

Sturm-Liouville Theory

In the three preceding lectures I demonstrated the utility of Fourier series in solving PDE/BVPs. As we'll now see, Fourier series are just the "tip of the iceberg" of the theory and utility of special functions. Before proceeding with the general theory, let me state clearly the basic properties of Fourier series we intend to generalize:

- (i) The Fourier sine functions $\{\sin(\frac{n\pi}{L}x) \mid n = 1, 2, 3, \dots\}$ and cosine functions $\{\cos(\frac{n\pi}{L}x) \mid n = 0, 1, 2, \dots\}$ are solutions of a second order linear homogeneous differential equation

$$(1) \quad y'' = -\lambda y$$

satisfying certain linear homogenous boundary conditions: viz, the Fourier-sine functions satisfy

$$(2a) \quad y(0) = 0 = y(L)$$

and the Fourier-cosine functions satisfy

$$(2b) \quad y'(0) = 0 = y'(L) \quad .$$

- (ii) The Fourier sine and cosine functions obey certain orthogonality relations

$$\begin{aligned} \frac{2}{L} \int_0^L \sin\left(\frac{n\pi}{L}x\right) \sin\left(\frac{m\pi}{L}x\right) dx &= \begin{cases} 1 & \text{if } n = m \\ 0 & \text{if } n \neq m \end{cases} \\ \frac{2}{L} \int_0^L \cos\left(\frac{n\pi}{L}x\right) \cos\left(\frac{m\pi}{L}x\right) dx &= \begin{cases} 1 & \text{if } n = m \\ 0 & \text{if } n \neq m \end{cases} \end{aligned}$$

- (iii) Any continuous, differentiable function on $[0, L]$ can be expressed in terms of a Fourier-sine or Fourier-cosine expansion

$$\begin{aligned} f(x) &= \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\left(\frac{n\pi}{L}x\right) \\ &= \sum_{n=1}^{\infty} b_n \sin\left(\frac{n\pi}{L}x\right) \end{aligned}$$

Moreover, the orthogonality relations (ii) allow us to determine the coefficients a_n and b_n

$$\begin{aligned} a_n &= \frac{2}{L} \int_0^L f(x) \cos\left(\frac{n\pi}{L}x\right) dx \\ b_n &= \frac{2}{L} \int_0^L f(x) \sin\left(\frac{n\pi}{L}x\right) dx \end{aligned}$$

1. Sturm-Liouville Problems

DEFINITION 15.1. A **Sturm-Liouville problem** is a second order homogeneous linear differential equation of the form

$$(3) \quad \frac{d}{dx} \left[p(x) \frac{dy}{dx} \right] - q(x)y + \lambda r(x)y = 0 \quad , \quad \forall x \in [0, 1]$$

together with boundary conditions of the form

$$(4a) \quad a_1 y(0) + a_2 y'(0) = 0$$

$$(4b) \quad b_1 y(1) + b_2 y'(1) = 0$$

In what follows, λ is to be regarded as a constant parameter and we shall always assume that $p(x)$ and $r(x)$ are positive functions on the interval $[0, 1]$

$$(5) \quad \begin{aligned} p(x) &> 0 \quad \forall x \in [0, 1] \\ r(x) &> 0 \quad \forall x \in [0, 1] \end{aligned}$$

EXAMPLE 15.2. Taking $p(x) = r(x) = 1$, $q(x) = 0$, $a_1 = b_1 = 1$, $a_2 = b_2 = 0$, we see that the corresponding Sturm-Liouville problem

$$\begin{aligned} y'' + \lambda y &= 0 \\ y(0) &= 0 \\ y(1) &= 0 \end{aligned}$$

has as its solutions

$$\lambda = (n\pi)^2 \quad n = 1, 2, 3, \dots \quad \text{and} \quad y(x) = \sin(n\pi x)$$

Note, in particular, that the solution of the Sturm-Liouville problem only exists for certain values of λ ; these values are “**eigenvalues**” of the Sturm-Liouville problem

NOTATION 15.3 (Linear differential operator notation). In what follows we shall denote by L the “linear differential operator”

$$(6) \quad L = \frac{d}{dx} \left[p(x) \frac{d}{dx} \right] - q(x)$$

which acts on a function $y(x)$ by

$$L[y] = \frac{d}{dx} \left[p(x) \frac{dy}{dx} \right] - q(x)y$$

In terms of L , the differential equation of a Sturm-Liouville problem can be expressed

$$L[y] = -\lambda r(x)y$$

THEOREM 15.4 (Lagrange’s identity). *Suppose u and v are two functions satisfying boundary conditions of the form*

$$(7a) \quad a_1 y(0) + a_2 y'(0) = 0$$

$$(7b) \quad b_1 y(1) + b_2 y'(1) = 0$$

then

$$\int_0^1 (L[u]v - uL[v]) dx = 0$$

Proof. We have

$$\begin{aligned}
\int_0^1 L[u]v dx &= \int_0^1 \left(\frac{d}{dx} \left[p \frac{du}{dx} \right] v - quv \right) dx \\
&= \int_0^1 \frac{d}{dx} \left[p \frac{du}{dx} \right] v dx - \int_0^1 quv dx \\
&= p \frac{du}{dx} v \Big|_0^1 - \int_0^1 \left(p \frac{du}{dx} \right) \frac{dv}{dx} dx - \int_0^1 quv dx \\
&= p \frac{du}{dx} v \Big|_0^1 - \int_0^1 \frac{du}{dx} \left(p \frac{dv}{dx} \right) dx - \int_0^1 quv dx \\
&= p \frac{du}{dx} v \Big|_0^1 - \left(up \frac{dv}{dx} \right) \Big|_0^1 + \int_0^1 u \frac{d}{dx} \left(p \frac{dv}{dx} \right) dx - \int_0^1 quv dx \\
(*) \quad &= p(u'v - uv') \Big|_0^1 + \int_0^1 uL[v] dv
\end{aligned}$$

where we have twice utilized the integration by parts formula

$$\int_a^b \frac{df}{dx} g dx = fg \Big|_a^b - \int_a^b f \frac{dg}{dx} dx$$

Now consider the first term in the last line

$$p(u'v - uv') \Big|_0^1$$

Since u and v satisfy (7a) and (7b) we have assuming $a_1, a_2, b_1, b_2 \neq 0$

$$\begin{aligned}
u'(0) &= -\frac{a_2}{a_1} u(0) & , & \quad v'(0) = -\frac{a_2}{a_1} v(0) \\
u'(1) &= -\frac{b_2}{b_1} u(1) & , & \quad v'(1) = -\frac{b_2}{b_1} v(1)
\end{aligned}$$

and

$$\begin{aligned}
u'(0)v(0) - u(0)v'(0) &= -\frac{a_2}{a_1} u(0)v(0) - u(0) \left(-\frac{a_2}{a_1} v(0) \right) = 0 \\
u'(1)v(1) - u(1)v'(1) &= -\frac{b_2}{b_1} u(1)v(1) - u(1) \left(-\frac{b_2}{b_1} v(1) \right) = 0
\end{aligned}$$

Hence

$$\begin{aligned}
p(u'v - uv') \Big|_0^1 &= p(1)(u'(1)v(1) - u(1)v'(1)) - p(0)(u'(0)v(0) - u(0)v'(0)) \\
&= p(1) \cdot 0 - p(0) \cdot 0 \\
&= 0
\end{aligned}$$

We can hence conclude from (*) that

$$\int_0^1 \left(L[u]v = \int_0^1 uL[v] \right) dx$$

and the theorem follows. \square

Recall that, for example, the Fourier-sine functions $\sin(\lambda x)$ not only satisfy certain a second order linear differential equation

$$f'' = -\lambda^2 f$$

but they also satisfy certain linear homogeneous boundary conditions

$$f(0) = 0 = f(L)$$

exactly when the parameter λ is “tuned” to the boundary conditions

$$\left. \begin{array}{l} f'' = -\lambda^2 f \\ f(0) = 0 \\ f(L) = 0 \end{array} \right\} \implies \left\{ \begin{array}{l} \lambda = \frac{n\pi}{L} \\ f(x) = \sin\left(\frac{n\pi}{L}x\right) \end{array} \right.$$

What may seem a little surprising at first is that the fact that the Fourier-sine functions are solutions to a Sturm-Liouville problem is also responsible for their orthogonality properties

$$\frac{2}{L} \int_0^L \sin\left(\frac{n\pi}{L}x\right) \sin\left(\frac{m\pi}{L}x\right) dx = \begin{cases} 1 & \text{if } n = m \\ 0 & \text{if } n \neq m \end{cases}$$

THEOREM 15.5. *If ϕ and ψ are two solutions of a Sturm-Liouville problem corresponding to different values of the parameter λ then*

$$\int_0^1 \phi(x) \psi(x) r(x) dx = 0 \quad .$$

Proof. Suppose

$$\begin{aligned} L[\phi] &= -\lambda_1 r(x) \phi \\ L[\psi] &= -\lambda_2 r(x) \psi \end{aligned}$$

with $\lambda_1 \neq \lambda_2$. Then since ϕ, ψ satisfy the Sturm-Liouville boundary conditions, we have by Theorem 11.1

$$\int_0^1 L[\phi] \psi dx = \int_0^1 \phi (L[\psi]) dx$$

and because they satisfy Sturm-Liouville-type differential equations

$$\int_0^1 (-\lambda_1 r(x) \phi(x)) \psi(x) dx = \int_0^1 \phi(x) (-\lambda_2 r(x) \psi(x)) dx$$

or

$$(\lambda_1 - \lambda_2) \int_0^1 \phi(x) \psi(x) r(x) dx = 0$$

By hypothesis, $\lambda_1 \neq \lambda_2$ and so we must conclude

$$\int_0^1 \phi(x) \psi(x) r(x) dx = 0$$

□

As remarked above, it is common to write (in shorthand) the differential equation of a Sturm-Liouville problem as

$$(8) \quad L[y] = -\lambda r(x) y$$

and refer to the parameter λ (corresponding to a particular solution $y(x)$) as an eigenvalue of the differential operator L . This language, of course, is derived from the nomenclature of linear algebra where if \mathbf{A} is an $n \times n$ matrix, \mathbf{v} is an $n \times 1$ column vector and λ is a number such that

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$$

then \mathbf{v} is called an eigenvector of \mathbf{A} and λ is the corresponding eigenvalue of \mathbf{A} . In fact, it is common to use such linear-algebraic-like nomenclature throughout Sturm-Liouville theory; and henceforth we shall refer to the functions $y(x)$ that satisfy (8) as *eigenfunctions* of L and the numbers λ as the corresponding *eigenvalues* of L .

THEOREM 15.6. *All the eigenvalues of a Sturm-Liouville problem are real.*

Proof. Consider the following inner product on the space of complex-valued solutions of a Sturm-Liouville problem (8) (for various λ)

$$\langle u, v \rangle = \int_0^1 u(x) \bar{v}(x) dx$$

Suppose

$$u(x) = v(x) = \mathcal{R}(x) + i\mathcal{I}(x)$$

that is, suppose we set u equal to v and split it up into its real and imaginary parts. Then

$$\begin{aligned} \langle u, u \rangle &= \int_0^1 (\mathcal{R}(x) + i\mathcal{I}(x)) (\mathcal{R}(x) - i\mathcal{I}(x)) dx \\ &= \int_0^1 [(\mathcal{R}(x))^2 + (\mathcal{I}(x))^2] dx \end{aligned}$$

Note now that the integrand is always non-negative. By a well-known result from Calculus, the integral of a continuous non-negative function over a finite interval is always non-negative, and zero only if the function is identically equal to zero. We conclude that for any continuous, non-zero, complex-valued function $u(x)$

$$\int_0^1 u(x) \overline{u(x)} dx > 0$$

Now consider Lagrange's identity (Theorem 11.4) when we set $v(x) = \overline{u(x)}$,

$$(9) \quad \int_0^1 u(x) L[\overline{u(x)}] dx = \int_0^1 \overline{u(x)} L[u(x)] dx$$

and suppose $u(x)$ is a Sturm-Liouville eigenfunction:

$$(10) \quad L[u] = -\lambda r u \quad \text{for some } \lambda \in \mathbb{C}$$

Taking the complex conjugate of this equation, we have, because $p(x)$, $q(x)$, and $r(x)$ are all assumed to be real-valued functions,

$$\begin{aligned} L[u] = -\lambda r u &\implies \overline{L[u]} = -\bar{\lambda} r \bar{u} \\ &\implies \overline{\frac{d}{dx} \left[p(x) \frac{du}{dx} \right] - q(x) u} = -\bar{\lambda} r \bar{u} \\ &\implies \frac{d}{dx} \left[p(x) \frac{d\bar{u}}{dx} \right] - q(x) \bar{u} = -\bar{\lambda} r \bar{u} \end{aligned}$$

or

$$(11) \quad L[\bar{u}] = -\bar{\lambda} r \bar{u}$$

In other words, if $u(x)$ is a solution of an S-L problem corresponding to eigenvalue λ , its complex conjugate function will be a solution of the same S-L problem, but with eigenvalues $\bar{\lambda}$.

Now plugging the right hand sides of (10) and (11) into (9), we can conclude that

$$\int_0^1 \lambda (ru) \bar{u} dx = \int_0^1 \bar{\lambda} u (r\bar{u}) dx$$

or

$$(12) \quad (\lambda - \bar{\lambda}) \int_0^1 ru\bar{u} dx = 0$$

We have argued above that $u(x) \overline{u(x)}$ is a non-negative function, and by assumption $r(x)$ is a positive function on $[0, 1]$. Thus, the integrand is a non-negative function and

$$0 < \int_0^1 ru\bar{u} dx$$

so long as $u(x)$ is not identically zero. We can therefore conclude from (12) that

$$\lambda - \bar{\lambda} = 0 \implies \lambda \in \mathbb{R} \quad .$$

□

In Linear Algebra, we say that an eigenvalue λ of a matrix \mathbf{A} has multiplicity m if the dimension of the corresponding eigenspace is m ; that is to say,

$$m = \dim \text{NullSp}(\mathbf{A} - \lambda \mathbf{I})$$

In Sturm-Liouville theory, we say that the multiplicity of an eigenvalue λ of a Sturm-Liouville problem

$$\begin{aligned} L[\phi] &= -\lambda r(x) \phi(x) \\ a_1 \phi(0) + a_2 \phi'(0) &= 0 \\ b_1 \phi(1) + b_2 \phi'(1) &= 0 \end{aligned}$$

if there are exactly m linearly independent solutions for that value of λ .

THEOREM 15.7. *The eigenvalues of a Sturm-Liouville problem are all of multiplicity one. Moreover, the eigenvalues form an infinite sequence and can be ordered according to increasing magnitude:*

$$\{\text{eigenvalues}\} = \{\lambda_1, \lambda_2, \lambda_3, \dots\} \quad , \quad \lambda_1 < \lambda_2 < \lambda_3 < \dots$$

Proof. By "all eigenvalues have multiplicity one", we mean that for each Sturm-Liouville eigenvalue λ , there is only one linearly independent solution $\phi(x)$ of the Sturm-Liouville problem. Suppose there were two, ϕ_1 and ϕ_2 , linearly independent solutions of

$$\begin{aligned} \frac{d}{dx} \left(p(x) \frac{d\phi}{dx} \right) - q(x) \phi &= -\lambda r(x) \phi(x) \\ a_1 \phi(0) + a_2 \phi'(0) &= 0 \\ b_1 \phi(1) + b_2 \phi'(1) &= 0 \end{aligned}$$

Then, as solutions of a second order, homogeneous, linear ODE, linear independence implies

$$W[\phi_1, \phi_2] \neq 0 \quad \text{for all } t \in [0, 1]$$

To see this explicitly, need some properties of the Wronskian.

LEMMA 15.8. *If $W[y_1, y_2](x) \equiv y_1(x)y_2'(x) - y_1'(x)y_2(x) = 0$, then $y_2(x) = Cy_1(x)$, with C a constant.*

Proof. We can regard the hypothesis of the lemma as a differential equation for y_2 . Putting it in the standard form of a first order linear ODE, it becomes

$$y_2' - \left(\frac{y_1'}{y_1} \right) y_2 = 0$$

which has as its general solution

$$y_2(x) = C \exp \left[- \int \frac{-y_1'(x)}{y_1(x)} dx \right] = C \exp \left(\int \frac{d}{dx} (\ln |y_1(x)|) \right) = C \exp (\ln |y_1(x)|) = Cy_1(x)$$

□

We also have the following lemma (which is in fact known as Abel's Theorem)

LEMMA 15.9. *If y_1 and y_2 are two solutions of $y'' + p(x)y' + qy = 0$ then*

$$(13) \quad W[y_1, y_2](x) = c \exp \left(- \int p(x) dx \right) \quad .$$

Proof. Since y_1 and y_2 are solutions

$$\begin{aligned} y_1'' + p(x)y_1' + qy_1 &= 0 \\ y_2'' + p(x)y_2' + qy_2 &= 0 \end{aligned}$$

If we multiply the first equation by $-y_2(x)$, the second equation by $y_1(x)$ and add the resulting equations, we have

$$(y_1 y_2'' - y_2 y_1'') - p(x)(y_1 y_2' - y_1' y_2) = 0$$

or

$$\frac{d}{dx}(y_1 y_2' - y_1' y_2) + p(x)(y_1 y_2' - y_1' y_2) = 0$$

which says the Wronskian $W(x) = y_1(x)y_2'(x) - y_1'(x)y_2(x)$ of y_1 and y_2 satisfies

$$\frac{d}{dx}W(x) + p(x)W(x) = 0 \quad .$$

This is a first order, homogeneous, linear differential equation whose solution is well-known to be

$$W(x) = C \exp \left[- \int p(x) dx \right]$$

□

COROLLARY 15.10. *If $y_1(x)$ and $y_2(x)$ are two solutions of a second order, homogeneous, linear ODE, and $W[y_1, y_2](x) = 0$ at any point x in their domain, then $y_1(x)$ and $y_2(x)$ are linearly dependent (i.e. $y_2(x) = cy_1(x)$ with c a constant).*

Proof. Note that the exponential factor in this formula (13) for $W(x)$ is never equal to 0 (so long as its argument is defined). Therefore, if $W(x) = 0$ at some point x , the constant C must equal zero, and hence $W(x)$ must vanish at all points x . But if $W(x) = 0$ for all x , we have the hypothesis of Lemma 15.8 and so $y_2(x) = Cy_1(x)$. □

Let's now return to the setting of the theorem we're trying to prove, where ϕ_1 and ϕ_2 are two solutions of a S-L problem with the same eigenvalue λ . In view of Corollary 15.10, if I can show that $W[\phi_1, \phi_2](0) = 0$, I can conclude that $W[\phi_1, \phi_2](x) = 0$ for all x , and hence ϕ_1 and ϕ_2 are linearly dependent.

Now since ϕ_1 and ϕ_2 also satisfy the S-L boundary conditions

$$\begin{aligned} a_1 \phi_1(0) + a_2 \phi_1'(0) &= 0 \\ a_1 \phi_2(0) + a_2 \phi_2'(0) &= 0 \end{aligned}$$

or

$$\begin{bmatrix} \phi_1(0) & \phi_1'(0) \\ \phi_2(0) & \phi_2'(0) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

This a homogeneous 2×2 system. As such it either has a unique solution (which is $a_1 = a_2 = 0$, which will not occur if we have a legitimate S-L boundary condition at $x = 0$) or

$$0 = \det \begin{pmatrix} \phi_1(0) & \phi_1'(0) \\ \phi_2(0) & \phi_2'(0) \end{pmatrix} = \phi_1(0)\phi_2'(0) - \phi_1'(0)\phi_2(0) = W[\phi_1, \phi_2](0)$$

We can now argue

$$W[\phi_1, \phi_2](0) = 0 \quad \Rightarrow \quad W[\phi_1, \phi_2](x) = 0 \text{ for all } x \quad \Rightarrow \quad \phi_2(x) = \lambda \phi_1(x) \quad \Rightarrow \quad \phi_1, \phi_2 \text{ are linearly dependent}$$

The theorem also includes the statement that the eigenvalues of a S-L can be listed in a natural order:

$$\lambda_0 < \lambda_1 < \lambda_2 < \dots$$

Since we already know the eigenvalues are real and distinct, the only thing we really have to show that there is a lowest eigenvalue λ_0 . This is actually not too hard to do, but the argument takes us well astray of the rest of the course. (It involves the use of a sort of "second derivative test" from variational calculus to show that a "minimal solution" exists.)

The ordering of the eigenvalues λ , and the fact that they multiplicity free, establishes a certain canonical ordering of the corresponding Sturm-Liouville eigenfunctions. We denote by ϕ_n the solution of the Sturm-Liouville problem with $\lambda = \lambda_n$. Actually, this only determines $\phi_n(x)$ up to a constant multiple (because if

$\phi(x)$ satisfies (3), (4a), (4b) then so does any constant multiple of $\phi(x)$. It is common practise to remove this ambiguity by demanding in addition

$$\int_0^1 \phi_n(x) \overline{\phi_n(x)} r(x) dx = 1$$

which fixes $\phi(x)$ up to a scalar factor of the form $e^{i\alpha}$.

Moreover, the facts that the eigenvalues are all real and multiplicity free, also implies that we can choose the $\phi_n(x)$ to be real-valued functions of x (because the complex span of the functions $\phi_n(x)$ must coincide with that of $\overline{\phi_n(x)}$). Thus, we can choose the eigenfunctions of a Sturm-Liouville problem to be real-valued functions

$$\int_0^1 \phi_n(x) \overline{\phi_n(x)} r(x) dx = 1$$

THEOREM 15.11. *Let $\phi_1, \phi_2, \phi_3, \dots$ be the normalized eigenfunctions of a Sturm-Liouville problem and suppose f is a piecewise continuous function on $[0, 1]$. Then if*

$$c_n := \int_0^1 f(x) \phi_n(x) r(x) dx$$

the series

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

converges to

$$\frac{f(x_+) + f(x_-)}{2}$$

at each point on $(0, 1)$.

Here

$$f(x_{\pm}) := \lim_{\varepsilon \rightarrow 0^+} f(x \pm \varepsilon)$$

Note that at any point x where f is continuous

$$\frac{f(x_+) + f(x_-)}{2} = f(x) \quad .$$

In summary, any time you have a differential equation of the form

$$\frac{d}{dx} \left[p(x) \frac{dy}{dx} \right] - q(x) y = -\lambda r(x) y$$

with homogeneous boundary conditions of the form

$$a_1 y(0) + a_2 y'(0) = 0 = b_1 y(1) + b_2 y'(1)$$

Then:

- Solutions will exist for only a discrete (but otherwise infinite) set of values for the parameter λ , all of which are real numbers. There is a unique lowest eigenvalue λ_0 and the other eigenvalues can be totally ordered

$$\lambda_0 < \lambda_1 < \lambda_2 < \lambda_3 < \dots$$

- For each λ_n , there is a solution ϕ_n of

$$\frac{d}{dx} \left[p(x) \frac{dy}{dx} \right] - q(x) y = -\lambda_n r(x) y$$

that is unique up to a scalar factor.

- The solutions ϕ_n , $n = 0, 1, 2, \dots$ can be normalized such that

$$\int_0^1 \phi_n(x) \phi_m(x) r(x) dx = \begin{cases} 1 & \text{if } n = m \\ 0 & \text{otherwise} \end{cases}$$

- Any continuous function $f(x)$ on the interval $[0, 1]$ can be expanded in terms of the Sturm-Liouville eigenfunctions ϕ_n , $n = 0, 1, 2, \dots$

$$f(x) = \sum_{n=0}^{\infty} a_n \phi_n(x)$$

with the coefficients a_n being determined by

$$a_n = \int_0^1 f(x) \phi_n(x) r(x) dx$$

If $f(x)$ is merely piece-wise continuous on $[0, 1]$, one has instead

$$\frac{f(x_+) + f(x_-)}{2} = \sum_{n=0}^{\infty} a_n \phi_n(x)$$

with the coefficients a_n determined by the same formula.