes; (c) no, $i\int f^*g dx \neq 0$; (d) yes.

17.9 The normalised eigenfunctions are $(2/\pi)^{1/2}\sin nx$, with n an integer. $y(x) = (4/\pi)\sum_{n\text{ odd}}[(-1)^{(n-1)/2}\sin nx]/[n^2(\kappa-n^2)]$.

17.11 $\lambda_n = (n+1/2)^2\pi^2$, $n = 0, 1, 2, \ldots$.
(a) Since $y_n(1)y_m'(1) \neq 0$, the Sturm–Liouville boundary conditions are not satisfied and the appropriate weight function has to be justified by inspection. The normalised eigenfunctions are $\sqrt{2}e^{-x/2}\sin[(n+1/2)\pi x]$, with $\rho(x) = e^x$.
(b) $y(x) = (-2/\pi^3)\sum_{n=0}^{\infty} e^{-x/2}\sin[(n+1/2)\pi x]/(n+1/2)^3$.

17.13 $y_n(x) = \sqrt{2}x^{-1/2}\sin(n\pi\ln x)$ with $\lambda_n = -n^2\pi^2$;

$$a_n = \begin{cases} -(n\pi)^{-2}\int_1^e \sqrt{2}x^{-1}\sin(n\pi\ln x)\,dx = -\sqrt{8}(n\pi)^{-3} & \text{for } n \text{ odd,} \\ 0 & \text{for } n \text{ even.}\end{cases}$$

17.15 Use the form of Green's function that is the integral over all eigenvalues of the 'outer product' of two eigenfunctions corresponding to the same eigenvalue, but with arguments $\mathbf{r}$ and $\mathbf{r}'$.