6

Sturm-Liouville Eigenvalue Problems

6.1 Introduction

In the last chapters we have explored the solution of boundary value problems that led to trigonometric eigenfunctions. Such functions can be used to represent functions in Fourier series expansions. We would like to generalize some of those techniques in order to solve other boundary value problems. A class of problems to which our previous examples belong and which have eigenfunctions with similar properties are the Sturm-Liouville Eigenvalue Problems. These problems involve self-adjoint (differential) operators which play an important role in the spectral theory of linear operators and the existence of the eigenfunctions we described in Section 4.3.2. These ideas will be introduced in this chapter.

In physics many problems arise in the form of boundary value problems involving second order ordinary differential equations. For example, we might want to solve the equation

$$a_2(x)y'' + a_1(x)y' + a_0(x)y = f(x) \tag{6.1}$$

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

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

Then, Equation (6.1) takes the form

$$Ly = f.$$

As we saw in the general boundary value problem (4.20) in Section 4.3.2, we can solve some equations using eigenvalue expansions. Namely, we seek solutions to the eigenvalue problem

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

with homogeneous boundary conditions and then seek a solution as an expansion of the eigenfunctions. Formally, we let

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

However, we are not guaranteed a nice set of eigenfunctions. We need an appropriate set to form a basis in the function space. Also, it would be nice to have orthogonality so that we can easily solve for the expansion coefficients as was done in Section 4.3.2. [Otherwise, we would have to solve a infinite coupled system of algebraic equations instead of an uncoupled and diagonal system.]

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

We define the *Sturm-Liouville operator* as

$$\mathcal{L} = \frac{d}{dx} p(x) \frac{d}{dx} + q(x). \tag{6.2}$$

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

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

or

$$\frac{d}{dx}\left(p(x)\frac{du}{dx}\right) + q(x)u + \lambda\sigma(x)u = 0, \tag{6.3}$$

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

We also need to impose the set of homogeneous boundary conditions

$$\begin{aligned} \alpha_1 u(a) + \beta_1 u'(a) &= 0, \\ \alpha_2 u(b) + \beta_2 u'(b) &= 0. \end{aligned} \tag{6.4}$$

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

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

As previously mentioned, equations of the form (6.1) occur often. We now show that Equation (6.1) can be turned into a *differential equation of Sturm-Liouville form:*

$$\frac{d}{dx}\left(p(x)\frac{dy}{dx}\right) + q(x)y = F(x). \tag{6.5}$$

Another way to phrase this is provided in the theorem:

Theorem 6.1. *Any second order linear operator can be put into the form of the Sturm-Liouville operator (6.2).*

The proof of this is straight forward, as we shall soon show. Consider the equation (6.1). If $a_1(x) = a_2'(x)$, then we can write the equation in the form

$$\begin{aligned} f(x) &= a_2(x)y'' + a_1(x)y' + a_0(x)y \\ &= (a_2(x)y')' + a_0(x)y. \end{aligned} \tag{6.6}$$

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

However, consider the differential equation

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

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

In the Sturm Liouville operator the derivative terms are gathered together into one perfect derivative. This is similar to what we saw in the first chapter when we solved linear first order equations. In that case we sought an integrating factor. We can do the same thing here. We seek a multiplicative function $\mu(x)$ that we can multiply through (6.1) so that it can be written in Sturm-Liouville form. We first divide out the $a_2(x)$, giving

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

Now, we multiply the differential equation by μ :

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

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

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

This is formally solved to give

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

Thus, the original equation can be multiplied by factor

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

to turn it into Sturm-Liouville form.

In summary,

Equation (6.1),

$$a_2(x)y'' + a_1(x)y' + a_0(x)y = f(x), \tag{6.7}$$

can be put into the Sturm-Liouville form

$$\frac{d}{dx}\left(p(x)\frac{dy}{dx}\right) + q(x)y = F(x), \tag{6.8}$$

where

$$\begin{aligned} p(x) &= e^{\int \frac{a_1(x)}{a_2(x)}\,dx}, \\ q(x) &= p(x)\frac{a_0(x)}{a_2(x)}, \\ F(x) &= p(x)\frac{f(x)}{a_2(x)}. \end{aligned} \tag{6.9}$$

Example 6.2. For the example above,

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

We need only multiply this equation by

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

to put the equation in Sturm-Liouville form:

$$\begin{aligned} 0 &= xy'' + y' + \frac{2}{x}y \\ &= (xy')' + \frac{2}{x}y. \end{aligned} \tag{6.10}$$

6.2 Properties of Sturm-Liouville Eigenvalue Problems

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

1. The eigenvalues are real, countable, ordered and there is a smallest eigenvalue. Thus, we can write them as $\lambda_1 < \lambda_2 < \dots$. However, there is no largest eigenvalue and $n \to \infty$, $\lambda_n \to \infty$.
2. For each eigenvalue λ_n there exists an eigenfunction ϕ_n with $n-1$ zeros on (a, b).
3. Eigenfunctions corresponding to different eigenvalues are orthogonal with respect to the weight function, $\sigma(x)$. Defining the inner product of $f(x)$ and $g(x)$ as
$$< f, g >= \int_a^b f(x)g(x)\sigma(x)\, dx, \tag{6.11}$$
then the orthogonality of the eigenfunctios can be written in the form
$$< \phi_n, \phi_m >=< \phi_n, \phi_n > \delta_{nm}, \quad n, m = 1, 2, \dots. \tag{6.12}$$
4. The set of eigenfunctions is complete; i.e., any piecewise smooth function can be represented by a generalized Fourier series expansion of the eigenfunctions,
$$f(x) \sim \sum_{n=1}^{\infty} c_n \phi_n(x),$$
where
$$c_n = \frac{< f, \phi_n >}{< \phi_n, \phi_n >}.$$
Actually, one needs $f(x) \in L^2_\sigma[a, b]$, the set of *square integrable functions* over $[a, b]$ with weight function $\sigma(x)$. By square integrable, we mean that $< f, f >< \infty$. One can show that such a space is isomorphic to a *Hilbert space*, a complete inner product space.
5. Multiply the eigenvalue problem
$$\mathcal{L}\phi_n = -\lambda_n \sigma(x)\phi_n$$
by ϕ_n and integrate. Solve this result for λ_n, to find the *Rayleigh Quotient*
$$\lambda_n = \frac{-p\phi_n \frac{d\phi_n}{dx}\Big|_a^b - \int_a^b \left[p\left(\frac{d\phi_n}{dx}\right)^2 - q\phi_n^2 \right] dx}{< \phi_n, \phi_n >}$$
The Rayleigh quotient is useful for getting estimates of eigenvalues and proving some of the other properties.

Example 6.3. We seek the eigenfunctions of the operator found in Example 6.2. Namely, we want to solve the eigenvalue problem

$$\mathcal{L}y = (xy')' + \frac{2}{x}y = -\lambda\sigma y \tag{6.13}$$

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

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

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

Expanding the derivative, we have

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

Multiply through by x to obtain

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

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

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

The characteristic equation is

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

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

$$y(x) = c_1 \cos(\sqrt{\lambda+2}\ln|x|) + c_2 \sin(\sqrt{\lambda+2}\ln|x|). \tag{6.14}$$

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

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

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

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

This will give nontrivial solutions when

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

In summary, the eigenfunctions for this eigenvalue problem are

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

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

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

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

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

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

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

$$\begin{aligned}
< y_n, y_m > &= \int_1^2 \cos\left(\frac{n\pi}{\ln 2} \ln x\right) \cos\left(\frac{m\pi}{\ln 2} \ln x\right) \frac{dx}{x} \\
&= \frac{\ln 2}{\pi} \int_0^\pi \cos ny \cos my \, dy \\
&= \frac{\ln 2}{2} \delta_{n,m}. \qquad (6.15)
\end{aligned}$$

6.2.1 Adjoint Operators

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

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

We first introduce some definitions.

Definition 6.4. *The domain of a differential operator L is the set of all $u \in L^2_\sigma[a, b]$ satisfying a given set of homogeneous boundary conditions.*

Definition 6.5. *The adjoint, $L^\dagger$, of operator L satisfies*

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

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

Example 6.6. As an example, we find the adjoint of second order linear differential operator $L = a_2(x)\frac{d^2}{dx^2} + a_1(x)\frac{d}{dx} + a_0(x)$.

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

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

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

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

$$\int_a^b u(x)a_1(x)v'(x)\, dx = a_1(x)u(x)v(x)\Big|_a^b - \int_a^b (u(x)a_1(x))' v(x)\, dx. \quad (6.16)$$

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

$$\begin{aligned} \int_a^b u(x)a_2(x)v''(x)\, dx &= a_2(x)u(x)v'(x)\Big|_a^b - \int_a^b (u(x)a_2(x))' v(x)'\, dx \\ &= \left[a_2(x)u(x)v'(x) - (a_2(x)u(x))'v(x)\right]\Big|_a^b \\ &\quad + \int_a^b (u(x)a_2(x))'' v(x)\, dx. \end{aligned} \quad (6.17)$$

Combining these results, we obtain

$$\begin{aligned} < u, Lv > &= \int_a^b u(a_2 v'' + a_1 v' + a_0 v)\, dx \\ &= \left[a_1(x)u(x)v(x) + a_2(x)u(x)v'(x) - (a_2(x)u(x))'v(x)\right]\Big|_a^b \\ &\quad + \int_a^b \left[(a_2 u)'' - (a_1 u)' + a_0 u\right] v\, dx. \end{aligned} \quad (6.18)$$

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

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

This leaves

$$< u, Lv >= \int_a^b [(a_2 u)'' - (a_1 u)' + a_0 u]\, v\, dx \equiv < L^\dagger u, v > .$$

Therefore,

$$L^\dagger = \frac{d^2}{dx^2} a_2(x) - \frac{d}{dx} a_1(x) + a_0(x). \tag{6.19}$$

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

Example 6.7. Determine $L^\dagger$ and its domain for operator $Lu = \frac{du}{dx}$ where u satisfies the boundary conditions $u(0) = 2u(1)$ on $[0, 1]$.

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

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

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

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

6.2.2 Lagrange's and Green's Identities

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

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

we have the two identities:

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

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

$$\begin{aligned}
u\mathcal{L}v - v\mathcal{L}u &= u\left[\frac{d}{dx}\left(p\frac{dv}{dx}\right) + qv\right] - v\left[\frac{d}{dx}\left(p\frac{du}{dx}\right) + qu\right] \\
&= u\frac{d}{dx}\left(p\frac{dv}{dx}\right) - v\frac{d}{dx}\left(p\frac{du}{dx}\right) \\
&= u\frac{d}{dx}\left(p\frac{dv}{dx}\right) + p\frac{du}{dx}\frac{dv}{dx} - v\frac{d}{dx}\left(p\frac{du}{dx}\right) - p\frac{du}{dx}\frac{dv}{dx} \\
&= \frac{d}{dx}\left[pu\frac{dv}{dx} - pv\frac{du}{dx}\right]. \qquad (6.20)
\end{aligned}$$

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

6.2.3 Orthogonality and Reality

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

Theorem 6.8. *The eigenvalues of the Sturm-Liouville problem are real.*

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

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

The complex conjugate of this equation is

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

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

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

Integrate both sides of this equation:

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

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

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

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

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

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

Theorem 6.9. *The eigenfunctions corresponding to different eigenvalues of the Sturm-Liouville problem are orthogonal.*

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

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

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

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

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

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

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

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

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

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

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

6.2.4 The Rayleigh Quotient

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

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

by ϕ_n and integrating. This gives

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

One can solve the last equation for λ to find

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

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

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

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

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

$$\lambda_n = \frac{-p\phi_n \frac{d\phi_n}{dx}\Big|_a^b + \int_a^b \left[p\left(\frac{d\phi_n}{dx}\right)^2 - q\phi_n^2\right] dx}{\int_a^b \phi_n^2 \sigma\, dx}. \tag{6.21}$$

In many applications the sign of the eigenvalue is important. As we had seen in the solution of the heat equation, $T' + k\lambda T = 0$. Since we expect the heat energy to diffuse, the solutions should decay in time. Thus, we would expect $\lambda > 0$. In studying the wave equation, one expects vibrations and these are only possible with the correct sign of the eigenvalue (positive again). Thus, in order to have nonnegative eigenvalues, we see from (6.21) that

a. $q(x) \leq 0$, and
b. $-p\phi_n \frac{d\phi_n}{dx}|_a^b \geq 0$.

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

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

$$\lambda_1 = \min_{y(x)} \frac{-py\frac{dy}{dx}|_a^b + \int_a^b \left[p\left(\frac{dy}{dx}\right)^2 - qy^2\right] dx}{\int_a^b y^2\sigma\, dx} \tag{6.22}$$

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

Example 6.10. We have already solved the eigenvalue problem $\phi'' + \lambda\phi = 0$, $\phi(0) = 0$, $\phi(1) = 0$. In this case, the lowest eigenvalue is $\lambda_1 = \pi^2$. We can pick a nice function satisfying the boundary conditions, say $y(x) = x - x^2$. Inserting this into Equation (6.22), we find

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

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

6.3 The Eigenfunction Expansion Method

In section 4.3.2 we saw generally how one can use the eigenfunctions of a differential operator to solve a nonhomogeneous boundary value problem. In this chapter we have seen that Sturm-Liouville eigenvalue problems have the requisite set of orthogonal eigenfunctions. In this section we will apply the eigenfunction expansion method to solve a particular nonhomogenous boundary value problem.

Recall that one starts with a nonhomogeneous differential equation

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

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

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

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

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

and inserts the expansion into the nonhomogeneous equation. This gives

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

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

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

Orthogonality yields

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

Solving for c_m, we have

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

Example 6.11. As an example, we consider the solution of the boundary value problem

$$(xy')' + \frac{y}{x} = \frac{1}{x}, \quad x \in [1, e], \tag{6.23}$$

$$y(1) = 0 = y(e). \tag{6.24}$$

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

$$(x\phi')' + \frac{\phi}{x} = -\lambda\sigma\phi, \quad \phi(1) = 0 = \phi(e). \tag{6.25}$$

Rearranging the terms and multiplying by x, we have that

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

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

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

This is easily solved. The characteristic equation is

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

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

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

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

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

$$\begin{aligned} 1 &= \int_1^e \phi_n(x)^2\sigma(x)\,dx \\ &= A^2 \int_1^e \sin(n\pi \ln x)\frac{1}{x}\,dx \\ &= A^2 \int_0^1 \sin(n\pi y)\,dy = \frac{1}{2}A^2. \end{aligned} \tag{6.26}$$

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

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

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

Inserting this solution into the differential equation, we have

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

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

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

Solving for c_m, we have

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

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

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

6.4 The Fredholm Alternative Theorem

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

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

Theorem 6.12. First Alternative

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

or

Theorem 6.13. Second Alternative
A solution of $Ax = b$, if it exists, is unique if and only if $x = 0$ is the only solution of $Ax = 0$.

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

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

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

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

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

Example 6.14. Determine the allowed forms of $\mathbf{b}$ for a solution of $A\mathbf{x} = \mathbf{b}$ to exist, where

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

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

$$\begin{aligned} < A\mathbf{x}, \mathbf{y} > &= < \mathbf{x}, A^*\mathbf{y} > \\ \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_j\bar{y}_i &= \sum_{j=1}^{n} x_j \sum_{j=1}^{n} a_{ij}y_i \\ &= \sum_{j=1}^{n} x_j \overline{\sum_{j=1}^{n} (\bar{a}^T)_{ji}\, y_i} \end{aligned} \tag{6.27}$$

For this example,

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

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

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

So, what does this say about solutions of boundary value problems? There is a more general theory for linear operators. The matrix formulations follows, since matrices are simply representations of linear transformations. A more general statement would be

Theorem 6.15. *If L is a bounded linear operator on a Hilbert space, then $Ly = f$ has a solution if and only if $< f, v >= 0$ for every v such that $L^\dagger v = 0$.*

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

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

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

Theorem 6.16. *The solution of the boundary value problem $\mathcal{L}u = f$ with boundary conditions $Bu = g$ exists if and only if*

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

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

Example 6.17. Consider the problem

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

Only certain values of α and β will lead to solutions. We first note that $L = L^\dagger$ $=$

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

Solutions of

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

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

Next one computes

$$\begin{aligned} S(u,v) &= [u'v - uv']_0^{2\pi} \\ &= u'(2\pi)v(2\pi) - u(2\pi)v'(2\pi) - u'(0)v(0) + u(0)v'(0). \end{aligned} \tag{6.28}$$

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

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

Similarly,

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

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

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

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

Problems

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

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

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

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

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

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

6.6. In Example 6.10 we found a bound on the lowest eigenvalue for the given eigenvalue problem.

a. Verify the computation in the example.

b. Apply the method using

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

Is this an upper bound on λ_1

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

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

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

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