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 as a *superposition*, generally containing an infinite number of terms, of some set of functions 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, 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 $\mathcal{L}$ and λ_i is then called the eigenvalue associated with the eigenfunction $y_i(x)$.

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

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

In this case the eigenfunctions are given by $y_n(t) = A_n e^{i\omega_n t}$, where $\omega_n = 2\pi n/T$, T is the period of oscillation, $n = 0, \pm 1, \pm 2, \ldots$ 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 this confusing terminology here.)

Another equation of the form (17.2) is Legendre's equation

$$\mathcal{L}y \equiv -(1-x^2)\frac{d^2y}{dx^2} + 2x\frac{dy}{dx} = \ell(\ell+1)y, \tag{17.5}$$

where

$$\mathcal{L} = -(1-x^2)\frac{d^2}{dx^2} + 2x\frac{d}{dx}. \tag{17.6}$$

We found the eigenfunctions of $\mathcal{L}$ by a series method in chapter 16, and for solutions to Legendre's equation that are regular at $x = \pm 1$ these are the Legendre polynomials, given by

$$y_\ell(x) = P_\ell(x) = \frac{1}{2^\ell \ell!}\frac{d^\ell}{dx^\ell}(x^2-1)^\ell \tag{17.7}$$

for $\ell = 0, 1, 2, \ldots$; they have associated eigenvalues $\ell(\ell+1)$. (Again, ℓ is sometimes, confusingly, referred to as the eigenvalue of this equation.)

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

$$\mathcal{L}y(x) = \lambda\rho(x)y(x), \tag{17.8}$$

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 'reasonable' function in the interval $a \le x \le b$ (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 $z_n(x)$ is chosen then the function can be expressed in terms of the new basis,

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

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

$$\langle f|g\rangle = \int_a^b f^*(x)g(x)\rho(x)\,dx, \tag{17.9}$$

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* on the interval $[a, b]$ if

$$\langle f|g\rangle = \int_a^b f^*(x)g(x)\rho(x)\,dx = 0, \tag{17.10}$$

and the *norm* of a function is defined as

$$\|f\| = \langle f|f\rangle^{1/2} = \left[\int_a^b f^*(x)f(x)\rho(x)\,dx\right]^{1/2} = \left[\int_a^b |f(x)|^2\rho(x)\,dx\right]^{1/2}. \tag{17.11}$$

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 $\phi_n(x)$, $n = 0, 1, 2, \ldots$, that are orthonormal, i.e. such that

$$\langle \phi_i|\phi_j\rangle = \int_a^b \phi_i^*(x)\phi_j(x)\rho(x)\,dx = \delta_{ij}. \tag{17.12}$$

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 ϕ_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, in which each of the new functions ψ_n is to be normalised, giving 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 $a \leq x \leq b$, so that it can, without loss of generality, be taken to be non-negative throughout. A function $y(x)$ that satisfies (17.8) 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 (17.8), since we have discussed these 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 operators in the simple harmonic oscillator equation and in Legendre's equation are examples) have particularly useful properties and these will be studied in detail. I turns out that many of the interesting operators met with in the physical sciences are Hermitian. Before continuing our discussion of the eigenfunctions of Hermitian operators, however, we will consider the 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 $f(x)$, $g(x)$, $h(x)$, ... on the interval $a \leq x \leq b$. That these functions form a linear vector space can be verified since the set is closed under

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

Espacios Vectoriales de Funciones

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

(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*

$\phi_n = \psi_n \langle \psi_n | \psi_n \rangle^{-1/2}$, before proceeding to the construction of the next one:

$$
\begin{aligned}
\psi_0 &= y_0, \\
\psi_1 &= y_1 - \phi_0 \langle \phi_0 | y_1 \rangle, \\
\psi_2 &= y_2 - \phi_1 \langle \phi_1 | y_2 \rangle - \phi_0 \langle \phi_0 | y_2 \rangle, \\
&\vdots \\
\psi_n &= y_n - \phi_{n-1} \langle \phi_{n-1} | y_n \rangle - \cdots - \phi_0 \langle \phi_0 | y_n \rangle \\
&\vdots
\end{aligned}
$$

It is straightforward to check that each $\phi_n = \psi_n \langle \psi_n | \psi_n \rangle^{-1/2}$ 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 ψ_n also 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 the first three orthonormal functions over the range $-1 < x < 1$.*

The first unnormalised function ψ_0 is simply equal to the first of the original functions, i.e.

$$\psi_0 = 1.$$

The normalisation is carried out by dividing by

$$\langle \psi_0 | \psi_0 \rangle^{1/2} = \left(\int_{-1}^{1} 1 \times 1 \, du \right)^{1/2} = \sqrt{2},$$

with the result that the first normalised function ϕ_0 is given by

$$\phi_0 = \frac{\psi_0}{\sqrt{2}} = \sqrt{\tfrac{1}{2}}.$$

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

$$\psi_1 = y_1 - \phi_0 \langle \phi_0 | y_1 \rangle.$$

It can easily be shown that $\langle \phi_0 | y_1 \rangle = 0$, and so $\psi_1 = x$. Normalising then gives

$$\phi_1 = \psi_1 \left(\int_{-1}^{1} u \times u \, du \right)^{-1/2} = \sqrt{\tfrac{3}{2}} x.$$

The third unnormalised function is similarly given by

$$
\begin{aligned}
\psi_2 &= y_2 - \phi_1 \langle \phi_1 | y_2 \rangle - \phi_0 \langle \phi_0 | y_2 \rangle \\
&= x^2 - 0 - \tfrac{1}{3},
\end{aligned}
$$

which, on normalising, gives

$$\phi_2 = \psi_2 \left(\int_{-1}^{1} \left(u^2 - \tfrac{1}{3} \right)^2 du \right)^{-1/2} = \tfrac{1}{2} \sqrt{\tfrac{5}{2}} (3x^2 - 1).$$

By comparing the functions ϕ_0, ϕ_1 and ϕ_2, with the list in subsection 16.6.1, 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 $\phi_n(x)$ as

$$f(x) = \sum_{n=0}^{\infty} a_n \phi_n(x) \tag{17.13}$$

then the coefficients a_n are given by

$$a_n = \langle \phi_n | f \rangle = \int_a^b \phi_n^*(x) f(x) \rho(x)\, dx. \tag{17.14}$$

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

$$|\langle f|g\rangle| \leq \langle f|f\rangle^{1/2} \langle g|g\rangle^{1/2}, \tag{17.15}$$

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

$$\|f+g\| \leq \|f\| + \|g\|, \tag{17.16}$$

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 $\phi_n(x)$ so that any function $f(x)$ can be written as

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

where $c_n = \langle \phi_n | f \rangle$. Bessel's inequality then states that

$$\langle f|f\rangle \geq \sum_n |c_n|^2. \tag{17.17}$$

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 and Hermitian operators

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

$$\int_a^b f(x)^* \left[\mathcal{L}g(x)\right] \rho(x)\, dx = \left\{ \int_a^b g^*(x) \left[\mathcal{L}^\dagger f(x)\right] \rho(x)\, dx \right\}^*, \qquad (17.18)$$

or, in inner product notation, $\langle f|\mathcal{L}g\rangle = \langle g|\mathcal{L}^\dagger f\rangle^*$. An operator is then said to be *self-adjoint* or *Hermitian* if $\mathcal{L}^\dagger = \mathcal{L}$, i.e. if

$$\int_a^b f^*(x) \left[\mathcal{L}g(x)\right] \rho(x)\, dx = \left\{ \int_a^b g^*(x) \left[\mathcal{L}f(x)\right] \rho(x)\, dx \right\}^*, \qquad (17.19)$$

or, in inner product notation, $\langle f|\mathcal{L}g\rangle = \langle g|\mathcal{L}f\rangle^*$. From (17.19) we note that, when applied to an Hermitian operator, the general property $\langle b|a\rangle^* = \langle a|b\rangle$ takes the form

$$\langle g|\mathcal{L}f\rangle^* = \langle \mathcal{L}f|g\rangle \qquad \Rightarrow \qquad \boxed{\langle \mathcal{L}f|g\rangle = \langle f|\mathcal{L}g\rangle = \langle f|\mathcal{L}|g\rangle,}$$

where the notation of the final equality emphasises that $\mathcal{L}$ can act on either f or g without changing the value of the inner product. 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. In general, however, an operator $\mathcal{L}$ is Hermitian over an interval $a \le x \le b$ only if certain boundary conditions are met by the functions f and g on which it acts.

► *Find the required boundary conditions for the linear operator $\mathcal{L} = d^2/dt^2$ to be Hermitian over the interval t_0 to $t_0 + T$.*

Substituting into the LHS of the definition of an Hermitian operator (17.19) and integrating by parts gives

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

where we have taken the weight function $\rho(x)$ to be unity. Integrating the second term on the RHS by parts yields

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

Remembering that the operator is real and taking the complex conjugate outside the integral gives

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

which, by comparison with (17.19), proves that $\mathcal{L}$ is Hermitian provided

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

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, so that it is possible to expand in an eigenfunction series any function $y(x)$ obeying the appropriate conditions:

$$y(x) = \sum_{n=0}^{\infty} c_n y_n(x), \tag{17.20}$$

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.

17.3 The 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 (17.8) is satisfied by at least two eigenfunctions $y_i(x)$ and $y_j(x)$, which have eigenvalues λ_i and λ_j respectively, so that

$$\mathcal{L}y_i = \lambda_i \rho(x) y_i, \tag{17.21}$$

$$\mathcal{L}y_j = \lambda_j \rho(x) y_j, \tag{17.22}$$

where $\rho(x)$ is the weight function. Multiplying (17.21) by y_j^* and (17.22) by y_i^*

† The proof of the completeness of the eigenfunctions of an Hermitian operator is beyond the scope of this book. The reader should refer to e.g. Courant and Hilbert, *Methods of Mathematical Physics* (Interscience Publishers, 1953).

and then integrating gives

$$\int_a^b y_j^* \mathcal{L} y_i \, dx = \lambda_i \int_a^b y_j^* y_i \rho \, dx, \tag{17.23}$$

$$\int_a^b y_i^* \mathcal{L} y_j \, dx = \lambda_j \int_a^b y_i^* y_j \rho \, dx. \tag{17.24}$$

Remembering that we have required $\rho(x)$ to be real, the complex conjugate of (17.23) becomes

$$\left[\int_a^b y_j^* \mathcal{L} y_i \, dx\right]^* = \lambda_i^* \int_a^b y_i^* y_j \rho \, dx, \tag{17.25}$$

and using the definition of an Hermitian operator (17.19) it follows that the LHS of (17.25) is equal to the LHS of (17.24). Thus

$$(\lambda_i^* - \lambda_j) \int_a^b y_i^* y_j \rho \, dx = 0. \tag{17.26}$$

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 of the eigenfunctions

From (17.26), it is immediately apparent that two eigenfunctions y_i and y_j that correspond to different eigenvalues, i.e. such that $\lambda_i \neq \lambda_j$, satisfy

$$\int_a^b y_i^* y_j \rho \, dx = 0, \tag{17.27}$$

which is a statement of the orthogonality of y_i and y_j. Because $\mathcal{L}$ is linear, the normalisation of the eigenfunctions $y_i(x)$ is arbitrary and we shall assume for definiteness that they are normalised so that $\int_a^b y_i^* y_i \rho \, dx = 1$. Thus we can write (17.27) in the form

$$\int_a^b y_i^* y_j \rho \, dx = \delta_{ij}, \tag{17.28}$$

which is valid for all pairs of values i, 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.

$$\mathcal{L}y_i = \lambda_0 \rho y_i \qquad \text{for } i = 0, 1, \ldots, k-1, \tag{17.29}$$

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

$$\mathcal{L}z \equiv \mathcal{L}\sum_{i=0}^{k-1} c_i y_i = \sum_{i=0}^{k-1} c_i \mathcal{L}y_i = \sum_{i=0}^{k-1} c_i \lambda_0 \rho y_i = \lambda_0 \rho z. \tag{17.30}$$

If the y_i defined in (17.29) are not already mutually orthogonal then consider the new eigenfunctions z_i constructed by the following procedure, in which each of the new functions w_i is to be normalised, to give z_i, before proceeding to the construction of the next one (the normalisation can be carried out by dividing the eigenfunction w_i by $(\int_a^b w_i^* w_i \rho\, dx)^{1/2}$):

$$\begin{aligned}
w_0 &= y_0, \\
w_1 &= y_1 - \left(z_0 \int_a^b z_0^* y_1 \rho\, dx \right), \\
w_2 &= y_2 - \left(z_1 \int_a^b z_1^* y_2 \rho\, dx \right) - \left(z_0 \int_a^b z_0^* y_2 \rho\, dx \right), \\
&\vdots \\
w_{k-1} &= y_{k-1} - \left(z_{k-2} \int_a^b z_{k-2}^* y_{k-1} \rho\, dx \right) - \cdots - \left(z_0 \int_a^b z_0^* y_{k-1} \rho\, dx \right).
\end{aligned}$$

Each of the integrals is just a number and thus each new function $z_i = w_i(\int_a^b w_i^* w_i \rho\, dx)^{-1/2}$ is, as can be shown from (17.30), 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 orthonormal set obtained, z_i, is not unique.

17.3.3 Construction of real eigenfunctions

Recall that the eigenfunction y_i satisfies

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

and that the complex conjugate of this gives

$$\mathcal{L}y_i^* = \lambda_i^* \rho y_i^* = \lambda_i \rho y_i^*, \tag{17.32}$$

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. Therefore the eigenfunctions can always be made real by taking suitable linear combinations. 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

$$p(x)\frac{d^2y}{dx^2} + r(x)\frac{dy}{dx} + q(x)y + \lambda\rho(x)y = 0, \quad \boxed{\text{where } r(x) = \frac{dp(x)}{dx}} \tag{17.33}$$

and p, q and r are real functions of x. (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 (17.33).) A variational approach to the Sturm–Liouville equation, which is useful in estimating the eigenvalues λ of the equation, is discussed in chapter 22. For now, however, we concentrate on a demonstration that the Sturm–Liouville equation can be solved by superposition methods.

It is clear that (17.33) can be written

$$\mathcal{L}y = \lambda\rho(x)y \qquad \text{where } \boxed{\mathcal{L} \equiv -\left[p(x)\frac{d^2}{dx^2} + r(x)\frac{d}{dx} + q(x)\right].} \tag{17.34}$$

An example is Legendre's equation (17.5), which is a Sturm–Liouville equation with $p(x) = 1 - x^2$, $r(x) = -2x = p'(x)$, $q(x) = 0$, $\rho(x) = 1$ and eigenvalues $\ell(\ell+1)$.

It will be seen that the general Sturm–Liouville equation (17.33) can be rewritten

$$\boxed{(py')' + qy + \lambda\rho y = 0,} \tag{17.35}$$

where primes denote differentiation with respect to x. Using (17.34) this may also be written $\mathcal{L}y \equiv -(py')' - qy = \lambda\rho y$. We will show in the next section that, under certain boundary conditions on the solutions $y(x)$, linear operators that can be written in this form are *self-adjoint*.

Whilst it is true that Sturm–Liouville equations represent only a small fraction of the differential equations encountered in practice, as we shall demonstrate in subsection 17.4.2 *any* second-order differential equation of the form

$$p(x)y'' + r(x)y' + q(x)y + \lambda\rho(x)y = 0 \tag{17.36}$$

can be converted into Sturm–Liouville form by multiplying through by a suitable factor; this is discussed in subsection 17.4.2.

17.4.1 Valid boundary conditions

For the linear operator of the Sturm–Liouville equation (17.34) to be Hermitian over the range $[a,b]$ requires certain boundary conditions to be met, namely, that any two eigenfunctions y_i and y_j of (17.34) must satisfy

$$\left[y_i^* p y_j'\right]_{x=a} = \left[y_i^* p y_j'\right]_{x=b} \qquad \text{for all } i, j. \tag{17.37}$$

Rearranging (17.37) we find that

$$\boxed{\left[y_i^* p y_j'\right]_{x=a}^{x=b} = 0,} \tag{17.38}$$

is 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 (17.37) and (17.38) 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 (17.38).*

Putting the Sturm–Liouville form $\mathcal{L}y = -(py')' - qy$ into the definition (17.19) 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 first term is zero because of the boundary conditions, and thus, integrating by parts again yields

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

The first term is once again zero. Thus

$$\begin{aligned}
-\int_a^b \left[y_i^*(py_j')' + y_i^* q y_j\right] dx &= \int_a^b \left[-((y_i^*)'p)' y_j - y_i^* q y_j\right] dx, \\
&= \left\{-\int_a^b \left[y_j^*(py_i')' + y_j^* q y_i\right] dx\right\}^*,
\end{aligned}$$

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

17.4.2 Putting an equation into Sturm–Liouville form

The Sturm–Liouville equation (17.33) requires that $r(x) = p'(x)$. However, any equation of the form

$$p(x)y'' + r(x)y' + q(x)y + \lambda\rho(x)y = 0, \tag{17.39}$$

can be put into self-adjoint form by multiplying through by the integrating factor

$$F(x) = \exp\left\{\int^x \frac{r(z) - p'(z)}{p(z)}\,dz\right\}. \qquad (17.40)$$

It is easily verified that (17.39) then takes the Sturm–Liouville form

$$[F(x)p(x)y']' + F(x)q(x)y + \lambda F(x)\rho(x)y = 0, \qquad (17.41)$$

with a different, but still non-negative, weight function $F(x)\rho(x)$.

►*Put the Hermite equation*

$$y'' - 2xy' + 2\alpha y = 0$$

into Sturm–Liouville form.

Using (17.40), with $p(z) = 1$, $p'(z) = 0$ and $r(z) = -2z$ gives the integrating factor

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

Thus, the Hermite equation becomes

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

which is clearly in Sturm–Liouville form with $p(x) = e^{-x^2}$, $q(x) = 0$, $\rho(x) = e^{-x^2}$ and $\lambda = 2\alpha$. ◄

17.5 Examples of Sturm–Liouville equations

In order to illustrate the wide applicability of Sturm–Liouville theory, in this section we present a short catalogue of some common equations of Sturm–Liouville form. Many of them have already been discussed in chapter 16. In particular the reader should note the orthogonality properties of the various solutions, which, in each case, follow because the differential operator is self-adjoint. For completeness we also quote the associated generating functions.

17.5.1 Legendre's equation

We have already met *Legendre's equation*,

$$(1 - x^2)y'' - 2xy' + \ell(\ell + 1)y = [(1 - x^2)y']' + \ell(\ell + 1)y = 0 \qquad (17.42)$$

and shown that it is a Sturm–Liouville equation with $p(x) = 1 - x^2$, $q(x) = 0$, $\rho(x) = 1$ and eigenvalues $\ell(\ell+1)$. In the previous chapter we found the solutions of Legendre's equation that are regular for all finite x. These are the Legendre polynomials $P_\ell(x)$, which are given by a Rodrigues' formula:

$$P_\ell(x) = \frac{1}{2^\ell \ell!}\frac{d^\ell}{dx^\ell}(x^2 - 1)^\ell.$$

The orthogonality and normalisation of the functions in the interval $-1 \leq x \leq 1$ is expressed by

$$\int_{-1}^{1} P_\ell(x) P_k(x)\, dx = \frac{2}{2\ell + 1} \delta_{\ell k}.$$

The generating function is

$$G(x, h) = (1 - 2xh + h^2)^{-1/2} = \sum_{n=0}^{\infty} P_n(x) h^n.$$

Legendre's equations appear in the analysis of physical situations involving the operator ∇^2 and axial symmetry, since the linear differential operator involved has the form of the polar-angle part of ∇^2, when the latter is expressed in spherical polar coordinates. Examples include the solution of Laplace's equation in axially symmetric situations and the solution of the Schrödinger equation for a quantum mechanical system involving a central potential.

17.5.2 The associated Legendre equation

Very closely related to the Legendre equation is the *associated Legendre equation*

$$[(1 - x^2) y']' + \left[\ell(\ell + 1) - \frac{m^2}{1 - x^2}\right] y = 0, \qquad (17.43)$$

which reduces to Legendre's equation when $m = 0$. In physical applications $-\ell \leq m \leq \ell$ and m is restricted to integer values. If $y(x)$ is a solution of Legendre's equation then

$$w(x) = (1 - x^2)^{|m|/2} \frac{d^{|m|} y}{dx^{|m|}}$$

is a solution of the associated equation. The solutions of the associated Legendre equation that are regular for all finite x are called the *associated Legendre functions* and are therefore given by

$$P_\ell^m(x) = (1 - x^2)^{|m|/2} \frac{d^{|m|} P_\ell}{dx^{|m|}}.$$

Note also that $P_\ell^m(x) = 0$ for $m > \ell$. Like the Legendre polynomials, the associated Legendre functions $P_\ell^m(x)$ are orthogonal in the range $-1 \leq x \leq 1$. This property, and their normalisation, is expressed by

$$\int_{-1}^{1} P_\ell^m(x) P_k^m(x)\, dx = \frac{2}{2\ell + 1} \frac{(\ell + m)!}{(\ell - m)!} \delta_{\ell k}.$$

They have the generating function

$$G(x, h) = \frac{(2m)!(1 - x^2)^{m/2}}{2^m m!(1 - 2hx + h^2)^{m+1/2}} = \sum_{n=0}^{\infty} P_{n+m}^m(x) h^n.$$

The associated Legendre equation arises in physical situations in which there is a dependence on azimuthal angle ϕ of the form $e^{im\phi}$ or $\cos m\phi$.

17.5.3 Bessel's equation

Physical situations that when described in spherical polar coordinates give rise to Legendre and associated Legendre equations lead to Bessel's equation when cylindrical polar coordinates are used. Bessel's equation has the form

$$x^2y'' + xy' + (x^2 - n^2)y = 0, \tag{17.44}$$

but on dividing by x and changing variables to $\xi = x/a$,† it takes on the Sturm-Liouville form

$$(\xi y')' + a^2\xi y + \frac{-n^2}{\xi}y = 0, \tag{17.45}$$

where a prime now indicates differentiation with respect to ξ.

We met Bessel's equation in chapter 16, where we saw that those of its solutions that are regular for finite x are the Bessel functions, given by

$$J_n(x) = \sum_{r=0}^{\infty} \frac{(-1)^r(\frac{1}{2}x)^{n+2r}}{r!\Gamma(n+r+1)}, \tag{17.46}$$

where Γ is the gamma function discussed in the Appendix. Their orthogonality and normalisation over the range $0 \leq x < \infty$ have been discussed in detail in chapter 16. The generating function for the Bessel functions is

$$G(x,h) = \exp\left[\frac{x}{2}\left(h - \frac{1}{h}\right)\right] = \sum_{n=-\infty}^{\infty} J_n(x)h^n. \tag{17.47}$$

17.5.4 The simple harmonic equation

The most trivial of Sturm–Liouville equations is the simple harmonic motion equation

$$y'' + \omega^2 y = 0, \tag{17.48}$$

which has $p(x) = 1$, $q(x) = 0$, $\rho(x) = 1$ and eigenvalue ω^2. We have already met the solutions of this equation in the Fourier analysis of chapter 12, and the properties of orthogonality and normalisation of the eigenfunctions given there can now be seen in the wider context of general Sturm–Liouville equations.

† This change of scale is required to give the conventional normalisation, but is not needed for the transformation into Sturm–Liouville form.

17.5.5 Hermite's equation

The Hermite equation appears in the description of the wavefunction of a harmonic oscillator and is given by

$$y'' - 2xy' + 2\alpha y = 0. \tag{17.49}$$

We have already seen that it can be converted to Sturm–Liouville form by multiplying by the integrating factor $\exp(-x^2)$, which yields

$$e^{-x^2}y'' - 2xe^{-x^2}y' + 2\alpha e^{-x^2}y = (e^{-x^2}y')' + 2\alpha e^{-x^2}y = 0. \tag{17.50}$$

The solutions, the Hermite polynomials $H_n(x)$, are given by a Rodrigues' formula:

$$H_n(x) = (-1)^n e^{x^2}\frac{d^n}{dx^n}\left(e^{-x^2}\right). \tag{17.51}$$

Their orthogonality over the range $-\infty < x < \infty$ and their normalisation are summarised by

$$\int_{-\infty}^{\infty} e^{-x^2}H_m(x)H_n(x)\,dx = 2^n n!\sqrt{\pi}\delta_{mn}, \tag{17.52}$$

and their generating function is

$$G(x,h) = e^{2hx-h^2} = \sum_{n=0}^{\infty}\frac{H_n(x)}{n!}h^n. \tag{17.53}$$

17.5.6 Laguerre's equation

The Laguerre equation appears in the description of the wavefunction of the hydrogen atom and is given by

$$xy'' + (1-x)y' + ny = 0. \tag{17.54}$$

It can be converted to Sturm–Liouville form by multiplying by the integrating factor $\exp(-x)$, which yields

$$xe^{-x}y'' + (1-x)e^{-x}y' + ne^{-x}y = (xe^{-x}y')' + ne^{-x}y = 0. \tag{17.55}$$

The solutions, the Laguerre polynomials $L_n(x)$, are again given by a Rodrigues' formula:

$$L_n(x) = e^x\frac{d^n}{dx^n}\left(x^n e^{-x}\right). \tag{17.56}$$

Their orthogonality over the range $0 \le x < \infty$ and their normalisation are expressed by

$$\int_0^{\infty} e^{-x}L_m(x)L_n(x)\,dx = (n!)^2\delta_{mn}, \tag{17.57}$$

and their generating function is

$$G(x,h) = \frac{e^{-xh/(1-h)}}{1-h} = \sum_{n=0}^{\infty} \frac{L_n(x)}{n!} h^n. \tag{17.58}$$

17.5.7 Chebyshev's equation

The Chebyshev equation

$$(1-x^2)y'' - xy' + n^2 y = 0 \tag{17.59}$$

can be converted to an equation of Sturm–Liouville form by multiplying by the integrating factor $(1-x^2)^{-1/2}$. Simplifying, this yields

$$\left[(1-x^2)^{1/2} y'\right]' + n^2(1-x^2)^{-1/2} y = 0. \tag{17.60}$$

The solutions, the Chebyshev polynomials $T_n(x)$, are once again given by a Rodrigues' formula:

$$T_n(x) = \frac{(-2)^n n!(1-x^2)^{1/2}}{(2n)!} \frac{d^n}{dx^n}(1-x^2)^{n-1/2}. \tag{17.61}$$

Their orthogonality over the range $-1 \le x \le 1$ and their normalisation are given by

$$\int_{-1}^{1} (1-x^2)^{-1/2} T_m(x) T_n(x)\, dx = \begin{cases} 0 & \text{for } m \neq n, \\ \pi/2 & \text{for } n = m \neq 0, \\ \pi & \text{for } n = m = 0, \end{cases} \tag{17.62}$$

and their generating function is

$$G(x,h) = \frac{1-xh}{1-2xh+h^2} = \sum_{n=0}^{\infty} T_n(x) h^n. \tag{17.63}$$

17.6 Superposition of eigenfunctions: Green's functions

We have already seen that if

$$\mathcal{L} y_n(x) = \lambda_n \rho(x) y_n(x), \tag{17.64}$$

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 (17.64) and some imposed boundary conditions (for which $\mathcal{L}$ is Hermitian).

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

$$\mathcal{L} y(x) = f(x), \tag{17.65}$$

subject to the same boundary conditions. Since the eigenfunctions of $\mathcal{L}$ form a complete set, the full solution, $y(x)$, to (17.65) may be written as a superposition of eigenfunctions, i.e.

$$y(x) = \sum_{n=0}^{\infty} c_n y_n(x), \tag{17.66}$$

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

$$f(x) = \mathcal{L}y(x) = \mathcal{L}\left(\sum_{n=0}^{\infty} c_n y_n(x)\right) = \sum_{n=0}^{\infty} c_n \mathcal{L} y_n(x) = \sum_{n=0}^{\infty} c_n \lambda_n \rho(x) y_n(x). \tag{17.67}$$

Multiplying the first and last terms of (17.67) by y_j^* and integrating, we obtain

$$\int_a^b y_j^*(z) f(z)\, dz = \sum_{n=0}^{\infty} \int_a^b c_n \lambda_n y_j^*(z) y_n(z) \rho(z)\, dz, \tag{17.68}$$

where we have used z as the integration variable for later convenience. Finally, using the orthogonality condition (17.28), we see that the integrals on the RHS are zero unless $n = j$, and so obtain

$$c_n = \frac{1}{\lambda_n} \frac{\int_a^b y_n^*(z) f(z)\, dz}{\int_a^b y_n^*(z) y_n(z) \rho(z)\, dz}. \tag{17.69}$$

Thus, if we can find all the eigenfunctions of a differential operator then (17.69) can be used to find the weighting coefficients for the superposition, to give as the full solution

$$y(x) = \sum_{n=0}^{\infty} \frac{1}{\lambda_n} \frac{\int_a^b y_n^*(z) f(z)\, dz}{\int_a^b y_n^*(z) y_n(z) \rho(z)\, dz} y_n(x). \tag{17.70}$$

If the eigenfunctions have already been normalised, so that

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

and we assume that we may interchange the order of summation and integration, then (17.70) can be written as

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

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,

$$y(x) = \int_a^b G(x, z) f(z)\, dz, \tag{17.71}$$

where

$$G(x,z) = \sum_{n=0}^{\infty} \frac{1}{\lambda_n} y_n(x) y_n^*(z). \tag{17.72}$$

We note that $G(x,z)$ is determined entirely by the boundary conditions and the eigenfunctions y_n, and hence by $\mathcal{L}$ itself, and that $f(z)$ depends purely on the RHS of the inhomogeneous equation (17.65). 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 (17.72) we also note that

$$G(x,z) = G^*(z,x). \tag{17.73}$$

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 self-adjoint 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. 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 (17.72) 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 (17.71), 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, i.e.

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

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. ◀

A useful relation between the eigenfunctions of $\mathcal{L}$ is given by writing

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

and hence

$$\rho(z) \sum_n y_n(x) y_n^*(z) = \delta(x - z). \qquad (17.74)$$

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 (17.74) 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.

$$\rho(z)\sum_n y_n(x)y_n^*(z) = \rho(x)\sum_n y_n(x)y_n^*(z). \tag{17.75}$$

17.7 A useful generalisation

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

$$\mathcal{L}y(x) - \lambda\rho(x)y(x) = f(x) \tag{17.76}$$

for some self-adjoint operator $\mathcal{L}$, with y subject to the appropriate boundary conditions and λ 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

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

Firstly, we expand $f(x)$ as follows:

$$\begin{aligned} f(x) &= \sum_{n=0}^{\infty} y_n(x)\int_a^b y_n^*(z)f(z)\rho(z)\,dz \\ &= \int_a^b \rho(z)\sum_{n=0}^{\infty} y_n(x)y_n^*(z)f(z)\,dz. \end{aligned} \tag{17.77}$$

Using (17.75) this becomes

$$\begin{aligned} f(x) &= \int_a^b \rho(x)\sum_{n=0}^{\infty} y_n(x)y_n^*(z)f(z)\,dz \\ &= \rho(x)\sum_{n=0}^{\infty} y_n(x)\int_a^b y_n^*(z)f(z)\,dz. \end{aligned} \tag{17.78}$$

Next, we expand $y(x)$ as $y = \sum_{n=0}^{\infty} c_n y_n(x)$ and seek the coefficients c_n. Substituting this and (17.78) in (17.76) we have

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

from which we find that

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

Hence the solution of (17.76) is given by

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

From this we may identify the Green's function

$$G(x, z) = \sum_{n=0}^{\infty} \frac{y_n(x) y_n^*(z)}{\lambda_n - \lambda}.$$

We note that if $\lambda = \lambda_n$, i.e. if λ 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 (17.76) satisfies the relation

$$\int_a^b y_n^*(x) f(x)\, dx = 0.$$

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

$$\int_a^b y_n^*(x) y_m(x) \rho(x)\, dx = \delta(n - m),$$

$$\int_0^{\infty} y_n^*(z) y_n(x) \rho(x)\, dn = \delta(x - z).$$

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

$$G(x, z) = \int_0^{\infty} \frac{y_n(x) y_n^*(z)}{\lambda_n - \lambda}\, dn.$$

17.8 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 (a) 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

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

(b) Demonstrate the validity of result (a) by direct calculation for the 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

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

in which $p(x)$, $q(x)$ and $\rho(x)$ are continuously differentiable real functions and $p(x)$ does not change sign in $a \le x \le b$. 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 zeroes 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 $P_n(z)$ given in subsection 16.6.1 for $n = 2, 3, 4, 5$.)

17.4 (a) 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

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

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

17.5 Express the hypergeometric equation

$$(x^2 - x)y'' + [(1 + \alpha + \beta)x - \gamma]y' + \alpha\beta y = 0$$

in Sturm–Liouville form, determining the conditions imposed on x and on the parameters α, β and γ by the boundary conditions and the allowed forms of weight function.

17.6 (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.

17.7 Use the generating function for the Legendre polynomials $P_n(x)$ to show that

$$\int_0^1 P_{2n+1}(x)\,dx = (-1)^n \frac{(2n)!}{2^{2n+1}n!(n+1)!}$$

and that, except for the case $n = 0$,

$$\int_0^1 P_{2n}(x)\,dx = 0.$$

17.8 The quantum mechanical wavefunction for a one-dimensional simple harmonic oscillator in its nth energy level is of the form

$$\psi(x) = \exp(-x^2/2)H_n(x),$$

where $H_n(x)$ is the nth Hermite polynomial. The generating function for the polynomials (17.53) is

$$G(x,h) = e^{2hx-h^2} = \sum_{n=0}^{\infty} \frac{H_n(x)}{n!}h^n.$$

(a) Find $H_i(x)$ for $i = 1, 2, 3, 4$.
(b) Evaluate by direct calculation

$$\int_{-\infty}^{\infty} e^{-x^2} H_p(x)H_q(x)\,dx,$$

(i) for $p = 2$, $q = 3$; (ii) for $p = 2$, $q = 4$; (iii) for $p = q = 3$. Check your answers against equation (17.52). (You will find it convenient to use

$$\int_{-\infty}^{\infty} x^{2n}e^{-x^2}\,dx = \frac{(2n)!\sqrt{\pi}}{2^{2n}n!}$$

for integer $n \geq 0$.)

17.9 The Laguerre polynomials, which are required for the quantum mechanical description of the hydrogen atom, can be defined by the generating function (equation (17.58))

$$G(x,h) = \frac{e^{-hx/(1-h)}}{1-h} = \sum_{n=0}^{\infty} \frac{L_n(x)}{n!}h^n.$$

By differentiating the equation separately with respect to x and h, and resubstituting for $G(x,h)$, prove that L_n and L'_n ($= dL_n(x)/dx$) satisfy the recurrence relations

$$L'_n - nL'_{n-1} + nL_{n-1} = 0,$$
$$L_{n+1} - (2n+1-x)L_n + n^2L_{n-1} = 0.$$

From these two equations and others derived from them, show that $L_n(x)$ satisfies the Laguerre equation

$$xL''_n + (1-x)L'_n + nL_n = 0.$$

17.10 Starting from the linearly independent functions 1, x, x^2, x^3, $\ldots$, in the range $0 \leq x < \infty$, 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 derived in exercise 17.9, deduce the form of $\phi_3(x)$.

17.11 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 $\{f(x)\}$:

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

17.12 The Chebyshev polynomials $T_n(x)$ can be written as

$$T_n(x) = \cos(n\cos^{-1} x).$$

(a) Verify that these functions do satisfy the Chebyshev equation.
(b) Use de Moivre's theorem to show that an alternative expression is

$$T_n(x) = \sum_{r \text{ even}}^{n} (-1)^{r/2} \frac{n!}{(n-r)!r!} x^{n-r}(1-x^2)^{r/2}.$$

17.13 A particle moves in a parabolic potential in which its natural angular frequency of oscillation is $1/2$. At time $t = 0$ it passes through the origin with velocity v and is suddenly subjected to an additional acceleration of $+1$ for $0 \leq t \leq \pi/2$, and then -1 for $\pi/2 < t \leq \pi$. At the end of this period it is at the origin again. Apply the results of the worked example in section 17.6 to show that

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

17.14 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

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

17.15 (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 the methods of subsection 15.2.5, showing that it is

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

(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).

17.16 (a) The differential operator $\mathcal{L}$ is defined by

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

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

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

(b) 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.
(c) 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.

17.17 Show that the linear operator

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

acting upon functions defined in $-1 \le x \le 1$ and vanishing at the endpoints 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

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

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

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

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

17.19 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.

17.20 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

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

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

(This integral can be evaluated using a contour integration (chapter 20) to give $(4\pi|\mathbf{r}-\mathbf{r}'|)^{-1}\exp(iK|\mathbf{r}-\mathbf{r}'|)$.)

17.9 Hints and answers

17.1 Express the condition $\langle h|h\rangle \ge 0$ 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.

17.2 (a) By twice integrating by parts the term containing p, show that $\int_a^b y\mathscr{L}(z;\lambda)\,dx = \int_a^b z\mathscr{L}(y;\lambda)dx$.
(b) $y(x) = A\cos(\sqrt{\lambda}x)$ with $\lambda = n^2\pi^2/4$, and $F(x) = \lambda - 2 - \lambda x^2$.

17.3 Follow an argument similar to that in subsection 17.3.1, but integrate from x_1 to x_2, rather than from a to b. Take x_1 and x_2 as two successive zeroes 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.

17.4 (a) Different combinations of sinusoids are needed for negative and positive ranges of x. (b) μ must satisfy $\tanh\mu\pi = 2\mu/a$, which requires $a > 2/\pi$.

17.5 $[x^\gamma(1-x)^{\alpha+\beta-\gamma+1}y']' = \alpha\beta x^{\gamma-1}(1-x)^{\alpha+\beta-\gamma}y$; $0 \le x \le 1$, $\alpha+\beta > \gamma > 1$.

17.6 (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 $y = 5(2x^3-x)/4$.

17.8 (a) $2x$, $4x^2-2$, $8x^3-12x$, $16x^4-48x^2+12$; (b) (i) 0, (ii) 0, (iii) $48\sqrt{\pi}$.

17.10 $\phi_0(x) = 1$, $\phi_1(x) = x-1$, $\phi_2(x) = (x^2-4x+2)/2$; $n!\phi_n(x) = (-1)^n L_n(x)$; $\phi_3(x) = (x^3-9x^2+18x-6)/6$.

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

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

17.15 (a) The normalised eigenfunctions are $(2/\pi)^{1/2}\sin nx$, with n an integer. $G(x,z) = (-2/\pi)\sum_{n=0}^{\infty}[\sin(nz)\sin(nx)]/n^2$.

17.16 (a) $\lambda_n = (n+1/2)^2\pi^2$, $n = 0, 1, 2, \ldots$.
(b) 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 $\rho(x) = e^x$.
(c) $y(x) = (-2/\pi^3)\sum_{n=0}^{\infty} e^{-x/2}\sin[(n+1/2)\pi x]/(n+1/2)^3$.

17.17 In terms of $\theta, \mathscr{L}$ is $d^2/d\theta^2 + a$ and has eigenfunctions $u(\theta) = \cos(\sqrt{a-\lambda}\theta)$, where $\sqrt{a-\lambda} = 2n+1$;
$f_1(x) = (1-x^2)/(1+x^2)$; $f_2(x) = 4[(1-x^2)/(1+x^2)]^3 - 3[(1-x^2)/(1+x^2)]$.

17.18 $y_n(x) = \sqrt{2}x^{-1/2}\sin(n\pi\ln x)$ with $\lambda_n = -n^2\pi^2$;

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

17.19 $G(\mathbf{r},\mathbf{r}') = (4\pi|\mathbf{r}-\mathbf{r}'|)^{-1}$.