5 Non-sinusoidal Harmonics and Special Functions

"To the pure geometer the radius of curvature is an incidental characteristic - like the grin of the Cheshire cat. To the physicist it is an indispensable characteristic. It would be going too far to say that to the physicist the cat is merely incidental to the grin. Physics is concerned with interrelatedness such as the interrelatedness of cats and grins. In this case the "cat without a grin" and the "grin without a cat" are equally set aside as purely mathematical phantasies."
Sir Arthur Stanley Eddington (1882-1944)

In this chapter we provide a glimpse into generalized Fourier series in which the normal modes of oscillation are not sinusoidal. In particular, we will explore Legendre polynomials and Bessel functions which will later arise in problems having cylindrical or spherical symmetry. For vibrating strings, we saw that the harmonics were sinusoidal basis functions for a large, infinite dimensional, function space. Now, we will extend these ideas to non-sinusoidal harmonics and explore the underlying structure behind these ideas.

The background for the study of generalized Fourier series is that of function spaces. We begin by exploring the general context in which one finds oneself when discussing Fourier series and (later) Fourier transforms. We can view the sine and cosine functions in the Fourier trigonometric series representations as basis vectors in an infinite dimensional function space. A given function in that space may then be represented as a linear combination over this infinite basis. With this in mind, we might wonder

- Do we have enough basis vectors for the function space?
- Are the infinite series expansions convergent?
- For other other bases, what functions can be represented by such expansions?

In the context of the boundary value problems which typically appear in physics, one is led to the study of boundary value problems in the form of Sturm-Liouville eigenvalue problems. These lead to an appropriate set of basis vectors for the function space under consideration. We will touch a little on these ideas, leaving some of the deeper results for more advanced courses in mathematics. For now, we will turn to the ideas of functions spaces and explore some typical basis functions who origins lie deep in physical problems. The common basis functions are often referred to as special functions in physics. Examples are the classical orthogonal polynomials (Legendre, Hermite, Laguerre, Tchebychef) and Bessel functions. But first we will introduce function spaces.

5.1 *Function Spaces*

EARLIER WE STUDIED finite dimensional vector spaces. Given a set of basis vectors, $\{\mathbf{a}_k\}_{k=1}^n$, in vector space V, we showed that we can expand any vector $\mathbf{v} \in \mathbf{V}$ in terms of this basis, $\mathbf{v} = \sum_{k=1}^n v_k \mathbf{a}_k$. We then spent some time looking at the simple case of extracting the components v_k of the vector. The keys to doing this simply were to have a scalar product and an orthogonal basis set. These are also the key ingredients that we will need in the infinite dimensional case. In fact, we had already done this when we studied Fourier series.

Recall when we found Fourier trigonometric series representations of functions, we started with a function (vector?) that we wanted to expand in a set of trigonometric functions (basis?) and we sought the Fourier coefficients (components?). In this section we will extend our notions from finite dimensional spaces to infinite dimensional spaces and we will develop the needed background in which to think about more general Fourier series expansions. This conceptual framework is very important in other areas in mathematics (such as ordinary and partial differential equations) and physics (such as quantum mechanics and electrodynamics).

We note that the above determination of vector components for finite dimensional spaces is precisely what we had done to compute the Fourier coefficients using trigonometric bases. Reading further, you will see how this works.

We will consider various infinite dimensional function spaces. Functions in these spaces would differ by what properties they satisfy. For example, we could consider the space of continuous functions on [0,1], the space of differentiably continuous functions, or the set of functions integrable from a to b. As you can see that there are many types of function spaces . In order to view these spaces as vector spaces, we will need to be able to add functions and multiply them by scalars in such as way that they satisfy the definition of a vector space as defined in Chapter 3.

We will also need a scalar product defined on this space of functions. There are several types of scalar products, or inner products, that we can define. For a real vector space, we define

Definition 5.1. An inner product $<,>$ on a real vector space V is a mapping from $V \times V$ into R such that for $u, v, w \in V$ and $\alpha \in R$ one has

1. $< v, v > \geq 0$ and $< v, v >= 0$ iff $v = 0$.
2. $< v, w > = < w, v >$.
3. $< \alpha v, w > = \alpha < v, w >$.
4. $< u + v, w > = < u, w > + < v, w >$.

A real vector space equipped with the above inner product leads to what is called a real inner product space. A more general definition with the third property replaced with $< v, w > = \overline{< w, v >}$ is needed for complex inner product spaces.

For the time being, we will only deal with real valued functions and, thus, we will need an inner product appropriate for such spaces. One such definition is the following. Let $f(x)$ and $g(x)$ be functions defined on $[a, b]$ and introduce the *weight function* $\sigma(x) > 0$. Then, we define the inner product, if the integral exists, as

$$< f, g > = \int_a^b f(x) g(x) \sigma(x)\, dx. \tag{5.1}$$

Spaces in which $< f, f > < \infty$ under this inner product are called the space of square integrable functions on (a, b) under weight σ and denoted as $L^2_\sigma(a, b)$. In what follows, we will assume for simplicity that $\sigma(x) = 1$. This is possible to do by using a change of variables.

The space of square integrable functions.

Now that we have functions spaces equipped with an inner product, we seek a basis for the space? For an n-dimensional space we need n basis vectors. For an infinite dimensional space, how many will we need? How do we know when we have enough? We will provide some answers to these questions later.

Let's assume that we have a basis of functions $\{\phi_n(x)\}_{n=1}^\infty$. Given a function $f(x)$, how can we go about finding the components of f in this basis? In other words, let

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

How do we find the c_n's? Does this remind you of the problem we had earlier for finite dimensional spaces? [You may want to review the discussion at the end of Section 3.1 as you read the next derivation.]

Formally, we take the inner product of f with each ϕ_j and use the properties of the inner product to find

$$\begin{aligned} < \phi_j, f > &= < \phi_j, \sum_{n=1}^\infty c_n \phi_n > \\ &= \sum_{n=1}^\infty c_n < \phi_j, \phi_n > . \end{aligned} \tag{5.2}$$

If the basis is an orthogonal basis, then we have

$$< \phi_j, \phi_n > = N_j \delta_{jn}, \tag{5.3}$$

where δ_{jn} is the Kronecker delta. Recall from Chapter 3 that the Kronecker delta is defined as

$$\delta_{ij} = \begin{cases} 0, & i \neq j \\ 1, & i = j. \end{cases} \tag{5.4}$$

Continuing with the derivation, we have

For the generalized Fourier series expansion $f(x) = \sum_{n=1}^{\infty} c_n \phi_n(x)$, we have determined the generalized Fourier coefficients to be $c_j = \frac{<\phi_j, f>}{<\phi_j, \phi_j>}$.

$$\begin{aligned} < \phi_j, f > &= \sum_{n=1}^{\infty} c_n < \phi_j, \phi_n > \\ &= \sum_{n=1}^{\infty} c_n N_j \delta_{jn} \\ &= c_1 N_j \delta_{j1} + c_2 N_j \delta_{j2} + \ldots + c_j N_j \delta_{jj} + \ldots \\ &= c_j N_j. \end{aligned} \tag{5.5}$$

So, the expansion coefficients are

$$c_j = \frac{< \phi_j, f >}{N_j} = \frac{< \phi_j, f >}{< \phi_j, \phi_j >} \quad j = 1, 2, \ldots.$$

We summarize this important result:

Generalized Basis Expansion

Let $f(x)$ be represented by an expansion over a basis of orthogonal functions, $\{\phi_n(x)\}_{n=1}^{\infty}$,

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

Then, the expansion coefficients are formally determined as

$$c_n = \frac{< \phi_n, f >}{< \phi_n, \phi_n >}.$$

This will be referred to as the general Fourier series expansion and the c_j's are called the Fourier coefficients. Technically, equality only holds when the infinite series converges to the given function on the interval of interest.

Example 5.1. *Find the coefficients of the Fourier sine series expansion of $f(x)$, given by*

$$f(x) = \sum_{n=1}^{\infty} b_n \sin nx, \quad x \in [-\pi, \pi].$$

In the last chapter we already established that the set of functions $\phi_n(x) = \sin nx$ for $n = 1, 2, \ldots$ is orthogonal on the interval $[-\pi, \pi]$. Recall that using trigonometric identities, we have for $n \neq m$

$$< \phi_n, \phi_m > \quad = \quad \int_{-\pi}^{\pi} \sin nx \sin mx \, dx$$

$$\begin{aligned} &= \frac{1}{2}\int_{-\pi}^{\pi}[\cos(n-m)x - \cos(n+m)x]\,dx \\ &= \frac{1}{2}\left[\frac{\sin(n-m)x}{n-m} - \frac{\sin(n+m)x}{n+m}\right]_{-\pi}^{\pi} \\ &= 0. \end{aligned} \tag{5.6}$$

[1] *So, we have determined that the set* $\phi_n(x) = \sin nx$ *for* $n = 1, 2, \ldots$ *is an orthogonal set of functions on the interval* $[-\pi, \pi]$. *Just as with vectors in three dimensions, we can normalize these basis functions to arrive at an orthonormal basis. This is simply done by dividing by the length of the vector. Recall that the length of a vector is obtained as* $v = \sqrt{\mathbf{v}\cdot\mathbf{v}}$. *In the same way, we define the norm of a function by*

$$\|f\| = \sqrt{< f, f >}.$$

Note, there are many types of norms, but this induced norm will be sufficient for us.

[1] There are many types of norms. The norm defined here is the natural, or induced, norm on the inner product space. Norms are a generalization of the concept of lengths of vectors. Denoting $\|\mathbf{v}\|$ the norm of $\mathbf{v}$, it needs to satisfy the properties

1. $\|\mathbf{v}\| \geq 0$. $\|\mathbf{v}\| = 0$ if and only if $\mathbf{v} = \mathbf{0}$.
2. $\|\alpha\mathbf{v}\| = |\alpha|\|\mathbf{v}\|$.
3. $\|\mathbf{u}+\mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$.

Examples of common norms are

1. Euclidean norm: $\|\mathbf{v}\| = \sqrt{v_1^2 + \cdots + v_n^2}$.
2. Taxicab norm: $\|\mathbf{v}\| = |v_1| + \cdots + |v_n|$.
3. L^p norm: $\|f\| = \left(\int [f(x)]^p\,dx\right)^{\frac{1}{p}}$.

For the above basis of sine functions, we want to first compute the norm of each function. Then we find a new basis from the original basis such that each new basis function has unit length. Of course, this is just an orthonormal basis. We first compute

$$\begin{aligned} \|\phi_n\|^2 &= \int_{-\pi}^{\pi} \sin^2 nx\,dx \\ &= \frac{1}{2}\int_{-\pi}^{\pi}[1-\cos 2nx]\,dx \\ &= \frac{1}{2}\left[x - \frac{\sin 2nx}{2n}\right]_{-\pi}^{\pi} = \pi. \end{aligned} \tag{5.7}$$

We have found for our example that

$$< \phi_j, \phi_n > = \pi\delta_{jn} \tag{5.8}$$

and that $\|\phi_n\| = \sqrt{\pi}$. *Defining* $\psi_n(x) = \frac{1}{\sqrt{\pi}}\phi_n(x)$, *we have normalized the* ϕ_n*'s and have obtained an orthonomal basis of functions on* $[-\pi, \pi]$.

Now, we can determine the expansion coefficients using

$$b_n = \frac{< \phi_n, f >}{N_n} = \frac{< \phi_n, f >}{< \phi_n, \phi_n >} = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin nx\,dx.$$

Does this result look familiar?

5.2 *Classical Orthogonal Polynomials*

FOR COMPLETENESS, we will next discuss series representations of functions using different bases. In this section we introduce the classical orthogonal polynomials. We begin by noting that the sequence of functions $\{1, x, x^2, \ldots\}$ is a basis of linearly independent functions. In fact, by the Stone-Weierstraß Approximation Theorem[2] this set is a basis of $L^2_\sigma(a,b)$, the space of square integrable functions over the interval $[a,b]$ relative to weight $\sigma(x)$. However, we will show that the sequence of functions $\{1, x, x^2, \ldots\}$ does not provide an orthogonal basis for these spaces. We will then proceed to find an appropriate orthogonal basis of functions.

[2] **Stone-Weierstraß Approximation Theorem** Suppose f is a continuous function defined on the interval $[a,b]$. For every $\epsilon > 0$, there exists a polynomial function $P(x)$ such that for all $x \in [a,b]$, we have $|f(x) - P(x)| < \epsilon$. Therefore, every continuous function defined on $[a,b]$ can be uniformly approximated as closely as we wish by a polynomial function.

We are familiar with being able to expand functions over the basis $\{1, x, x^2, \ldots\}$, since these expansions are just power series representation of the functions,[3]

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

[3] The reader may recognize this series expansion as a Maclaurin series expansion, or Taylor series expansion about $x = 0$. For a review of Taylor series, see the Appendix.

However, this basis is not an orthogonal set of basis functions. One can easily see this by integrating the product of two even, or two odd, basis functions with $\sigma(x) = 1$ and (a,b)=$(-1,1)$. For example,

$$\int_{-1}^{1} x^0 x^2\, dx = \frac{2}{3}.$$

Since we have found that orthogonal bases have been useful in determining the coefficients for expansions of given functions, we might ask if it is possible to obtain an orthogonal basis involving powers of x. Of course, finite combinations of these basis elements are just polynomials!

OK, we will ask. “Given a set of linearly independent basis vectors, can one find an orthogonal basis of the given space?" The answer is yes. We recall from introductory linear algebra, which mostly covers finite dimensional vector spaces, that there is a method for carrying this out called the **Gram-Schmidt Orthogonalization Process.** We will review this process for finite dimensional vectors and then generalize to function spaces.

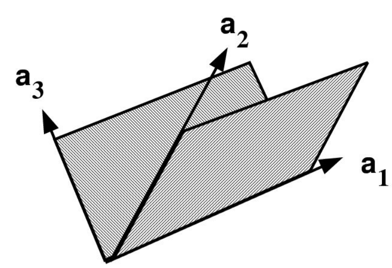

Figure 5.1: The basis $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$, of $\mathbf{R}^3$.

Let’s assume that we have three vectors that span $\mathbf{R}^3$, given by $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$ and shown in Figure 5.1. We seek an orthogonal basis $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_3$, beginning one vector at a time.

First we take one of the original basis vectors, say $\mathbf{a}_1$, and define

$$\mathbf{e}_1 = \mathbf{a}_1.$$

It is sometimes useful to normalize these basis vectors, denoting such a normalized vector with a ’hat’:

$$\hat{\mathbf{e}}_1 = \frac{\mathbf{e}_1}{e_1},$$

where $e_1 = \sqrt{\mathbf{e}_1 \cdot \mathbf{e}_1}$.

Next, we want to determine an $\mathbf{e}_2$ that is orthogonal to $\mathbf{e}_1$. We take another element of the original basis, $\mathbf{a}_2$. In Figure 5.2 we show the orientation of the vectors. Note that the desired orthogonal vector is $\mathbf{e}_2$. We can now write $\mathbf{a}_2$ as the sum of $\mathbf{e}_2$ and the projection of $\mathbf{a}_2$ on $\mathbf{e}_1$. Denoting this projection by $\mathbf{pr}_1\mathbf{a}_2$, we then have

$$\mathbf{e}_2 = \mathbf{a}_2 - \mathbf{pr}_1\mathbf{a}_2. \tag{5.9}$$

Recall the projection of one vector onto another from your vector calculus class.

$$\mathbf{pr}_1\mathbf{a}_2 = \frac{\mathbf{a}_2 \cdot \mathbf{e}_1}{e_1^2}\mathbf{e}_1. \tag{5.10}$$

This is easily proven by writing the projection as a vector of length $a_2 \cos\theta$ in direction $\hat{\mathbf{e}}_1$, where θ is the angle between $\mathbf{e}_1$ and $\mathbf{a}_2$. Using the definition of the dot product, $\mathbf{a} \cdot \mathbf{b} = ab\cos\theta$, the projection formula follows.

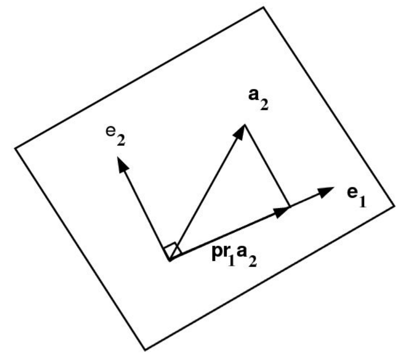

Figure 5.2: A plot of the vectors $\mathbf{e}_1$, $\mathbf{a}_2$, and $\mathbf{e}_2$ needed to find the projection of $\mathbf{a}_2$, on $\mathbf{e}_1$.

Combining Equations (5.9)-(5.10), we find that

$$\mathbf{e}_2 = \mathbf{a}_2 - \frac{\mathbf{a}_2 \cdot \mathbf{e}_1}{e_1^2}\mathbf{e}_1. \tag{5.11}$$

It is a simple matter to verify that $\mathbf{e}_2$ is orthogonal to $\mathbf{e}_1$:

$$\begin{aligned} \mathbf{e}_2 \cdot \mathbf{e}_1 &= \mathbf{a}_2 \cdot \mathbf{e}_1 - \frac{\mathbf{a}_2 \cdot \mathbf{e}_1}{e_1^2}\mathbf{e}_1 \cdot \mathbf{e}_1 \\ &= \mathbf{a}_2 \cdot \mathbf{e}_1 - \mathbf{a}_2 \cdot \mathbf{e}_1 = 0. \end{aligned} \tag{5.12}$$

Next, we seek a third vector $\mathbf{e}_3$ that is orthogonal to both $\mathbf{e}_1$ and $\mathbf{e}_2$. Pictorially, we can write the given vector $\mathbf{a}_3$ as a combination of vector projections along $\mathbf{e}_1$ and $\mathbf{e}_2$ with the new vector. This is shown in Figure 5.3. Thus, we can see that

$$\mathbf{e}_3 = \mathbf{a}_3 - \frac{\mathbf{a}_3 \cdot \mathbf{e}_1}{e_1^2}\mathbf{e}_1 - \frac{\mathbf{a}_3 \cdot \mathbf{e}_2}{e_2^2}\mathbf{e}_2. \tag{5.13}$$

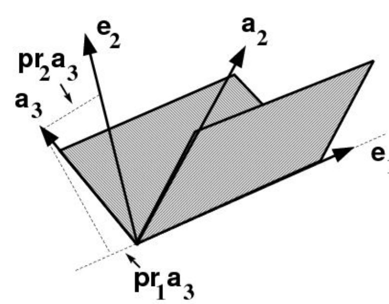

Figure 5.3: A plot of vectors for determining $\mathbf{e}_3$.

Again, it is a simple matter to compute the scalar products with $\mathbf{e}_1$ and $\mathbf{e}_2$ to verify orthogonality.

We can easily generalize this procedure to the N-dimensional case. Let $\mathbf{a}_n$, $n = 1, ..., N$ be a set of linearly independent vectors in $\mathbf{R}^N$. Then, an orthogonal basis can be found by setting $\mathbf{e}_1 = \mathbf{a}_1$ and for $n > 1$,

$$\boxed{\mathbf{e}_n = \mathbf{a}_n - \sum_{j=1}^{n-1} \frac{\mathbf{a}_n \cdot \mathbf{e}_j}{e_j^2}\mathbf{e}_j.} \tag{5.14}$$

Now, we can generalize this idea to (real) function spaces. Let $f_n(x)$, $n \in N_0 = \{0, 1, 2, \ldots\}$, be a linearly independent sequence of continuous functions defined for $x \in [a, b]$. Then, an orthogonal basis of functions, $\phi_n(x)$, $n \in N_0$ can be found and is given by

$$\phi_0(x) = f_0(x)$$

and

$$\phi_n(x) = f_n(x) - \sum_{j=0}^{n-1} \frac{< f_n, \phi_j >}{\|\phi_j\|^2} \phi_j(x), \quad n = 1, 2, \ldots. \tag{5.15}$$

Here we are using inner products relative to weight $\sigma(x)$,

$$< f, g >= \int_a^b f(x)g(x)\sigma(x)\, dx. \tag{5.16}$$

Note the similarity between the orthogonal basis in (5.15) and the expression for the finite dimensional case in Equation (5.14).

Example 5.2. *Apply the Gram-Schmidt Orthogonalization process to the set $f_n(x) = x^n$, $n \in N_0$, when $x \in (-1, 1)$ and $\sigma(x) = 1$.*

First, we have $\phi_0(x) = f_0(x) = 1$. Note that

$$\int_{-1}^{1} \phi_0^2(x)\, dx = 2.$$

We could use this result to fix the normalization of our new basis, but we will hold off on doing that for now.

Now, we compute the second basis element:

$$\begin{aligned} \phi_1(x) &= f_1(x) - \frac{< f_1, \phi_0 >}{\|\phi_0\|^2} \phi_0(x) \\ &= x - \frac{< x, 1 >}{\|1\|^2} 1 = x, \end{aligned} \tag{5.17}$$

since $< x, 1 >$ is the integral of an odd function over a symmetric interval.

For $\phi_2(x)$, we have

$$\begin{aligned} \phi_2(x) &= f_2(x) - \frac{< f_2, \phi_0 >}{\|\phi_0\|^2} \phi_0(x) - \frac{< f_2, \phi_1 >}{\|\phi_1\|^2} \phi_1(x) \\ &= x^2 - \frac{< x^2, 1 >}{\|1\|^2} 1 - \frac{< x^2, x >}{\|x\|^2} x \\ &= x^2 - \frac{\int_{-1}^{1} x^2\, dx}{\int_{-1}^{1} dx} \\ &= x^2 - \frac{1}{3}. \end{aligned} \tag{5.18}$$

So far, we have the orthogonal set $\{1, x, x^2 - \frac{1}{3}\}$. If one chooses to normalize these by forcing $\phi_n(1) = 1$, then one obtains the classical Legendre[4] polynomials, $P_n(x)$. Thus,

$$P_2(x) = \frac{1}{2}(3x^2 - 1).$$

Note that this normalization is different than the usual one. In fact, we see the $P_2(x)$ does not have a unit norm,

$$\|P_2\|^2 = \int_{-1}^{1} P_2^2(x)\, dx = \frac{2}{5}.$$

[4] Adrien-Marie Legendre (1752-1833) was a French mathematician who made many contributions to analysis and algebra.

The set of Legendre polynomials is just one set of classical orthogonal polynomials that can be obtained in this way. Many of these special functions had originally appeared as solutions of important boundary value problems in physics. They all have similar properties and we will just elaborate some of these for the Legendre functions in the next section. Others in this group are shown in Table 5.1.

Table 5.1: Common classical orthogonal polynomials with the interval and weight function used to define them.

Polynomial	Symbol	Interval	$\sigma(x)$
Hermite	$H_n(x)$	$(-\infty, \infty)$	e^{-x^2}
Laguerre	$L_n^\alpha(x)$	$[0, \infty)$	e^{-x}
Legendre	$P_n(x)$	(-1,1)	1
Gegenbauer	$C_n^\lambda(x)$	(-1,1)	$(1-x^2)^{\lambda-1/2}$
Tchebychef of the 1st kind	$T_n(x)$	(-1,1)	$(1-x^2)^{-1/2}$
Tchebychef of the 2nd kind	$U_n(x)$	(-1,1)	$(1-x^2)^{-1/2}$
Jacobi	$P_n^{(\nu,\mu)}(x)$	(-1,1)	$(1-x)^\nu(1-x)^\mu$

5.3 *Fourier-Legendre Series*

In the last chapter we saw how useful Fourier series expansions were for solving the heat and wave equations. In Chapter 9 we will investigate partial differential equations in higher dimensions and find that problems with spherical symmetry may lead to the series representations in terms of a basis of Legendre polynomials. For example, we could consider the steady state temperature distribution inside a hemispherical igloo, which takes the form

$$\phi(r,\theta) = \sum_{n=0}^{\infty} A_n r^n P_n(\cos\theta)$$

in spherical coordinates. Evaluating this function at the surface $r = a$ as $\phi(a,\theta) = f(\theta)$, leads to a Fourier-Legendre series expansion of function f :

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

where $c_n = A_n a^n$

In this section we would like to explore Fourier-Legendre series expansions of functions $f(x)$ defined on $(-1,1)$:

$$f(x) \sim \sum_{n=0}^{\infty} c_n P_n(x). \tag{5.19}$$

As with Fourier trigonometric series, we can determine the expansion coefficients by multiplying both sides of Equation (5.19) by $P_m(x)$ and integrating for $x \in [-1,1]$. Orthogonality gives the usual form for the generalized Fourier coefficients,

$$c_n = \frac{< f, P_n >}{\|P_n\|^2}, n = 0, 1, \ldots .$$

We will later show that

$$\|P_n\|^2 = \frac{2}{2n+1}.$$

Therefore, the Fourier-Legendre coefficients are

$$c_n = \frac{2n+1}{2} \int_{-1}^{1} f(x) P_n(x)\, dx. \tag{5.20}$$

Rodrigues Formula

We can do examples of Fourier-Legendre expansions given just a few facts about Legendre polynomials. The first property that the Legendre polynomials have is the *Rodrigues formula*:

$$\boxed{P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n}(x^2-1)^n, \quad n \in N_0.} \tag{5.21}$$

From the Rodrigues formula, one can show that $P_n(x)$ is an nth degree polynomial. Also, for n odd, the polynomial is an odd function and for n even, the polynomial is an even function.

Example 5.3. *Determine $P_2(x)$ from Rodrigues formula:*

$$\begin{aligned} P_2(x) &= \frac{1}{2^2 2!} \frac{d^2}{dx^2}(x^2-1)^2 \\ &= \frac{1}{8} \frac{d^2}{dx^2}(x^4 - 2x^2 + 1) \\ &= \frac{1}{8} \frac{d}{dx}(4x^3 - 4x) \\ &= \frac{1}{8}(12x^2 - 4) \\ &= \frac{1}{2}(3x^2 - 1). \end{aligned} \tag{5.22}$$

Note that we get the same result as we found in the last section using orthogonalization.

The first several Legendre polynomials are given in Table 5.2. In Figure 5.4 we show plots of these Legendre polynomials.

Three Term Recursion Formula

All of the classical orthogonal polynomials satisfy a *three term recursion formula* (or, recurrence relation or formula). In the case of the Legendre polynomials, we have

$$(n+1)P_{n+1}(x) = (2n+1)xP_n(x) - nP_{n-1}(x), \quad n = 1, 2, \ldots . \tag{5.23}$$

n	$(x^2-1)^n$	$\frac{d^n}{dx^n}(x^2-1)^n$	$\frac{1}{2^n n!}$	$P_n(x)$
0	1	1	1	1
1	x^2-1	$2x$	$\frac{1}{2}$	x
2	x^4-2x^2+1	$12x^2-4$	$\frac{1}{8}$	$\frac{1}{2}(3x^2-1)$
3	$x^6-3x^4+3x^2-1$	$120x^3-72x$	$\frac{1}{48}$	$\frac{1}{2}(5x^3-3x)$

Table 5.2: Tabular computation of the Legendre polynomials using the Rodrigues formula.

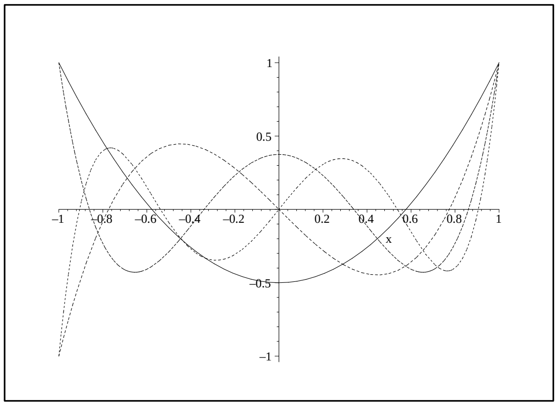

Figure 5.4: Plots of the Legendre polynomials $P_2(x)$, $P_3(x)$, $P_4(x)$, and $P_5(x)$.

This can also be rewritten by replacing n with $n-1$ as

$$(2n-1)xP_{n-1}(x) = nP_n(x) + (n-1)P_{n-2}(x), \quad n = 1, 2, \ldots. \tag{5.24}$$

Example 5.4. *Use the recursion formula to find $P_2(x)$ and $P_3(x)$, given that $P_0(x) = 1$ and $P_1(x) = x$.*

We first begin by inserting $n = 1$ into Equation (5.23):

$$2P_2(x) = 3xP_1(x) - P_0(x) = 3x^2 - 1.$$

So, $P_2(x) = \frac{1}{2}(3x^2-1)$.

For $n = 2$, we have

$$\begin{aligned} 3P_3(x) &= 5xP_2(x) - 2P_1(x) \\ &= \frac{5}{2}x(3x^2-1) - 2x \\ &= \frac{1}{2}(15x^3 - 9x). \end{aligned} \tag{5.25}$$

This gives $P_3(x) = \frac{1}{2}(5x^3 - 3x)$. These expressions agree with the earlier results.

We will prove the three term recursion formula in two ways. First we use the orthogonality properties of Legendre polynomials and the following lemma.

The first proof of the three term recursion formula is based upon the nature of the Legendre polynomials as an orthogonal basis, while the second proof is derived using generating functions.

Lemma 5.1. *The leading coefficient of x^n in $P_n(x)$ is $\frac{1}{2^n n!}\frac{(2n)!}{n!}$.*

Proof. We can prove this using Rodrigues formula. first, we focus on the leading coefficient of $(x^2-1)^n$, which is x^{2n}. The first derivative of x^{2n} is $2nx^{2n-1}$. The second derivative is $2n(2n-1)x^{2n-2}$. The jth derivative is

$$\frac{d^j x^{2n}}{dx^j} = [2n(2n-1)\ldots(2n-j+1)]x^{2n-j}.$$

Thus, the nth derivative is given by

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

This proves that $P_n(x)$ has degree n. The leading coefficient of $P_n(x)$ can now be written as

$$\begin{aligned}\frac{1}{2^n n!}[2n(2n-1)\ldots(n+1)] &= \frac{1}{2^n n!}[2n(2n-1)\ldots(n+1)]\frac{n(n-1)\ldots 1}{n(n-1)\ldots 1}\\ &= \frac{1}{2^n n!}\frac{(2n)!}{n!}. \qquad (5.26)\end{aligned}$$

□

Theorem 5.1. *Legendre polynomials satisfy the three term recursion formula*

$$(2n-1)xP_{n-1}(x) = nP_n(x) + (n-1)P_{n-2}(x), \quad n = 1, 2, \ldots. \qquad (5.27)$$

Proof. In order to prove the three term recursion formula we consider the expression $(2n-1)xP_{n-1}(x) - nP_n(x)$. While each term is a polynomial of degree n, the leading order terms cancel. We need only look at the coefficient of the leading order term first expression. It is

$$(2n-1)\frac{1}{2^{n-1}(n-1)!}\frac{(2n-2)!}{(n-1)!} = \frac{1}{2^{n-1}(n-1)!}\frac{(2n-1)!}{(n-1)!} = \frac{(2n-1)!}{2^{n-1}\left[(n-1)!\right]^2}.$$

The coefficient of the leading term for $nP_n(x)$ can be written as

$$n\frac{1}{2^n n!}\frac{(2n)!}{n!} = n\left(\frac{2n}{2n^2}\right)\left(\frac{1}{2^{n-1}(n-1)!}\right)\frac{(2n-1)!}{(n-1)!}\frac{(2n-1)!}{2^{n-1}\left[(n-1)!\right]^2}.$$

It is easy to see that the leading order terms in $(2n-1)xP_{n-1}(x) - nP_n(x)$ cancel.

The next terms will be of degree $n-2$. This is because the P_n's are either even or odd functions, thus only containing even, or odd, powers of x. We conclude that

$$(2n-1)xP_{n-1}(x) - nP_n(x) = \text{ polynomial of degree } n-2.$$

Therefore, since the Legendre polynomials form a basis, we can write this polynomial as a linear combination of of Legendre polynomials:

$$(2n-1)xP_{n-1}(x) - nP_n(x) = c_0P_0(x) + c_1P_1(x) + \ldots + c_{n-2}P_{n-2}(x). \qquad (5.28)$$

Multiplying Equation (5.28) by $P_m(x)$ for $m = 0, 1, \ldots, n-3$, integrating from -1 to 1, and using orthogonality, we obtain

$$0 = c_m \|P_m\|^2, \quad m = 0, 1, \ldots, n-3.$$

[Note: $\int_{-1}^{1} x^k P_n(x)\, dx = 0$ for $k \leq n-1$. Thus, $\int_{-1}^{1} x P_{n-1}(x) P_m(x)\, dx = 0$ for $m \leq n-3$.]

Thus, all of these c_m's are zero, leaving Equation (5.28) as

$$(2n-1)xP_{n-1}(x) - nP_n(x) = c_{n-2}P_{n-2}(x).$$

The final coefficient can be found by using the normalization condition, $P_n(1) = 1$. Thus, $c_{n-2} = (2n-1) - n = n-1$. □

Generating Functions

A second proof of the three term recursion formula can be obtained from the *generating function* of the Legendre polynomials. Many special functions have such generating functions. In this case it is given by

$$g(x,t) = \frac{1}{\sqrt{1-2xt+t^2}} = \sum_{n=0}^{\infty} P_n(x)t^n, \quad |x| \leq 1, |t| < 1. \tag{5.29}$$

This generating function occurs often in applications. In particular, it arises in potential theory, such as electromagnetic or gravitational potentials. These potential functions are $\frac{1}{r}$ type functions. For example, the gravitational potential between the Earth and the moon is proportional to the reciprocal of the magnitude of the difference between their positions relative to some coordinate system. An even better example, would be to place the origin at the center of the Earth and consider the forces on the non-pointlike Earth due to the moon. Consider a piece of the Earth at position $\mathbf{r}_1$ and the moon at position $\mathbf{r}_2$ as shown in Figure 5.5. The tidal potential Φ is proportional to

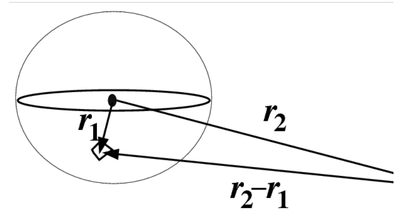

Figure 5.5: The position vectors used to describe the tidal force on the Earth due to the moon.

$$\Phi \propto \frac{1}{|\mathbf{r}_2 - \mathbf{r}_1|} = \frac{1}{\sqrt{(\mathbf{r}_2 - \mathbf{r}_1)\cdot(\mathbf{r}_2 - \mathbf{r}_1)}} = \frac{1}{\sqrt{r_1^2 - 2r_1r_2\cos\theta + r_2^2}},$$

where θ is the angle between $\mathbf{r}_1$ and $\mathbf{r}_2$.

Typically, one of the position vectors is much larger than the other. Let's assume that $r_1 \ll r_2$. Then, one can write

$$\Phi \propto \frac{1}{\sqrt{r_1^2 - 2r_1r_2\cos\theta + r_2^2}} = \frac{1}{r_2}\frac{1}{\sqrt{1 - 2\frac{r_1}{r_2}\cos\theta + \left(\frac{r_1}{r_2}\right)^2}}.$$

Now, define $x = \cos\theta$ and $t = \frac{r_1}{r_2}$. We then have that the tidal potential is proportional to the generating function for the Legendre polynomials! So, we can write the tidal potential as

$$\Phi \propto \frac{1}{r_2}\sum_{n=0}^{\infty} P_n(\cos\theta)\left(\frac{r_1}{r_2}\right)^n.$$

The first term in the expansion, $\frac{1}{r_2}$, is the gravitational potential that gives the usual force between the Earth and the moon. [Recall that the gravitational potential for mass m at distance r from M is given by $\Phi = -\frac{GMm}{r}$ and that the force is the gradient of the potential, $\mathbf{F} = -\nabla\Phi \propto \nabla\left(\frac{1}{r}\right)$.] The next terms will give expressions for the tidal effects.

Now that we have some idea as to where this generating function might have originated, we can proceed to use it. First of all, the generating function can be used to obtain special values of the Legendre polynomials.

Example 5.5. *Evaluate $P_n(0)$ using the generating function. $P_n(0)$ is found by considering $g(0,t)$. Setting $x = 0$ in Equation (5.29), we have*

$$\begin{aligned} g(0,t) &= \frac{1}{\sqrt{1+t^2}} \\ &= \sum_{n=0}^{\infty} P_n(0)t^n \\ &= P_0(0) + P_1(0)t + P_2(0)t^2 + P_3(0)t^3 + \dots . \end{aligned} \tag{5.30}$$

We can use the binomial expansion to find the final answer. [See Chapter 1 for a review of the binomial expansion.] Namely, we have

$$\frac{1}{\sqrt{1+t^2}} = 1 - \frac{1}{2}t^2 + \frac{3}{8}t^4 + \dots .$$

Comparing these expansions, we have the $P_n(0) = 0$ for n odd and for even integers one can show (see Problem 12) that[5]

$$P_{2n}(0) = (-1)^n \frac{(2n-1)!!}{(2n)!!}, \tag{5.31}$$

where $n!!$ is the double factorial,

$$n!! = \begin{cases} n(n-2)\dots(3)1, & n > 0, odd, \\ n(n-2)\dots(4)2, & n > 0, even, \\ 1 & n = 0, -1 \end{cases} .$$

[5] This example can be finished by first proving that

$$(2n)!! = 2^n n!$$

and

$$(2n-1)!! = \frac{(2n)!}{(2n)!!} = \frac{(2n)!}{2^n n!}.$$

Example 5.6. *Evaluate $P_n(-1)$. This is a simpler problem. In this case we have*

$$g(-1,t) = \frac{1}{\sqrt{1+2t+t^2}} = \frac{1}{1+t} = 1 - t + t^2 - t^3 + \dots .$$

Therefore, $P_n(-1) = (-1)^n$.

Second proof of the three term recursion formula.

Proof of the three term recursion formula using $\frac{\partial g}{\partial t}$.

Proof. We can also use the generating function to find recurrence relations. To prove the three term recursion (5.23) that we introduced above, then we need only differentiate the generating function with respect to t in Equation (5.29) and rearrange the result. First note that

$$\frac{\partial g}{\partial t} = \frac{x-t}{(1-2xt+t^2)^{3/2}} = \frac{x-t}{1-2xt+t^2} g(x,t).$$

Combining this with

$$\frac{\partial g}{\partial t} = \sum_{n=0}^{\infty} nP_n(x)t^{n-1},$$

we have

$$(x-t)g(x,t) = (1-2xt+t^2) \sum_{n=0}^{\infty} nP_n(x)t^{n-1}.$$

Inserting the series expression for $g(x,t)$ and distributing the sum on the right side, we obtain

$$(x-t)\sum_{n=0}^{\infty} P_n(x)t^n = \sum_{n=0}^{\infty} nP_n(x)t^{n-1} - \sum_{n=0}^{\infty} 2nxP_n(x)t^n + \sum_{n=0}^{\infty} nP_n(x)t^{n+1}.$$

Multiplying out the $x-t$ factor and rearranging, leads to three separate sums:

$$\sum_{n=0}^{\infty} nP_n(x)t^{n-1} - \sum_{n=0}^{\infty}(2n+1)xP_n(x)t^n + \sum_{n=0}^{\infty}(n+1)P_n(x)t^{n+1} = 0. \tag{5.32}$$

Each term contains powers of t that we would like to combine into a single sum. This is done by reindexing. For the first sum, we could use the new index $k = n-1$. Then, the first sum can be written

$$\sum_{n=0}^{\infty} nP_n(x)t^{n-1} = \sum_{k=-1}^{\infty}(k+1)P_{k+1}(x)t^k.$$

Using different indices is just another way of writing out the terms. Note that

$$\sum_{n=0}^{\infty} nP_n(x)t^{n-1} = 0 + P_1(x) + 2P_2(x)t + 3P_3(x)t^2 + \dots$$

and

$$\sum_{k=-1}^{\infty}(k+1)P_{k+1}(x)t^k = 0 + P_1(x) + 2P_2(x)t + 3P_3(x)t^2 + \dots$$

actually give the same sum. The indices are sometimes referred to as *dummy indices* because they do not show up in the expanded expression and can be replaced with another letter.

If we want to do so, we could now replace all of the k's with n's. However, we will leave the k's in the first term and now reindex the next sums in Equation (5.32). The second sum just needs the replacement $n = k$ and the last sum we reindex using $k = n + 1$. Therefore, Equation (5.32) becomes

$$\sum_{k=-1}^{\infty}(k+1)P_{k+1}(x)t^k - \sum_{k=0}^{\infty}(2k+1)xP_k(x)t^k + \sum_{k=1}^{\infty}kP_{k-1}(x)t^k = 0. \tag{5.33}$$

We can now combine all of the terms, noting the $k = -1$ term is automatically zero and the $k = 0$ terms give

$$P_1(x) - xP_0(x) = 0. \tag{5.34}$$

Of course, we know this already. So, that leaves the $k > 0$ terms:

$$\sum_{k=1}^{\infty}\left[(k+1)P_{k+1}(x) - (2k+1)xP_k(x) + kP_{k-1}(x)\right]t^k = 0. \tag{5.35}$$

Since this is true for all t, the coefficients of the t^k's are zero, or

$$(k+1)P_{k+1}(x) - (2k+1)xP_k(x) + kP_{k-1}(x) = 0, \quad k = 1, 2, \ldots.$$

While this is the standard form for the three term recurrence relation, the earlier form is obtained by setting $k = n - 1$. □

There are other recursion relations which we list in the box below. Equation (5.36) was derived using the generating function. Differentiating it with respect to x, we find Equation (5.37). Equation (5.38) can be proven using the generating function by differentiating $g(x,t)$ with respect to x and rearranging the resulting infinite series just as in this last manipulation. This will be left as Problem 4. Combining this result with Equation (5.36), we can derive Equations (5.39)-(5.40). Adding and subtracting these equations yields Equations (5.41)-(5.42).

Recursion Formulae for Legendre Polynomials for $n = 1, 2, \ldots$.

$$\begin{aligned}
(n+1)P_{n+1}(x) &= (2n+1)xP_n(x) - nP_{n-1}(x) && (5.36)\\
(n+1)P'_{n+1}(x) &= (2n+1)[P_n(x) + xP'_n(x)] - nP'_{n-1}(x) && (5.37)\\
P_n(x) &= P'_{n+1}(x) - 2xP'_n(x) + P'_{n-1}(x) && (5.38)\\
P'_{n-1}(x) &= xP'_n(x) - nP_n(x) && (5.39)\\
P'_{n+1}(x) &= xP'_n(x) + (n+1)P_n(x) && (5.40)\\
P'_{n+1}(x) + P'_{n-1}(x) &= 2xP'_n(x) + P_n(x). && (5.41)\\
P'_{n+1}(x) - P'_{n-1}(x) &= (2n+1)P_n(x). && (5.42)\\
(x^2-1)P'_n(x) &= nxP_n(x) - nP_{n-1}(x) && (5.43)
\end{aligned}$$

Finally, Equation (5.43) can be obtained using Equations (5.39) and (5.40). Just multiply Equation (5.39) by x,

$$x^2 P_n'(x) - nxP_n(x) = xP_{n-1}'(x).$$

Now use Equation (5.40), but first replace n with $n-1$ to eliminate the $xP_{n-1}'(x)$ term:

$$x^2 P_n'(x) - nxP_n(x) = P_n'(x) - nP_{n-1}(x).$$

Rearranging gives the result.

Legendre Polynomials as Solutions of a Differential Equation

The Legendre polynomials satisfy a second order linear differential equation. This differential equation occurs naturally in the solution of initial-boundary value problems in three dimensions which possess some spherical symmetry. We will see this in the last chapter. There are two approaches we could take in showing that the Legendre polynomials satisfy a particular differential equation. Either we can write down the equations and attempt to solve it, or we could use the above properties to obtain the equation. For now, we will seek the differential equation satisfied by $P_n(x)$ using the above recursion relations.

We begin by differentiating Equation (5.43) and using Equation (5.39) to simplify:

$$\begin{aligned} \frac{d}{dx}\left((x^2-1)P_n'(x)\right) &= nP_n(x) + nxP_n'(x) - nP_{n-1}'(x) \\ &= nP_n(x) + n^2 P_n(x) \\ &= n(n+1)P_n(x). \end{aligned} \tag{5.44}$$

Therefore, Legendre polynomials, or Legendre functions of the first kind, are solutions of the differential equation

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

As this is a linear second order differential equation, we expect two linearly independent solutions. The second solution, called the Legendre function of the second kind, is given by $Q_n(x)$ and is not well behaved at $x = \pm 1$. For example,

$$Q_0(x) = \frac{1}{2}\ln\frac{1+x}{1-x}.$$

A generalization of the Legendre equation is given by $(1-x^2)y'' - 2xy' + \left[n(n+1) - \frac{m^2}{1-x^2}\right]y = 0$. Solutions to this equation, $P_n^m(x)$ and $Q_n^m(x)$, are called the associated Legendre functions of the first and second kind.

We will not need these for physically interesting examples in this book.

Normalization Constant Another use of the generating function is to obtain the normalization constant. Namely, we want to evaluate

$$\|P_n\|^2 = \int_{-1}^{1} P_n(x)P_n(x)\,dx.$$

This can be done by first squaring the generating function in order to get the products $P_n(x)P_m(x)$, and then integrating over x.

Squaring the generating function has to be done with care, as we need to make proper use of the dummy summation index. So, we first write

$$\begin{aligned}\frac{1}{1-2xt+t^2} &= \left[\sum_{n=0}^{\infty} P_n(x)t^n\right]^2 \\ &= \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} P_n(x)P_m(x)t^{n+m}. \qquad (5.45)\end{aligned}$$

Integrating from -1 to 1 and using the orthogonality of the Legendre polynomials, we have

$$\begin{aligned}\int_{-1}^{1} \frac{dx}{1-2xt+t^2} &= \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} t^{n+m}\int_{-1}^{1} P_n(x)P_m(x)\,dx \\ &= \sum_{n=0}^{\infty} t^{2n}\int_{-1}^{1} P_n^2(x)\,dx. \qquad (5.46)\end{aligned}$$

However, one can show that[6]

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

Expanding this expression about $t=0$, we obtain[7]

$$\frac{1}{t}\ln\left(\frac{1+t}{1-t}\right) = \sum_{n=0}^{\infty}\frac{2}{2n+1}t^{2n}.$$

Comparing this result with Equation (5.46), we find that

$$\|P_n\|^2 = \int_{-1}^{1} P_n^2(x)\,dx = \frac{2}{2n+1}. \qquad (5.47)$$

Fourier-Legendre Series

With these properties of Legendre functions we are now prepared to compute the expansion coefficients for the Fourier-Legendre series representation of a given function.

Example 5.7. *Expand $f(x) = x^3$ in a Fourier-Legendre series.*

We simply need to compute

$$c_n = \frac{2n+1}{2}\int_{-1}^{1} x^3 P_n(x)\,dx. \qquad (5.48)$$

We first note that

$$\int_{-1}^{1} x^m P_n(x)\,dx = 0 \quad \text{for } m < n.$$

As a result, we will have for this example that $c_n = 0$ for $n > 3$. We could just compute $\int_{-1}^{1} x^3 P_m(x)\,dx$ for $m = 0, 1, 2, \ldots$ outright by looking up Legendre

[6] You will need the integral

$$\int \frac{dx}{a+bx} = \frac{1}{b}\ln(a+bx) + C.$$

[7] From Appendix A you will need the series expansion

$$\begin{aligned}\ln(1+x) &= \sum_{n=1}^{\infty}(-1)^{n+1}\frac{x^n}{n} \\ &= x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots.\end{aligned}$$

polynomials. But, note that x^3 *is an odd function,* $c_0 = 0$ *and* $c_2 = 0$*. This leaves us with only two coefficients to compute. We refer to Table 5.2 and find that*

$$c_1 = \frac{3}{2}\int_{-1}^{1} x^4\,dx = \frac{3}{5}$$

$$c_3 = \frac{7}{2}\int_{-1}^{1} x^3\left[\frac{1}{2}(5x^3-3x)\right]\,dx = \frac{2}{5}.$$

Thus,

$$x^3 = \frac{3}{5}P_1(x) + \frac{2}{5}P_3(x).$$

Of course, this is simple to check using Table 5.2:

$$\frac{3}{5}P_1(x) + \frac{2}{5}P_3(x) = \frac{3}{5}x + \frac{2}{5}\left[\frac{1}{2}(5x^3-3x)\right] = x^3.$$

Well, maybe we could have guessed this without doing any integration. Let's see,

$$\begin{aligned} x^3 &= c_1 x + \frac{1}{2}c_2(5x^3-3x) \\ &= (c_1 - \frac{3}{2}c_2)x + \frac{5}{2}c_2 x^3. \end{aligned} \tag{5.49}$$

Equating coefficients of like terms, we have that $c_2 = \frac{2}{5}$ *and* $c_1 = \frac{3}{2}c_2 = \frac{3}{5}$*.*

Example 5.8. *Expand the Heaviside*[8] *function in a Fourier-Legendre series.*

The Heaviside function is defined as

$$H(x) = \begin{cases} 1, & x > 0, \\ 0, & x < 0. \end{cases} \tag{5.50}$$

In this case, we cannot find the expansion coefficients without some integration. We have to compute

$$\begin{aligned} c_n &= \frac{2n+1}{2}\int_{-1}^{1} f(x)P_n(x)\,dx \\ &= \frac{2n+1}{2}\int_{0}^{1} P_n(x)\,dx. \end{aligned} \tag{5.51}$$

We can make use of the identity

$$P'_{n+1}(x) - P'_{n-1}(x) = (2n+1)P_n(x), \quad n > 1. \tag{5.52}$$

We have for $n > 0$

$$c_n = \frac{1}{2}\int_0^1 [P'_{n+1}(x) - P'_{n-1}(x)]\,dx = \frac{1}{2}[P_{n-1}(0) - P_{n+1}(0)].$$

For $n = 0$*, we have*

$$c_0 = \frac{1}{2}\int_0^1 dx = \frac{1}{2}.$$

[8] Oliver Heaviside (1850-1925) was an English mathematician, physicist and engineer who used complex analysis to study circuits and was a co-founder of vector analysis. The Heaviside function is also called the step function.

This leads to the expansion

$$f(x) \sim \frac{1}{2} + \frac{1}{2}\sum_{n=1}^{\infty}[P_{n-1}(0) - P_{n+1}(0)]P_n(x).$$

We still need to evaluate the Fourier-Legendre coefficients. Since $P_n(0) = 0$ *for* n *odd, the* c_n*'s vanish for* n *even. Letting* $n = 2k-1$*, we can re-index the sum,*

$$f(x) \sim \frac{1}{2} + \frac{1}{2}\sum_{k=1}^{\infty}[P_{2k-2}(0) - P_{2k}(0)]P_{2k-1}(x).$$

We can compute the Fourier coefficients, $c_{2k-1} = \frac{1}{2}[P_{2k-2}(0) - P_{2k}(0)]$*, using a result from Problem 12:*

$$P_{2k}(0) = (-1)^k \frac{(2k-1)!!}{(2k)!!}. \tag{5.53}$$

Namely, we have

$$\begin{aligned} c_{2k-1} &= \frac{1}{2}[P_{2k-2}(0) - P_{2k}(0)] \\ &= \frac{1}{2}\left[(-1)^{k-1}\frac{(2k-3)!!}{(2k-2)!!} - (-1)^k\frac{(2k-1)!!}{(2k)!!}\right] \\ &= -\frac{1}{2}(-1)^k\frac{(2k-3)!!}{(2k-2)!!}\left[1 + \frac{2k-1}{2k}\right] \\ &= -\frac{1}{2}(-1)^k\frac{(2k-3)!!}{(2k-2)!!}\frac{4k-1}{2k}. \end{aligned} \tag{5.54}$$

Thus, the Fourier-Legendre series expansion for the Heaviside function is given by

$$f(x) \sim \frac{1}{2} - \frac{1}{2}\sum_{n=1}^{\infty}(-1)^n\frac{(2n-3)!!}{(2n-2)!!}\frac{4n-1}{2n}P_{2n-1}(x). \tag{5.55}$$

The sum of the first 21 terms of this series are shown in Figure 5.6. We note the slow convergence to the Heaviside function. Also, we see that the Gibbs phenomenon is present due to the jump discontinuity at $x = 0$*. [See Section 4.12.]*

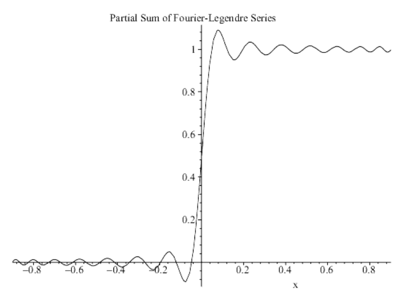

Figure 5.6: Sum of first 21 terms for Fourier-Legendre series expansion of Heaviside function.

5.4 Gamma Function

The name and symbol for the Gamma function were first given by Legendre in 1811. However, the search for a generalization of the factorial extends back to the 1720's when Euler provided the first representation of the factorial as an infinite product, later to be modified by others like Gauß, Weierstraß, and Legendre.

A FUNCTION THAT OFTEN OCCURS in the study of special functions is the Gamma function. We will need the Gamma function in the next section on Fourier-Bessel series.

For $x > 0$ we define the Gamma function as

$$\Gamma(x) = \int_0^{\infty} t^{x-1}e^{-t}\,dt, \quad x > 0. \tag{5.56}$$

The Gamma function is a generalization of the factorial function and a plot is shown in Figure 5.7. In fact, we have

$$\Gamma(1) = 1$$

and

$$\Gamma(x+1) = x\Gamma(x).$$

The reader can prove this identity by simply performing an integration by parts. (See Problem 7.) In particular, for integers $n \in Z^+$, we then have

$$\Gamma(n+1) = n\Gamma(n) = n(n-1)\Gamma(n-2) = n(n-1)\cdots 2\Gamma(1) = n!.$$

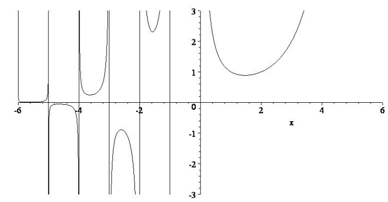

Figure 5.7: Plot of the Gamma function.

We can also define the Gamma function for negative, non-integer values of x. We first note that by iteration on $n \in Z^+$, we have

$$\Gamma(x+n) = (x+n-1)\cdots(x+1)x\Gamma(x), \quad x+n > 0.$$

Solving for $\Gamma(x)$, we then find

$$\Gamma(x) = \frac{\Gamma(x+n)}{(x+n-1)\cdots(x+1)x}, \quad -n < x < 0$$

Note that the Gamma function is undefined at zero and the negative integers.

Example 5.9. *We now prove that*

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}.$$

This is done by direct computation of the integral:

$$\Gamma\left(\frac{1}{2}\right) = \int_0^\infty t^{-\frac{1}{2}} e^{-t}\, dt.$$

Letting $t = z^2$, we have

$$\Gamma\left(\frac{1}{2}\right) = 2\int_0^\infty e^{-z^2}\, dz.$$

Due to the symmetry of the integrand, we obtain the classic integral

$$\Gamma\left(\frac{1}{2}\right) = \int_{-\infty}^\infty e^{-z^2}\, dz,$$

which can be performed using a standard trick.[9] *Consider the integral*

$$I = \int_{-\infty}^\infty e^{-x^2}\, dx.$$

Then,

$$I^2 = \int_{-\infty}^\infty e^{-x^2}\, dx \int_{-\infty}^\infty e^{-y^2}\, dy.$$

[9] In Example 7.4 we show the more general result:

$$\int_{-\infty}^\infty e^{-\beta y^2}\, dy = \sqrt{\frac{\pi}{\beta}}.$$

Note that we changed the integration variable. This will allow us to write this product of integrals as a double integral:

$$I^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-(x^2+y^2)}\, dxdy.$$

This is an integral over the entire xy-plane. We can transform this Cartesian integration to an integration over polar coordinates. The integral becomes

$$I^2 = \int_0^{2\pi}\int_0^{\infty} e^{-r^2}\, rdrd\theta.$$

This is simple to integrate and we have $I^2 = \pi$. So, the final result is found by taking the square root of both sides:

$$\Gamma\left(\frac{1}{2}\right) = I = \sqrt{\pi}.$$

In Problem 12 the reader will prove the identity

$$\Gamma(n + \frac{1}{2}) = \frac{(2n-1)!!}{2^n}\sqrt{\pi}.$$

Another useful relation, which we only state, is

$$\Gamma(x)\Gamma(1-x) = \frac{\pi}{\sin \pi x}.$$

The are many other important relations, including infinite products, which we will not need at this point. The reader is encouraged to read about these elsewhere. In the meantime, we move on to the discussion of another important special function in physics and mathematics.

5.5 *Fourier-Bessel Series*

BESSEL FUNCTIONS ARISE in many problems in physics possessing cylindrical symmetry such as the vibrations of circular drumheads and the radial modes in optical fibers. They provide us with another orthogonal set of functions. You might have seen in a course on differential equations that Bessel functions are solutions of the differential equation

$$x^2y'' + xy' + (x^2 - p^2)y = 0. \tag{5.57}$$

Solutions to this equation are obtained in the form of series expansions. Namely, one seeks solutions of the form

$$y(x) = \sum_{j=0}^{\infty} a_j x^{j+n}$$

by determining the for the coefficients must take. We will leave this for a homework exercise and simply report the results.

The history of Bessel functions, does not originate in the study of partial differential equations. These solutions originally came up in the study of the Kepler problem, describing planetary motion. According to G. N. Watson in his *Treatise on Bessel Functions*, the formulation and solution of Kepler's Problem was discovered by Joseph-Louis Lagrange (1736-1813), in 1770. Namely, the problem was to express the radial coordinate and what is called the eccentric anomaly, E, as functions of time. Lagrange found expressions for the coefficients in the expansions of r and E in trigonometric functions of time. However, he only computed the first few coefficients. In 1816 Friedrich Wilhelm Bessel (1784-1846) had shown that the coefficients in the expansion for r could be given an integral representation. In 1824 he presented a thorough study of these functions, which are now called Bessel functions.

One solution of the differential equation is the *Bessel function of the first kind of order* p, given as

$$y(x) = J_p(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{\Gamma(n+1)\Gamma(n+p+1)} \left(\frac{x}{2}\right)^{2n+p}. \qquad (5.58)$$

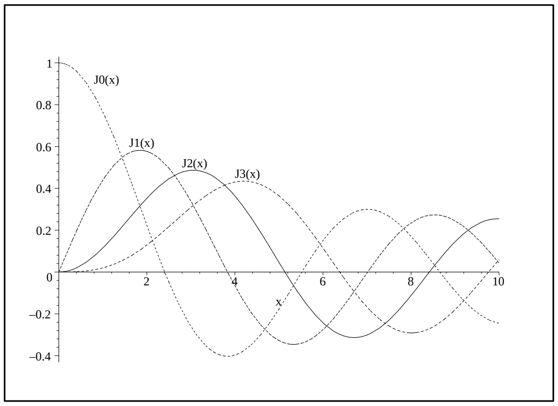

Figure 5.8: Plots of the Bessel functions $J_0(x)$, $J_1(x)$, $J_2(x)$, and $J_3(x)$.

In Figure 5.8 we display the first few Bessel functions of the first kind of integer order. Note that these functions can be described as decaying oscillatory functions.

A second linearly independent solution is obtained for p not an integer as $J_{-p}(x)$. However, for p an integer, the $\Gamma(n+p+1)$ factor leads to evaluations of the Gamma function at zero, or negative integers, when p is negative. Thus, the above series is not defined in these cases.

Another method for obtaining a second linearly independent solution is through a linear combination of $J_p(x)$ and $J_{-p}(x)$ as

$$N_p(x) = Y_p(x) = \frac{\cos \pi p J_p(x) - J_{-p}(x)}{\sin \pi p}. \qquad (5.59)$$

These functions are called the Neumann functions, or Bessel functions of the second kind of order p.

In Figure 5.9 we display the first few Bessel functions of the second kind of integer order. Note that these functions are also decaying oscillatory functions. However, they are singular at $x = 0$.

In many applications one desires bounded solutions at $x = 0$. These functions do not satisfy this boundary condition. For example, we will later study one standard problem is to describe the oscillations of a circular drumhead. For this problem one solves the two dimensional wave equation using separation of variables in cylindrical coordinates.

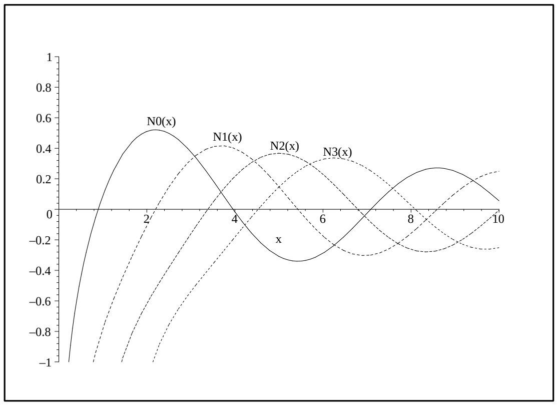

Figure 5.9: Plots of the Neumann functions $N_0(x)$, $N_1(x)$, $N_2(x)$, and $N_3(x)$.

The r equation leads to a Bessel equation. The Bessel function solutions describe the radial part of the solution and one does not expect a singular solution at the center of the drum. The amplitude of the oscillation must remain finite. Thus, only Bessel functions of the first kind can be used.

Bessel functions satisfy a variety of properties, which we will only list at this time for Bessel functions of the first kind. The reader will have the opportunity to prove these for homework.

Derivative Identities These identities follow directly from the manipulation of the series solution.

$$\frac{d}{dx}\left[x^p J_p(x)\right] = x^p J_{p-1}(x). \tag{5.60}$$

$$\frac{d}{dx}\left[x^{-p} J_p(x)\right] = -x^{-p} J_{p+1}(x). \tag{5.61}$$

Recursion Formulae The next identities follow from adding, or subtracting, the derivative identities.

$$J_{p-1}(x) + J_{p+1}(x) = \frac{2p}{x} J_p(x). \tag{5.62}$$

$$J_{p-1}(x) - J_{p+1}(x) = 2J_p'(x). \tag{5.63}$$

Orthogonality As we will see in the next chapter, one can recast the Bessel equation into an eigenvalue problem whose solutions form an orthogonal basis of functions on $L_x^2(0,a)$. Using Sturm-Liouville theory, one can show that

$$\int_0^a xJ_p(j_{pn}\frac{x}{a})J_p(j_{pm}\frac{x}{a})\,dx = \frac{a^2}{2}\left[J_{p+1}(j_{pn})\right]^2\delta_{n,m}, \qquad (5.64)$$

where j_{pn} is the nth root of $J_p(x)$, $J_p(j_{pn}) = 0$, $n = 1, 2, \ldots$. A list of some of these roots are provided in Table 5.3.

Table 5.3: The zeros of Bessel Functions

n	$p=0$	$p=1$	$p=2$	$p=3$	$p=4$	$p=5$
1	2.405	3.832	5.135	6.379	7.586	8.780
2	5.520	7.016	8.147	9.760	11.064	12.339
3	8.654	10.173	11.620	13.017	14.373	15.700
4	11.792	13.323	14.796	16.224	17.616	18.982
5	14.931	16.470	17.960	19.410	20.827	22.220
6	18.071	19.616	21.117	22.583	24.018	25.431
7	21.212	22.760	24.270	25.749	27.200	28.628
8	24.353	25.903	27.421	28.909	30.371	31.813
9	27.494	29.047	30.571	32.050	33.512	34.983

Generating Function

$$e^{x(t-\frac{1}{t})/2} = \sum_{n=-\infty}^{\infty} J_n(x)t^n, \quad x > 0, t \neq 0. \qquad (5.65)$$

Integral Representation

$$J_n(x) = \frac{1}{\pi}\int_0^{\pi} \cos(x\sin\theta - n\theta)\,d\theta, \quad x > 0, n \in \mathrm{Z}. \qquad (5.66)$$

Fourier-Bessel Series

Since the Bessel functions are an orthogonal set of functions of a Sturm-Liouville problem,[10] we can expand square integrable functions in this basis. In fact, the Sturm-Liouville problem is given in the form

$$x^2y'' + xy' + (\lambda x^2 - p^2)y = 0, \quad x \in [0, a], \qquad (5.67)$$

satisfying the boundary conditions: $y(x)$ is bounded at $x = 0$ and $y(a) = 0$. The solutions are then of the form $J_p(\sqrt{\lambda}x)$, as can be shown by making the substitution $t = \sqrt{\lambda}x$ in the differential equation. Namely, we let $y(x) = u(t)$ and note that

$$\frac{dy}{dx} = \frac{dt}{dx}\frac{du}{dt} = \sqrt{\lambda}\frac{du}{dt}.$$

Then,

$$t^2u'' + tu' + (t^2 - p^2)u = 0,$$

which has a solution $u(t) = J_p(t)$.

[10] In the study of boundary value problems in differential equations, Sturm-Liouville problems are a bountiful source of basis functions for the space of square integrable functions as will be seen in the next section.

Using Sturm-Liouville theory, one can show that $J_p(j_{pn}\frac{x}{a})$ is a basis of eigenfunctions and the resulting *Fourier-Bessel series expansion* of $f(x)$ defined on $x \in [0, a]$ is

$$f(x) = \sum_{n=1}^{\infty} c_n J_p(j_{pn}\frac{x}{a}), \tag{5.68}$$

where the Fourier-Bessel coefficients are found using the orthogonality relation as

$$c_n = \frac{2}{a^2\left[J_{p+1}(j_{pn})\right]^2}\int_0^a x f(x) J_p(j_{pn}\frac{x}{a})\,dx. \tag{5.69}$$

Example 5.10. *Expand $f(x) = 1$ for $0 < x < 1$ in a Fourier-Bessel series of the form*

$$f(x) = \sum_{n=1}^{\infty} c_n J_0(j_{0n}x)$$

.

We need only compute the Fourier-Bessel coefficients in Equation (5.69):

$$c_n = \frac{2}{[J_1(j_{0n})]^2}\int_0^1 x J_0(j_{0n}x)\,dx. \tag{5.70}$$

From the identity

$$\frac{d}{dx}\left[x^p J_p(x)\right] = x^p J_{p-1}(x). \tag{5.71}$$

we have

$$\begin{aligned}
\int_0^1 x J_0(j_{0n}x)\,dx &= \frac{1}{j_{0n}^2}\int_0^{j_{0n}} y J_0(y)\,dy \\
&= \frac{1}{j_{0n}^2}\int_0^{j_{0n}} \frac{d}{dy}\left[y J_1(y)\right]\,dy \\
&= \frac{1}{j_{0n}^2}\left[y J_1(y)\right]_0^{j_{0n}} \\
&= \frac{1}{j_{0n}} J_1(j_{0n}).
\end{aligned} \tag{5.72}$$

As a result, the desired Fourier-Bessel expansion is given as

$$1 = 2\sum_{n=1}^{\infty} \frac{J_0(j_{0n}x)}{j_{0n}J_1(j_{0n})}, \quad 0 < x < 1. \tag{5.73}$$

In Figure 5.10 we show the partial sum for the first fifty terms of this series. Note once again the slow convergence due to the Gibbs phenomenon.

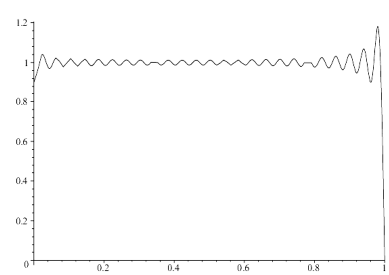

Figure 5.10: Plot of the first 50 terms of the Fourier-Bessel series in Equation (5.73) for $f(x) = 1$ on $0 < x < 1$.

5.6 Sturm-Liouville Eigenvalue Problems

In the last chapter we explored the solutions of differential equations that led to solutions in the form of trigonometric functions and special functions. Such solutions can be used to represent functions in generalized Fourier series expansions. We would like to generalize some of those techniques we had first used to solve the heat equation in order to solve other boundary value problems. A class of problems to which our previous examples belong and which have eigenfunctions with similar properties are the Sturm-Liouville Eigenvalue Problems. These problems involve self-adjoint (differential) operators which play an important role in the spectral theory of linear operators and the existence of the eigenfunctions. These ideas will be introduced in this section.

5.6.1 Sturm-Liouville Operators

In physics many problems arise in the form of boundary value problems involving second order ordinary differential equations. For example, we will explore the wave equation and the heat equation in three dimensions. Separating out the time dependence leads to a three dimensional boundary value problem in both cases. Further separation of variables leads to a set of boundary value problems involving second order ordinary differential equations.

In general, we might obtain equations of the form

$$a_2(x)y'' + a_1(x)y' + a_0(x)y = f(x) \tag{5.74}$$

subject to boundary conditions. We can write such an equation in operator form by defining the differential operator

$$L = a_2(x)\frac{d^2}{dx^2} + a_1(x)\frac{d}{dx} + a_0(x).$$

Then, Equation (5.74) takes the form

$$Ly = f.$$

Recall that we had solved such nonhomogeneous differential equations in Chapter 2. In this chapter we will show that these equations can be solved using eigenfunction expansions. Namely, we seek solutions to the eigenvalue problem

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

with homogeneous boundary conditions on ϕ and then seek a solution of the nonhomogeneous problem, $Ly = f$, as an expansion over these eigenfunctions. Formally, we let

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

However, we are not guaranteed a nice set of eigenfunctions. We need an appropriate set to form a basis in the function space. Also, it would be nice to have orthogonality so that we can easily solve for the expansion coefficients.

It turns out that any linear second order differential operator can be turned into an operator that possesses just the right properties (self-adjointedness) to carry out this procedure. The resulting operator is referred to as a Sturm-Liouville operator. We will highlight some of the properties of such operators and prove a few key theorems, though this will not be an extensive review of Sturm-Liouville theory. The interested reader can review the literature and advanced texts for a more in depth analysis.

We define the *Sturm-Liouville operator* as

The Sturm-Liouville operator.

$$\mathcal{L} = \frac{d}{dx} p(x) \frac{d}{dx} + q(x). \tag{5.75}$$

The *Sturm-Liouville eigenvalue problem* is given by the differential equation

The Sturm-Liouville eigenvalue problem.

$$\mathcal{L}y = -\lambda \sigma(x) y,$$

or

$$\frac{d}{dx}\left(p(x)\frac{dy}{dx}\right) + q(x)y + \lambda\sigma(x)y = 0, \tag{5.76}$$

for $x \in (a, b)$, $y = y(x)$, plus boundary conditions. The functions $p(x)$, $p'(x)$, $q(x)$ and $\sigma(x)$ are assumed to be continuous on (a, b) and $p(x) > 0$, $\sigma(x) > 0$ on $[a, b]$. If the interval is finite and these assumptions on the coefficients are true on $[a, b]$, then the problem is said to be *regular*. Otherwise, it is called *singular*.

We also need to impose the set of homogeneous boundary conditions

Types of boundary conditions.

$$\begin{aligned} \alpha_1 y(a) + \beta_1 y'(a) &= 0, \\ \alpha_2 y(b) + \beta_2 y'(b) &= 0. \end{aligned} \tag{5.77}$$

The α's and β's are constants. For different values, one has special types of boundary conditions. For $\beta_i = 0$, we have what are called *Dirichlet boundary conditions*. Namely, $y(a) = 0$ and $y(b) = 0$. For $\alpha_i = 0$, we have *Neumann boundary conditions*. In this case, $y'(a) = 0$ and $y'(b) = 0$. In terms of the heat equation example, Dirichlet conditions correspond to maintaining a fixed temperature at the ends of the rod. The Neumann boundary conditions would correspond to no heat flow across the ends, or insulating conditions, as there would be no temperature gradient at those points. The more general boundary conditions allow for partially insulated boundaries.

Another type of boundary condition that is often encountered is the *periodic boundary condition*. Consider the heated rod that has been bent to form a circle. Then the two end points are physically the same. So, we would expect that the temperature and the temperature gradient should agree at those points. For this case we write $y(a) = y(b)$ and $y'(a) = y'(b)$. Boundary value problems using these conditions have to be handled differently than the above homogeneous conditions. These conditions leads to different types of eigenfunctions and eigenvalues.

As previously mentioned, equations of the form (5.74) occur often. We now show that Equation (5.74) can be turned into a *differential equation of Sturm-Liouville form:*

$$\frac{d}{dx}\left(p(x)\frac{dy}{dx}\right) + q(x)y = F(x). \tag{5.78}$$

Another way to phrase this is provided in the theorem:

Theorem 5.2. *Any second order linear operator can be put into the form of the Sturm-Liouville operator (5.76).*

The proof of this is straight forward, as we shall soon show. Consider the equation (5.74). If $a_1(x) = a_2'(x)$, then we can write the equation in the form

$$\begin{aligned} f(x) &= a_2(x)y'' + a_1(x)y' + a_0(x)y \\ &= (a_2(x)y')' + a_0(x)y. \end{aligned} \tag{5.79}$$

This is in the correct form. We just identify $p(x) = a_2(x)$ and $q(x) = a_0(x)$.

However, consider the differential equation

$$x^2y'' + xy' + 2y = 0.$$

In this case $a_2(x) = x^2$ and $a_2'(x) = 2x \neq a_1(x)$. The linear differential operator in this equation is not of Sturm-Liouville type. But, we can change it to a Sturm Liouville operator.

Proof. In the Sturm Liouville operator the derivative terms are gathered together into one perfect derivative. This is similar to what we saw in the Chapter 2 when we solved linear first order equations. In that case we sought an integrating factor. We can do the same thing here. We seek a multiplicative function $\mu(x)$ that we can multiply through (5.74) so that it can be written in Sturm-Liouville form. We first divide out the $a_2(x)$, giving

$$y'' + \frac{a_1(x)}{a_2(x)}y' + \frac{a_0(x)}{a_2(x)}y = \frac{f(x)}{a_2(x)}.$$

Now, we multiply the differential equation by μ :

$$\mu(x)y'' + \mu(x)\frac{a_1(x)}{a_2(x)}y' + \mu(x)\frac{a_0(x)}{a_2(x)}y = \mu(x)\frac{f(x)}{a_2(x)}.$$

The first two terms can now be combined into an exact derivative $(\mu y')'$ if $\mu(x)$ satisfies

$$\frac{d\mu}{dx} = \mu(x)\frac{a_1(x)}{a_2(x)}.$$

This is formally solved to give

$$\mu(x) = e^{\int \frac{a_1(x)}{a_2(x)}\,dx}.$$

Thus, the original equation can be multiplied by factor

$$\frac{\mu(x)}{a_2(x)} = \frac{1}{a_2(x)}e^{\int \frac{a_1(x)}{a_2(x)}\,dx}$$

to turn it into Sturm-Liouville form. □

In summary,

Conversion of a linear second order differential equation to Sturm Liouville form.

Equation (5.74),

$$a_2(x)y'' + a_1(x)y' + a_0(x)y = f(x), \tag{5.80}$$

can be put into the Sturm-Liouville form

$$\frac{d}{dx}\left(p(x)\frac{dy}{dx}\right) + q(x)y = F(x), \tag{5.81}$$

where

$$\begin{aligned} p(x) &= e^{\int \frac{a_1(x)}{a_2(x)}\,dx}, \\ q(x) &= p(x)\frac{a_0(x)}{a_2(x)}, \\ F(x) &= p(x)\frac{f(x)}{a_2(x)}. \end{aligned} \tag{5.82}$$

Example 5.11. *For the example above,*

$$x^2y'' + xy' + 2y = 0.$$

We need only multiply this equation by

$$\frac{1}{x^2}e^{\int \frac{dx}{x}} = \frac{1}{x},$$

to put the equation in Sturm-Liouville form:

$$\begin{aligned} 0 &= xy'' + y' + \frac{2}{x}y \\ &= (xy')' + \frac{2}{x}y. \end{aligned} \tag{5.83}$$

5.6.2 *Properties of Sturm-Liouville Eigenvalue Problems*

There are several properties that can be proven for the (regular) Sturm-Liouville eigenvalue problem in (5.76). However, we will not prove them all here. We will merely list some of the important facts and focus on a few of the properties.

Real, countable eigenvalues.

1. The eigenvalues are real, countable, ordered and there is a smallest eigenvalue. Thus, we can write them as $\lambda_1 < \lambda_2 < \ldots$. However, there is no largest eigenvalue and $n \to \infty$, $\lambda_n \to \infty$.

Oscillatory eigenfunctions.

2. For each eigenvalue λ_n there exists an eigenfunction ϕ_n with $n-1$ zeros on (a, b).

Orthogonality of eigenfunctions.

3. Eigenfunctions corresponding to different eigenvalues are orthogonal with respect to the weight function, $\sigma(x)$. Defining the inner product of $f(x)$ and $g(x)$ as

$$< f, g >= \int_a^b f(x)g(x)\sigma(x)\,dx, \tag{5.84}$$

then the orthogonality of the eigenfunctions can be written in the form

$$< \phi_n, \phi_m >=< \phi_n, \phi_n > \delta_{nm}, \quad n, m = 1, 2, \ldots. \tag{5.85}$$

Complete basis of eigenfunctions.

4. The set of eigenfunctions is complete; i.e., any piecewise smooth function can be represented by a generalized Fourier series expansion of the eigenfunctions,

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

where

$$c_n = \frac{< f, \phi_n >}{< \phi_n, \phi_n >}.$$

Actually, one needs $f(x) \in L^2_\sigma(a, b)$, the set of *square integrable functions* over $[a, b]$ with weight function $\sigma(x)$. By square integrable, we mean that $< f, f >< \infty$. One can show that such a space is isomorphic to a *Hilbert space*, a complete inner product space. Hilbert spaces play a special role in quantum mechanics.

Rayleigh Quotient.

5. Multiply the eigenvalue problem

$$\mathcal{L}\phi_n = -\lambda_n \sigma(x)\phi_n$$

by ϕ_n and integrate. Solve this result for λ_n, to find the *Rayleigh Quotient*

$$\lambda_n = \frac{-p\phi_n \frac{d\phi_n}{dx}\Big|_a^b - \int_a^b \left[p\left(\frac{d\phi_n}{dx}\right)^2 - q\phi_n^2 \right] dx}{< \phi_n, \phi_n >}$$

The Rayleigh quotient is useful for getting estimates of eigenvalues and proving some of the other properties.

Example 5.12. *We seek the eigenfunctions of the operator found in Example 5.11. Namely, we want to solve the eigenvalue problem*

$$\mathcal{L}y = (xy')' + \frac{2}{x}y = -\lambda\sigma y \tag{5.86}$$

subject to a set of homogeneous boundary conditions. Let's use the boundary conditions

$$y'(1) = 0, \quad y'(2) = 0.$$

[Note that we do not know $\sigma(x)$ yet, but will choose an appropriate function to obtain solutions.]

Expanding the derivative, we have

$$xy'' + y' + \frac{2}{x}y = -\lambda\sigma y.$$

Multiply through by x to obtain

$$x^2y'' + xy' + (2 + \lambda x\sigma)\, y = 0.$$

Notice that if we choose $\sigma(x) = x^{-1}$, then this equation can be made a Cauchy-Euler type equation. Thus, we have

$$x^2y'' + xy' + (\lambda + 2)\, y = 0.$$

The characteristic equation is

$$r^2 + \lambda + 2 = 0.$$

For oscillatory solutions, we need $\lambda + 2 > 0$. Thus, the general solution is

$$y(x) = c_1 \cos(\sqrt{\lambda + 2}\ln|x|) + c_2 \sin(\sqrt{\lambda + 2}\ln|x|). \tag{5.87}$$

Next we apply the boundary conditions. $y'(1) = 0$ forces $c_2 = 0$. This leaves

$$y(x) = c_1 \cos(\sqrt{\lambda + 2}\ln x).$$

The second condition, $y'(2) = 0$, yields

$$\sin(\sqrt{\lambda+2}\ln 2) = 0.$$

This will give nontrivial solutions when

$$\sqrt{\lambda+2}\ln 2 = n\pi, \quad n = 0, 1, 2, 3 \ldots.$$

In summary, the eigenfunctions for this eigenvalue problem are

$$y_n(x) = \cos\left(\frac{n\pi}{\ln 2}\ln x\right), \quad 1 \le x \le 2$$

and the eigenvalues are $\lambda_n = \left(\frac{n\pi}{\ln 2}\right)^2 - 2$ for $n = 0, 1, 2, \ldots$.

Note: We include the $n = 0$ case because $y(x) =$ constant is a solution of the $\lambda = -2$ case. More specifically, in this case the characteristic equation reduces to $r^2 = 0$. Thus, the general solution of this Cauchy-Euler equation is

$$y(x) = c_1 + c_2 \ln|x|.$$

Setting $y'(1) = 0$, forces $c_2 = 0$. $y'(2)$ automatically vanishes, leaving the solution in this case as $y(x) = c_1$.

We note that some of the properties listed in the beginning of the section hold for this example. The eigenvalues are seen to be real, countable and ordered. There is a least one, $\lambda_0 = -2$. Next, one can find the zeros of each eigenfunction on [1,2]. Then the argument of the cosine, $\frac{n\pi}{\ln 2}\ln x$, takes values 0 to $n\pi$ for $x \in [1, 2]$. The cosine function has $n - 1$ roots on this interval.

Orthogonality can be checked as well. We set up the integral and use the substitution $y = \pi \ln x / \ln 2$. This gives

$$\begin{aligned} < y_n, y_m > &= \int_1^2 \cos\left(\frac{n\pi}{\ln 2}\ln x\right)\cos\left(\frac{m\pi}{\ln 2}\ln x\right)\frac{dx}{x} \\ &= \frac{\ln 2}{\pi}\int_0^\pi \cos ny \cos my \, dy \\ &= \frac{\ln 2}{2}\delta_{n,m}. \end{aligned} \tag{5.88}$$

5.6.3 *Adjoint Operators*

In the study of the spectral theory of matrices, one learns about the adjoint of the matrix, $A^\dagger$, and the role that self-adjoint, or Hermitian, matrices play in diagonalization. Also, one needs the concept of adjoint to discuss the existence of solutions to the matrix problem $\mathbf{y} = A\mathbf{x}$. In the same spirit, one is interested in the existence of solutions of the operator equation $Lu = f$ and solutions of the corresponding eigenvalue problem. The study of linear operators on a Hilbert space is a generalization of what the reader had seen in a linear algebra course.

Just as one can find a basis of eigenvectors and diagonalize Hermitian, or self-adjoint, matrices (or, real symmetric in the case of real matrices), we will see that the Sturm-Liouville operator is self-adjoint. In this section we will define the *domain of an operator* and introduce the notion of *adjoint operators*. In the last section we discuss the role the adjoint plays in the existence of solutions to the operator equation $Lu = f$.

We first introduce some definitions.

Definition 5.2. The *domain* of a differential operator L is the set of all $u \in L^2_\sigma(a,b)$ satisfying a given set of homogeneous boundary conditions.

The *adjoint*, $L^\dagger$, of operator L.

Definition 5.3. The *adjoint*, $L^\dagger$, of operator L satisfies

$$< u, Lv >=< L^\dagger u, v >$$

for all v in the domain of L and u in the domain of $L^\dagger$.

Example 5.13. *As an example, we find the adjoint of second order linear differential operator* $L = a_2(x)\frac{d^2}{dx^2} + a_1(x)\frac{d}{dx} + a_0(x)$.

In order to find the adjoint, we place the operator under an integral. So, we consider the inner product

$$< u, Lv >= \int_a^b u(a_2 v'' + a_1 v' + a_0 v)\,dx.$$

We have to move the operator L from v and determine what operator is acting on u in order to formally preserve the inner product. For a simple operator like $L = \frac{d}{dx}$*, this is easily done using integration by parts. For the given operator, we will need to apply several integrations by parts to the individual terms. We will consider the individual terms.*

First we consider the $a_1 v'$ *term. Integration by parts yields*

$$\int_a^b u(x)a_1(x)v'(x)\,dx = a_1(x)u(x)v(x)\Big|_a^b - \int_a^b (u(x)a_1(x))'v(x)\,dx. \tag{5.89}$$

Now, we consider the $a_2 v''$ *term. In this case it will take two integrations by parts:*

$$\begin{aligned}\int_a^b u(x)a_2(x)v''(x)\,dx &= a_2(x)u(x)v'(x)\Big|_a^b - \int_a^b (u(x)a_2(x))'v(x)'\,dx \\ &= \left[a_2(x)u(x)v'(x) - (a_2(x)u(x))'v(x)\right]\Big|_a^b \\ &\quad + \int_a^b (u(x)a_2(x))''v(x)\,dx. \end{aligned} \tag{5.90}$$

Combining these results, we obtain

$$< u, Lv > \;=\; \int_a^b u(a_2 v'' + a_1 v' + a_0 v)\,dx$$

$$\begin{aligned} &= \left[a_1(x)u(x)v(x) + a_2(x)u(x)v'(x) - (a_2(x)u(x))'v(x)\right]\Big|_a^b \\ &\quad + \int_a^b \left[(a_2u)'' - (a_1u)' + a_0u\right] v\,dx. \end{aligned} \tag{5.91}$$

Inserting the boundary conditions for v, one has to determine boundary conditions for u such that

$$\left[a_1(x)u(x)v(x) + a_2(x)u(x)v'(x) - (a_2(x)u(x))'v(x)\right]\Big|_a^b = 0.$$

This leaves

$$< u, Lv >= \int_a^b \left[(a_2u)'' - (a_1u)' + a_0u\right] v\,dx \equiv < L^\dagger u, v > .$$

Therefore,

$$L^\dagger = \frac{d^2}{dx^2}a_2(x) - \frac{d}{dx}a_1(x) + a_0(x). \tag{5.92}$$

Self-adjoint operators.

When $L^\dagger = L$, the operator is called *formally self-adjoint*. When the domain of L is the same as the domain of $L^\dagger$, the term *self-adjoint* is used. As the domain is important in establishing self-adjointness, we need to do a complete example in which the domain of the adjoint is found.

Example 5.14. *Determine $L^\dagger$ and its domain for operator $Lu = \frac{du}{dx}$ where u satisfies the boundary conditions $u(0) = 2u(1)$ on $[0,1]$.*

We need to find the adjoint operator satisfying $< v, Lu >=< L^\dagger v, u >$. Therefore, we rewrite the integral

$$< v, Lu >= \int_0^1 v\frac{du}{dx}\,dx = uv|_0^1 - \int_0^1 u\frac{dv}{dx}\,dx =< L^\dagger v, u > .$$

From this we have the adjoint problem consisting of an adjoint operator and the associated boundary condition:

1. $L^\dagger = -\frac{d}{dx}$.
2. $uv\Big|_0^1 = 0 \Rightarrow 0 = u(1)[v(1) - 2v(0)] \Rightarrow v(1) = 2v(0)$.

5.6.4 Lagrange's and Green's Identities

Before turning to the proofs that the eigenvalues of a Sturm-Liouville problem are real and the associated eigenfunctions orthogonal, we will first need to introduce two important identities. For the Sturm-Liouville operator,

$$\mathcal{L} = \frac{d}{dx}\left(p\frac{d}{dx}\right) + q,$$

we have the two identities:

Lagrange's Identity:	$u\mathcal{L}v - v\mathcal{L}u$	$= [p(uv' - vu')]'.$
Green's Identity:	$\int_a^b (u\mathcal{L}v - v\mathcal{L}u)\,dx$	$= [p(uv' - vu')]\|_a^b.$

Proof. The proof of Lagrange's identity follows by a simple manipulations of the operator:

$$\begin{aligned} u\mathcal{L}v - v\mathcal{L}u &= u\left[\frac{d}{dx}\left(p\frac{dv}{dx}\right) + qv\right] - v\left[\frac{d}{dx}\left(p\frac{du}{dx}\right) + qu\right] \\ &= u\frac{d}{dx}\left(p\frac{dv}{dx}\right) - v\frac{d}{dx}\left(p\frac{du}{dx}\right) \\ &= u\frac{d}{dx}\left(p\frac{dv}{dx}\right) + p\frac{du}{dx}\frac{dv}{dx} - v\frac{d}{dx}\left(p\frac{du}{dx}\right) - p\frac{du}{dx}\frac{dv}{dx} \\ &= \frac{d}{dx}\left[pu\frac{dv}{dx} - pv\frac{du}{dx}\right]. \end{aligned} \tag{5.93}$$

Green's identity is simply proven by integrating Lagrange's identity. □

5.6.5 Orthogonality and Reality

WE ARE NOW READY to prove that the eigenvalues of a Sturm-Liouville problem are real and the corresponding eigenfunctions are orthogonal. These are easily established using Green's identity, which in turn is a statement about the Sturm-Liouville operator being self-adjoint.

Theorem 5.3. *The eigenvalues of the Sturm-Liouville problem (5.76) are real.*

Proof. Let $\phi_n(x)$ be a solution of the eigenvalue problem associated with λ_n:

$$\mathcal{L}\phi_n = -\lambda_n \sigma \phi_n.$$

The complex conjugate of this equation is

$$\mathcal{L}\overline{\phi}_n = -\overline{\lambda}_n \sigma \overline{\phi}_n.$$

Now, multiply the first equation by $\overline{\phi}_n$ and the second equation by ϕ_n and then subtract the results. We obtain

$$\overline{\phi}_n \mathcal{L}\phi_n - \phi_n \mathcal{L}\overline{\phi}_n = (\overline{\lambda}_n - \lambda_n)\sigma\phi_n\overline{\phi}_n.$$

Integrate both sides of this equation:

$$\int_a^b \left(\overline{\phi}_n \mathcal{L}\phi_n - \phi_n \mathcal{L}\overline{\phi}_n\right) dx = (\overline{\lambda}_n - \lambda_n)\int_a^b \sigma\phi_n\overline{\phi}_n\,dx.$$

Apply Green's identity to the left hand side to find

$$[p(\overline{\phi}_n\phi_n' - \phi_n\overline{\phi}_n')]|_a^b = (\overline{\lambda}_n - \lambda_n)\int_a^b \sigma\phi_n\overline{\phi}_n\,dx.$$

Using the homogeneous boundary conditions (5.77) for a self-adjoint operator, the left side vanishes to give

$$0 = (\bar{\lambda}_n - \lambda_n) \int_a^b \sigma \|\phi_n\|^2 \, dx.$$

The integral is nonnegative, so we must have $\bar{\lambda}_n = \lambda_n$. Therefore, the eigenvalues are real. □

Theorem 5.4. *The eigenfunctions corresponding to different eigenvalues of the Sturm-Liouville problem (5.76) are orthogonal.*

Proof. This is proven similar to the last theorem. Let $\phi_n(x)$ be a solution of the eigenvalue problem associated with λ_n,

$$\mathcal{L}\phi_n = -\lambda_n \sigma \phi_n,$$

and let $\phi_m(x)$ be a solution of the eigenvalue problem associated with $\lambda_m \neq \lambda_n$,

$$\mathcal{L}\phi_m = -\lambda_m \sigma \phi_m,$$

Now, multiply the first equation by ϕ_m and the second equation by ϕ_n. Subtracting the results, we obtain

$$\phi_m \mathcal{L}\phi_n - \phi_n \mathcal{L}\phi_m = (\lambda_m - \lambda_n)\sigma \phi_n \phi_m$$

Similar to the previous proof, we integrate both sides of the equation and use Green's identity and the boundary conditions for a self-adjoint operator. This leaves

$$0 = (\lambda_m - \lambda_n) \int_a^b \sigma \phi_n \phi_m \, dx.$$

Since the eigenvalues are distinct, we can divide by $\lambda_m - \lambda_n$, leaving the desired result,

$$\int_a^b \sigma \phi_n \phi_m \, dx = 0.$$

Therefore, the eigenfunctions are orthogonal with respect to the weight function $\sigma(x)$. □

5.6.6 The Rayleigh Quotient - optional

The Rayleigh quotient is useful for getting estimates of eigenvalues and proving some of the other properties associated with Sturm-Liouville eigenvalue problems. We begin by multiplying the eigenvalue problem

$$\mathcal{L}\phi_n = -\lambda_n \sigma(x) \phi_n$$

by ϕ_n and integrating. This gives

$$\int_a^b \left[\phi_n \frac{d}{dx}\left(p\frac{d\phi_n}{dx}\right) + q\phi_n^2\right] dx = -\lambda \int_a^b \phi_n^2\, dx.$$

One can solve the last equation for λ to find

$$\lambda = \frac{-\int_a^b \left[\phi_n \frac{d}{dx}\left(p\frac{d\phi_n}{dx}\right) + q\phi_n^2\right] dx}{\int_a^b \phi_n^2 \sigma\, dx}.$$

It appears that we have solved for the eigenvalue and have not needed the machinery we had developed in Chapter 4 for studying boundary value problems. However, we really cannot evaluate this expression because we do not know the eigenfunctions, $\phi_n(x)$ yet. Nevertheless, we will see what we can determine.

One can rewrite this result by performing an integration by parts on the first term in the numerator. Namely, pick $u = \phi_n$ and $dv = \frac{d}{dx}\left(p\frac{d\phi_n}{dx}\right) dx$ for the standard integration by parts formula. Then, we have

$$\int_a^b \phi_n \frac{d}{dx}\left(p\frac{d\phi_n}{dx}\right) dx = p\phi_n \frac{d\phi_n}{dx}\Big|_a^b - \int_a^b \left[p\left(\frac{d\phi_n}{dx}\right)^2 - q\phi_n^2\right] dx.$$

Inserting the new formula into the expression for λ, leads to the *Rayleigh Quotient*

$$\lambda_n = \frac{-p\phi_n \frac{d\phi_n}{dx}\Big|_a^b + \int_a^b \left[p\left(\frac{d\phi_n}{dx}\right)^2 - q\phi_n^2\right] dx}{\int_a^b \phi_n^2 \sigma\, dx}. \tag{5.94}$$

In many applications the sign of the eigenvalue is important. As we had seen in the solution of the heat equation, $T' + k\lambda T = 0$. Since we expect the heat energy to diffuse, the solutions should decay in time. Thus, we would expect $\lambda > 0$. In studying the wave equation, one expects vibrations and these are only possible with the correct sign of the eigenvalue (positive again). Thus, in order to have nonnegative eigenvalues, we see from (5.94) that

a. $q(x) \leq 0$, and

b. $-p\phi_n \frac{d\phi_n}{dx}\big|_a^b \geq 0$.

Furthermore, if λ is a zero eigenvalue, then $q(x) \equiv 0$ and $\alpha_1 = \alpha_2 = 0$ in the homogeneous boundary conditions. This can be seen by setting the numerator equal to zero. Then, $q(x) = 0$ and $\phi_n'(x) = 0$. The second of these conditions inserted into the boundary conditions forces the restriction on the type of boundary conditions.

One of the (unproven here) properties of Sturm-Liouville eigenvalue problems with homogeneous boundary conditions is that the eigenvalues are ordered, $\lambda_1 < \lambda_2 < \ldots$. Thus, there is a smallest eigenvalue. It turns out that for any continuous function, $y(x)$,

$$\lambda_1 = \min_{y(x)} \frac{-py\frac{dy}{dx}|_a^b + \int_a^b \left[p\left(\frac{dy}{dx}\right)^2 - qy^2 \right] dx}{\int_a^b y^2 \sigma \, dx} \tag{5.95}$$

and this minimum is obtained when $y(x) = \phi_1(x)$. This result can be used to get estimates of the minimum eigenvalue by using trial functions which are continuous and satisfy the boundary conditions, but do not necessarily satisfy the differential equation.

Example 5.15. *We have already solved the eigenvalue problem $\phi'' + \lambda\phi = 0$, $\phi(0) = 0$, $\phi(1) = 0$. In this case, the lowest eigenvalue is $\lambda_1 = \pi^2$. We can pick a nice function satisfying the boundary conditions, say $y(x) = x - x^2$. Inserting this into Equation (5.95), we find*

$$\lambda_1 \leq \frac{\int_0^1 (1-2x)^2 \, dx}{\int_0^1 (x-x^2)^2 \, dx} = 10.$$

Indeed, $10 \geq \pi^2$.

5.6.7 The Eigenfunction Expansion Method - optional

In this section we show how one can solve the nonhomogeneous problem $\mathcal{L}y = f$ using expansions over the basis of Sturm-Liouville eigenfunctions. In this chapter we have seen that Sturm-Liouville eigenvalue problems have the requisite set of orthogonal eigenfunctions. In this section we will apply the eigenfunction expansion method to solve a particular nonhomogenous boundary value problem.

Recall that one starts with a nonhomogeneous differential equation

$$\mathcal{L}y = f,$$

where $y(x)$ is to satisfy given homogeneous boundary conditions. The method makes use of the eigenfunctions satisfying the eigenvalue problem

$$\mathcal{L}\phi_n = -\lambda_n \sigma \phi_n$$

subject to the given boundary conditions. Then, one assumes that $y(x)$ can be written as an expansion in the eigenfunctions,

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

and inserts the expansion into the nonhomogeneous equation. This gives

$$f(x) = \mathcal{L}\left(\sum_{n=1}^{\infty} c_n \phi_n(x)\right) = -\sum_{n=1}^{\infty} c_n \lambda_n \sigma(x) \phi_n(x).$$

The expansion coefficients are then found by making use of the orthogonality of the eigenfunctions. Namely, we multiply the last equation by $\phi_m(x)$ and integrate. We obtain

$$\int_a^b f(x)\phi_m(x)\,dx = -\sum_{n=1}^{\infty} c_n\lambda_n \int_a^b \phi_n(x)\phi_m(x)\sigma(x)\,dx.$$

Orthogonality yields

$$\int_a^b f(x)\phi_m(x)\,dx = -c_m\lambda_m \int_a^b \phi_m^2(x)\sigma(x)\,dx.$$

Solving for c_m, we have

$$c_m = -\frac{\int_a^b f(x)\phi_m(x)\,dx}{\lambda_m \int_a^b \phi_m^2(x)\sigma(x)\,dx}.$$

Example 5.16. *As an example, we consider the solution of the boundary value problem*

$$(xy')' + \frac{y}{x} = \frac{1}{x}, \quad x \in [1, e], \tag{5.96}$$

$$y(1) = 0 = y(e). \tag{5.97}$$

This equation is already in self-adjoint form. So, we know that the associated Sturm-Liouville eigenvalue problem has an orthogonal set of eigenfunctions. We first determine this set. Namely, we need to solve

$$(x\phi')' + \frac{\phi}{x} = -\lambda\sigma\phi, \quad \phi(1) = 0 = \phi(e). \tag{5.98}$$

Rearranging the terms and multiplying by x, we have that

$$x^2\phi'' + x\phi' + (1 + \lambda\sigma x)\phi = 0.$$

This is almost an equation of Cauchy-Euler type. Picking the weight function $\sigma(x) = \frac{1}{x}$, we have

$$x^2\phi'' + x\phi' + (1 + \lambda)\phi = 0.$$

This is easily solved. The characteristic equation is

$$r^2 + (1 + \lambda) = 0.$$

One obtains nontrivial solutions of the eigenvalue problem satisfying the boundary conditions when $\lambda > -1$. The solutions are

$$\phi_n(x) = A\sin(n\pi \ln x), \quad n = 1, 2, \ldots.$$

where $\lambda_n = n^2\pi^2 - 1$.

It is often useful to normalize the eigenfunctions. This means that one chooses A so that the norm of each eigenfunction is one. Thus, we have

$$\begin{aligned} 1 &= \int_1^e \phi_n(x)^2 \sigma(x)\, dx \\ &= A^2 \int_1^e \sin(n\pi \ln x)\frac{1}{x}\, dx \\ &= A^2 \int_0^1 \sin(n\pi y)\, dy = \frac{1}{2}A^2. \end{aligned} \tag{5.99}$$

Thus, $A = \sqrt{2}$.

We now turn towards solving the nonhomogeneous problem, $\mathcal{L}y = \frac{1}{x}$. *We first expand the unknown solution in terms of the eigenfunctions,*

$$y(x) = \sum_{n=1}^{\infty} c_n \sqrt{2} \sin(n\pi \ln x).$$

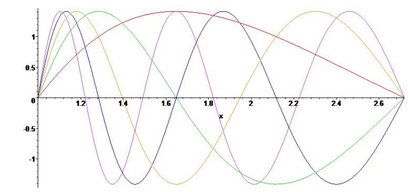

Figure 5.11: Plots of the first five eigenfunctions, $y(x) = \sqrt{2}\sin(n\pi \ln x)$.

Inserting this solution into the differential equation, we have

$$\frac{1}{x} = \mathcal{L}y = -\sum_{n=1}^{\infty} c_n \lambda_n \sqrt{2} \sin(n\pi \ln x)\frac{1}{x}.$$

Next, we make use of orthogonality. Multiplying both sides by $\phi_m(x) = \sqrt{2}\sin(m\pi \ln x)$ *and integrating, gives*

$$\lambda_m c_m = \int_1^e \sqrt{2} \sin(m\pi \ln x)\frac{1}{x}\, dx = \frac{\sqrt{2}}{m\pi}[(-1)^m - 1].$$

Solving for c_m, *we have*

$$c_m = \frac{\sqrt{2}}{m\pi}\frac{[(-1)^m - 1]}{m^2\pi^2 - 1}.$$

Finally, we insert our coefficients into the expansion for $y(x)$. *The solution is then*

$$y(x) = \sum_{n=1}^{\infty} \frac{2}{n\pi}\frac{[(-1)^n - 1]}{n^2\pi^2 - 1} \sin(n\pi \ln(x)).$$

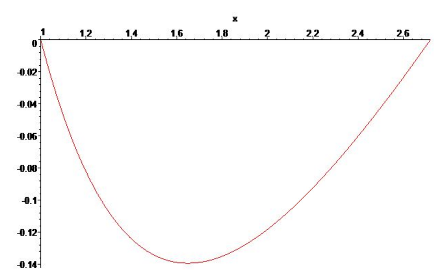

Figure 5.12: Plot of the solution in Example 5.16.

5.7 *Appendix: The Least Squares Approximation*

IN THE FIRST SECTION of this chapter we showed that we can expand functions over an infinite set of basis functions as

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

and that the generalized Fourier coefficients are given by

$$c_n = \frac{< \phi_n, f >}{< \phi_n, \phi_n >}.$$

In this section we turn to a discussion of approximating $f(x)$ by the partial sums $\sum_{n=1}^{N} c_n \phi_n(x)$ and showing that the Fourier coefficients are the best coefficients minimizing the deviation of the partial sum from $f(x)$. This will lead us to a discussion of the convergence of Fourier series.

More specifically, we set the following goal:

Goal

To find the best approximation of $f(x)$ on $[a, b]$ by $S_N(x) = \sum_{n=1}^{N} c_n \phi_n(x)$ for a set of fixed functions $\phi_n(x)$; i.e., to find the c_n's such that $S_N(x)$ approximates $f(x)$ in the *least squares sense*.

We want to measure the deviation of the finite sum from the given function. Essentially, we want to look at the error made in the approximation. This is done by introducing the *mean square deviation:*

$$E_N = \int_a^b [f(x) - S_N(x)]^2 \rho(x)\, dx,$$

where we have introduced the weight function $\rho(x) > 0$. It gives us a sense as to how close the Nth partial sum is to $f(x)$.

We want to minimize this deviation by choosing the right c_n's. We begin by inserting the partial sums and expand the square in the integrand:

$$\begin{aligned} E_N &= \int_a^b [f(x) - S_N(x)]^2 \rho(x)\, dx \\ &= \int_a^b \left[f(x) - \sum_{n=1}^{N} c_n \phi_n(x) \right]^2 \rho(x)\, dx \\ &= \int_a^b f^2(x)\rho(x)\, dx - 2\int_a^b f(x) \sum_{n=1}^{N} c_n \phi_n(x) \rho(x)\, dx \\ &\quad + \int_a^b \sum_{n=1}^{N} c_n \phi_n(x) \sum_{m=1}^{N} c_m \phi_m(x) \rho(x)\, dx \end{aligned} \tag{5.100}$$

Looking at the three resulting integrals, we see that the first term is just the inner product of f with itself. The other integrations can be rewritten after interchanging the order of integration and summation. The double sum can be reduced to a single sum using the orthogonality of the ϕ_n's. Thus, we have

$$\begin{aligned} E_N &= < f, f > -2 \sum_{n=1}^{N} c_n < f, \phi_n > + \sum_{n=1}^{N} \sum_{m=1}^{N} c_n c_m < \phi_n, \phi_m > \\ &= < f, f > -2 \sum_{n=1}^{N} c_n < f, \phi_n > + \sum_{n=1}^{N} c_n^2 < \phi_n, \phi_n > . \end{aligned} \tag{5.101}$$

We are interested in finding the coefficients, so we will complete the square in c_n. Focusing on the last two terms, we have

$$\begin{aligned} E_N - < f, f > &= -2\sum_{n=1}^{N} c_n < f, \phi_n > + \sum_{n=1}^{N} c_n^2 < \phi_n, \phi_n > \\ &= \sum_{n=1}^{N} < \phi_n, \phi_n > c_n^2 - 2 < f, \phi_n > c_n \\ &= \sum_{n=1}^{N} < \phi_n, \phi_n > \left[c_n^2 - \frac{2 < f, \phi_n >}{< \phi_n, \phi_n >} c_n \right] \\ &= \sum_{n=1}^{N} < \phi_n, \phi_n > \left[\left(c_n - \frac{< f, \phi_n >}{< \phi_n, \phi_n >} \right)^2 - \left(\frac{< f, \phi_n >}{< \phi_n, \phi_n >} \right)^2 \right]. \end{aligned} \tag{5.102}$$

To this point we have shown that the mean square deviation is given as

$$E_N = < f, f > + \sum_{n=1}^{N} < \phi_n, \phi_n > \left[\left(c_n - \frac{< f, \phi_n >}{< \phi_n, \phi_n >} \right)^2 - \left(\frac{< f, \phi_n >}{< \phi_n, \phi_n >} \right)^2 \right].$$

So, E_N is minimized by choosing $c_n = \frac{< f, \phi_n >}{< \phi_n, \phi_n >}$. However, these are the Fourier Coefficients. This minimization is often referred to as **Minimization in Least Squares Sense**.

Inserting the Fourier coefficients into the mean square deviation yields

$$0 \leq E_N = < f, f > - \sum_{n=1}^{N} c_n^2 < \phi_n, \phi_n > .$$

Thus, we obtain *Bessel's Inequality*:

$$< f, f > \geq \sum_{n=1}^{N} c_n^2 < \phi_n, \phi_n > .$$

For convergence, we next let N get large and see if the partial sums converge to the function. In particular, we say that the infinite series *converges in the mean* if

$$\int_a^b [f(x) - S_N(x)]^2 \rho(x)\, dx \ \to 0 \text{ as } N \to \infty.$$

Letting N get large in Bessel's inequality shows that $\sum_{n=1}^{N} c_n^2 <$ $\phi_n, \phi_n >$ converges if

$$(< f, f > = \int_a^b f^2(x) \rho(x)\, dx < \infty.$$

The space of all such f is denoted $L^2_\rho(a, b)$, the space of square integrable functions on (a, b) with weight $\rho(x)$.

From the nth term divergence theorem we know that $\sum a_n$ converges implies that $a_n \to 0$ as $n \to \infty$. Therefore, in this problem the terms $c_n^2 < \phi_n, \phi_n >$ approach zero as n gets large. This is only possible if the c_n's go to zero as n gets large. Thus, if $\sum_{n=1}^N c_n\phi_n$ *converges in the mean* to f, then $\int_a^b [f(x) - \sum_{n=1}^N c_n\phi_n]^2 \rho(x)\, dx$ approaches zero as $N \to \infty$. This implies from the above derivation of Bessel's inequality that

$$< f, f > - \sum_{n=1}^{N} c_n^2 (\phi_n, \phi_n) \to 0.$$

This leads to *Parseval's equality*:

$$< f, f >= \sum_{n=1}^{\infty} c_n^2 < \phi_n, \phi_n > .$$

Parseval's equality holds if and only if

$$\lim_{N\to\infty} \int_a^b \left(f(x) - \sum_{n=1}^{N} c_n\phi_n(x)\right)^2 \rho(x)\; dx = 0.$$

If this is true for every square integrable function in $L^2_\rho(a, b)$, then the set of functions $\{\phi_n(x)\}_{n=1}^\infty$ is said to be **complete**. One can view these functions as an infinite dimensional basis for the space of square integrable functions on (a, b) with weight $\rho(x) > 0$.

One can extend the above limit $c_n \to 0$ as $n \to \infty$, by assuming that $\frac{\phi_n(x)}{\|\phi_n\|}$ is uniformly bounded and that $\int_a^b |f(x)|\rho(x)\; dx < \infty$. This is the **Riemann-Lebesque Lemma**, but will not be proven now.

5.8 *Appendix: The Fredholm Alternative Theorem*

GIVEN THAT $Ly = f$, when can one expect to find a solution? Is it unique? These questions are answered by the Fredholm Alternative Theorem. This theorem occurs in many forms from a statement about solutions to systems of algebraic equations to solutions of boundary value problems and integral equations. The theorem comes in two parts, thus the term "alternative". Either the equation has exactly one solution for all f, or the equation has many solutions for some f's and none for the rest.

The reader is familiar with the statements of the Fredholm Alternative for the solution of systems of algebraic equations. One seeks solutions of the system $Ax = b$ for A an $n \times m$ matrix. Defining the matrix adjoint, A^* through $< Ax, y >=< x, A^*y >$ for all $x, y, \in \mathcal{C}^n$, then either

Theorem 5.5. First Alternative

The equation $Ax = b$ has a solution if and only if $< b, v >= 0$ for all v such that $A^*v = 0$.

or

Theorem 5.6. Second Alternative

A solution of $Ax = b$, if it exists, is unique if and only if $x = 0$ is the only solution of $Ax = 0$.

The second alternative is more familiar when given in the form: The solution of a nonhomogeneous system of n equations and n unknowns is unique if the only solution to the homogeneous problem is the zero solution. Or, equivalently, A is invertible, or has nonzero determinant.

Proof. We prove the second theorem first. Assume that $Ax = 0$ for $x \neq 0$ and $Ax_0 = b$. Then $A(x_0 + \alpha x) = b$ for all α. Therefore, the solution is not unique. Conversely, if there are two different solutions, x_1 and x_2, satisfying $Ax_1 = b$ and $Ax_2 = b$, then one has a nonzero solution $x = x_1 - x_2$ such that $Ax = A(x_1 - x_2) = 0$.

The proof of the first part of the first theorem is simple. Let $A^*v = 0$ and $Ax_0 = b$. Then we have

$$< b, v >=< Ax_0, v >=< x_0, A^*v >= 0.$$

For the second part we assume that $< b, v >= 0$ for all v such that $A^*v = 0$. Write b as the sum of a part that is in the range of A and a part that in the space orthogonal to the range of A, $b = b_R + b_O$. Then, $0 =< b_O, Ax >=< A^*b, x >$ for all x. Thus, A^*b_O. Since $< b, v >= 0$ for all v in the nullspace of A^*, then $< b, b_O >= 0$.

Therefore, $< b, v >= 0$ implies that

$$0 =< b, b_O >=< b_R + b_O, b_O >=< b_O, b_O > .$$

This means that $b_O = 0$, giving $b = b_R$ is in the range of A. So, $Ax = b$ has a solution. □

Example 5.17. *Determine the allowed forms of* **b** *for a solution of $A\mathbf{x} = \mathbf{b}$ to exist, where*

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix}.$$

First note that $A^* = \overline{A}^T$. This is seen by looking at

$$\begin{aligned} < A\mathbf{x}, \mathbf{y} > &= < \mathbf{x}, A^*\mathbf{y} > \\ \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_j\bar{y}_i &= \sum_{j=1}^{n} x_j \sum_{j=1}^{n} a_{ij}\bar{y}_i \\ &= \sum_{j=1}^{n} x_j \overline{\sum_{j=1}^{n} (\bar{a}^T)_{ji}\, y_i}. \end{aligned} \tag{5.103}$$

For this example,

$$A^* = \begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix}.$$

We next solve $A^*\mathbf{v} = 0$*. This means,* $v_1 + 3v_2 = 0$*. So, the nullspace of* A^* *is spanned by* $\mathbf{v} = (3, -1)^T$*. For a solution of* $A\mathbf{x} = \mathbf{b}$ *to exist,* $\mathbf{b}$ *would have to be orthogonal to* $\mathbf{v}$*. Therefore, a solution exists when*

$$\mathbf{b} = \alpha \begin{pmatrix} 1 \\ 3 \end{pmatrix}.$$

So, what does this say about solutions of boundary value problems? We need a more general theory for linear operators. A more general statement would be

Theorem 5.7. *If* L *is a bounded linear operator on a Hilbert space, then* $Ly = f$ *has a solution if and only if* $< f, v >= 0$ *for every* v *such that* $L^\dagger v = 0$.

The statement for boundary value problems is similar. However, we need to be careful to treat the boundary conditions in our statement. As we have seen, after several integrations by parts we have that

$$< \mathcal{L}u, v >= S(u, v) + < u, \mathcal{L}^\dagger v >,$$

where $S(u, v)$ involves the boundary conditions on u and v. Note that for nonhomogeneous boundary conditions, this term may no longer vanish.

Theorem 5.8. *The solution of the boundary value problem* $\mathcal{L}u = f$ *with boundary conditions* $Bu = g$ *exists if and only if*

$$< f, v > -S(u, v) = 0$$

for all v *satisfying* $\mathcal{L}^\dagger v = 0$ *and* $B^\dagger v = 0$.

Example 5.18. *Consider the problem*

$$u'' + u = f(x), \quad u(0) - u(2\pi) = \alpha, u'(0) - u'(2\pi) = \beta.$$

Only certain values of α *and* β *will lead to solutions. We first note that*

$$L = L^\dagger = \frac{d^2}{dx^2} + 1.$$

Solutions of

$$L^\dagger v = 0, \quad v(0) - v(2\pi) = 0, v'(0) - v'(2\pi) = 0$$

are easily found to be linear combinations of $v = \sin x$ *and* $v = \cos x$.

Next one computes

$$\begin{aligned} S(u,v) &= \left[u'v - uv'\right]_0^{2\pi} \\ &= u'(2\pi)v(2\pi) - u(2\pi)v'(2\pi) - u'(0)v(0) + u(0)v'(0). \end{aligned} \tag{5.104}$$

For $v(x) = \sin x$, this yields

$$S(u, \sin x) = -u(2\pi) + u(0) = \alpha.$$

Similarly,

$$S(u, \cos x) = \beta.$$

Using $< f, v > -S(u,v) = 0$, this leads to the conditions that we were seeking,

$$\int_0^{2\pi} f(x) \sin x \, dx = \alpha,$$

$$\int_0^{2\pi} f(x) \cos x \, dx = \beta.$$

Problems

1. Consider the set of vectors $(-1,1,1)$, $(1,-1,1)$, $(1,1,-1)$.

a. Use the Gram-Schmidt process to find an orthonormal basis for R^3 using this set in the given order.

b. What do you get if you do reverse the order of these vectors?

2. Use the Gram-Schmidt process to find the first four orthogonal polynomials satisfying the following:

a. Interval: $(-\infty, \infty)$ Weight Function: e^{-x^2}.

b. Interval: $(0, \infty)$ Weight Function: e^{-x}.

3. Find $P_4(x)$ using

a. The Rodrigues Formula in Equation (5.21).

b. The three term recursion formula in Equation (5.23).

4. In Equations (5.36)-(5.43) we provide several identities for Legendre polynomials. Derive the results in Equations (5.37)-(5.43) as decribed in the text. Namely,

a. Differentiating Equation (5.36) with respect to x, derive Equation (5.37).

b. Derive Equation (5.38) by differentiating $g(x,t)$ with respect to x and rearranging the resulting infinite series.

c. Combining the last result with Equation (5.36), derive Equations (5.39)-(5.40).

d. Adding and subtracting Equations (5.39)-(5.40), obtain Equations (5.41)-(5.42).

e. Derive Equation (5.43) using some of the other identities.

5. Use the recursion relation (5.23) to evaluate $\int_{-1}^{1} xP_n(x)P_m(x)\,dx$, $n \leq m$.

6. Expand the following in a Fourier-Legendre series for $x \in (-1, 1)$.

a. $f(x) = x^2$.

b. $f(x) = 5x^4 + 2x^3 - x + 3$.

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

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

7. Use integration by parts to show $\Gamma(x+1) = x\Gamma(x)$.

8. Prove the double factorial identities:

$$(2n)!! = 2^n n!$$

and

$$(2n-1)!! = \frac{(2n)!}{2^n n!}.$$

9. Express the following as Gamma functions. Namely, noting the form $\Gamma(x+1) = \int_0^\infty t^x e^{-t}\,dt$ and using an appropriate substitution, each expression can be written in terms of a Gamma function.

a. $\int_0^\infty x^{2/3} e^{-x}\,dx$.

b. $\int_0^\infty x^5 e^{-x^2}\,dx$

c. $\int_0^1 \left[\ln\left(\frac{1}{x}\right)\right]^n\,dx$

10. The coefficients C_k^p in the binomial expansion for $(1+x)^p$ are given by

$$C_k^p = \frac{p(p-1)\cdots(p-k+1)}{k!}.$$

a. Write C_k^p in terms of Gamma functions.

b. For $p = 1/2$ use the properties of Gamma functions to write $C_k^{1/2}$ in terms of factorials.

c. Confirm you answer in part b by deriving the Maclaurin series expansion of $(1+x)^{1/2}$.

11. The Hermite polynomials, $H_n(x)$, satisfy the following:

i. $< H_n, H_m >= \int_{-\infty}^{\infty} e^{-x^2} H_n(x) H_m(x)\, dx = \sqrt{\pi} 2^n n! \delta_{n,m}$.

ii. $H_n'(x) = 2n H_{n-1}(x)$.

iii. $H_{n+1}(x) = 2x H_n(x) - 2n H_{n-1}(x)$.

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

Using these, show that

a. $H_n'' - 2x H_n' + 2n H_n = 0$. [Use properties ii. and iii.]

b. $\int_{-\infty}^{\infty} x e^{-x^2} H_n(x) H_m(x)\, dx = \sqrt{\pi} 2^{n-1} n! \left[\delta_{m,n-1} + 2(n+1)\delta_{m,n+1}\right]$. [Use properties i. and iii.]

c. $H_n(0) = \begin{cases} 0, & n \text{ odd}, \\ (-1)^m \frac{(2m)!}{m!}, & n = 2m. \end{cases}$ [Let $x = 0$ in iii. and iterate. Note from iv. that $H_0(x) = 1$ and $H_1(x) = 2x$.]

12. In Maple one can type **simplify(LegendreP(2*n-2,0)-LegendreP(2*n,0));** to find a value for $P_{2n-2}(0) - P_{2n}(0)$. It gives the result in terms of Gamma functions. However, in Example 5.8 for Fourier-Legendre series, the value is given in terms of double factorials! So, we have

$$P_{2n-2}(0) - P_{2n}(0) = \frac{\sqrt{\pi}(4n-1)}{2\Gamma(n+1)\Gamma\left(\frac{3}{2}-n\right)} = (-1)^n \frac{(2n-3)!!}{(2n-2)!!} \frac{4n-1}{2n}.$$

You will verify that both results are the same by doing the following:

a. Prove that $P_{2n}(0) = (-1)^n \frac{(2n-1)!!}{(2n)!!}$ using the generating function and a binomial expansion.

b. Prove that $\Gamma\left(n + \frac{1}{2}\right) = \frac{(2n-1)!!}{2^n}\sqrt{\pi}$ using $\Gamma(x) = (x-1)\Gamma(x-1)$ and iteration.

c. Verify the result from Maple that $P_{2n-2}(0) - P_{2n}(0) = \frac{\sqrt{\pi}(4n-1)}{2\Gamma(n+1)\Gamma\left(\frac{3}{2}-n\right)}$.

d. Can either expression for $P_{2n-2}(0) - P_{2n}(0)$ be simplified further?

13. A solution Bessel's equation, $x^2 y'' + x y' + (x^2 - n^2) y = 0$, , can be found using the guess $y(x) = \sum_{j=0}^{\infty} a_j x^{j+n}$. One obtains the recurrence relation $a_j = \frac{-1}{j(2n+j)} a_{j-2}$. Show that for $a_0 = (n! 2^n)^{-1}$ we get the Bessel function of the first kind of order n from the even values $j = 2k$:

$$J_n(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!(n+k)!} \left(\frac{x}{2}\right)^{n+2k}.$$

14. Use the infinite series in the last problem to derive the derivative identities (5.71) and (5.61):

a. $\frac{d}{dx}\left[x^n J_n(x)\right] = x^n J_{n-1}(x)$.

b. $\frac{d}{dx}\left[x^{-n}J_n(x)\right] = -x^{-n}J_{n+1}(x)$.

15. Prove the following identities based on those in the last problem.

a. $J_{p-1}(x) + J_{p+1}(x) = \frac{2p}{x}J_p(x)$.
b. $J_{p-1}(x) - J_{p+1}(x) = 2J_p'(x)$.

16. Use the derivative identities of Bessel functions,(5.71)-(5.61), and integration by parts to show that

$$\int x^3 J_0(x)\, dx = x^3 J_1(x) - 2x^2 J_2(x) + C.$$

17. Use the generating function to find $J_n(0)$ and $J_n'(0)$.

18. Bessel functions $J_p(\lambda x)$ are solutions of $x^2y'' + xy' + (\lambda^2x^2 - p^2)y = 0$. Assume that $x \in (0,1)$ and that $J_p(\lambda) = 0$ and $J_p(0)$ is finite.

a. Show that this equation can be written in the form

$$\frac{d}{dx}\left(x\frac{dy}{dx}\right) + (\lambda^2 x - \frac{p^2}{x})y = 0.$$

This is the standard Sturm-Liouville form for Bessel's equation.

b. Prove that

$$\int_0^1 xJ_p(\lambda x)J_p(\mu x)\, dx = 0, \quad \lambda \neq \mu$$

by considering

$$\int_0^1 \left[J_p(\mu x)\frac{d}{dx}\left(x\frac{d}{dx}J_p(\lambda x)\right) - J_p(\lambda x)\frac{d}{dx}\left(x\frac{d}{dx}J_p(\mu x)\right)\right] dx.$$

Thus, the solutions corresponding to different eigenvalues (λ, μ) are orthogonal.

c. Prove that

$$\int_0^1 x\left[J_p(\lambda x)\right]^2\, dx = \frac{1}{2}J_{p+1}^2(\lambda) = \frac{1}{2}J_p'^2(\lambda).$$

19. We can rewrite Bessel functions, $J_\nu(x)$, in a form which will allow the order to be non-integer by using the gamma function. You will need the results from Problem 12b for $\Gamma\left(k + \frac{1}{2}\right)$.

a. Extend the series definition of the Bessel function of the first kind of order ν, $J_\nu(x)$, for $\nu \geq 0$ by writing the series solution for $y(x)$ in Problem 13 using the gamma function.
b. Extend the series to $J_{-\nu}(x)$, for $\nu \geq 0$. Discuss the resulting series and what happens when ν is a positive integer.

c. Use these results to obtain the closed form expressions

$$J_{1/2}(x) = \sqrt{\frac{2}{\pi x}} \sin x,$$

$$J_{-1/2}(x) = \sqrt{\frac{2}{\pi x}} \cos x.$$

d. Use the results in part c with the recursion formula for Bessel functions to obtain a closed form for $J_{3/2}(x)$.

20. In this problem you will derive the expansion

$$x^2 = \frac{c^2}{2} + 4\sum_{j=2}^{\infty} \frac{J_0(\alpha_j x)}{\alpha_j^2 J_0(\alpha_j c)}, \quad 0 < x < c,$$

where the α_j's are the positive roots of $J_1(\alpha c) = 0$, by following the below steps.

a. List the first five values of α for $J_1(\alpha c) = 0$ using the Table 5.3 and Figure 5.8. [Note: Be careful determining α_1.]

b. Show that $\|J_0(\alpha_1 x)\|^2 = \frac{c^2}{2}$. Recall,

$$\|J_0(\alpha_j x)\|^2 = \int_0^c x J_0^2(\alpha_j x)\, dx.$$

c. Show that $\|J_0(\alpha_j x)\|^2 = \frac{c^2}{2}\left[J_0(\alpha_j c)\right]^2, j = 2, 3, \ldots$. (This is the most involved step.) First note from Problem 18 that $y(x) = J_0(\alpha_j x)$ is a solution of

$$x^2 y'' + xy' + \alpha_j^2 x^2 y = 0.$$

i. Verify the Sturm-Liouville form of this differential equation: $(xy')' = -\alpha_j^2 xy$.

ii. Multiply the equation in part i. by $y(x)$ and integrate from $x = 0$ to $x = c$ to obtain

$$\begin{aligned}\int_0^c (xy')' y\, dx &= -\alpha_j^2 \int_0^c xy^2\, dx \\ &= -\alpha_j^2 \int_0^c x J_0^2(\alpha_j x)\, dx. \end{aligned} \quad (5.105)$$

iii. Noting that $y(x) = J_0(\alpha_j x)$, integrate the left hand side by parts and use the following to simplify the resulting equation.

1. $J_0'(x) = -J_1(x)$ from Equation (5.61).
2. Equation (5.64).
3. $J_2(\alpha_j c) + J_0(\alpha_j c) = 0$ from Equation (5.62).

iv. Now you should have enough information to complete this part.

d. Use the results from parts b and c to derive the expansion coefficients for

$$x^2 = \sum_{j=1}^{\infty} c_j J_0(\alpha_j x)$$

in order to obtain the desired expansion.

21. Prove the if $u(x)$ and $v(x)$ satisfy the general homogeneous boundary conditions

$$\begin{aligned} \alpha_1 u(a) + \beta_1 u'(a) &= 0, \\ \alpha_2 u(b) + \beta_2 u'(b) &= 0 \end{aligned} \tag{5.106}$$

at $x = a$ and $x = b$, then

$$p(x)[u(x)v'(x) - v(x)u'(x)]_{x=a}^{x=b} = 0.$$

22. Prove Green's Identity $\int_a^b (u\mathcal{L}v - v\mathcal{L}u)\,dx = [p(uv' - vu')]|_a^b$ for the general Sturm-Liouville operator $\mathcal{L}$.

23. Find the adjoint operator and its domain for $Lu = u'' + 4u' - 3u$, $u'(0) + 4u(0) = 0$, $u'(1) + 4u(1) = 0$.

24. Show that a Sturm-Liouville operator with periodic boundary conditions on $[a, b]$ is self-adjoint if and only if $p(a) = p(b)$. [Recall, periodic boundary conditions are given as $u(a) = u(b)$ and $u'(a) = u'(b)$.]

25. The Hermite differential equation is given by $y'' - 2xy' + \lambda y = 0$. Rewrite this equation in self-adjoint form. From the Sturm-Liouville form obtained, verify that the differential operator is self adjoint on $(-\infty, \infty)$. Give the integral form for the orthogonality of the eigenfunctions.

26. Find the eigenvalues and eigenfunctions of the given Sturm-Liouville problems.

a. $y'' + \lambda y = 0$, $y'(0) = 0 = y'(\pi)$.
b. $(xy')' + \frac{\lambda}{x} y = 0$, $y(1) = y(e^2) = 0$.

27. The eigenvalue problem $x^2 y'' - \lambda x y' + \lambda y = 0$ with $y(1) = y(2) = 0$ is not a Sturm-Liouville eigenvalue problem. Show that none of the eigenvalues are real by solving this eigenvalue problem.

28. In Example 5.15 we found a bound on the lowest eigenvalue for the given eigenvalue problem.

a. Verify the computation in the example.
b. Apply the method using

$$y(x) = \begin{cases} x, & 0 < x < \frac{1}{2} \\ 1 - x, & \frac{1}{2} < x < 1. \end{cases}$$

Is this an upper bound on λ_1

c. Use the Rayleigh quotient to obtain a good upper bound for the lowest eigenvalue of the eigenvalue problem: $\phi'' + (\lambda - x^2)\phi = 0$, $\phi(0) = 0$, $\phi'(1) = 0$.

29. Use the method of eigenfunction expansions to solve the problem:

$$y'' + 4y = x^2, \quad y(0) = y(1) = 0.$$

30. Determine the solvability conditions for the nonhomogeneous boundary value problem: $u'' + 4u = f(x)$, $u(0) = \alpha$, $u'(1) = \beta$.