47

Sturm-Liouville Problems

"Sturm-Liouville problems" are boundary-value problems that naturally arise when solving certain partial differential equation problems using a "separation of variables" method that will be discussed in a later chapter. It is the theory behind Sturm-Liouville problems that, ultimately, justifies the "separation of variables" method for these partial differential equation problems. The simplest applications lead to the various Fourier series, and less simple applications lead to generalizations of Fourier series involving Bessel functions, Hermite polynomials, etc.

Unfortunately, there are several difficulties with our studying Sturm-Liouville problems:

1. Motivation: Had we the time, we would first discuss partial differential equation problems and develop the separation of variables method for solving certain important types of problems involving partial differential equations. We would then see how these "Sturm-Liouville problems" arise and why they are so important. But we don't have time. Instead, I'll briefly remind you of some results from linear algebra that are analogous to the results we will, eventually, obtain.

2. Another difficulty is that the simplest examples (which are very important since they lead to the Fourier series) are too simple to really illustrate certain elements of the theory, while the other standard examples tend to get complicated and require additional tricks which distract from illustrating the theory. We'll deal with this as well as we can.

3. The material gets theoretical. Sorry, there is no way around this. The end results, however, are very useful in computations, especially now that we have computers to do the tedious computations. I hope we get to that point.

4. Finally, I must warn you that, in most texts, the presentation of the "Sturm-Liouville theory" stinks. In most introductory ordinary differential equation texts, this material is usually near the end, which usually means that the author just wants to finish the damn book, and get it published. In texts on partial differential equations, the results are usually quoted with some indication that the proofs can be found in a good text on differential equations.

47.1 Linear Algebraic Antecedents

Let us briefly review some elements from linear algebra which we will be “recreating” using functions instead of finite-dimensional vectors. But first, let me remind you of a little “complex variable algebra”. This is because our column vectors and matrices may have complex values.

Recollections From (Real) Linear Algebra

For the rest of this section, let N be some positive integer. Following fairly standard convention, we'll let $\mathbb{R}^N$ denote the vector space of all $N \times 1$ matrices with real components. That is, saying that $\mathbf{v}$ and $\mathbf{w}$ are ‘vectors’ in $\mathbb{R}^N$ will simply mean

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_N \end{bmatrix} \qquad \text{and} \qquad \mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_N \end{bmatrix}$$

where the v_k's and w_k's are real numbers.

Recall that the norm (or ‘length’) of the above $\mathbf{v}$, $\|\mathbf{v}\|$, is computed by

$$\|\mathbf{v}\| = \sqrt{|v_1|^2 + |v_2|^2 + \cdots |v_N|^2}$$

and that, computationally, the classic dot product of the above $\mathbf{v}$ and $\mathbf{w}$ is

$$\begin{aligned} \mathbf{v} \cdot \mathbf{w} &= v_1 w_1 + v_2 w_2 + \cdots + v_N w_N \\ &= [v_1 , v_2 , \ldots , v_N] \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_N \end{bmatrix} = \mathbf{v}^{\mathsf{T}}\mathbf{w} \end{aligned}$$

where $\mathbf{v}^{\mathsf{T}}\mathbf{w}$ is the matrix product with $\mathbf{v}^{\mathsf{T}}$ being the transpose of matrix $\mathbf{v}$ (i.e., the matrix constructed from $\mathbf{v}$ by switching rows with columns).

Observe that, since $v^2 = |v|^2$ when v is real,

$$\begin{aligned} \mathbf{v} \cdot \mathbf{v} &= v_1 v_1 + v_2 v_2 + \cdots + v_N v_N \\ &= |v_1|^2 + |v_2|^2 + \cdots + |v_N|^2 = \|\mathbf{v}\|^2 \quad . \end{aligned}$$

Also recall that a set of vectors $\{\mathbf{v}^1, \mathbf{v}^2, \mathbf{v}^2, \ldots\}$ is said to be orthogonal if and only if

$$\mathbf{v}^j \cdot \mathbf{v}^k = 0 \qquad \text{whenever} \quad j \neq k \quad .$$

And finally, recall that a matrix $\mathbf{A}$ is *symmetric* if and only if $\mathbf{A}^{\mathsf{T}} = \mathbf{A}$, and that, for symmetric matrices, we have the following theorem from linear algebra (and briefly mentioned in chapter 39):

Theorem 47.1
Let $\mathbf{A}$ be a symmetric $N \times N$ matrix (with real-valued components). Then both of the following hold:

1. *All the eigenvalues of* $\mathbf{A}$ *are real.*

2. *There is an orthogonal basis for* $\mathbb{R}^N$ *consisting of eigenvectors for* $\mathbf{A}$.

Ultimately, we will obtain an analog to this theorem involving a linear differential operator instead of a matrix $\mathbf{A}$.

Linear Algebra with Complex Components

Complex Conjugates and Magnitudes

We've certainly used complex numbers before in this text, but I should remind you that a complex number z is something that can be written as

$$z = x + iy$$

where x and y are real numbers — the *real* and *imaginary parts*, respectively of z. The corresponding complex conjugate z^* and magnitude $|z|$ of z are then given by

$$z^* = x - iy \qquad \text{and} \qquad |z| = \sqrt{x^2 + y^2} \quad .$$

Now if x is a real number — positive, negative or zero — then $x^2 = |x|^2$. However, you can easily verify that $z^2 \neq |z|^2$. Instead, we have

$$|z|^2 = x^2 + y^2 = x^2 - (iy)^2 = (x - iy)(x + iy) = z^*z \quad .$$

Vectors and Matrices with Complex-Valued Components

Everything mentioned above can be generalized to $\mathbb{C}^N$, the vector space of all $N \times 1$ matrices with complex components. In this case, saying that $\mathbf{v}$ and $\mathbf{w}$ are 'vectors' in $\mathbb{C}^N$ simply means

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_N \end{bmatrix} \qquad \text{and} \qquad \mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_N \end{bmatrix}$$

where the v_k's and w_k's are complex numbers. Since the components are complex, we define the complex conjugate of $\mathbf{v}$ in the obvious way,

$$\mathbf{v}^* = \begin{bmatrix} v_1^* \\ v_2^* \\ \vdots \\ v_N^* \end{bmatrix} \quad .$$

The natural norm (or 'length') of $\mathbf{v}$ is still given by

$$\|\mathbf{v}\| = \sqrt{|v_1|^2 + |v_2|^2 + \cdots |v_N|^2}$$

However, when we try to relate this to the classic dot product $\mathbf{v} \cdot \mathbf{v}$, we get

$$\begin{aligned} \|\mathbf{v}\|^2 &= |v_1|^2 + |v_2|^2 + \cdots + |v_N|^2 \\ &= {v_1}^* v_1 + {v_2}^* v_2 + \cdots + {v_N}^* v_N \\ &= (\mathbf{v}^*) \cdot \mathbf{v} \neq \mathbf{v} \cdot \mathbf{v} \quad . \end{aligned}$$

This suggests that, instead of using the classic dot product, we use the *(standard vector) inner product* of $\mathbf{v}$ with $\mathbf{w}$, which is denoted by $\langle \mathbf{v} \mid \mathbf{w} \rangle$ and defined by by

$$\begin{aligned} \langle \mathbf{v} \mid \mathbf{w} \rangle = \mathbf{v}^* \cdot \mathbf{w} &= {v_1}^* w_1 + {v_2}^* w_2 + \cdots + {v_N}^* w_N \\ &= [v_1^*, v_2^*, \ldots, v_N^*] \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_N \end{bmatrix} = (\mathbf{v}^*)^{\mathsf{T}} \mathbf{w} \quad . \end{aligned}$$

Then

$$\langle \mathbf{v} \mid \mathbf{v} \rangle = \mathbf{v}^* \cdot \mathbf{v} = \|\mathbf{v}\|^2 \quad .$$

We also adjust our notion of "orthogonality" by saying that any set $\{\mathbf{v}^1, \mathbf{v}^2, \ldots\}$ of vectors in $\mathbb{C}^N$ is *orthogonal* if and only if

$$\left\langle \mathbf{v}^m \mid \mathbf{v}^n \right\rangle = 0 \qquad \text{whenever} \quad m \neq n \quad .$$

The inner product for vectors with complex components is the mathematically natural extension of the standard dot product for vectors with real components. Some easily verified (and useful) properties of this inner product are given in the next theorem. Verifying it will be left as an exercise (see exercise 47.6).

Theorem 47.2 (properties of the inner product)
Suppose α and β are two (possibly complex) constants, and $\mathbf{v}$, $\mathbf{w}$, and $\mathbf{u}$ are vectors in $\mathbb{C}^N$. Then

1. $\langle \mathbf{v} \mid \mathbf{w} \rangle = \langle \mathbf{w} \mid \mathbf{v} \rangle^*$,
2. $\langle \mathbf{u} \mid \alpha\mathbf{v} + \beta\mathbf{w} \rangle = \alpha \langle \mathbf{u} \mid \mathbf{v} \rangle + \beta \langle \mathbf{u} \mid \mathbf{w} \rangle$,
3. $\langle \alpha\mathbf{v} + \beta\mathbf{w} \mid \mathbf{u} \rangle = \alpha^* \langle \mathbf{v} \mid \mathbf{u} \rangle + \beta^* \langle \mathbf{w} \mid \mathbf{u} \rangle$,

and

4. $\langle \mathbf{v} \mid \mathbf{v} \rangle = \|\mathbf{v}\|$.

Later, we'll define other "inner products" for functions. These inner products will have very similar properties to those in given in the last theorem.

Adjoints

The *adjoint* of any matrix $\mathbf{A}$ — denoted $\mathbf{A}^\dagger$ — is the transpose of the complex conjugate of $\mathbf{A}$,

$$\mathbf{A}^\dagger = (\mathbf{A}^*)^{\mathsf{T}} \qquad \text{(equivalently, } \left(\mathbf{A}^{\mathsf{T}}\right)^* \text{)} \quad .$$

That is, $\mathbf{A}^\dagger$ is the matrix obtained from matrix $\mathbf{A}$ by replacing each entry in $\mathbf{A}$ with its complex conjugate, and then switching the rows and columns (or first switch the rows and columns and then replace the entries with their complex conjugates — you get the same result either way).

!►*Example 47.1:* *If*

$$\mathbf{A} = \begin{bmatrix} 1+2i & 3-4i & 5i \\ -6i & 7 & -8i \end{bmatrix} \quad ,$$

then

$$\mathbf{A}^\dagger = \left(\begin{bmatrix} 1+2i & 3-4i & 5i \\ -6i & 7 & -8i \end{bmatrix}^* \right)^\mathsf{T} = \begin{bmatrix} 1-2i & 3+4i & -5i \\ 6i & 7 & 8i \end{bmatrix}^\mathsf{T} = \begin{bmatrix} 1-2i & 6i \\ 3+4i & 7 \\ -5i & 8i \end{bmatrix} \quad .$$

This “adjoint” turns out to be more useful than the transpose when we allow vectors to have complex components.

A matrix $\mathbf{A}$ is *self adjoint*[1] if and only if $\mathbf{A}^\dagger = \mathbf{A}$. Note that:

1. A self-adjoint matrix is automatically square.

2. If $\mathbf{A}$ is a square matrix with just real components, then $\mathbf{A}^\dagger = \mathbf{A}^\mathsf{T}$, and “$\mathbf{A}$ is self adjoint” means the same as “$\mathbf{A}$ is symmetric”.

If you take the proof of theorem 47.1 and modify it to take into account the possibility of complex-valued components, you get

Theorem 47.3

Let $\mathbf{A}$ be a self-adjoint $N \times N$ matrix. Then both of the following hold:

1. *All the eigenvalues of $\mathbf{A}$ are real.*

2. *There is an orthogonal basis for $\mathbb{C}^N$ consisting of eigenvectors for $\mathbf{A}$.*

This theorem is noteworthy because it will help explain the source of some of the terminology that we will later be using.

What is more noteworthy is what the above theorem says about computing $\mathbf{Av}$ when $\mathbf{A}$ is self adjoint. To see this, let

$$\left\{ \mathbf{b}^1 \, , \, \mathbf{b}^2 \, , \, \mathbf{b}^3 \, , \, \dots \, , \, \mathbf{b}^N \right\}$$

be any orthogonal basis for $\mathbb{C}^N$ consisting of eigenvectors for $\mathbf{A}$ (remember, the theorem says there is such a basis), and let

$$\{ \lambda_1 \, , \, \lambda_2 \, , \, \lambda_3 \, , \, \dots \, , \, \lambda_N \}$$

be the corresponding set of eigenvectors (so $\mathbf{Ab}^k = \lambda_k \mathbf{b}^k$ for $k = 1, 2, \dots, N$). Since the set of $\mathbf{b}^k$'s is a basis, we can express $\mathbf{v}$ as a linear combination of these basis vectors.

$$\mathbf{v} = v_1 \mathbf{b}^1 + v_2 \mathbf{b}^2 + v_3 \mathbf{b}^3 + \cdots + v_N \mathbf{b}^N = \sum_{k=1}^{N} v_k \mathbf{b}^k \quad , \tag{47.1}$$

[1] the term *Hermitian* is also used

and can reduce the computation of $\mathbf{Av}$ to

$$\begin{aligned}\mathbf{Av} &= \mathbf{A}\left[v_1\mathbf{b}^1 + v_2\mathbf{b}^2 + v_3\mathbf{b}^3 + \cdots + v_N\mathbf{b}^N\right] \\ &= v_1\mathbf{Ab}^1 + v_2\mathbf{Ab}^2 + v_3\mathbf{Ab}^3 + \cdots + v_N\mathbf{Ab}^N \\ &= v_1\lambda_1\mathbf{b}^1 + v_2\lambda_2\mathbf{b}^2 + v_3\lambda_3\mathbf{b}^3 + \cdots + v_N\lambda_N\mathbf{b}^N \quad .\end{aligned}$$

If N is large, this could be a lot faster than doing the basic component-by-component matrix multiplication. (And in our analog with functions, N will be infinite.)

Finding Vector Components

One issue is finding the components $(v_1, v_2, v_3, \ldots, v_N)$ in formula (47.1) for $\mathbf{v}$. But a useful formula is easily derived. First, take the inner product of both sides of (47.1) with one of the $\mathbf{b}^k$'s, say, $\mathbf{b}^1$ and then repeatedly use the linearity property described in theorem 47.2,

$$\begin{aligned}\left\langle \mathbf{b}^1 \mid \mathbf{v} \right\rangle &= \left\langle \mathbf{b}^1 \mid v_1\mathbf{b}^1 + v_2\mathbf{b}^2 + v_3\mathbf{b}^3 + \cdots + v_N\mathbf{b}^N \right\rangle \\ &= v_1\left\langle \mathbf{b}^1 \mid \mathbf{b}^1 \right\rangle + v_2\left\langle \mathbf{b}^1 \mid \mathbf{b}^2 \right\rangle + v_3\left\langle \mathbf{b}^1 \mid \mathbf{b}^3 \right\rangle + \cdots + v_N\left\langle \mathbf{b}^1 \mid \mathbf{b}^N \right\rangle \quad .\end{aligned}$$

Remember, this set of $\mathbf{b}^k$'s is orthogonal. That means

$$\left\langle \mathbf{b}^1 \mid \mathbf{b}^k \right\rangle = 0 \qquad \text{if} \quad 1 \neq k \quad .$$

On the other hand,

$$\left\langle \mathbf{b}^1 \mid \mathbf{b}^1 \right\rangle = \left\|\mathbf{b}^1\right\|^2 \quad .$$

Taking this into account, we continue our computation of $\mathbf{b}^1 \cdot \mathbf{v}$:

$$\begin{aligned}\left\langle \mathbf{b}^1 \mid \mathbf{v} \right\rangle &= v_1\left\langle \mathbf{b}^1 \mid \mathbf{b}^1 \right\rangle + v_2\left\langle \mathbf{b}^1 \mid \mathbf{b}^2 \right\rangle + v_3\left\langle \mathbf{b}^1 \mid \mathbf{b}^3 \right\rangle + \cdots + v_N\left\langle \mathbf{b}^1 \mid \mathbf{b}^N \right\rangle \\ &= v_1 \cdot \left\|\mathbf{b}^1\right\|^2 + v_2 \cdot 0 + v_3 \cdot 0 + \cdots + v_N \cdot 0 \quad .\end{aligned}$$

So,

$$\left\langle \mathbf{b}^1 \mid \mathbf{v} \right\rangle = v_1 \left\|\mathbf{b}^1\right\|^2 \quad .$$

Solving for v_1 gives us

$$v_1 = \frac{\left\langle \mathbf{b}^1 \mid \mathbf{v} \right\rangle}{\left\|\mathbf{b}^1\right\|^2} \quad .$$

Repeating this using each $\mathbf{b}^k$ yields the next theorem:

Theorem 47.4

Let

$$\left\{ \mathbf{b}^1 , \mathbf{b}^2 , \mathbf{b}^3 , \ldots , \mathbf{b}^N \right\}$$

be an orthogonal basis for $\mathbb{C}^N$ *. Then, for any vector* $\mathbf{v}$ *in* $\mathbb{R}^N$ *,*

$$\mathbf{v} = v_1\mathbf{b}^1 + v_2\mathbf{b}^2 + v_3\mathbf{b}^3 + \cdots + v_N\mathbf{b}^N = \sum_{k-1}^{N} v_k\mathbf{b}^k \tag{47.2a}$$

where

$$v_k = \frac{\left\langle \mathbf{b}^k \mid \mathbf{v} \right\rangle}{\left\|\mathbf{b}^k\right\|^2} \qquad \textit{for} \quad k = 1,\ 2,\ 3,\ \ldots,\ N \quad . \tag{47.2b}$$

If you know a little about "Fourier series", then you may recognize formula set (47.2) as a finite-dimensional analog of the Fourier series formulas. If you know nothing about "Fourier series", then I'll tell you that a "Fourier series for a function f" is an infinite-dimensional analog of formula set (47.2). We will eventually see this.

About 'Normalizing' a Basis

In the above, we assumed that

$$\left\{ \mathbf{b}^1 , \mathbf{b}^2 , \mathbf{b}^3 , \dots , \mathbf{b}^N \right\}$$

is an orthogonal basis for $\mathbb{C}^N$. If it had been ortho*normal*, then we would also have had

$$\left\| \mathbf{b}^k \right\| = 1 \qquad \text{for} \quad k = 1, 2, 3, \dots, N \quad ,$$

and the formulas in our last theorem would have simplified somewhat. In fact, given any orthogonal set

$$\left\{ \mathbf{b}^1 , \mathbf{b}^2 , \mathbf{b}^3 , \dots , \mathbf{b}^N \right\} \quad ,$$

we can construct a corresponding ortho*normal* set

$$\left\{ \mathbf{n}^1 , \mathbf{n}^2 , \mathbf{n}^3 , \dots , \mathbf{n}^N \right\}$$

by just letting

$$\mathbf{n}^k = \frac{\mathbf{b}^k}{\left\| \mathbf{b}^k \right\|} \qquad \text{for} \quad k = 1, 2, 3, \dots, N \quad .$$

Moreover, if $\mathbf{b}^k$ is an eigenvector for a matrix $\mathbf{A}$ with corresponding eigenvalue λ_k, so is $\mathbf{n}^k$.

When we so compute the $\mathbf{n}^k$'s from the $\mathbf{b}^k$'s, we are said to be *normalizing* our $\mathbf{b}^k$'s. Some authors like to normalize their orthogonal bases because it does yield simpler formulas for computing with such a basis. These authors, typically, are only deriving pretty results and are not really using them in applications. Those that really use the results rarely normalize, especially when the dimension is infinite (as it will be for us), because normalizing leads to artificial formulas for the basis vectors, and dealing with these artificial formulas for basis vectors usually complicates matters enough to completely negate the advantages of having the 'simpler' formulas for computing with these basis vectors.

We won't normalize.

Inner Products and Adjoints

You should recall (or be able to quickly confirm) that

$$(\mathbf{AB})^{\mathsf{T}} = \mathbf{B}^{\mathsf{T}}\mathbf{A}^{\mathsf{T}} \quad , \quad \left(\mathbf{A}^{\mathsf{T}}\right)^{\mathsf{T}} = \mathbf{A} \quad , \quad (\mathbf{AB})^* = \mathbf{A}^* \mathbf{B}^* \quad \text{and} \quad \left(\mathbf{A}^*\right)^* = \mathbf{A} \quad .$$

With these and the definition of the adjoint, you can easily verify that

$$(\mathbf{AB})^{\dagger} = \mathbf{B}^{\dagger}\mathbf{A}^{\dagger} \qquad \text{and} \qquad \left(\mathbf{A}^{\dagger}\right)^{\dagger} = \mathbf{A} \quad .$$

Also observe that, if

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_N \end{bmatrix} \qquad \text{and} \qquad \mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_N \end{bmatrix} \quad ,$$

then

$$\mathbf{v}^\dagger\mathbf{w} = \left[v_1{}^* \ , \ v_2{}^* \ , \ \ldots \ , \ v_N{}^*\right]\begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_N \end{bmatrix} = v_1{}^*w_1 + v_2{}^*w_2 + \cdots + v_N{}^*w_N = \langle\, \mathbf{v} \mid \mathbf{w} \,\rangle \quad .$$

So, for any $N \times N$ matrix $\mathbf{A}$ and any two vectors $\mathbf{v}$ and $\mathbf{u}$ in $\mathbb{C}^N$,

$$\left\langle\, \mathbf{v} \,\middle|\, \mathbf{A}^\dagger\mathbf{u} \,\right\rangle = \mathbf{v}^\dagger\left(\mathbf{A}^\dagger\mathbf{u}\right) = \mathbf{v}^\dagger\mathbf{A}^\dagger\mathbf{u} = (\mathbf{A}\mathbf{v})^\dagger\mathbf{u} = \langle\, \mathbf{A}\mathbf{v} \mid \mathbf{u} \,\rangle \quad .$$

Consequently, if $\mathbf{A}$ is self adjoint (i.e., $\mathbf{A}^\dagger = \mathbf{A}$), then

$$\langle\, \mathbf{v} \mid \mathbf{A}\mathbf{u} \,\rangle = \langle\, \mathbf{A}\mathbf{v} \mid \mathbf{u} \,\rangle \qquad \text{for every } \mathbf{v}, \mathbf{u} \text{ in } \mathbb{C}^N$$

It's a simple exercise in linear algebra to show that the above completely characterizes "adjointness" and "self adjointness" for matrices. That is, you should be able to finish proving the next theorem. We'll use the results to extend the these concepts to things other than matrices.

Theorem 47.5 (characterization of "adjointness")
Let $\mathbf{A}$ *be an* $N \times N$ *matrix. Then,*

$$\left\langle\, \mathbf{v} \,\middle|\, \mathbf{A}^\dagger\mathbf{u} \,\right\rangle = \langle\, \mathbf{A}\mathbf{v} \mid \mathbf{u} \,\rangle \qquad \textit{for every } \mathbf{v} \textit{ and } \mathbf{u} \textit{ in } \mathbb{C}^N \quad .$$

Moreover, both of the following hold:

1. $\mathbf{B} = \mathbf{A}^\dagger$ *if and only if*

$$\langle\, \mathbf{v} \mid \mathbf{B}\mathbf{u} \,\rangle = \langle\, \mathbf{A}\mathbf{v} \mid \mathbf{u} \,\rangle \qquad \textit{for every } \mathbf{v} \textit{ and } \mathbf{u} \textit{ in } \mathbb{C}^N \quad .$$

2. $\mathbf{A}$ *is self adjoint if and only if*

$$\langle\, \mathbf{v} \mid \mathbf{A}\mathbf{u} \,\rangle = \langle\, \mathbf{A}\mathbf{v} \mid \mathbf{u} \,\rangle \qquad \textit{for every } \mathbf{v} \textit{ and } \mathbf{u} \textit{ in } \mathbb{C}^N \quad .$$

Comments

To be honest, we are not going to directly use the material we've developed over the past several pages. The reason we went over the "theory of self-adjoint matrices" and related material concerning vectors in $\mathbb{C}^N$ is that the Sturm-Liouville theory we'll be developing is a functional analog of what we just discussed, using differential operators and functions instead of matrices and vectors in $\mathbb{C}^N$. Understanding the theory and computations we've just developed should expedite learning the theory and computations we will be developing.

47.2 Boundary-Value Problems with Parameters

In each boundary-value problem of interest to us, the differential equation will involve both the unknown function — which (following tradition) we will denote by ϕ or $\phi(x)$ — and a parameter λ (basically, λ is just some yet undetermined constant). We'll assume that each of these equations can be written as

$$A(x)\frac{d^2\phi}{dx^2} + B(x)\frac{d\phi}{dx} + [C(x)+\lambda]\phi = 0 \qquad (47.3a)$$

or, equivalently, as

$$A(x)\frac{d^2\phi}{dx^2} + B(x)\frac{d\phi}{dx} + C(x)\phi = -\lambda\phi \qquad (47.3b)$$

where A, B and C are known functions.

A solution to such a problem is a pair (λ, ϕ) that satisfies the given problem — both the differential equation and the boundary conditions. To avoid triviality, we will insist that the function ϕ be nontrivial (i.e., not always zero). Traditionally, the λ is called an eigenvalue, and the corresponding ϕ is called an eigenfuction. We'll often refer to the two together, (λ, ϕ), as an eigen-pair. The entire problem can be called an eigen-problem, a boundary-value problem or a "Sturm-Liouville problem" (though we won't completely define just what a "Sturm-Liouville problem" is until later).

In our problems, we will need to find the general solution $\phi = \phi_\lambda$ to the given differential equation for each possible value of λ, and then apply the boundary conditions to find all possible eigen-pairs. Technically, at this point, λ can be any complex number. However, thanks to the foreknowledge of the author, we can assume λ is real. Why this is a safe assumption will be one of the things we will later need to verify (it's analogous to the fact that self-adjoint matrices have only real eigenvalues). What we cannot yet do, though, is assume the λ's all come from some particular subinterval of the real line. This means you must consider all possible real values for λ, and take into account that the form of the solution ϕ_λ may be quite different for different ranges of these λ's.

!► ***Example 47.2:*** *The simplest (and possibly most important) example of such an boundary-value problem with parameter is*

$$\phi'' + \lambda\phi = 0 \qquad (47.4a)$$

with boundary conditions

$$\phi(0) = 0 \quad \text{and} \quad \phi(L) = 0 \qquad (47.4b)$$

where L is some positive number (think of it as a given 'length').

The above differential equation is a simple second-order homogeneous differential equation with constant coefficients. Its characteristic equation is

$$r^2 + \lambda = 0 \quad ,$$

with solution

$$r = \pm\sqrt{-\lambda} \quad .$$

In this example, the precise formula for $\phi_\lambda(x)$, the equation's general solution corresponding to a particular value of λ, depends on whether $\lambda > 0$, $\lambda = 0$ or $\lambda < 0$. Let us go through all the cases:

$\lambda < 0$: *In this case,* $-\lambda > 0$. *For convenience, let* $\nu = \sqrt{-\lambda}$. *Then*

$$r = \pm\nu \quad ,$$

and

$$\phi_\lambda(x) = c_1 e^{\nu x} + c_2 e^{-\nu x}$$

where c_1 *and* c_2 *are arbitrary constants, and, as already stated,* $\nu = \sqrt{-\lambda}$.

Applying the first boundary condition:

$$0 = \phi_\lambda(0) = c_1 e^{\nu \cdot 0} + c_2 e^{-\nu \cdot 0} = c_1 + c_2 \quad ,$$

which tells us that

$$c_2 = -c_1 \quad ,$$

and thus,

$$\phi_\lambda(x) = c_1 e^{\nu x} - c_1 e^{-\nu x} = c_1\left[e^{\nu x} - e^{-\nu x}\right] \quad .$$

Combining this with the boundary condition at $x = L$ *yields*

$$0 = \phi_\lambda(L) = c_1\left[e^{\nu L} - e^{-\nu L}\right] \quad .$$

Now $e^{\nu L} > e^0 = 1$ *and* $e^{-\nu L} < e^0 = 1$. *So*

$$e^{\nu L} - e^{-\nu L} > 0 \quad ,$$

and the only way we can have

$$0 = c_1\left[e^{\nu L} - e^{-\nu L}\right]$$

is for $c_1 = 0$. *Thus, the only solution to this problem with* $\lambda < 0$ *is*

$$\phi_\lambda(x) = 0\left[e^{\nu L} - e^{-\nu L}\right] = 0 \quad ,$$

the trivial solution (which we don't really care about).

$\lambda = 0$: *Instead of using the characteristic equation, let's just note that, if* $\lambda = 0$, *the differential equation for* $\phi = \phi_0(x)$ *reduces to*

$$\phi''(x) = 0 \quad .$$

Integrating this twice yields

$$\phi_0(x) = \phi_0(x) = c_1 x + c_2 \quad .$$

Applying the first boundary condition gives us

$$0 = \phi_0(0) = c_1 \cdot 0 + c_2 = c_2 \quad .$$

Combined with the boundary condition at $x = L$, *we then have*

$$0 = \phi_0(L) = c_1 L + 0 = c_1 L \quad ,$$

which says that $c_1 = 0$ *(since* $L > 0$*). Thus, the only solution to the differential equation that satisfies the boundary conditions when* $\lambda = 0$ *is*

$$\phi_0(x) = 0 \cdot x + 0 = 0 \quad ,$$

again, just the trivial solution.

$\lambda > 0$: *This time, it is convenient to let* $\nu = \sqrt{\lambda}$*, so that the solution to the characteristic equation becomes*

$$r = \pm\sqrt{-\lambda} = \pm i\nu \quad .$$

While the general formula for ϕ_λ *can be written in terms of complex exponentials, it is better to recall that these complex exponentials can be written in terms of sines and cosines, and that the general formula for* ϕ_λ *can then be given as*

$$\phi_\lambda(x) = c_1 \cos(\nu x) + c_2 \sin(\nu x)$$

where, again, c_1 *and* c_2 *are arbitrary constants, and, as already stated,* $\nu = \sqrt{\lambda}$.

Applying the first boundary condition:

$$\begin{aligned} 0 &= \phi_\lambda(0) \\ &= c_1 \cos(\nu \cdot 0) + c_2 \sin(\nu \cdot 0) \\ &= c_1 \cdot 1 + c_2 \cdot 0 = c_1 \quad , \end{aligned}$$

telling us that

$$\phi_\lambda(x) = c_2 \sin(\nu x) \quad .$$

With the boundary condition at $x = L$*, this gives*

$$0 = \phi_\lambda(L) = c_2 \sin(\nu L) \quad .$$

To avoid triviality, we want c_2 *to be nonzero. So, for the boundary condition at* $x = L$ *to hold, we must have*

$$\sin(\nu L) = 0 \quad ,$$

which means that

$$\nu L = \text{an integral multiple of } \pi \quad ,$$

Moreover, since $\nu = \sqrt{\lambda} > 0$, νL *must be a positive integral multiple of* π *. Thus, we have a list of allowed values of* ν *,*

$$\nu_k = \frac{k\pi}{L} \qquad \text{with} \quad k = 1, 2, 3, \ldots \quad ,$$

a corresponding list of allowed values for λ *,*

$$\lambda_k = (\nu_k)^2 = \left(\frac{k\pi}{L}\right)^2 \qquad \text{with} \quad k = 1, 2, 3, \ldots$$

(these are the eigenvalues), and a corresponding list of $\phi(x)$*'s (the corresponding eigenfunctions),*

$$\phi_k(x) = c_k \sin(\nu_k x) = c_k \sin\left(\frac{k\pi}{L} x\right) \qquad \text{with} \quad k = 1, 2, 3, \ldots$$

where the c_k*'s are arbitrary constants. For our example, these are the only nontrivial eigenfunctions.*

In summary, we have a list of solutions to our "Sturm-Liouville problem", namely,

$$(\lambda_k, \phi_k) \qquad \text{for} \quad k = 1, 2, 3, \ldots$$

where, for each k *,*

$$\lambda_k = \left(\frac{k\pi}{L}\right)^2$$

and

$$\phi_k(x) = c_k \sin(\nu_k x) = c_k \sin\left(\frac{k\pi}{L} x\right)$$

(with the c_k*'s being arbitrary constants).*

In particular, if $L = 1$ *, then the solutions to our "Sturm-Liouville problem" are given by*

$$(\lambda_k, \phi_k(x)) = \left(k^2\pi^2 , c_k \sin(k\pi x)\right) \qquad \text{for} \quad k = 1, 2, 3, \ldots \quad .$$

47.3 The Sturm-Liouville Form for a Differential Equation

To solve these differential equations with parameters, you will probably want to use form (47.3a), as we did in our example. However, to develop and use the theory that we will be developing and using, we will want to rewrite each differential equation in its *Sturm-Liouville form*

$$\frac{d}{dx}\left[p(x)\frac{d\phi}{dx}\right] + q(x)\phi = -\lambda w(x)\phi$$

where p , q and w are known functions.

!►*Example 47.3:* *Technically, the equation*

$$\phi'' + \lambda\phi = 0$$

is not in Sturm-Liouville form, but the equivalent equation

$$\phi'' = -\lambda\phi$$

is, with $p = 1$ *,* $q = 0$ *and* $w = 1$ *.*

!►*Example 47.4:* *The equation*

$$\frac{d}{dx}\left[\sin(x)\frac{d\phi}{dx}\right] + \cos(x)\phi = -\lambda x^2\phi$$

is in Sturm-Liouville form, with

$$p(x) = \sin(x) \quad , \quad q(x) = \cos(x) \quad \text{and} \quad w(x) = x^2 \quad .$$

On the other hand,

$$x\frac{d^2\phi}{dx^2} + 2\frac{d\phi}{dx} + [\sin(x) + \lambda]\phi = 0$$

is not in Sturm-Liouville form.

Fortunately, just about any differential equation in form (47.3a) or (47.3b) can converted to Sturm-Liouville form using a procedure similar to that used to solve first-order linear equations. To describe the procedure in general, let's assume we have at least gotten our equation to the form

$$A(x)\frac{d^2\phi}{dx^2} + B(x)\frac{d\phi}{dx} + C(x)\phi = -\lambda\phi \quad .$$

To illustrate the procedure, we'll use the equation

$$x\frac{d^2\phi}{dx^2} + 2\frac{d\phi}{dx} + \sin(x)\phi = -\lambda\phi$$

(with $(0, \infty)$ being our interval of interest).

Here is what you do:

1. Divide through by $A(x)$, obtaining

$$\frac{d^2\phi}{dx^2} + \frac{B(x)}{A(x)}\phi + \frac{C(x)}{A(x)}\phi = -\lambda\frac{1}{A(x)}\phi \quad .$$

Doing that with our example yields

$$\frac{d^2\phi}{dx^2} + \frac{2}{x}\frac{d\phi}{dx} + \frac{\sin(x)}{x}\phi = -\lambda\frac{1}{x}\phi \quad .$$

2. Compute the "integrating factor"

$$p(x) = e^{\int \frac{B(x)}{A(x)}\,dx} \quad ,$$

ignoring any arbitrary constants.

In our example,

$$\int \frac{B(x)}{A(x)}\,dx = \int \frac{2}{x}\,dx = 2\ln x + c \quad .$$

So (ignoring c),

$$p(x) = e^{\int \frac{B(x)}{A(x)}\,dx} = e^{2\ln x} = e^{\ln x^2} = x^2 \quad .$$

(As an exercise, you should verify that

$$\frac{dp}{dx} = p\frac{B}{A} \quad .$$

In a moment, we'll use this fact.)

3. Using the $p(x)$ just found:

(a) Multiply the differential equation resulting from step 1 by $p(x)$, obtaining

$$p\frac{d^2\phi}{dx^2} + p\frac{B}{A}\frac{d\phi}{dx} + p\frac{C}{A}\phi = -\lambda\frac{p}{A}\phi \quad ,$$

(b) observe that (via the product rule)

$$\frac{d}{dx}\left[p\frac{d\phi}{dx}\right] = p\frac{d^2\phi}{dx^2} + p\frac{B}{A}\frac{d\phi}{dx} \quad ,$$

(c) and rewrite the differential equation according to this observation,

$$\frac{d}{dx}\left[p\frac{d\phi}{dx}\right] + p\frac{C}{A}\phi = -\lambda\frac{p}{A}\phi \quad .$$

In our case, we have

$$x^2\left[\frac{d^2\phi}{dx^2} + \frac{2}{x}\frac{d\phi}{dx} + \frac{\sin(x)}{x}\phi\right] = x^2\left[-\lambda\frac{1}{x}\phi\right]$$

$$\hookrightarrow \qquad x^2\frac{d^2\phi}{dx^2} + 2x\frac{d\phi}{dx} + x\sin(x)\phi = -\lambda x\phi \quad .$$

Oh look! By the product rule,

$$\frac{d}{dx}\left[p\frac{d\phi}{dx}\right] = \frac{d}{dx}\left[x^2\frac{d\phi}{dx}\right] = x^2\frac{d^2\phi}{dx^2} + 2x\frac{d\phi}{dx} \quad .$$

Using this to 'simplify' the first two terms of our last differential equation above, we get

$$\frac{d}{dx}\left[x^2\frac{d\phi}{dx}\right] + x\sin(x)\phi = -\lambda x\phi \quad .$$

It is now in the desired form, with

$$p(x) = x^2 \quad , \quad q(x) = x\sin(x) \quad \text{and} \quad w(x) = x \quad .$$

It will often be convenient to abbreviate the left side of

$$\frac{d}{dx}\left[p(x)\frac{d\phi}{dx}\right] + q(x)\phi = -\lambda w(x)\phi$$

as $\mathcal{L}[\phi]$. In fact, given any sufficiently differentiable function f, we will let

$$\mathcal{L}[f] = \frac{d}{dx}\left[p(x)\frac{df}{dx}\right] + q(x)\phi \quad ,$$

whether or not f is a solution to the corresponding differential equation. With this definition, we can view $\mathcal{L}$ as a process for converting any function f into another function (namely, the function you get after computing $\mathcal{L}[f]$). If we ever need to name this process, we'll call it the *Sturm-Liouville operator* corresponding to a given differential equation in Sturm-Liouville form).[2] Do note that our differential equation can be written more concisely as

$$\mathcal{L}[\phi] = -\lambda w\phi \quad .$$

It will also be worth noting that this operator is linear, that is, if α and β are any two constants, and f and g are any two twice-differentiable functions, then, as you can easily verify yourself,

$$\mathcal{L}[\alpha f + \beta g] = \alpha\mathcal{L}[f] + \beta\mathcal{L}[g] \quad .$$

?►*Exercise 47.1:* *Verify the last claim.*

---

[2] More generally, an *operator* on a set of functions is something that converts each function in the given set to some other function. The derivative, which converts each differentiable function f to its derivative f' is an operator on the set of differentiable functions. A Sturm-Liouville operator, as just defined, is an operator on functions that can be differentiated twice.

47.4 Boundary Conditions for Sturm-Liouville Problems

The boundary conditions for a Sturm-Liouville problem will have to be homogeneous (as defined in the previous chapter). They will also have to satisfy a condition that we will derive in this section allowing us to view the associated Sturm-Liouville operator as being"self adjoint".

Green's Formula, and "Sturm-Liouville Appropriate" Boundary Conditions

To describe the additional boundary conditions needed, we need to derive an identity of Green. You should start this derivation by integration by parts to verify the following lemma:

Lemma 47.6 (the preliminary Green's formula)
Let (a, b) be some finite interval, and let $\mathcal{L}$ be the operator given by

$$\mathcal{L}[\phi] = \frac{d}{dx}\left[p(x)\frac{d\phi}{dx}\right] + q(x)\phi$$

where p and q are any suitably smooth and integrable functions on (a, b). Then

$$\int_a^b f\,\mathcal{L}[g]\,dx = p(x)f(x)\frac{dg}{dx}\bigg|_a^b - \int_a^b p\frac{df}{dx}\frac{dg}{dx}\,dx + \int_a^b qfg\,dx \tag{47.5}$$

whenever f and g are suitably smooth and integrable functions on (a, b).

?►*Exercise 47.2:* *Verify the above lemma. (Hint: Write out the integral on the left of equation (47.5) using the given formula for $\mathcal{L}[\phi]$ and then integrate by parts.)*

Now suppose we have two functions u and v on (a, b) (assumed "suitably smooth and integrable", but, possibly, complex valued), and suppose we want to compare

$$\int_a^b u^*\,\mathcal{L}[v]\,dx \qquad \text{and} \qquad \int_a^b \mathcal{L}[u]^* v\,dx \quad .$$

(Why? Because this will lead to our extending the notion of "self adjointness" as characterized in theorem 47.5 on page 47–8)

If p and q are real-valued functions, then it is trivial to verify that

$$\mathcal{L}[u]^* = \mathcal{L}\big[u^*\big]$$

Using this and the above preliminary Green's formula, we see that

$$\begin{aligned}\int_a^b u^*\,\mathcal{L}[v]\,dx - \int_a^b \mathcal{L}[u]^* v\,dx &= \int_a^b u^*\,\mathcal{L}[v]\,dx - \int_a^b v\,\mathcal{L}\big[u^*\big]\,dx \\ &= \left[pu^*\frac{dv}{dx}\bigg|_a^b - \int_a^b p\frac{du^*}{dx}\frac{dv}{dx}\,dx + \int_a^b qu^*v\,dx\right] \\ &\quad - \left[pv\frac{du^*}{dx}\bigg|_a^b - \int_a^b p\frac{dv}{dx}\frac{du^*}{dx}\,dx + \int_a^b qvu^*\,dx\right] \quad .\end{aligned}$$

Nicely enough, most of the terms on the right cancel out, leaving us with:

Theorem 47.7 (Green's formula)
Let (a,b) be some interval, p and q any suitably smooth and integrable real-valued functions on (a,b), and $\mathcal{L}$ the operator given by

$$\mathcal{L}[\phi] = \frac{d}{dx}\left[p(x)\frac{d\phi}{dx}\right] + q(x)\phi \quad .$$

Then

$$\int_a^b u^* \,\mathcal{L}[v]\, dx - \int_a^b \mathcal{L}[u]^*\, v\, dx = p\left[u^*\frac{dv}{dx} - v\frac{du^*}{dx}\right]\Bigg|_a^b \tag{47.6}$$

for any two suitably smooth and differentiable functions u and v.

Equation (47.6) is known as Green's formula. (Strictly speaking, the right side of the equation is the "Green's formula" for the left side).

Now we can state what sort of boundary conditions are appropriate for our discussions. We will refer to a pair of homogeneous boundary conditions at $x = a$ and $x = b$ as being *Sturm-Liouville appropriate* if and only if

$$p\left[u^*\frac{dv}{dx} - v\frac{du^*}{dx}\right]\Bigg|_a^b = 0 \quad . \tag{47.7}$$

whenever both u and v satisfy these boundary conditions.

For brevity, we may say "appropriate" when we mean "Sturm-Liouville appropriate". Do note that the function p in equation (47.7) comes from the differential equation in the Sturm-Liouville problem. Consequently, it is possible that a particular pair of boundary conditions is "Sturm-Liouville appropriate" when using one differential equation, and not "Sturm-Liouville appropriate" when using a different differential equation. This is particularly true when the boundary conditions are of the periodic or boundedness types.

!►*Example 47.5:* *Consider the boundary conditions*

$$\phi(a) = 0 \quad \text{and} \quad \phi(b) = 0 \quad .$$

If u and v satisfy these conditions; that is,

$$u(a) = 0 \quad \text{and} \quad u(b) = 0$$

and

$$v(a) = 0 \quad \text{and} \quad v(b) = 0 \quad ,$$

then, assuming $p(a)$ and $p(b)$ are finite numbers,

$$p\left[u^*\frac{dv}{dx} - v\frac{du^*}{dx}\right]\Bigg|_a^b = p(b)\left[(u(b))^*\,\frac{dv}{dx}\bigg|_{x=b} - v(b)\,\frac{du^*}{dx}\bigg|_{x=b}\right] - p(a)\left[(u(a))^*\,\frac{dv}{dx}\bigg|_{x=a} - v(a)\,\frac{du^*}{dx}\bigg|_{x=a}\right]$$

$$\begin{aligned} &= p(b)\left[0^* \frac{dv}{dx}\bigg|_{x=b} - 0\,\frac{du^*}{dx}\bigg|_{x=b}\right] \\ &\quad - p(a)\left[0^* \frac{dv}{dx}\bigg|_{x=a} - 0\,\frac{du^*}{dx}\bigg|_{x=a}\right] \\ &= 0 \quad . \end{aligned}$$

So

$$\phi(a) = 0 \qquad \text{and} \qquad \phi(b) = 0 \quad .$$

are "Sturm-Liouville appropriate" boundary conditions, at least whenever $p(a)$ and $p(b)$ are finite.

A really significant observation is that, whenever u and v both satisfy "appropriate" boundary conditions, then Green's formula reduces to

$$\int_a^b u^*\,\mathcal{L}[v]\,dx - \int_a^b \mathcal{L}[u]^*\,v\,dx = 0 \quad ,$$

which, in turn, gives us the following lemma, which, in turn, suggests just what we will be using for "inner products" and "self adjointness" in the near future.

Lemma 47.8
Let (a,b) be some interval, p and q any suitably smooth and integrable real-valued functions on (a,b), and $\mathcal{L}$ the operator given by

$$\mathcal{L}[\phi] = \frac{d}{dx}\left[p(x)\frac{d\phi}{dx}\right] + q(x)\phi \quad .$$

Assume, further, that $u(x)$ and $v(x)$ satisfy Sturm-Liouville appropriate boundary conditions at a and b. Then

$$\int_a^b u^*\,\mathcal{L}[v]\,dx = \int_a^b \mathcal{L}[u]^*\,v\,dx \quad .$$

47.5 Sturm-Liouville Problems Full Definition

Finally, we can fully define the "Sturm-Liouville problem": A *Sturm-Liouville problem* is a boundary-value problem consisting of both of the following:

1. A differential equation that can be written in the form

$$\frac{d}{dx}\left[p(x)\frac{d\phi}{dx}\right] + q(x)\phi = -\lambda w(x)\phi \qquad \text{for} \quad a < x < b \tag{47.8}$$

where p, q and w are sufficiently smooth and integrable functions on the finite interval (a,b), with p and q being real valued, and w being positive on this interval.

2. A corresponding pair of Sturm-Liouville appropriate homogeneous boundary conditions at $x = a$ and $x = b$.

As with any boundary-value problem with parameter, a *solution* to a Sturm-Liouville problem consists of a pair (λ, ϕ) where λ is a constant — called an eigenvalue — and ϕ is a nontrivial (i.e, nonzero) function — called an eigenfunction — which, together, satisfy the given Sturm-Liouville problem. When convenient, we will refer to (λ, ϕ) as an *eigen-pair* for the given Sturm-Liouville problem.

A Important Class of Sturm-Liouville Problems

There are several classes of Sturm-Liouville problems. One particularly important class goes by the name of "regular" Sturm-Liouville problems. A Sturm-Liouville problem is said to be *regular* if and only if all the following hold:

1. The functions p, q and w are all real valued and continuous on the *closed* interval $[a, b]$, with p being differentiable on (a, b), and both p and w being positive on the closed interval $[a, b]$.

2. We have homogeneous regular boundary conditions at both $x = a$ and $x = b$. That is

$$\alpha_a \phi(a) \ + \ \beta_a \phi'(a) \ = \ 0$$

where α_a and β_a are constants, with at least one being nonzero, and

$$\alpha_b \phi(b) \ + \ \beta_b \phi'(b) \ = \ 0$$

where α_a and β_b are constants, with at least one being nonzero.

Many, but not all, of the Sturm-Liouville problems generated in solving partial differential equation problems are "regular". For example, the problem considered in example 47.2 on page 47–9 is a regular Sturm-Liouville problem.

So What?

It turns out that the eigenfunctions from a Sturm-Liouville problem on an interval (a, b) can often be used in much the same way as the eigenvectors from a self-adjoint matrix to form an "orthogonal basis" for a large set of functions. Consider, for example, the eigenfunctions from the example 47.2

$$\phi_k(x) \ = \ c_k \sin\left(\frac{k\pi}{L}\, x\right) \qquad \text{for} \quad k \ = \ 1,\ 2,\ 3,\ \ldots \quad .$$

If f is any reasonable function on $(0, L)$ (say, any continuous function on this interval), then we will discover that there are constants c_1, c_2, c_3, $\ldots$ such that

$$f(x) \ = \ \sum_{k=1}^{\infty} c_k \sin\left(\frac{k\pi}{L} x\right) \qquad \text{for all } x \text{ in } (0, L) \quad .$$

This infinite series, called the (Fourier) sine series for f on $(0, L)$, turns out the be extremely useful in many applications. For one thing, it expresses any function on $(0, L)$ in terms of the well-understood sine functions.

Our next goal is to develop enough of the necessary theory to "discover" what I just said we will discover. In particular we want to learn how to compute the c_k's in the above expression for $f(x)$. It turns out to be remarkable similar to the formula for computing the v_k's in theorem 47.4 on page 47–6. Take a look at it right now. Of course, before we can verify this claim, we will have to find out just what we are using for an "inner product".

All this, alas, will take a few pages.

47.6 The Eigen-Spaces

Suppose we have some Sturm-Liouville problem. For convenience, let us write the differential equation in that problem as

$$\mathcal{L}[\phi] = -\lambda w \phi$$

with, as usual,

$$\mathcal{L}[\phi] = \frac{d}{dx}\left[p(x)\frac{d\phi}{dx}\right] + q(x)\phi \quad .$$

Now let $\phi(x)$ and $\psi(x)$ be two eigenfunctions corresponding to the same eigenvalue λ. Using the linearity you verified earlier, we have

$$\begin{aligned}\mathcal{L}[c_1\phi + c_2\psi] &= c_1\mathcal{L}[\phi] + c_2\mathcal{L}[\psi] \\ &= c_1[-\lambda w\phi] + c_2[-\lambda w\psi] = -\lambda w[c_1\phi + c_2\psi] \quad .\end{aligned}$$

This shows that the linear combination $c_1\phi + c_2\psi$ also satisfies the differential equation with that particular value of λ. Does it also satisfy the boundary conditions? Of course. Remember, the set of boundary conditions in a Sturm-Liouville problem is "homogeneous", meaning that, if $\phi(x)$ and $\psi(x)$ satisfy the given boundary conditions, so does any linear combination of them. Hence, any linear combination of eigenfunctions corresponding to a single eigenvalue is also an eigenfunction for our Sturm-Liouville problem, a fact significant enough to write as a lemma.

Lemma 47.9

Assume (λ, ϕ) and (λ, ψ) are both eigen-pairs with the same eigenvalue λ for some Sturm-Liouville problem. Then any nonzero linear combination of $\phi(x)$ and $\psi(x)$ is also an eigenfunction for the Sturm-Liouville problem corresponding to eigenvalue λ.

This lemma tells us that the set of all eigenfunctions for our Sturm-Liouville problem corresponding to any single eigenvector λ is a vector space of functions (after throwing in the zero function). Naturally, we call this the *eigenspace* corresponding to eigenvalue λ. Keep in mind that these functions are all solutions to the second-order homogeneous linear equation

$$\frac{d}{dx}\left[p(x)\frac{d\phi}{dx}\right] + [q(x) + \lambda w(x)]\phi = 0 \quad ,$$

and that the general solution to such a differential equation can be written as

$$\phi(x) = c_1\phi_1(x) + c_2\phi_2(x)$$

where c_1 and c_2 are arbitrary constants and $\{\phi_1, \phi_2\}$ is any linearly independent pair of solutions to the differential equation. One of these two solutions, say, ϕ_1, can be chosen as one of the eigenfunctions. Whether or not ϕ_2 can be chosen to be an eigenfunction depends on whether or not there is a linearly independent pair of eigenfunctions corresponding to this λ. That gives us exactly two possibilities:

1. There is a linearly independent pair of eigenvectors corresponding to λ. This means the eigenspace corresponding to eigenvalue λ is two dimensional (i.e., λ is a 'double' eigenvalue), and every solution to the differential equation (with the given λ) is an eigenfunction.

2. There is not a linearly independent pair of eigenvectors corresponding to λ. This means the eigenspace corresponding to eigenvalue λ is one dimensional (i.e., λ is a 'simple' eigenvalue), and every eigenfunction is a constant multiple of ϕ_1.

47.7 Inner Products, Orthogonality and Generalized Fourier Series

We must, briefly, forget about Sturm-Liouville problems to develop the basic analog of the inner product described for finite dimensional vectors at the beginning of this chapter. Throughout this discussion, (a, b) is an interval, and $w = w(x)$ is a function on (a, b) satisfying

$$w(x) > 0 \qquad \text{whenever} \quad a < x < b \quad .$$

This function, w, will be called the *weight function* for our inner product. Often, it is simply the constant 1 function (i.e., $w(x) = 1$ for every x).

Also, throughout this discussion, let us assume that our functions are at least piecewise continuous and "sufficiently integrable" for the computations being described. We'll discuss what "suitably integrable" means if I decide it is relevant. Otherwise, we'll slop over issues of integrability.

Inner Products for Functions

Let f and g be two functions on the interval (a, b). We define the *inner product (with weight function w)* of f with g — denoted $\langle f \mid g \rangle$ — by

$$\langle f \mid g \rangle = \int_a^b f(x)^* g(x) w(x)\, dx \quad .$$

For convenience, let's say that the *standard inner product* is simply the inner product with weight function $w = 1$,

$$\langle f \mid g \rangle = \int_a^b f(x)^* g(x)\, dx \quad .$$

For now, you can consider the inner product to be simply shorthand for some integrals that will often arise in our work.

!►*Example 47.6:* *Let* $(a,b) = (0,2)$ *and* $w(x) = x^2$ *. Then*

$$\langle f \mid g \rangle = \int_0^2 f(x)^* g(x) x^2 \, dx \quad .$$

In particular

$$\left\langle 6x + 2i \,\middle|\, \frac{1}{x} \right\rangle = \int_0^2 (6x+2i)^* \cdot \frac{1}{x} \cdot x^2 \, dx$$
$$= \int_0^2 (6x - 2i) x \, dx$$
$$= \int_0^2 \left[6x^2 - i2x \right] dx$$
$$= 2x^3 - ix^2 \Big|_0^2 = 16 - 4i \quad .$$

The inner product of functions just defined is, in many ways, analogous to the inner product defined for finite dimensional vectors at the beginning of this chapter. To see this, we'll verify the following theorem, which is very similar to theorem 47.2 on page 47–4.

Theorem 47.10 (properties of the inner product)
Let $\langle \cdot | \cdot \rangle$ *be an inner product as just defined above. Suppose* α *and* β *are two (possibly complex) constants, and* f , g *, and* h *are functions on* (a,b) *. Then*

1. $\langle f \mid g \rangle = \langle g \mid f \rangle^*$,
2. $\langle h \mid \alpha f + \beta g \rangle = \alpha \langle h \mid f \rangle + \beta \langle h \mid g \rangle$,
3. $\langle \alpha f + \beta g \mid h \rangle = \alpha^* \langle f \mid h \rangle + \beta^* \langle g \mid h \rangle$,

and

4. $\langle f \mid f \rangle \geq 0$ *with* $\langle f \mid f \rangle = 0$ *if and only if* $f = 0$ *on* (a,b) *.*

PROOF: `To Be Written (Someday) --- Take Notes in Class`

Norms

Recall that the norm of any vector $\mathbf{v}$ in $\mathbb{C}^N$ is related to the inner product of $\mathbf{v}$ with itself by

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v} \mid \mathbf{v} \rangle} \quad .$$

In turn, for each inner product $\langle \cdot \mid \cdot \rangle$, we define the corresponding *norm* of a function f by

$$\|f\| = \sqrt{\langle f \mid f \rangle} \quad .$$

In general

$$\|f\|^2 = \langle f \mid f \rangle = \int_a^b f(x)^* f(x) w(x) \, dx = \int_a^b |f(x)|^2 \, w(x) \, dx \quad .$$

Do note that, in a loose sense,

$$\text{“ } f \text{ is generally small over } (a,b) \text{”} \iff \text{“ } \|f\| \text{ is small”} \quad .$$

!►Example 47.7: *Let* $(a,b) = (0,2)$ *and* $w(x) = x^2$ *. Then*

$$\|f\| = \sqrt{\langle\, f \mid f \,\rangle} = \sqrt{\int_0^2 (f(x))^* f(x) x^2\, dx} = \sqrt{\int_0^2 |f(x)|^2\, x^2\, dx} \quad .$$

In particular,

$$\begin{aligned}
\|5x+6i\|^2 &= \int_0^2 (5x+6i)^*(5x+6i)x^2\, dx \\
&= \int_0^2 (5x-6i)(5x+6i)x^2\, dx \\
&= \int_0^2 \left[25x^2+36\right] x^2\, dx \\
&= \int_0^2 \left[25x^4+36x^2\right]\, dx \\
&= 5x^5 + 12x^3\Big|_0^2 \\
&= 256 \quad .
\end{aligned}$$

So

$$\|5x+6i\| = \sqrt{256} = 16 \quad .$$

Orthogonality

Recall that any two vectors $\mathbf{v}$ and $\mathbf{w}$ in $\mathbb{C}^N$ are orthogonal if and only if

$$\langle\, \mathbf{v} \mid \mathbf{w} \,\rangle = 0 \quad .$$

Analogously, we say that any pair of functions f and g is *orthogonal* (over the interval) (with respect to the inner product, or with respect to the weight function) if and only if

$$\langle\, f \mid g \,\rangle = 0 \quad .$$

More generally, we will refer to any indexed set of nonzero functions

$$\{\, \phi_1,\ \phi_2,\ \phi_3,\ \dots \,\}$$

as being *orthogonal* if and only if

$$\langle\, \phi_k \mid \phi_n \,\rangle = 0 \qquad \text{whenever} \quad k \neq n \quad .$$

If, in addition, we have

$$\|\phi_k\| = 1 \qquad \text{for each} \quad k \quad ,$$

then we say the set is *orthonormal*. For our work, orthogonality will be important, but we won't spend time or effort making the sets orthonormal.

!▶*Example 47.8:* *Consider the set*

$$\left\{ \sin\left(\frac{k\pi}{L}x\right) : k = 1, 2, 3, \dots \right\}$$

which is the set of sine functions (without the arbitrary constants) obtained as eigenfunctions in example 47.2 on page 47–9. The interval is $(0, L)$. For the weight function, we'll use $w(x) = 1$. Observe that, if k and n are two different positive integers, then, using the trigonometric identity

$$2\sin(A)\sin(B) = \cos(A - B) - \cos(A + B) \quad ,$$

we have

$$\begin{aligned} \left\langle \sin\left(\frac{k\pi}{L}x\right) \,\middle|\, \sin\left(\frac{n\pi}{L}x\right) \right\rangle &= \int_0^L \left(\sin\left(\frac{k\pi}{L}x\right)\right)^* \sin\left(\frac{n\pi}{L}x\right) dx \\ &= \int_0^L \sin\left(\frac{k\pi}{L}x\right) \sin\left(\frac{n\pi}{L}x\right) dx = \cdots = 0 \quad . \end{aligned}$$

So

$$\left\{ \sin\left(\frac{k\pi}{L}x\right) : k = 1, 2, 3, \dots \right\}$$

is an orthogonal set of functions on $(0, L)$ with respect to the weight function $w(x) = 1$.

Generalized Fourier Series

Now suppose $\{\phi_1, \phi_2, \phi_3, \dots\}$ is some orthogonal set of nonzero functions on (a, b), and f is some function that can be written as a (possibly infinite) linear combination of the ϕ_k's,

$$f(x) = \sum_k c_k \phi_k(x) \qquad \text{for} \quad a < x < b \quad .$$

To find each constant c_k, first observe what happens when we take the inner product of both sides of the above with one of the ϕ_k's, say, ϕ_3. Using the linearity of the inner product and the orthogonallity of our functions, we get

$$\begin{aligned} \langle \phi_3 \mid f \rangle &= \left\langle \phi_3 \,\middle|\, \sum_k c_k \phi_k \right\rangle \\ &= \sum_k c_k \langle \phi_3 \mid \phi_k \rangle \\ &= \sum_k c_k \left\{ \begin{matrix} \|\phi_3\|^2 & \text{if} \quad k = 3 \\ 0 & \text{if} \quad k \neq 3 \end{matrix} \right\} = c_3 \|\phi_3\|^2 \quad . \end{aligned}$$

So

$$c_3 = \frac{\langle \phi_3 \mid f \rangle}{\|\phi_3\|^2} \quad .$$

Since there is nothing special about $k = 3$, we clearly have

$$c_k = \frac{\langle \phi_k \mid f \rangle}{\|\phi_k\|^2} \qquad \text{for all} \quad k \quad .$$

More generally, whether or not f can be expressed as a linear combination of the ϕ_k's, we define the *generalized Fourier series* for f (with respect to the given inner product and orthogonal set $\{\phi_1, \ldots, \}$) to be

$$G.F.S.[f]|_x = \sum_k c_k \phi_k(x)$$

where, for each k

$$c_k = \frac{\langle \phi_k \mid f \rangle}{\|\phi_k\|^2} \quad .$$

The c_k's are called the corresponding *generalized Fourier coefficients* of f.[3] Don't forget:

$$\langle \phi_k \mid f \rangle = \int_a^b \phi_k(x)^* f(x) w(x)\, dx$$

and

$$\|\phi_k\|^2 = \int_a^b |\phi_k(x)|^2 w(x)\, dx \quad .$$

We will also refer to $G.F.S.[f]$ as the *expansion* of f in terms of the ϕ_k's. If the ϕ_k's just happen to be eigenfunctions from some Sturm-Liouville problem, we will even refer to $G.F.S.[f]$ as the *eigenfunction expansion* of f.

!►Example 47.9: *From previous exercises, we know*

$$\{\phi_1, \phi_2, \phi_3, \ldots\} = \left\{ \sin\left(\frac{k\pi}{L}x\right) : k = 1, 2, 3, \ldots \right\}$$

is an orthogonal set of functions on $(0, L)$ with respect to the weight function $w(x) = 1$. (Recall that this is a set of eigenfunctions for the Sturm-Liouville problem from example 47.2 on page 47–9.) Using this set,

$$G.F.S.[f]|_x = \sum_k c_k \phi_k(x) \qquad \text{with} \quad c_k = \frac{\langle \phi_k \mid f \rangle}{\|\phi_k\|^2}$$

becomes

$$G.F.S.[f]|_x = \sum_{k=1}^{\infty} c_k \sin\left(\frac{k\pi}{L}x\right)$$

with

$$c_k = \frac{\langle \phi_k \mid f \rangle}{\|\phi_k\|^2} = \frac{\int_0^L \left(\sin\left(\frac{k\pi}{L}x\right)\right)^* f(x)\, dx}{\int_0^L \left|\sin\left(\frac{k\pi}{L}x\right)\right|^2 dx} = \frac{\int_0^L f(x) \sin\left(\frac{k\pi}{L}x\right) dx}{\int_0^L \sin^2\left(\frac{k\pi}{L}x\right) dx} \quad .$$

Since

$$\int_0^L \sin^2\left(\frac{k\pi}{L}x\right) dx = \cdots = \frac{L}{2} \quad .$$

the above reduces to

$$G.F.S.[f]|_x = \sum_{k=1}^{\infty} c_k \sin\left(\frac{k\pi}{L}x\right) \qquad \text{with} \quad c_k = \frac{2}{L}\int_0^L f(x) \sin\left(\frac{k\pi}{L}x\right) dx \quad .$$

[3] Compare the formula for the generalized Fourier coefficients with formula in theorem (47.4) on page 47–6 for the components of a vector with respect to any orthogonal basis. They are virtually the same!

In particular, suppose $f(x) = x$ *for* $0 < x < L$ *. Then the above formula for* c_k *yields*

$$\begin{aligned} c_k &= \frac{2}{L}\int_0^L x \sin\left(\frac{k\pi}{L}x\right) dx \\ &= \frac{2}{L}\left[\frac{-2x}{k\pi}\cos\left(\frac{k\pi}{L}x\right)\Big|_0^L + \frac{2}{k\pi}\int_0^L \cos\left(\frac{k\pi}{L}x\right) dx\right] \\ &= \frac{2}{L}\left[0 + \frac{2L}{k\pi}\cos\left(\frac{k\pi}{L}L\right) + \left(\frac{2}{k\pi}\right)^2 \sin\left(\frac{k\pi}{L}x\right)\Big|_0^L\right] = (-1)^k\frac{4}{k\pi} \quad . \end{aligned}$$

So, using the given interval, weight function and orthogonal set, the generalized Fourier series for $f(x) = x$ *, is*

$$\sum_{k=1}^{\infty}(-1)^k\frac{4}{k\pi}\sin\left(\frac{k\pi}{L}x\right) \quad .$$

(In fact, this is the classic "Fourier sine series for $f(x) = x$ *on* $(0, L)$ *".)*

Approximations and Completeness

Let's say we have an orthogonal set of functions $\{\phi_1, \phi_2, \phi_3, \ldots\}$ and some function f which is "reasonably behaved" on (a, b) (we'll specify just what "reasonably behaved" means later). Let

$$G.F.S.[f]|_x = \sum_{k=1}^{\infty} c_k\,\phi_k(x)$$

In practice, to avoid using the entire series, we may simply wish to approximate f using the N^{th} partial sum,

$$f(x) \approx \sum_{k=1}^{N} c_k\,\phi_k(x) \quad .$$

The error in using this is

$$E_N(x) = f(x) - \sum_{k=1}^{N} c_k\,\phi_k(x) \quad ,$$

and the square of its norm,

$$\begin{aligned} \|E_N\|^2 &= \left\| f(x) - \sum_{k=1}^{N} c_k\,\phi_k(x)\right\|^2 \\ &= \int_a^b \left| f(x) - \sum_{k=1}^{N} c_k\,\phi_k(x)\right|^2 w(x)\,dx \quad , \end{aligned}$$

give a convenient measure of how good this approximation is. The above integral is sometimes known as the "(weighted) mean square error in using $\sum_{k=1}^{N} c_k\,\phi_k(x)$ for $f(x)$ on the interval (a, b)" .

We, of course, hope the error shrinks to zero (as measured by $\|E_N\|$) as $N \to \infty$. If we can be sure this happens no matter what (reasonably behaved) function f we start with, then we say the orthogonal set $\{\phi_1, \phi_2, \phi_3, \ldots\}$ is *complete*.

Now if $\{\phi_1, \phi_2, \phi_3, \ldots\}$ is complete, then taking the limits above yield

$$\|f - G.F.S.[f]\| = 0 \quad .$$

Equivalently

$$\int_a^b \left| f(x) - \sum_{k=1}^{\infty} c_k \phi_k(x) \right|^2 w(x)\,dx = 0 \quad .$$

In practice, this usually means that the infinite series $\sum_k c_k\phi_k(x)$ converges to $f(x)$ at every x in (a,b) at which f is continuous. In any case, if the set $\{\phi_1, \phi_2, \phi_3, \ldots\}$ is complete, then we can view the corresponding generalized Fourier series for a function f as being the same as that function, and can write

$$f(x) = \sum_k c_k \phi_k(x) \qquad \text{for} \quad a < x < b$$

where

$$c_k = \frac{\langle \phi_k \mid f \rangle}{\|\phi_k\|^2} \quad .$$

In other words, a complete orthogonal set of (nonzero) functions can be viewed as a basis for the vector space of (reasonably behaved) functions on the interval.

47.8 Sturm-Liouville Problems and Eigenfunction Expansions

Suppose we have some Sturm-Liouville problem with differential equation

$$\frac{d}{dx}\left[p(x)\frac{d\phi}{dx}\right] + q(x)\phi = -\lambda w\phi \qquad \text{for} \quad a < x < b$$

and Sturm-Liouville appropriate boundary conditions. As usual, we'll let $\mathcal{L}$ be the operator given by

$$\mathcal{L}[\phi] = \frac{d}{dx}\left[p(x)\frac{d\phi}{dx}\right] + q(x)\phi \quad .$$

Remember, by definition, p and q are real-valued functions and w is a positive function on (a,b).

Self-Adjointness and Some Immediate Results

Glance back at lemma 47.8 on page 47–17. It tells us that, whenever u and v are sufficiently differentiable functions satisfying the boundary conditions,

$$\int_a^b u^* \mathcal{L}[v]\,dx = \int_a^b \mathcal{L}[u]^* v\,dx \quad .$$

Using the standard inner product for functions on (a,b),

$$\langle f \mid g \rangle = \int_a^b f(x)^* g(x)\,dx \quad ,$$

we can write this equality as

$$\langle\, u \mid \mathcal{L}[v]\,\rangle \;=\; \langle\, \mathcal{L}[u] \mid v\,\rangle \quad ,$$

which looks very similar to the equation characterizing self-adjointness for a matrix $\mathbf{A}$ in theorem 47.5 on page 47–8. Because of this we often say either that Sturm-Liouville problems are *self adjoint*, or that the operator $\mathcal{L}$ is *self adjoint*.[4] The terminology is important for communication, but what is even more important is that functional analogs to the results obtained for self-adjoint matrices also hold for here.

To derive two important results, let (λ_1, ϕ_1) and (λ_2, ϕ_2) be two solutions to our Sturm-Liouville problem (i.e., λ_1 and λ_2 are two eigenvalues, and ϕ_1 and ϕ_2 are corresponding eigenfunctions). From a corollary to Green's formula (lemma 47.8, noted just above), we know

$$\int_a^b \phi_1{}^* \, \mathcal{L}[\phi_2] \, dx \;=\; \int_a^b \mathcal{L}[\phi_1]^* \, \phi_2 \, dx \quad .$$

But, from the differential equation in the problem, we also have

$$\mathcal{L}[\phi_2] \;=\; -\lambda_2 w \phi_2 \quad ,$$

$$\mathcal{L}[\phi_1] \;=\; -\lambda_1 w \phi_1$$

and thus,

$$\mathcal{L}[\phi_1]^* \;=\; (-\lambda_1 w \phi_1)^* \;=\; -\lambda_1{}^* w \phi_1{}^*$$

(remember w is a positive function). So,

$$\int_a^b \phi_1{}^* \, \mathcal{L}[\phi_2] \, dx \;=\; \int_a^b \mathcal{L}[\phi_1]^* \, \phi_2 \, dx$$

$$\hookrightarrow \qquad \int_a^b \phi_1{}^*(-\lambda_2 w \phi_2) \, dx \;=\; \int_a^b (-\lambda_1{}^* w \phi_1{}^*)\phi_2 \, dx$$

$$\hookrightarrow \qquad -\lambda_2 \int_a^b \phi_1{}^* \phi_2 w \, dx \;=\; -\lambda_1{}^* \int_a^b \phi_1{}^* \phi_2 w \, dx \quad .$$

Since the integrals on both sides of the last equation are the same, we must have either

$$\lambda_2 \;=\; \lambda_1{}^* \qquad \text{or} \qquad \int_a^b \phi_1{}^* \phi_2 w \, dx \;=\; 0 \quad . \tag{47.9}$$

Now, we did not necessarily assume the solutions were different. If they are the same,

$$(\lambda_1, \phi_1) \;=\; (\lambda_2, \phi_2) \;=\; (\lambda, \phi)$$

and the above reduces to

$$\lambda \;=\; \lambda^* \qquad \text{or} \qquad \int_a^b \phi^* \phi w \, dx \;=\; 0 \quad .$$

But, since w is a positive function and ϕ is necessarily nontrivial,

$$\int_a^b \phi^* \phi w \, dx \;=\; \int_a^b |\phi(x)|^2 \, w(x) \, dx \;>\; 0 \quad (\neq 0) \quad .$$

[4] The correct terminology is that "$\mathcal{L}$ is a *self-adjoint operator* on the vector space of twice-differentiable functions satisfying the given homogeneous boundary conditions".

So we must have

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

which is only possible if λ is a *real* number. Thus,

> ***FACT:*** *The eigenvalues are all real numbers.*

Now suppose λ_1 and λ_2 are not the same. Then, since they are different *real* numbers, we certainly do not have

$$\lambda_2 = {\lambda_1}^* \quad .$$

Line (47.9) then tells us that we must have

$$\int_a^b {\phi_1}^* \phi_2 w \, dx = 0 \quad ,$$

which we can also write as

$$\langle \phi_1 \mid \phi_2 \rangle = 0$$

using the inner product with weight function w,

$$\langle f \mid g \rangle = \int_a^b f(x)^* g(x) w(x) \, dx \quad .$$

Thus,

> ***FACT:*** *Any pair of eigenfunctions corresponding to different eigenvalues are orthogonal with respect to the inner product with weight function* $w(x)$.

By the way, since the eigenvalues are real, it is fairly easy to show that the real part and the imaginary part of each eigenfunction is also an eigenfunction. From this it follows that we can always choose real-valued functions as our basis for each eigenspace.

What all the above means, at least in part, is that we will be constructing generalized Fourier series using eigenfunctions from Sturm-Liouville problems, and that the inner product used will be based on the weight function $w(x)$ from the differential equation

$$\frac{d}{dx}\left[p(x)\frac{d\phi}{dx}\right] + q(x)\phi = -\lambda w \phi \qquad \text{for} \quad a < x < b \quad .$$

in the Sturm-Liouville problem. Accordingly, we may refer to $w(x)$ and the corresponding inner product on (a, b) as the *natural weight function* and the *natural inner product* corresponding to the given Sturm-Liouville problem.

Other Results Concerning Eigenvalues

It can be shown that there is a *smallest* eigenvalue λ_0 for each Sturm-Liouville problem normally encountered in practice. The rest of the eigenvalues are larger. Unfortunately, verifying this is beyond our ability (unless the Sturm-Liouville problem is sufficiently simple). Another fact you will just have to accept without proof is that, for each Sturm-Liouville problem normally encountered in practice, the eigenvalues form an infinite increasing sequence

$$\lambda_0 < \lambda_1 < \lambda_2 < \lambda_3 < \cdots$$

with

$$\lim_{k \to \infty} \lambda_k = \infty \quad .$$

The Eigenfunctions

Let us assume that

$$\mathcal{E} = \{\lambda_0, \lambda_1, \lambda_2, \lambda_3, \ldots\}$$

is the set of all distinct eigenvalues for our Sturm-Liouville problem (indexed so that λ_0 is the smallest and $\lambda_k < \lambda_{k+1}$ in general). Remember, each eigenvalue will be either a simple or a double eigenvalue. Next, choose a set of eigenfunctions

$$\mathcal{B} = \{\phi_0, \phi_1, \phi_2, \phi_3, \ldots\}$$

as follows:

1. For each simple eigenvalue, choose exactly one corresponding eigenfunction for $\mathcal{B}$.
2. For each double eigenvalue, choose exactly one orthogonal pair of corresponding (real-valued) eigenfunctions for $\mathcal{B}$.

Remember, this set of functions will be orthogonal with respect to the weight function w from the differential equation in the Sturm-Liouville problem. (Note: Each ϕ_k is an eigenfunction corresponding to eigenvalue λ_k only if all the eigenvalues are simple.)

Now let f be a function on (a, b). Since $\mathcal{B}$ is an orthogonal set, we can construct the corresponding generalized Fourier series for f

$$G.F.S.[f(x)] = \sum_{k=0}^{\infty} c_k \phi_k(x)$$

with

$$c_k = \frac{\langle \phi_k \mid f \rangle}{\|\phi_k\|^2} = \frac{\int_a^b \phi_k(x)^* f(x) w(x)\, dx}{\int_a^b |\phi_k(x)|^2\, w(x)\, dx} .$$

The obvious question to now ask is "Is this orthogonal set of eigenfunctions complete?" That is, can we assume

$$f = \sum_{k=0}^{\infty} c_k \phi_k \qquad \text{on} \quad (a, b) \quad ?$$

The answer is yes, at least for the Sturm-Liouville problems normally encountered in practice. But you will have to trust me on this. There are also some issues regarding the convergence of $\sum_{k=0}^{\infty} c_k \phi_k(x)$ when x is a point at which f is discontinuous or when x is an endpoint of (a, b) and f does not satisfy the same boundary conditions as in the Sturm-Liouville problem, but we will gloss over those issues for now.

Finally (assuming "reasonable" assumptions concerning the functions in the differential equation), it can be shown that the graphs of the eigenfunctions corresponding to higher values of the eigenvalues "wiggle" more than those corresponding to the lower-valued eigenvalues. To be precise, eigenfunctions corresponding to higher values of the eigenvalues must cross the X–axis (i.e., be zero) more oftern that do those corresponding to the lower-valued eigenvalues. To see this (sort of), suppose ϕ_0 is never zero on (a, b) (so, it hardly wiggles — this is typically the case with ϕ_0). So ϕ_0 is either always positive or always negative on (a, b). Since $\pm\phi_0$ will also be an eigenfunction, we can assume we've chosen ϕ_0 to always be positive on the interval. Now let ϕ_k be an eigenfunction corresponding to another eigenvalue. If it, too, is never zero on

(a, b), then, as with ϕ_0, we can assume we've chosen ϕ_k to always be positive on (a, b). But then,

$$\langle\, \phi_0 \mid \phi_k \,\rangle \;=\; \int_a^b \underbrace{\phi_0(x)\phi_k(x)w(x)}_{>0}\, dx \;>\; 0 \quad ,$$

contrary to the known orthogonality of eigenfunctions corresponding to different eigenvalues. Thus, each ϕ_k other than ϕ_0 must be zero at least at one point in (a, b). This idea can be extended, showing that eigenfunctions corresponding to high-valued eigenvalues cross the X–axis more often than do those corresponding to lower-valued eigenvalues, but requires developing much more differential equation theory than we have time (or patience) for.

47.9 The Main Results Summarized (Sort of)

A Mega-Theorem

We have just gone through a general discussion of the general things that can (often) be derived regarding the solutions to Sturm-Liouville problems. Precisely what can be proven depends somewhat on the problem. Here is one standard theorem that can be found (usually unproven) in many texts on partial differential equations and mathematical physics. It concerns the regular Sturm-Liouville problems (see page 47–18).

Theorem 47.11 (Mega-Theorem on Regular Sturm-Liouville problems)
Consider a regular Sturm-Liouville problem[5] with differential equation

$$\frac{d}{dx}\left[p(x)\frac{d\phi}{dx}\right] + q(x)\phi \;=\; -\lambda w(x)\phi \qquad \text{for} \quad a < x < b$$

and homogeneous regular boundary conditions at the endpoints of the finite interval (a, b). Then, all of the following hold:

1. *All the eigenvalues are real.*
2. *The eigenvalues form an ordered sequence*

   $$\lambda_0 \;<\; \lambda_1 \;<\; \lambda_2 \;<\; \lambda_3 \;<\; \cdots \quad ,$$

   with a smallest eigenvalue (denoted here by λ_0) and no largest eigenvalue.[6] In fact,

   $$\lambda_k \;\to\; \infty \qquad \text{as} \quad k \;\to\; \infty \quad .$$

3. *All the eigenvalues are simple.*
4. *If, for each eigenvalue λ_k, we choose a corresponding eigenfunction ϕ_k, then the set*

   $$\mathcal{B} \;=\; \{\phi_0, \phi_1, \phi_2, \phi_3, \ldots\}$$

[5] i.e., p, q and w are real valued and continuous on the *closed* interval $[a, b]$, with p being differentiable on (a, b), and both p and w being positive on the closed interval $[a, b]$.

[6] In practice, it may be more convenient to index the eigenvalues so that λ_1 is the smallest.

is a complete, orthonormal set of functions relative to the inner product

$$\langle\, u \mid v \,\rangle \;=\; \int_a^b u(x)^* v(x) w(x)\, dx \quad .$$

Moreover, if f *is any piecewise smooth function on* (a, b) *, then the corresponding generalized Fourier series of* f *,*

$$G.F.S.[f] \;=\; \sum_{k=0}^{\infty} c_k\, \phi_k(x) \qquad \text{with} \quad c_k \;=\; \frac{\langle\, \phi_k \mid f \,\rangle}{\|\phi_k\|^2} \quad ,$$

converges for each x *in* (a, b) *, and it converges to*

(a) $f(x)$ *if* f *is continuous at* x *.*

(b) *the "midpoint of the jump in* f *",*

$$\lim_{\epsilon \to 0^+} \frac{1}{2}\,[f(x+\epsilon) + f(x-\epsilon)] \quad ,$$

if f *is discontinuous at* x *.*

5. *The eigenfunction* ϕ_0 *is never zero on* (a, b) *. Moreover, for each* k *, there are exactly* k *points in* (a, b) *at which* ϕ_k *is zero.*[7]

Similar mega-theorems can be proven for other Sturm-Liouville problems. The main difference occurs when we have periodic boundary conditions. Then most of the eigenvalues are *double* eigenvalues, and our complete set of eigenfunctions looks like

$$\{\, \ldots,\; \phi_k,\; \psi_k,\; \ldots \}$$

where $\{\phi_k, \psi_k\}$ is an orthogonal pair of eigenfunctions corresponding to eigenvalue λ_k . (Typically, though, the smallest eigenvalue is still simple.)

⋆ ⋆ ⋆

[7] This assumes the indexing given in this theorem. If the indexing is so that the smallest eigenvalue is λ_1 , then ϕ_k will be zero at exactly $k-1$ points in (a, b) .

Additional Exercises

47.3. *Let* $\lambda_1 = 3$, $\lambda_2 = -2$,

$$\mathbf{b}^1 = \begin{bmatrix} 2 \\ 3 \end{bmatrix} \quad \text{and} \quad \mathbf{b}^2 = \begin{bmatrix} -3 \\ 2 \end{bmatrix} \quad ,$$

and assume $(\lambda_1, \mathbf{b}^1)$ *and* $(\lambda_2, \mathbf{b}^2)$ *are eigenpairs for a matrix* $\mathbf{A}$.

a. *Verify that* $\{\mathbf{b}^1, \mathbf{b}^2\}$ *is an orthogonal set.*

b. *Compute* $\|\mathbf{b}^1\|$ *and* $\|\mathbf{b}^2\|$.

c. *Express each of the following vectors in terms of* $\mathbf{b}^1$ *and* $\mathbf{b}^2$ *using the formulas in theorem 47.4 on page 47–6.*

i. $\mathbf{u} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$ ***ii.*** $\mathbf{v} = \begin{bmatrix} 12 \\ 5 \end{bmatrix}$ ***iii.*** $\mathbf{w} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

d. *Compute the following using the vectors from the previous part. Leave your answers in terms of* $\mathbf{b}^1$ *and* $\mathbf{b}^2$.

i. $\mathbf{Au}$ ***ii.*** $\mathbf{Av}$ ***iii.*** $\mathbf{Aw}$

47.4. *Let* $\lambda_1 = 2$, $\lambda_2 = 4$, $\lambda_3 = 6$,

$$\mathbf{b}^1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \quad , \quad \mathbf{b}^2 = \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{b}^3 = \begin{bmatrix} 1 \\ 2 \\ -5 \end{bmatrix} \quad .$$

and assume $(\lambda_1, \mathbf{b}^1)$, $(\lambda_2, \mathbf{b}^2)$ *and* $(\lambda_3, \mathbf{b}^3)$ *are eigenpairs for a matrix* $\mathbf{A}$.

a. *Verify that* $\{\mathbf{b}^1, \mathbf{b}^2, \mathbf{b}^3\}$ *is an orthogonal set.*

b. *Compute* $\|\mathbf{b}^1\|$, $\|\mathbf{b}^2\|$ *and* $\|\mathbf{b}^3\|$.

c. *Express each of the following vectors in terms of* $\mathbf{b}^1$, $\mathbf{b}^2$ *and* $\mathbf{b}^3$ *using the formulas in theorem 47.4 on page 47–6.*

i. $\mathbf{u} = \begin{bmatrix} 4 \\ -7 \\ 22 \end{bmatrix}$ ***ii.*** $\mathbf{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ ***iii.*** $\mathbf{w} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$

d. *Compute the following using the vectors from the previous part. Leave your answers in terms of* $\mathbf{b}^1$, $\mathbf{b}^2$ *and* $\mathbf{b}^3$.

i. $\mathbf{Au}$ ***ii.*** $\mathbf{Av}$ ***iii.*** $\mathbf{Aw}$

47.5. *Let* $\lambda_1 = -1$, $\lambda_2 = 0$, $\lambda_3 = 1$,

$$\mathbf{b}^1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad , \quad \mathbf{b}^2 = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \quad \text{and} \quad \mathbf{b}^3 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \quad .$$

and assume $(\lambda_1, \mathbf{b}^1)$, $(\lambda_2, \mathbf{b}^2)$ *and* $(\lambda_3, \mathbf{b}^3)$ *are eigenpairs for a matrix* $\mathbf{A}$.

a. *Verify that* $\{\mathbf{b}^1, \mathbf{b}^2, \mathbf{b}^3\}$ *is an orthogonal set.*

b. *Compute* $\left\|\mathbf{b}^1\right\|$, $\left\|\mathbf{b}^2\right\|$ *and* $\left\|\mathbf{b}^3\right\|$.

c. *Express each of the following vectors in terms of* $\mathbf{b}^1$, $\mathbf{b}^2$ *and* $\mathbf{b}^3$ *using the formulas in theorem 47.4 on page 47–6.*

i. $\mathbf{u} = \begin{bmatrix} -5 \\ 8 \\ 3 \end{bmatrix}$ **ii.** $\mathbf{v} = \begin{bmatrix} 0 \\ -3 \\ 6 \end{bmatrix}$ **iii.** $\mathbf{w} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

d. *Compute the following using the vectors from the previous part. Leave your answers in terms of* $\mathbf{b}^1$, $\mathbf{b}^2$ *and* $\mathbf{b}^3$.

i. $\mathbf{Au}$ **ii.** $\mathbf{Av}$ **iii.** $\mathbf{Aw}$

47.6. *Verify each of the claims in theorem 47.2 on page 47–4.*

47.7. *Finish verifying the claims in theorem 47.5 on page 47–8. In particular show that, if* $\mathbf{B}$ *and* $\mathbf{A}$ *are* $N \times N$ *matrices such that*

$$\langle\, \mathbf{v} \mid \mathbf{Bu} \,\rangle = \langle\, \mathbf{Av} \mid \mathbf{u} \,\rangle \qquad \text{for every} \quad \mathbf{v}, \mathbf{u} \text{ in } \mathbb{R}^N \quad ,$$

then $\mathbf{B} = \mathbf{A}^\dagger$.

47.8. *Find the general solutions for each of the following corresponding to each real value* λ . *Be sure the state the values of* λ *for which each general solution is valid. (Note: Some of these are Euler equations — see chapter 19.)*

a. $\phi'' + 4\phi = -\lambda\phi$

b. $\phi'' + 2\phi' = -\lambda\phi$

c. $x^2\phi'' + x\phi' = -\lambda\phi \qquad \text{for} \quad 0 < x$

d. $x^2\phi'' + 3x\phi' = -\lambda\phi \qquad \text{for} \quad 0 < x$

47.9. *Rewrite each of the following equations in Sturm-Liouville form.*

a. $\phi'' + 4\phi + \lambda\phi = 0$

b. $\phi'' + 2\phi' = -\lambda\phi$

c. $x^2\phi'' + x\phi' = -\lambda\phi \qquad \text{for} \quad 0 < x$

d. $x^2\phi'' + 3x\phi' = -\lambda\phi \qquad \text{for} \quad 0 < x$

e. $\phi'' - 2x\phi' = -\lambda\phi$ *(Hermite's Equation)*

f. $\left(1 - x^2\right)\phi'' - x\phi' = -\lambda\phi \qquad \text{for} \quad -1 < x < 1$ *(Chebyshev's Equation)*

g. $x\phi'' + (1 - x)\phi' = -\lambda\phi \qquad \text{for} \quad 0 < x$ *(Laguerre's Equation)*

47.10. *Solve the Sturm-Liouville problems consisting of the differential equation*

$$\phi'' + \lambda\phi = 0 \qquad \text{for} \quad 0 < x < L$$

(where L *is some finite positive length) and each of the following sets of boundary conditons:*

a. $\phi'(0) = 0$ *and* $\phi'(L) = 0$

b. $\phi(0) = 0$ *and* $\phi'(L) = 0$

47.11. *Solve the following Sturm-Liouville problems:*

a. $\phi'' = -\lambda\phi$ *for* $0 < x < 2\pi$ *with* $\phi(0) = \phi(2\pi)$ *and* $\phi'(0) = \phi'(2\pi)$

b. $x^2\phi'' + x\phi' = -\lambda\phi$ *for* $1 < x < e^\pi$ *with* $\phi(1) = 0$ *and* $\phi(e^\pi) = 0$

c. $x^2\phi'' + 3x\phi' = -\lambda\phi$ *for* $1 < x < e^\pi$ *with* $\phi(1) = 0$ *and* $\phi(e^\pi) = 0$

47.12. *Let* $\mathcal{L}$ *be the Sturm-Liouville operator*

$$\mathcal{L}[\phi] = \frac{d}{dx}\left[p(x)\frac{d\phi}{dx}\right] + q(x)\phi$$

where p *and* q *are continuous functions on the closed interval* $[a, b]$ *.*

a. *Verify that the boundary conditions*

$$\phi'(a) = 0 \qquad \text{and} \qquad \phi'(b) = 0$$

is a Sturm-Liouville appropriate set of boundary conditions.

b. *Show that any pair of homogeneous regular boundary conditions at* a *and* b *is Sturm-Liouville appropriate.*

c. *Suppose our boundary conditions are periodic; that is,*

$$\phi(a) = \phi(b) \qquad \text{and} \qquad \phi'(a) = \phi'(b) \quad .$$

What must p *satisfy for these boundary conditions to be Sturm-Liouville appropriate?*

47.13. *Compute the following, assuming the interval is* $(0, 3)$ *and the weight function is* $w(x) = 1$ *.*

a. $\langle\, x \mid \sin(2\pi x) \,\rangle$

b. $\left\langle\, x^2 \,\middle|\, 9 + i8x \,\right\rangle$

c. $\left\langle\, 9 + i8x \,\middle|\, x^2 \,\right\rangle$

d. $\left\langle\, e^{i2\pi x} \,\middle|\, x \,\right\rangle$

e. $\|x\|$

f. $\|9 + i8x\|$

g. $\|\sin(2\pi x)\|$

h. $\left\|e^{i2\pi x}\right\|$

47.14. *Compute the following, assuming the interval is* $(0, 1)$ *and the weight function is* $w(x) = x$ *.*

a. $\langle\, x \mid \sin(2\pi x) \,\rangle$

b. $\left\langle\, x^2 \,\middle|\, 9 + i8x \,\right\rangle$

c. $\left\langle\, 9 + i8x \,\middle|\, x^2 \,\right\rangle$

d. $\left\langle\, e^{i2\pi x} \,\middle|\, x \,\right\rangle$

e. $\|x\|$

f. $\|4 + i8x\|$

g. $\|\sin(2\pi x)\|$

h. $\left\|e^{i2\pi x}\right\|$

47.15. *Determine all the values for* β *so that*

$$\left\{ e^{i2\pi x^2} , e^{i2\pi\beta x^2} \right\}$$

is an orthogonal set on $(0, 1)$ *with weight function* $w(x) = x$.

47.16. *Let* L *be a positive value. Verify that each of the following sets of functions is orthogonal on* $(0, L)$ *with respect to the weight function* $w(x) = 1$.

a. $\left\{ \cos\left(\frac{k\pi}{L}x\right) : k = 1, 2, 3, \ldots \right\}$ **b.** $\left\{ e^{i2k\pi x/L} : k = 0, \pm1, \pm2, \pm3, \ldots \right\}$

47.17. *Consider*

$$\{\phi_1, \phi_2, \phi_3, \ldots\} = \left\{ \cos\left(\frac{1\pi}{L}x\right) , \cos\left(\frac{2\pi}{L}x\right) , \cos\left(\frac{3\pi}{L}x\right) , \ldots \right\} \quad ,$$

which, in the exercise above, you showed is an orthogonal set of functions on $(0, L)$ *with respect to the weight function* $w(x) = 1$. *Using this interval, weight function and orthogonal set, do the following:*

a. *Show that, in this case, the generalized Fourier series*

$$G.F.S.[f] = \sum_{k=1}^{\infty} c_k \phi_k(x) \qquad \text{with} \quad c_k = \frac{\langle \phi_k \mid f \rangle}{\|\phi_k\|^2}$$

is given by

$$G.F.S.[f] = \sum_{k=1}^{\infty} c_k \cos\left(\frac{k\pi}{L}x\right) \qquad \text{with} \quad c_k = \frac{2}{L}\int_0^L f(x)\cos\left(\frac{k\pi}{L}x\right) dx \quad .$$

b. *Find the generalized Fourier series for*

i. $f(x) = x$

ii. $f(x) = \begin{cases} 1 & \text{if } 0 < x < L/2 \\ 0 & \text{if } L/2 < x < L \end{cases}$

iii. $f(x) = \begin{cases} x & \text{if } 0 < x < L/2 \\ L - x & \text{if } L/2 < x < L \end{cases}$

c. *Show that this set of cosines is* not *a complete set by showing*

$$f(x) \neq \sum_{k=1}^{\infty} c_k \cos\left(\frac{k\pi}{L}x\right) \qquad \text{with} \quad c_k = \frac{2}{L}\int_0^L f(x)\cos\left(\frac{k\pi}{L}x\right) dx$$

when f *is the constant function* $f(x) = 1$ *on* $(0, L)$.

47.18. *In several exercises above, you considered either the Sturm-Liouville problem*

$$\phi'' + \lambda\phi = 0 \qquad \text{for} \quad 0 < x < L \qquad \text{with} \quad \phi'(0) = 0 \quad \text{and} \quad \phi'(L) = 0$$

or, at least, the above differential equation. You may use the results from those exercises to help answer the ones below:

a. *What is the natural weight function* $w(x)$ *and inner product* $\langle f \mid g \rangle$ *corresponding to this Sturm-Liouville problem?*

b. *We know that*

$$\{\lambda_0,\ \lambda_1,\ \lambda_2,\ \ldots,\ \lambda_k,\ \ldots\} = \left\{0,\ \left(\frac{1\pi}{L}\right)^2,\ \left(\frac{2\pi}{L}\right)^2,\ \ldots,\ \left(\frac{k\pi}{L}\right)^2,\ \ldots\right\}$$

is the complete set of eigenvalues for this Sturm-Liouville problem. Now write out a corresponding orthogonal set of eigenfunctions (without arbitrary constants),

$$\{\phi_0,\ \phi_1,\ \phi_2,\ \ldots,\ \phi_k,\ \ldots\} \quad .$$

c. *Compute the norm of each of your* ϕ_k*'s from your answer to the last part.*

d. *To what does each coefficient in the generalized Fourier series*

$$G.F.S.[f] = \sum_{k=0}^{\infty} c_k\,\phi_k(x) \qquad \text{with} \quad c_k = \frac{\langle \phi_k \mid f \rangle}{\|\phi_k\|^2} \quad ,$$

reduce to using the eigenfunctions and inner products from previous parts of this exercise?

e. *Using the results from the last part, find the generalized Fourier series for the following:*

i. $f(x) = 1$ **ii.** $f(x) = x$

iii. $f(x) = \begin{cases} 1 & \text{if} \quad 0 < x < L/2 \\ 0 & \text{if} \quad L/2 < x < L \end{cases}$

47.19. *In several exercises above, you considered either the Sturm-Liouville problem*

$$x^2\phi'' + x\phi' = -\lambda\phi \qquad \text{for} \quad 1 < x < e^{\pi} \qquad \text{with} \quad \phi(1) = 0 \quad \text{and} \quad \phi(e^{\pi}) = 0$$

or, at least, the above differential equation. You may use the results from those exercises to answer the ones below:

a. *What is the natural weight function* $w(x)$ *and inner product* $\langle f \mid g \rangle$ *corresponding to this Sturm-Liouville problem?*

b. *We know that*

$$\{\lambda_1,\ \lambda_2,\ \lambda_2,\ \ldots,\ \lambda_k,\ \ldots\} = \left\{1,\ 4,\ 9,\ \ldots,\ k^2,\ \ldots\right\}$$

is the complete set of eigenvalues for this Sturm-Liouville problem. Now write out a corresponding orthogonal set of eigenfunctions (without arbitrary constants),

$$\{\phi_1,\ \phi_2,\ \phi_3,\ \ldots,\ \phi_k,\ \ldots\} \quad .$$

c. *Compute the norm of each of your* ϕ_k*'s from your answer to the last part. (A simple substitution may help.)*

d. *To what does each coefficient in the generalized Fourier series*

$$G.F.S.[f] = \sum_{k=0}^{\infty} c_k\,\phi_k(x) \qquad \text{with} \quad c_k = \frac{\langle\, \phi_k \mid f \,\rangle}{\|\phi_k\|^2} \quad ,$$

reduce to using the eigenfunctions and inner products from previous parts of this exercise?

e. *Using the results from the last part, find the generalized Fourier series for the following:*

i. $f(x) = 1$ **ii.** $f(x) = \ln|x|$ **iii.** $f(x) = x^{\alpha}$

47.20. *In several exercises above, you considered either the Sturm-Liouville problem*

$$x^2\phi'' + 3x\phi' = -\lambda\phi \qquad \text{for} \quad 1 < x < e^{\pi} \qquad \text{with} \quad \phi(1) = 0 \quad \text{and} \quad \phi(e^{\pi}) = 0$$

or, at least, the above differential equation. You may use the results from those exercises to answer the ones below:

a. *What is the natural weight function* $w(x)$ *and inner product* $\langle\, f \mid g \,\rangle$ *corresponding to this Sturm-Liouville problem?*

b. *We know that*

$$\{\lambda_1,\, \lambda_2,\, \lambda_2,\, \ldots,\, \lambda_k,\, \ldots\} = \left\{2,\, 5,\, 10,\, \ldots,\, k^2+1,\, \ldots\right\}$$

is the complete set of eigenvalues for this Sturm-Liouville problem. Now write out a corresponding orthogonal set of eigenfunctions (without arbitrary constants),

$$\{\phi_1,\, \phi_2,\, \phi_3,\, \ldots,\, \phi_k,\, \ldots\}$$

c. *Compute the norm of each of your* ϕ_k*'s from your answer to the last part. (A simple substitution may help.)*

d. *To what does each coefficient in the generalized Fourier series*

$$G.F.S.[f] = \sum_{k=0}^{\infty} c_k\,\phi_k(x) \qquad \text{with} \quad c_k = \frac{\langle\, \phi_k \mid f \,\rangle}{\|\phi_k\|^2} \quad ,$$

reduce to using the eigenfunctions and inner products from previous parts of this exercise?

e. *Using the results from the last part, find the generalized Fourier series for the following:*

i. $f(x) = 1$ **ii.** $f(x) = x^{\alpha}$

47.21. *Consider the Sturm-Liouville problem*

$$\left(1 - x^2\right)\phi'' - x\phi' + \lambda\phi = 0 \qquad \text{for} \quad -1 < x < 1$$

with

$$|\phi(-1)| < \infty \qquad \text{and} \qquad |\phi(-1)| < \infty \quad .$$

This is not a regular Sturm-Liouville problem, but "it can be shown" that the results claimed in "mega-theorem" 47.11 hold for this Sturm-Liouville problem. In particular, it can be shown that the Chebyshev polynomials (type I),

$$\phi_0(x) = 1 \quad , \quad \phi_1(x) = x \quad , \quad \phi_2(x) = 2x^2 - 1 \quad ,$$
$$\phi_3(x) = 4x^3 - 3x \quad , \quad \dots \quad ,$$

make up a complete, orthogonal set of eigenfunctions for this Sturm-Liouville problem. Assuming all this, do the following:

a. *Using the above differential equation and formulas for ϕ_0 , ϕ_1 , ϕ_2 and ϕ_3 , find the corresponding eigenvalues λ_0 , λ_1 , λ_2 and λ_3 .*

b. *Rewrite the above differential equation in Sturm-Liouville form.*

c. *Based on the above, what is the appropriate inner product $\langle f \mid g \rangle$ when using the type I Chebyshev polynomials?*

d. *Using the inner product from the previous part, and the above ϕ_k's, compute the following norms[8]:*

i. $\|\phi_0\|$ **ii.** $\|\phi_1\|$ **iii.** $\|\phi_2\|$ **iv.** $\|\phi_3\|$

e. *According to our theory, any reasonable function $f(x)$ on $(-1,1)$ can be approximated by the N^{th} partial sum*

$$\sum_{k=0}^{N} c_k \phi_k(x)$$

where

$$\sum_{k=0}^{\infty} c_k \phi_k(x)$$

is the generalized Fourier series for $f(x)$ using the type I Chebyshev polynomials. For the following, let

$$f(x) = x\sqrt{1 - x^2}$$

and find the above mentioned N^{th} partial sum when

i. $N = 0$ **ii.** $N = 1$ **iii.** $N = 2$ **iv.** $N = 3$

---

[8] the substitution $x = \sin(\theta)$ may be useful

Some Answers to Some of the Exercises

WARNING! Most of the following answers were prepared hastily and late at night. They have not been properly proofread! Errors are likely!

3b. $\|\mathbf{b}^1\| = \sqrt{13}$ and $\|\mathbf{b}^2\| = \sqrt{13}$
3c i. $\mathbf{u} = 2\mathbf{b}^1$
3c ii. $\mathbf{v} = 3\mathbf{b}^1 - 2\mathbf{b}^2$
3c iii. $\mathbf{w} = \frac{3}{13}\mathbf{b}^1 + \frac{2}{13}\mathbf{b}^2$
3d i. $\mathbf{Au} = 6\mathbf{b}^1$
3d ii. $\mathbf{Av} = 9\mathbf{b}^1 + 4\mathbf{b}^2$
3d iii. $\mathbf{Aw} = \frac{9}{13}\mathbf{b}^1 - \frac{4}{13}\mathbf{b}^2$
4b. $\|\mathbf{b}^1\| = \sqrt{6}$, $\|\mathbf{b}^2\| = \sqrt{5}$ and $\|\mathbf{b}^3\| = \sqrt{30}$
4c i. $\mathbf{u} = 2\mathbf{b}^1 + 3\mathbf{b}^2 - 4\mathbf{b}^3$
4c ii. $\mathbf{v} = \frac{4}{3}\mathbf{b}^1 - \frac{1}{3}\mathbf{b}^3$
4c iii. $\mathbf{w} = \frac{1}{3}\mathbf{b}^1 - \frac{1}{5}\mathbf{b}^2 + \frac{1}{15}\mathbf{b}^3$
4d i. $\mathbf{Au} = 4\mathbf{b}^1 + 12\mathbf{b}^2 - 24\mathbf{b}^3$
4d ii. $\mathbf{Av} = \frac{8}{3}\mathbf{b}^1 - 2\mathbf{b}^3$
4d iii. $\mathbf{Aw} = \frac{2}{3}\mathbf{b}^1 - \frac{4}{5}\mathbf{b}^2 + \frac{2}{5}\mathbf{b}^3$
5b. $\|\mathbf{b}^1\| = \sqrt{3}$, $\|\mathbf{b}^2\| = \sqrt{6}$ and $\|\mathbf{b}^3\| = \sqrt{2}$
5c i. $\mathbf{u} = 2\mathbf{b}^1 - 3\mathbf{b}^2 + 4\mathbf{b}^3$
5c ii. $\mathbf{v} = \mathbf{b}^1 + 2\mathbf{b}^2 + 3\mathbf{b}^3$
5c iii. $\mathbf{w} = \frac{1}{3}\mathbf{b}^1 + \frac{1}{6}\mathbf{b}^2 - \frac{1}{2}\mathbf{b}^3$
5d i. $\mathbf{Au} = -2\mathbf{b}^1 + 4\mathbf{b}^3$
5d ii. $\mathbf{Av} = -\mathbf{b}^1 + 3\mathbf{b}^3$
5d iii. $\mathbf{Aw} = -\frac{1}{3}\mathbf{b}^1 - \frac{1}{2}\mathbf{b}^3$
8a. If $\lambda < -4$, $\phi_\lambda(x) = ae^{\nu x} + be^{-\nu x}$ with $\nu = \sqrt{-(4+\lambda)}$; If $\lambda = -4$, $\phi_4(x) = ax + b$; If $\lambda > -4$, $\phi_\lambda(x) = a\cos(\nu x) + b\sin(\nu x)$ with $\nu = \sqrt{4+\lambda}$
8b. If $\lambda < 1$, $\phi_\lambda(x) = ae^{(-1+\nu)x} + be^{(-1-\nu)x}$ with $\nu = \sqrt{1-\lambda}$; If $\lambda = 1$, $\phi_1(x) = ae^{-x} + bxe^{-x}$; If $\lambda > 1$, $\phi_\lambda(x) = ae^{-x}\cos(\nu x) + be^{-x}\sin(\nu x)$ with $\nu = \sqrt{\lambda - 1}$
8c. If $\lambda < 0$, $\phi_\lambda(x) = ax^\nu + bx^{-\nu}$ with $\nu = \sqrt{-\lambda}$; If $\lambda = 0$, $\phi_0(x) = a\ln|x| + b$; If $\lambda > 0$, $\phi_\lambda(x) = a\cos(\nu\ln|x|) + b\sin(\nu\ln|x|)$ with $\nu = \sqrt{\lambda}$
8d. If $\lambda < 1$, $\phi_\lambda(x) = ax^{\nu-1} + bx^{-\nu-1}$ with $\nu = \sqrt{1-\lambda}$; If $\lambda = 1$, $\phi_1(x) = ax^{-1}\ln|x| + bx^{-1}$; If $\lambda > 1$, $\phi_\lambda(x) = ax^{-1}\cos(\nu\ln|x|) + bx^{-1}\sin(\nu\ln|x|)$ with $\nu = \sqrt{\lambda - 1}$
9a. $\phi'' + 4\phi = -\lambda\phi$
9b. $\frac{d}{dx}\left[e^{2x}\frac{d\phi}{dx}\right] = -\lambda e^{2x}\phi$
9c. $\frac{d}{dx}\left[x\frac{d\phi}{dx}\right] = -\lambda\frac{1}{x}\phi$
9d. $\frac{d}{dx}\left[x^3\frac{d\phi}{dx}\right] = -\lambda x\phi$
9e. $\frac{d}{dx}\left[e^{-x^2}\frac{d\phi}{dx}\right] = -\lambda e^{-x^2}\phi$
9f. $\frac{d}{dx}\left[\sqrt{1-x^2}\frac{d\phi}{dx}\right] = -\lambda\frac{1}{\sqrt{1-x^2}}\phi$

9g. $\frac{d}{dx}\left[xe^{-x}\frac{d\phi}{dx}\right] = -\lambda e^{-x}\phi$

10a. $(\lambda_0, \phi_0(x)) = (0, c_0)$ and $(\lambda_k, \phi_k(x)) = \left(\left(\frac{k\pi}{L}\right)^2 \, , \; c_k \cos\left(\frac{k\pi}{L}x\right)\right)$ for $k = 1, 2, 3, \dots$

10b. $(\lambda_k, \phi_k(x)) = \left(\left(\frac{(2k+1)\pi}{2L}\right)^2 \, , \; c_k \sin\left(\frac{(2k+1)\pi}{2L}x\right)\right)$ for $k = 1, 2, 3, \dots$

11a. $(\lambda_0, \phi_0(x)) = (0, c_0)$ and $(\lambda_k, \phi_k(x)) = \left(k^2 \, , \; c_{k,1}\cos(kx) + c_{k,2}\sin(kx)\right)$ for $k = 1, 2, 3, \dots$

11b. $(\lambda_k, \phi_k(x)) = \left(k^2 \, , \; c_k \sin(k \ln|x|)\right)$ for $k = 1, 2, 3, \dots$

11c. $(\lambda_k, \phi_k(x)) = \left(k^2 + 1 \, , \; c_k x^{-1} \sin(k \ln|x|)\right)$ for $k = 1, 2, 3, \dots$

12c. $p(b) = p(a)$

13a. $-\frac{3}{2\pi}$

13b. $81 + 162i$

13c. $81 - 162i$

13d. $\frac{3}{2\pi}i$

13e. 3

13f. $\sqrt{819}$

13g. $\sqrt{\frac{3}{2}}$

13h. $\sqrt{3}$

14a. $-\frac{1}{2\pi}$

14b. $\frac{9}{4} + i\frac{8}{5}$

14c. $\frac{9}{4} - i\frac{8}{5}$

14d. $\frac{1}{2\pi^2} + \frac{1}{2\pi}i$

14e. $\frac{1}{2}$

14f. $2\sqrt{6}$

14g. $\frac{1}{2}\sqrt{1 + \frac{1}{2\pi}}$

14h. $\frac{1}{\sqrt{2}}$

15. β can be any integer except 1 .

17b i. $\sum_{k=1}^{\infty} \frac{2L}{(k\pi)^2}\left[(-1)^k - 1\right]\cos\left(\frac{k\pi}{L}x\right)$

17b ii. $\sum_{k=1}^{\infty} \frac{2}{k\pi}\sin\left(\frac{k\pi}{2}\right)\cos\left(\frac{k\pi}{L}x\right)$

18a. $w(x) = 1$, $\langle\, f \mid g \,\rangle = \int_0^L (f(x))^* g(x)\, dx$

18b. $\left\{1, \; \cos\left(\frac{1\pi}{L}x\right), \cos\left(\frac{2\pi}{L}x\right), \; \dots, \; \cos\left(\frac{k\pi}{L}x\right), \; \dots\right\}$

18c. $\|1\| = L$, $\left\|\cos\left(\frac{k\pi}{L}x\right)\right\| = \frac{L}{2}$

18d. $c_0 = \frac{1}{L}\int_0^L f(x)\, dx$, $c_k = \frac{2}{L}\int_0^L f(x)\cos\left(\frac{k\pi}{L}x\right)\, dx$ for $k > 0$

18e i. 1

18e ii. $\frac{L}{2} + \sum_{k=1}^{\infty} \frac{2L}{(k\pi)^2}\left[(-1)^k - 1\right]\cos\left(\frac{k\pi}{L}x\right)$

18e iii. $\frac{1}{2} + \sum_{k=1}^{\infty} \frac{2}{k\pi} \sin\left(\frac{k\pi}{2}\right) \cos\left(\frac{k\pi}{L}x\right)$

19a. $w(x) = \frac{1}{x}$, $\langle\, f \mid g \,\rangle = \int_1^{e^\pi} (f(x))^* g(x) \frac{1}{x}\, dx$

19b. $\{\sin(1 \ln|x|)\,,\ \sin(2 \ln|x|)\,,\ \sin(3 \ln|x|)\,,\ \ldots,\ \sin(k \ln|x|)\,,\ \ldots\}$

19c. $\|\sin(k \ln|x|)\| = \frac{\pi}{2}$

19d. $c_k = \frac{2}{\pi} \int_1^{e^\pi} f(x) \sin(k \ln|x|) \frac{1}{x}\, dx$

19e i. $\sum_{k=1}^{\infty} \frac{2}{k\pi} \left[1 - (-1)^k\right] \sin(k \ln|x|)$

19e ii. $\sum_{k=1}^{\infty} (-1)^{k+1} \frac{2}{k} \sin(k \ln|x|)$

19e iii. $\sum_{k=1}^{\infty} \frac{2k}{\pi(\alpha^2 + k^2)} \left[1 - (-1)^k e^{\alpha\pi}\right] \sin(k \ln|x|)$

20a. $w(x) = x$, $\langle\, f \mid g \,\rangle = \int_1^{e^\pi} (f(x))^* g(x) x\, dx$

20b. $\left\{x^{-1} \sin(1 \ln|x|)\,,\ x^{-1} \sin(2 \ln|x|)\,,\ x^{-1} \sin(3 \ln|x|)\,,\ \ldots,\ x^{-1} \sin(k \ln|x|)\,,\ \ldots\right\}$

20c. $\left\|x^{-1} \sin(k \ln|x|)\right\| = \sqrt{\frac{\pi}{2}}$

20d. $c_k = \frac{2}{\pi} \int_1^{e^\pi} f(x) \sin(k \ln|x|)\ dx$

20e i. $\sum_{k=1}^{\infty} \frac{2k}{\pi\left(1 + k^2\right)} \left[1 - (-1)^k e^{\pi}\right] \frac{\sin(k \ln|x|)}{x}$

20e ii. $\sum_{k=1}^{\infty} \frac{2k}{\pi\left((\alpha+1)^2 + k^2\right)} \left[1 - (-1)^k e^{(\alpha+1)\pi}\right] \frac{\sin(k \ln|x|)}{x}$

21a. $\lambda_0 = 0\,,\ \lambda_1 = 1\,,\ \lambda_2 = 4\,,\ \lambda_3 = 9$

21b. $\frac{d}{dx}\left[\sqrt{1 - x^2} \frac{d\phi}{dx}\right] = -\lambda \frac{1}{\sqrt{1 - x^2}} \phi$

21c. $\langle\, f \mid g \,\rangle = \int_{-1}^{1} (f(x))^* g(x) \left(1 - x^2\right)^{-1/2}\ dx$

21d i. $\sqrt{\pi}$

21d ii. $\sqrt{\frac{\pi}{2}}$

21d iii. $\sqrt{\frac{3\pi}{8}}$

21d iv. $\sqrt{\frac{\pi}{2}}$

21e i. 0

21e ii. $\frac{4}{3\pi} x$

21e iii. $\frac{4}{3\pi} x$

21e iv. $\frac{4}{3\pi} x - \frac{4}{5\pi} x^3$