19

Sturm-Liouville Theory

To make use of the separable solutions to a partial differential equation problem, we need to know some of the properties of the solution sets to the boundary value problem arising in the separation of variables procedure. It is the Sturm-Liouville theory that tells us those properties. This theory basically concerns the solutions to a boundary-value problem involving an ordinary differential equation (but can — and will — be extended to cover solutions to boundary-value problems involving partial differential equations). Much of it is very similar to the theory we developed for self-adjoint (i.e., Hermitian) operators last term. Indeed, much of the Sturm-Liouville theory is essentially the theory we developed for Hermitian/self-adjoint operators in chapter 7, with our big "Sturm-Liouville theorems" being reminiscent of the "Big Theorem on Hermitian Operators", theorem 7.8 on page 7–19.

By the way, you should be aware that the first boundary-value problem for which a "Sturm-Liouville theory" was developed was the one leading to the classical Fourier series. The theory that will be discussed here is a good deal more general; so we won't just be developing Fourier series, we will be developing "generalized" Fourier series.

19.1 Preliminaries to the Sturm-Liouville Theory

Before developing the theory for the eigenproblems (the "Sturm-Liouville theory), we need to review some linear algebraic concepts discussed last term, and extend them a little to take into account the fact that our basis will contain infinitely many elements.

Preliminaries to the Preliminaries

?► ***Exercise 19.1:*** *Quickly review chapter 3 on general vector spaces, especially the "basics" in §3.2; the material on inner products, norms and orthogonality in §3.3, and the material on vector components when the basis is orthogonal in §3.4.*

The Weight Function and Interval of Interest

Throughout this section we will assume that we have a finite interval (a, b) and a positive function $w(x)$ on this interval.[1] The w will be called the *weight function*, and we will occasionally refer to (a, b) as the "interval of interest". Both of these will be determined by the eigenproblem: the interval in the obvious way, and the weight function by means discussed later. The underlying vector space for our discussions will be some set of 'suitably integrable functions of interest' on (a, b). We won't define this set precisely, but will assume it at least contains all the bounded, continuous functions on the interval which satisfy "appropriate boundary conditions". These "appropriate" boundary conditions will be similar to the "suitable" boundary conditions discussed in the last chapter, and will be described more completely in a few pages. We will also assume, unless otherwise stated, that all of the functions mentioned are from this vector space.

The Inner Product and Norm

Let f and g be two suitably integrable functions on (a, b). Their inner product $\langle f \mid g \rangle$ is defined by

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

(Remember f^* is the complex conjugate of f, and w is that 'weight function' mentioned above.)[2] In most of the cases of interest to us, f and g are real valued, and this definition reduces to

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

Remember from last term that an 'inner product' is a generalization of the idea of a vector 'dot product'. That the above does satisfy those properties we expect of an inner product is the claim of the next theorem.

Theorem 19.1 (properties of inner products)
Assume f, g and h are 'suitably integrable' functions, and let α and β be any two constants. Then the following all hold using the above inner product:

1. *(linearity)* $\langle f \mid \alpha g + \beta h \rangle = \alpha \langle f \mid g \rangle + \beta \langle f \mid h \rangle$.
2. *(positive-definiteness)* $\langle f \mid f \rangle \geq 0$ *with* $\langle f \mid f \rangle = 0$ *if and only if* $f = 0$ *on* (a, b).
3. *(conjugate symmetry)* $\langle f \mid g \rangle = \langle g \mid f \rangle^*$.

?►*Exercise 19.2:* *Verify each of the claims in the above theorem.*

---

[1] Actually, we only need $w(x)$ to be positive "almost everywhere" on (a, b). We can, for example, allow $w(x)$ to be zero at a finite number of points in the interval.

[2] Be warned that other authors may, instead, denote this inner product by $\langle f, g \rangle$ or (f, g), and may, instead, use

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

?►*Exercise 19.3:* *Using the properties claimed in the theorem, show that*

$$\langle\, \alpha f + \beta g \mid h \,\rangle \;=\; \alpha^* \langle\, f \mid h \,\rangle \;+\; \beta^* \langle\, g \mid h \,\rangle \quad .$$

With the inner product defined, we now have corresponding notions of "norm" and "orthogonality".

In particular, the *norm* of a function f (with respect to the given inner product, or with respect to the weight function) is given by

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

Equivalently,

$$\|f\| \;=\; \left[\int_a^b f^*(x) f(x) w(x)\, dx\right]^{1/2} \;=\; \left[\int_a^b |f(x)|^2\, w(x)\, dx\right]^{1/2} \quad .$$

Do note that, in a loose sense,

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

As with any inner product, we say that two functions f and g are *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.

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

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

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 eigenproblem, we will even refer to $G.F.S.[f]$ as the *eigenfunction expansion* of f.

Completeness

An orthogonal set of functions $\{\phi_1, \phi_2, \phi_3, \ldots\}$ is said to be *complete* if and only if, for each function f 'of interest' (i.e., in our underlying vector space of functions),

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

That is

$$\left\| f(x) - \sum_k c_k \phi_k(x) \right\| = 0 \qquad \text{where} \quad c_k = \frac{\langle \phi_k \mid f \rangle}{\|\phi_k\|^2} \quad .$$

Actually, since the summations may be infinite, we should write this as

$$\lim_{N\to\infty} \left\| f(x) - \sum_k^N c_k \phi_k(x) \right\| = 0 \quad .$$

Note that this can be written as

$$\lim_{N\to\infty} \int_a^b \left| f(x) - \sum_k^N c_k \phi_k(x) \right|^2 w(x)\, dx = 0 \quad .$$

[3] Compare the formula for the generalized Fourier coefficients with formula (3.7) on page 3–16 for the components of a vector with respect to any orthogonal basis. They are virtually the same!

The above integral is sometimes known as the "(weighted) mean square error in using $\sum_k^N c_k\,\phi_k(x)$ for $f(x)$ on the interval (a,b)."

In practice, all of the above usually means that the infinite series $\sum_k c_k\phi_k(x)$ converges pointwise 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 all functions 'of interest'.

19.2 The Sturm-Liouville Theory

The Problem and Basic Terminology

In this section we are going to look at boundary-value problems involving an equation of the form

$$\mathcal{L}[\phi] = -\lambda w \phi \qquad \text{on some interval } (a,b)$$

where $\mathcal{L}$ is a linear, second-order, ordinary differential operator, $w = w(x)$ is some known function, and λ and ϕ are the unknowns to be determined. λ is a constant (the *eigenvalue*) and $\phi = \phi(x)$ is a function (the corresponding *eigenfunction*) that will also be required to satisfy "appropriate" boundary conditions (which we will discuss in a few pages).[4]

Any similarity between these problems and the eigenvalue/eigenvector problems of the first term should be noted. As already mentioned (perhaps overly often) "Sturm-Liouville theory" is basically the function version of the "Hermitian operator theory" we discussed then. This will require that we restrict our choices of $\mathcal{L}$, $w(x)$ and the boundary conditions somewhat.

The function $w(x)$ will end up being the weight function for the inner product discussed earlier, so it will have to be positive (almost everywhere) on (a,b).

The operator $\mathcal{L}$ will have to be *self-adjoint* (equivalently, *Hermitian* or *Sturm-Liouvillian*). For no apparent reason, we will define these terms to mean that $\mathcal{L}[\phi]$ can be written as

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

where p and q are known *real-valued* functions on whatever interval is of interest to us at the time. Just why we say that such an operator is "self-adjoint" will be explained later. Do note that, by insisting p and q be real-valued, we automatically have

$$\mathcal{L}[\phi]^* = \mathcal{L}\left[\phi^*\right] \quad .$$

This will be important.

[4] In this section, the variable will be denoted by x. In practice, we will also use the theory developed on functions of y, r, θ, ...

!►*Example 19.1:* *Our favorite second-order differential operator,*

$$\mathcal{L} = \frac{d^2}{dx^2}$$

is self-adjoint, with $p(x) = 1$ *and* $q(x) = 0$ *.*

A less trivial example would be

$$\mathcal{L} = x^2 \frac{d^2}{dx^2} + 2x\frac{d}{dx} + 3x \quad .$$

Given any sufficiently differentiable $\phi(x)$ *, you can easily verify that*

$$\begin{aligned}\mathcal{L}[\phi] &= x^2 \frac{d^2\phi}{dx^2} + 2x\frac{d\phi}{dx} + 3x\phi \\ &= \frac{d}{dx}\left[x^2 \frac{d\phi}{dx}\right] + 3x\phi \quad .\end{aligned}$$

So this operator is self-adjoint, with $p(x) = x^2$ *and* $q(x) = 3x$ *.*

Along the same lines, any homogeneous, second-order, linear ordinary differential equation will be said to be in *self-adjoint form* if it is written as

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

if no eigenvalue is involved, or as

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

if an eigenvalue λ is involved (in which case, we may also say the equation is in *Sturm-Liouville form*). As before p , q and w are known real-valued functions on the interval of interest (we won't insist w be positive for this definition). Do note that such equations can be written, respectively, as

$$\mathcal{L}[\phi] = 0 \qquad \text{and} \qquad \mathcal{L}[\phi] = -\lambda w(x)\phi$$

where $\mathcal{L}$ is a self-adjoint differential operator.

Fortunately, just about any homogneous, second-order, linear ordinary differential equation (with real-valued coefficients) can be written in self-adjoint 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

$$A(x)\frac{d^2\phi}{dx^2} + B(x)\frac{d\phi}{dx} + C(x)\phi = 0$$

where A and B are real-valued, and A is never zero on our interval of interest. To illustrate the procedure, we'll use the equation

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

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

Here is what you do:

1. Divide through by $A(x)$

Doing that with our example yields

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

2. Compute the "integrating factor"

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

(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),

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

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

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

$$p\frac{d^2\phi}{dx^2} + p\frac{B}{A}\frac{d\phi}{dx} + p\frac{C}{A}\phi = 0 \quad ,$$

(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,

$$\frac{d}{dx}\left[p\frac{d\phi}{dx}\right] + p\frac{C}{A}\phi = 0 \quad .$$

In our case, we have

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

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

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

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

4. If there is no eigenvalue λ involved, you are done — the equation is in the desired form. Otherwise, finish getting the equation into desired form by moving the term with λ to the right side of the equation. You may also want to note just which formulas are $p(x)$, $q(x)$, and, if it is there, $w(x)$.

   We have a 'λ term'. Moving it to the other side yields

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

   It is now in the desired form, with

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

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

Green's Formula, Boundary Conditions and Self-Adjointness

In the following exercise, you will verify a basic identity involving self-adjoint differential operators. This identity is the starting point for the Green's formula which, in turn, leads to many of our most important results.

?►Exercise 19.4 (the preliminary Green's formula): *Let (a, b) be some interval; let p and q be any suitably smooth and integrable functions on (a, b), and let $\mathcal{L}$ be the operator given by*

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

Using integration by parts, show that

$$\int_a^b f\,\mathcal{L}[g]\,dx = p(x)f(x)\frac{dg}{dx}\bigg|_a^b - \int_a^b p\frac{df}{dx}\frac{dg}{dx}\,dx + \int_a^b qfg\,dx \tag{19.1}$$

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

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

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

Assuming p and q are real-valued functions (so $\mathcal{L}[u]^* = \mathcal{L}[u^*]$) and using equation (19.1), we can easily write out the difference between these inner products,

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

Notice that the integrals on the right all cancel out, leaving us with:

Theorem 19.2 (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

$$\int_a^b u^* \mathcal{L}[v]\, dx - \int_a^b \mathcal{L}[u]^* v\, dx = p(x)\left[u^*(x)\frac{dv}{dx} - v(x)\frac{du^*}{dx}\right]\Bigg|_a^b \tag{19.2}$$

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

Equation (19.2) 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 boundary conditions at $x = a$ and $x = b$ as being *Sturm-Liouville appropriate* if and only if both of the following hold:

1. The boundary conditions are "suitable" as defined on page 18–12. That is, if two functions satisfy the given boundary conditions, then so does every linear combination of them.

2. Whenever both u and v satisfy these boundary conditions, then

$$p(x)\left[u^*(x)\frac{dv}{dx} - v(x)\frac{du^*}{dx}\right]\Bigg|_a^b = 0 \quad . \tag{19.3}$$

For brevity, we may say "appropriate" when we mean "Sturm-Liouville appropriate". Let's make two quick observations:

1. Because the function p, which comes from the differential equation, appears in equation (19.3) defining our notion of Sturm-Liouville appropriateness, 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, even if the boundary conditions are "suitable" as defined on page 18–12.

2. It is not hard to verify that the set of all "sufficiently differentiable" functions satisfying a given set of Sturm-Liouville appropriate boundary conditions is a vector space of functions.

!▶*Example 19.2:* *Consider the boundary conditions*

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

If u and v satisfy these conditions; that is,

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

and

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

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

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

So

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

are "Sturm-Liouville appropriate" boundary conditions as far as we are concerned.

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

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

from which we immediately get

Corollary 19.3
If $\mathcal{L}$ *is a self-adjoint operator as in theorem 19.2, and if* u *and* v *satisfy the corresponding Sturm-Liouville appropriate boundary conditions, then*

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

In terms of the inner product

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

the corollary tells us that

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

Now recall: by our definitions from last term, the adjoint of $\mathcal{L}$ is the operator $\mathcal{L}^\dagger$ such that

$$\langle u \mid \mathcal{L}[v] \rangle = \langle \mathcal{L}^\dagger[u] \mid v \rangle \quad .$$

So, if $\mathcal{L}$ is as we defined *and* we insist on appropriate boundary conditions, then we really do have $\mathcal{L}^\dagger = \mathcal{L}$, which means $\mathcal{L}$ is self-adjoint in the linear algebraic sense (i.e., as defined in chapter 4).

Sturm-Liouville Problems, Defined

Finally, we can state with reasonable precision the sort of problems Sturm-Liouville theory is concerned with:

A *Sturm-Liouville problem* consists of the following:

1. A differential equation of the form

$$\frac{d}{dx}\left[p(x)\frac{d\phi}{dx}\right] + q(x)\phi = -\lambda w(x)\phi \qquad \text{for} \quad a < x < b \tag{19.4}$$

   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. (The sign of p and w turn out to be relevant for many results. We will usually assume they are positive functions on (a,b). This is what invariably happens in real applications.)

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

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 function (called an eigenfunction) which, together, satisfy the given Sturm-Liouville problem.

Whenever we have a Sturm-Liouville problem, we will automatically let $\mathcal{L}$ be the self-adjoint operator defined by the left side of the differential equation (19.4),

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

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 regular/homogeneous 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.

Most, but not all, of the Sturm-Liouville problems that we will generate in solving partial differential equation problems are "regular".

19.3 The Main Results

Throughout this section, we assume we have some given Sturm-Liouville problem

$$\underbrace{\frac{d}{dx}\left[p(x)\frac{d\phi}{dx}\right] + q(x)\phi}_{\mathcal{L}[\phi]} = -\lambda w(x)\phi \qquad \text{for} \quad a < x < b \tag{19.5a}$$

with Sturm-Liouville appropriate boundary conditions at a and b. (19.5b)

The functions p and w will be assumed to be real valued and continuous on (a, b), w will also be assumed positive on (a, b), and q will be real valued and continuous on the interval.

Our goal here is to develop the results we will need to finish solving partial differential equation problems using separation of variables. Some of the results immediately follow from the theory of Hermitian operators developed last term. Other results will be derived or justified to the extent possible. Unfortunately, it is not practical for us to derive all the results of importance. Ultimately, I will just have to state the big theorem, with enough of it verified that (with luck):

1. You understand the material well enough to understand the theorem.

2. The unverified parts seem reasonable.

Using (Mainly) Hermitian Operator Theory from Last Term*

For our vector space $\mathcal{V}$, let us use (for now) the set of all "sufficiently differentiable" functions on the interval that also satisfy the boundary conditions given for our Sturm-Liouville problem (recall the claim on page 19–9 that this set of functions is a vector space). For the inner product, we will use

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

Now let's try to match our development so far with the discussion in chapter 7 regarding eigen-problems involving self-adjoint/Hermitian operators. Unfortunately there are difficulties in making that match work using our Sturm-Liouville operator $\mathcal{L}$. For one thing, strictly speaking, λ is *not* the eigenvalue for this operator as defined there — the sign is wrong and there is that function w sitting there. Moreover, the "self-adjointness" of $\mathcal{L}$,

$$\langle\, \mathcal{L}[f] \mid g \,\rangle = \langle\, f \mid \mathcal{L}[g] \,\rangle \quad ,$$

was in terms of the wrong inner product, namely,

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

and not

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

* The development in this subsection is rather nonstandard. You probably won't find it in any pde or math/physics text. If you want to see a more standard derivation of some of the results, see the appendix, section 19.5, starting on page 19–21.

To make the match, we can introduce a new operator $\mathcal{H}$ defined by

$$\mathcal{H}[f] = -\frac{1}{w}\mathcal{L}[f] \quad .$$

You can easily verify that $\mathcal{H}$ is a linear operator on $\mathcal{V}$, and that

$$\mathcal{H}[f]^* = \mathcal{H}\big[f^*\big]$$

since w is real valued. Note also that, if (λ, ϕ) is a solution to the Sturm-Liouville problem, then

$$\mathcal{H}[\phi] = -\frac{1}{w}\mathcal{L}[\phi] = -\frac{1}{w}[-\lambda w\phi] = \lambda\phi \quad .$$

So λ is an eigenvalue for $\mathcal{H}$ with corresponding eigenvector/eigenfunction ϕ. Conversely, it should be clear that any eigen-pair (λ, ϕ) for $\mathcal{H}$ is also a solution to the given Sturm-Liouville problem. Consequently, everything we learned about eigenvalues/eigenvectors last term (see chapter 7) applies here. In particular, if λ is a single eigenvalue, then the set of all corresponding eigenfunctions must form a vector space. That is, if ϕ_1 and ϕ_2 are eigenfunctions corresponding to the eigenvalue λ, so is any linear combination

$$c_1\phi_1(x) + c_2\phi_2(x) \quad .$$

Keep in mind, though, that these ϕ's are solutions to a differential equation that can be rewritten as

$$a(x)\frac{d^2\phi}{dx^2} + b(x)\frac{d\phi}{dx} + c(x)\phi = 0$$

where

$$a(x) = p(x) \quad , \quad b(x) = p'(x) \quad \text{and} \quad c(x) = q(x) + \lambda w(x) \quad .$$

This is a second-order, homogeneous linear differential equation, and its general solution can be written as

$$\phi(x) = c_1\psi_1(x) + c_2\psi_2(x)$$

where ψ_1 and ψ_2 are any two independent solutions to the differential equation. Clearly, for ψ_1 we can use one of our eigenfunctions, say, ϕ_1. Whether we can use a second eigenfunction for ψ_2 depends on whether there is a linearly independent pair of eigenfunctions for this one eigenvalue. This means we have exactly two possibilities:

1. There is not an independent pair of eigenvectors corresponding to λ. This means that 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.

2. There is an independent pair of eigenvectors corresponding to λ. This means that 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.

   Remember that, if this is the case, then the second eigenfunction, ϕ_2 can be chosen to be orthogonal to the first eigenfunction (if necessary, use the Gram-Schmidt procedure (see page 3–19) — since the space is two dimensional, there won't be much to the computations!).

What about the self-adjointness of $\mathcal{H}$? Using the inner product with weight function w and corollary 19.3, we have, for each pair of functions u and v in $\mathcal{V}$

$$
\begin{aligned}
\langle\, \mathcal{H}[u] \mid v \,\rangle &= \int_a^b \mathcal{H}[u(x)]^* \, v(x) w(x)\, dx \\
&= \int_a^b \left[-\frac{1}{w(x)} \mathcal{L}\,[u(x)]^* \right] v(x) w(x)\, dx \\
&= -\int_a^b \mathcal{L}\left[u^*(x)\right] v(x)\, dx \\
&= -\int_a^b u^*(x) \mathcal{L}\,[v(x)]\, dx \\
&= \int_a^b u^*(x) \left[-\frac{1}{w(x)} \mathcal{L}\,[v(x)] \right] w(x)\, dx \\
&= \int_a^b u^*(x) \mathcal{H}[v(x)]\, w(x)\, dx \;=\; \langle\, u \mid \mathcal{H}[v] \,\rangle \quad .
\end{aligned}
$$

Thus, $\mathcal{H}$ is a self-adjoint/Hermitian operator on $\mathcal{V}$ using the inner product with weight function w.

What this means is that we can apply our theory from chapter 7 using $\mathcal{H}$ as the Hermitian operator on $\mathcal{V}$. Recall that the big theorem from that section, theorem 7.8 on page 7–19, was

Theorem 19.4 (Big Theorem on Hermitian Operators)
Let $\mathcal{H}$ be a Hermitian (i.e., self-adjoint) operator on a vector space $\mathcal{V}$. Then:

1. *All eigenvalues of $\mathcal{H}$ are real.*
2. *Every pair of eigenvectors corresponding to different eigenvalues are orthogonal.*

Moreover, if $\mathcal{V}$ is finite dimensional, then

1. *If λ is an eigenvalue of algebraic multiplicity m in the characteristic polynomial, then we can find an orthonormal set of exactly m eigenvectors whose linear combinations generate all other eigenvectors corresponding to λ.*
2. *$\mathcal{V}$ has an orthonormal basis consisting of eigenvectors for $\mathcal{H}$.*

From this we immediately get the following two facts concerning the solutions to our Sturm-Liouville problem:

1. Each eigenvalue λ is a real number.[5]
2. If ϕ and ψ are eigenfunctions corresponding to different eigenvalues, then ϕ and ψ are orthogonal; that is,

$$
\langle\, \phi \mid \psi \,\rangle = \int_a^b \phi^*(x) \psi(x) w(x)\, dx = 0 \quad .
$$

[5] It may be worth noting that, since λ is a real number and p, q and w are real-valued functions, you can show that we can choose real-valued eigenfunctions. Moreover, if ϕ is a complex-valued eigenfunction, then its real and imaginary parts are, themselves, eigenfunctions.

Now suppose we have the set of all eigenvalues for our Sturm-Liouville problem. For each simple eigenvalue, choose one corresponding eigenfunction, and for each double eigenvalue, choose an orthogonal pair of corresponding eigenfunctions. The resulting set of eigenfunctions, according to the above, will be an orthogonal set. That is important. At the very least, we can use them as the start for an orthogonal basis for $\mathcal{V}$ (i.e., a "complete orthogonal set of functions for $\mathcal{V}$" — see page 19–4).

Now, how many eigenfunctions will be in our chosen set? Well, recall that, in the one example we've seen (the boundary value problem arising in the separation of variables procedure discussed in the previous chapter), there were infinitely many eigenvalues, namely,

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

So our corresponding set of chosen eigenfunctions will also be infinite. Thus, in this case at least, our vector space $\mathcal{V}$ will be infinite dimensional, and we cannot immediately assume the theorem's claim about there being a basis of eigenfunctions. Still, it seems reasonable to expect that this claim does extend (in some sense) to the case where the vector space is not finite dimensional; so it should seem reasonable to expect that there is a complete orthogonal set of eigenfunctions for our space of functions. In other words, we should expect that we can construct an orthogonal basis of eigenfunctions for our space. And remember (see the discussion regarding generalized Fourier series and completeness starting on page 19–3), this means that we can express any function in $\mathcal{V}$ as a generalized Fourier series of these eigenfunctions.

Other Results

The Rayleigh Quotient and the Eigenvalues

If (λ, ϕ) is an eigen-pair for our Sturm-Liouville problem, then you can show that the eigenvalue λ can be computed from the eigenfunction ϕ via the *Rayleigh quotient*

$$\lambda = \frac{-p\phi^* \frac{d\phi}{dx}\bigg|_a^b + \int_a^b \left[p \left| \frac{d\phi}{dx} \right|^2 - q\,|\phi|^2 \right] dx}{\|\phi\|^2} \quad . \tag{19.6}$$

In fact, you will show it.

?►*Exercise 19.5:* *Derive equation 19.6 assuming λ is an eigenvalue and ϕ is a corresponding eigenfunction for our Sturm-Liouville problem. (Hint: Try using the preliminary Green's formula, equation (19.1) on page 19–8, along with the fact that $\mathcal{L}[\phi] = -\lambda w \phi$.)*

Since one rarely has found eigenfunctions without also having found the corresponding eigenvalues, the Rayleigh quotient is not normally used to compute eigenvalues corresponding to some list of known eigenfunctions. It's value is in finding lower bounds on the possible values of the eigenvalues. Another exercise will illustrate this:

?►*Exercise 19.6:* *Assume that (λ, ϕ) is an eigen-pair for a Sturm-Liouville problem*

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

with Sturm-Liouville appropriate boundary conditions at a and b .

Assume, further, that

$$p(x) > 0 \quad \textit{and} \quad q(x) \leq 0 \qquad \textit{for} \quad a < x < b$$

and that the boundary conditions require that either ϕ *or* ϕ' *be zero at* a *and* b *. Using the Rayleigh quotient, show that*

$$\lambda \geq 0 \quad .$$

Further show that there is a zero eigenvalue if and only if q *is the zero function, and that, in this case, the corresponding eigenfunction must be a constant.*

More generally, using the Rayleigh quotient it can be shown that there is a *smallest* eigenvalue λ_0 for each Sturm-Liouville problem, at least whenever p and w are "reasonable" positive functions on (a, b) . The rest of the eigenvalues are larger.

Unfortunately, verifying the last statement is beyond our ability (unless the Sturm-Liouville problem is sufficiently simple, as in the above exercise). Another fact you will just have to accept without proof is that 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 eigenfunctions for $\mathcal{B}$.

Remember, this set will be an orthogonal set of functions in $\mathcal{V}$, the vector space of sufficiently differentiable functions satisfying the boundary conditions in the Sturm-Liouville problem. (Observe that we may assume each ϕ_k is an eigenfunction corresponding to eigenvalue λ_k only if all the eigenvalues are simple.)

As already noted after restating the big theorem on Hermitian operators (theorem 19.4 on page 19–14), we should also expect $\mathcal{B}$ to be a complete set for $\mathcal{V}$. That is, we expect that, if f is any function in $\mathcal{V}$, then we can express f as the generalized Fourier series of the ϕ_k's ,

$$f(x) = \sum_{k=0}^{\infty} c_k \, \phi_k(x) \qquad \text{where} \quad c_k = \frac{\langle \phi_k \mid f \rangle}{\|\phi_k\|^2} \quad .$$

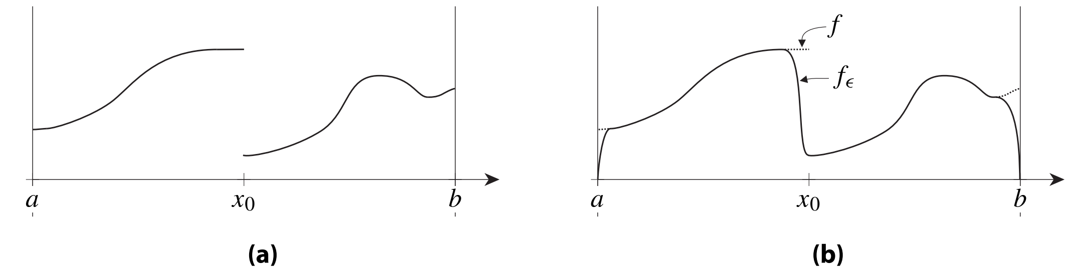

Figure 19.1: The graphs of **(a)** a discontinuous function f on an interval (a, b), and **(b)** a smooth function f_ϵ that vanishes at the interval's endpoint, but differs from f on such small intervals that $\|f - f_\epsilon\| < \epsilon$ for some small value ϵ.

(More precisely,

$$\lim_{N\to\infty} \left\| f - \sum_{k=0}^{N} c_k \phi_k \right\| = 0 \quad .)$$

Rigorously proving this "completeness" is beyond our abilities (in this course), so you will have to trust me that, under reasonable assumptions regarding the functions in the differential equation of our Sturm-Liouville probem, it can be proven that $\mathcal{B}$ is a complete orthogonal set for $\mathcal{V}$.[6]

Now consider just about any other function f whose graph you can sketch. Maybe it has jumps. Maybe it doesn't satisfy the given boundary conditions. For simplicity, let's pretend our boundary conditions are that the functions vanish at $x = a$ and $x = b$. But suppose, instead, that $f(a)$ and $f(b)$ are two other finite numbers, and that f is, say, twice-differentiable except for a jump discontinuity at one point x_0 (as illustrated in figure 19.1a). If you think about it, given any $\epsilon > 0$, you can (as illustrated in figure 19.1b) construct a corresponding function f_ϵ in $\mathcal{V}$ (i.e., f_ϵ is sufficiently differentiable (and continuous) and satisfies the desired boundary conditions) such that

$$\|f - f_\epsilon\| < \epsilon \quad .$$

So f_ϵ closely approximates f. This means that the generalized Fourier series for f_ϵ also closely approximates f. Using this, another minimization result concerning the generalized Fourier series for f, and letting $\epsilon \to 0$, "you can show" that the generalized Fourier series for f converges (in norm) to f, even though f is not in $\mathcal{V}$. That is, whether or not f is in $\mathcal{V}$,

$$\lim_{N\to\infty} \left\| f - \sum_{k=0}^{N} c_k \phi_k \right\| = 0 \qquad \text{where} \quad c_k = \frac{\langle \phi_k \mid f \rangle}{\|\phi_k\|^2} \quad .$$

Thus, "for all practical purposes"

$$f(x) = \sum_{k=0}^{\infty} c_k \phi_k(x) \qquad \text{where} \quad c_k = \frac{\langle \phi_k \mid f \rangle}{\|\phi_k\|^2}$$

(at least for all x not being a, b, points of discontinuity, etc.). Using advanced ideas from real analysis, you can actually show this holds for the set of all functions having finite norm (which, if w is at least bounded on (a, b), includes every piecewise smooth function f on (a, b)).

[6] One approach would be to consider the problem of minimizing the Rayleigh quotient over the vector subspace $\mathcal{U}$ of $\mathcal{V}$ which is orthogonal to all the eigenfunctions. You can then (I think) show that, if $\mathcal{U}$ is not empty, this minimization problem has a solution (λ, ϕ) and that this solution must also satisfy the Sturm-Liouville problem. Hence ϕ is an eigenfunction in $\mathcal{U}$, contrary to the definition of $\mathcal{U}$. So $\mathcal{U}$ must be empty.

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 do eigenfunctions 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 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.

19.4 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 19–11).

Theorem 19.5 (Mega-Theorem on Regular Sturm-Liouville problems)
Consider a regular Sturm-Liouville problem[7] *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 regular/homogeneous 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*

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

[7] 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 $[a, b]$.

with a smallest eigenvalue (usually denoted by λ_0 or λ_1) and no largest eigenvalue. In fact,

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

3. *All the eigenvalues are simple.*

4. *If, for each eigenvalue λ_k , we choose a corresponding (nontrivial) eigenfunction ϕ_k , then the set*

$$\{ \phi_0, \phi_1, \phi_2, \phi_3, \ldots \}$$

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

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

on the set of all piecewise smooth functions on (a, b) . Thus, if f is any such function, then

$$f(x) = \sum_{k=0}^{\infty} c_k\,\phi_k(x) \qquad \text{where} \quad c_k = \frac{\langle \phi_k \mid f \rangle}{\|\phi_k\|^2} \quad .$$

5. *The eigenfunction(s) corresponding to the smallest eigenvalue is never zero on (a, b) . Moreover, for each k , ϕ_{k+1} has exactly one more zero in (a, b) than does ϕ_k .*

6. *Each eigenvalue λ is related to any corresponding eigenfunction ϕ via the Rayleigh quotient*

$$\lambda = \frac{-p\phi^*\dfrac{d\phi}{dx}\bigg|_a^b + \displaystyle\int_a^b \left[p\left|\frac{d\phi}{dx}\right|^2 - q\,|\phi|^2 \right] dx}{\|\phi\|^2} \quad .$$

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.)

Illustrating the Mega-Theorem

In solving our pde problem in the previous chapter, we obtained the eigen-problem

$$\frac{d^2\phi}{dx^2} = -\lambda\phi \qquad \text{for} \quad 0 < x < L$$

with

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

Observe that this is a regular Sturm-Liouville problem with

$$p \equiv 1 \quad , \quad q \equiv 0 \quad \text{and} \quad w \equiv 1 \quad ,$$

and recall that the solutions to this problem were found to be

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

?►***Exercise 19.7:*** *Use the above Sturm-Liouville problem and its solution set to illustrate all the claims made in theorem 19.5 (except, possibly, the claim of the completeness of the set of eigenfunctions).*

19.5 APPENDIX
(Re)Deriving Some Results Already Derived

Since some of the results concerning the values of the eigenvectors and the orthogonality of eigenfunctions were derived in a "nonstandard" manner and using stuff from last term which you may have forgotten, let us re-derive those results here.[8] These derivations correspond more closely to those commonly found in texts on partial differential equations and mathematical physics.

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 the corollary to Green's formula, 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

$$\begin{aligned} \mathcal{L}[\phi_2] &= -\lambda_2 w \phi_2 \\ \mathcal{L}[\phi_1] &= -\lambda_1 w \phi_1 \end{aligned}$$

and thus,

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

(remember w is a positive function). So,

$$\begin{aligned} && \int_a^b {\phi_1}^* \mathcal{L}[\phi_2]\, dx &= \int_a^b \mathcal{L}[\phi_1]^* \phi_2\, dx \\ \Longrightarrow && \int_a^b {\phi_1}^*(-\lambda_2 w \phi_2)\, dx &= \int_a^b (-{\lambda_1}^* w {\phi_1}^*)\phi_2\, dx \\ \Longrightarrow && -\lambda_2 \int_a^b {\phi_1}^* \phi_2 w\, dx &= -{\lambda_1}^* \int_a^b {\phi_1}^* \phi_2 w\, dx \quad . \end{aligned}$$

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

$$\lambda_2 \;=\; {\lambda_1}^* \qquad \text{or} \qquad \int_a^b {\phi_1}^* \phi_2 w\, dx \;=\; 0 \quad . \tag{19.7}$$

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,

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

[8] Besides, I already had these notes written.

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 (19.7) 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:*** *Eigenfunctions corresponding to different eigenvalues are orthogonal with respect to the inner product with weight function* $w(x)$ *.*