Introduction to Sturm-Liouville Problem

Saiyue Lyu
School of Computer Science
University of Waterloo
s6lyu@uwaterloo.ca

Abstract

In 1800's, Jacques Charles Francois Sturm and Joseph Liouville worked on a particular second-order linear differential operator $\mathcal{L} = \frac{d}{dx}\big(p(x)\frac{d}{dx}\big) + \rho(x)$. The Sturm-Liouville Problem was essentially solving eigenvalue problems for the differential operator $\mathcal{L}$. It is still an active area of research as researchers are trying to solve it with weaker and weaker assumptions. The Sturm-Liouville Theory is powerful in applied mathematics as all second-order linear ordinary differential equations can be reduced to this form. This report is a literature review of the Sturm-Liouville Problem.

1 General Linear Second-Order Equations

Consider the ordinary differential equation of second order on the real interval I given by :

$$a(x)y'' + b(x)y' + c(x)y = f(x) \tag{1}$$

where a, b, c, f are complex functions on I.

Definition 1.1. *For equation (1), when $f = 0$ on I, the equation is called* **homogeneous**, *otherwise it is* **nonhomogeneous**. *Any complex function $g \in \mathcal{C}^2(I)$ is a* **solution** *of equation (1) if we substitute y by g, the equation still holds.*

Denote the second-order differential operator $a(x)\frac{d^2}{dx^2} + b(x)\frac{d}{dx} + c(x)$ by $\mathcal{L}$, then equation (1) can be written as $\mathcal{L}y = f$. Note that $\mathcal{L}$ is linear in the sense that $\mathcal{L}(cg_1 + g_2) = c\mathcal{L}g_1 + \mathcal{L}g_2$, $\forall c \in \mathbb{C}, g_1, g_2 \in \mathcal{C}^2(I)$. And this is why equation (1) is a linear differential equation.

If $a(x) \neq 0$ on I, then we can obtain a "monic" equation by dividing $a(x)$ for equation (1):

$$y'' + q(x)y' + \rho(x)y = h(x) \tag{2}$$

where $q = b/a,\ \rho = c/a,\ h = f/a$.

Definition 1.2. *Equation (1) is said to be* **regular** *on I if $a(x)$ does not vanish at any point on I, in which case, equation (1) and equation (2) have the same solution sets. Otherwise, if there exists $x_0 \in I$ such that $a(x_0) = 0$, equation (1) is said to be* **singular** *with a singular point x_0.*

Definition 1.3. *If $q, \rho, h \in \mathcal{C}(I)$ and $x_0 \in I$, $\forall \sigma, \tau \in \mathbb{C}$, there is a unique solution ϕ of equation (2) such that $\phi(x_0) = \sigma$ and $\phi'(x_0) = \tau$, which are called* **initial conditions**. *The system of this condition and equation (2) is called an* **initial value problem**.

Definition 1.4. *When $I = [a, b]$, the boundary conditions have the forms :*

• $y(a) = \sigma,\ y(b) = \tau.$ • $y'(a) = \sigma,\ y'(b) = \tau.$ • $y(c) = \sigma,\ y'(c) = \tau$, *where $c = a$ or b.*

which can be generalized by :

$$\alpha_1 y(a) + \alpha_2 y'(a) + \alpha_3 y(b) + \alpha_4 y'(b) = \sigma, \quad \beta_1 y(b) + \beta_2 y'(b) + \beta_3 y(a) + \beta_4 y'(a) = \tau \tag{3}$$

Submitted to Functional Analysis (PMath 453/653, 2020 Fall).

where $\alpha_i, \beta_i \in \mathbb{C}$ *and* $\sum_{i=1}^4 |\alpha_i| > 0$ *and* $\sum_{i=1}^4 |\beta_i| > 0$. *The system of equation (1) and (3) is called a* **boundary value problem**.

Definition 1.5. *Boundary condition is called :*

• **homogeneous** *if* $\sigma = \tau = 0$.

• **separated** *if* $\alpha_3 = \alpha_4 = \beta_3 = \beta_4 = 0$, *i.e.* $\alpha_1 y(a) + \alpha_2 y'(a) = \sigma$ *and* $\beta_1 y(b) + \beta_2 y'(b) = \tau$.

Definition 1.6. $\forall f, g \in \mathcal{C}^1$, *the* **Wronskian** *of* f *and* g *is the determinant (also denoted as* $W(x)$*):*

$$W(f,g)(x) = \begin{vmatrix} f(x) & g(x) \\ f'(x) & g'(x) \end{vmatrix} = f(x)g'(x) - g(x)f'(x)$$

Theorem 1.7. *If* y_1 *and* y_2 *are solutions of homogeneous equation*

$$y'' + q(x)y' + \rho(x)y = 0,\ x \in I \tag{4}$$

where $q \in \mathcal{C}(I)$, *then either* $W(y_1, y_2)(x) = 0,\ \forall x \in I$, *or* $W(y_1, y_2)(x) \neq 0,\ \forall x \in I$.

Moreover y_1, y_2 *are linearly independent if and only if* $W(y_1, y_2)(x) \neq 0$.

Proof. Since y_1, y_2 are solutions of equation (4),thus

$$y_1'' + qy_1' + \rho y_1 = 0,\ \ y_2'' + qy_2' + \rho y_2 = 0$$

Hence have :

$$\begin{aligned} W' + qW &= y_1 y_2'' - y_2 y_1'' + q(y_1 y_2' - y_2 y_1') \\ &= y_1(y_2'' + qy_2' + \rho y_2) - y_2(y_1'' + qy_1' + \rho y_1) = 0 \end{aligned}$$

Integrating both sides to get :

$$W(x) = c \cdot \exp\Big(-\int_a^x q(t)dt\Big),\ x \in I$$

Thus $W(x) = 0$ if and only if $c = 0$, this proves the first part of the theorem.

For the second part, if y_1, y_2 are linearly dependent, one is a multiple of the other,

Thus $W(x) = 0$ on I.

Conversely, if $W(x) = 0$, thus $W(x) \equiv 0$

Then the vectors (y_1, y_1') and (y_2, y_2') are linearly independent, thus so is y_1, y_2. ■

Definition 1.8. *A function* $f : I \to \mathbb{C}$ *is said to have an* **isolated zero** *at* $x_0 \in I$ *if* $f(x_0) = 0$ *and there exists a neighbor* U *of* x_0 *such that* $f(x) \neq 0, \forall x \in (U - \{x_0\}) \cap I$.

Note that if y is a solution of equation (4) and $y(x_0) = 0$ for some $x_0 \in I$. If $y'(x_0) = 0$, this will forces $y = 0$. If $y'(x_0) \neq 0$, note $y' \in \mathcal{C}(I)$, then $\exists$ neighbor U of x_0 such that $y' \neq 0$ on $U \cap I$, then y is strictly increasing or strictly decreasing. This means the zeros of y are isolated on I.

Theorem 1.9 (Sturm Separation Theorem). *If* y_1, y_2 *are linearly independent solutions of equation (4), then zeros of* y_1 *are distinct from zeros of* y_2. *And there is exactly one zero of* y_1 *between any two consecutive zeros of* y_2, *there is exactly one zero of* y_2 *between any two consecutive zeros of* y_1.

Proof. By Theorem 1.7, $W(y_1, y_2)(x) = y_1 y_2' - y_2 y_1' \neq 0$, then y_1, y_2 cannot have common zeros.

Let a, b be two consecutive zeros of y_2, then

$$\begin{aligned} 0 \neq W(a) = y_1(a)y_2'(a) - y_2(a)y_1'(a) = y_1(a)y_2'(a) \\ 0 \neq W(b) = y_1(b)y_2'(b) - y_2(b)y_1'(b) = y_1(b)y_2'(b) \end{aligned}$$

Since $y_2' \in \mathcal{C}(I)$, $\exists U$ of a and V of b such that sign of y_2' does not change on U and V.

Note a, b are consecutive zeros, so y_2' will have opposite signs on $U \cap I$ and $V \cap I$.

Hence $y_1(a)$ and $y_1(b)$ have opposite signs on I to not change the sign of $W(x)$.

Thus there exists at least one zero of y_1 between a and b.

But if α, β are two zeros of y_1 between a and b

Then using above argument, $\exists$ a zero of y_2 between α, β, which is a contradiction. ■

A contrapositive statement of above theorem will tell us that if y_1, y_2 are two solutions of equation (4) with common zeros in I, then y_1, y_2 are linearly dependent.

Now consider if $y = uv$, then

$$y' = u'v + uv', \quad y'' = u''v + 2u'v' + uv''$$

Then equation (4) becomes :

$$0 = y'' + qy' + \rho y = vu'' + (2v' + qv)u' + (v'' + qv' + \rho v)u$$

Now choose v such that $2v' + qv = 0$, i.e.

$$v(x) = \exp\big(-\frac{1}{2}\int_a^x q(t)dt\big)$$
$$r(x) = v'' + qv' + \rho v = \rho(x) - \frac{1}{4}q^2(x) - \frac{1}{2}q'(x)$$

Note $v \neq 0$, then zeros of u are equivalent to zeros of y, we transform equation (4) to :

$$u'' + ru = 0 \tag{5}$$

Theorem 1.10 (Sturm Comparison Theorem). *Let u and v be such that $u'' + r_1 u = 0$ and $v'' + r_2 v = 0$ over I. If $r_1(x) \geq r_2(x), \forall x \in I$. Then either u has at least one zero between two consecutive zeros of v, or $r_1 \equiv r_2$ and u, v are linearly dependent.*

Proof. Let a, b be any two consecutive zeros of v and assume u has no zeros between a and b.

WLOG, assume $u > 0, v > 0$ on (a, b), also note $v'(a) \geq 0$ and $v'(b) \leq 0$.

Hence consider the Wronskian of u and v :

$$W(a) = u(a)v'(a) \geq 0, \quad W(b) = u(b)v'(b) \leq 0$$
$$W'(x) = u(x)v''(x) - u''(x)v(x) = \Big(r_1(x) - r_2(x)\Big)u(x)v(x) \geq 0, \ \forall x \in (a, b)$$

To make $W'(x) \geq 0$ and $W(a) \geq 0 \geq W(b)$, this forces $r_1 \equiv r_2$, i.e. $W(x) \equiv 0$.

By Theorem 1.7, u and v are linearly dependent. ■

If u is a nontrivial solution of equation (5) on I, if $r(x) \leq 0$, with above theorem, we can conclude that u has at most one zero on I.

2 Differential Equations and Self-Adjoint Differential Operator

We can modify equation (4) as :

$$p(x)y'' + q(x)y' + \rho(x)y = 0 \tag{6}$$

Introduce a linear second-order differential operator :

$$\mathcal{L} = p(x)\frac{d^2}{dx^2} + q(x)\frac{d}{dx} + \rho(x)$$

Equation (6) can be written as :

$$\mathcal{L}y = 0,\ y \in L_2(I) \cap \mathcal{C}^2(I)$$

Recall that if X is an inner product space, the **adjoint** of a linear operator T, if exists, is the operator T' such that $\langle Tx, y\rangle = \langle x, T'y\rangle$, $\forall x, y \in X$. And T is **self-adjoint** if $T = T'$. If X is finite dimensional, then T can be represented as a matrix with respect to the orthonormal basis $\{e_i\}$ and the corresponding matrix of T' is the transpose of the complex conjugate of the matrix of T, i.e. :

$$T = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix},\ T' = \begin{bmatrix} \bar{a}_{11} & \cdots & \bar{a}_{n1} \\ \vdots & \ddots & \vdots \\ \bar{a}_{1n} & \cdots & \bar{a}_{nn} \end{bmatrix} = \bar{T}^\top$$

Proposition 2.1. *If T is self-adjoint, in matrix language, T is a Hermitian matrix, then*

- *The eigenvalues of T are real numbers.*
- *The eigenvectors of T corresponding to distinct eigenvalues are orthogonal.*
- *The eigenvectors of T forms a basis of X.* ■

We would like to study $\mathcal{L}'$, the adjoint of $\mathcal{L}$, where $\langle \mathcal{L}f, g\rangle = \langle f, \mathcal{L}'g\rangle$, $\forall f, g \in L_2(I)$. Have :

$$\begin{aligned}
\langle \mathcal{L}f, g\rangle &= \int_a^b (pf'' + qf' + \rho f)\bar{g}dx \\
&= \Big(pf'\bar{g} - f(p\bar{g})'\Big)\Big|_a^b + \int_a^b f(p\bar{g})''dx + qf\bar{g}\Big|_a^b - \int_a^b f(q\bar{g})'dx + \int_a^b f\rho\bar{g}dx \\
&= \langle f, (\bar{p}g)' - (\bar{q}g)' + \bar{\rho}g\rangle + \Big(p(f'\bar{g} - f\bar{g}') + (q - p')f\bar{g}\Big)\Big|_a^b \\
&= \langle f, \bar{p}g'' + (2\bar{p}' - \bar{q})g' + (\bar{p}'' - \bar{q}' + \bar{\rho})g\rangle + \Big(p(f'\bar{g} - f\bar{g}') + (q - p')f\bar{g}\Big)\Big|_a^b
\end{aligned}$$

where the integral is improper if $I = (a, b)$ is infinite or any of the integrands is unbounded at a or b. And it is well-defined if $p \in \mathcal{C}^2(I), q \in \mathcal{C}^1(I)$ and $\rho \in \mathcal{C}(I)$.

Definition 2.2. *Define the* **formal adjoint** *of $\mathcal{L}$ to be :*

$$\mathcal{L}^* = \bar{p}\frac{d^2}{dx^2} + (2\bar{p}' - \bar{q})\frac{d}{dx} + (\bar{p}'' - \bar{q}' + \bar{\rho})$$

$\mathcal{L}$ is said to be **formally self-adjoint** *if $\mathcal{L}^* = \mathcal{L}$.*

$\mathcal{L}$ is formally self-adjoint if $\mathcal{L}^* = \mathcal{L}$, i.e.

$$\bar{p} = p, \quad 2\bar{p}' - \bar{q} = q, \quad \bar{p}'' - \bar{q} + \bar{\rho} = \rho$$

This forces $p, q, \rho \in \mathbb{R}[x]$ and $p' = q$, thus

$$\mathcal{L}f = pf'' + p'f' + \rho f = (pf')' + \rho f$$

Therefore $\mathcal{L}$ is formally self-adjoint if

$$\mathcal{L} = \frac{d}{dx}(p\frac{d}{dx}) + \rho$$

Then we obtain :

$$\langle \mathcal{L}f, g\rangle = \langle f, \mathcal{L}^*g\rangle + \Big(p(f'\bar{g} - f\bar{g}') + (q - p')f\bar{g}\Big)\Big|_a^b = \langle f, \mathcal{L}g\rangle + \Big(p(f'\bar{g} - f\bar{g}')\Big)\Big|_a^b$$

Hence $\Big(p(f'\bar{g} - f\bar{g}')\Big)\Big|_a^b = 0$, $\forall f, g \in L_2(I)$. This yields the following theorem :

Theorem 2.3. *If $p \in \mathcal{C}^2(I), q \in \mathcal{C}^1(I), \rho \in \mathcal{C}(I)$, then :*

- *If $p, q, \rho \in \mathbb{R}[x]$ and $q = p'$, then $\mathcal{L}$ is formally self-adjoint.*
- *If $\mathcal{L}$ is formally self-adjoint and $\Big(p(f'\bar{g} - f\bar{g}')\Big)\Big|_a^b = 0$, $\forall f, g \in L_2(I)$, then $\mathcal{L}$ is self-adjoint.*

Now consider the eigenvalue problem :

$$\mathcal{L}u + \lambda u = 0 \tag{7}$$

Any $u \neq 0 \in L_2$ satisfying equation (7) for some eigenvalue λ is an eigenfunction of $-\mathcal{L}$.But why we consider $-\mathcal{L}$ instead of $\mathcal{L}$? It can be observed that when $p > 0$, eigenvalues of $\mathcal{L}$ will be negative. Equation (7) leads us one more step closer to the Sturm-Liouville Problem. We still need to do some generalization and we will need some boundary conditions.

Theorem 2.4. *If $\mathcal{L}$ is self-adjoint, then eigenvalues of equation (7) are all real and any pair of eigenfunctions f, g associated with distinct eigenvalues are orthogonal in $L_2(I)$.*

Proof. Assume $\lambda \in \mathbb{C}$ is an eigenvalue of $-\mathcal{L}$, i.e. $\exists\, f \in L_2(I) - \{0\}$ s.t. $\mathcal{L}f + \lambda f = 0$, then :

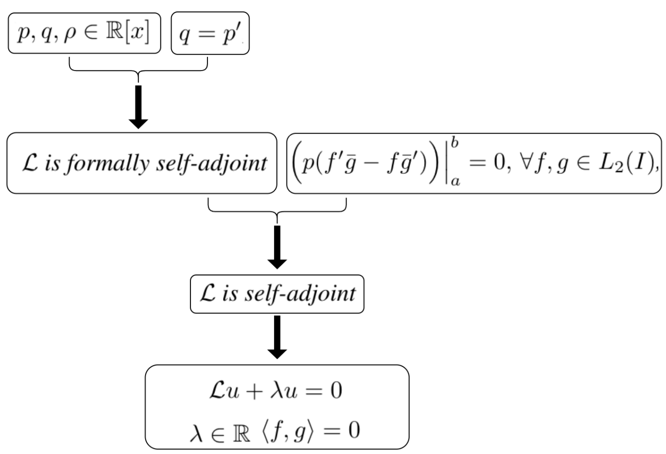
$$\lambda \|f\|^2 = \langle \lambda f, f\rangle = -\langle \mathcal{L}f, f\rangle = -\langle f, \mathcal{L}f\rangle(\text{by self-adjoint}) = \langle f, \lambda f\rangle = \bar{\lambda}\|f\|^2$$

Thus $\lambda = \bar{\lambda}$, i.e. $\lambda \in \mathbb{R}$. And if (μ, g) is another eigenvalue eigenfunction pair, then

$$\lambda\langle f, g\rangle = -\langle \mathcal{L}f, g\rangle = -\langle f, \mathcal{L}g\rangle = \mu\langle f, g\rangle$$

Therefore $\langle f, g\rangle = 0$ ■

Therefore what we have obtained so far for $\mathcal{L}$ is as follows :



But what if $q \neq p'$? WLOG assume $p > 0$, define a new operator via multiplying $\mathcal{L}$ by a function r :

$$\tilde{\mathcal{L}} = r\mathcal{L} = rp\frac{d^2}{dx^2} + rq\frac{d}{dx} + r\rho$$

By Theorem 2.3, $\tilde{\mathcal{L}}$ is formally self-adjoint if

$$rq = (rp)' = r'p + rp'$$

Solving above first-order differential equation, we obtain :

$$r(x) = \frac{c}{p(x)}\exp\Big(\int_a^x \frac{q(t)}{p(t)}dt\Big)$$

Note that $r \in \mathcal{C}^2(I)$ and $r > 0$ on I. When $q = p'$, r will be a constant and $\mathcal{L}$ will be formally self-adjoint. Therefore when $q \neq p'$ and $\mathcal{L}$ is not formally self-adjoint, modify equation (7) to get :

$$r\mathcal{L}u + \lambda ru = \tilde{\mathcal{L}}u + \lambda ru = 0 \tag{8}$$

where $\tilde{\mathcal{L}}$ is formally self-adjoint. Also note with $r > 0$, we have :

$$\Big(rp(f'\bar{g} - f\bar{g}')\Big)\Big|_a^b = 0 \iff \Big(p(f'\bar{g} - f\bar{g}')\Big)\Big|_a^b = 0,\ \forall f, g \in L_2(I)$$

So with above condition, $\tilde{\mathcal{L}}$ is self-adjoint.

Now we may need more definitions for the generalization of eigenvalues and the eigenfunctions.

Definition 2.5. *Let $r \in \mathcal{C}(I)$ be positive, where $I = (a, b)$, the* **inner product** *of two functions $f, g \in \mathcal{C}(I)$* **with respect to the weight function** *r is :*

$$\langle f, g \rangle_r = \int_a^b f(x)\bar{g}(x)r(x)dx$$

This can be verified to be indeed an inner product.

We say f is orthogonal to g with respect to r if $\langle f, g \rangle_r = 0$.

Definition 2.6. *The* **induced norm** *of previous definition is :*

$$\|f\|_r = \Big(\int_a^b |f(x)|^2 r(x)dx \Big)^{1/2}$$

This can be verified to be indeed a norm.

The **corresponding space** *is defined to be :*

$$L_2^r(I) = \{ f : I \to \mathbb{C} \ | \ \|f\|_r < \infty \}$$

This can also be verified to be an inner product space. Note $L_2(I)$ is a special case when $r \equiv 1$.

If $f \in L_2^r(I)$ is an eigenfunction of $\mathcal{L}$ with eigenvalue λ, then

$$\lambda \|f\|_r^2 = \langle \lambda r f, f \rangle = -\langle \tilde{\mathcal{L}} f, f \rangle = -\langle f, \tilde{\mathcal{L}} f \rangle = \langle f, \lambda r f \rangle = \bar{\lambda} \|f\|_r^2$$

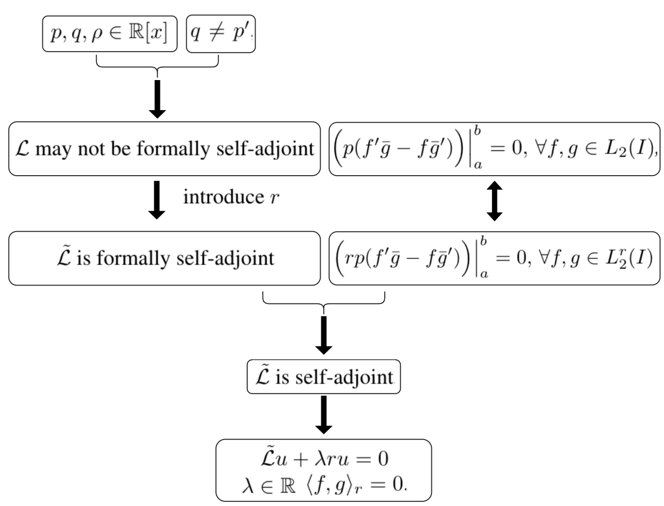
Thus $\lambda \in \mathbb{R}$. And let (μ, g) be another eigenvalue eigenfunction pair, then

$$(\lambda - \mu)\langle f, g \rangle_r = \lambda \langle r f, g \rangle - \mu \langle r f, g \rangle = \langle \lambda r f, g \rangle - \langle f, \mu r g \rangle = \langle -\tilde{\mathcal{L}} f, g \rangle - \langle f, -\tilde{\mathcal{L}} g \rangle = 0$$

Note that $\tilde{\mathcal{L}}$ is self-adjoint, so $\langle f, g \rangle_r = 0$.

Above yields the following theorem, which is a generalization of Theorem 2.4 :

Theorem 2.7. *If $\tilde{\mathcal{L}}$ is self-adjoint and $r \in \mathcal{C}(I)$ is positive on I, then eigenvalues of equation (8) are all real and any pair of eigenfunctions f, g associated with distinct eigenvalues are orthogonal in $L_2^r(I)$. I.e. :*



Note that in fact, eigenvalues and eigenfunctions of equation (8) are the eigenvalues and eigenfuncctions of the operator $-r^{-1}\tilde{\mathcal{L}}$.

When I is finite, then r attains its min α and max β in the sense that $0 < \alpha \le r(x) \le \beta < \infty$. Hence $\sqrt{\alpha}\|f\| \le \|f\|_r \le \sqrt{\beta}\|f\|$. This actually shows that $\|\cdot\|$ and $\|\cdot\|_r$ are equivalent, and thus $L_2(I)$ and $L_2^r(I)$ are the same even induced by different norms.

3 The Regular Sturm-Liouville Problem

Consider the formally self-adjoint second-order linear differential operator :

$$\mathcal{L} = \tfrac{d}{dx}\big(p(x)\tfrac{d}{dx}\big) + \rho(x)$$

The eigenvalue equation :

$$\mathcal{L}u + \lambda r(x)u = 0, \;\; x \in (a,b) \tag{9}$$

subject to the separated homogeneous boundary conditions :

$$\begin{aligned} R_1 u &= \alpha_1 u(a) + \alpha_2 u'(a) = 0, \qquad & |\alpha_1| + |\alpha_2| > 0, \\ R_2 u &= \beta_1 u(b) + \beta_2 u'(b) = 0, \qquad & |\beta_1| + |\beta_2| > 0, \end{aligned}$$

where $\alpha_i, \beta_i \in \mathbb{R}$, is called a **Sturm-Liouville Problem**.

If $I = (a,b)$ is bounded and $p(x) \neq 0$ over I, then it is a **regular Sturm-Liouville Problem**, otherwise it is **singular**. Here we consider the regular Sturm-Liouville Problem and we assume $p(x)$ to be positive on I.

Note that $\mathcal{L}$ is linear and equation (9) is linear in λ. So we could do eigenvalue shifting. Thus WLOG we could assume 0 is not an eigenvalue of $\mathcal{L}$.

Theorem 3.1. *The eigenvalues of $-\mathcal{L}$ are bounded below by a real number.*

Proof. Let $u \in \mathcal{C}^2([a,b])$ be an eigenfunction with boundary conditions and λ be its eigenvalue.

Then $u(a) = u(b) = 0$, have :

$$\begin{aligned} \lambda \|u\|^2 &= \langle -\mathcal{L}u, u\rangle \\ &= \int_a^b \Big(-(pu')'\bar{u} - \rho|u|^2\Big)dx \\ &= \int_a^b p(x)|u'(x)|^2 dx - \int_a^b \rho(x)|u'(x)^2 dx + p(b)\frac{\beta_1}{\beta_2}u^2(b) - p(a)\frac{\alpha_1}{\alpha_2}u^2(a) \\ &= \int_a^b p(x)|u'(x)|^2 dx - \int_a^b \rho(x)|u'(x)^2 dx \\ &\geq -\|u\|^2 \cdot \max_{a\leq x\leq b} |r(x)| \end{aligned}$$

Therefore $\lambda \geq c$, where $c = \max_{a\leq x\leq b} |r(x)| \in \mathbb{R}$.

Also there exists at most two linearly independent eigenfunctions with eigenvalues $< c$.

If so, say u, v, w with eigenvalues $\lambda, \mu, \nu < c$. WLOG assume u, v, w are orthonormal.

With bounday conditions, the six below vectors lie in a 1-dimensional subspace of $\mathbb{R}^2$:

$$\Big(u(a), u'(a)\Big), \Big(u(b), u'(b)\Big), \Big(v(a), v'(a)\Big), \Big(v(b), v'(b)\Big), \Big(w(a), w'(a)\Big), \Big(w(b), w'(b)\Big)$$

Then the below three vectors lie in a 2-dimensional subspace of $\mathbb{R}^4$:

$$\Big(u(a), u'(a), u(b), u'(b)\Big), \Big(v(a), v'(a), v(b), v'(b)\Big), \Big(w(a), w'(a), w(b), w'(b)\Big)$$

Then $\exists s, t, l \in \mathbb{R}$ not all zero such that :

$$s\Big(u(a), u'(a), u(b), u'(b)\Big) + t\Big(v(a), v'(a), v(b), v'(b)\Big) + l\Big(w(a), w'(a), w(b), w'(b)\Big) = 0$$

Therefore $f = su + tv + lw$ is an eigenfunction of $-\mathcal{L}$ such that $f(a) = f(b) = 0$, then :

$$\langle -\mathcal{L}f, f\rangle = \lambda|s|^2 + \mu|t|^2 + \nu|l|^2 < c\big(|s|^2 + |t|^2 + |l|^2\big) = c\,\|f\|^2$$

which is a contradiction. ■

Definition 3.2. **Green's Function** *for the self-adjoint operator* $\mathcal{L} = p\frac{d^2}{dx^2} + p'\frac{d}{dx} + \rho$ *under previous boundary conditions is a function* $G : \mathcal{C}([a,b]^2) \to \mathbb{R}$ *such that*

- *G is symmetric, i.e. $G(x,y) = G(y,x), \forall x, y \in [a,b]$.*
- *G satisfies the boundary conditions for x and y.*
- *$\mathcal{L}_x G(x,y) = 0$, for $x \neq y$.*
- *Derivative $\partial G/\partial x$ has a "jump" discontinuity at $x = y$ by $\frac{\partial G}{\partial x}(y^+, y) - \frac{\partial G}{\partial x}(y^-, y) = \frac{1}{p(y)}$*

Theorem 3.3. *With $\mathcal{L} = \frac{d}{dx}(p\frac{d}{dx}) + \rho$, we can easily verify the* **Lagrange identity** *:*

$$u\mathcal{L}v - v\mathcal{L}u = \Big(p(uv' - vu')\Big)' = \Big(pW\Big)'$$

And this gives the **Green's formula** *:*

$$\int_a^b (u\mathcal{L}v - v\mathcal{L}u)dx = \Big(p(uv' - vu')\Big)\Big|_a^b = \Big(pW\Big)\Big|_a^b$$

Consider $\mathcal{L}u = 0$ has two unique solutions u, v such that

$$\begin{aligned} u(a) &= \alpha_2, & u'(a) &= -\alpha_1 \\ v(b) &= \beta_2, & v'(b) &= -\beta_1 \end{aligned}$$

Note that boundary conditions are satisfied. And u, v must be linearly independent otherwise 0 would be an eigenvalue.

Consider the Wronskian $W(u,v)(x) = u(x)v'(x) - u'(x)v(x)$,which is non-zero over $[a,b]$. And note $p(x) \neq 0$, we define the Green's Function for the regular Sturm-Liouville Problem to be :

$$G(x,y) = \frac{1}{p(x)W(x)} \cdot \begin{cases} u(y)v(x), & a \leq y \leq x \leq b \\ u(x)v(y), & a \leq x \leq y \leq b \end{cases}$$

And the Lagrange identity gives $\Big(p(uv' - u'v)\Big)' = u\mathcal{L}v - v\mathcal{L}u = 0$.

Theorem 3.4. *$G(x,y)$ defined above is indeed a Green's function.*

Proof. The first three bullets are easy to verify, we consider the derivative.

Let $\epsilon > 0$ and differentiate $G(x,y)$, we get :

$$\frac{\partial G}{\partial x}(y, y+\epsilon) - \frac{\partial G}{\partial y}(y, y-\epsilon) = \frac{1}{p(y)W(y)}\Big(u(y)v'(y+\epsilon) - u'(y-\epsilon)v(y)\Big)$$

Then the above expression $\to \frac{1}{p(y)}$ as $\epsilon \to 0$ since $u', v' \in \mathcal{C}(I)$. ∎

Now we define an operator T on $\mathcal{C}(I)$ by :

$$Tf(x) = \int_a^b G(x,y)f(y)dy$$

Theorem 3.5. *The function $Tf \in \mathcal{C}^2(I)$ and solves the differential equation $\mathcal{L}u = f$.*

Proof. Rewrite Tf and differentiate it, we get :

$$\begin{aligned} (Tf)''(x) &= \int_a^x \frac{\partial^2 G(x,y)}{\partial x^2} f(y)dy + \frac{\partial G(x,x^-)}{\partial x} f(x^-) + \int_x^b \frac{\partial^2 G(x,y)}{\partial x^2} f(y)dy - \frac{\partial G(x,x^+)}{\partial x} f(x^+) \\ &= \int_a^x \frac{\partial^2 G(x,y)}{\partial x^2} f(y)dy + \int_x^b \frac{\partial^2 G(x,y)}{\partial x^2} f(y)dy + \frac{f(x)}{p(x)} \end{aligned}$$

Hence $Tf \in \mathcal{C}^2(I)$ and with $\mathcal{L}_x G(x,y) = 0, \forall x \neq y$ we have :

$$\begin{aligned}\mathcal{L}(Tf)(x) &= p(x)(Tf)''(x) + p'(x)(Tf)'(x) + \rho(x)(Tf)(x) \\ &= \int_a^x \mathcal{L}_x G(x,y) f(y) dy + \int_x^b \mathcal{L}_x G(x,y) f(y) dy + f(x) \\ &= f(x)\end{aligned}$$

Therefore Tf is a solution of $\mathcal{L}u = f$. ∎

Note G is symmetric, then Tf satisfies the boundary conditions. Consider the following theorem :

Theorem 3.6. *If $u \in \mathcal{C}^2(I)$ satisfies the boundary condition, then $T(\mathcal{L}u) = u$.*

Proof. Since $p, u, u' \in \mathcal{C}(I)$, we have :

$$\begin{aligned}&T(\mathcal{L}u)(x) \\ =& \int_a^x G(x,y)\mathcal{L}u(y)dy + \int_x^b G(x,y)\mathcal{L}u(y)dy \\ =& \int_a^x u(y)\mathcal{L}_y G(x,y)dy + \int_x^b u(y)\mathcal{L}_y G(x,y)dy \\ &+ p(y)\Big(u'(y)G(x,y) - u(x,y)G_y(x,y)\Big)\Big|_a^b \\ =& p(y)u(y)G_y(x,y)(x,y)\big|_{x^-}^{x^+} + p(y)\Big(u'(y)G(x,y) - u(x,y)G_y(x,y)\Big)\Big|_a^b \\ =& u(x)\end{aligned}$$

From this we can view T as $\mathcal{L}^{-1}$, the inverse of $\mathcal{L}$. ∎

Therefore consider the Sturm-Liouville system :

$$\begin{aligned}\mathcal{L}u + \lambda u = 0 \\ \alpha_1 u(a) + \alpha_2 u'(a) = 0 \\ \beta_1 u(b) + \beta_2 u'(b) = 0\end{aligned}$$

is equivalent to the eigenvalue equation :

$$Tu = \mu u, \quad \text{where } \mu = -\frac{1}{\lambda}$$

So now we can focus on the spectral properties of T.

Definition 3.7. *Let F be an infinite set of continuous function on I. F is* **equicontinuous** *on I if $\forall \epsilon > 0,\ \exists \delta > 0$, which depends on ϵ only, such that*

$$x, y \in I, \quad d(x,y) < \delta \implies d\big(f(x), f(y)\big) < \epsilon, \quad \forall f \in F$$

This actually means $\forall f \in F$, f is uniformly continuous on I, moreover the same δ works for all f.

F is **uniformly bounded** *if there exists $M > 0$ such that $|f(x)| \leq M$, $\forall f \in F$.*

By Ascoli-Arzela Theorem, if F is infinite, equicontinuous and uniformly bounded, and $I = [a,b]$ is bounded, then there is a uniformly convergent sequence $\{f_n\}$ in F with limit f being continuous.

Theorem 3.8. *The set of functions $F = \{Tu : u \in \mathcal{C}([a,b]), \|u\| \leq 1\}$ is equicontinuous and uniformly bounded.*

Proof. Note that Green's function G is continuous on $[a,b]^2$.

Hence $|G(x,y)|$ is uniformly continuous and bounded by some $M > 0$, then

$$|Tu(x)| = |\langle G(x,y), u(y)\rangle| \leq \|G\| \, \|u\| \leq \sqrt{b-a}M \, \|u\| \leq \sqrt{b-a}M$$

Thus F is uniformly bounded.

Since G is uniformly continuous on $[a,b]^2$, then $\forall\epsilon,\ \exists\,\delta>0$ such that

$$x_1,x_2\in[a,b],\ \ d(x_1,x_2)<\delta\implies d\big(G(x_1,y),G(x_2,y)\big)<\epsilon,\ \forall y\in[a,b]$$

Since $u\in\mathcal{C}([a,b])$, then

$$\begin{aligned}d(x_1,x_2)<\delta\implies d\big(Tu(x_1),Tu(x_2)\big)&=|\,\langle G(x_1,y),u(y)\rangle-\langle G(x_2,y),u(y)\rangle\,|\\&<\sqrt{b-a}\,\epsilon\,\|u\|\le\sqrt{b-a}\,\epsilon\end{aligned}$$

Therefore F is equicontinuous. ■

Recall the norm of the operator T, denoted by $\|T\|$:

$$\|T\|=\sup\{\|Tu\|\ :\ u\in\mathcal{C}(I),\|u\|=1\}=\sup_{\|u\|=1}|\,\langle Tu,u\rangle\,|$$

Theorem 3.9. *Either $\|T\|$ or $-\|T\|$ is an eigenvalue of T.*

Proof. Note either $\|T\|=\sup_{\|u\|=1}\langle Tu,u\rangle$ or $\|T\|=-\inf_{\|u\|=1}\langle Tu,u\rangle$.

WLOG assume $\|T\|=\sup_{\|u\|=1}\langle Tu,u\rangle$, proof for the other case would be symmetric.

Then there is a sequence of functions $u_k\in\mathcal{C}(I)$ with norm 1 such that $\langle Tu_k,u_k\rangle\to\|T\|$.

By Ascoli-Arzela Thm, subsequence $\{Tu_{k_i}\}$ uniformly convergent to $\phi_0\in\mathcal{C}(I)$.

As $i\to\infty$, we have :

$$\sup_{x\in[a,b]}|Tu_{k_i}-\phi_0(x)|\to 0\implies\|Tu_{k_i}-\phi_0\|\to 0\implies\|Tu_{k_i}\|\to\|\phi_0\|$$

Let $\mu_0>0$ denote the limit of $\langle Tu_{k_i},u_{k_i}\rangle$, thus :

$$\|Tu_{k_i}-\mu_0u_{k_i}\|^2=\|Tu_{k_i}\|^2+\mu_0^2-2\mu_0\langle Tu_{k_i},u_{k_i}\rangle\to\|\phi_0\|^2-\mu_0^2$$

$$\begin{aligned}\|Tu_{k_i}-\mu_0u_{k_i}\|^2&=\|Tu_{k_i}\|^2+\mu_0^2-2\mu_0\langle Tu_{k_i},u_{k_i}\rangle\\&\le\|T\|^2\,\|u_{k_i}\|^2+\mu_0^2-2\mu_0\langle Tu_{k_i},u_{k_i}\rangle\\&\le 2\mu_0^2-2\mu_0\langle Tu_{k_i},u_{k_i}\rangle\to 2\mu_0^2-2\mu_0^2=0\end{aligned}$$

Thus $\|\phi_0\|>0$ and $\|Tu_{k_i}-\mu_0u_{k_i}\|\to 0$, then :

$$\begin{aligned}&\le\|T\phi_0-\mu_0\phi_0\|\\&\le\|T\phi_0-T(Tu_{k_i})\|+\|T(Tu_{k_i})-\mu_0Tu_{k_i}\|+\|\mu_0Tu_{k_i}-\mu_0\phi_0\|\\&\le\|T\|\,\|\phi_0-Tu_{k_i}\|+\|T\|\,\|Tu_{k_i}-\mu_0u_{k_i}\|+|\mu_0|\cdot\|Tu_{k_i}-\phi_0\|\\&\to\ 0\end{aligned}$$

Therefore $\|T\phi_0-\mu_0\phi_0\|=0$, i.e. $T\phi_0(x)=\mu_0\phi_0(x),\forall x\in[a,b]$.

Thus ϕ_0 is an eigenfunction of T and $\mu_0=\|T\|$ is the corresponding eigenvalue. ■

Theorem 3.10. *T has an infinite sequence of eigenfunctions $\{\psi_n\}$ orthonormal in $L_2(I)$.*

Proof. For any $u\in\mathcal{C}(I)$, following the previous theorem, define :

$$\psi_0=\frac{\phi_0}{\|\phi_0\|},\qquad G_1(x,y)=G(x,y)-\mu_0\psi_0(x)\bar{\psi}_0(y)$$

$$(T_1u)(x)=\int_a^b G_1(x,y)u(y)dy=Tu(x)-\mu_0\langle u,\psi_0\rangle\psi_0(x)$$

Note G_1 and G behave similarly, then Theorem 3.8 and 3.9 can apply to T_1.

If $\|T_1\| \neq 0$, define :

$$|\mu_1| = \sup\{|\langle T_1 u, u\rangle| : u \in \mathcal{C}(I), \|u\| = 1\}$$

By Theorem 3.9, μ_1 is an eigenvalue of T_1 corresponding to eigenfunction $\phi_1 \in \mathcal{C}(I)$, i.e.

$$T_1\phi_1 = \mu_1\phi_1$$

Define $\psi_1 = \frac{\phi_1}{\|\phi_1\|}$, then :

$$\langle T_1 u, \psi_0\rangle = \langle Tu, \psi_0\rangle - \mu_0\langle \langle u, \psi_0\rangle\psi_0, \psi_0 \rangle = \langle u, T\psi_0\rangle - \langle u, \mu_0\psi_0\rangle = 0$$

Note $\langle T_1\psi_1, \psi_0\rangle = \langle \mu_1\psi_1, \psi_0\rangle = 0$, thus ψ_1 is orthogonal to ψ_0. So :

$$T\psi_1 = T_1\psi_1 = \mu_1\psi_1$$

Therefore ψ_1 is an eigenfunction of T with eigenvalue μ_1 s.t. $|\mu_1| = \|T\psi_1\| \leq \|T\| = |\mu_0|$.

Repeat the above procedure, define :

$$G_2(x,y) = G_1(x,y) - \mu_1\psi_1(x)\bar{\psi}_1(y) = G(x,y) - \sum_{k=0}^{1}\mu_k\psi_k(x)\bar{\psi}(y)$$

$$T_2 u = T_1 u - \mu_1\langle u, \psi_1\rangle\psi_1 = Tu - \sum_{k=0}^{1}\mu_k\langle u, \psi_k\rangle\psi_k$$

We have an orthonormal sequence of eigenfunctions $\{\psi_n\}$ with eigenvalues $|\mu_0| \geq |\mu_1| \cdots$.

Note the procedure terminate only when $\|T_n\| = 0$ for some n, in which case :

$$0 = \mathcal{L}T_n u = \mathcal{L}Tu - \sum_{k=0}^{n-1}\mu_k\langle u, \psi_k\rangle\mathcal{L}\psi_k = u - \sum_{k=0}^{n-1}\mu_k\langle u, \psi_k\rangle\mathcal{L}\psi_k$$

$$u = \sum_{k=0}^{n-1}\mu_k\langle u, \psi_k\rangle\mathcal{L}\psi_k = \sum_{k=0}^{n-1}\langle u, \psi_k\rangle\mathcal{L}T\psi_k = \sum_{k=0}^{n-1}\langle u, \psi_k\rangle\psi_k$$

which is a contradiction as no finite set of eigenfunctions can span $\mathcal{C}(I)$.

Therefore the sequence of ψ_n will be infinite. ■

Above proves the existence of the sequence of eigenfunctions, moreover, it is complete. For any $f \in L_2(I)$, by bessel's inequality, $\sum_{k=0}^{\infty}|\langle f, \psi_k\rangle|^2 \leq \|f\|^2$, we need to show the equality holds. The idea is to show $f = \sum_{k=0}^{\infty}\langle f, \psi_k\rangle\psi_k$ for any $f \in \mathcal{C}^2(I)$ and extend this to $L_2(I)$ using the density of $\mathcal{C}^2(I)$.

Theorem 3.11. *For any $f \in \mathcal{C}^2(I)$ satisfying the boundary conditions, the infinite series $\sum\langle f, \psi_k\rangle\psi_k$ converges uniformly to f on I.*

Proof. For any $x \in I$, have : $\langle G(x,\cdot), \psi_k\rangle = T\bar{\psi}_k(x) = \mu_k\bar{\psi}_k(x)$.

Apply Bessel's inequality to $G(\cdot, y)$, have : $\sum_{k=0}^{n}\mu_k^2|\psi_k(x)|^2 \leq \int_a^b |G(x,y)|^2 dy$.

Let $M = \max\{|G(x,y), (x,y) \in [a,b]^2\}$, integrate w.r.t x and let $n \to \infty$, we have :

$$\sum_{k=0}^{\infty}\mu_k^2 \leq (b-a)^2M^2 \implies \lim_{n\to\infty}|\mu_n| = 0$$

For any $u \in \mathcal{I}$, have :

$$\|T_n u\| = \left\|Tu - \sum_{k=0}^{n-1}\mu_k\langle u, \psi_k\rangle\psi_k\right\| \leq |\mu_n|\,\|u\| \to 0$$

For $n > m$, with $|Tu| \leq \sqrt{b-a}M \, \|u\|$, we have :

$$
\begin{aligned}
|\sum_{k=m}^{n} \mu_k \langle u, \psi_k \rangle \psi_k| = |T\Big(\sum_{k=m}^{n} \langle u, \psi_k \rangle \psi_k\Big)| \leq \|T\| \left\| \sum_{k=m}^{n} \langle u, \psi_k \rangle \psi_k \right\| \\
\leq \sqrt{b-a}M \Big(\sum_{k=m}^{n} |\langle u, \psi_k \rangle|^2\Big)^{1/2}
\end{aligned}
$$

By Bessel's inequality, $\Big(\sum_{k=m}^{n} |\langle u, \psi_k \rangle|^2\Big)^{1/2} \to 0$ as $m, n \to \infty$.

Thus $\sum_{k=0}^{\infty} \mu_k \langle u, \psi_k \rangle \psi_k$ converges uniformly on I to a continuous function, therefore :

$$
Tu(x) = \sum_{k=0}^{\infty} \mu_k \langle u, \psi_k \rangle \psi_k(x)
$$

Since f satisfies the boundary conditions, then $u = \mathcal{L}f$ is continuous and $f = Tu$. Hence :

$$
\begin{aligned}
f(x) = Tu(x) &= \sum_{k=0}^{\infty} \mu_k \langle u, \psi_k \rangle \psi_k(x) \\
&= \sum_{k=0}^{\infty} \langle u, \mu_k \psi_k \rangle \psi_k(x) = \sum_{k=0}^{\infty} \langle u, T\psi_k \rangle \psi_k(x) = \sum_{k=0}^{\infty} \langle Tu, \psi_k \rangle \psi_k(x) \\
&= \sum_{k=0}^{\infty} \langle f, \psi_k \rangle \psi_k(x)
\end{aligned}
$$

Therefore the infinite series $\sum \langle f, \psi_k \rangle \psi_k$ converges uniformly to f on I. ■

It can be shown that $\mathcal{C}^2(I)$ is dense in $L_2(I)$ (details omitted in this report). So any function $f \in L_2(I)$ can be approximated in the L_2 norm by a function in $\mathcal{C}^2(I)$ with boundary conditions. This yields the following theorem :

Theorem 3.12. *For any $f \in L_2(I)$ satisfying the boundary conditions, the infinite series $\sum \langle f, \psi_k \rangle \psi_k$ converges uniformly to f on I in L_2 norm, i.e. :*

$$
\lim_{n \to \infty} \left\| f - \sum_{k=0}^{n} \langle f, \psi_k \rangle \psi_k \right\| = 0 \iff f = \sum_{k=0}^{\infty} \langle f, \psi_k \rangle \psi_k \iff \|f\|^2 = \sum_{k=0}^{\infty} |\langle f, \psi_k \rangle|^2
$$

Thus the orthonormal eigenfunction sequence $\{\psi_k\}$ of T forms a complete set in $L_2(I)$.

Back to the regular Sturm-Liouville Problem :

$$
\mathcal{L}u + \lambda u = 0, \; R_1 u = R_2 u = 0
$$

which is equivalent to the single integral equation :

$$
Tu = \mu u = -\frac{1}{\lambda} u
$$

Each eigenvalue λ of $-\mathcal{L}$ corresponds to a unique eigenfunction u, then each eigenvalue μ of T corresponds to a unique eigenfunction u. Note that in the proof of Theorem 3.11, we have :

$$
\lim_{n \to \infty} \frac{1}{|\lambda_n|} = \lim_{n \to \infty} |\mu_n| = 0 \implies \lim_{n \to \infty} \lambda_n = \infty \text{ by Theorem 3.1} \tag{10}
$$

With the weight function r, we have the following fundamental theorem :

Theorem 3.13. *If $p', \rho, r \in \mathcal{C}(I)$ and $p, r > 0$, then the Sturm-Liouville problem has an infinite sequence of real eigenvalues $\lambda_0 < \lambda_1 < \cdots$ such that $\lambda_n \to \infty$.Each eigenvalue λ_n corresponds a unique eigenfunction f_n and the sequence of eigenfunctions $\{f_n\}$ forms an orthonormal basis of $L_2^r(I)$.*

4 The Singular Sturm-Liouville Problem

In the singular Sturm-Liouville Problem, $p(x) = 0$ for some point $x \in [a, b]$. And here we consider the cases that $p(x) = 0$ at $x = a$ and/or $x = b$ or that the interval (a, b) is infinite. Then the expression $rp(f'g - fg')$ vanishes at the endpoint, thus no boundary conditions is required at the endpoint. If (a, b) is infinite, then $\sqrt{r}u(x) \to 0$ as $|x| \to \infty$ to make $u \in \mathcal{L}_2^r$. Theorem 3.13 will remain valid but it will require some more argument. More details involving Lebesgue measure and integration, smoothness conditions can be applied to this problem with weaker conditions. To solve the regular case of the Sturm-Liouville Problem, we may need to use Fourier series for the trigonometric functions. And it allows us to generalize the Sturm-Liouville theory to other orthonormal functions.

References

[1] Vladimir Aleksandrovich Marchenko. *Sturm-Liouville operators and applications*, volume 373. American Mathematical Soc., 2011.

[2] Anton Zettl. *Sturm-liouville theory*. Number 121. American Mathematical Soc., 2010.

[3] Werner O Amrein, Andreas M Hinz, and David B Pearson. *Sturm-Liouville theory: past and present*. Springer Science & Business Media, 2005.

[4] Mohammed A Al-Gwaiz. *Sturm-Liouville theory and its applications*, volume 7. Springer, 2008.

[5] Frederick V Atkinson and Angelo B Mingarelli. *Multiparameter eigenvalue problems: Sturm-Liouville theory*. CRC Press, 2010.