3
Generalized Fourier Series and Function Spaces

"Understanding is, after all, what science is all about and science is a great deal more than mindless computation." Sir Roger Penrose (1931-)

In this chapter we provide a glimpse into more general notions for generalized Fourier series and the convergence of Fourier series. It is useful to think about the general context in which one finds oneself when discussing Fourier series and 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 this chapter we touch a little on these ideas, leaving some of the deeper results for more advanced courses.

3.1 Finite Dimensional Vector Spaces

Much of the discussion and terminology that we will use comes from the theory of vector spaces . Until now you may only have dealt with finite dimensional vector spaces. Even then, you might only be comfortable with two and three dimensions. We will review a little of what we know about finite dimensional spaces so that we can introduce more general function spaces later.

The notion of a vector space is a generalization of three dimensional vectors and operations on them. In three dimensions, we have objects called vectors,[1] which are represented by arrows of a specific length and pointing

[1] In introductory physics one defines a vector as any quantity that has both magnitude and direction.

in a given direction. To each vector, we can associate a point in a three dimensional Cartesian system. We just attach the tail of the vector $\mathbf{v}$ to the origin and the head lands at the point (x, y, z).[2] We then use unit vectors $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ along the coordinate axes to write

[2] In multivariate calculus one concentrates on the component form of vectors. These representations are easily generalized as we will see.

$$\mathbf{v} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}.$$

Having defined vectors, we then learned how to add vectors and multiply vectors by numbers, or scalars. We then learned that there were two types of multiplication of vectors. We could multiply them to get a scalar or a vector. This led to dot products and cross products, respectively. The dot product is useful for determining the length of a vector, the angle between two vectors, if the vectors are perpendicular, or projections of one vector onto another. The cross product is used to produce orthogonal vectors, areas of parallelograms, and volumes of parallelepipeds.

In physics you first learned about vector products when you defined work, $W = \mathbf{F} \cdot \mathbf{r}$. Cross products were useful in describing things like torque, $\tau = \mathbf{r} \times \mathbf{F}$, or the force on a moving charge in a magnetic field, $\mathbf{F} = q\mathbf{v} \times \mathbf{B}$. We will return to these more complicated vector operations later when we need them.

Properties and definition of vector spaces.

The properties three dimensional vectors are generalized to spaces of more than three dimensions in linear algebra courses. The properties roughly outlined above need to be preserved. So, we will start with a space of vectors and the operations of addition and scalar multiplication. We will need a set of scalars, which generally come from some field. However, in our applications the field will either be the set of real numbers or the set of complex numbers.

A field is a set together with two operations, usually addition and multiplication, such that we have

- Closure under addition and multiplication
- Associativity of addition and multiplication
- Commutativity of addition and multiplication
- Additive and multiplicative identity
- Additive and multiplicative inverses
- Distributivity of multiplication over addition

A vector space V over a field F is a set that is closed under addition and scalar multiplication and satisfies the following conditions:

For any $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and $a, b \in F$

1. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$.
2. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$.
3. There exists a $\mathbf{0}$ such that $\mathbf{0} + \mathbf{v} = \mathbf{v}$.
4. There exists an additive inverse, $-\mathbf{v}$, such that $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$.

   There are also several distributive properties:
5. $a(b\mathbf{v}) = (ab)\mathbf{v}$.
6. $(a + b)\mathbf{v} = a\mathbf{v} + b\mathbf{v}$.
7. $a(\mathbf{u} + \mathbf{v}) = a\mathbf{u} + a\mathbf{v}$.
8. There exists a multiplicative identity, 1, such that $1(\mathbf{v}) = \mathbf{v}$.

For now, we will restrict our examples to two and three dimensions and the field will consist of the set of real numbers.

Basis vectors.

In three dimensions the unit vectors $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ play an important role. Any vector in the three dimensional space can be written as a linear combination of these vectors,

$$\mathbf{v} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}.$$

In fact, given any three non-coplanar vectors, $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$, all vectors can be written as a linear combination of those vectors,

$$\mathbf{v} = c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + c_3\mathbf{a}_3.$$

Such vectors are said to span the space and are called a basis for the space.

We can generalize these ideas. In an n-dimensional vector space any vector in the space can be represented as the sum over n linearly independent vectors (the equivalent of non-coplanar vectors). Such a linearly independent set of vectors $\{\mathbf{v}_j\}_{j=1}^n$ satisfies the condition

n-dimensional vector spaces.

Linearly independent vectors.

$$\sum_{j=1}^{n} c_j\mathbf{v}_j = \mathbf{0} \quad \Leftrightarrow \quad c_j = 0.$$

Note that we will often use summation notation instead of writing out all of the terms in the sum. Also, the symbol $\Leftrightarrow$ means "if and only if," or "is equivalent to." Each side of the symbol implies the other side.

Now we can define a basis for an n-dimensional vector space. We begin with the standard basis in an n-dimensional vector space. It is a generalization of the standard basis in three dimensions ($\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$).

The standard basis vectors, $\mathbf{e}_k$ are a natural generalization of $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$.

We define the standard basis with the notation

$$\mathbf{e}_k = (0, \ldots, 0, \underbrace{1}_{k\text{th space}}, 0, \ldots, 0), \quad k = 1, \ldots, n. \tag{3.1}$$

We can expand any $\mathbf{v} \in V$ as

$$\mathbf{v} = \sum_{k=1}^{n} v_k\mathbf{e}_k, \tag{3.2}$$

where the v_k's are called the components of the vector in this basis. Sometimes we will write $\mathbf{v}$ as an n-tuple $(v_1, v_2, \ldots, v_n)$. This is similar to the ambiguous use of (x, y, z) to denote both vectors and points in the three dimensional space.

The only other thing we will need at this point is to generalize the dot product. Recall that there are two forms for the dot product in three dimensions. First, one has that

For more general vector spaces the term inner product is used to generalize the notions of dot and scalar products as we will see below.

$$\mathbf{u} \cdot \mathbf{v} = uv\cos\theta, \tag{3.3}$$

where u and v denote the length of the vectors. The other form is the component form:

$$\mathbf{u} \cdot \mathbf{v} = u_1v_1 + u_2v_2 + u_3v_3 = \sum_{k=1}^{3} u_kv_k. \tag{3.4}$$

Of course, this form is easier to generalize. So, we define the scalar product between two n-dimensional vectors as

$$\langle \mathbf{u}, \mathbf{v} \rangle = \sum_{k=1}^{n} u_k v_k. \tag{3.5}$$

Actually, there are a number of notations that are used in other texts. One can write the scalar product as $(\mathbf{u}, \mathbf{v})$ or even in the Dirac bra-ket notation[3] $\langle \mathbf{u} | \mathbf{v} \rangle$.

[3] The bra-ket notation was introduced by Paul Adrien Maurice Dirac (1902-1984) in order to facilitate computations of inner products in quantum mechanics. In the notation $\langle \mathbf{u}|\mathbf{v}\rangle$, $\langle \mathbf{u}|$ is the bra and $|\mathbf{v}\rangle$ is the ket. The kets live in a vector space and represented by column vectors with respect to a given basis. The bras live in the dual vector space and are represented by row vectors. The correspondence between bra and kets is $|\mathbf{v}\rangle = \overline{|\mathbf{v}\rangle^T}$. One can operate on kets, $A|\mathbf{v}\rangle$, and make sense out of operations like$\langle \mathbf{u}|A|\mathbf{v}\rangle$, which are used to obtain expectation values associated with the operator. Finally, the outer product, $|\mathbf{v}\rangle\langle\mathbf{v}|$ is used to perform vector space projections.

We note that the (real) scalar product satisfies some simple properties. For vectors $\mathbf{v}$, $\mathbf{w}$ and real scalar α we have

1. $\langle \mathbf{v}, \mathbf{v} \rangle \geq 0$ and $\langle \mathbf{v}, \mathbf{v} \rangle = 0$ if and only if $\mathbf{v} = \mathbf{0}$.
2. $\langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{w}, \mathbf{v} \rangle$.
3. $\langle \alpha\mathbf{v}, \mathbf{w} \rangle = \alpha\langle \mathbf{v}, \mathbf{w} \rangle$.

While it does not always make sense to talk about angles between general vectors in higher dimensional vector spaces, there is one concept that is useful. It is that of orthogonality, which in three dimensions is another way of saying the vectors are perpendicular to each other. So, we also say that vectors $\mathbf{u}$ and $\mathbf{v}$ are orthogonal if and only if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$. If $\{\mathbf{a}_k\}_{k=1}^n$, is a set of basis vectors such that

$$\langle \mathbf{a}_j, \mathbf{a}_k \rangle = 0, \quad k \neq j,$$

Orthogonal basis vectors.

then it is called an *orthogonal basis*.

If in addition each basis vector is a unit vector, then one has an orthonormal basis. This generalization of the unit basis can be expressed more compactly. We will denote such a basis of unit vectors by $\mathbf{e}_j$ for $j = 1 \ldots n$. Then,

$$\langle \mathbf{e}_j, \mathbf{e}_k \rangle = \delta_{jk}, \tag{3.6}$$

where we have introduced the Kronecker delta (named after Leopold Kronecker (1823-1891))

$$\delta_{jk} \equiv \begin{cases} 0, & j \neq k \\ 1, & j = k \end{cases} \tag{3.7}$$

Normalization of vectors.

The process of making vectors have unit length is called normalization. This is simply done by dividing by the length of the vector. Recall that the length of a vector, $\mathbf{v}$, is obtained as $v = \sqrt{\mathbf{v} \cdot \mathbf{v}}$. So, if we want to find a unit vector in the direction of $\mathbf{v}$, then we simply normalize it as

$$\hat{\mathbf{v}} = \frac{\mathbf{v}}{v}.$$

Notice that we used a hat to indicate that we have a unit vector. Furthermore, if $\{\mathbf{a}_j\}_{j=1}^n$, is a set of orthogonal basis vectors, then

$$\hat{\mathbf{a}}_j = \frac{\mathbf{a}_i}{\sqrt{\langle \mathbf{a}_j, \mathbf{a}_j \rangle}}, \quad j = 1 \ldots n.$$

Example 3.1. *Find the angle between the vectors* $\mathbf{u} = (-2, 1, 3)$ *and* $\mathbf{v} = (1, 0, 2)$. *we need the lengths of each vector,*

$$u = \sqrt{(-2)^2 + 1^2 + 3^2} = \sqrt{14},$$

$$v = \sqrt{1^2 + 0^2 + 2^2} = \sqrt{5}.$$

We also need the scalar product of these vectors,

$$\mathbf{u} \cdot \mathbf{v} = -2 + 6 = 4.$$

This gives

$$\cos\theta = \frac{\mathbf{u} \cdot \mathbf{v}}{uv} = \frac{4}{\sqrt{5}\sqrt{14}}.$$

So, $\theta = 61.4^\circ$.

Example 3.2. *Normalize the vector* $\mathbf{v} = 2\mathbf{i} + \mathbf{j} - 2\mathbf{k}$.

The length of the vector is $v = \sqrt{2^2 + 1^2 + (-2)^2} = \sqrt{9} = 3$. *So, the unit vector in the direction of* $\mathbf{v}$ *is* $\hat{\mathbf{v}} = \frac{2}{3}\mathbf{i} + \frac{1}{3}\mathbf{j} - \frac{2}{3}\mathbf{k}$.

Let $\{\mathbf{a}_k\}_{k=1}^n$, be a set of orthogonal basis vectors for vector space V. We know that any vector $\mathbf{v}$ can be represented in terms of this basis, $\mathbf{v} = \sum_{k=1}^n v_k \mathbf{a}_k$. If we know the basis and vector, can we find the components, v_k? The answer is yes. We can use the scalar product of $\mathbf{v}$ with each basis element $\mathbf{a}_j$. Using the properties of the scalar product, we have for $j = 1, \ldots, n$

$$\begin{aligned} \langle \mathbf{a}_j, \mathbf{v} \rangle &= \langle \mathbf{a}_j, \sum_{k=1}^n v_k \mathbf{a}_k \rangle \\ &= \sum_{k=1}^n v_k \langle \mathbf{a}_j, \mathbf{a}_k \rangle. \end{aligned} \tag{3.8}$$

Since we know the basis elements, we can easily compute the numbers

$$A_{jk} \equiv \langle \mathbf{a}_j, \mathbf{a}_k \rangle$$

and

$$b_j \equiv \langle \mathbf{a}_j, \mathbf{v} \rangle.$$

Therefore, the system (3.8) for the v_k's is a linear algebraic system, which takes the form

$$b_j = \sum_{k=1}^n A_{jk} v_k. \tag{3.9}$$

We can write this set of equations in a more compact form. The set of numbers A_{jk}, $j, k = 1, \ldots n$ are the elements of an $n \times n$ matrix A with A_{jk} being an element in the jth row and kth column. We write such matrices with the n^2 entries A_{ij} as

$$A = \begin{pmatrix} A_{11} & A_{12} & \ldots & A_{1n} \\ A_{21} & A_{22} & \ldots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \ldots & A_{nn} \end{pmatrix}. \tag{3.10}$$

Also, v_j and b_j can be written as column vectors, $\mathbf{v}$ and $\mathbf{b}$, respectively. Thus, system (3.8) can be written in matrix form as

$$A\mathbf{v} = \mathbf{b}.$$

However, if the basis is orthogonal, then the matrix $A_{jk} \equiv \langle \mathbf{a}_j, \mathbf{a}_k \rangle$ is diagonal,

$$A = \begin{pmatrix} A_{11} & 0 & \dots & \dots & 0 \\ 0 & A_{22} & \ddots & \ddots & 0 \\ \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \dots & 0 & A_{nn} \end{pmatrix}. \tag{3.11}$$

and the system is easily solvable. Recall that two vectors are orthogonal if and only if

$$\langle \mathbf{a}_i, \mathbf{a}_j \rangle = 0, \quad i \neq j. \tag{3.12}$$

Thus, in this case we have that

$$\langle \mathbf{a}_j, \mathbf{v} \rangle = v_j \langle \mathbf{a}_j, \mathbf{a}_j \rangle, \quad j = 1, \dots, n. \tag{3.13}$$

or

$$v_j = \frac{\langle \mathbf{a}_j, \mathbf{v} \rangle}{\langle \mathbf{a}_j, \mathbf{a}_j \rangle}. \tag{3.14}$$

In fact, if the basis is orthonormal, i.e., the basis consists of an orthogonal set of unit vectors, then A is the identity matrix and the solution takes on a simpler form:

$$v_j = \langle \mathbf{a}_j, \mathbf{v} \rangle. \tag{3.15}$$

Example 3.3. *Consider the set of vectors* $\mathbf{a}_1 = \mathbf{i} + \mathbf{j}$ *and* $\mathbf{a}_2 = \mathbf{i} - 2\mathbf{j}$.

1. *Determine the matrix elements* $A_{jk} = \langle \mathbf{a}_j, \mathbf{a}_k \rangle$.
2. *Is this an orthogonal basis?*
3. *Expand the vector* $\mathbf{v} = 2\mathbf{i} + 3\mathbf{j}$ *in the basis* $\{\mathbf{a}_1, \mathbf{a}_2\}$.

First, we compute the matrix elements of A:

$$\begin{aligned} A_{11} &= \langle \mathbf{a}_1, \mathbf{a}_1 \rangle = 2 \\ A_{12} &= \langle \mathbf{a}_1, \mathbf{a}_2 \rangle = -1 \\ A_{21} &= \langle \mathbf{a}_2, \mathbf{a}_1 \rangle = -1 \\ A_{22} &= \langle \mathbf{a}_2, \mathbf{a}_2 \rangle = 5 \end{aligned} \tag{3.16}$$

So,

$$A = \begin{pmatrix} 2 & -1 \\ -1 & 5 \end{pmatrix}.$$

Since $A_{12} = A_{21} \neq 0$, *the vectors are not orthogonal. However, they are linearly independent. Obviously, if* $c_1 = c_2 = 0$, *then the linear combination*

$c_1\mathbf{a}_1 + c_2\mathbf{a}_2 = \mathbf{0}$. *Conversely, we assume that* $c_1\mathbf{a}_1 + c_2\mathbf{a}_2 = \mathbf{0}$ *and solve for the coefficients. Inserting the given vectors, we have*

$$\begin{aligned} \mathbf{0} &= c_1(\mathbf{i}+\mathbf{j}) + c_2(\mathbf{i}-2\mathbf{j}) \\ &= (c_1+c_2)\mathbf{i} + (c_1-2c_2)\mathbf{j}. \end{aligned} \tag{3.17}$$

This implies that

$$\begin{aligned} c_1 + c_2 &= 0 \\ c_1 - 2c_2 &= 0. \end{aligned} \tag{3.18}$$

Solving this system, one has $c_1 = 0$, $c_2 = 0$. *Therefore, the two vectors are linearly independent.*

In order to determine the components of $\mathbf{v}$ *with respect to the new basis, we need to set up the system (3.8) and solve for the* v_k*'s. We have first,*

$$\begin{aligned} \mathbf{b} &= \begin{pmatrix} \langle \mathbf{a}_1, \mathbf{v}\rangle \\ \langle \mathbf{a}_2, \mathbf{v}\rangle \end{pmatrix} \\ &= \begin{pmatrix} \langle \mathbf{i}+\mathbf{j}, 2\mathbf{i}+3\mathbf{j}\rangle \\ \langle \mathbf{i}-2\mathbf{j}, 2\mathbf{i}+3\mathbf{j}\rangle \end{pmatrix} \\ &= \begin{pmatrix} 5 \\ -4 \end{pmatrix}. \end{aligned} \tag{3.19}$$

So, now we have to solve the system $A\mathbf{v} = \mathbf{b}$ *for* $\mathbf{v}$:

$$\begin{pmatrix} 2 & -1 \\ -1 & 5 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 5 \\ -4 \end{pmatrix}. \tag{3.20}$$

We can solve this with matrix methods, $\mathbf{v} = A^{-1}\mathbf{b}$, *or rewrite it as a system of two equations and two unknowns as*

$$\begin{aligned} 2v_1 - v_2 &= 5 \\ -v_1 + 5v_2 &= -4. \end{aligned} \tag{3.21}$$

The solution of this set of algebraic equations is $v_1 = \frac{7}{3}$, $v_2 = -\frac{1}{3}$. *Thus,* $\mathbf{v} = \frac{7}{3}\mathbf{a}_1 - \frac{1}{3}\mathbf{a}_2$. *We will return later to using matrix methods to solve such systems.*

3.2 *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 already did 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 did 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 their properties. 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 will see, there are many types of function spaces . In order to view these spaces as vector spaces, we must 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. An inner product $\langle , \rangle$ 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. $\langle v, v \rangle \geq 0$ and $\langle v, v \rangle = 0$ iff $v = 0$.
2. $\langle v, w \rangle = \langle w, v \rangle$.
3. $\langle \alpha v, w \rangle = \alpha \langle v, w \rangle$.
4. $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$.

A real vector space equipped with the above inner product leads to what is called a real inner product space. For complex inner product spaces, the above properties hold with the third property replaced with $\langle v, w \rangle = \overline{\langle w, v \rangle}$.

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

$$\langle f, g \rangle = \int_a^b f(x) g(x) \sigma(x)\, dx. \tag{3.22}$$

The space of square integrable functions.

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

Now that we have function 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 Fourier series expansions? Does it 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

$$\begin{aligned} \langle \phi_j, f \rangle &= \langle \phi_j, \sum_{n=1}^{\infty} c_n \phi_n \rangle \\ &= \sum_{n=1}^{\infty} c_n \langle \phi_j, \phi_n \rangle. \end{aligned} \tag{3.23}$$

If the basis is an orthogonal basis, then we have

$$\langle \phi_j, \phi_n \rangle = N_j \delta_{jn}, \tag{3.24}$$

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

$$\delta_{ij} = \begin{cases} 0, & i \neq j \\ 1, & i = j. \end{cases} \tag{3.25}$$

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 $c_j = \langle \phi_j, f \rangle / \langle \phi_j, \phi_j \rangle$.

$$\begin{aligned} \langle \phi_j, f \rangle &= \sum_{n=1}^{\infty} c_n \langle \phi_j, \phi_n \rangle \\ &= \sum_{n=1}^{\infty} c_n N_j \delta_{jn}. \end{aligned} \tag{3.26}$$

Expanding the sum, we see that the Kronecker delta picks out one nonzero term:

$$\begin{aligned} \langle \phi_j, f \rangle &= c_1 N_j \delta_{j1} + c_2 N_j \delta_{j2} + \ldots + c_j N_j \delta_{jj} + \ldots \\ &= c_j N_j. \end{aligned} \tag{3.27}$$

So, the expansion coefficients are

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

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

$$c_n = \frac{\langle \phi_n, f\rangle}{\langle \phi_n, \phi_n\rangle}.$$

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 3.4. *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 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 $n \neq m$

$$\langle \phi_n, \phi_m\rangle = \int_{-\pi}^{\pi} \sin nx \sin mx \, dx = \pi\delta_{nm}. \tag{3.28}$$

Therefore, the set $\phi_n(x) = \sin nx$ for $n = 1, 2, \ldots$ is an orthogonal set of functions on the interval $[-\pi, \pi]$.

We determine the expansion coefficients using

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

Does this result look familiar?

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

$$\|f\| = \sqrt{\langle f, f\rangle}.$$

Note that there are many types of norms, but this induced norm will be sufficient.[4]

For this example, the norms of the basis functions are $\|\phi_n\| = \sqrt{\pi}$. Defining $\psi_n(x) = \frac{1}{\sqrt{\pi}}\phi_n(x)$, we can normalize the ϕ_n's and have obtained an orthonormal basis of functions on $[-\pi, \pi]$.

We can also use the normalized basis to determine the expansion coefficients. In this case we have

$$b_n = \frac{\langle \psi_n, f\rangle}{N_n} = \langle \psi_n, f\rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x) \sin nx \, dx.$$

[4] 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:
$$\|f\| = \left(\int [f(x)]^p \, dx\right)^{\frac{1}{p}}.$$

3.3 *Classical Orthogonal Polynomials*

There are other basis functions that can be used to develop series representations of functions. 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[5] 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.

[5] **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 a basis of powers of x, $\{1, x, x^2, \ldots\}$, since these expansions are just Maclaurin series representations of the functions about $x = 0$,

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

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,

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

The Gram-Schmidt Orthogonalization Process.

Since we have found that orthogonal bases have been useful in determining the coefficients for expansions of given functions, we might 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 out this so-called Gram-Schmidt Orthogonalization Process. We will review this process for finite dimensional vectors and then generalize to function spaces.

Let's assume that we have three vectors that span the usual three-dimensional space, $\mathbf{R}^3$, given by $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$ and shown in Figure 3.1. We seek an orthogonal basis $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_3$, beginning one vector at a time.

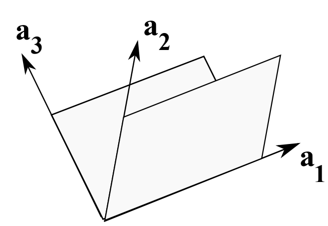

Figure 3.1: The basis $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$, of $\mathbf{R}^3$.

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 3.2 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

$$\mathbf{e}_2 = \mathbf{a}_2 - \mathbf{pr}_1\mathbf{a}_2. \tag{3.29}$$

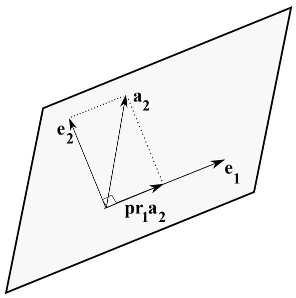

Figure 3.2: A plot of the vectors $\mathbf{e}_1$, $\mathbf{a}_2$, and $\mathbf{e}_2$ needed to find the projection of $\mathbf{a}_2$, on $\mathbf{e}_1$.

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

$$\mathbf{pr}_1\mathbf{a}_2 = \frac{\mathbf{a}_2 \cdot \mathbf{e}_1}{e_1^2}\mathbf{e}_1. \tag{3.30}$$

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.

Combining Equations (3.29) and (3.30), we find that

$$\mathbf{e}_2 = \mathbf{a}_2 - \frac{\mathbf{a}_2 \cdot \mathbf{e}_1}{e_1^2}\mathbf{e}_1. \tag{3.31}$$

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

$$\begin{aligned} \mathbf{e}_2 \cdot \mathbf{e}_1 &= \mathbf{a}_2 \cdot \mathbf{e}_1 - \frac{\mathbf{a}_2 \cdot \mathbf{e}_1}{e_1^2}\mathbf{e}_1 \cdot \mathbf{e}_1 \\ &= \mathbf{a}_2 \cdot \mathbf{e}_1 - \mathbf{a}_2 \cdot \mathbf{e}_1 = 0. \end{aligned} \tag{3.32}$$

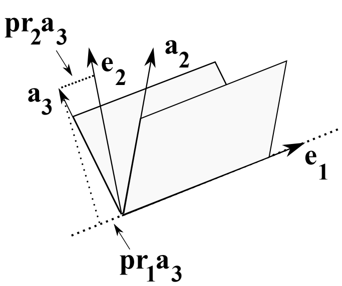

Figure 3.3: A plot of vectors for determining $\mathbf{e}_3$.

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 3.3. Thus, we can see that

$$\mathbf{e}_3 = \mathbf{a}_3 - \frac{\mathbf{a}_3 \cdot \mathbf{e}_1}{e_1^2}\mathbf{e}_1 - \frac{\mathbf{a}_3 \cdot \mathbf{e}_2}{e_2^2}\mathbf{e}_2. \tag{3.33}$$

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 defining

$$\boxed{\mathbf{e}_n = \mathbf{a}_n - \sum_{j=1}^{n-1} \frac{\mathbf{a}_n \cdot \mathbf{e}_j}{e_j^2}\mathbf{e}_j, \quad n = 2, 3, \ldots, N.} \tag{3.34}$$

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

$$\boxed{\phi_n(x) = f_n(x) - \sum_{j=0}^{n-1} \frac{\langle f_n, \phi_j \rangle}{\|\phi_j\|^2}\phi_j(x), \quad n = 1, 2, \ldots.} \tag{3.35}$$

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

$$\langle f, g\rangle = \int_a^b f(x)g(x)\sigma(x)\,dx. \tag{3.36}$$

Note the similarity between the orthogonal basis in Equation (3.35) and the expression for the finite dimensional case in Equation (3.34).

Example 3.5. *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 the new basis, but we will hold off doing that for now.

Now we compute the second basis element:

$$\begin{aligned}\phi_1(x) &= f_1(x) - \frac{\langle f_1, \phi_0\rangle}{\|\phi_0\|^2}\phi_0(x)\\ &= x - \frac{\langle x, 1\rangle}{\|1\|^2}1 = x,\end{aligned} \tag{3.37}$$

since $\langle x, 1\rangle$ is the integral of an odd function over a symmetric interval.

For $\phi_2(x)$, we have

$$\begin{aligned}\phi_2(x) &= f_2(x) - \frac{\langle f_2, \phi_0\rangle}{\|\phi_0\|^2}\phi_0(x) - \frac{\langle f_2, \phi_1\rangle}{\|\phi_1\|^2}\phi_1(x)\\ &= x^2 - \frac{\langle x^2, 1\rangle}{\|1\|^2}1 - \frac{\langle x^2, x\rangle}{\|x\|^2}x\\ &= x^2 - \frac{\int_{-1}^1 x^2\,dx}{\int_{-1}^1 dx}\\ &= x^2 - \frac{1}{3}.\end{aligned} \tag{3.38}$$

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 polynomials, $P_n(x)$. Thus,

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

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

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

The set of Legendre[6] 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 3.1.

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

Table 3.1: 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)	$(1-x)^\nu(1-x)^\mu$

3.4 *Fourier-Legendre Series*

In the last chapter we saw how useful Fourier series expansions were for solving the heat and wave equations. In the study of partial differential equations in higher dimensions and one finds that problems with spherical symmetry can 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 $(-1,1)$:

$$f(x) \sim \sum_{n=0}^{\infty} c_n P_n(x). \tag{3.39}$$

As with Fourier trigonometric series, we can determine the expansion coefficients by multiplying both sides of Equation (3.39) by $P_m(x)$ and integrating for $x \in [-1,1]$. Orthogonality gives the usual form for the generalized Fourier coefficients,

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

We will later show that

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

Therefore, the Fourier-Legendre coefficients are

$$c_n = \frac{2n+1}{2}\int_{-1}^{1} f(x)P_n(x)\,dx. \tag{3.40}$$

3.4.1 Properties of Legendre Polynomials

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:

$$P_n(x) = \frac{1}{2^n n!}\frac{d^n}{dx^n}(x^2-1)^n, \quad n \in N_0. \tag{3.41}$$

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 3.6. *Determine $P_2(x)$ from the Rodrigues Formula:*

$$\begin{aligned} P_2(x) &= \frac{1}{2^2 2!}\frac{d^2}{dx^2}(x^2-1)^2 \\ &= \frac{1}{8}\frac{d^2}{dx^2}(x^4-2x^2+1) \\ &= \frac{1}{8}\frac{d}{dx}(4x^3-4x) \\ &= \frac{1}{8}(12x^2-4) \\ &= \frac{1}{2}(3x^2-1). \end{aligned} \tag{3.42}$$

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

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 3.2: Tabular computation of the Legendre polynomials using the Rodrigues Formula.

The first several Legendre polynomials are given in Table 3.2. In Figure 3.4 we show plots of these Legendre polynomials.

The 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

$$(n+1)P_{n+1}(x) = (2n+1)xP_n(x) - nP_{n-1}(x), \quad n = 1, 2, \ldots. \tag{3.43}$$

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

$$(2n-1)xP_{n-1}(x) = nP_n(x) + (n-1)P_{n-2}(x), \quad n = 1, 2, \ldots. \tag{3.44}$$

Example 3.7. *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 (3.43):

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

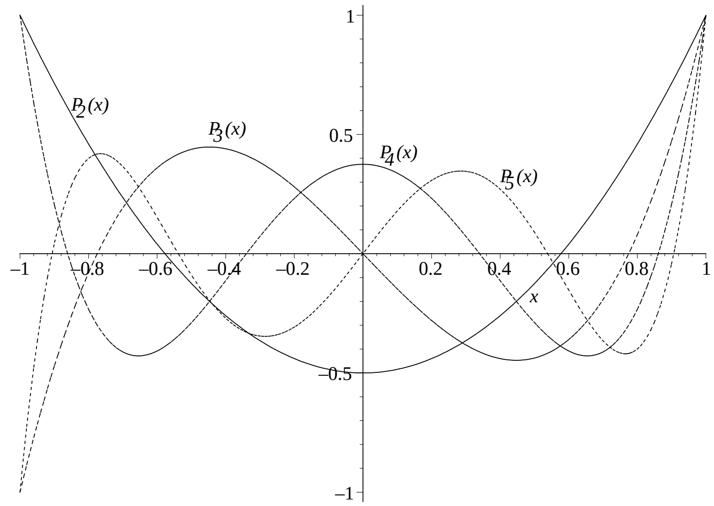

Figure 3.4: Plots of the Legendre polynomials $P_2(x)$, $P_3(x)$, $P_4(x)$, and $P_5(x)$.

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

For $n = 2$, *we have*

$$\begin{aligned} 3P_3(x) &= 5xP_2(x) - 2P_1(x) \\ &= \frac{5}{2}x(3x^2-1) - 2x \\ &= \frac{1}{2}(15x^3 - 9x). \end{aligned} \tag{3.45}$$

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

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.

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

Lemma 3.1. *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 the 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

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

Thus, the nth derivative is given by

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

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

$$\begin{aligned} \frac{[2n(2n-1)\dots(n+1)]}{2^n n!} &= \frac{[2n(2n-1)\dots(n+1)]}{2^n n!}\frac{n(n-1)\dots 1}{n(n-1)\dots 1} \\ &= \frac{1}{2^n n!}\frac{(2n)!}{n!}. \end{aligned} \tag{3.46}$$

□

Theorem 3.1. *Legendre polynomials satisfy the three-term recursion formula*

$$(2n-1)xP_{n-1}(x) = nP_n(x) + (n-1)P_{n-2}(x), \quad n = 1, 2, \ldots. \tag{3.47}$$

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

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

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 the expression $(2n-1)xP_{n-1}(x) - nP_n(x)$ 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 Legendre polynomials:

$$(2n-1)xP_{n-1}(x) - nP_n(x) = c_0P_0(x) + c_1P_1(x) + \ldots + c_{n-2}P_{n-2}(x). \tag{3.48}$$

Multiplying Equation (3.48) 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, $\int_{-1}^{1} xP_{n-1}(x)P_m(x)\,dx = 0$ for $m \leq n-3$.]

Thus, all these c_m's are zero, leaving Equation (3.48) as

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

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

3.4.2 *The Generating Function for Legendre Polynomials*

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

$$g(x,t) = \frac{1}{\sqrt{1-2xt+t^2}} = \sum_{n=0}^{\infty} P_n(x)t^n, \quad |x| \leq 1, |t| < 1. \tag{3.49}$$

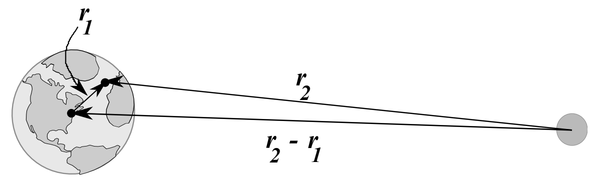

Figure 3.5: The position vectors used to describe the tidal force on the Earth due to the moon.

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 3.5. The tidal potential Φ is proportional to

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

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 3.8. *Evaluate $P_n(0)$ using the generating function. $P_n(0)$ is found by considering $g(0,t)$. Setting $x = 0$ in Equation (3.49), we have*

$$g(0,t) = \frac{1}{\sqrt{1+t^2}}$$

$$= \sum_{n=0}^{\infty} P_n(0)t^n$$
$$= P_0(0) + P_1(0)t + P_2(0)t^2 + P_3(0)t^3 + \dots. \quad (3.50)$$

We can use the binomial expansion to find the final answer. 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[7]

$$P_{2n}(0) = (-1)^n \frac{(2n-1)!!}{(2n)!!}, \quad (3.51)$$

where $n!!$ is the double factorial,

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

[7] 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 3.9. *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, $P_n(-1) = (-1)^n$.

Proof of the three-term recursion formula using the generating function.

Example 3.10. *Prove the three-term recursion formula,*

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

using the generating function.

We can also use the generating function to find recurrence relations. To prove the three term recursion (3.43) that we introduced above, then we need only differentiate the generating function with respect to t in Equation (3.49) 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:

$$\sum_{n=0}^{\infty} nP_n(x)t^{n-1} - \sum_{n=0}^{\infty} (2n+1)xP_n(x)t^n + \sum_{n=0}^{\infty} (n+1)P_n(x)t^{n+1} = 0. \quad (3.52)$$

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 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 (3.52). The second sum just needs the replacement $n = k$ and the last sum we re-index using $k = n + 1$. Therefore, Equation (3.52) becomes

$$\sum_{k=-1}^{\infty} (k+1)P_{k+1}(x)t^k - \sum_{k=0}^{\infty} (2k+1)xP_k(x)t^k + \sum_{k=1}^{\infty} kP_{k-1}(x)t^k = 0. \quad (3.53)$$

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

$$P_1(x) - xP_0(x) = 0. \quad (3.54)$$

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

$$\sum_{k=1}^{\infty} \left[(k+1)P_{k+1}(x) - (2k+1)xP_k(x) + kP_{k-1}(x)\right] t^k = 0. \quad (3.55)$$

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

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

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 that we list in the box below. Equation (3.56) was derived using the generating function. Differentiating it with respect to x, we find Equation (3.57). Equation (3.58) 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 (3.56), we can derive Equations (3.59) and (3.60). Adding and subtracting these equations yields Equations (3.61) and (3.62).

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

$$\begin{aligned}
(n+1)P_{n+1}(x) &= (2n+1)xP_n(x) - nP_{n-1}(x) && (3.56)\\
(n+1)P'_{n+1}(x) &= (2n+1)[P_n(x) + xP'_n(x)] - nP'_{n-1}(x) && (3.57)\\
P_n(x) &= P'_{n+1}(x) - 2xP'_n(x) + P'_{n-1}(x) && (3.58)\\
P'_{n-1}(x) &= xP'_n(x) - nP_n(x) && (3.59)\\
P'_{n+1}(x) &= xP'_n(x) + (n+1)P_n(x) && (3.60)\\
P'_{n+1}(x) + P'_{n-1}(x) &= 2xP'_n(x) + P_n(x). && (3.61)\\
P'_{n+1}(x) - P'_{n-1}(x) &= (2n+1)P_n(x). && (3.62)\\
(x^2-1)P'_n(x) &= nxP_n(x) - nP_{n-1}(x) && (3.63)
\end{aligned}$$

Finally, Equation (3.63) can be obtained using Equations (3.59) and (3.60). Just multiply Equation (3.59) by x,

$$x^2P'_n(x) - nxP_n(x) = xP'_{n-1}(x).$$

Now use Equation (3.60), but first replace n with $n-1$ to eliminate the $xP'_{n-1}(x)$ term:

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

Rearranging gives the Equation (3.63).

Example 3.11. *Use the generating function to prove*

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

Another use of the generating function is to obtain the normalization constant. 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.

The normalization constant.

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

$$\begin{aligned}
\frac{1}{1-2xt+t^2} &= \left[\sum_{n=0}^{\infty} P_n(x)t^n\right]^2\\
&= \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} P_n(x)P_m(x)t^{n+m}. && (3.64)
\end{aligned}$$

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

$$\int_{-1}^{1} \frac{dx}{1-2xt+t^2} = \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} t^{n+m} \int_{-1}^{1} P_n(x)P_m(x)\,dx$$

$$= \sum_{n=0}^{\infty} t^{2n} \int_{-1}^{1} P_n^2(x)\, dx. \tag{3.65}$$

However, one can show that[8]

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

Expanding this expression about $t = 0$, we obtain[9]

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

Comparing this result with Equation (3.65), we find that

$$\|P_n\|^2 = \int_{-1}^{1} P_n^2(x)\, dx = \frac{2}{2n+1}. \tag{3.66}$$

[8] You will need the integral

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

[9] You will need the series expansion

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

3.4.3 The Differential Equation for Legendre Polynomials

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. 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 (3.63) and using Equation (3.59) to simplify:

$$\begin{aligned} \frac{d}{dx}\left((x^2-1)P_n'(x)\right) &= nP_n(x) + nxP_n'(x) - nP_{n-1}'(x) \\ &= nP_n(x) + n^2 P_n(x) \\ &= n(n+1)P_n(x). \end{aligned} \tag{3.67}$$

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,

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

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

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.

3.4.4 Fourier-Legendre Series Examples

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

We simply need to compute

$$c_n = \frac{2n+1}{2} \int_{-1}^{1} x^3 P_n(x)\, dx. \tag{3.68}$$

We first note that

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

As a result, we have 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 polynomials. We note that x^3 is an odd function. So, $c_0 = 0$ and $c_2 = 0$.

This leaves us with only two coefficients to compute. We refer to Table 3.2 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 3.2:

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

We could have obtained this result without doing any integration. Write x^3 as a linear combination of $P_1(x)$ and $P_3(x)$:

$$\begin{aligned} x^3 &= c_1 x + \frac{1}{2} c_2 (5x^3 - 3x) \\ &= (c_1 - \frac{3}{2} c_2) x + \frac{5}{2} c_2 x^3. \end{aligned} \tag{3.69}$$

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 3.13. *Expand the Heaviside[10] function in a Fourier-Legendre series.*

The Heaviside function is defined as

$$H(x) = \begin{cases} 1, & x > 0, \\ 0, & x < 0. \end{cases} \tag{3.70}$$

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

$$\begin{aligned} c_n &= \frac{2n+1}{2} \int_{-1}^{1} f(x) P_n(x)\, dx \\ &= \frac{2n+1}{2} \int_{0}^{1} P_n(x)\, dx. \end{aligned} \tag{3.71}$$

[10] 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.

We can make use of identity (3.62),

$$P'_{n+1}(x) - P'_{n-1}(x) = (2n+1)P_n(x), \quad n > 0. \tag{3.72}$$

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*

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

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

$$c_n = \frac{1}{2}[P_{n-1}(0) - P_{n+1}(0)].$$

Since $P_n(0) = 0$ *for n odd, the* c_n*'s vanish for n even. Letting* $n = 2k-1$*, we re-index the sum, obtaining*

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

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

$$P_{2k}(0) = (-1)^k \frac{(2k-1)!!}{(2k)!!}. \tag{3.73}$$

Namely, we have

$$\begin{aligned} c_{2k-1} &= \frac{1}{2}[P_{2k-2}(0) - P_{2k}(0)] \\ &= \frac{1}{2}\left[(-1)^{k-1}\frac{(2k-3)!!}{(2k-2)!!} - (-1)^k \frac{(2k-1)!!}{(2k)!!}\right] \\ &= -\frac{1}{2}(-1)^k \frac{(2k-3)!!}{(2k-2)!!}\left[1 + \frac{2k-1}{2k}\right] \\ &= -\frac{1}{2}(-1)^k \frac{(2k-3)!!}{(2k-2)!!}\frac{4k-1}{2k}. \end{aligned} \tag{3.74}$$

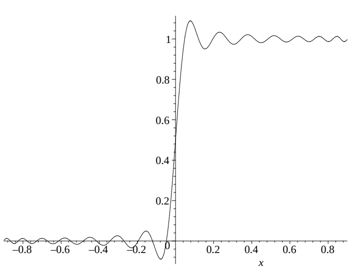

Figure 3.6: Sum of first 21 terms for Fourier-Legendre series expansion of Heaviside function.

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

$$f(x) \sim \frac{1}{2} - \frac{1}{2}\sum_{n=1}^{\infty}(-1)^n \frac{(2n-3)!!}{(2n-2)!!}\frac{4n-1}{2n}P_{2n-1}(x). \tag{3.75}$$

The sum of the first 21 terms of this series are shown in Figure 3.6. 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 2.5.]*

3.5 Gamma Function

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.

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.

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

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t}\, dt, \quad x > 0. \tag{3.76}$$

The Gamma function is a generalization of the factorial function and a plot is shown in Figure 3.7. In fact, we have

$$\Gamma(1) = 1$$

and

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

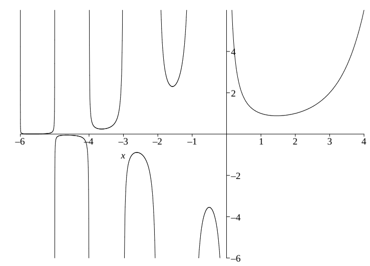

Figure 3.7: Plot of the Gamma function.

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

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

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 3.14. *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. Consider the integral

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

Then,

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

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

In Example 5.5 we show the more general result:

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

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

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.

3.6 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 also provide us with another orthogonal set of basis functions.

Bessel functions have a long history and were named after Friedrich Wilhelm Bessel (1784-1846).

The first occurrence of Bessel functions (zeroth order) was in the work of Daniel Bernoulli on heavy chains (1738). More general Bessel functions were studied by Leonhard Euler in 1781 and in his study of the vibrating membrane in 1764. Joseph Fourier found them in the study of heat conduction in solid cylinders and Siméon Poisson (1781-1840) in heat conduction of spheres (1823).

The history of Bessel functions, did not just originate in the study of the wave and heat 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.

You might have seen Bessel functions in a course on differential equations as solutions of the differential equation

$$x^2y'' + xy' + (x^2 - p^2)y = 0. \tag{3.77}$$

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

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

$$y(x) = J_p(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{\Gamma(n+1)\Gamma(n+p+1)} \left(\frac{x}{2}\right)^{2n+p}. \tag{3.78}$$

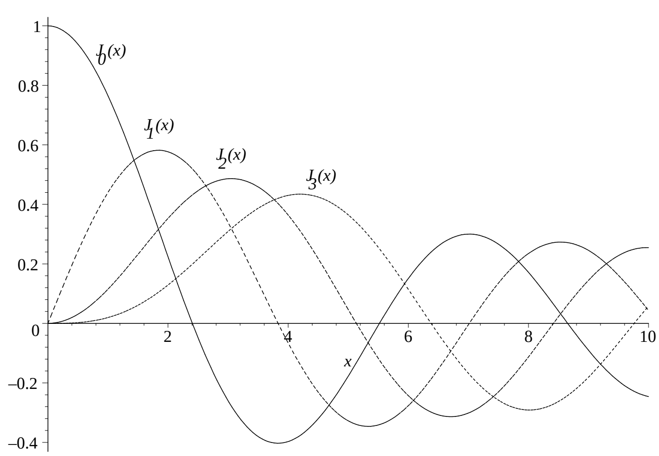

Figure 3.8: Plots of the Bessel functions $J_0(x)$, $J_1(x)$, $J_2(x)$, and $J_3(x)$.

In Figure 3.8, 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

$$N_p(x) = Y_p(x) = \frac{\cos \pi p J_p(x) - J_{-p}(x)}{\sin \pi p}. \tag{3.79}$$

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

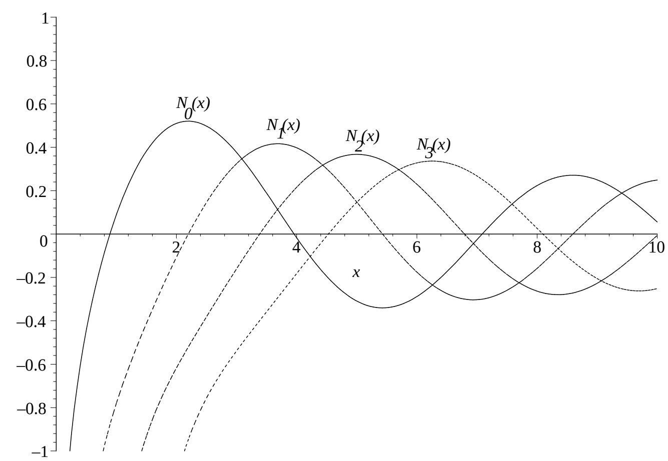

Figure 3.9: Plots of the Neumann functions $N_0(x)$, $N_1(x)$, $N_2(x)$, and $N_3(x)$.

In Figure 3.9, 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, 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. The radial 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.

$$\frac{d}{dx}\left[x^p J_p(x)\right] = x^p J_{p-1}(x). \tag{3.80}$$

$$\frac{d}{dx}\left[x^{-p} J_p(x)\right] = -x^{-p} J_{p+1}(x). \tag{3.81}$$

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

$$J_{p-1}(x) + J_{p+1}(x) = \frac{2p}{x} J_p(x). \tag{3.82}$$

$$J_{p-1}(x) - J_{p+1}(x) = 2J_p'(x). \tag{3.83}$$

Orthogonality One can recast the Bessel equation into an eigenvalue problem whose solutions form an orthogonal basis of functions on $L^2_x(0,a)$. Using Sturm-Liouville Theory, one can show that

$$\int_0^a x J_p(j_{pn}\frac{x}{a}) J_p(j_{pm}\frac{x}{a})\,dx = \frac{a^2}{2}\left[J_{p+1}(j_{pn})\right]^2 \delta_{n,m}, \tag{3.84}$$

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 is provided in Table 3.3.

n	$m=0$	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$
1	2.405	3.832	5.136	6.380	7.588	8.771
2	5.520	7.016	8.417	9.761	11.065	12.339
3	8.654	10.173	11.620	13.015	14.373	15.700
4	11.792	13.324	14.796	16.223	17.616	18.980
5	14.931	16.471	17.960	19.409	20.827	22.218
6	18.071	19.616	21.117	22.583	24.019	25.430
7	21.212	22.760	24.270	25.748	27.199	28.627
8	24.352	25.904	27.421	28.908	30.371	31.812
9	27.493	29.047	30.569	32.065	33.537	34.989

Table 3.3: The zeros of Bessel Functions, $J_m(j_{mn}) = 0$.

Generating Function

$$e^{x(t-\frac{1}{t})/2} = \sum_{n=-\infty}^{\infty} J_n(x) t^n, \quad x > 0, t \neq 0. \tag{3.85}$$

Integral Representation

$$J_n(x) = \frac{1}{\pi}\int_0^\pi \cos(x\sin\theta - n\theta)\,d\theta, \quad x > 0, n \in Z. \tag{3.86}$$

Fourier-Bessel Series

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

$$x^2 y'' + xy' + (\lambda x^2 - p^2)y = 0, \quad x \in [0,a], \tag{3.87}$$

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 $u(t) = J_p(t)$.

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

$$f(x) = \sum_{n=1}^{\infty} c_n J_p(j_{pn}\frac{x}{a}), \tag{3.88}$$

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

$$c_n = \frac{2}{a^2\left[J_{p+1}(j_{pn})\right]^2}\int_0^a xf(x)J_p(j_{pn}\frac{x}{a})\,dx. \tag{3.89}$$

Example 3.15. *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 (3.89):

$$c_n = \frac{2}{[J_1(j_{0n})]^2}\int_0^1 xJ_0(j_{0n}x)\,dx. \tag{3.90}$$

From the identity

$$\frac{d}{dx}\left[x^pJ_p(x)\right] = x^pJ_{p-1}(x), \tag{3.91}$$

we have

$$\begin{aligned}\int_0^1 xJ_0(j_{0n}x)\,dx &= \frac{1}{j_{0n}^2}\int_0^{j_{0n}} yJ_0(y)\,dy \\ &= \frac{1}{j_{0n}^2}\int_0^{j_{0n}} \frac{d}{dy}\left[yJ_1(y)\right]\,dy \\ &= \frac{1}{j_{0n}^2}\left[yJ_1(y)\right]_0^{j_{0n}} \\ &= \frac{1}{j_{0n}}J_1(j_{0n}).\end{aligned} \tag{3.92}$$

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

$$1 = 2\sum_{n=1}^{\infty}\frac{J_0(j_{0n}x)}{j_{0n}J_1(j_{0n})}, \quad 0 < x < 1. \tag{3.93}$$

In Figure 3.10, we show the partial sum for the first fifty terms of this series. Note once again the slow convergence due to the Gibbs phenomenon.

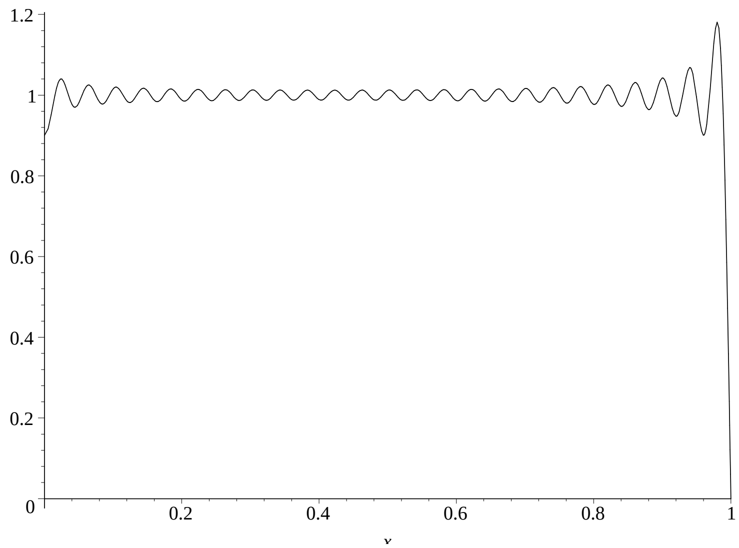

Figure 3.10: Plot of the first 50 terms of the Fourier-Bessel series in Equation (3.93) for $f(x) = 1$ on $0 < x < 1$.

3.7 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 $S_N(x) = \sum\limits_{n=1}^{N} c_n \phi_n(x)$ for a set of fixed functions $\phi_n(x)$; i.e., to find the expansion coefficients, c_n, 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:

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

The mean square deviation.

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:

$$\begin{aligned} E_N &= \int_a^b [f(x) - S_N(x)]^2 \rho(x)\, dx \\ &= \int_a^b \left[f(x) - \sum_{n=1}^{N} c_n \phi_n(x) \right]^2 \rho(x)\, dx \\ &= \int_a^b f^2(x)\rho(x)\, dx - 2\int_a^b f(x) \sum_{n=1}^{N} c_n \phi_n(x) \rho(x)\, dx \\ &\quad + \int_a^b \sum_{n=1}^{N} c_n \phi_n(x) \sum_{m=1}^{N} c_m \phi_m(x) \rho(x)\, dx. \end{aligned} \tag{3.94}$$

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

$$\begin{aligned} E_N &= < f, f > -2 \sum_{n=1}^{N} c_n < f, \phi_n > + \sum_{n=1}^{N} \sum_{m=1}^{N} c_n c_m < \phi_n, \phi_m > \\ &= < f, f > -2 \sum_{n=1}^{N} c_n < f, \phi_n > + \sum_{n=1}^{N} c_n^2 < \phi_n, \phi_n > . \end{aligned} \tag{3.95}$$

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

$$\begin{aligned} & -2 \sum_{n=1}^{N} c_n < f, \phi_n > + \sum_{n=1}^{N} c_n^2 < \phi_n, \phi_n > \\ &= \sum_{n=1}^{N} < \phi_n, \phi_n > c_n^2 - 2 < f, \phi_n > c_n \\ &= \sum_{n=1}^{N} < \phi_n, \phi_n > \left[c_n^2 - \frac{2 < f, \phi_n >}{< \phi_n, \phi_n >} c_n \right] \\ &= \sum_{n=1}^{N} < \phi_n, \phi_n > \left[\left(c_n - \frac{< f, \phi_n >}{< \phi_n, \phi_n >} \right)^2 - \left(\frac{< f, \phi_n >}{< \phi_n, \phi_n >} \right)^2 \right]. \end{aligned} \tag{3.96}$$

Up 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.

Minimization in Least Squares Sense

Inserting the Fourier coefficients into the mean square deviation yields

Bessel's Inequality.

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

Thus, we obtain Bessel's Inequality:

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

Convergence in the mean.

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

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

Letting N get large in Bessel's inequality shows that the sum $\sum_{n=1}^{N} c_n^2 < \phi_n, \phi_n >$ 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 test from calculus, we know that the convergence of $\sum a_n$ implies that $a_n \rightarrow 0$ as $n \rightarrow \infty$. 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 $N \rightarrow \infty$. This implies from the above derivation of Bessel's inequality that

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

This leads to Parseval's equality:

Parseval's equality.

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

Parseval's equality holds if and only if

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

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 $\int\limits_a^b |f(x)|\rho(x)\, dx < \infty$. This is the Riemann-Lebesgue Lemma, but will not be proven here.

3.8 Appendix: Convergence of Trigonometric Fourier Series

In this section we list definitions, lemmas and theorems needed to provide convergence arguments for trigonometric Fourier series. We will not attempt to discuss the derivations in depth, but provide enough for the interested reader to see what is involved in establishing convergence.

Definitions

1. For any nonnegative integer k, a function u is C^k if every k-th order partial derivative of u exists and is continuous.
2. For two functions $f$and g defined on an interval $[a, b]$, we will define the **inner product** as $< f, g >= \int_a^b f(x)g(x)\, dx$.
3. A function f is **periodic with period *p*** if $f(x + p) = f(x)$ for all x.
4. Let f be a function defined on $[-L, L]$ such that $f(-L) = f(L)$. The **periodic extension** $\tilde{f}$of f is the unique periodic function of period $2L$ such that $\tilde{f}(x) = f(x)$for all $x \in [-L, L]$.
5. The expression
$$D_N(x) = \frac{1}{2} + \sum_{n=1}^{N} \cos \frac{n\pi x}{L}$$
is called the ***N*-th Dirichlet Kernel.** [This will be summed later and the sequences of kernels converges to what is called the **Dirac Delta function**.]
6. A sequence of functions $\{s_1(x), s_2(x), \ldots\}$ is said to **converge pointwise** to $f(x)$ on the interval $[-L, L]$ if for each fixed x in the interval,
$$\lim_{N\to\infty} |f(x) - s_N(x)| = 0.$$
7. A sequence of functions $\{s_1(x), s_2(x), \ldots\}$ is said to **converge uniformly** to $f(x)$ on the interval $[-L, L]$ if
$$\lim_{N\to\infty} \left(\max_{|x|\le L} |f(x) - s_N(x)| \right) = 0.$$
8. **One-sided limits**: $f(x_0^+) = \lim_{x\downarrow x_0} f(x)$ and $f(x_0^-) = \lim_{x\uparrow x_0} f(x)$.
9. A function f is **piecewise continuous** on $[a, b]$ if the function satisfies
   a. f is defined and continuous at all but a finite number of points of $[a, b]$.

b. For all $x \in (a,b)$, the limits $f(x^+)$ and $f(x^-)$ exist.

c. $f(a^+)$ and $f(b^-)$ exist.

10. A function is **piecewise** C^1 on $[a,b]$ if $f(x)$ and $f'(x)$ are piecewise continuous on $[a,b]$.

Lemmas

1. **Bessel's Inequality:** Let $f(x)$ be defined on $[-L,L]$ and $\int\_{-L}^{L} f^2(x)\,dx < \infty$. If the trigonometric Fourier coefficients exist, then $a_0^2 + \sum_{n=1}^{N}(a_n^2+b_n^2) \le \frac{1}{L}\int_{-L}^{L} f^2(x)\,dx$. This follows from the earlier section on the Least Squares Approximation.
2. **Riemann-Lebesgue Lemma:** Under the conditions of Bessel's Inequality, the Fourier coefficients approach zero as $n \to 0$. This is based upon some earlier convergence results seen in Calculus in which one learns for a series of nonnegative terms, $\sum c_n$with $c_n \ge 0$, *if* c_n*does not approach 0 as* $n \to \infty$, *then* $\sum c_n$*does not converge*. Therefore, the contrapositive holds, if $\sum c_n$ *converges, then* $c_n \to 0$ *as* $n \to \infty$. From Bessel's Inequality, we see that when f is square integrable, the series formed by the sums of squares of the Fourier coefficients converges. Therefore, the Fourier coefficients must go to zero as n increases. This is also referred to in the earlier section on the Least Squares Approximation. However, an extension to absolutely integrable functions exists, which is called the Riemann-Lebesgue Lemma.
3. **Green's Formula:** Let $f$and $g$be C^2 functions on $[a,b]$. Then $< f'', g > - < f, g'' >= [f'(x)g(x) - f(x)g'(x)]|_a^b$. [Note: This is just an iteration of integration by parts.]
4. **Special Case of Green's Formula:** Let $f$and $g$be C^2 functions on $[-L,L]$ and both functions satisfy the conditions $f(-L) = f(L)$ and $f'(-L) = f'(L)$. Then $< f'', g >=< f, g'' >$.
5. **Lemma 1:** If g is a periodic function of period $2L$and c any real number, then $\int_{-L+c}^{L+c} g(x)\,dx = \int_{-L}^{L} g(x)\,dx$.
6. **Lemma 2:** Let f be a C^2 function on $[-L,L]$ such that $f(-L) = f(L)$ and $f'(-L) = f'(L)$. Then for $M = \max_{|x|\le L}|f''(x)|$ and $n \ge 1$,

$$|a_n| = \left|\frac{1}{L}\int_{-L}^{L} f(x)\cos\frac{n\pi x}{L}\,dx\right| \le \frac{2L^2M}{n^2\pi^2} \quad (3.97)$$

$$|b_n| = \left|\frac{1}{L}\int_{-L}^{L} f(x)\sin\frac{n\pi x}{L}\,dx\right| \le \frac{2L^2M}{n^2\pi^2}. \quad (3.98)$$

7. **Lemma 3:** For any real θ such that $\sin\frac{\theta}{2} \ne 0$,

$$\frac{1}{2} + \cos\theta + \cos 2\theta + \cdots + \cos n\theta = \frac{\sin((n+\frac{1}{2})\theta)}{2\sin\frac{\theta}{2}}$$

8. **Lemma 4:** Let $h(x)$ be C^1 on $[-L, L]$. Then

$$\lim_{n\to\infty} \frac{1}{L} \int_{-L}^{L} D_n(x)h(x)\, dx = h(0).$$

Convergence Theorems

1. **Theorem 1.** (Pointwise Convergence) Let f be C^1 on $[-L, L]$ with $f(-L) = f(L), f'(-L) = f'(L)$. Then $FS\, f(x) = f(x)$ for all x in $[-L, L]$.
2. **Theorem 2.** (Uniform Convergence) Let f be C^2 on $[-L, L]$ with $f(-L) = f(L), f'(-L) = f'(L)$. Then $FS\, f(x)$ converges uniformly to $f(x)$. In particular,

$$|f(x) - S_N(x)| \le \frac{4L^2 M}{\pi^2 N}$$

for all x in $[-L, L]$, where $M = \max_{|x|\le L} |f''(x)|$.
3. **Theorem 3.** (Piecewise C^1 - Pointwise Convergence) Let f be a piecewise C^1 function on $[-L, L]$. Then $FS\, f(x)$ converges to the periodic extension of

$$f(x) = \begin{cases} \frac{1}{2}[f(x^+) + f(x^-)], & -L < x < L \\ \frac{1}{2}[f(L^+) + f(L^-)], & x = \pm L \end{cases}$$

for all x in $[-L, L]$.
4. **Theorem 4.** (Piecewise C^1 - Uniform Convergence) Let f be a piecewise C^1 function on $[-L, L]$ such that $f(-L) = f(L)$. Then $FS\, f(x)$ converges uniformly to $f(x)$.

Proof of Convergence
We are considering the Fourier series of $f(x)$:

$$FS\, f(x) = a_0 + \sum_{n=1}^{\infty} \left[a_n \cos\frac{n\pi x}{L} + b_n \sin\frac{n\pi x}{L} \right],$$

where the Fourier coefficients are given by

$$\begin{aligned} a_0 &= \tfrac{1}{2L} \int_{-L}^{L} f(x)\, dx, \\ a_n &= \tfrac{1}{L} \int_{-L}^{L} f(x) \cos\tfrac{n\pi x}{L}\, dx, \\ b_n &= \tfrac{1}{L} \int_{-L}^{L} f(x) \sin\tfrac{n\pi x}{L}\, dx. \end{aligned}$$

We are first interested in the pointwise convergence of the infinite series. Thus, we need to look at the partial sums for each x. Writing out the partial sums, inserting the Fourier coefficients and rearranging, we have

$$\begin{aligned}
S_N(x) &= a_0 + \sum_{n=1}^{N}\left[a_n \cos\frac{n\pi x}{L} + b_n \sin\frac{n\pi x}{L}\right] \\
&= \frac{1}{2L}\int_{-L}^{L} f(y)\,dy + \sum_{n=1}^{N}\left[\left(\frac{1}{L}\int_{-L}^{L} f(y)\cos\frac{n\pi y}{L}\,dy\right)\cos\frac{n\pi x}{L}\right. \\
&\quad \left. + \left(\frac{1}{L}\int_{-L}^{L} f(y)\sin\frac{n\pi y}{L}\,dy.\right)\sin\frac{n\pi x}{L}\right] \\
&= \frac{1}{L}\int_{-L}^{L}\left\{\frac{1}{2} + \sum_{n=1}^{N}\left(\cos\frac{n\pi y}{L}\cos\frac{n\pi x}{L} + \sin\frac{n\pi y}{L}\sin\frac{n\pi x}{L}\right)\right\} f(y)\,dy \\
&= \frac{1}{L}\int_{-L}^{L}\left\{\frac{1}{2} + \sum_{n=1}^{N}\cos\frac{n\pi(y-x)}{L}\right\} f(y)\,dy \\
&\equiv \frac{1}{L}\int_{-L}^{L} D_N(y-x) f(y)\,dy \qquad (3.99)
\end{aligned}$$

Here

$$D_N(x) = \frac{1}{2} + \sum_{n=1}^{N}\cos\frac{n\pi x}{L}$$

is called the ***N*-th Dirichlet Kernel**. What we seek to prove is (**Lemma 4**) that

$$\lim_{N\to\infty}\frac{1}{L}\int_{-L}^{L} D_N(y-x) f(y)\,dy = f(x).$$

[Technically, we need the periodic extension of f.] So, we need to consider the Dirichlet kernel. Then pointwise convergence follows, as $\lim_{N\to\infty} S_N(x) = f(x)$.

Proposition:

$$D_n(x) = \begin{cases} \frac{\sin((n+\frac{1}{2})\frac{\pi x}{L})}{2\sin\frac{\pi x}{2L}}, & \sin\frac{\pi x}{2L} \neq 0 \\ n + \frac{1}{2}, & \sin\frac{\pi x}{2L} = 0 \end{cases}.$$

Proof: Actually, this follows from **Lemma 3**. Let $\theta = \frac{\pi x}{L}$ and multiply $D_n(x)$ by $2\sin\frac{\theta}{2}$ to obtain:

$$\begin{aligned}
2\sin\frac{\theta}{2}D_n(x) &= 2\sin\frac{\theta}{2}\left[\frac{1}{2} + \cos\theta + \cdots + \cos n\theta\right] \\
&= \sin\frac{\theta}{2} + 2\cos\theta\sin\frac{\theta}{2} + 2\cos 2\theta\sin\frac{\theta}{2} + \cdots + 2\cos n\theta\sin\frac{\theta}{2} \\
&= \sin\frac{\theta}{2} + \left(\sin\frac{3\theta}{2} - \sin\frac{\theta}{2}\right) + \left(\sin\frac{5\theta}{2} - \sin\frac{3\theta}{2}\right) + \cdots \\
&\quad + \left(\sin\left((n+\frac{1}{2})\theta\right) - \sin\left((n-\frac{1}{2})\theta\right)\right) \\
&= \sin\left((n+\frac{1}{2})\theta\right). \qquad (3.100)
\end{aligned}$$

Thus,

$$2\sin\frac{\theta}{2}D_n(x) = \sin\left((n+\frac{1}{2})\theta\right),$$

or if $\sin\frac{\theta}{2} \neq 0$,

$$D_n(x) = \frac{\sin\left((n+\frac{1}{2})\theta\right)}{2\sin\frac{\theta}{2}},\ \theta = \frac{\pi x}{L}.$$

If $\sin\frac{\theta}{2} = 0$, then one needs to apply L'Hospital's Rule:

$$\begin{aligned}\lim_{\theta\to 2m\pi}\frac{\sin\left((n+\frac{1}{2})\theta\right)}{2\sin\frac{\theta}{2}} &= \lim_{\theta\to 2m\pi}\frac{(n+\frac{1}{2})\cos\left((n+\frac{1}{2})\theta\right)}{\cos\frac{\theta}{2}} \\ &= \frac{(n+\frac{1}{2})\cos(2mn\pi+m\pi)}{\cos m\pi} \\ &= n+\frac{1}{2}. \end{aligned} \tag{3.101}$$

As $n \to \infty$, $D_n(x) \to \delta(x)$, the **Dirac delta function,** on the interval $[-L, L]$. In Figures 5.13-5.14 are some plots for $L = \pi$and $n = 25, 50, 100$. Note how the central peaks of $D_N(x)$ grow as N gets large and the values of $D_N(x)$ tend towards zero for nonzero x.

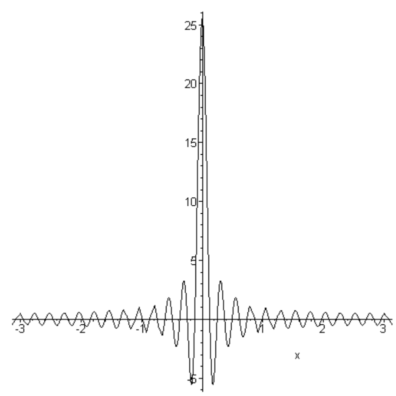

Figure 3.11: *N*th Dirichlet Kernel for N=25.

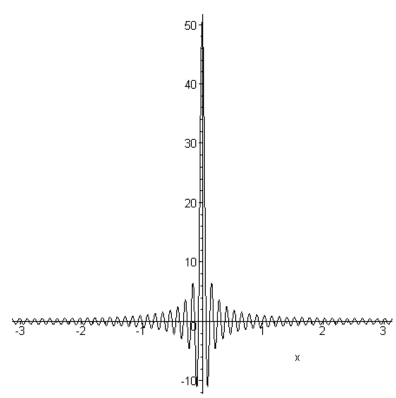

Figure 3.12: *N*th Dirichlet Kernel for N=50.

The Dirac delta function can be defined as that quantity satisfying

a. $\delta(x) = 0,\ x \neq 0;$

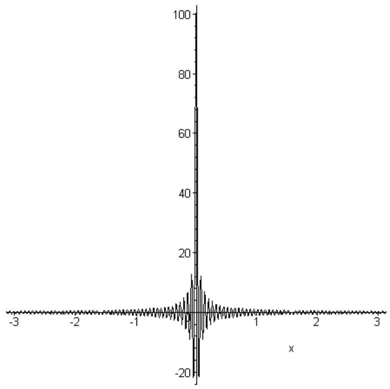

Figure 3.13: Nth Dirichlet Kernel for N=100.

b. $\int_{-\infty}^{\infty} \delta(x)\,dx = 1.$

This generalized function, or **distribution**, also has the property:

$$\int_{-\infty}^{\infty} f(x)\delta(x-a)\,dx = f(a).$$

Thus, under the appropriate conditions on f, one can show

$$\lim_{N\to\infty} \frac{1}{L}\int_{-L}^{L} D_N(y-x)f(y)\,dy = f(x).$$

We need to prove **Lemma 4** first.

Proof: Since $\frac{1}{L}\int_{-L}^{L} D_N(x)\,dx = \frac{1}{2L}\int_{-L}^{L} dx = 1$, we have that

$$\begin{aligned}
\frac{1}{L}\int_{-L}^{L} D_N(x)h(x)\,dx - h(0) &= \frac{1}{L}\int_{-L}^{L} D_N(x)\left[h(x)-h(0)\right]\,dx \\
&= \frac{1}{2L}\int_{-L}^{L}\left[\cos\frac{n\pi x}{L} + \cot\frac{\pi x}{L}\sin\frac{n\pi x}{L}\right]\left[h(x)-h(0)\right]\,dx.
\end{aligned} \tag{3.102}$$

The two terms look like the Fourier coefficients. An application of the Riemann-L:ebesgue Lemma indicates that these coefficients tend to zero as $n\to\infty$, provided the functions being expanded are square integrable and the integrals above exist. The cosine integral follows, but a little work is needed for the sine integral. One can use L'Hospital's Rule with $h \in C^1$.

Now we apply **Lemma 4** to get the convergence from

$$\lim_{N\to\infty} \frac{1}{L}\int_{-L}^{L} D_N(y-x)f(y)\,dy = f(x).$$

Due to periodicity, we have

$$\frac{1}{L}\int_{-L}^{L} D_N(y-x)f(y)\,dy = \frac{1}{L}\int_{-L}^{L} D_N(y-x)\tilde{f}(y)\,dy$$

$$= \frac{1}{L} \int_{-L+x}^{L+x} D_N(y-x) \tilde{f}(y)\, dy$$
$$= \frac{1}{L} \int_{-L}^{L} D_N(z) \tilde{f}(x+z)\, dz. \qquad (3.103)$$

We can apply **Lemma 4** providing $\tilde{f}(z+x)$ is C^1 in z, which is true since f is C^1 and behaves well at $\pm L$.

To prove **Theorem 2** on uniform convergence, we need only combine **Theorem 1** with **Lemma 2**. Then we have,

$$\begin{aligned} |f(x) - S_N(x)| &= |f(x) - S_N(x)| \\ &\leq \sum_{n=N+1}^{\infty} \left[\left| a_n \cos \frac{n\pi x}{L} \right| + \left| b_n \sin \frac{n\pi x}{L} \right| \right] \\ &\leq \sum_{n=N+1}^{\infty} [|a_n| + |b_n|] \qquad (3.104) \\ &\leq \frac{4L^2 M}{\pi^2} \sum_{n=N+1}^{\infty} \frac{1}{n^2} \\ &\leq \frac{4L^2 M}{\pi^2 N}. \qquad (3.105) \end{aligned}$$

This gives the uniform convergence.

These Theorems can be relaxed to include piecewise C^1 functions. **Lemma 4** needs to be changed for such functions to the result that

$$\lim_{n\to\infty} \frac{1}{L} \int_{-L}^{L} D_n(x) h(x)\, dx = \frac{1}{2} [h(0^+) + h(0^-)]$$

by splitting the integral into integrals over $[-L, 0]$, $[0, L]$ and applying a one-sided L'Hospital's Rule. Proving uniform convergence under the conditions in **Theorem 4** takes a little more effort, but it can be done.

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 (3.41).

b. The three-term recursion formula in Equation (3.43).

4. In Equations (3.56) through (3.63) we provide several identities for Legendre polynomials. Derive the results in Equations (3.57) through (3.63) as described in the text. Namely,

a. Differentiating Equation (3.56) with respect to x, derive Equation (3.57).

b. Derive Equation (3.58) by differentiating $g(x,t)$ with respect to x and rearranging the resulting infinite series.

c. Combining the previous result with Equation (3.56), derive Equations (3.59) and (3.60).

d. Adding and subtracting Equations (3.59) and (3.60), obtain Equations (3.61) and (3.62).

e. Derive Equation (3.63) using some of the other identities.

5. Use the recursion relation (3.43) to evaluate $\int_{-1}^{1} xP_n(x)P_m(x)\,dx$, $n \le m$.

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 your 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 3.13 for Fourier-Legendre series, the value is given in terms of double factorials! So, we have

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

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 of 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 Problem 13 to derive the derivative identities (3.80) and (3.81):

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 Problem 14.

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, (3.80) and (3.81), 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 $x^2y'' + xy' + (\lambda^2 x^2 - p^2)y = 0$. 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 $\alpha_j's$ 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 Table 3.3 and Figure 3.8. [Note: Be careful in 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

$$\begin{aligned} \int_0^c (xy')' y\, dx &= -\alpha_j^2 \int_0^c xy^2\, dx \\ &= -\alpha_j^2 \int_0^c x J_0^2(\alpha_j x)\, dx. \end{aligned} \qquad (3.106)$$

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 (3.81).
2. Equation (3.84).
3. $J_2(\alpha_j c) + J_0(\alpha_j c) = 0$ from Equation (3.82).

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

d. Use the results from parts b and c and Problem 16 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.