3

Orthogonal Functions and Sturm–Liouville Problems

In Fourier series we have seen that a function can be expressed in terms of an infinite series of sines and cosines. This is possible mainly because these trigonometrical functions form a complete orthogonal set.

The concept of an orthogonal set of functions is a natural generalization of the concept of an orthogonal set of vectors. In fact, a function can be considered as a generalized vector in an infinite dimensional vector space and sines and cosines as basis vectors of this space. This make us ask where does such basis come from. Are there other bases as well? In this chapter we discover that such bases arise as the eigenfunctions of self-adjoint (Hermitian) linear differential operators, just as Hermitian $n \times n$ matrices provide us with sets of eigenvectors that are orthogonal bases for n-dimensional space.

Many important physical problems are described by differential equations which can be put into a form known as Sturm–Liouville equation. We will show that under certain boundary conditions of the solution of the equation, the Sturm–Liouville operators are self-adjoint. Therefore many basis sets of orthogonal functions can be generated by Sturm–Liouville equations. Viewed from a broader Sturm–Liouville theory, Fourier series is only a special case.

Some Sturm–Liouville equations are of great importance, we give names to them. Solutions of these equations are known as special functions. In this chapter we will discuss the origin and properties of some special functions that are frequently encountered in mathematical physics. A more detailed discussion of the most important ones will be given in Chap. 4.

3.1 Functions as Vectors in Infinite Dimensional Vector Space

3.1.1 Vector Space

When we construct our number system, first we find that additions and multiplications of positive integers satisfy certain rules concerning the order in which the computation can proceed. Then we use these rules to define a wider class of numbers.

Here we are going to do the same thing with vectors. Based on the properties of ordinary three-dimensional vectors, we abstract a set of rules that these vectors satisfy. Then we use this set of rules as the definition of a vector space. Any set of objects that satisfies these rules is said to form a linear vector space.

As a consequence of the definition of ordinary vectors, it can be easily shown that they satisfy the following set of rules:

– Vector addition is commutative and associative

$$\mathbf{a}+\mathbf{b}=\mathbf{b}+\mathbf{a},$$
$$(\mathbf{a}+\mathbf{b})+\mathbf{c}=\mathbf{a}+(\mathbf{b}+\mathbf{c}).$$

– Multiplication by a scalar is distributive and associative

$$\alpha(\mathbf{a}+\mathbf{b})=\alpha\mathbf{a}+\alpha\mathbf{b},$$
$$(\alpha+\beta)\mathbf{a}=\alpha\mathbf{a}+\beta\mathbf{a},$$
$$\alpha(\beta\mathbf{a})=(\alpha\beta)\mathbf{a},$$

  where α and β are arbitrary scalars.
– There exists a null vector $\mathbf{0}$, such that

$$\mathbf{a}+\mathbf{0}=\mathbf{a}.$$

– All vectors $\mathbf{a}$ have a corresponding negative vector $-\mathbf{a}$, such that

$$\mathbf{a}+(-\mathbf{a})=\mathbf{0}.$$

– Multiplication by unit scalar leaves any vector unchanged,

$$1\mathbf{a}=\mathbf{a}.$$

– Multiplication by zero gives a null vector,

$$0\mathbf{a}=\mathbf{0}.$$

Now let us consider all well behaved functions $f(x),\ g(x),\ h(x),\ldots$ defined in the interval $a\leq x\leq b$. Clearly, they form a linear vector space, since it can be readily verified that

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

$$\alpha\,[f(x)+g(x)] = \alpha f(x) + \alpha g(x),$$
$$(\alpha+\beta)f(x) = \alpha f(x) + \beta f(x),$$
$$\alpha(\beta f(x)) = (\alpha\beta)f(x).$$

$$f(x)+0 = f(x).$$
$$f(x)+(-f(x)) = 0.$$
$$1 \times f(x) = f(x).$$
$$0 \times f(x) = 0.$$

Therefore a collection of all functions of x defined in a certain interval of x constitutes a vector space.

Dimension of a Vector Space. A three-dimensional ordinary vector $\mathbf{v}$ is described by its three components $(v_1,\ v_2,\ v_3)$. It can be regarded a function with three distinct values $[v(1),\ v(2),\ v(3)]$. A n-dimensional vector is defined by n-tuples $[v(1),\ v(2), \ldots, v(n)]$, as we have seen in the matrix theory. Now the function $f(x)$ is a vector, what is its dimension?

Let us imagine approximating the function $f(x)$ between $a \le x \le b$ in a piecewise constant manner. Divide the x interval $(a \le x \le b)$ into n equal parts. Approximate the function by a sequence of values $(f_1,\ f_2, \ldots, f_n)$, where f_i is the value of $f(x)$ at the left endpoint of the ith subinterval, except f_n which is the value of $f(b)$. For example, if we approximate $f(x) = 1 + x$ in $0 \le x \le 1$ by dividing the interval into two equal parts, then $f(x)$ is approximated by $[f(0),\ f(0.5),\ f(1)]$, or $(1,\ 1.5,\ 2.0)$. Of course this is a very poor approximation. A better approximation would be to divide the interval in ten equal parts and approximate the function with 11 tuples of numbers $(1,\ 1.1,\ 1.2, \ldots, 2)$. Since the function is actually defined by all possible values of x between 0 and 1, which consists of infinite number of values of x from 0 to 1, the function is described by n-tuples of numbers with $n \to \infty$. In this sense, we say that the function is a vector in an infinite dimensional vector space.

3.1.2 Inner Product and Orthogonality

So far we have not mentioned dot product of vectors. Dot product is also called inner product or scalar product. Often it is written as $\mathbf{u} \cdot \mathbf{v}$, or as $\langle \mathbf{u} |\ \mathbf{v} \rangle$, or $\langle \mathbf{u}, \mathbf{v} \rangle$.

$$\mathbf{u} \cdot \mathbf{v} = \langle \mathbf{u} |\ \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle .$$

A vector space does not need to have a dot product. But a function space without an inner product defined is too large a vector space to be useful in physical applications.

If we choose to introduce an inner product for the function space, how is it to be defined? Again we elevate the properties of dot product of familiar vectors to axioms and require the inner product of any vector space to satisfy these axioms.

From the definition of the dot product of two three-dimensional vectors $\mathbf{u}$ and $\mathbf{v}$:

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + u_3 v_3 = \sum_{j=1}^{3} u_j v_j,$$

it can be easily deduced that dot product is

- commutative
$$\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u},$$
- and linear
$$(\alpha \mathbf{u} + \beta \mathbf{v}) \cdot \mathbf{w} = \alpha (\mathbf{u} \cdot \mathbf{w}) + \beta (\mathbf{v} \cdot \mathbf{w}).$$

The norm (or length) of vector is defined as

$$\|\mathbf{u}\| = (\mathbf{u} \cdot \mathbf{u})^{1/2} = \left(\sum_{j=1}^{3} u_j u_j \right)^{1/2}.$$

- Therefore the norm is non-negative
$$\mathbf{u} \cdot \mathbf{u} > 0 \quad \text{for} \quad \text{all} \quad \mathbf{u} \neq \mathbf{0}.$$

In complex space, the components of a vector can assume complex values. As we have seen in matrix theory, the inner product in complex space is defined as

$$\mathbf{u} \cdot \mathbf{v} = u_1^* v_1 + u_2^* v_2 + u_3^* v_3 = \sum_{j=1}^{3} u_j^* v_j,$$

where u^* is the complex conjugate of u. Therefore in complex space,

- The commutative rule $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$ is replaced by
$$\mathbf{u} \cdot \mathbf{v} = (\mathbf{v} \cdot \mathbf{u})^*, \tag{3.1}$$

This follows from the fact that

$$\mathbf{u} \cdot \mathbf{v} = \sum_{j=1}^{3} u_j^* v_j = \sum_{j=1}^{3} \left(u_j v_j^* \right)^* = \left(\sum_{j=1}^{3} v_j^* u_j \right)^* = (\mathbf{v} \cdot \mathbf{u})^*.$$

Thus, if α is a complex number, then

$$(\alpha \mathbf{u} \cdot \mathbf{v}) = \alpha^* (\mathbf{u} \cdot \mathbf{v}), \tag{3.2}$$
$$(\mathbf{u} \cdot \alpha \mathbf{v}) = \alpha (\mathbf{u} \cdot \mathbf{v}). \tag{3.3}$$

Now if we use these properties as axioms to define a wider class of inner products, then we can see that for two n-dimensional vectors $\mathbf{u}$ and $\mathbf{v}$ in complex space, the expression

$$\mathbf{u} \cdot \mathbf{v} = u_1^* v_1 w_1 + u_2^* v_2 w_2 + \cdots + \; u_n^* v_n w_n = \sum_{j=1}^{n} u_j^* v_j w_j \qquad (3.4)$$

is also a legitimate inner product as long as w_j is a fixed real positive constant for each j.

Let us use two-dimensional real space for illustration. Suppose that $\mathbf{u} = (1, 2)$ and $\mathbf{v} =(3, -4)$, with $w_1 = 2, \; w_2 = 3$, then

$$\mathbf{u} \cdot \mathbf{v} = (1)(3)(2) + (2)(-4)(3) = -18$$

$$\mathbf{v} \cdot \mathbf{u} = (3)(1)(2) + (-4)(2)(3) = -18,$$

in agreement with the axiom $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$.

On the other hand, if $w_1 = 2, \; w_2 = -3$, then

$$\mathbf{u} \cdot \mathbf{u} =(1)(1)(2) + (2)(2)(-3) = -10,$$

in violation of the axiom $\mathbf{u} \cdot \mathbf{u} > 0$ for $\mathbf{u} \neq \mathbf{0}$.

It can be readily verified that with real positive w_j, (3.4) satisfies all the axioms of inner product. The w_js are known as "weights" because they attach more or less weight to the different components of the vector. Of course, w_j can all be equal to one. In many applications, this is indeed the case.

To define an inner product in a function space in the interval $a \leq x \leq b$, let us divide the interval into $n-1$ equal parts and imagine that the functions $f(x)$ and $g(x)$ are approximated in a piecewise constant manner as discussed before:

$$f(x) = (f_1, \; f_2, \ldots, f_n),$$

$$g(x) = (g_1, \; g_2, \ldots, g_n).$$

We can adopt the inner product as

$$\langle f \; | g \rangle = \sum_{j=1}^{n} f_j^* g_j \Delta x_j,$$

where Δx_j is the width of the subinterval. Regarding Δx_j as the weights, this definition is in accordance with (3.4). If we let $n \to \infty$, this sum becomes an integral

$$\langle f \; | g \rangle = \int_a^b f^*(x) g(x) \mathrm{d}x.$$

The weight could also be $w(x)\mathrm{d}x$, as long as $w(x)$ is a real positive function. In that case, the inner product is defined to be

$$\langle f \; | g \rangle = \int_a^b f^*(x) g(x) w(x) \mathrm{d}x.$$

This is the general definition of an inner product of an infinite dimensional vector space of functions. It can be readily shown that this definition satisfies all the axioms of an inner product. As mentioned before, in many problems the weight function $w(x)$ is equal to one for all x. It is to be emphasized that our heuristic approach is neither a derivation nor a proof, it only provides the motivation for this definition.

Two functions are said to be orthogonal in the interval between a and b if

$$\langle f \mid g \rangle = \int_a^b f^*(x) g(x) w(x) \mathrm{d}x = 0.$$

The norm of a function is defined as

$$\| f \| = \langle f \mid f \rangle^{1/2} = \left[\int_a^b f^*(x) f(x) w(x) \mathrm{d}x \right]^{1/2} = \left[\int_a^b |f(x)|^2 w(x) \mathrm{d}x \right]^{1/2}.$$

The function is said to be normalized if

$$\| f \| = 1.$$

An infinite dimensional vector space of functions, for which an inner product is defined is called a Hilbert space. In quantum mechanics, all legitimate wavefunctions live in Hilbert space.

3.1.3 Orthogonal Functions

Orthonormal Set. A collection of functions $\{\psi_n(x)\}$, where $n = 1, 2, \ldots$ is called an orthogonal set if $\langle \psi_n \mid \psi_m \rangle = 0$ whenever $n \neq m$.

Dividing each function by its norm

$$\phi_n(x) = \frac{1}{\| \psi_n \|} \psi_n(x),$$

we have an orthonormal set $\{\phi_n(x)\}$, which satisfies the relation

$$\langle \phi_n \mid \phi_m \rangle = \begin{cases} 0 & n \neq m \\ 1 & n = m \end{cases}.$$

It is to be noted that the functions in the set and their inner products are to be defined in the same interval of x.

For example, with a unit weight function $w(x) = 1$, the set of functions $\left\{ \sin \frac{n\pi x}{L} \right\}$ $(n = 1, 2, \ldots)$ is orthogonal on the interval $0 \leq x \leq L$, since

$$\int_0^L \sin \frac{n\pi x}{L} \sin \frac{m\pi x}{L} \mathrm{d}x = \begin{cases} 0 & n \neq m \\ \frac{L}{2} & n = m \end{cases}.$$

Furthermore, $\{\phi_n(x)\}$ where

$$\phi_n(x) = \sqrt{\frac{2}{L}} \sin \frac{n\pi x}{L},$$

is an orthonormal set in the interval of $t(0, L)$.

Gram–Schmidt Orthogonalization. Out of a linearly independent (but not orthogonal) set of functions $\{u_n(x)\}$, an orthonormal set $\{\phi_n\}$ over an arbitrary interval and with respect to an arbitrary weight function can be constructed by the Gram–Schmidt orthogonalization method. The procedure is similar to that we have used in the construction of a set of orthogonal eigenvectors of a Hermitian matrix.

From a given linearly independent set $\{u_n\}$, an orthogonal set $\{\psi_n\}$ can be constructed. We start with $n = 0$. Let

$$\psi_0(x) = u_0(x)$$

and normalized it to unity and denote the result as ϕ_0

$$\phi_0(x) = \frac{1}{\left[\int |\psi_0(x)|^2 w(x)\mathrm{d}x\right]^{1/2}} \psi_0(x).$$

Clearly,

$$\int |\phi_0(x)|^2 w \ \mathrm{d}x = \frac{1}{\left[\int |\psi_0(x)|^2 w(x)\mathrm{d}x\right]} \int |\psi_0(x)|^2 w(x) \ \mathrm{d}x = 1.$$

For $n = 1$, let

$$\psi_1(x) = u_1(x) + a_{10}\phi_0 \ (x).$$

we require $\psi_1(x)$ to be orthogonal to $\phi_0 \ (x)$,

$$\begin{aligned} &\int \phi_0^*(x)\psi_1(x)w(x) \ \mathrm{d}x \\ &= \int \phi_0^*(x)u_1(x)w(x) \ \mathrm{d}x + a_{10} \int |\phi_0(x)|^2 w(x) \ \mathrm{d}x = 0. \end{aligned}$$

Since ϕ_0 is normalized to unity, we have

$$a_{10} = -\int \phi_0^*(x)u_1(x)w(x)\mathrm{d}x.$$

With a_{10} so determined, $\psi_1(x)$ is a known function, which can be normalized. Let

$$\phi_1(x) = \frac{1}{\left[\int |\psi_1(x)|^2 w(x)\mathrm{d}x\right]^{1/2}} \psi_1(x).$$

For $n = 2$, let

$$\psi_2(x) = u_2(x) + a_{21}\phi_1(x) + a_{20}\phi_0\ (x).$$

The requirement that $\psi_2(x)$ be orthogonal to $\phi_1(x)$ and to $\phi_0(x)$ leads to

$$a_{21} = -\int \phi_1^*(x)u_2(x)w(x)\,\mathrm{d}x,$$

$$a_{20} = -\int \phi_0^*(x)u_2(x)w(x)\,\mathrm{d}x.$$

Thus $\psi_2(x)$ is determined. Clearly this process can be continued. We take ψ_i as the ith function of $\{\psi_n\}$ and set it to equal u_i plus an unknown linear combination of the previously determined ϕ_j, $j = 0, 1, \ldots i-1$. The requirement that ψ_i be orthogonal to each of the previous ϕ_j yields just enough constraints to determine each of the unknown coefficients. Then the fully determined ψ_i can be normalized to unity and the steps are repeated for $\psi_{i+1.}$ In terms of the inner products, the procedure can be expressed as:

$$\begin{aligned}
\psi_0 &= u_0 & \phi_0 &= \psi_0 \langle \psi_0 | \psi_0 \rangle^{-1/2} \\
\psi_1 &= u_1 - \phi_0 \langle \phi_0 | u_1 \rangle & \phi_1 &= \psi_1 \langle \psi_1 | \psi_1 \rangle^{-1/2} \\
\psi_2 &= u_2 - \phi_1 \langle \phi_1 | u_2 \rangle - \phi_0 \langle \phi_0 | u_2 \rangle & \phi_2 &= \psi_2 \langle \psi_2 | \psi_2 \rangle^{-1/2} \\
&\cdot \\
&\cdot \\
\psi_i &= u_i - \phi_{i-1} \langle \phi_{i-1} | u_i \rangle - \cdots & \phi_i &= \psi_i \langle \psi_i | \psi_i \rangle^{-1/2}.
\end{aligned}$$

Clearly $\{\psi_n\}$ is an orthogonal set and $\{\phi_n\}$ is an orthonormal set.

Example 3.1.1. Legendre Polynomials. Construct an orthonormal set from the linear independent functions $u_n(x) = x^n$, $n = 0, 1, 2, \ldots$ in the interval of $-1 \le x \le 1$ with a weight function $w(x) = 1$.

Solution 3.1.1. According to the Gram–Schmidt process, the first unnormalized function of the orthogonal set $\{\psi_n\}$ is simply u_0,

$$\psi_0 = u_0 = 1.$$

The first normalized function of the orthonormal set $\{\phi_n\}$ is

$$\phi_0 = \psi_0 \langle \psi_0 | \psi_0 \rangle^{-1/2} = \psi_0 \left[\int_{-1}^{1} \mathrm{d}x\right]^{-1/2} = \frac{1}{\sqrt{2}}.$$

The next function in the orthogonal set is

$$\psi_1 = u_1 - \phi_0 \langle \phi_0 | u_1 \rangle.$$

Since

$$\langle \phi_0 \, | u_1 \rangle = \int_{-1}^{1} \frac{1}{\sqrt{2}} x \, \mathrm{d}x = 0,$$

so

$$\psi_1 = x$$

and

$$\phi_1 = \psi_1 \langle \psi_1 \, | \psi_1 \rangle^{-1/2} = x \left[\int_{-1}^{1} x^2 \, \mathrm{d}x \right]^{-1/2} = \sqrt{\frac{3}{2}} x.$$

Continue the process

$$\psi_2 = u_2 - \phi_1 \langle \phi_1 \, | u_2 \rangle - \phi_0 \langle \phi_0 \, | u_2 \rangle .$$

Since

$$\langle \phi_1 \, | u_2 \rangle = \int_{-1}^{1} \sqrt{\frac{3}{2}} x^3 \, \mathrm{d}x = 0, \qquad \langle \phi_0 \, | u_2 \rangle = \int_{-1}^{1} \sqrt{\frac{1}{2}} x^2 \, \mathrm{d}x = \frac{\sqrt{2}}{3},$$

so

$$\psi_2 = x^2 - 0 - \frac{1}{\sqrt{2}} \frac{\sqrt{2}}{3} = x^2 - \frac{1}{3},$$

and

$$\phi_2 = \psi_2 \langle \psi_2 \, | \psi_2 \rangle^{-1/2} = \left(x^2 - \frac{1}{3} \right) \left[\int_{-1}^{1} \left(x^2 - \frac{1}{3} \right)^2 \mathrm{d}x \right]^{-1/2}$$

$$= \left(x^2 - \frac{1}{3} \right) \sqrt{\frac{45}{8}} = \sqrt{\frac{5}{2}} \left(\frac{3}{2} x^2 - \frac{1}{2} \right).$$

The next normalized function is

$$\phi_3 = \sqrt{\frac{7}{2}} \left(\frac{5}{2} x^3 - \frac{3}{2} x \right).$$

It is straight-forward, although tedious, to show that

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

where $P_n(x)$ is a polynomial of order n, and

$$P_n(1) = 1,$$

$$\int_{-1}^{1} P_n(x) P_m(x) \, \mathrm{d}x = \frac{2}{2n+1} \delta_{nm}.$$

These polynomials are known as Legendre polynomials. They are one of the most useful and most frequently encountered special functions in mathematical physics. Fortunately, as we shall see later, there are much easier methods to derive them.

In this example, we have used the Gram–Schmidt procedure to rearrange the set of linear independent functions $\{x^n\}$ into an orthonormal set for the given interval $-1 \leq x \leq 1$ and given weight function $w(x) = 1$. With other choices of intervals and weight functions, we will get other sets of orthogonal polynomials. For example, with the same set of functions $\{x^n\}$ and the same weight function $w(x) = 1$, if the interval is chosen to be $[0, 1]$, instead of $[-1, 1]$, the Gram–Schmidt process will lead to another set of orthogonal polynomials known as shifted Legendre polynomial $\{P_n^{\mathrm{s}}(x)\}$. With $P_n^{\mathrm{s}}(x)$ normalized in such a way that $P_n^{\mathrm{s}}(1) = 1$,

$$P_n^{\mathrm{s}}(x) = P_n\left(2\left(x - \frac{1}{2}\right)\right).$$

The first few shifted Legendre polynomials are

$$P_0^{\mathrm{s}}(x) = 1, \quad P_1^{\mathrm{s}}(x) = 2x - 1, \quad P_2^{\mathrm{s}}(x) = 6x^2 - 6x + 1.$$

As another example, with the weight function chosen as $w(x) = \mathrm{e}^{-x}$ in the interval of $0 \leq x < \infty$, the orthonormal set constructed from $\{x^n\}$ is known as the Laguerre polynomial $\{L_n(x)\}$. The first three Laguerre polynomials are

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

It can be readily verified that

$$\int_0^\infty L_n(x) L_m(x) \mathrm{e}^{-x} \mathrm{d}x = \delta_{nm}.$$

Sometimes Laguerre polynomials are defined with a normalization

$$\int_0^\infty L_n(x) L_m(x) \mathrm{e}^{-x} \mathrm{d}x = \delta_{nm} (n!)^2.$$

In that case, the first three Laguerre polynomials are

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

Obviously infinitely many orthogonal sets of functions can be generated from $\{x^n\}$ by the Gram–Schmidt process. With a given weight function and a specified interval, the Gram–Schmidt process is unique up to a multiplication constant, positive or negative. This process is rather cumbersome. Fortunately, almost all interesting orthogonal polynomials constructed by this method are solutions of particular differential equations. Therefore they can be discussed from the perspective of differential equations.

3.2 Generalized Fourier Series

By analogy with finite dimensional vector space, we can consider an orthonormal set of functions $\{\phi_n(x)\}$ $(n = 0, 1, 2, \ldots)$ on the interval $a \leq x \leq b$ as basis vectors in an infinite dimensional vector space of functions, in which

$$\langle \phi_m | \phi_n \rangle = \int_a^b \phi_m^*(x)\phi_n(x)w(x)\,\mathrm{d}x = \delta_{nm}.$$

If any arbitrary piecewise continuous bounded function $f(x)$ in the same interval can be represented as the linear sum of these functions

$$f(x) = c_0\phi_0(x) + c_1\phi_1(x) + \cdots = \sum_{n=0}^{\infty} c_n\phi_n(x), \tag{3.5}$$

then $\{\phi_n(x)\}$ is said to be complete. If this equation is valid, taking the inner product with $\phi_m(x)$, we have

$$\langle \phi_m | f \rangle = \sum_{n=0}^{\infty} c_n \langle \phi_m | \phi_n \rangle = \sum_{n=0}^{\infty} c_n\delta_{nm} = c_m.$$

The coefficients c_n

$$c_n = \langle \phi_n | f \rangle = \int_a^b \phi_m^*(x)f(x)w(x)\,\mathrm{d}x \tag{3.6}$$

are called Fourier coefficients and the series (3.5) with these coefficients

$$f(x) = \sum_{n=0}^{\infty} \langle \phi_n | f \rangle \phi_n(x)$$

is called the generalized Fourier series. Clearly if a different set of basis $\{\varphi_n\}$ is chosen, then the function can be expressed in terms of the new basis with a different set of coefficients.

The nature of the representation of $f(x)$ by a generalized Fourier series is that the series representation converges to the mean. Let us use real functions to illustrate. Select M equally spaced points in the interval $a \leq x \leq b$ at $x_1 = a,\ x_2 = a + \Delta x,\ x_3 = a + 2\,\Delta x, \ldots$ where $\Delta x = (b - a)/(M - 1)$. Then approximate the function at any one of these M points by the finite series

$$f(x_i) = \sum_{n=0}^{N} A_n\phi_n(x_i).$$

In order to make this approximation as good as possible in the least square sense, we have to minimize the mean square error. This means we have to differentiate the mean square error D,

$$D = \sum_{i=1}^{M} \left[f(x_i) - \sum_{n=0}^{N} A_n \phi_n(x_i) \right]^2 w(x_i)\ \Delta x$$

with respect to each of coefficient A_n and set it zero. Let A_k be one of the A_ns. The differentiation with respect to A_k

$$\frac{\partial D}{\partial A_k} = 0$$

leads to

$$\sum_{i=1}^{M} 2 \left[f(x_i) - \sum_{n=0}^{N} A_n \phi_n(x_i) \right] [-\phi_k(x_i)]\, w(x_i)\ \Delta x = 0,$$

or

$$\sum_{i=1}^{M} \phi_k(x_i) f(x_i) w(x_i) \Delta x - \sum_{n=0}^{N} A_n \sum_{i=1}^{M} \phi_k(x_i) \phi_n(x_i) w(x_i)\ \Delta x = 0.$$

Now if we take the limit as $M \to \infty$ and $\Delta x \to 0$, we see this approaching the limit

$$\int_a^b \phi_k(x) f(x) w(x)\ \mathrm{d}x - \sum_{n=0}^{N} A_n \int_a^b \phi_k(x) \phi_n(x) w(x)\ \mathrm{d}x = 0.$$

With real functions, the orthogonality condition is

$$\int_a^b \phi_k(x) \phi_n(x) w(x)\ \mathrm{d}x = \delta_{nk}.$$

Therefore

$$A_k = \int_a^b \phi_k(x) f(x) w(x)\ \mathrm{d}x$$

which is exactly the same as the Fourier coefficient. In this approximation, the mean square error is minimized. For the generalized Fourier series, in which $\{\phi_n\}$ is a complete set and $N \to \infty$, the integral of the error squared goes to zero.

Of crucial importance is that the basis set must be complete. The set $\{\phi_n\}$ is complete in the function space if there is no nonzero function that is orthogonal to each of the function ϕ_n. For example, $\left\{ \frac{1}{\sqrt{\pi}} \sin nx \right\}$ $(n = 1, 2, \ldots)$ is an orthonormal set on the interval $-\pi \le x \le \pi$. But it is not complete since any even function in that interval is orthogonal to any of ϕ_n in the set.

It is not always that easy to use the definition to test if a set is complete. Fortunately, complete sets of orthogonal functions are provided by the eigenfunctions of certain type of differential operators known as Hermitian (or self-adjoint) operators.

3.3 Hermitian Operators

3.3.1 Adjoint and Self-adjoint (Hermitian) Operators

If the functions $f(x)$ and $g(x)$ in the vector space of functions, satisfy certain boundary conditions, the adjoint of a linear differential operator L, denoted by L^{+}, is defined by the relation

$$\langle Lf \mid g\rangle = \left\langle f \mid L^{+}g\right\rangle .$$

For example, in an infinite dimensional vector space consisting of all square-integrable functions with the inner product defined as

$$\langle f \mid f\rangle = \int_{-\infty}^{\infty} |f|^{2}\,\mathrm{d}x < \infty,$$

all functions must satisfy the boundary conditions

$$f(x) \to 0, \text{ as } x \to \pm\infty.$$

If the differential operator L in this space, in which $w(x) = 1$, is $\mathrm{d}/\mathrm{d}x$; ($L = \mathrm{d}/\mathrm{d}x$), then the inner product $\langle Lf \mid g\rangle$ is given by

$$\langle Lf \mid g\rangle = \left\langle \frac{\mathrm{d}}{\mathrm{d}x}f \middle| g \right\rangle = \int_{-\infty}^{\infty} \left(\frac{\mathrm{d}}{\mathrm{d}x}f\right)^{*} g\,\mathrm{d}x = \int_{-\infty}^{\infty} \frac{\mathrm{d}}{\mathrm{d}x}f^{*}g\,\mathrm{d}x.$$

With integration by parts,

$$\int_{-\infty}^{\infty} \frac{\mathrm{d}}{\mathrm{d}x}f^{*}g\,\mathrm{d}x = f^{*}(x)g(x)|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} f^{*}\frac{\mathrm{d}}{\mathrm{d}x}g\,\mathrm{d}x = \left\langle f \middle| -\frac{\mathrm{d}}{\mathrm{d}x}g \right\rangle = \left\langle f \mid L^{+}g\right\rangle,$$

since the integrated part is equal to zero because of the boundary conditions $f(\pm\infty) \to 0$. Thus, the adjoint of the operator $L = \mathrm{d}/\mathrm{d}x$ is $L^{+} = -\mathrm{d}/\mathrm{d}x$ in this space.

Example 3.3.1. In the space of square integrable functions $f(x)$ on the interval $-\infty < x < \infty$, find the adjoint of the operators (a) $L = \mathrm{d}^2/\mathrm{d}x^2$, and (b) $L = \frac{1}{\mathrm{i}}\dfrac{\mathrm{d}}{\mathrm{d}x}$.

Solution 3.3.1. (a) $L = \dfrac{\mathrm{d}^2}{\mathrm{d}x^2}$,

$$\langle Lf \mid g\rangle = \left\langle \frac{\mathrm{d}^2}{\mathrm{d}x^2}f \middle| g \right\rangle = \left\langle \frac{\mathrm{d}}{\mathrm{d}x}f \middle| -\frac{\mathrm{d}}{\mathrm{d}x}g \right\rangle = \left\langle f \middle| \frac{\mathrm{d}^2}{\mathrm{d}x^2}g \right\rangle = \left\langle f \mid L^{+}g\right\rangle.$$

Therefore the adjoint of $\mathrm{d}^2/\mathrm{d}x^2$ is $L^{+} = \mathrm{d}^2/\mathrm{d}x^2$.

(b) $L = \frac{1}{\mathrm{i}}\frac{\mathrm{d}}{\mathrm{d}x}$,

$$\langle Lf \,| g\rangle = \left\langle \frac{1}{\mathrm{i}}\frac{\mathrm{d}}{\mathrm{d}x}f \middle| g \right\rangle = \frac{1}{-\mathrm{i}}\left\langle \frac{\mathrm{d}}{\mathrm{d}x}f \middle| g \right\rangle = \frac{1}{-\mathrm{i}}\left\langle f \middle| -\frac{\mathrm{d}}{\mathrm{d}x}g \right\rangle = \left\langle f \middle| \frac{1}{\mathrm{i}}\frac{\mathrm{d}}{\mathrm{d}x}g \right\rangle,$$

where we have used (3.2) and (3.3). Therefore the adjoint of $L = \frac{1}{\mathrm{i}}\frac{\mathrm{d}}{\mathrm{d}x}$ is $L^+ = \frac{1}{\mathrm{i}}\frac{\mathrm{d}}{\mathrm{d}x}$.

An operator is said to be self-adjoint (or Hermitian) if $L^+ = L$. Thus, in the above example, the operators $\frac{\mathrm{d}^2}{\mathrm{d}x^2}$ and $\frac{1}{\mathrm{i}}\frac{\mathrm{d}}{\mathrm{d}x}$ are Hermitian, but $\mathrm{d}/\mathrm{d}x$ is not Hermitian since $L^+ = -\mathrm{d}/\mathrm{d}x$ which is not the same as $L = \mathrm{d}/\mathrm{d}x$.

In this example, the weight function $w(x)$ is taken to be unity. In general, $w(x)$ can be any real and positive function. Furthermore, the space can be defined in any interval. If x is specified to be on the interval $a \leq x \leq b$, the general expressions of inner products take the following forms.

$$\langle Lf \,| g\rangle = \int_{-\infty}^{\infty} (Lf(x))^* g(x)w(x)\mathrm{d}x,$$

$$\langle f \,| Lg\rangle = \int_{-\infty}^{\infty} f^*(x)Lg(x)w(x)\mathrm{d}x.$$

Since $w(x)$ is real, and

$$\int_{-\infty}^{\infty} (Lf(x))^* g(x)w(x)\ \mathrm{d}x = \left(\int_{-\infty}^{\infty} g^*(x)Lf(x)w(x)\ \mathrm{d}x\right)^*,$$

a self-adjoint operator L can also be expressed as

$$\int_{-\infty}^{\infty} f^*(x)Lg(x)w(x)\ \mathrm{d}x = \left(\int_{-\infty}^{\infty} g^*(x)Lf(x)w(x)\mathrm{d}x\right)^*.$$

Symbolically, this also follows from the fact that $\langle Lf \,| g\rangle = \langle f \,| Lg\rangle$ and $\langle Lf \,| g\rangle = \langle g \,| Lf\rangle^*$, so

$$\langle f \,| Lg\rangle = \langle g \,| Lf\rangle^*. \tag{3.7}$$

In a finite dimensional space, the eigenvalues of a Hermitian matrix are real and the eigenvectors form an orthogonal basis. In an infinite dimensional space, the Hermitian differential operator plays the same role as the Hermitian matrix in the finite dimensional space. Corresponding to the matrix eigenvalue problem, we have the eigenvalue problem of differential operator

$$L\phi(x) = \lambda\phi(x),$$

where λ is a constant. For a given choice of λ, a function which satisfies the equation and the imposed boundary conditions is called an eigenfunction corresponding to λ. The constant λ is then called an eigenvalue. There is no guarantee the eigenfunction $\phi(x)$ will exist for any arbitrary choice of the parameter λ. The requirement that there be an eigenfunction often restricts the acceptable values of λ to a discrete set. We shall see in Sect. 3.3.2 that the eigenvalues of a Hermitian operator are real and the eigenfunctions form a complete orthogonal basis set.

Furthermore, the elements a_{ij} of a Hermitian matrix are characterized by the relation

$$a_{ij} = a_{ji}^{*}. \tag{3.8}$$

In analogy, we often define a "matrix element" L_{ij} of a Hermitian operator

$$L_{ij} = \langle \phi_i \left| L\phi_j \right\rangle .$$

By (3.7),

$$\langle \phi_i \left| L\phi_j \right\rangle = \left\langle \phi_j \left| L\phi_i \right\rangle^{*} .$$

Therefore

$$L_{ij} = L_{ji}^{*}. \tag{3.9}$$

The similarity between (3.8) and (3.9) is obvious.

In quantum mechanics, the expectation value of an observable (a physical quantity that can be observed), such as energy and momentum, is the average value of many measurements of that quantity. The outcome of a measurement is of course a real number. Furthermore, the observable is represented by an operator O and the expectation value is given by $\langle \Psi \left| O\, \Psi \right\rangle$ where Ψ is the wave function describing the state of the system. Thus $\langle \Psi \left| O\, \Psi \right\rangle$ must be real, that is

$$\langle \Psi \left| O\, \Psi \right\rangle^{*} = \langle \Psi \left| O\, \Psi \right\rangle .$$

Since

$$\langle \Psi \left| O\, \Psi \right\rangle^{*} = \langle O\Psi \left| \Psi \right\rangle ,$$

it follows

$$\langle O\Psi \left| \Psi \right\rangle = \langle \Psi \left| O\, \Psi \right\rangle .$$

Therefore any operator representing an observable must be Hermitian.

3.3.2 Properties of Hermitian Operators

The Eigenvalues of a Hermitian Operator are Real. Let λ be an eigenvalue of the operator L and ϕ be the corresponding eigenfunction

$$L\phi = \lambda\phi.$$

So

$$\langle L\phi \left| \phi \right\rangle = \langle \lambda\phi \left| \phi \right\rangle = \lambda^{*} \langle \phi \left| \phi \right\rangle .$$

Since L is Hermitian, it follows that

$$\langle L\phi \,|\phi\rangle = \langle \phi \,|L\phi\rangle = \langle \phi \,|\lambda\phi\rangle = \lambda \langle \phi \,|\phi\rangle .$$

Thus

$$\lambda^* \langle \phi \,|\phi\rangle = \lambda \langle \phi \,|\phi\rangle .$$

Therefore

$$\lambda^* = \lambda,$$

the eigenvalue of a Hermitian operator must be real.

It is interesting to note that the Hermitian operator can be imaginary. Even if the operator is real, the eigenfunction can be complex. But in all cases, the eigenvalues must be real.

Because the eigenvalues are real, the eigenfunctions of a real Hermitian operator can always be made real by taking a suitable linear combinations. Since by definition

$$L\phi_i = \lambda_i\phi_i,$$

the complex conjugate is given by

$$L\phi_i^* = \lambda_i^*\phi_i^* = \lambda_i\phi_i^*,$$

where we have used the fact $\lambda^* = \lambda$. Thus both ϕ_i and ϕ_i^* are eigenfunctions corresponding to the same eigenvalue. Because of the linearity of L, any linear combination of ϕ_i and ϕ_i^* must also be an eigenfunction. Now both $\phi_i + \phi_i^*$ and $\mathrm{i}(\phi_i - \phi_i^*)$ are real, so we can take them as eigenfunctions for the eigenvalue λ_i. So for a real operator, we can assume both eigenvalues and eigenfunctions are real.

The Eigenfunctions of a Hermitian Operator are Orthogonal. Let ϕ_i and ϕ_j be eigenfunctions corresponding to two different eigenvalues λ_i and λ_j,

$$L\phi_i = \lambda_i\phi_i,$$
$$L\phi_j = \lambda_j\phi_j.$$

It follows that

$$\langle L\phi_i \,|\phi_j\rangle = \langle \lambda_i\phi_i \,|\phi_j\rangle = \lambda_i^* \langle \phi_i \,|\phi_j\rangle = \lambda_i \langle \phi_i \,|\phi_j\rangle ,$$

the last equality follows from the fact that the eigenvalues are real. Since L is Hermitian,

$$\langle L\phi_i \,|\phi_j\rangle = \langle \phi_i \,|L\phi_j\rangle = \langle \phi_i \,|\lambda_j\phi_j\rangle = \lambda_j \langle \phi_i \,|\phi_j\rangle .$$

Thus

$$\lambda_i \langle \phi_i \,|\phi_j\rangle = \lambda_j \langle \phi_i \,|\phi_j\rangle ,$$
$$(\lambda_i - \lambda_j) \langle \phi_i \,|\phi_j\rangle = 0.$$

Since $\lambda_i \neq \lambda_j$, we must have

$$\langle \phi_i | \phi_j \rangle = 0.$$

Therefore ϕ_i and ϕ_j are orthogonal.

Degeneracy. If n linear independent eigenfunctions correspond to the same eigenvalue, the eigenvalue is said to be n-fold degenerate. If this is the case, we cannot use the above argument to show that these eigenfunctions are orthogonal and they may not be. However, if they are not orthogonal, we can use the Gram–Schmidt process to construct n-orthogonal functions out of the n linearly independent eigenfunctions. These newly constructed functions will satisfy the same equation and be orthogonal to each other and to other eigenfunctions belonging to different eigenvalues.

The Eigenfunctions of an Hermitian Operator form a Complete Set. Recall that a Hermitian matrix can always be diagonalized. The eigenvector of a diagonalized matrix is a column vector with only one nonzero element. For example

$$\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \lambda_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \lambda_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Any vector in this two-dimensional space can be expressed in terms of these two eigenvectors

$$\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = c_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

We say that these two eigenvectors form a complete orthogonal basis. Clearly, the eigenvectors of a $n \times n$ Hermitian matrix will form a complete orthogonal basis for the n-dimensional space.

One would expect that in an infinite dimensional vector space of functions, the eigenfunctions of a Hermitian operator will form a complete set of orthogonal basis. This is indeed the case. A proof of this fact can be found in "Methods of Mathematical Physics", Chap. 6, by Courant and Hilbert, Interscience Publishers (1953), Reprinted by Wiley (1989).

Thus, in the interval where the linear operator L is Hermitian, any piecewise continuous function $f(x)$ can be expressed in a generalized Fourier series of eigenfunctions of L, that is, if the set of eigenfunctions $\{\phi_n\}$ $(n = 0, 1, 2, \ldots)$ is normalized, then

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

where $L\phi_n = \lambda_n \phi_n$.

It is to be emphasized that in the space where L is Hermitian, the functions in this space have to satisfy certain boundary conditions. It is these boundary conditions that determine the eigenfunctions. Let us illustrate this point with the following example.

Example 3.3.2. (a) Let the weight function be equal to unity $w(x) = 1$, find the required boundary conditions for the differential operator $L = \mathrm{d}^2/\mathrm{d}x^2$ to be Hermitian over the interval $a \le x \le b$. (b) Show that if the solutions of $Ly = \lambda y$ in the interval $0 \le x \le 2\pi$ satisfy the boundary conditions $y(0) = y(2\pi),\ y'(0) = y'(2\pi)$, (where y' means the derivative of y with respect to x), then the operator L in this interval is Hermitian. (c) Find the complete set of eigenfunctions of L.

Solution 3.3.2. (a) Let $y_i(x)$ and $y_j(x)$ be two functions in this space. Integrating the inner product $\langle y_i | L y_j \rangle$ by parts gives

$$\langle y_i | L y_j \rangle = \int_a^b y_i^* \frac{\mathrm{d}^2 y_j}{\mathrm{d}x^2} \mathrm{d}x = \left[y_i^* \frac{\mathrm{d}y_j}{\mathrm{d}x} \right]_a^b - \int_a^b \frac{\mathrm{d}y_i^*}{\mathrm{d}x} \frac{\mathrm{d}y_j}{\mathrm{d}x} \mathrm{d}x.$$

Integrating the second term on the right-hand side by parts again yields

$$\int_a^b \frac{\mathrm{d}y_i^*}{\mathrm{d}x} \frac{\mathrm{d}y_j}{\mathrm{d}x} \mathrm{d}x = \left[\frac{\mathrm{d}y_i^*}{\mathrm{d}x} y_j \right]_a^b - \int_a^b \frac{\mathrm{d}^2 y_i^*}{\mathrm{d}x^2} y_j \mathrm{d}x.$$

Thus

$$\langle y_i | L y_j \rangle = \left[y_i^* \frac{\mathrm{d}y_j}{\mathrm{d}x} \right]_a^b - \left[\frac{\mathrm{d}y_i^*}{\mathrm{d}x} y_j \right]_a^b + \langle L y_i | y_j \rangle .$$

Therefore L is Hermitian provided

$$\left[y_i^* \frac{\mathrm{d}y_j}{\mathrm{d}x} \right]_a^b - \left[\frac{\mathrm{d}y_i^*}{\mathrm{d}x} y_j \right]_a^b = 0.$$

(b) Because of the boundary conditions $y(0) = y(2\pi),\ y'(0) = y'(2\pi)$,

$$\left[y_i^* \frac{\mathrm{d}y_j}{\mathrm{d}x} \right]_0^{2\pi} = y_i^*(2\pi) y_j'(2\pi) - y_i^*(0) y_j'(0) = 0,$$

$$\left[\frac{\mathrm{d}y_i^*}{\mathrm{d}x} y_j \right]_0^{2\pi} = y_i^{*\prime}(2\pi) y_j(2\pi) - y_i^{*\prime}(0) y_j(0) = 0.$$

Therefore L is Hermitian in this interval, since

$$\langle y_i | L y_j \rangle = \left[y_i^* \frac{\mathrm{d}y_j}{\mathrm{d}x} \right]_a^b - \left[\frac{\mathrm{d}y_i^*}{\mathrm{d}x} y_j \right]_a^b + \langle L y_i | y_j \rangle = \left\langle y_i \middle| L^+ y_j \right\rangle .$$

(c) To find the eigenfunctions of L, we must solve the differential equation

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

subject to the boundary conditions

$$y(0) = y(2\pi), \qquad y'(0) = y'(2\pi).$$

The solution of the differential equation is

$$y(x) = A\cos\sqrt{\lambda}x + B\sin\sqrt{\lambda}x,$$

where A and B are two arbitrary constants. So

$$y'(x) = -\sqrt{\lambda}A\sin\sqrt{\lambda}x + \sqrt{\lambda}B\cos\sqrt{\lambda}x,$$

and

$$\begin{aligned} y(0) &= A, & y(2\pi) &= A\cos\sqrt{\lambda}2\pi + B\sin\sqrt{\lambda}2\pi, \\ y'(0) &= \sqrt{\lambda}B, & y'(2\pi) &= -\sqrt{\lambda}A\sin\sqrt{\lambda}2\pi + \sqrt{\lambda}B\cos\sqrt{\lambda}2\pi. \end{aligned}$$

Because of the boundary conditions $y(0) = y(2\pi),\ y'(0) = y'(2\pi)$,

$$\begin{aligned} A &= A\cos\sqrt{\lambda}2\pi + B\sin\sqrt{\lambda}2\pi, \\ \sqrt{\lambda}B &= -\sqrt{\lambda}A\sin\sqrt{\lambda}2\pi + \sqrt{\lambda}B\cos\sqrt{\lambda}2\pi, \end{aligned}$$

or

$$\begin{aligned} A\left(1-\cos\sqrt{\lambda}2\pi\right) - B\sin\sqrt{\lambda}2\pi &= 0 \\ A\sin\sqrt{\lambda}2\pi + B\left(1-\cos\sqrt{\lambda}2\pi\right) &= 0. \end{aligned}$$

A and B will have nontrivial solutions if and only if

$$\begin{vmatrix} 1-\cos\sqrt{\lambda}2\pi & -\sin\sqrt{\lambda}2\pi \\ \sin\sqrt{\lambda}2\pi & 1-\cos\sqrt{\lambda}2\pi \end{vmatrix} = 0.$$

It follows that

$$1 - 2\cos\sqrt{\lambda}2\pi + \cos^2\sqrt{\lambda}2\pi + \sin^2\sqrt{\lambda}2\pi = 0,$$

or

$$2 - 2\cos\sqrt{\lambda}2\pi = 0.$$

Thus

$$\cos\sqrt{\lambda}2\pi = 1$$

and

$$\sqrt{\lambda} = n, \qquad n = 0, 1, 2, \ldots.$$

Hence, for each integer n, the solution is

$$y_n(x) = A_n\cos nx + B_n\sin nx.$$

In other words, for this periodic boundary conditions, the eigenfunctions of this Hermitian operator $\mathrm{d}^2/\mathrm{d}x^2$ are $\cos nx$ and $\sin nx$. This means that the collection of $\{\cos nx,\ \sin nx\}$ $(n = 0, 1, 2, \ldots)$ is a complete basis set for this space. Therefore, any piecewise continuous periodic function with period of 2π can be expanded in terms of these eigenfunctions. This expansion is, of course, just the regular Fourier series.

A systematic account of the relations between the boundary conditions and the eigenfunctions of the second-order differential equations is provided by the Sturm–Liouville theory.

3.4 Sturm–Liouville Theory

In the last example, we have seen that the eigenfunctions of the differential operator $\mathrm{d}^2/\mathrm{d}x^2$ with some boundary conditions form a complete set of orthogonal basis. A far more general eigenvalue problem of second-order differential operators is the Sturm–Liouville problem.

3.4.1 Sturm–Liouville Equations

A linear second-order differential equation

$$A(x)\frac{\mathrm{d}^2}{\mathrm{d}x^2}y + B(x)\frac{\mathrm{d}}{\mathrm{d}x}y + C(x)y + \lambda D(x)y = 0, \tag{3.10}$$

where λ is a parameter to be determined by the boundary conditions, can be put in the form of

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2}y + b(x)\frac{\mathrm{d}}{\mathrm{d}x}y + c(x)y + \lambda\,\mathrm{d}(x)y = 0 \tag{3.11}$$

by dividing every term by $A(x)$, provided $A(x) \neq 0$. Let us define an integrating factor $p(x)$,

$$p(x) = \mathrm{e}^{\int^x b(x')\mathrm{d}x'}.$$

Multiplying (3.11) by $p(x)$, we have

$$p(x)\frac{\mathrm{d}^2}{\mathrm{d}x^2}y + p(x)b(x)\frac{\mathrm{d}}{\mathrm{d}x}y + p(x)c(x)y + \lambda p(x)d(x)y = 0. \tag{3.12}$$

Since

$$\frac{\mathrm{d}p(x)}{\mathrm{d}x} = \frac{\mathrm{d}}{\mathrm{d}x}\mathrm{e}^{\int^x b(x')\mathrm{d}x'} = \mathrm{e}^{\int^x b(x')\mathrm{d}x'}\frac{\mathrm{d}}{\mathrm{d}x}\int^x b(x')\mathrm{d}x = p(x)b(x),$$

so

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[p(x)\frac{\mathrm{d}}{\mathrm{d}x}y\right]=p(x)\frac{\mathrm{d}^2}{\mathrm{d}x^2}y+\frac{\mathrm{d}p(x)}{\mathrm{d}x}\frac{\mathrm{d}}{\mathrm{d}x}y=p(x)\frac{\mathrm{d}^2}{\mathrm{d}x^2}y+p(x)b(x)\frac{\mathrm{d}}{\mathrm{d}x}y.$$

Thus, (3.12) can be written as

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[p(x)\frac{\mathrm{d}}{\mathrm{d}x}y\right]+q(x)y+\lambda w(x)y=0, \tag{3.13}$$

where $q(x)=p(x)c(x)$ and $w(x)=p(x)d(x)$. Since the factor $p(x)$ is everywhere nonzero, the solutions of (3.10)–(3.13) are identical, so these equations are equivalent.

Under the general conditions that p, q, w are real and continuous, and both $p(x)$ and $w(x)$ are positive on certain interval, equations in the form of (3.13) are known as Sturm–Liouville equations, named after French mathematicians Sturm (1803 − 1855) and Liouville (1809 − 1882), who first developed an extensive theory of these equations.

These equations can be put in the usual eigenvalue problem form

$$Ly=\lambda y$$

by defining a Sturm–Liouville operator

$$L=-\frac{1}{w(x)}\left[\frac{\mathrm{d}}{\mathrm{d}x}\left(p(x)\frac{\mathrm{d}}{\mathrm{d}x}\right)+q(x)\right]. \tag{3.14}$$

Sturm–Liouville theory is very important in engineering and physics, because under a variety of boundary conditions on the solution $y(x)$, linear operators that can be written in this form are Hermitian. Therefore the eigenfunctions of the Sturm–Liouville equations form complete sets of orthogonal bases for the function space in which the weight function is $w(x)$. The set of cosine and sine functions of Fourier series is just one example within a broader Sturm–Liouville theory.

We note that the definitions of the Sturm–Liouville operator vary; some authors use

$$L=\frac{\mathrm{d}}{\mathrm{d}x}\left(p\frac{\mathrm{d}}{\mathrm{d}x}\right)+q(x)$$

and write the eigenvalue equation as

$$Ly=-\lambda wy.$$

As long as it is consistent, the difference is just a matter of convention. We will use (3.14) as the definition of the Sturm–Liouville operator.

3.4.2 Boundary Conditions of Sturm–Liouville Problems

Sturm–Liouville Operators as Hermitian Operators. Let L be the Sturm–Liouville operator in (3.14), and $f(x)$ and $g(x)$ be two functions having continuous second derivatives on the interval $a \le x \le b$, then

$$\langle Lf \mid g\rangle = \int_a^b \left\{ -\frac{1}{w}\left[\frac{\mathrm{d}}{\mathrm{d}x}\left(p\frac{\mathrm{d}}{\mathrm{d}x}\right)+q\right] f\right\}^* gw \,\mathrm{d}x.$$

Since $p,\ q,\ w$ are real, the integral can be written as

$$\langle Lf \mid g\rangle = -\int_a^b \frac{\mathrm{d}}{\mathrm{d}x}\left(p\frac{\mathrm{d}}{\mathrm{d}x}f^*\right) g\,\mathrm{d}x - \int_a^b qf^*g\,\mathrm{d}x.$$

With integration by parts,

$$\int_a^b \frac{\mathrm{d}}{\mathrm{d}x}\left(p\frac{\mathrm{d}f^*}{\mathrm{d}x}\right) g\,\mathrm{d}x = p\frac{\mathrm{d}f^*}{\mathrm{d}x}g\bigg|_a^b - \int_a^b p\frac{\mathrm{d}f^*}{\mathrm{d}x}\frac{\mathrm{d}g}{\mathrm{d}x}\mathrm{d}x,$$

and

$$\int_a^b p\frac{\mathrm{d}f^*}{\mathrm{d}x}\frac{\mathrm{d}g}{\mathrm{d}x}\mathrm{d}x = \int_a^b \frac{\mathrm{d}f^*}{\mathrm{d}x}p\frac{\mathrm{d}g}{\mathrm{d}x}\mathrm{d}x = f^*p\frac{\mathrm{d}g}{\mathrm{d}x}\bigg|_a^b - \int_a^b f^*\frac{\mathrm{d}}{\mathrm{d}x}\left(p\frac{\mathrm{d}g}{\mathrm{d}x}\right)\mathrm{d}x.$$

It follows that

$$\langle Lf \mid g\rangle = -\,p\frac{\mathrm{d}f^*}{\mathrm{d}x}g\bigg|_a^b + f^*p\frac{\mathrm{d}g}{\mathrm{d}x}\bigg|_a^b - \int_a^b f^*\frac{\mathrm{d}}{\mathrm{d}x}\left(p\frac{\mathrm{d}g}{\mathrm{d}x}\right)\mathrm{d}x - \int_a^b qf^*g\,\mathrm{d}x,$$

or

$$\begin{aligned}\langle Lf \mid g\rangle &= \left[p\left(f^*\frac{\mathrm{d}g}{\mathrm{d}x} - \frac{\mathrm{d}f^*}{\mathrm{d}x}g\right)\right]_a^b + \int_a^b f^*\left\{-\frac{1}{w}\left[\frac{\mathrm{d}}{\mathrm{d}x}\left(p\frac{\mathrm{d}}{\mathrm{d}x}\right)+q\right]g\right\} w\,\mathrm{d}x \\ &= \left[p\left(f^*\frac{\mathrm{d}g}{\mathrm{d}x} - \frac{\mathrm{d}f^*}{\mathrm{d}x}g\right)\right]_a^b + \langle f \mid Lg\rangle .\end{aligned}$$

It is clear that if

$$\left[p\left(f^*\frac{\mathrm{d}g}{\mathrm{d}x} - \frac{\mathrm{d}f^*}{\mathrm{d}x}g\right)\right]_a^b = 0, \tag{3.15}$$

then

$$\langle Lf \mid g\rangle = \langle f \mid Lg\rangle .$$

In other words, if the function space consists of functions that satisfy (3.15), then the Sturm–Liouville operator L is Hermitian in that space.

Sturm–Liouville Problems. It is customary to refer to the Sturm–Liouville equation and the boundary conditions together as the Sturm–Liouville problem. Since the operator is Hermitian, the eigenfunctions of the Sturm–Liouville problem are orthogonal to each other with respect to the weight function $w(x)$ and they are complete. Therefore they can be used as basis for the generalized Fourier series, which is also called eigenfunction expansion.

If any two solutions $y_n(x)$ and $y_m(x)$ of the linear homogeneous second-order differential equation

$$[p(x)y'(x)]' + q(x)y(x) + \lambda wy(x) = 0, \qquad a \leq x \leq b$$

satisfy the boundary condition (3.15), then the equation together with its boundary conditions is called a Sturm–Liouville problem. Since the operator is real, the eigenfunctions can also be taken as real. Therefore the boundary condition (3.15) can be conveniently written as

$$p(b) \begin{vmatrix} y_n(b) & y_n'(b) \\ y_m(b) & y_m'(b) \end{vmatrix} - p(a) \begin{vmatrix} y_n(a) & y_n'(a) \\ y_m(a) & y_m'(a) \end{vmatrix} = 0. \tag{3.16}$$

Depending on how the boundary conditions are met, Sturm–Liouville problems are divided into the following subgroups.

3.4.3 Regular Sturm–Liouville Problems

In this case, $p(a) \neq 0$ and $p(b) \neq 0$. The Sturm–Liouville problem consists of the equation

$$Ly(x) = \lambda y(x)$$

with L given by (3.14), and the boundary conditions

$$\alpha_1 y(a) + \alpha_2 y'(a) = 0,$$
$$\beta_1 y(b) + \beta_2 y'(b) = 0,$$

where the constants α_1 and α_2 cannot both be zero, and β_1 and β_2 also cannot both be zero.

Let us show that these boundary conditions satisfy (3.16). If $y_n(x)$ and $y_m(x)$ are two different solutions of the problem, both have to satisfy the boundary conditions. The first boundary condition requires

$$\alpha_1 y_n(a) + \alpha_2 y_n'(a) = 0,$$
$$\alpha_1 y_m(a) + \alpha_2 y_m'(a) = 0.$$

This is a system of two simultaneous equations in α_1 and α_2. Since α_1 and α_2 cannot both be zero, the determinant of the coefficients must be zero,

$$\begin{vmatrix} y_n(a) & y_n'(a) \\ y_m(a) & y_m'(a) \end{vmatrix} = 0.$$

Similarly, the second boundary condition requires

$$\begin{vmatrix} y_n(b) & y_n'(b) \\ y_m(b) & y_m'(b) \end{vmatrix} = 0.$$

Clearly,

$$p(b)\begin{vmatrix} y_n(b) & y_n'(b) \\ y_m(b) & y_m'(b) \end{vmatrix} - p(a)\begin{vmatrix} y_n(a) & y_n'(a) \\ y_m(a) & y_m'(a) \end{vmatrix} = 0.$$

Therefore the boundary condition of (3.16) is satisfied.

Example 3.4.1. (a) Show that for $0 \le x \le 1$,

$$y'' + \lambda y = 0$$
$$y(0) = 0, \quad y(1) = 0,$$

constitute a regular Sturm–Liouville problem.

(b) Find the eigenvalues and eigenfunctions of the problem.

Solution 3.4.1. (a) With $p(x) = 1,\ q(x) = 0,\ w(x) = 1$, the Sturm–Liouville equation

$$(py')' + qy + \lambda wy = 0$$

becomes

$$y'' + \lambda y = 0.$$

Furthermore, with $a = 0,\ b = 1,\ \alpha_1 = 1,\ \alpha_2 = 0,\ \beta_1 = 1,\ \beta_2 = 0$, the boundary conditions

$$\alpha_1 y(a) + \alpha_2 y'(a) = 0,$$

$$\beta_1 y(b) + \beta_2 y'(b) = 0,$$

become

$$y(0) = 0, \qquad y(1) = 0.$$

Therefore the given equation and the boundary conditions constitute a regular Sturm–Liouville problem.

(b) To find the eigenvalues, let us look at the possibilities of $\lambda = 0, \lambda < 0$, $\lambda > 0$.

If $\lambda = 0$, the solution of the equation is given by

$$y(x) = c_1 x + c_2.$$

Applying the boundary conditions, we have

$$y(0) = c_2 = 0, \quad y(1) = c_1 + c_2 = 0,$$

so $c_1 = 0$ and $c_2 = 0$. This is a trivial solution. Therefore $\lambda = 0$ is not an eigenvalue.

If $\lambda < 0$, let $\lambda = -\mu^2$ with real μ, so the solution of the equation is

$$y(x) = c_1 \mathrm{e}^{\mu x} + c_2 \mathrm{e}^{-\mu x}.$$

The condition $y(0) = 0$ makes $c_2 = -c_1$. The condition $y(1) = 0$ requires

$$y(1) = c_1(\mathrm{e}^{u} - \mathrm{e}^{-\mu}) = 0.$$

Since $\mu \neq 0$, so $c_1 = 0$. Again this gives the trivial solution.

This leaves the only possibility that $\lambda > 0$. Let $\lambda = \mu^2$ with real μ, so the solution of the equation becomes

$$y(x) = c_1 \cos \mu x + c_2 \sin \mu x.$$

Applying the boundary condition $y(0) = 0$ leads to

$$y(0) = c_1 = 0.$$

Therefore we are left with

$$y(x) = c_2 \sin \mu x.$$

The boundary condition $y(1) = 0$ requires

$$c_2 \sin \mu = 0.$$

For the nontrivial solution, we must have

$$\sin \mu = 0.$$

This will occur if μ is an integer multiple of π,

$$\mu = n\pi, \quad n = 1,\ 2, \ldots.$$

Thus the eigenvalues are

$$\lambda_n = \mu^2 = (n\pi)^2, \qquad n = 1,\ 2, \ldots,$$

and the corresponding eigenfunctions are

$$y_n(x) = \sin n\pi x.$$

Of course, we can solve this problem without knowing that it is a Sturm–Liouville problem. The advantage of knowing that $\{\sin n\pi x\}$ $(n = 1, 2, \ldots)$ are eigenfunctions of a Sturm–Liouville problem is that immediately we know that they are orthogonal to each other. More importantly, we know that they form a complete set in the interval $0 \leq x \leq 1$.

Example 3.4.2. (a) Put the following problem into the Sturm–Liouville form,

$$y'' - 2y' + \lambda y = 0, \qquad 0 \le x \le \pi$$
$$y(0) = 0, \quad y(\pi) = 0.$$

(b) Find the eigenvalues and eigenfunctions of the problem.
(c) Find the eigenfunction expansion of a given function $f(x)$ on the interval $0 \le x \le \pi$.

Solution 3.4.2. (a) Let us first find the integrating factor p,

$$p(x) = \mathrm{e}^{\int^x (-2)dx'} = \mathrm{e}^{-2x}.$$

Multiplying the differential equation by $p(x)$, we have

$$\mathrm{e}^{-2x}y'' - 2\mathrm{e}^{-2x}y' + \lambda \mathrm{e}^{-2x}y = 0,$$

which can be written as

$$(\mathrm{e}^{-2x}y')' + \lambda\, \mathrm{e}^{-2x}y = 0.$$

This is a Sturm–Liouville equation with $p(x) = \mathrm{e}^{-2x},\ q(x) = 0$, and $w(x) = \mathrm{e}^{-2x}$.

(b) Since the original differential equation is an equation with constant coefficients, we seek the solution in the form of $y(x) = \mathrm{e}^{mx}$. With this trial solution, the equation becomes

$$(m^2 - 2m + \lambda)\mathrm{e}^{mx} = 0.$$

The roots of the characteristic equation $m^2 - 2m + \lambda = 0$ are

$$m = 1 \pm \sqrt{1-\lambda},$$

therefore

$$y(x) = \mathrm{e}^{x}\left(c_1\mathrm{e}^{\sqrt{1-\lambda}x} + c_2\mathrm{e}^{-\sqrt{1-\lambda}x}\right)$$

for $\lambda \neq 1$.

For $\lambda = 1$, the characteristic equation has a double root at $m = 1$, and the solution becomes

$$y_2(x) = c_3 x + c_4.$$

The boundary conditions $y_2(0) = 0$ and $y_2(\pi) = 0$ require that $c_3 = c_4 = 0$. Therefore there is no nontrivial solution in this case, so $\lambda = 1$ is not an eigenvalue.

For $\lambda \neq 1$, the boundary condition $y(0) = 0$ requires

$$y(0) = c_1 + c_2 = 0.$$

Therefore the solution becomes

$$y(x) = c_1 \mathrm{e}^x \left(\mathrm{e}^{\sqrt{1-\lambda}x} - \mathrm{e}^{-\sqrt{1-\lambda}x} \right).$$

If $\lambda < 1$, the other boundary condition $y(\pi) = 0$ requires

$$y(\pi) = c_1 \mathrm{e}^\pi \left(\mathrm{e}^{\sqrt{1-\lambda}\pi} - \mathrm{e}^{-\sqrt{1-\lambda}\pi} \right) = 0.$$

This is possible only for the trivial solution of $c_1 = 0$. Therefore there is no eigenvalue less than 1.

For $\lambda > 1$, the solution can be written in the form of

$$\begin{aligned} y(x) &= c_1 \mathrm{e}^x \left(\mathrm{e}^{\mathrm{i}\sqrt{\lambda-1}x} - \mathrm{e}^{-\mathrm{i}\sqrt{\lambda-1}x} \right) \\ &= 2\mathrm{i}c_1 \mathrm{e}^x \sin\sqrt{\lambda-1}x. \end{aligned}$$

The boundary condition $y(\pi) = 0$ is satisfied if

$$\sin\sqrt{\lambda-1}\pi = 0.$$

This can occur if

$$\sqrt{\lambda-1} = n, \quad n = 1,\ 2, \ldots.$$

Therefore the eigenvalues are

$$\lambda_n = n^2 + 1, \quad n = 1,\ 2, \ldots,$$

and the eigenfunction associated with each eigenvalue λ_n is

$$\phi_n(x) = \mathrm{e}^x \sin nx.$$

Any arbitrary constant can be multiplied to $\phi_n(x)$ to give a solution for the problem with $\lambda = \lambda_n$.

(c) For a given function $f(x)$ on the interval $0 \le x \le \pi$, the eigenfunction expansion is

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

Since $\{\phi_n\}$ $(n = 1, 2, \ldots)$ is a set of eigenfunctions of the Sturm–Liouville problem, it is an orthogonal set with respect to the weight function $w(x) = \mathrm{e}^{-2x}$,

$$\langle \phi_n | \phi_m \rangle = \int_0^\pi (\mathrm{e}^x \sin nx)(\mathrm{e}^x \sin mx)\mathrm{e}^{-2x} \mathrm{d}x = 0, \quad for \ \ n \neq m.$$

For $n = m$,

$$\langle \phi_n | \phi_n \rangle = \int_0^\pi (\mathrm{e}^x \sin nx)(\mathrm{e}^x \sin nx)\mathrm{e}^{-2x} \mathrm{d}x = \int_0^\pi \sin^2 nx \ \mathrm{d}x = \frac{\pi}{2}.$$

Therefore

$$\langle \phi_n | \phi_m \rangle = \frac{\pi}{2}\delta_{nm}.$$

Taking the inner product of both sides of the eigenfunction expansion with ϕ_m, we have

$$\langle f | \phi_m \rangle = \sum_{n=1}^{\infty} c_n \langle \phi_n | \phi_m \rangle = \sum_{n=1}^{\infty} c_n \frac{\pi}{2}\delta_{nm} = \frac{\pi}{2} c_m.$$

Therefore

$$c_n = \frac{2}{\pi} \langle f | \phi_n \rangle,$$

where

$$\langle f | \phi_n \rangle = \int_0^{\pi} f(x) \mathrm{e}^{x} \sin nx \ \mathrm{e}^{-2x} \mathrm{d}x = \int_0^{\pi} f(x) \mathrm{e}^{-x} \sin nx \ \mathrm{d}x.$$

It follows that

$$f(x) = \sum_{n=1}^{\infty} \frac{2}{\pi} \langle f | \phi_n \rangle \phi_n$$

$$= \sum_{n=1}^{\infty} \frac{2}{\pi} \left(\int_0^{\pi} f(x) \mathrm{e}^{-x} \sin nx \ \mathrm{d}x \right) \mathrm{e}^{x} \sin nx.$$

Example 3.4.3. (a) Find the eigenvalues and eigenfunctions of the following Sturm–Liouville problem:

$$y'' + \lambda y = 0,$$

$$y(0) = 0, \quad y(1) - y'(1) = 0.$$

(b) Show that the eigenfunctions are orthogonal by explicit integration,

$$\int_0^1 y_n(x) y_m(x) \mathrm{d}x = 0, \quad n \neq m.$$

(c) Find the orthonormal set of the eigenfunctions.

Solution 3.4.3. (a) It can be easily shown that for $\lambda < 0$, no solution can satisfy the equation and the boundary conditions. For $\lambda = 0$, it is actually an eigenvalue with an associated eigenfunction $y_0(x) = x$, since it satisfies both the equation and the boundary conditions

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2} x = 0, \quad y_0(0) = 0, \quad y_0(1) - y_0'(1) = 1 - 1 = 0.$$

Most of the eigenvalues come from the branch where $\lambda = \alpha^2 > 0$. In that case, the solution of

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2} y(x) + \alpha^2 y(x) = 0$$

is given by

$$y(x) = A \cos \alpha x + B \sin \alpha x.$$

The boundary condition $y(0) = A = 0$ leaves us with

$$y(x) = B \sin \alpha x.$$

The other boundary condition $y(1) - y'(1) = 0$ requires that

$$\sin \alpha - \alpha \cos \alpha = 0. \tag{3.17}$$

Therefore α has to be the positive roots of

$$\tan \alpha = \alpha.$$

These roots are labeled as α_n in Fig. 3.1. The roots of the equation $\tan x = \mu x$ are frequently needed in many applications, and they are listed in Tables 4.19 and 4.20 in "Handbook of Mathematical Functions" by M. Abramowitz and I.A. Stegun, Dover Publications, 1970. For example, in our case $\mu = 1$, $\alpha_1 = 4.49341$, $\alpha_2 = 7.72525$, $\alpha_3 = 10.90412$, $\alpha_4 = 14.06619 \ldots$. Thus the eigenvalues of this Sturm–Liouville problem are $\lambda_0 = 0, \lambda_n = \alpha_n^2 (n = 1, 2, \ldots)$, the corresponding eigenfunctions are

$$y_0(x) = x, \ y_n(x) = \sin \alpha_n x \ (n = 1, 2, \ldots).$$

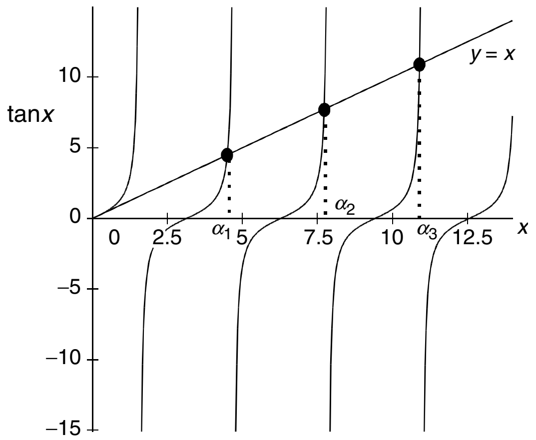


Fig. 3.1. Roots of $\tan x = x$, α_n is the nth root. $\alpha_1 = 4.49341, \alpha_2 = 7.72525$, $\alpha_3 = 10.90412, \ldots$ as listed in Table 4.19 of "Handbook of Mathematical Functions", by M. Abramowitz and I.A. Stegun, Dover Publications, 1970

(b) According to the Sturm–Liouville theory, these eigenfunctions are orthogonal to each other. It is instructive to show this explicitly. First,

$$\begin{aligned}\int_0^1 x\sin\alpha_n x\;\mathrm{d}x &= \left[-\frac{x}{\alpha_n}\cos\alpha_n x + \frac{1}{\alpha_n^2}\sin\alpha_n x\right]_0^1\\ &= \frac{1}{\alpha_n^2}\left[-\alpha_n\cos\alpha_n + \sin\alpha_n\right] = 0,\end{aligned}$$

since α_n satisfies (3.17). Next

$$\begin{aligned}\int_0^1 \sin\alpha_n x\sin\alpha_m x\;\mathrm{d}x &= \frac{1}{2}\int_0^1\left[\cos(\alpha_n-\alpha_m)x - \cos(\alpha_n+\alpha_m)x\right]\;\mathrm{d}x\\ &= \frac{1}{2}\left[\frac{\sin(\alpha_n-\alpha_m)}{\alpha_n-\alpha_m} - \frac{\sin(\alpha_n+\alpha_m)}{\alpha_n+\alpha_m}\right].\end{aligned}$$

Now

$$\begin{aligned}\alpha_n-\alpha_m = \tan\alpha_n - \tan\alpha_m &= \frac{\sin\alpha_n}{\cos\alpha_n} - \frac{\sin\alpha_m}{\cos\alpha_m}\\ &= \frac{\sin\alpha_n\cos\alpha_m - \cos\alpha_n\sin\alpha_m}{\cos\alpha_n\cos\alpha_m} = \frac{\sin(\alpha_n-\alpha_m)}{\cos\alpha_n\cos\alpha_m},\end{aligned}$$

thus

$$\frac{\sin(\alpha_n-\alpha_m)}{\alpha_n-\alpha_m} = \cos\alpha_n\cos\alpha_m.$$

Similarly

$$\frac{\sin(\alpha_n+\alpha_m)}{\alpha_n+\alpha_m} = \cos\alpha_n\cos\alpha_m.$$

It follows that

$$\int_0^1 \sin\alpha_n x\sin\alpha_m x\;\mathrm{d}x = \frac{1}{2}\left[\cos\alpha_n\cos\alpha_m - \cos\alpha_n\cos\alpha_m\right] = 0.$$

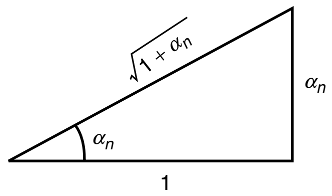
(c) To find the normalization constant $\beta_n^2 = \int_0^1 y_n^2(x)\;\mathrm{d}x$:

$$\beta_0^2 = \int_0^1 x^2\mathrm{d}x = \frac{1}{3}.$$

$$\begin{aligned}\beta_n^2 &= \int_0^1 \sin^2\alpha_n x\;\mathrm{d}x = \frac{1}{2}\int_0^1 (1-\cos 2\alpha_n x)\;\mathrm{d}x\\ &= \frac{1}{2}\left[x - \frac{\sin 2\alpha_n x}{2\alpha_n}\right]_0^1 = \frac{1}{2} - \frac{\sin 2\alpha_n}{4\alpha_n}\\ &= \frac{1}{2} - \frac{\sin\alpha_n\cos\alpha_n}{2\alpha_n}.\end{aligned}$$

Since $\tan\alpha_n = \alpha_n$, from the following diagram, we see that

$$\sin\alpha_n = \frac{\alpha_n}{\sqrt{1+\alpha_n^2}},\qquad \cos\alpha_n = \frac{1}{\sqrt{1+\alpha_n^2}}.$$



Thus

$$\beta_n^2 = \frac{1}{2}\left(1 - \frac{1}{\alpha_n}\frac{\alpha_n}{\sqrt{1+\alpha_n^2}}\frac{1}{\sqrt{1+\alpha_n^2}}\right) = \frac{\alpha_n^2}{2(1+\alpha_n^2)}$$

Therefore, the orthonormal set of the eigenfunctions is as follows:

$$\left\{\sqrt{3}x,\ \frac{\sqrt{2(1+\alpha_n^2)}}{\alpha_n}\sin\alpha_n x\right\}\ (n = 1,\ 2,\ 3, \ldots).$$

3.4.4 Periodic Sturm–Liouville Problems

On the interval $a \le x \le b$, if $p(a) = p(b)$, then the periodic boundary conditions

$$y(a) = y(b), \qquad y'(a) = y'(b)$$

also satisfy the condition (3.16). This is very easy to show. Let $y_n(x)$ and $y_m(x)$ be two functions that satisfy these boundary conditions, that is

$$y_n(a) = y_n(b), \qquad y_n'(a) = y_n'(b),$$
$$y_m(a) = y_m(b), \qquad y_m'(a) = y_m'(b).$$

Clearly

$$p(b)\begin{vmatrix} y_n(b) & y_n'(b) \\ y_m(b) & y_m'(b) \end{vmatrix} - p(a)\begin{vmatrix} y_n(a) & y_n'(a) \\ y_m(a) & y_m'(a) \end{vmatrix} = 0,$$

since the two terms are equal.

Therefore, a Sturm–Liouville equation plus these periodic boundary conditions also constitute a Sturm–Liouville problem. Note that the difference between the regular and periodic Sturm–Liouville problems is that the boundary conditions in the regular Sturm–Liouville problem are separated, with one condition applying at $x = a$ and the other at $x = b$, whereas the boundary conditions in the periodic Sturm–Liouville problem relate the values at $x = a$ to the values at $x = b$. In addition, in the periodic Sturm–Liouville problem, $p(a)$ must equal to $p(b)$.

For example,

$$y'' + \lambda y = 0, \qquad a \le x \le b$$

is a Sturm–Liouville equation with $p = 1,\ q = 0$, and $w = 1$. Since $p(a) = p(b) = 1$, a periodic boundary condition will make this a Sturm–Liouville problem. As we have seen, if $y(0) = y(2\pi),\ y'(0) = y'(2\pi)$, the eigenfunctions are $\{\cos nx,\ \sin nx\}\ (n = 0, 1, 2, \ldots)$, which is the basis of the ordinary Fourier series for any periodic function of period 2π.

Note that, within the interval of $0 \le x \le 2\pi$, any piece-wise continuous function $f(x)$, not necessarily periodic, can be expanded into a Fourier series of cosines and sines. However, outside the interval, since the trigonometric functions are periodic, $f(x)$ will also be periodic with period 2π.

If the period is not 2π, we can either make change of scale in the Fourier series, or change the boundary in the Sturm–Liouville problem. The result will be the same.

3.4.5 Singular Sturm–Liouville Problems

In this case, $p(x)$ (and possibly $w(x)$) vanishes at one or both endpoints. We call it singular, because Sturm–Liouville equation

$$(py')' + qy + \lambda wy = 0$$

can be written as

$$py'' + p'y' + qy + \lambda wy = 0,$$

or

$$y'' + \frac{1}{p}p'y' + \frac{1}{p}qy + \lambda\frac{1}{p}wy = 0.$$

If $p(a) = 0$, then clearly at $x = a$, this equation is singular.

If both $p(a)$ and $p(b)$ are zero, $p(a) = 0$ and $p(b) = 0$, the boundary condition (3.16) is automatically satisfied. This may suggest that there is no restriction on the eigenvalue λ. However, for an arbitrary λ, the equation may have no meaningful solution. The requirement that the solution and its derivative must remain bounded even at the singular points often restricts the acceptable values of λ to a discrete set. In other words, the boundary conditions in this case are replaced by the requirement that $y(x)$ must be bounded at $x = a$ and $x = b$.

If $p(a) = 0$ and $p(b) \neq 0$, then the boundary condition (3.16) becomes

$$\begin{vmatrix} y_n(b) & y_n'(b) \\ y_m(b) & y_m'(b) \end{vmatrix} = 0.$$

This condition will be met, if all solutions of the equation satisfy the boundary condition

$$\beta_1 y(b) + \beta_2 y'(b) = 0,$$

where constants β_1 and β_2 are not both zero. In addition, solutions must be bounded at $x = a$.

Similarly, if $p(b) = 0$ and $p(a) \neq 0$, then $y(x)$ must be bounded at $x = b$, and

$$\alpha_1 y(a) + \alpha_2 y'(a) = 0,$$

where α_1 and α_2 are not both equal to zero.

Many physically important and named differential equations are singular Sturm–Liouville problems. The following are a few examples.

Legendre Equation. The Legendre differential equation

$$(1 - x^2)y'' - 2xy' + \lambda y = 0, \qquad (-1 \leq x \leq 1)$$

is one of the most important equations in mathematical physics. The details of the solutions of this equation will be studied in Chap. 4. Here we only want to note that it is a singular Sturm–Liouville problem because this equation can be obviously written as

$$\left[(1 - x^2)y'\right]' + \lambda y = 0,$$

which is in the form of Sturm–Liouville equation with $p(x) = 1 - x^2$, $q = 0, w = 1$. Since $p(x)$ vanishes at both ends, $p(1) = p(-1) = 0$, it is a singular Sturm–Liouville problem. As we will see in Chap. 4, in order to have a bounded solution on $-1 \leq x \leq 1$, λ has to assume one of the following values

$$\lambda_n = n(n+1), \qquad n = 0,\ 1,\ 2,\ \ldots.$$

Corresponding to each λ_n, the eigenfunction is the Legendre function $P_n(x)$, which is a polynomial of order n. We have met these functions when we constructed an orthogonal set out of $\{x^n\}$ in the interval $-1 \leq x \leq 1$, with a unit weight function. The properties of this function will be discussed again in Chap. 4. Since $P_n(x)$ are eigenfunctions of a Sturm–Liouville problem, they are orthogonal to each other in the interval $-1 \leq x \leq 1$ with respect to a unit weight function $w(x) = 1$. Furthermore, the set $\{P_n(x)\}$ $(n = 0, 1, 2, \ldots)$ is complete. Therefore, any piece-wise continuous function $f(x)$ in the interval $-1 \leq x \leq 1$ can be expressed as

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

where

$$c_n = \frac{\langle f \mid P_n \rangle}{\langle P_n \mid P_n \rangle} = \frac{\int_{-1}^{1} f(x) P_n(x) \mathrm{d}x}{\int_{-1}^{1} P_n^2(x) \mathrm{d}x}.$$

This series is known a Fourier–Legendre series, which is very important in solving partial differential equations with spherical symmetry, as we shall see.

Bessel Equation. The problem consists of the differential equation

$$x^2 y''(x) + x y'(x) - \nu^2 y + \lambda^2 x^2 y(x) = 0, \qquad 0 \leq x \leq L \tag{3.18}$$

and the boundary condition

$$y(L) = 0.$$

It is a singular Sturm–Liouville problem. In the equation, ν^2 is a given constant and λ^2 is a parameter that can be chosen to fit the boundary condition. To convert this equation into the standard Sturm–Liouville form, let us first divide the equation by x^2,

$$y''(x) + \frac{1}{x}y'(x) - \frac{1}{x^2}\nu^2 y + \lambda^2 y(x) = 0, \tag{3.19}$$

and then find the integrating factor

$$p(x) = \mathrm{e}^{\int^x \frac{1}{x'}\mathrm{d}x'} = \mathrm{e}^{\ln x} = x.$$

Multiplying (3.19) by this integrating factor, we have

$$xy''(x) + y'(x) - \frac{1}{x}\nu^2 y(x) + \lambda^2 xy(x) = 0, \tag{3.20}$$

which can be written as

$$[xy']' - \frac{1}{x}\nu^2 y + \lambda^2 xy = 0.$$

This is a Sturm–Liouville equation with $p(x) = x$, $q(x) = -\nu^2/x$, $w(x) = x$. Of course, (3.20) can be obtained directly from (3.18) by dividing (3.18) by x. However, a step by step approach will enable us to handle more complicated equations, as we shall soon see.

Since $p(0) = 0$, there is a singular point at $x = 0$. So we only need the other boundary condition $y(L) = 0$ at $x = L$ to make it a Sturm–Liouville problem.

Equation (3.18) is closely related to the well known Bessel equation. To see this connection, let us make a change of variable, $t = \lambda x$,

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y}{\mathrm{d}t}\frac{\mathrm{d}t}{\mathrm{d}x} = \lambda\frac{\mathrm{d}y}{\mathrm{d}t},$$

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right) = \lambda\frac{\mathrm{d}}{\mathrm{d}t}\left(\lambda\frac{\mathrm{d}y}{\mathrm{d}t}\right) = \lambda^2\frac{\mathrm{d}^2y}{\mathrm{d}t^2}.$$

Thus

$$x\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{t}{\lambda}\lambda\frac{\mathrm{d}y}{\mathrm{d}t} = t\frac{\mathrm{d}y}{\mathrm{d}t},$$

$$x^2\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = \left(\frac{t}{\lambda}\right)^2\lambda^2\frac{\mathrm{d}^2y}{\mathrm{d}t^2} = t^2\frac{\mathrm{d}^2y}{\mathrm{d}t^2}.$$

It follows that (3.18) can be written as

$$t^2\frac{\mathrm{d}^2y}{\mathrm{d}t^2}+t\frac{\mathrm{d}y}{\mathrm{d}t}-\nu^2y+t^2y=0.$$

This is the Bessel equation which is very important in both pure mathematics and applied sciences. A great deal of information about this equation is known. We shall discuss some of its properties in Chap. 4.

There are two linearly independent solutions of this equation. One is known as the Bessel function $J_\nu(t)$, and the other, the Neumann function $N_\nu(t)$. The Bessel function is everywhere bounded, but the Neumann function goes to infinity as $t\to 0$.

Since $t=\lambda x$, the solution $y(x)$ of (3.18) must be

$$y(x)=AJ_\nu(\lambda x)+BN_\nu(\lambda x).$$

Since the solution must be bounded at $x=0$, therefore the constant B must be zero. Now the values of the Bessel functions $J_\nu(t)$ can be calculated, as we shall see in Chap. 4. As an example, we show in Fig. 3.2 the Bessel function of zeroth order $J_0(t)$ as a function t. Note that at certain values of t, it becomes zero. These values are known as the zeros of the Bessel functions, they are tabulated for many values of ν. For example, the first zero of $J_0(t)$ occur at $t=2.405$, the second zero at $t=5.520,\ldots$. These values are listed as $z_{01}=2.405,\ z_{02}=5.520,\ldots$.

The boundary condition $y(L)=0$ requires that

$$J_v(\lambda L)=0.$$

This means that λ can only assume certain discrete values such that

$$\lambda_1L=z_{\nu 1},\quad \lambda_2L=z_{\nu 2},\quad \lambda_3L=z_{\nu 3},\ldots\ .$$

That is,

$$\lambda_n=\frac{z_{\nu n}}{L}.$$

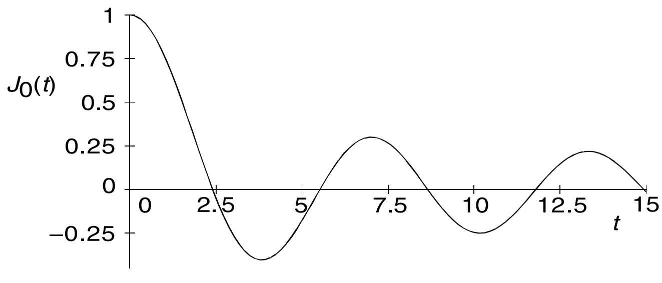


Fig. 3.2. Bessel function of zeroth order $J_0(t)$

It follows that the eigenfunctions of our Sturm–Liouville problem are

$$y_n(x) = J_\nu(\lambda_n x).$$

Now $J_\nu(\lambda_n x)$ and $J_\nu(\lambda_m x)$ are two different eigenfunctions corresponding two different eigenvalues λ_n and λ_m. The eigenfunctions are orthogonal to each other with respect to the weight function $w(x) = x$. Furthermore, $\{J_\nu(\lambda_n x)\}$ $(n = 1, 2, 3, \ldots.)$ is a complete set in the interval $0 \le x \le L$. Therefore any piece-wise continuous function $f(x)$ in this interval can be expanded in terms of these eigenfunctions,

$$f(x) = \sum_{n=1}^{\infty} c_n J_v(\lambda_n x),$$

where

$$c_n = \frac{\langle f(x) \, | J_\nu(\lambda_n x) \rangle}{\langle J_\nu(\lambda_n x) \, | J_\nu(\lambda_n x) \rangle} = \frac{\int_0^L f(x) J_\nu(\lambda_n x) x \, \mathrm{d}x}{\int_0^L \left[J_\nu(\lambda_n x)\right]^2 x \, \mathrm{d}x}.$$

This expansion is known as Fourier–Bessel series. It is needed in solving partial differential equations with cylindrical symmetry.

Example 3.4.4. Hermite Equation. Show that the following differential equation

$$y'' - 2xy' + 2\alpha y = 0, \qquad -\infty < x < \infty$$

forms a singular Sturm–Liouville problem. If $H_n(x)$ and $H_m(x)$ are two solutions of this problem, show that

$$\int_{-\infty}^{\infty} H_n(x) H_m(x) \mathrm{e}^{-x^2} \mathrm{d}x = 0 \quad \text{for} \quad n \neq m.$$

Solution 3.4.4. To put the equation into the Sturm–Liouville form, let us first calculate the integrating factor

$$p(x) = \mathrm{e}^{-\int^x 2x' \mathrm{d}x'} = \mathrm{e}^{-x^2}.$$

Multiplying the equation by this integrating factor, we have

$$\mathrm{e}^{-x^2} y'' - 2x \, \mathrm{e}^{-x^2} y' + 2\alpha \, \mathrm{e}^{-x^2} y = 0.$$

Since

$$\left[\mathrm{e}^{-x^2} y'\right]' = \mathrm{e}^{-x^2} y'' - 2x \, \mathrm{e}^{-x^2} y',$$

the equation can be written as

$$\left[\mathrm{e}^{-x^2} y'\right]' + 2\alpha \, \mathrm{e}^{-x^2} y = 0.$$

This is in the form of a Sturm–Liouville equation with $p(x) = \mathrm{e}^{-x^2}, q = 0$, $w(x) = \mathrm{e}^{-x^2}$. Since $p(\infty) = p(-\infty) = 0$, this is a singular Sturm–Liouville problem. Therefore, if $H_n(x)$ and $H_m(x)$ are two solutions of this problem, then they must be orthogonal with respect to the weight function $w(x) = \mathrm{e}^{-x^2}$, that is

$$\int_{-\infty}^{\infty} H_n(x)H_m(x)\mathrm{e}^{-x^2}\mathrm{d}x = 0 \quad \text{for} \quad n \neq m.$$

Example 3.4.5. Laguerre Equation. Show that the following differential equation

$$xy'' + (1-x)y' + ny = 0, \qquad 0 < x < \infty$$

forms a singular Sturm–Liouville problem. If $L_n(x)$ and $L_m(x)$ are two solutions of this problem, show that

$$\int_{0}^{\infty} L_n(x)L_m(x)\mathrm{e}^{-x}\mathrm{d}x = 0 \quad \text{for} \quad n \neq m.$$

Solution 3.4.5. To put the equation into the Sturm–Liouville form, let us first divide the equation by x

$$y'' + \frac{1-x}{x}y' + n\frac{1}{x}y = 0$$

and then calculate the integrating factor

$$p(x) = \mathrm{e}^{\int^x \frac{1-x'}{x'}\mathrm{d}x'} = \mathrm{e}^{\ln x - x} = x\ \mathrm{e}^{-x}.$$

Multiplying the last equation by this integrating factor, we have

$$x\ \mathrm{e}^{-x}y'' + (1-x)\mathrm{e}^{-x}y' + n\ \mathrm{e}^{-x}y = 0.$$

Since

$$\left[x\ \mathrm{e}^{-x}y'\right]' = x\ \mathrm{e}^{-x}y'' + (1-x)\mathrm{e}^{-x}y',$$

the equation can be written as

$$\left[x\ \mathrm{e}^{-x}y'\right]' + n\ \mathrm{e}^{-x}y = 0.$$

This is in the form of a Sturm–Liouville equation with $p(x) = x\ \mathrm{e}^{-x}, q = 0$, $w(x) = \mathrm{e}^{-x}$. Since $p(0) = p(\infty) = 0$, this is a singular Sturm–Liouville problem. Therefore, if $L_n(x)$ and $L_m(x)$ are two solutions of this problem, then they must be orthogonal with respect to the weight function $w(x) = \mathrm{e}^{-x}$, that is

$$\int_{-\infty}^{\infty} L_n(x)L_m(x)\mathrm{e}^{-x}\mathrm{d}x = 0 \quad \text{for} \quad n \neq m.$$

Example 3.4.6. Chebyshev Equation. Show that the following differential equation

$$(1-x^2)y'' - xy' + n^2 y = 0, \qquad -1 < x < 1$$

forms a singular Sturm–Liouville problem. If $T_n(x)$ and $T_m(x)$ are two solutions of this problem, show that

$$\int_0^{\infty} T_n(x) T_m(x) \frac{1}{\sqrt{1-x^2}} \mathrm{d}x = 0 \quad \text{for} \quad n \neq m.$$

Solution 3.4.6. To put the equation into the Sturm–Liouville form, let us first divide the equation by $(1-x^2)$

$$y'' - \frac{x}{1-x^2} y' + n^2 \frac{1}{1-x^2} y = 0$$

and then calculate the integrating factor

$$p(x) = \mathrm{e}^{-\int^x \frac{x'}{1-x'^2} \mathrm{d}x'}.$$

To evaluate the integral, let $u = 1-x^2, \mathrm{d}u = -2x \ \mathrm{d}x$, so

$$\int^x \frac{x'}{1-x'^2} \mathrm{d}x' = -\frac{1}{2} \int \frac{\mathrm{d}u}{u} = -\frac{1}{2} \ln u = -\frac{1}{2} \ln(1-x^2).$$

Thus,

$$p(x) = \mathrm{e}^{-\int^x \frac{x'}{1-x'^2} \mathrm{d}x'} = \mathrm{e}^{\frac{1}{2} \ln(1-x^2)} = \left[\mathrm{e}^{\ln(1-x^2)} \right]^{1/2} = (1-x^2)^{1/2}.$$

Multiplying the last equation by this integrating factor, we have

$$(1-x^2)^{1/2} y'' - (1-x^2)^{-1/2} xy' + n^2 (1-x^2)^{-1/2} y = 0.$$

Since

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

the equation can be written as

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

This is in the form of a Sturm–Liouville equation with $p(x) = (1-x^2)^{1/2}$, $q = 0$, $w(x) = (1-x^2)^{-1/2}$. Since $p(-1) = p(1) = 0$, this is a singular Sturm–Liouville problem. Therefore, if $T_n(x)$ and $T_m(x)$ are two solutions of this problem, then they must be orthogonal with respect to the weight function $w(x) = (1-x^2)^{-1/2}$, that is

$$\int_{-\infty}^{\infty} T_n(x) T_m(x) \frac{1}{\sqrt{1-x^2}} \mathrm{d}x = 0 \quad \text{for} \quad n \neq m.$$

3.5 Green's Function

3.5.1 Green's Function and Inhomogeneous Differential Equation

So far we have shown that if the solutions of the Sturm–Liouville equation satisfy certain boundary conditions, then they become a set of orthogonal eigenfunctions $y_n(x)$, with associated eigenvalues λ_n.

Now suppose that we want to solve the following inhomogeneous differential equation in the interval $a \leq x \leq b$,

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[p(x)\frac{\mathrm{d}}{\mathrm{d}x}y\right] + \mathrm{q}(x)y + kw(x)y = f(x), \tag{3.21}$$

where $f(x)$ is a given function. The boundary conditions to be satisfied by the solution $y(x)$ are the same as that satisfied by eigenfunctions $y_n(x)$ of the Sturm–Liouville problem

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

Note that $k \neq \lambda_n$. In fact, k can even be zero.

It is more convenient to work with the normalized eigenfunctions. If $y_n(x)$ is not yet normalized, we can define

$$\phi_n(x) = \frac{1}{\langle y_n \mid y_n\rangle^{1/2}} y_n(x),$$

so that

$$\langle \phi_m \mid \phi_n\rangle = \int_a^b \phi_m(x)\phi_n(x)w(x)\mathrm{d}x = \delta_{nm}.$$

Since $\{\phi_n\}$ $(n = 1, 2, \ldots)$ is a complete orthonormal set, the solution $y(x)$ of (3.21) can be expanded in terms of ϕ_n,

$$y(x) = \sum_{n=1}^{\infty} c_n\phi_n(x).$$

Putting it into (3.21), we have

$$\sum_{n=1}^{\infty} c_n \left\{\frac{\mathrm{d}}{\mathrm{d}x}\left[p(x)\frac{\mathrm{d}}{\mathrm{d}x}\right] + q(x)\right\}\phi_n(x) + kw(x)\sum_{n=1}^{\infty} c_n\phi_n(x) = f(x).$$

Since

$$\left\{\frac{\mathrm{d}}{\mathrm{d}x}\left[p(x)\frac{\mathrm{d}}{\mathrm{d}x}\right] + q(x)\right\}\phi_n(x) = -\lambda_n w(x)\phi_n(x),$$

so

$$\sum_{n=1}^{\infty} c_n(-\lambda_n + k)w(x)\phi_n(x) = f(x).$$

Multiplying both sides by $\phi_m(x)$ and integrating,

$$\sum_{n=1}^{\infty} c_n(-\lambda_n + k)\int_a^b w(x)\phi_n(x)\phi_m(x)\mathrm{d}x = \int_a^b f(x)\phi_m(x)\mathrm{d}x.$$

Because of the orthogonality condition, we have

$$c_m(-\lambda_m + k) = \int_a^b f(x)\phi_m(x)\mathrm{d}x,$$

or

$$c_n = \frac{1}{k-\lambda_n}\int_a^b f(x)\phi_n(x)\mathrm{d}x.$$

Hence the solution $y(x)$ is given by

$$y(x) = \sum_{n=1}^{\infty} c_n\phi_n(x) = \sum_{n=1}^{\infty}\left[\frac{1}{k-\lambda_n}\int_a^b f(x')\phi_n(x')\mathrm{d}x'\right]\phi_n(x).$$

Since $f(x)$ is a given function, presumably this series can be computed. However, we want to put it in a somewhat different form, and introduce a conceptually important function, known as the Green's function. Assuming the summation and the integration can be interchanged, we can write the last expression as:

$$y(x) = \int_a^b f(x')\sum_{n=1}^{\infty}\frac{\phi_n(x')\phi_n(x)}{k-\lambda_n}\mathrm{d}x'.$$

Now if we define the Green's function as:

$$G(x', x) = \sum_{n=1}^{\infty}\frac{\phi_n(x')\phi_n(x)}{k-\lambda_n}, \tag{3.22}$$

then the solution $y(x)$ can be written as:

$$y(x) = \int_a^b f(x')G(x', x)\mathrm{d}x'.$$

3.5.2 Green's Function and Delta Function

To appreciate the meaning of the Green's function, we will first show that $G(x', x)$ is the solution of (3.21), except with $f(x)$ replaced by the delta function $\delta(x' - x)$. That is, we will show that

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[p(x)\frac{\mathrm{d}}{\mathrm{d}x}G(x',x)\right]+q(x)G(x',x)+kw(x)G(x',x)=\delta(x'-x), \quad (3.23)$$

where the delta function $\delta(x'-x)$ is defined by the relation

$$F(x)=\int_a^b F(x')\delta(x'-x)\mathrm{d}x', \qquad a<x<b.$$

With $G(x',x)$ given by (3.22),

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[p(x)\frac{\mathrm{d}}{\mathrm{d}x}G(x',x)\right]+q(x)G(x',x)+kw(x)G(x',x)$$
$$=\left\{\frac{\mathrm{d}}{\mathrm{d}x}\left[p(x)\frac{\mathrm{d}}{\mathrm{d}x}\right]+q(x)\right\}\sum_{n=1}^{\infty}\frac{\phi_n(x')\phi_n(x)}{k-\lambda_n}+kw(x)\sum_{n=1}^{\infty}\frac{\phi_n(x')\phi_n(x)}{k-\lambda_n}$$
$$=\sum_{n=1}^{\infty}\frac{-\lambda_n w(x)\phi_n(x')\phi_n(x)}{k-\lambda_n}+kw(x)\sum_{n=1}^{\infty}\frac{\phi_n(x')\phi_n(x)}{k-\lambda_n}=w(x)\sum_{n=1}^{\infty}\phi_n(x')\phi_n(x),$$

which can be shown as the eigenfunction expansion of the delta function. Let

$$\delta(x'-x)=\sum_{n=1}^{\infty}a_n\phi_n(x).$$

The inner product of both sides with one of the eigenfunctions shows that

$$a_n=\langle\delta(x'-x)\,|\phi_n(x)\rangle\,.$$

Therefore

$$\delta(x'-x)=\sum_{n=1}^{\infty}a_n\phi_n(x)=\sum_{n=1}^{\infty}\langle\delta(x'-x)\,|\phi_n(x)\rangle\,\phi_n(x)$$
$$=\sum_{n=1}^{\infty}\left[\int_a^b\delta(x'-x)\phi_n(x)w(x)\mathrm{d}x\right]\phi_n(x)=w(x')\sum_{n=1}^{\infty}\phi_n(x')\phi_n(x).$$

Furthermore, since $\delta(x'-x)$ is nonzero only for $x=x'$,

$$\delta(x'-x)=\delta(x-x')=w(x)\sum_{n=1}^{\infty}\phi_n(x)\phi_n(x'). \quad (3.24)$$

Equation (3.23) is thus established.

Now the Green's function can be interpreted as follows. The linear differential equation, such as (3.21), can be used to describe a linear physical system. The function $f(x)$ in the right-hand side of the equation represents the "force," or forcing function applied to the system. In other words, $f(x)$

is the input to the system. The solution $y(x)$ of the equation represents the response of the system.

The Green's function $G(x', x)$ describes the response of the physical system to a unit delta function, which represents the impulse of a point source at x' with a unit strength.

We can model any input $f(x)$ as the sum of a set of point inputs. This is expressed as

$$f(x) = \int f(x')\delta(x' - x)\mathrm{d}x'.$$

The value of $f(x')$ is simply the strength of the delta function at x'. Since $G(x', x)$ is the response of a unit delta function, if the strength of the delta function is $f(x')$ times larger, the response will also be larger by that amount. That is, the response will be $f(x')\ G(x', x)$. Since the system is linear, we can find the response of the system to the input $f(x)$ by adding up the responses of the point inputs. That is

$$y(x) = \int f(x')G(x', x)\mathrm{d}x'.$$

Example 3.5.1. (a) Determine the eigenfunction expansion of the Green's function $G(x', x)$ for

$$y'' + y = x$$

$$y(0) = 0, \quad y(1) = 0.$$

(b) Find the solution $y(x)$ of the inhomogeneous differential equation through

$$y(x) = \int_0^1 x'G(x', x)\mathrm{d}x'.$$

Solution 3.5.1. (a) To solve this inhomogeneous differential equation, let us first look at the related eigenvalue problem,

$$y'' + y + \lambda y = 0.$$
$$y(0) = 0, \quad y(1) = 0,$$

which is a regular Sturm–Liouville problem, with $p(x) = 1, \ \ q(x) = 1, \ w(x) = 1$. The solution of the equation

$$y'' = -(1 + \lambda)y$$

is

$$y(x) = A\cos\sqrt{1+\lambda}x + B\sin\sqrt{1+\lambda}x.$$

The boundary condition $y(0) = 0$ requires

$$y(0) = A = 0,$$

so

$$y(1) = B \sin \sqrt{1+\lambda}.$$

Thus the other boundary condition $y(1) = 0$ makes it necessary that

$$\sqrt{1+\lambda} = n\pi, \qquad n = 1,\ 2,\ 3,\ \ldots.$$

It follows that the eigenvalues are

$$\lambda_n = n^2\pi^2 - 1,$$

and the corresponding eigenfunctions are

$$y_n(x) = \sin n\pi x.$$

Therefore the normalized eigenfunctions are

$$\phi_n(x) = \frac{\sin n\pi x}{\left[\int_0^1 \sin^2 n\pi x \; \mathrm{d}x\right]^{1/2}} = \sqrt{2} \sin n\pi x.$$

Hence the Green's function can be written as

$$G(x', x) = \sum_{n=1}^{\infty} \frac{\phi_n(x')\phi_n(x)}{0 - \lambda_n} = 2 \sum_{n=1}^{\infty} \frac{\sin n\pi x' \sin n\pi x}{1 - n^2\pi^2}.$$

(b) The solution $y(x)$ is therefore given by

$$y(x) = \int_0^1 x' G(x', x) \mathrm{d}x' = 2 \sum_{n=1}^{\infty} \frac{\sin n\pi x}{1 - n^2\pi^2} \int_0^1 x' \sin n\pi x' \; \mathrm{d}x'.$$

Since

$$\begin{aligned}\int_0^1 x' \sin n\pi x' \; \mathrm{d}x' &= \left[x'\left(-\frac{1}{n\pi}\cos n\pi x'\right)\right]_0^1 + \frac{1}{n\pi}\int_0^1 \cos n\pi x' \; \mathrm{d}x' \\ &= -\frac{1}{n\pi}\cos n\pi = \frac{(-1)^{n+1}}{n\pi},\end{aligned}$$

the solution can be expressed as:

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

In this example, with the eigenfunction expansion of the Green's function, we have found the solution of the problem expressed in a Fourier series of sine functions. To show that the Green's function is the response of the system to a unit delta function, it is instructive to solve the same problem with a Green's function obtained directly from the equation

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2} G\left(x', x\right) + G\left(x', x\right) = \delta\left(x' - x\right). \tag{3.25}$$

This we will do in the following example.

Example 3.5.2. (a) Solve the problem of the previous example with a Green's function obtained from the fact that it is the response of the system to a unit delta function. (b) Solve the inhomogeneous differential equation of the previous example, with the Green's function obtained in (a).

Solution 3.5.2. (a) Since the Green's function is the response of the system to a delta function, we require it to be continuous and bounded in the interval of interest. For $x \neq x'$, the Green's function satisfies the equation

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2} G\left(x', x\right) + G\left(x', x\right) = 0.$$

The solution of this equation is given by

$$G(x', x) = A(x') \cos x + B(x') \sin x.$$

As far as x is concerned, $A(x')$ and $B(x')$ are two arbitrary constants. But there is no reason that these constants are the same for $x < x'$ as for $x > x'$, in fact they are not. So let us write $G(x', x)$ as

$$G(x', x) = \begin{cases} a \cos x + b \sin x & x < x', \\ c \cos x + d \sin x & x > x'. \end{cases}$$

Since the Green's function must satisfy the same boundary conditions as the original differential equation. At $x = 0$, $G(x', 0) = 0$. Since $x = 0$ is certainly less than x', therefore we require

$$G(x', 0) = a \cos 0 + b \sin 0 = a = 0.$$

Furthermore, because at $x = 1$, $G(x', 1) = 0$, we have

$$G(x', 1) = c \cos 1 + d \sin 1 = 0.$$

It follows that

$$d = -c \frac{\cos 1}{\sin 1}.$$

Thus, for $x > x'$,

$$\begin{aligned} G(x', x) &= c \cos x - c \frac{\cos 1}{\sin 1} \sin x = c \frac{1}{\sin 1} (\sin 1 \cos x - \cos 1 \sin x) \\ &= c \frac{1}{\sin 1} \sin(1 - x). \end{aligned}$$

Hence, with boundary conditions, we are left with two constants in the Green's function to be determined,

$$G(x', x) = \begin{cases} b \sin x & x < x', \\ c \frac{1}{\sin 1} \sin(1 - x) & x > x'. \end{cases}$$

To determine b and c, we invoke the condition that $G(x', x)$ must be continuous at $x = x'$, so

$$b \sin x' = c \frac{1}{\sin 1} \sin(1 - x')$$

Thus,

$$G(x', x) = \begin{cases} G^-(x', x) = b \sin x & x < x', \\ G^+(x', x) = b \frac{\sin x'}{\sin(1-x')} \sin(1 - x) & x > x'. \end{cases}$$

Next we integrate both sides of (3.25) over a small interval across x',

$$\int_{x'-\epsilon}^{x'+\epsilon} \frac{\mathrm{d}^2}{\mathrm{d}x^2} G(x', x)\mathrm{d}x + \int_{x'-\epsilon}^{x'+\epsilon} G(x', x)\mathrm{d}x = \int_{x'-\epsilon}^{x'+\epsilon} \delta(x' - x)\mathrm{d}x.$$

The integral on the right-hand side is equal to 1, by the definition of the delta function. As $\epsilon \to 0$,

$$\lim_{\epsilon \to 0} \int_{x'-\epsilon}^{x'+\epsilon} G(x', x)\mathrm{d}x = 0.$$

This integral is equal to 2ϵ times the average value of $G(x', x)$ over 2ϵ at $x = x'$. Since $G(x', x)$ is bounded, this integral is equal to zero as ϵ goes to zero. Now

$$\int_{x'-\epsilon}^{x'+\epsilon} \frac{\mathrm{d}^2}{\mathrm{d}x^2} G(x', x)\mathrm{d}x = \left. \frac{\mathrm{d}G(x', x)}{\mathrm{d}x} \right|_{x'-\epsilon}^{x'+\epsilon}.$$

It follows that as $\epsilon \to 0$,

$$\lim_{\epsilon \to 0} \int_{x'-\epsilon}^{x'+\epsilon} \frac{\mathrm{d}^2}{\mathrm{d}x^2} G(x', x)\mathrm{d}x = \left. \frac{\mathrm{d}G^+(x', x)}{\mathrm{d}x} \right|_{x=x'} - \left. \frac{\mathrm{d}G^-(x', x)}{\mathrm{d}x} \right|_{x=x'}.$$

Hence

$$-b \frac{\sin x'}{\sin(1 - x')} \cos(1 - x') - b \cos x' = 1,$$

or

$$-\frac{b}{\sin(1 - x')} \left[\sin(1 - x') \cos x' + \cos(1 - x') \sin x'\right] = 1.$$

Since

$$\left[\sin(1 - x') \cos x' + \cos(1 - x') \sin x'\right] = \sin(1 - x' + x') = \sin 1,$$

so

$$b = -\frac{\sin(1 - x')}{\sin 1}.$$

Thus, the Green's function is given by

$$G(x', x) = \begin{cases} -\frac{\sin(1-x')}{\sin 1} \sin x & x < x', \\ -\frac{\sin x'}{\sin 1} \sin(1 - x) & x > x'. \end{cases}$$

(b)

$$\begin{aligned}
y(x) &= \int_0^1 x' G(x', x)\mathrm{d}x' \\
&= -\int_0^x x' \frac{\sin x'}{\sin 1} \sin(1-x)\mathrm{d}x' - \int_x^1 x' \frac{\sin(1-x')}{\sin 1} \sin x \ \mathrm{d}x' \\
&= -\frac{\sin(1-x)}{\sin 1} \int_0^x x' \sin x' \ \mathrm{d}x' - \frac{\sin x}{\sin 1} \int_x^1 x' \sin(1-x')\mathrm{d}x'.
\end{aligned}$$

Since

$$\int_0^x x' \sin x' \ \mathrm{d}x' = \left[-x' \cos x' + \sin x'\right]_0^x = -x\cos x + \sin x,$$

$$\begin{aligned}
\int_x^1 x' \sin(1-x')\mathrm{d}x' &= \left[x' \cos(1-x') + \sin(1-x')\right]_x^1 \\
&= 1 - x\cos(1-x) - \sin(1-x),
\end{aligned}$$

so

$$\begin{aligned}
y\left(x\right) &= -\frac{1}{\sin 1}\left[-x\sin\left(1-x\right)\cos x + \sin x - x \sin x \cos\left(1-x\right)\right] \\
&= -\frac{1}{\sin 1}\left[-x \sin\left(1-x+x\right) + \sin x\right] = x - \frac{1}{\sin 1}\sin x.
\end{aligned}$$

To see if this result is the same as the solution obtained in the previously example, we can expand it in terms of the Fourier sine series in the range of $0 \leq x \leq 1$,

$$x - \frac{1}{\sin 1}\sin x = \sum_{n=1}^{\infty} a_n \sin n\pi x,$$

$$a_n = 2\int_0^1 \left(x - \frac{1}{\sin 1}\sin x\right)\sin n\pi x \ \mathrm{d}x.$$

It can be readily shown that

$$\int_0^1 x \sin n\pi x \ \mathrm{d}x = \frac{(-1)^{n+1}}{n\pi},$$

$$\begin{aligned}
\int_0^1 \sin x \sin n\pi x \ \mathrm{d}x &= \frac{1}{2}\left[\frac{1}{n\pi - 1}\sin\left(n\pi - 1\right) - \frac{1}{n\pi + 1}\sin\left(n\pi + 1\right)\right] \\
&= \frac{1}{2}\left[\frac{(-1)^{n+1}}{n\pi - 1}\sin 1 + \frac{(-1)^{n+1}}{n\pi + 1}\sin 1\right] = \frac{(-1)^{n+1} n\pi}{n^2\pi^2 - 1}\sin 1.
\end{aligned}$$

Thus

$$a_n = 2\left[\frac{(-1)^{n+1}}{n\pi} - \frac{(-1)^{n+1}n\pi}{n^2\pi^2 - 1}\right] = \frac{2(-1)^{n+1}}{n\pi\left(1 - n^2\pi^2\right)},$$

and

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

which is identical to the result of the previous example.

This problem can be simply solved by the "ordinary method." Clearly x is a particular solution, and the complementary function is $y_{\rm c} = A\cos x + B\sin x$. Applying the boundary conditions $y(0) = 0$ and $y(1) = 1$ to the solution

$$y(x) = y_{\rm p} + y_{\rm c} = x + A\cos x + B\sin x,$$

we get

$$y\left(x\right) = x - \frac{1}{\sin 1}\sin x.$$

We used this problem to illustrate how the Green's function works. For such simple problems, the Green's function may not offer any advantage, but the idea of Green's function is a powerful one in dealing with boundary conditions and introducing approximations in solving partial differential equations. We shall see these aspects of the Green's function in a later chapter.

Exercises

1. (a) Use the explicit expressions of the first six Legendre polynomials

$$P_0\left(x\right) = 1,\ P_1\left(x\right) = x,\ P_2\left(x\right) = \frac{1}{2}\left(3x^2 - 1\right),\ P_3\left(x\right) = \frac{1}{2}\left(5x^3 - 3x\right),$$
$$P_4\left(x\right) = \frac{1}{8}\left(35x^4 - 30x^2 + 3\right),\quad P_5\left(x\right) = \frac{1}{8}\left(63x^5 - 70x^3 + 15x\right),$$

to show that the conditions

$$P_n\left(1\right) = 1,$$

$$\int_{-1}^{1} P_n(x)P_m(x){\rm d}x = \frac{2}{2n+1}\delta_{nm},$$

are satisfied by $P_n(x)$ at least for $n = 0$ to $n = 5$.

(b) Show that if $y_n = P_n(x)$ for $n = 0, 1, \ldots, 5$, then

$$(1 - x^2)y_n'' - 2xy_n' + n(n+1)y_n = 0.$$

2. Express the "ramp" function $f(x)$

$$f(x) = \begin{cases} 0 & -1 \le x \le 0 \\ x & 0 \le x \le 1 \end{cases},$$

in terms of the Legendre polynomials in the interval $-1 \le x \le 1$. Find the first four nonvanishing terms explicitly.

Ans. $f(x) = \sum_{n=0}^{\infty} a_n P_n(x)$, $a_n = \dfrac{2n+1}{2} \int_0^1 x P_n(x) \mathrm{d}x$,
$f(x) = \dfrac{1}{4} P_0(x) + \dfrac{1}{2} P_1(x) + \dfrac{5}{16} P_2(x) - \dfrac{3}{32} P_4(x) + \cdots$.

3. *Laguerre polynomial.* (a) Use the Gram–Schmidt procedure to generate from the set $\{x^n\}$ $(n = 0, 1, \ldots)$ the first three polynomials $L_n(x)$ that are orthogonal over the interval $0 \le x < \infty$ with the weight function e^{-x}. Use the convention that $L_n(0) = 1$.
(b) Show, by direct integration, that

$$\int_0^{\infty} L_n(x) L_m(x) \mathrm{e}^{-x} \mathrm{d}x = \delta_{nm}.$$

(c) Show that if $y_n = L_n(x)$, then y_n satisfies the Laguerre differential equation

$$x y_n'' + (1 - x) y_n' + n y_n = 0.$$

(you may need $\int_0^{\infty} x^n \mathrm{e}^{-x} \mathrm{d}x = n!$)

Ans. $L_0(x) = 1,\ L_1(x) = 1 - x,\ L_2(x) = 1 - 2x + \frac{1}{2} x^2$.

4. *Hermite polynomial.* (a) Use the Gram–Schmidt procedure to generate from the set $\{x^n\}$ $(n = 0, 1, \ldots)$ the first three polynomials $H_n(x)$ that are orthogonal over the interval $-\infty \le x < \infty$ with the weight function e^{-x^2}. Fix the multiplicative constant by the requirement

$$\int_{-\infty}^{\infty} H_n(x) H_m(x) \mathrm{e}^{-x^2} \mathrm{d}x = \delta_{nm} n! 2^n \sqrt{\pi}.$$

(b) Show that if $y_n = H_n(x)$, then y_n satisfies the Hermite differential equation

$$y_n'' + -2x y_n' + 2n y_n = 0.$$

(you may need $\int_{-\infty}^{\infty} \mathrm{e}^{-x^2} \mathrm{d}x = \sqrt{\pi}$)

Ans. $H_0(x) = 1,\ H_1(x) = 2x,\ H_2(x) = 4x^2 - 2$.

5. *Associated laguerre equation.* (a) Express associated Laguerre's differential equation

$$x y''(x) + (k + 1 - x) y'(x) + n y(x) = 0$$

in the form of a Sturm–Liouville equation.

(b) Show that in the interval $0 \le x < \infty$, it is a singular Sturm–Liouville problem.
(c) Find the orthogonality condition of its eigenfunctions.

Ans. (a) $\left[x^{k+1}\mathrm{e}^{-x}y'(x)\right]' + nx^k\mathrm{e}^{-x}y(x) = 0$
(c) $\int_0^\infty x^k\mathrm{e}^{-x}y_n(x)y_m(x)\mathrm{d}x = 0, \quad n \ne m.$

6. *Associated laguerre polynomial.* (a) Use the Gram–Schmidt procedure to generate from the set $\{x^n\}$ $(n = 0, 1, \ldots)$ the first three polynomials $L_n^1(x)$ that are orthogonal over the interval $0 \le x < \infty$ with the weight function $x\ e^{-x}$. Fix the multiplicative constant by the requirement

$$\int_0^\infty L_n^1(x)L_m^1(x)x\ \mathrm{e}^{-x}\mathrm{d}x = \delta_{nm}.$$

(b) Show that if $y_n = L_n^1(x)$, then y_n satisfies the Associated Laguerre differential equation with $k = 1$

$$xy_n'' + (k + 1 + x)y_n' + ny_n = 0.$$

Ans. $L_0^1(x) = 1,\ L_1^1(x) = \frac{1}{\sqrt{2}}(x - 2),\ L_2^1(x) = \frac{1}{2\sqrt{3}}(x^2 - 6x + 6).$

7. *Chebyshev polynomial.* (a) Show that the Chebyshev equation

$$(1 - x^2)y''(x) - xy'(x) + \lambda y(x) = 0, \qquad -1 \le x \le 1$$

can be converted into

$$\frac{\mathrm{d}^2}{\mathrm{d}\theta^2}\Theta(\theta) + \lambda\Theta(\theta) = 0, \quad (0 \le \theta \le \pi)$$

with a change of variable $x = \cos\theta$. $(\Theta(\theta) = y(x(\theta)) = y(\cos\theta))$.
(b) Show that in terms of $\theta, dy/dx$ can be expressed as

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \left[A\sqrt{\lambda}\sin\sqrt{\lambda}\theta - B\sqrt{\lambda}\cos\sqrt{\lambda}\theta\right]\frac{1}{\sin\theta}.$$

Hint: $\Theta(\theta) = A\cos\sqrt{\lambda}\theta + B\sin\sqrt{\lambda}\theta$;

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}\Theta}{\mathrm{d}x} = \frac{\mathrm{d}\Theta}{\mathrm{d}\theta}\frac{\mathrm{d}\theta}{\mathrm{d}x}; \quad \frac{\mathrm{d}\theta}{\mathrm{d}x} = -\frac{1}{\sin\theta}.$$

(c) Show that the conditions for y and $\mathrm{d}y/\mathrm{d}x$ to be bounded are

$$B = 0, \qquad \lambda = n^2,\ n = 0, 1, 2, \ldots.$$

Therefore the eigenvalues and eigenfunctions are

$$\lambda_n = n^2, \quad \Theta_n(\theta) = \cos n\theta$$

(d) The eigenfunctions of the Chebyshev equation are known as Chebyshev polynomial, usually labeled as $T_n(x)$. Find $T_n(x)$ with the condition $T_n(1) = 1$ for $n = 0, 1, 2, 3, 4$.

Hint: $T_n(x) = y_n(x) = \Theta_n(\theta) = \cos n\theta$; $\cos 2\theta = 2\cos^2\theta - 1$, $\cos 3\theta = 4\cos^3\theta - 3\cos\theta$, $\cos 4\theta = 8\cos^4\theta - 8\cos^2\theta + 1$.
(e) Show that for any integer n and m,

$$\int_{-1}^{1} T_n(x)\, T_m(x) \frac{1}{\sqrt{1-x^2}} \mathrm{d}x = \begin{cases} 0 & n \neq m \\ \pi & n = m = 0 \\ \pi/2 & n = m \neq 0 \end{cases}.$$

Ans. (d) $T_0 = 1,\ T_1 = x,\ T_2 = 2x^2 - 1,\ T_3 = 4x^3 - 3x,\ T_4 = 8x^4 - 8x^2 + 1$.

8. *Hypergeometric equation.* Express the hypergeometric equation

$$(x - x^2)y'' + [c - (1 + a + b)x]\, y' - aby = 0$$

in a Sturm–Liouville form. For it to be a singular Sturm–Liouville problem in the range of $0 \le x \le 1$, what conditions must be imposed on a, b and c, if the weight function is required to satisfy the conditions $w(0) = 0$ and $w(1) = 0$?

Hint: Use partial fraction of $\frac{c-(1+a+b)x}{x(1-x)}$ to evaluate $\exp \int^x \frac{c-(1+a+b)x}{x(1-x)}\mathrm{d}x$.

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

9. Show that if L is a linear operator and

$$\langle h \,|Lh\rangle = \langle Lh \,|h\rangle$$

for all functions h in the complex function space, then

$$\langle f \,|Lg\rangle = \langle Lf \,|g\rangle$$

for all f and g.

Hint: First let $h = f + g$, then let $h = f + \mathrm{i}g$.

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

$$\text{(a)} \frac{\mathrm{d}}{\mathrm{d}x} + x, \quad \text{(b)}\ \frac{\mathrm{d}^2}{\mathrm{d}x^2}, \quad \text{(c)}\ -\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}x} + x^2, \quad \text{(d)}\ \mathrm{i}x\frac{\mathrm{d}}{\mathrm{d}x}.$$

Ans. (a) no, (b) yes, (c) yes, (d) no.

11. (a) Express the bounded solution of the following inhomogeneous differential equation

$$(1-x^2)y'' - 2xy' + ky = f(x), \qquad -1 \le x \le 1,$$

in terms of Legendre polynomials with the help of a Green's function.
(b) If $k = 14$ and $f(x) = 5x^3$, find the explicit solution.

Ans. $G(x', x) = \sum_{n=0}^{\infty} \frac{2n+1}{2} \frac{P_n(x') P_n(x)}{k - \lambda_n}$
(a) $y(x) = \sum_{n=0}^{\infty} a_n P_n(x), \quad a_n = \frac{2n+1}{2} \frac{1}{k-n(n+1)} \int_{-1}^{1} f(x') P_n(x') \, dx'$.
(b) $y(x) = (10x^3 - 5x)/4$.

12. Determine the eigenvalues and corresponding eigenfunctions for the following problems.

$$\begin{aligned}
&\text{(a) } y'' + \lambda y = 0, \qquad y(0) = 0, \quad y'(1) = 0. \\
&\text{(b) } y'' + \lambda y = 0, \qquad y'(0) = 0, \quad y'(\pi) = 0. \\
&\text{(c) } y'' + \lambda y = 0, \qquad y(0) = y(2\pi), \quad y'(0) = y'(2\pi).
\end{aligned}$$

Ans. (a) $\lambda_n = [(2n+1)\pi/2]^2, \quad n = 0, 1, 2, \ldots; \quad y_n(x) = \sin \frac{\pi}{2}(2n+1)x$.
(b) $\lambda_n = n^2, \quad n = 0, 1, 2, \ldots; \quad y_n(x) = \cos nx$.
(c) $\lambda_n = n^2, \quad n = 0, 1, 2, \ldots; \quad y_n(x) = \cos nx, \quad \sin nx$.

13. (a) Show that the following differential equation together with the boundary conditions is a Sturm–Liouville problem. What is the weight function?

$$\begin{gathered}
y'' - 2y' + \lambda y = 0, \quad 0 \le x \le 1, \\
y(0) = 0, \quad y(1) = 0.
\end{gathered}$$

(b) Determine the eigenvalues and corresponding eigenfunctions of the problem. Fix the multiplication constant by the requirement

$$\int_0^1 y_n(x) y_m(x) w(x) \mathrm{d}x = \frac{1}{2} \delta_{nm}.$$

Ans. (a) $\left[\mathrm{e}^{-2x} y'\right]' + \lambda \mathrm{e}^{-2x} y = 0, \quad w(x) = \mathrm{e}^{-2x}$.
(b) $\lambda_n = n^2\pi^2 + 1, \qquad y_n(x) = \mathrm{e}^x \sin n\pi x$.

14. (a) Show that if $\alpha_1, \ \alpha_2, \alpha_3, \ldots$ are positive roots of

$$\tan \alpha = \frac{h}{\alpha},$$

then $\lambda_n = \alpha_n^2$ and $y_n(x) = \cos\alpha_n x,\ n = 0, 1, 2, \ldots$ are the eigenvalues and eigenfunctions of the following Sturm–Liouville problem:

$$y'' + \lambda y = 0, \qquad 0 \le x \le 1$$
$$y'(0) = 0, \quad y'(1) + hy(1) = 0.$$

(b) Show that

$$\int_0^1 \cos\alpha_n x \cos\alpha_m x \, \mathrm{d}x = \beta_n^2 \ \delta_{nm}.$$

$$\beta^2 = \frac{\alpha_n^2 + h^2 + h}{2(\alpha_n^2 + h^2)}$$

Hint: $\beta^2 = \frac{1}{2} + \frac{2\sin 2\alpha_n}{4\alpha_n}$, $\quad \sin 2\alpha_n = 2\sin\alpha_n \cos\alpha_n = \dfrac{2\alpha_n h}{\alpha_n^2 + h^2}$.

15. Find the eigenfunction expansion for the solution with boundary conditions $y(0) = y(\pi) = 0$ of the inhomogeneous differential equation

$$y'' + ky = f(x),$$

where k is a constant and

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

Ans. $y\,(x) = \frac{4}{\pi} \sum_{n=\text{odd}} \frac{(-1)^{(n-1)/2}}{n^2(k-n^2)} \sin nx.$

16. (a) Find the normalized eigenfunctions $y_n(x)$ of the Hermitian operator $\mathrm{d}^2/\mathrm{d}x^2$ that satisfy the boundary conditions $y_n(0) = y_n(\pi) = 0$. Construct the Green's function of this operator $G(x', x)$.
(b) Show that the Green's function obtained from

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2} G(x'x) = \delta(x' - x)$$

is

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

(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).

Ans. (a) $y_n\,(x) = \left(\frac{2}{\pi}\right)^{1/2} \sin nx, \ \ n = 1, 2, \ldots.$
$G\,(x', x) = -\frac{2}{\pi} \sum_{n=1}^{\infty} \frac{1}{n^2} \sin nx' \sin nx.$