Notes on Sturm-Liouville Differential Equations

Charles Byrne (Charles_Byrne@uml.edu)
Department of Mathematical Sciences
University of Massachusetts at Lowell
Lowell, MA 01854, USA

April 9, 2009

1 Recalling the Wave Equation

The one-dimensional wave equation is

$$\phi_{tt}(x,t) = c^2\phi_{xx}(x,t), \tag{1.1}$$

where $c > 0$ is the propagation speed. Separating variables, we seek a solution of the form $\phi(x,t) = f(t)y(x)$. Inserting this into Equation (1.1), we get

$$f''(t)y(x) = c^2 f(t)y''(x),$$

or

$$f''(t)/f(t) = c^2 y''(x)/y(x) = -\omega^2,$$

where $\omega > 0$ is the separation constant. We then have the separated differential equations

$$f''(t) + \omega^2 f(t) = 0, \tag{1.2}$$

and

$$y''(x) + \frac{\omega^2}{c^2}y(x) = 0. \tag{1.3}$$

The solutions to Equation (1.3) are

$$y(x) = \alpha \sin\Big(\frac{\omega}{c}x\Big).$$

For each arbitrary ω, the corresponding solution of Equation (1.2) is

$$f(t) = \beta\sin(\omega t),$$

or

$$f(t) = \gamma \cos(\omega t).$$

In the vibrating string problem, the string is fixed at both ends, $x = 0$ and $x = L$, so that

$$\phi(0, t) = \phi(L, t) = 0,$$

for all t. Therefore, we must have $y(0) = y(L) = 0$, so that the solutions must have the form

$$y(x) = A_m \sin\Big(\frac{\omega_m}{c} x\Big) = A_m \sin\Big(\frac{\pi m}{L} x\Big),$$

where $\omega_m = \frac{\pi c m}{L}$, for any positive integer m. Therefore, the boundary conditions limit the choices for the separation constant ω. In addition, if the string is not moving at time $t = 0$, then

$$f(t) = \gamma \cos(\omega_m t).$$

We want to focus on Equation (1.3).

Equation (1.3) can be written as

$$y''(x) + \lambda y(x) = 0, \tag{1.4}$$

which is an *eigenvalue problem*. What we have just seen is that the boundary conditions $y(0) = y(L) = 0$ limit the possible values of λ for which there can be solutions: we must have

$$\lambda = \lambda_m = \Big(\frac{\omega_m}{c}\Big)^2 = \Big(\frac{\pi m}{L}\Big)^2,$$

for some positive integer m. The corresponding solutions

$$y_m(x) = \sin\Big(\frac{\pi m}{L} x\Big)$$

are the *eigenfunctions*.

In the vibrating string problem, we typically have the condition $\phi(x, 0) = h(x)$, where $h(x)$ describes the initial position of the string. The problem that remains is to find a linear combination of the eigenfunctions that satisfies this additional initial condition. Therefore, we need to find coefficients A_m so that

$$h(x) = \sum_{m=1}^{\infty} A_m \sin\Big(\frac{\pi m}{L} x\Big). \tag{1.5}$$

Orthogonality will help.

We multiply the equation

$$y_m'' = -\lambda_m y_m$$

by y_n and the equation

$$y_n'' = -\lambda_n y_n$$

by y_m and subtract, to get

$$y_m'' y_n - y_n'' y_m = (\lambda_n - \lambda_m)(y_m y_n).$$

Using

$$y_m'' y_n - y_n'' y_m = (y_n y_m' - y_m y_n')',$$

and integrating, we get

$$0 = y_n(L)y_m'(L) - y_m(L)y_n'(L) - y_n(0)y_m'(0) - y_m(0)y_n'(0) = (\lambda_n - \lambda_m)\int_0^L y_m(x)y_n(x)dx,$$

so that

$$\int_0^L y_m(x)y_n(x)dx = 0,$$

for $m \neq n$. Using this orthogonality of the $y_m(x)$, we can easily find the coefficients A_m.

2 Overview

In what follows we shall study the Sturm-Liouville equations, a class of second-order ordinary differential equations that contains, as a special case, the eigenvalue problem in Equation (1.4). As we shall see, the theory follows closely what we have just discovered about the one-dimensional wave equation. The general form for the Sturm-Liouville Problem is

$$\frac{d}{dx}\Big(p(x)y'(x)\Big) + \lambda w(x)y(x) = 0. \tag{2.1}$$

As with the one-dimensional wave equation, boundary conditions, such as $y(a) = y(b) = 0$, where $a = -\infty$ and $b = +\infty$ are allowed, restrict the possible eigenvalues λ to an increasing sequence of positive numbers λ_m. The corresponding eigenfunctions $y_m(x)$ will be $w(x)$-orthogonal, meaning that

$$0 = \int_a^b y_m(x)y_n(x)w(x)dx,$$

for $m \neq n$. As we shall see later, for various choices of $w(x)$ and $p(x)$ and various choices of a and b, we obtain several famous sets of “orthogonal” functions.

We called the problem

$$y''(x) + \lambda y(x) = 0$$

an eigenvalue problem, which suggests that a theory similar to that for matrices might be possible. This leads to the notion of *self-adjoint* differential operators and helps to motivate the particular form of Sturm-Liouville problems. As we shall see, the pleasant properties of the solutions of the boundary-value problem involving Equation (1.4) stem from the fact that the operator $Ly = y''$ is self-adjoint on functions that are zero at the end points. Many of these properties hold, as well, for solutions to other self-adjoint problems, in particular, to solutions of Sturm-Liouville problems.

3 Self-Adjoint Linear Differential Operators

Separation of variables in partial differential equations often leads to eigenvalue problems associated with linear differential operators. Self-adjoint linear differential operators, which generalize the notion of real symmetric matrices, are a convenient class of operators for which the theory of eigenvalue problems is particularly fruitful.

3.1 Self-Adjoint Matrices

The usual inner product for real (column) vectors u and v is just the dot product, written variously as

$$\langle u, v\rangle = u \cdot v = u^T v.$$

For any real square matrix A and any inner product, the *adjoint* matrix A^* is defined by the property

$$\langle Au, v\rangle = \langle u, A^* v\rangle,$$

for all u and v. Since, for the dot product, we have

$$\langle Au, v\rangle = (Au)^T v = u^T (A^T v) = \langle u, A^T v\rangle,$$

it follows that $A^* = A^T$ for this inner product. Therefore, the matrices that are self-adjoint for the usual inner product are just the symmetric matrices.

If λ_n and λ_m are distinct eigenvalues of a real symmetric matrix A then their corresponding eigenvectors, u_n and u_m, are orthogonal: we have

$$(Au_n)^T u_m = u_n^T A^T u_m = u_n^T (Au_m) = \lambda_m u_n^T u_m,$$

and

$$(Au_n)^T u_m = \lambda_n u_n^T u_m.$$

Since $\lambda_n \neq \lambda_m$, it follows that $u_n^T u_m = 0$.

3.2 Self-Adjoint Operators

We want to extend this idea of being self-adjoint to linear differential operators and inner products of functions.

Given any inner product on functions, written $\langle y, z\rangle$, and any linear operator on these functions, Ly, the adjoint of L is defined by the identity

$$\langle Ly, z\rangle = \langle y, L^* z\rangle,$$

for all functions y and z. The operator L is *self-adjoint* on a certain class of functions if $L^* = L$ for those functions.

3.2.1 The Operator $Dy = y'$

For example, consider the linear differential operator $Dy = \frac{dy}{dx}$. We take for the inner product of two functions $y(x)$ and $z(x)$ the integral

$$\langle y, z\rangle = \int_0^1 y(x)z(x)dx.$$

For functions y and z that are zero at the end points, we have, using integration by parts,

$$\langle Dy, z\rangle = \int_0^1 y'(x)z(x)dx = -\int_0^1 y(x)z'(x)dx = \langle y, D^* z\rangle,$$

from which we conclude that $D^* z = -\frac{dz}{dx}$.

3.2.2 The Operator $Ly = y''$

Now consider the linear differential operator $Ly = \frac{d^2y}{dx^2}$. Using the same inner product, restricting to functions y and z that are zero at the end points, and again using integration by parts, we find that $L^* z = \frac{d^2z}{dx^2} = Lz$; therefore, we say that this operator is *self-adjoint*. Self-adjoint operators generalize real symmetric matrices.

3.2.3 General Second-Order Linear ODE's

We are concerned, in these notes, with second-order linear differential equations with (possibly) non-constant coefficients, that is, differential equations of the form

$$a_2(x)y''(x) + a_1(x)y'(x) + a_0(x)y(x) = 0. \tag{3.1}$$

Now we ask when the linear differential operator

$$Ly = [a_2(x)y'' + a_1(x)y']$$

is self-adjoint. Once again, we consider functions that are zero at end points $x = a$ and $x = b$ and define the inner product of y and z to be

$$\langle y, z\rangle = \int_a^b y(x)z(x)dx.$$

Using integration by parts several times, we find that

$$L^*z = a_2(x)z'' + (2a_2'(x) - a_1(x))z' + (a_2''(x) - a_1'(x))z.$$

Therefore, if it is the case that $a_2'(x) = a_1(x)$, then $L^* = L$ and L is self-adjoint. In this case, we can write Equation (3.1) as

$$(a_2(x)y'(x))' + a_0(x)y(x) = 0,$$

which has the form of the Sturm-Liouville problem,

$$\frac{d}{dx}\Big(p(x)y'(x)\Big) + w(x)y(x) = 0.$$

A similar calculation shows that, for any weight function $w(x) > 0$, the linear differential operator

$$Ty = \frac{1}{w(x)}(p(x)y'(x))'$$

is self-adjoint with respect to the inner product defined by

$$\langle y, z\rangle = \int_a^b y(x)z(x)w(x)dx.$$

Since we can write Equation (2.1) as

$$\frac{1}{w(x)}(p(x)y'(x))' + \lambda y(x) = 0,$$

this tells us that we are dealing with an eigenvalue problem associated with a self-adjoint linear differential operator.

4 Qualitative Analysis of ODE

We are interested in second-order linear differential equations with possibly varying coefficients, as given in equation (3.1), which we can also write as

$$y'' + P(x)y' + Q(x)y = 0. \tag{4.1}$$

Although we can find explicit solutions of Equation (4.1) in special cases, such as

$$y'' + y = 0, \tag{4.2}$$

generally, we will not be able to do this. Instead, we can try to answer certain questions about the behavior of the solution, without actually finding the solution; such an approach is called *qualitative analysis*. The discussion here is based on that in Simmons [1].

4.1 A Simple Example

We know that the solution to Equation (4.2) satisfying $y(0) = 0$, and $y'(0) = 1$ is $y(x) = \sin x$; with $y(0) = 1$ and $y'(0) = 0$, the solution is $y(x) = \cos x$. But, suppose that we did not know these solutions; what could we find out without solving for them?

Suppose that $y(x) = s(x)$ satisfies Equation (4.2), with $s(0) = 0$, $s(\pi) = 0$, and $s'(0) = 1$. As the graph of $s(x)$ leaves the point $(0, 0)$ with x increasing, the slope is initially $s'(0) = 1$, so the graph climbs above the x-axis. But since $y''(x) = -y(x)$, the second derivative is negative for $y(x) > 0$, and becomes increasingly so as $y(x)$ climbs higher; therefore, the derivative is decreasing from $s'(0) = 1$, eventually equaling zero, at say $x = m$, and continuing to become negative. The function $s(x)$ will be zero again at $x = \pi$, and, by symmetry, we have $m = \frac{\pi}{2}$.

Now let $y(x) = c(x)$ solve Equation (4.2), but with $c(0) = 1$, and $c'(0) = 0$. Since $y(x) = s(x)$ satisfies Equation (4.2), so does $y(x) = s'(x)$, with $s'(0) = 1$ and $s''(0) = 0$. Therefore, $c(x) = s'(x)$. Since the derivative of the function $s(x)^2 + c(x)^2$ is zero, this function must be equal to one for all x. In the section that follows, we shall investigate the zeros of solutions.

4.2 Sturm Separation Theorem

Theorem 4.1 *Let $y_1(x)$ and $y_2(x)$ be linearly independent solutions of Equation (4.1). Then their zeros are distinct and occur alternately.*

Proof: Since, for each x, $a = b = 0$ is the only solution of the system

$$ay_1(x) + by_2(x) = 0,$$

$$ay_1'(x) + by_2'(x) = 0,$$

the Wronskian,

$$W(y_1, y_2) = y_1(x)y_2'(x) - y_2(x)y_1'(x),$$

which is the determinant of this two-by-two linear system of equations, can never be zero, so must have constant sign, as x varies. Therefore, the two functions $y_1(x)$ and $y_2(x)$ have no common zero. Suppose that $y_2(x_1) = y_2(x_2) = 0$, with $x_1 < x_2$ successive zeros of $y_2(x)$. Suppose, in addition, that $y_2(x) > 0$ in the interval (x_1, x_2). Therefore, we have $y_2'(x_1) > 0$ and $y_2'(x_2) < 0$. It follows that $y_1(x_1)$ and $y_1(x_2)$ have opposite signs, and there must be a zero between x_1 and x_2. ∎

4.3 From Standard to Normal Form

Equation (4.1) is called the *standard form* of the differential equation. To put the equation into *normal form*, by which we mean an equation of the form

$$u''(x) + q(x)u(x) = 0, \tag{4.3}$$

we write $y(x) = u(x)v(x)$. Inserting this product into Equation (4.1), we obtain

$$vu'' + (2v' + Pv)u + (v'' + Pv' + Qv)u = 0.$$

With

$$v = \exp(-\frac{1}{2}\int Pdx),$$

the coefficient of u' becomes zero. Now we set

$$q(x) = Q(x) - \frac{1}{4}P(x)^2 - \frac{1}{2}P'(x),$$

to get

$$u''(x) + q(x)u(x) = 0.$$

It can be shown that, if $q(x) < 0$ and $u(x)$ satisfies Equation (4.3), then either $u(x) = 0$, for all x, or $u(x)$ has at most one zero. Since we are interested in oscillatory solutions, we restrict $q(x)$ to be (eventually) positive. With $q(x) > 0$ and

$$\int_1^\infty q(x)dx = \infty,$$

the solution $u(x)$ will have infinitely many zeros, but only finitely many on any bounded interval.

4.4 Sturm Comparison Theorem

Solutions to

$$y'' + 4y = 0$$

oscillate faster than solutions of Equation (4.2). This leads to the Sturm Comparison Theorem:

Theorem 4.2 *Let $y'' + q(x)y = 0$ and $z'' + r(x)z = 0$, with $0 < r(x) < q(x)$, for all x. Then between any two zeros of $z(x)$ is a zero of $y(x)$.*

4.4.1 Bessel's Equation

Bessel's Equation is

$$x^2y'' + xy' + (x^2 - p^2)y = 0. \tag{4.4}$$

In normal form, it becomes

$$u'' + \Big(1 + \frac{1-4p^2}{4x^2}\Big)u = 0. \tag{4.5}$$

Information about the zeros of solutions of Bessel's Equation can be obtained by using Sturm's Comparison Theorem and comparing with solutions of Equation (4.2).

5 Sturm-Liouville Equations

The Sturm-Liouville Equations have the form

$$\frac{d}{dx}\Big(p(x)\frac{dy}{dx}\Big) + \lambda w(x)y = 0. \tag{5.1}$$

Here we assume that $p(x) > 0$ and $w(x) > 0$ are continuous, and $p'(x)$ is continuous.

5.1 Special Cases

The problem of the vibrations of a hanging chain leads to the equation

$$\frac{\partial}{\partial x}\Big(gx\frac{\partial y}{\partial x}\Big) = \frac{\partial^2 y}{\partial t^2}, \tag{5.2}$$

and, after separating variables, to

$$\frac{d}{dx}\Big(gx\frac{du}{dx}\Big) + \lambda u = 0. \tag{5.3}$$

The problem of the vibrations of a non-homogeneous string leads to

$$\frac{\partial^2 y}{\partial x^2} = \frac{m(x)}{T}\frac{\partial^2 y}{\partial t^2}, \tag{5.4}$$

and, after separating the variables, to

$$u'' + \lambda m(x)u = 0, \tag{5.5}$$

with $u(0) = u(\pi) = 0$.

5.2 Normal Form

We can put an equation in the Sturm-Liouville form into normal form by first writing it in standard form. There is a better way, though. With the change of variable from x to μ, where

$$\mu(x) = \int_a^x \frac{1}{p(t)} dt,$$

and

$$\mu'(x) = 1/p(x),$$

we can show that

$$\frac{dy}{dx} = \frac{1}{p(x)} \frac{dy}{d\mu}$$

and

$$\frac{d^2y}{dx^2} = \frac{1}{p^2} \frac{d^2y}{d\mu^2} - \frac{p'(x)}{p(x)} \frac{dy}{d\mu}.$$

It follows that

$$\frac{d^2y}{d\mu^2} + q_1(\mu)y = 0. \tag{5.6}$$

For that reason, we study equations of the form

$$y'' + q(x)y = 0. \tag{5.7}$$

6 Analysis of $y'' + q(x)y = 0$

Using the Sturm Comparison Theorem, we have the following lemma.

Lemma 6.1 *Let $y'' + q(x)y = 0$, and $z'' + r(x)z = 0$, with $0 < r(x) < q(x)$. Let $y(b_0) = z(b_0) = 0$ and $z(b_j) = 0$, and $b_j < b_{j+1}$, for $j = 1, 2, ...$ Then, y has at least as many zeros as z in $[b_0, b_n]$. If $y(a_j) = 0$, for $b_0 < a_1 < a_2 < ...$, then $a_n < b_n$.*

Lemma 6.2 *Suppose that $0 < m^2 < q(x) < M^2$ on $[a, b]$, and $y(x)$ solves $y''+q(x)y = 0$ on $[a, b]$. If x_1 and x_2 are successive zeros of $y(x)$ then*

$$\frac{\pi}{M} < x_2 - x_1 < \frac{\pi}{m}.$$

If $y(a) = y(b) = 0$ and $y(x) = 0$ for $n - 1$ other points in (a, b), then

$$\frac{m(b-a)}{\pi} < n < \frac{M(b-a)}{\pi}.$$

Lemma 6.3 *Let y_λ solve*

$$y'' + \lambda q(x)y = 0,$$

with $y_\lambda(a) = 0$, and $y'_\lambda(a) = 1$. Then, there exist $\lambda_1 < \lambda_2 < \ldots$, converging to $+\infty$, such that $y_\lambda(b) = 0$ if and only if $\lambda = \lambda_n$, for some n. The solution y_{λ_n} has exactly $n-1$ roots in (a, b).

6.1 The Original Problem

Returning to the original Sturm-Liouville Equation (5.1), we let $y_n(x)$ be the solution corresponding to λ_n, with $y_n(a) = y_n(b) = 0$; these are the *eigenfunction solutions*. We have the following orthogonality theorem.

Theorem 6.1 *For m not equal to n, we have*

$$\int_a^b y_m(x)y_n(x)w(x)dx = 0. \tag{6.1}$$

Proof: We multiply the equation

$$\frac{d}{dx}\Big(p(x)y'_m(x)\Big) + \lambda_m w(x)y_m(x) = 0$$

by $y_n(x)$, the equation

$$\frac{d}{dx}\Big(p(x)y'_n(x)\Big) + \lambda_n w(x)y_n(x) = 0$$

by $y_m(x)$, and subtract, to get

$$\frac{d}{dx}\Big(p(x)(y_n(x)y'_m(x) - y_m(x)y'_n(x))\Big) = (\lambda_n - \lambda_m)y_m(x)y_n(x)w(x).$$

Integrating, and using the fact that

$$y_m(a) = y_n(a) = y_m(b) = y_n(b) = 0,$$

we get

$$0 = \int_a^b y_m(x)y_n(x)w(x)dx,$$

for $m \neq n$. ∎

We can then associate with (most) functions $h(x)$ on $[a, b]$ an expansion in terms of the eigenfunctions:

$$h(x) = \sum_{n=1}^{\infty} a_n y_n(x), \tag{6.2}$$

with

$$a_n = \int_a^b h(x)y_n(x)w(x)dx \Big/ \int_a^b y_n(x)^2 w(x)dx. \tag{6.3}$$

6.2 Famous Examples

Well known examples of Sturm-Liouville problems include

- **Legendre:**
$$\frac{d}{dx}\Big((1-x^2)\frac{dy}{dx}\Big) + \lambda y = 0;$$
- **Chebyshev:**
$$\frac{d}{dx}\Big(\sqrt{1-x^2}\frac{dy}{dx}\Big) + \lambda(1-x^2)^{-1/2}y = 0;$$
- **Hermite:**
$$\frac{d}{dx}\Big(e^{-x^2}\frac{dy}{dx}\Big) + \lambda e^{-x^2}y = 0;$$
and
- **Laguerre:**
$$\frac{d}{dx}\Big(xe^{-x}\frac{dy}{dx}\Big) + \lambda e^{-x}y = 0.$$

References

[1] Simmons, G. (1972) *Differential Equations, with Applications and Historical Notes*. New York: McGraw-Hill.