LECTURE 31: THE STURM-LIOUVILLE BVP

In the preceding lecture, we derived the Sturm-Liouville type boundary value problems as a generalization of the heat equations in a rod. Naturally, the next questions are:
(1) What can we say about the eigenvalues and eigenfunctions of a Sturm-Liouville problem?
(2) If the eigenvalues and eigenfunctions are found, would they help us express solutions of the original heat conduction problem?

Recall that in the simple case

$$X'' + \lambda X = 0, \qquad X(0) = X(1) = 0,$$

which is encountered in the process of solving the heat equation $u_t = \alpha^2 u_{xx}$ in a rod of length 1 with the zero boundary conditions , the eigenvalues are $(n\pi)^2$ and the eigenfunctions are $X_n(x) = \sin n\pi x$. Then we considered the superposition

$$u(x,t) = \sum_{n=1}^{\infty} c_n X_n(x) T_n(t),$$

and used the initial condition $u(x,0) = f(x)$ to determine the coefficients c_n. That the sum $\sum_{n=1}^{\infty} c_n X_n(x) T_n(0)$ equals $f(x)$ at all continuous points is a result of the fact that the series expansion of $f(x)$ is a sine series and the Fourier convergence theorem.

Now in the more general setting, even if we could find all eigenfunctions of a Sturm-Liouville problem, say $X_n(x)$, we are not sure whether the coefficients in the sum $\sum_{n=1}^{\infty} c_n X_n(x) T_n(t)$ can be determined by the initial condition, nor are we sure that, if c_n's can be found, the sum $\sum_{n=1}^{\infty} c_n X_n(x) T_n(0)$ converges to $f(x)$ nicely (i.e., an analogue of the Fourier convergence theorem).

It turns out that the crucial properties of the simple case are shared by the more general Sturm-Liouville problems. They are the *orthogonality between the eigenfunctions* $y_n(x)$ and *the convergence theorem of expanding a "nice" function* $f(x)$ *as a* $y_n(x)$*-series.* These results can be summarized in five theorems. Before stating them, let us fix some definitions and notations.

1. Definitions and Notations

(SL): The Sturm-Liouville boundary value problem

$$\begin{cases} [p(x)y']' - q(x)y + \lambda r(x)y = 0, & 0 < x < 1, \\ \alpha_1 y(0) + \alpha_2 y'(0) = \beta_1 y(1) + \beta_2 y'(1) = 0. \end{cases}$$

We also assume that $p(x), r(x) > 0$ for all $0 < x < 1$; $(\alpha_1, \alpha_2), (\beta_1, \beta_2) \neq (0, 0)$.

Date: 11/14/16.

$\underline{L[y]}$: We define the linear operator $L[y]$ as

$$L[y] = -[p(x)y']' + q(x)y,$$

so that the equation in (SL) can be written as

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

$\underline{\langle \cdot, \cdot \rangle_H}$: The Hermitian L^2-inner product of two complex-valued functions defined on the interval $[0, 1]$, defined as

$$\langle u, v \rangle_H := \int_0^1 u(x)\overline{v(x)}dx,$$

where the over-bar denotes the complex conjugate: $\overline{a + ib} = a - ib$.

$\underline{\langle \cdot, \cdot \rangle_{r(x)}}$: For $r(x) > 0$ on $[0, 1]$, we define the L^2-inner product of two real functions $u(x), v(x)$ on $[0, 1]$, with weight function $r(x)$, as

$$\langle u, v \rangle_{r(x)} = \int_0^1 r(x)u(x)v(x)dx.$$

Functions orthonormal under $\langle \cdot, \cdot \rangle_{r(x)}$: A sequence of functions $\phi_1(x), \phi_2(x), ...$ is said to be *orthonormal* under the inner product $\langle \cdot, \cdot \rangle_{r(x)}$ if

$$\langle \phi_n(x), \phi_m(x) \rangle_{r(x)} = 0, \qquad n \neq m,$$

and

$$\langle \phi_n(x), \phi_n(x) \rangle_{r(x)} = 1, \qquad n = 1, 2, 3, ...$$

In particular, a list of functions satisfying the first condition above is said to be *orthogonal.* Functions satisfying the second condition are said to be *normalized.*

Exercise. Given function $u(x)$ defined on $[0, 1]$, show that the *normalization* of $u(x)$ in the sense of the inner product $\langle \cdot, \cdot \rangle_{r(x)}$ is

$$(\langle u(x), u(x) \rangle_{r(x)})^{-\frac{1}{2}} u(x).$$

Hint: Use the bilinearity of the inner product.

2. Theorems

Theorem 1. If $u(x), v(x)$ both satisfy the boundary condition in (SL), then

$$\langle L[u], v \rangle_H = \langle u, L[v] \rangle_H.$$

In other words, the linear operator L is self-adjoint, restricted to the vector space of functions satisfying the BV of (SL).

Theorem 2. All eigenvalues of (SL) are real.

Theorem 3. If ϕ_m, ϕ_n are eigenfunctions of L with eigenvalues λ_m, λ_n, i.e.,

$$L[\phi_m] = \lambda_m r(x)\phi_m, \qquad L[\phi_n] = \lambda_n r(x)\phi_n,$$

then

$$\langle \phi_m, \phi_n \rangle_{r(x)} = 0.$$

Theorem 4. All eigenvalues of (SL) form an infinite sequence:

$$\lambda_1 < \lambda_2 < \lambda_3 < ... < \lambda_n ...$$

with the property

$$\lim_{n\to\infty} \lambda_n = \infty.$$

Moreover, the eigenspace of each eigenvalue is 1-dimensional.

Theorem 5. Given $f(x)$ with f, f' piecewise continuous on $[0,1]$. Let $\psi_1, \psi_2, ...$ be all the *normalized* eigenfunctions of (SL), then the sum

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

converges to $\frac{f(x^+)+f(x^-)}{2}$ on $(0,1)$, where c_n's are calculated by

$$c_n = \langle f(x), \psi_n \rangle_{r(x)}.$$

3. Examples

In this section, we give two examples to help us understand what the five theorems are saying.

Example 1. In (SL), take $p(x) = r(x) = 1,\ q(x) = 0$, and $\alpha_1 = \beta_1 = 1, \alpha_2 = \beta_2 = 0$. We thus have the familiar two-point boundary value problem:

$$\begin{cases} y'' + \lambda y = 0. \\ y(0) = y(1) = 0. \end{cases}$$

Previously, we only checked for the real λ's and found the eigenvalues to be $(n\pi)^2$ and the eigenfunctions $\sin n\pi$. This matches the result of Theorem 4 about the distribution of eigenvalues and the dimension of each eigenspace.

Furthermore, since we have set $r(x) = 1$, the inner product $\langle \cdot, \cdot \rangle_{r(x)}$ is just the usual L^2-inner product. We know that $\{\sin n\pi x\}_{n=1}^{\infty}$ form an orthogonal set in the sense of L^2-inner product on $[0,1]$. Thus, the conclusion of Theorem 3 is checked.

Now we normalize $\sin n\pi x$. Note that

$$\langle \sin n\pi x, \sin n\pi x \rangle_{r(x)} = \int_0^1 \sin^2 n\pi x dx = \int_0^1 \frac{1-\cos 2n\pi x}{2} dx = \frac{1}{2}.$$

Therefore, the normalized eigenfunctions are

$$\psi_n(x) = \sqrt{2} \sin n\pi x.$$

According to Theorem 5, given a function $f(x)$, we could calculate the coefficients c_n:

$$c_n = \langle f(x), \psi_n(x) \rangle_{r(x)} = \int_0^1 \sqrt{2} \sin n\pi x f(x) dx.$$

Therefore,

$$\sum_{n=1}^{\infty} c_n \psi_n(x) = \sum_{n=1}^{\infty} \Big(2\int_0^1 \sin n\pi f(x)dx\Big) \sin n\pi x$$

converges to $f(x)$ at all continuous points of f. You may realize that this conclusion is identical to the one given by the Fourier convergence theorem and that $2\int_0^1 \sin n\pi f(x)dx$ are exactly the Fourier coefficients when we expand $f(x)$ as a sine series.

You may ask now, what about Theorem 1 and 2? In fact, theorem 1, once proven, leads to theorem 2, which further leads to theorem 3. We will see the proofs in the next section. Of course, it is a good exercise to try to prove Theorem 1 in the setting of Example 1.

Exercise. As in Example 1, verify the conclusions of Theorem 3-5 for the following choice of functions/constants in (SL):

$$p(x) = r(x) = 1,\ q(x) = 0. \qquad \alpha_1 = \beta_1 = 0,\ \alpha_2 = \beta_2 = 1.$$

You may realize that this two-point boundary value problem arises in the heat conduction problem when both ends are insulated.

Example 2. In (SL), set $p(x) = r(x) = 1,\ q(x) = 0$. Also set $\alpha_1 = 1, \alpha_2 = 0,\ \beta_1 = \beta_2 = 1$. We thus have the two-point boundary value problem, which we studied as an example in Lecture 23:

$$\begin{cases} y'' + \lambda y = 0, \\ y(0) = 0,\ y(1) + y'(1) = 0. \end{cases}$$

Previously, we found all the eigenvalues of this problem to be $\lambda_n = \mu_n^2$ where μ_n are all the positive values which satisfy

$$\mu_n = -\tan \mu_n.$$

And the eigenfunctions are

$$y_n(x) = \sin \mu_n x.$$

This verifies Theorem 4.

Theorem 3 says that all the $y_n(x)$'s are mutually orthogonal under $\langle \cdot, \cdot \rangle_{r(x)}$, which is just the L^2-inner product since $r(s) = 1$. Let us see if this is true by direct calculation. By the definition of L^2-inner product, we have

$$\begin{aligned} \langle y_n(x), y_m(x) \rangle &= \int_0^1 \sin \mu_n x \sin \mu_m x dx \\ &= \int_0^1 \frac{1}{2}[\cos(\mu_n - \mu_m)x - \cos(\mu_n + \mu_m)x]dx \\ &= \frac{1}{2}\Big[\frac{1}{\mu_n - \mu_m}\sin(\mu_n - \mu_m) - \frac{1}{\mu_n + \mu_m}\sin(\mu_n + \mu_m)\Big] \\ &= \ldots \end{aligned}$$

We could have continued the calculation and by using $\mu_i = -\tan \mu_i$ $(i = m, n)$, obtaining $\langle y_n(x), y_m(x) \rangle = 0$. However, the theorem exempts us from carrying out all the calculations. As you will see later, the orthogonality result follows just from simple linear algebra.

As mentioned before, Theorem 5 enables us to expand a function $f(x)$ as a y_n-series. First, we normalize $y_n(x)$. Note that

$$\begin{aligned}\langle y_n(x), y_n(x)\rangle_{r(x)} &= \int_0^1 \sin^2 \mu_n x dx \\ &= \int_0^1 \frac{1-\cos 2\mu_n x}{2} dx \\ &= \frac{1}{2} - \frac{1}{4\mu_n}\sin 2\mu_n.\end{aligned}$$

Therefore, the normalized eigenfunctions are

$$\psi_n(x) = \frac{\sin \mu_n x}{\left(\frac{1}{2} - \frac{1}{4\mu_n}\sin 2\mu_n\right)^{\frac{1}{2}}}.$$

If we denote $\left(\frac{1}{2} - \frac{1}{4\mu_n}\sin 2\mu_n\right)^{-\frac{1}{2}}$ as k_n, then

$$c_n = \langle f(x), \psi_n(x)\rangle_{r(x)} = \int_0^1 f(x) k_n y_n(x) dx = k_n \int_0^1 f(x) \sin \mu_n x dx,$$

and formally,

$$f(x) = \sum_{n=1}^{\infty} c_n \psi_n(x) = \sum_{n=1}^{\infty} \left(k_n^2 \int_0^1 f(x)\sin \mu_n x dx\right) \sin \mu_n x.$$

4. Proofs of Theorems 1-3

In this section, we prove Theorem 1-3 in section 2. Proofs of Theorems 4 and 5 is beyond our current scope.

Proof of Theorem 1. The proof is pure calculation, in which we are going to use integration by parts twice. Given u, v satisfying the boundary conditions of (SL), we have

$$\begin{aligned}\langle L[u], v\rangle_H &= \int_0^1 (-[p(x)u']' + q(x)u)\overline{v} dx \\ &= -\int_0^1 \overline{v} d(p(x)u') + \int_0^1 q(x)u\overline{v} dx \\ &= -\overline{v}p(x)u'\Big|_0^1 + \int_0^1 \overline{v}' p(x) u' dx + \int_0^1 q(x) u\overline{v} dx \\ &= -\overline{v}p(x)u'\Big|_0^1 + \int_0^1 \overline{v}' p(x) du(x) + \int_0^1 q(x) u\overline{v} dx \\ &= -\overline{v}p(x)u'\Big|_0^1 + \overline{v}'p(x)u\Big|_0^1 - \int_0^1 [p(x)\overline{v}']' u dx + \int_0^1 q(x)u\overline{v} dx \\ &= -\overline{v}p(x)u'\Big|_0^1 + \overline{v}'p(x)u\Big|_0^1 + \langle u, L[v]\rangle_H.\end{aligned}$$

Thus, to prove Theorem 1, it suffices to show that

$$-\overline{v}p(x)u'\Big|_0^1 + \overline{v}'p(x)u\Big|_0^1 = 0,$$

which is just

$$p(x)[-\overline{v}u'(1) + \overline{v}u'(0) + \overline{v}'u(1) - \overline{v}'u(0)] = 0.$$

Note that

$$-\overline{v}u'(1) + \overline{v}u'(0) + \overline{v}'u(1) - \overline{v}'u(0) = \det\begin{pmatrix} \overline{v}' & \overline{v} \\ u' & u \end{pmatrix}(1) - \det\begin{pmatrix} \overline{v}' & \overline{v} \\ u' & u \end{pmatrix}(0).$$

On the other hand,

$$\begin{pmatrix} \overline{v}' & \overline{v} \\ u' & u \end{pmatrix}\Big|_{x=0}\begin{pmatrix} \alpha_2 \\ \alpha_1 \end{pmatrix} = 0, \qquad \begin{pmatrix} \overline{v}' & \overline{v} \\ u' & u \end{pmatrix}\Big|_{x=1}\begin{pmatrix} \beta_2 \\ \beta_1 \end{pmatrix} = 0,$$

where $(\alpha_1, \alpha_2), (\beta_1, \beta_2) \neq (0, 0)$. Therefore,

$$\det\begin{pmatrix} \overline{v}' & \overline{v} \\ u' & u \end{pmatrix}(1) = \det\begin{pmatrix} \overline{v}' & \overline{v} \\ u' & u \end{pmatrix}(0) = 0.$$

This completes the proof of Theorem 1.

Proof of Theorem 2. Let λ be an eigenvalue and ϕ an associated eigenfunction. We have

$$\langle L[\phi], \phi\rangle_H = \int_0^1 \lambda r(x)\phi(x)\overline{\phi(x)}dx = \lambda\int_0^1 r(x)|\phi(x)|^2 dx.$$

Also, since $r(x)$ is real-valued,

$$\langle \phi, L[\phi]\rangle_H = \int_0^1 \phi(x)\overline{\lambda r(x)\phi(x)}dx = \overline{\lambda}\int_0^1 r(x)|\phi(x)^2|dx.$$

By Theorem 1, $\langle \mathcal{L}[\phi], \phi\rangle_H = \langle \phi, L[\phi]\rangle_H$. Hence

$$(\lambda - \overline{\lambda})\int_0^1 r(x)|\phi(x)|^2 dx = 0.$$

By $r(x) > 0$ and $\phi(x) \neq 0$, we have $\int_0^1 r(x)|\phi(x)|^2 dx \neq 0$. Therefore, $\lambda - \overline{\lambda} = 0$. In other words, λ is real.

Proof of Theorem 3. Let $\lambda_m \neq \lambda_n$ be eigenvalues. Let ϕ_m, ϕ_n be the corresponding (real-valued) eigenfunctions. We have,

$$\langle L[\phi_m], \phi_n\rangle_H = \langle \lambda_m r(x)\phi_m, \phi_n\rangle_H = \lambda_m\langle r(x)\phi_m, \phi_n\rangle_H = \lambda_m\langle \phi_m, \phi_n\rangle_{r(x)},$$

where we used the fact that for real-valued functions u, v, $\langle r(x)u, v\rangle_H = \langle u, r(x)v\rangle_H = \langle u, v\rangle_{r(x)}$.

Similarly, because λ_n is real,

$$\langle \phi_m, L[\phi_n]\rangle_H = \langle \phi_m, \lambda_n r(x)\phi_n\rangle_H = \lambda_n\langle \phi_m, r(x)\phi_n\rangle_H = \lambda_n\langle \phi_m, \phi_n\rangle_{r(x)}.$$

By Theorem 1,

$$\langle L[\phi_m], \phi_n\rangle_H = \langle \phi_m, L[\phi_n]\rangle_H.$$

Therefore,

$$\lambda_m\langle \phi_m, \phi_n\rangle_{r(x)} = \lambda_n\langle \phi_m, \phi_n\rangle_{r(x)}.$$

Since $\lambda_m \neq \lambda_n$, we must have

$$\langle \phi_m, \phi_n\rangle_{r(x)} = 0.$$

This completes the proof.