5 Sturm-Liouville Theory

Let's begin with an example to get the feel of the kind of problems we'll tackle with these techniques.

Consider a uniform string with fixed ends. The displacement of this string obeys the wave equation

$$\frac{1}{c^2}\frac{\partial^2 y}{\partial t^2} = \frac{\partial^2 y}{\partial x^2} \ , \tag{1}$$

with boundary conditions $y = 0$ at $x = 0$ and at $x = l$ for all time. To start, we separate variables, making the ansatz $y(x,t) = X(x)T(t)$. This yields

$$-\frac{1}{c^2}\frac{\ddot{T}}{T} = -\frac{X''}{X} = \text{const} = \lambda \ , \tag{2}$$

say. Vibrations correspond to $T \sim e^{i\omega t}$, so $\lambda = \omega^2/c^2$. We seek solutions of

$$X'' = -\lambda X \ , \tag{3}$$

subject to $X = 0$ at $x = 0$ and at $x = l$. The solutions are of the form

$$X_n(x) = B_n \sin\left(\frac{n\pi x}{l}\right) \ , \tag{4}$$

with $n = 1, 2, 3, \ldots$.

Note that there exist solutions only for $\lambda \in S$, a discretely distributed set of real eigenvalues. S is referred to as the *spectrum* of eigenvalues. The corresponding eigenfunctions X_n are orthogonal on $[0, l]$. The equation of motion for $X(x)$ is typical of a wide class of eigenvalue problems which arise from some partial differential equations of physics.

5.1 General Remarks

We will study (at first) second order linear differential equations for $x \in I = [a, b]$. Restrict attention to linear differential operators of the form $L = -a_2 D^2 - a_1 D + a_0$ for given functions $a_i(x)$ with $a_2(x) > 0$ on I. Actually, it will prove sufficient to restrict our attention to L of the *self-adjoint* form:

$$L = -DpD + q \ , \tag{5}$$

with $p(x) > 0$, $q(x)$ given functions on I.

the *Sturm-Liouville Problem* is specified by the differential equation

$$Ly(x) = \lambda w(x) y(x) \ , \tag{6}$$

to be solved for $y(x)$ for $x \in I$, subject to the boundary conditions (to be specified), with $w(x) > 0$ on I, and λ an eigenvalue parameter. The solution has some general features

1. $\exists$ nontrivial solutions which obey the boundary conditions in use iff $\lambda \in S$, the spectrum of the problem. S is a monotonic set of discretely distributed real eigenvalues λ_n, so that $\lambda_0 < \lambda_1 < \lambda_2 < \cdots$, with $\lambda_n \to \infty$ like n^2 as $n \to \infty$.

2. The eigenfunctions y_n corresponding to the $\lambda_n \in S$ are unique to within normalization. Also, y_n and y_m are orthogonal in a sense that we shall see soon, if $n \neq m$.

3. The y_n provide a basis in the infinite dimensional vector space of functions on I, which obey the boundary conditions in use, and suitable smoothness properties.

5.2 Orthogonality and Boundary Conditions

Suppose y_1 and y_2 obey the Sturm-Liouville equation. Then

$$\begin{aligned} -(py_1')' + qy_1 &= \lambda_1 w y_1 \ , & (7) \\ -(py_2')' + qy_2 &= \lambda_2 w y_2 \ , & (8) \end{aligned}$$

and suitable boundary conditions, for distinct values λ_1 and λ_2 of the eigenvalue parameter. Form the object

$$\int_a^b dx \, [y_2 \times (7) - y_1 \times (8)] \ . \tag{9}$$

This yields

$$\begin{aligned} (\lambda_1 - \lambda_2) \int_a^b dx \, w(x) y_1(x) y_2(x) &= \int_a^b dx \, [-(py_1')' y_2 + (py_2')' y_1] \ , \\ &= [-(py_1') y_2 + (py_2') y_1]_a^b \ . & (10) \end{aligned}$$

Now, appropriate boundary conditions are those that make this vanish. For then, since $\lambda_1 \neq \lambda_2$, we have

$$\int_a^b dx \, w(x) y_1(x) y_2(x) = 0 \ . \tag{11}$$

This is the sense in which y_1 and y_2 are *orthogonal* with respect to the *weight function* $w(x)$.

A good example is given by Legendre's equation and polynomials. This arises from the Laplace equation in cylindrical coordinates (r, θ, z). Writing $x \equiv \cos\theta$, the Legendre equation is

$$-\frac{d}{dx}\left[(1-x^2)\frac{d}{dx}P(x)\right] = \lambda P(x) \ , \tag{12}$$

with $I = (-1, 1)$. Note that this is a Sturm-Liouville problem with $w(x) = 1$, and $p(x) = (1 - x^2)$. The suitable boundary conditions are automatically imposed if $P(x)$ is finite at $x = \pm 1$, since $p(x) \to 0$ at the endpoints. The solutions are a set of polynomials, the first few of which are

$$P_0 = 1 \quad , \quad \lambda_0 = 0 \ , \tag{13}$$

$$P_1 = x \quad , \quad \lambda_1 = 2 \ , \tag{14}$$

$$P_2 = x^2 - \frac{1}{3} \quad , \quad \lambda_2 = 6 \ . \tag{15}$$

More generally, there exists a unique $P_n = x^n + \cdots$ with $\lambda_n = n(n+1)$. It can also be checked that the P_n are orthogonal on I.

5.3 Real Eigenvalues

Let us allow the possibility the λ_n and y_n are complex. Then, we may write two Sturm-Liouville equations as

$$-(py')' + qy = \lambda w y \ , \tag{16}$$

$$-(p(y^*)')' + qy^* = \lambda^* w y^* \ . \tag{17}$$

We now form the object

$$\int_a^b dx\,[y^* \times (16) - y \times (17)] \ . \tag{18}$$

This yields

$$\begin{aligned} (\lambda - \lambda^*)\int_a^b dx\, w(x)y(x)y(x)^* &= \int_a^b dx\,[-(p(y^*)')'y + (py')'y^*] \ , \\ &= [-(p(y^*)')y + (py')y^*]_a^b \ , \\ &= 0 \ , \end{aligned} \tag{19}$$

for the suitable boundary conditions. Now, since $w(x) > 0$ on I, this implies that $\int_a^b dx\, w|y|^2$ is strictly positive on I. Therefore,

$$\lambda^* = \lambda \ . \tag{20}$$

i.e., λ is real.

5.4 Formal Vector Space View

Let us regard suitably behaved functions $f(x)$, $g(x)$, which obey the boundary conditions of our Sturm-Liouville problem as elements of an infinite dimensional vector space $\mathcal{V}$, spanned by the $y_n(x)$. Thus, we can write

$$f(x) = \sum_{n=0}^{\infty} c_n y_n(x) \ . \tag{21}$$

Define a scalar (inner) product on $\mathcal{V}$ by

$$(f, g) \equiv \int_a^b dx\, w(x) f(x) g(x) \ , \tag{22}$$

and note that "suitably well-behaved" requires that $||f|| = (f, f)^{1/2}$ exists. Then, $(y_n, y_m) = 0$ for $n \neq m$. Further, choose the scale of y_n to achieve orthonormality:

$$(y_n, y_n) = \delta_{nm} \ . \tag{23}$$

Also

$$\begin{aligned} (y_m, f) &= \sum_n c_n (y_n, y_m) \ , \\ &= \sum_n cm \delta_{nm} \ , \\ &= c_m \ . \end{aligned} \tag{24}$$

This is to be compared with the formulae for Fourier series. Thus, if we assume that the y_n are known, then for a given f we can obtain the c_m via

$$c_n = \int_a^b d\xi\, w(\xi) y_n(\xi) f(\xi) \ . \tag{25}$$

An important and useful result follows if we substitute this into the expansion for $f(x)$. We obtain

$$\begin{aligned} f(x) &= \sum_{n=0}^{\infty} \left[\int_a^b d\xi\, w(\xi) y_n(\xi) f(\xi) \right] y_n(x) \ , \\ &= \int_a^b d\xi \left[\sum_{n=0}^{\infty} w(\xi) y_n(\xi) y_n(x) \right] f(\xi) \ . \end{aligned} \tag{26}$$

Since this is true $\forall\ f \in \mathcal{V}$, we may therefore infer

$$\sum_{n=0}^{\infty} w(\xi) y_n(\xi) y_n(x) = \delta(x - \xi) \ . \tag{27}$$

This is a formal *completeness relation* for the Sturm-Liouville problem. Note that we can check this if we assume that $\delta(x-\xi)$ obeys the boundary conditions for $a<\xi<b$. We can then expand the delta function as

$$\delta(x-\xi)=\sum_{n=0}^{\infty} c_n(\xi) y_n(\xi) \ . \tag{28}$$

Then, our expression for c_n gives

$$\begin{aligned} c_n(\xi) &= \int_a^b dx\, w(x) y_n(x) \delta(x-\xi) \ , \\ &= w(\xi) y_n(\xi) \ . \end{aligned} \tag{29}$$

5.5 Inhomogeneous Equations and Green's Functions

Suppose we have solved the Sturm-Liouville problem

$$L y_n(x)=\lambda_n w(x) y_n(x) \ , \tag{30}$$

where the y_n obey suitable boundary conditions on $I=[a,b]$, and $w(x)>0$ on I. By *solved* I mean that the λ_n and y_n are determined, and $(y_n, y_m)=\delta_{nm}$ has been arranged.

We would like to solve the problem

$$[L-\lambda w(x)] y(x)=f(x) \ , \tag{31}$$

for $y(x)$, for $x\in I$, subject to the same boundary conditions. Here λ is a fixed real number, and $f(x)$ is a given function; naturally assumed to obey the same boundary conditions. To proceed, write

$$\begin{aligned} f(x) &= w(x) h(x) \ , \\ h(x) &= \sum_{n=0}^{\infty} h_n y_n(x) \ , \end{aligned} \tag{32}$$

where we can calculate the coefficients h_n when f (and hence h) is given, and the y_n are known.

We now posit the expansion

$$y(x)=\sum_{n=0}^{\infty} a_n y_n(x) \ , \tag{33}$$

and seek the unknowns a_n to complete the specification of the solution. Substituting in to the problem (31), the left hand side becomes

$$\begin{aligned}\sum_{n=0}^{\infty} a_n (L - \lambda w) y_n(x) &= \sum_{n=0}^{\infty} a_n (\lambda_n w - \lambda w) y_n(x) \ , \\ &= w(x) \sum_{n=0}^{\infty} a_n (\lambda_n - \lambda) y_n(x) \ . \ ,\end{aligned}$$

Now the right hand side is

$$w(x) \sum_{n=0}^{\infty} h_n y_n(x) \ . \tag{34}$$

Finally, equating these, multiplying both sides by $y_m(x)$, and integrating over $a < x < b$, we obtain

$$a_m (\lambda_m - \lambda) = h_m \ . \tag{35}$$

Now, if $\lambda \notin S$, we have

$$a_n = \frac{h_n}{\lambda_n - \lambda} \ . \tag{36}$$

We now have the solution in terms of quantities calculated for the homogeneous equation. It is instructive to write this in another way. The solution is

$$\begin{aligned} y(x) &= \sum_{n=0}^{\infty} y_n(x) \frac{h_n}{\lambda_n - \lambda} \ , \\ &= \sum_{n=0}^{\infty} y_n(x) \frac{1}{\lambda_n - \lambda} \left[\int_a^b d\xi \, w(\xi) y_n(\xi) h(\xi) \right] \ , \\ &= \int_a^b d\xi \left[\sum_{n=0}^{\infty} \frac{y_n(x) y_n(\xi)}{\lambda_n - \lambda} \right] f(\xi) \ , \end{aligned} \tag{37}$$

In this form it is easy to identify the *Green's Function* in the form

$$G(x, \xi) = \sum_{n=0}^{\infty} \frac{y_n(x) y_n(\xi)}{\lambda_n - \lambda} \ , \tag{38}$$

We can check that this Green's function behaves as it should, by acting on it with the left hand side of (31):

$$\begin{aligned} [L - \lambda w(x)] G(x, \xi) &= \sum_{n=0}^{\infty} \frac{y_n(\xi)}{\lambda_n - \lambda} (L - \lambda w) y_n(x) \ , \\ &= \sum_{n=0}^{\infty} \frac{y_n(\xi)}{\lambda_n - \lambda} (\lambda_n - \lambda) w(x) y_n(x) \ , \\ &= \sum_{n=0}^{\infty} y_n(\xi) w(x) y_n(x) \ , \end{aligned}$$

which, using our earlier results, is

$$[L - \lambda w(x)] G(x, \xi) = \delta(x - \xi) \ , \tag{39}$$

as expected.

5.6 Self-Adjointness

It has probably not escaped your notice that the Sturm-Liouville problem and its solutions are related to things that you've seen in quantum mechanics. Here, we shall see precisely what this relation is. In particular, we shall compare the terms *self-adjoint* (in Sturm-Liouville theory) and *Hermitian* (in quantum mechanics).

I will begin by considering the case of $w(x) = 1$. Consider a typical Sturm-Liouville problem

$$\begin{aligned} Ly &= \lambda y \ , \\ L &= -DpD + q \ , \end{aligned} \tag{40}$$

with boundary conditions $y(a) = y(b) = 0$. In the vector space $\mathcal{V}$ of (possibly complex) functions f, g, obeying the boundary conditions, we will use the inner product

$$(f,g) = \int_a^b dx\, f(x)^* g(x) \ . \tag{41}$$

We are interested in operators A such that $Af \in \mathcal{V}$, $\forall f \in \mathcal{V}$. Define the *Hermitian conjugate* (or Hermitian adjoint) operator, $A^\dagger$ by

$$(A^\dagger f, g) = (f, Ag) \ , \tag{42}$$

$\forall\ f,\ g \in \mathcal{V}$. We say that the operator A is *hermitian* if $A^\dagger = A$, or

$$(Af, g) = (f, Ag) \ , \tag{43}$$

With this definition, note that our self-adjoint operator L is hermitian (for the special case $w = 1$). You can check this trivially using the definition of the inner product. Do this as a brief exercise.

Now turn to the case of $w \neq 1$. In this course, and in Sturm-Liouville theory in general, we refer to the operator L as self-adjoint, even when $w \neq 1$. we also use $Ly = \lambda w y$ and $(f,g) = \int_a^b dx\, w f g$, from which real λ_n and orthogonal y_n follow. However, with this definition, $(Lf, g) \neq (f, Lg)$, because of w. To understand the relationship to hermiticity, we define

$$M \equiv -w^{-1/2} DpDw^{1/2} + q \ , \tag{44}$$

which is hermitian. To see this:

$$\begin{aligned}
(Mf, g) &= -\int_a^b dx\, ww^{-1/2}[DpDw^{1/2}f]g + \cdots \ , \\
&= \int_a^b dx\, (pDw^{1/2}f)D(w^{1/2}g) + \cdots \ , \\
&= (f, Mg) \ . \qquad (45)
\end{aligned}$$

Therefore, care is needed with terminology, even though either definition leads to real eigenvalues and orthogonal eigenfunctions.