MATH 461: Fourier Series and Boundary Value Problems

Chapter V: Sturm–Liouville Eigenvalue Problems

Greg Fasshauer

Department of Applied Mathematics
Illinois Institute of Technology

Fall 2015

Outline

So far all of our problems have involved the two-point BVP

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

which – depending on the boundary conditions – leads to a certain set of eigenvalues and eigenfunctions: e.g.,

1. $\varphi(0) = \varphi(L) = 0$:

$$\lambda_n = \left(\frac{n\pi}{L}\right)^2, \quad \varphi_n(x) = \sin\frac{n\pi x}{L}, \quad n = 1, 2, 3, \ldots$$

2. $\varphi'(0) = \varphi'(L) = 0$:

$$\lambda_n = \left(\frac{n\pi}{L}\right)^2, \quad \varphi_n(x) = \cos\frac{n\pi x}{L}, \quad n = 0, 1, 2, \ldots$$

3. $\varphi(-L) = \varphi(L)$ and $\varphi'(-L) = \varphi'(L)$:

$$\lambda_n = \left(\frac{n\pi}{L}\right)^2, \quad \varphi_n(x) = c_1\cos\frac{n\pi x}{L} + c_2\sin\frac{n\pi x}{L}, \quad n = 0, 1, 2, \ldots$$

Remark

- *The eigenfunctions in the examples on the previous slide were subsequently used to generate*
  1. *Fourier sine series,*
  2. *Fourier cosine series, or*
  3. *Fourier series.*

- *In this chapter we will study problems which involve more general BVPs and then lead to generalized Fourier series.*

Heat Flow in a Nonuniform Rod

Recall the general form of the 1D heat equation:

$$c(x)\rho(x)\frac{\partial u}{\partial t}(x,t) = \frac{\partial}{\partial x}\left(K_0(x)\frac{\partial u}{\partial x}(x,t)\right) + Q(x,t).$$

The separation of variables technique is likely to be applicable if the PDE is linear and homogeneous. Therefore, we assume

$$Q(x,t) = \alpha(x)u(x,t)$$

with x-dependent proportionality factor α.
The resulting PDE

$$c(x)\rho(x)\frac{\partial u}{\partial t}(x,t) = \frac{\partial}{\partial x}\left(K_0(x)\frac{\partial u}{\partial x}(x,t)\right) + \alpha(x)u(x,t) \tag{1}$$

is linear and homogeneous and we will derive the corresponding BVP resulting from separation of variables below.

Remark

Note that

$$\frac{\partial}{\partial x}\left(K_0(x)\frac{\partial u}{\partial x}(x,t)\right) = K_0'(x)\frac{\partial u}{\partial x}(x,t) + K_0(x)\frac{\partial^2 u}{\partial x^2}(x,t).$$

Therefore, a PDE such as

$$c(x)\rho(x)\frac{\partial u}{\partial t}(x,t) = \frac{\partial}{\partial x}\left(K_0(x)\frac{\partial u}{\partial x}(x,t)\right) + \alpha(x)u(x,t)$$

arises, e.g., as convection-diffusion-reaction equation in the modeling of chemical reactions (such as air pollution models) with

convection term: $K_0'(x)\frac{\partial u}{\partial x}(x,t)$

diffusion term: $K_0(x)\frac{\partial^2 u}{\partial x^2}(x,t)$

reaction term: $\alpha(x)u(x,t)$

We now assume $u(x,t) = \varphi(x)T(t)$ and apply separation of variables to

$$c(x)\rho(x)\frac{\partial u}{\partial t}(x,t) = \frac{\partial}{\partial x}\left(K_0(x)\frac{\partial u}{\partial x}(x,t)\right) + \alpha(x)u(x,t).$$

This results in

$$c(x)\rho(x)\varphi(x)T'(t) = \frac{\mathrm{d}}{\mathrm{d}x}\left(K_0(x)\varphi'(x)T(t)\right) + \alpha(x)\varphi(x)T(t).$$

Division by $c(x)\rho(x)\varphi(x)T(t)$ gives

$$\frac{T'(t)}{T(t)} = \frac{1}{c(x)\rho(x)\varphi(x)}\frac{\mathrm{d}}{\mathrm{d}x}\left(K_0(x)\varphi'(x)\right) + \frac{\alpha(x)}{c(x)\rho(x)} = -\lambda.$$

Remark

As always, we choose the minus sign with λ so that the resulting ODE $T'(t) = -\lambda T(t)$ has a decaying solution for positive λ.

From

$$\frac{T'(t)}{T(t)} = \frac{1}{c(x)\rho(x)\varphi(x)} \frac{\mathrm{d}}{\mathrm{d}x}\left(K_0(x)\varphi'(x)\right) + \frac{\alpha(x)}{c(x)\rho(x)} = -\lambda.$$

we see that the resulting ODE for the spatial BVP is

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(K_0(x)\varphi'(x)\right) + \alpha(x)\varphi(x) + \lambda c(x)\rho(x)\varphi(x) = 0$$

and it is in general not known how to solve this ODE eigenvalue problem analytically.

Circularly Symmetric Heat Flow in 2D

The standard 2D-heat equation in polar coordinates is given by

$$\frac{\partial u}{\partial t}(r,\theta,t) = k\nabla^2 u(r,\theta,t),$$

where

$$\nabla^2 u = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial u}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 u}{\partial \theta^2}.$$

If we assume circular symmetry, i.e., no dependence on θ, then $\frac{\partial^2 u}{\partial \theta^2} = 0$ and we have (see also HW 1.5.5)

$$\frac{\partial u}{\partial t}(r,t) = \frac{k}{r}\frac{\partial}{\partial r}\left(r\frac{\partial u}{\partial r}(r,t)\right).$$

We assume $u(r,t) = \varphi(r)T(t)$ and apply separation of variables to

$$\frac{\partial u}{\partial t}(r,t) = \frac{k}{r}\frac{\partial}{\partial r}\left(r\frac{\partial u}{\partial r}(r,t)\right)$$

to get

$$\varphi(r)T'(t) = \frac{k}{r}\frac{\mathrm{d}}{\mathrm{d}r}\left(r\varphi'(r)T(t)\right)$$

or

$$\frac{1}{k}\frac{T'(t)}{T(t)} = \frac{1}{r\varphi(r)}\frac{\mathrm{d}}{\mathrm{d}r}\left(r\varphi'(r)\right) = -\lambda.$$

From

$$\frac{1}{k}\frac{T'(t)}{T(t)} = \frac{1}{r\varphi(r)}\frac{\mathrm{d}}{\mathrm{d}r}\left(r\varphi'(r)\right) = -\lambda$$

we see that the ODE for the spatial (radial) BVP problem is

$$\frac{\mathrm{d}}{\mathrm{d}r}\left(r\varphi'(r)\right) + \lambda r\varphi(r) = 0.$$

Again, we don't yet know how to solve this ODE. Contrary to the previous problem, this equation can be solved using Bessel functions (more later).

In earlier work (see Chapter 2.5) we encountered the steady-state solution of this equation, i.e., Laplace's equation.
Potential BCs therefore are:

- On an annulus, with BCs $u(a,t) = u(b,t) = 0$ or $\varphi(a) = \varphi(b) = 0$.
- On a circular disk, with BCs $u(b,t) = 0$ and $|u(0,t)| < \infty$, i.e., $\varphi(b) = 0$ and $|\varphi(0)| < \infty$.

A general form of an ODE that captures all of the examples discussed so far is the Sturm–Liouville differential equation

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(p(x)\varphi'(x)\right) + q(x)\varphi(x) + \lambda\sigma(x)\varphi(x) = 0$$

with given coefficient functions p, q and σ, and parameter λ.

We now show how this equation covers all of our examples.

Example

- If we let $p(x) = 1$, $q(x) = 0$ and $\sigma(x) = 1$ in

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(p(x)\varphi'(x)\right) + q(x)\varphi(x) + \lambda\sigma(x)\varphi(x) = 0$$

  we get

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

  which led to the standard Fourier series earlier.

Example

- If we let $p(x) = K_0(x)$, $q(x) = \alpha(x)$ and $\sigma(x) = c(x)\rho(x)$ in

$$\frac{d}{dx}\left(p(x)\varphi'(x)\right) + q(x)\varphi(x) + \lambda\sigma(x)\varphi(x) = 0$$

we get

$$\frac{d}{dx}\left(K_0(x)\varphi'(x)\right) + \alpha(x)\varphi(x) + \lambda c(x)\rho(x)\varphi(x) = 0$$

which is the ODE for the heat equation in a nonuniform rod.

Example

- If we let $p(x) = x$, $q(x) = 0$ and $\sigma(x) = x$ in

$$\frac{d}{dx}\left(p(x)\varphi'(x)\right) + q(x)\varphi(x) + \lambda\sigma(x)\varphi(x) = 0$$

and then replace x by r we get

$$\frac{d}{dr}\left(r\varphi'(r)\right) + \lambda r\varphi(r) = 0$$

which is the ODE for the circularly symmetric heat equation.

Example

- If we let $p(x) = T_0$, $q(x) = \alpha(x)$ and $\sigma(x) = \rho_0(x)$ in

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(p(x)\varphi'(x)\right) + q(x)\varphi(x) + \lambda\sigma(x)\varphi(x) = 0$$

  we get

$$T_0\varphi''(x) + \alpha(x)\varphi(x) + \lambda\rho_0(x)\varphi(x) = 0$$

  which is the ODE for vibrations of a nonuniform string (see HW 5.3.1).

Boundary Conditions

A nice summary is provided by the table on p.156 of [Haberman]:

	Heat flow	Vibrating string	Mathematical terminology
$\phi = 0$	Fixed (zero) temperature	Fixed (zero) displacement	First kind or Dirichlet condition
$\frac{d\phi}{dx} = 0$	Insulated	Free	Second kind or Neumann condition
$\frac{d\phi}{dx} = \pm h\phi$ $\left(\begin{array}{l} +\text{left end} \\ -\text{right end} \end{array}\right)$	(Homogeneous) Newton's law of cooling 0° outside temperature, $h = H/K_0$, $h > 0$ (physical)	(Homogeneous) elastic boundary condition $h = k/T_0$, $h > 0$ (physical)	Third kind or Robin condition
$\phi(-L) = \phi(L)$ $\frac{d\phi}{dx}(-L) = \frac{d\phi}{dx}(L)$	Perfect thermal contact	—	Periodicity condition (example of mixed type)
$\lvert\phi(0)\rvert < \infty$	Bounded temperature	—	Singularity condition

Regular Sturm–Liouville Eigenvalue Problems

We will now consider the ODE

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(p(x)\varphi'(x)\right) + q(x)\varphi(x) + \lambda\sigma(x)\varphi(x) = 0, \qquad x \in (a,b) \tag{2}$$

with boundary conditions

$$\begin{aligned} \beta_1\varphi(a) + \beta_2\varphi'(a) &= 0 \\ \beta_3\varphi(b) + \beta_4\varphi'(b) &= 0 \end{aligned} \tag{3}$$

where the β_i are real numbers.

Definition

If p, q, σ and p' in (2) are real-valued and continuous on $[a,b]$ and if $p(x)$ and $\sigma(x)$ are positive for all x in $[a,b]$, then (2) with (3) is called a regular Sturm–Liouville problem.

Remark

Note that the BCs don't capture those of the periodic or singular type.

Facts for Regular Sturm–Liouville Problems

We pick the well-known example

$$\begin{aligned} \varphi''(x) + \lambda\varphi(x) &= 0 \\ \varphi(0) = \varphi(L) &= 0 \end{aligned}$$

with eigenvalues $\lambda_n = \left(\frac{n\pi}{L}\right)^2$ and eigenfunctions $\varphi_n(x) = \sin\frac{n\pi x}{L}$, $n = 1, 2, 3, \ldots$ to illustrate the following facts which hold for **all** regular Sturm–Liouville problems.

Later we will study the properties and prove that they hold in more generality.

1. All eigenvalues of a regular SL problem are real.
   For our example, obviously $\lambda_n = \left(\frac{n\pi}{L}\right)^2$ is real for any value of the integer n.

Remark

This property ensures that when we search for eigenvalues of a regular SL problem it suffices to consider the three cases

$$\lambda > 0, \quad \lambda = 0 \quad \textit{and } \lambda < 0.$$

Complex values of λ are not possible.

We will later prove this fact.

2. Every regular SL problem has infinitely many eigenvalues which can be strictly ordered (after a possible renumbering)

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

For our example, clearly $\lambda_n = \left(\frac{n\pi}{L}\right)^2$ satisfy this property. We have

$$\lambda_1 = \frac{\pi^2}{L^2} \qquad \text{and} \qquad \lambda_n \to \infty \text{ as } n \to \infty.$$

We will not prove this fact.

3. Every eigenvalue λ_n of a regular SL problem has an associated eigenfunction φ_n which is unique up to a constant factor. Moreover, φ_n has exactly $n-1$ zeros in the open interval (a, b).

   For our example, $\lambda_n = \left(\frac{n\pi}{L}\right)^2$ is uniquely associated with $\varphi_n(x) = \sin\frac{n\pi x}{L}$ which has $n-1$ zeros in $(0, L)$.

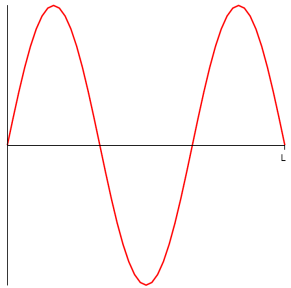


Figure: $\varphi_3(x) = \sin\frac{3\pi x}{L}$ has two zeros in $(0, L)$.

We will later prove the first part of this fact.

4. The set of eigenfunctions, $\{\varphi_n\}_{n=1}^{\infty}$, of a regular SL problem is complete, i.e., any piecewise smooth function f can be represented by a generalized Fourier series (or eigenfunction expansion)

$$f(x) \sim \sum_{n=1}^{\infty} a_n \varphi_n(x)$$

which converges to $\frac{1}{2}\left[f(x+) + f(x-)\right]$ for $a < x < b$.
The generalized Fourier coefficients a_n are addressed below in property 5.
For our example, we have the Fourier sine series

$$f(x) \sim \sum_{n=1}^{\infty} a_n \sin \frac{n\pi x}{L}$$

with the stated convergence properties.
We do not prove the completeness claim.

5. The eigenfunctions associated with different eigenvalues of a regular SL problem are orthogonal on (a, b) with respect to the weight σ, i.e.,

$$\int_a^b \varphi_n(x)\varphi_m(x)\sigma(x)\,\mathrm{d}x = 0 \quad \text{provided } \lambda_n \neq \lambda_m.$$

For our example (where $\sigma(x) = 1$) we have

$$\int_0^L \sin\frac{n\pi x}{L}\sin\frac{m\pi x}{L}\,\mathrm{d}x = \begin{cases} 0 & \text{provided } n \neq m, \\ \frac{L}{2} & \text{if } n = m. \end{cases}$$

These orthogonality relations give us the generalized Fourier coefficients.

We will later prove this fact.

In order to derive the formula for the generalized Fourier coefficients we begin with the generalized Fourier series

$$f(x) \sim \sum_{n=1}^{\infty} a_n \varphi_n(x)$$

Next, we multiply both sides by $\varphi_m(x)\sigma(x)$ and integrate from a to b (assuming we can interchange infinite summation and integration):

$$\int_a^b f(x)\varphi_m(x)\sigma(x)\,\mathrm{d}x = \sum_{n=1}^{\infty} a_n \underbrace{\int_a^b \varphi_n(x)\varphi_m(x)\sigma(x)\,\mathrm{d}x}_{\neq 0 \text{ only if } n=m}$$

Therefore

$$\int_a^b f(x)\varphi_m(x)\sigma(x)\,\mathrm{d}x = a_m \int_a^b \varphi_m^2(x)\sigma(x)\,\mathrm{d}x$$

or

$$a_n = \frac{\int_a^b f(x)\varphi_n(x)\sigma(x)\,\mathrm{d}x}{\int_a^b \varphi_n^2(x)\sigma(x)\,\mathrm{d}x}, \qquad n = 1, 2, 3, \ldots$$

Remark

Note that the formula

$$a_n = \frac{\int_a^b f(x)\varphi_n(x)\sigma(x)\,\mathrm{d}x}{\int_a^b \varphi_n^2(x)\sigma(x)\,\mathrm{d}x}, \qquad n = 1, 2, 3, \ldots$$

for the generalized Fourier coefficients is well-defined, i.e., the denominator

$$\int_a^b \varphi_n^2(x)\sigma(x)\,\mathrm{d}x \neq 0$$

since

- *for a regular SL problem we demanded that* $\sigma(x) > 0$ *on* $[a, b]$
- *and we always have* $\varphi_n^2(x) \geq 0$. *In fact, we know that* $\varphi_n \not\equiv 0$ *due to the properties of its zeros (see fact 3).*

6. The Rayleigh quotient provides a way to express the eigenvalues of a regular SL problem in terms of their associated eigenfunctions:

$$\lambda = \frac{-\,p(x)\varphi(x)\varphi'(x)\big|_a^b + \int_a^b \left(p(x)\left[\varphi'(x)\right]^2 - q(x)\varphi^2(x)\right)\,\mathrm{d}x}{\int_a^b \varphi^2(x)\sigma(x)\,\mathrm{d}x}$$

The Rayleigh quotient is obtained by integrating the SL-ODE by parts.

We will prove this fact in Chapter 5.6.

In our example, we have $p(x) = 1$, $q(x) = 0$, $\sigma(x) = 1$, $a = 0$ and $b = L$, so that

$$\lambda = \frac{-\varphi(x)\varphi'(x)|_0^L + \int_0^L [\varphi'(x)]^2 \, dx}{\int_0^L \varphi^2(x) \, dx}.$$

Using the boundary conditions $\varphi(0) = \varphi(L) = 0$ we get

$$\lambda = \frac{\int_0^L [\varphi'(x)]^2 \, dx}{\int_0^L \varphi^2(x) \, dx}.$$

Remark

Note that this formula gives us information about the relationship between the eigenvalue and eigenfunction – even though in general neither λ nor φ is known.

For example, since

- $\varphi^2(x) \geq 0$,
- $\varphi \not\equiv 0$, and
- $[\varphi'(x)]^2 \geq 0$

we can conclude from the Rayleigh quotient

$$\lambda = \frac{\int_0^L [\varphi'(x)]^2 \, dx}{\int_0^L \varphi^2(x) \, dx},$$

i.e., for our example, that $\lambda \geq 0$

Therefore, the Rayleigh quotient shows – without any detailed calculations – that our example can not have any negative eigenvalues.

Moreover, we can also conclude that $\lambda = 0$ is not possible for our example.
If we had $\lambda = 0$ then the Rayleigh quotient would imply

$$\int_0^L \left[\varphi'(x)\right]^2 \mathrm{d}x = 0$$

or

$$\varphi'(x) = 0 \quad \text{for all } x \text{ in } [0, L] \quad \Longrightarrow \quad \varphi(x) = \text{const.}$$

However, the BCs $\varphi(0) = \varphi(L) = 0$ would then imply $\varphi \equiv 0$, but this is not an eigenfunction, and so $\lambda = 0$ is not an eigenvalue.

As discussed at the beginning of this chapter, the PDE used to model heat flow in a 1D rod without sources (i.e., $Q(x,t) = 0$) is

$$c(x)\rho(x)\frac{\partial u}{\partial t}(x,t) = \frac{\partial}{\partial x}\left(K_0(x)\frac{\partial u}{\partial x}(x,t)\right).$$

We add boundary conditions

$$u(0,t) = 0 \quad \text{and} \quad \frac{\partial u}{\partial x}(L,t) = 0$$

to model fixed temperature zero at the left end and perfect insulation at $x = L$.
The initial temperature is

$$u(x,0) = f(x).$$

Note that this corresponds to (1) studied earlier with $\alpha(x) = 0$, and therefore we will use separation of variables.

The *Ansatz* $u(x,t) = \varphi(x)T(t)$ gives us the two ODEs

$$T'(t) = -\lambda T(t) \tag{4}$$

and

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(K_0(x)\varphi'(x)\right) + \lambda c(x)\rho(x)\varphi(x) = 0. \tag{5}$$

The boundary conditions for (5) are

$$\varphi(0) = 0 \quad \text{and} \quad \varphi'(L) = 0. \tag{6}$$

We also know that solutions of (4) are given by

$$T_n(t) = c_1 \mathrm{e}^{-\lambda_n t},$$

where λ_n, $n = 1, 2, 3, \ldots$, are the eigenvalues of the Sturm–Liouville problem (5)-(6).

Note that the boundary value problem (5)-(6)

$$\frac{d}{dx}\left(K_0(x)\varphi'(x)\right) + \lambda c(x)\rho(x)\varphi(x) = 0$$
$$\varphi(0) = 0 \quad \text{and} \quad \varphi'(L) = 0$$

is indeed a regular Sturm–Liouville problem:

- $p(x) = K_0(x)$, the thermal conductivity, is positive, real-valued and continuous on $[0, L]$,
- $q(x) = 0$, so it is also real-valued and continuous on $[0, L]$,
- $\sigma(x) = c(x)\rho(x)$, the product of specific heat and density, is positive, real-valued and continuous on $[0, L]$, and
- $p'(x) = K_0'(x)$ is real-valued and (hopefully) continuous on $[0, L]$.

Remark

Note that the above assertions are true only for "nice enough" functions K_0, c and ρ.

The boundary value problem

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(K_0(x)\varphi'(x)\right) + \lambda c(x)\rho(x)\varphi(x) = 0$$
$$\varphi(0) = 0 \quad \text{and} \quad \varphi'(L) = 0$$

is quite a bit more complicated than the eigenvalue problems we studied earlier and for general K_0, c and ρ an analytical solution does not exist.

- Instead, we use the properties of regular Sturm–Liouville problems to obtain as much qualitative information about the solution u as possible.
- One could use numerical methods to find approximate eigenvalues and eigenfunctions.

From the general Sturm–Liouville properties we know:

1. & 2. There are infinitely many real eigenvalues satisfying

$$\lambda_1 < \lambda_2 < \lambda_3 < \ldots$$

3. & 4. A complete set of associated eigenfunctions $\{\varphi_n\}_{n=1}^{\infty}$ exists (but in general we don't know the explicit form of φ_n).

5. The eigenfunctions are orthogonal on $[0, L]$ with respect to the weight

$$\sigma(x) = c(x)\rho(x).$$

Therefore, using superposition and $T_n(t) = e^{-\lambda_n t}$, the solution will be of the form

$$u(x,t) = \sum_{n=1}^{\infty} a_n \varphi_n(x) e^{-\lambda_n t}.$$

As we have done before, the generalized Fourier coefficients a_n can be determined using the orthogonality of the eigenfunctions and the initial condition:

$$u(x,0) = \sum_{n=1}^{\infty} a_n \varphi_n(x) \stackrel{!}{=} f(x).$$

Multiplication by $\varphi_m(x)c(x)\rho(x)$ and integration from 0 to L yield:

$$\int_0^L f(x)\varphi_m(x)c(x)\rho(x)\,\mathrm{d}x = \sum_{n=1}^{\infty} a_n \underbrace{\int_0^L \varphi_n(x)\varphi_m(x)c(x)\rho(x)\,\mathrm{d}x}_{=0 \text{ when } n \neq m}$$

Therefore

$$a_n = \frac{\int_0^L f(x)\varphi_n(x)c(x)\rho(x)\,\mathrm{d}x}{\int_0^L \varphi_n^2(x)c(x)\rho(x)\,\mathrm{d}x}, \qquad n = 1, 2, 3, \ldots$$

Qualitative analysis of the solution for large values of t

First, we use the Rayleigh quotient

$$\lambda = \frac{-\, p(x)\varphi(x)\varphi'(x)\big|_a^b + \int_a^b \left(p(x) \left[\varphi'(x)\right]^2 - q(x)\varphi^2(x) \right) \mathrm{d}x}{\int_a^b \varphi^2(x)\sigma(x)\,\mathrm{d}x},$$

which for us – using $[a, b] = [0, L]$, $p(x) = K_0(x)$, $q(x) = 0$, and $\sigma(x) = c(x)\rho(x)$ – becomes

$$\lambda = \frac{-\, K_0(x)\varphi(x)\varphi'(x)\big|_0^L + \int_0^L K_0(x) \left[\varphi'(x)\right]^2 \mathrm{d}x}{\int_0^L \varphi^2(x)c(x)\rho(x)\,\mathrm{d}x},$$

to show that all eigenvalues are positive.

We have

$$\begin{aligned}\lambda &= \frac{-K_0(x)\varphi(x)\varphi'(x)\big|_0^L + \int_0^L K_0(x)\left[\varphi'(x)\right]^2 \mathrm{d}x}{\int_0^L \varphi^2(x)c(x)\rho(x)\,\mathrm{d}x} \\ &= \frac{K_0(0)\underbrace{\varphi(0)}_{=0}\varphi'(0) - K_0(L)\varphi(L)\underbrace{\varphi'(L)}_{=0} + \int_0^L K_0(x)\left[\varphi'(x)\right]^2 \mathrm{d}x}{\int_0^L \varphi^2(x)c(x)\rho(x)\,\mathrm{d}x} \\ &= \frac{\int_0^L K_0(x)\left[\varphi'(x)\right]^2 \mathrm{d}x}{\int_0^L \varphi^2(x)c(x)\rho(x)\,\mathrm{d}x}\end{aligned}$$

Since $K_0(x)$, $c(x)$ and $\rho(x)$ are all positive we have $\lambda \geq 0$.

The only way for $\lambda = 0$ would be to have $\varphi'(x) = 0$.
This, however, is not possible since this would imply $\varphi(x) = \text{const}$ and the BC $\varphi(0) = 0$ would force $\varphi(x) \equiv 0$ (which is not a possible eigenfunction).
Therefore, $\lambda > 0$.

The fact that all eigenvalues $\lambda_n > 0$ implies that the solution

$$u(x,t) = \sum_{n=1}^{\infty} a_n \varphi_n(x) e^{-\lambda_n t}$$

will decay over time.

Moreover, since

- the decay is exponential
- and since λ_n increases with n,

the most significant contribution to the solution for large values of t comes from the first term of the series, i.e.,

$$u(x,t) \approx a_1 \varphi_1(x) e^{-\lambda_1 t}, \qquad t \text{ large}.$$

Since

$$a_1 = \frac{\int_0^L f(x)\varphi_1(x)c(x)\rho(x)\,\mathrm{d}x}{\int_0^L \varphi_1^2(x)c(x)\rho(x)\,\mathrm{d}x}$$

we can conclude that $a_1 \neq 0$ provided $f(x) \geq 0$ (and not identically equal zero).

This is true since

- $c(x) > 0$, $\rho(x) > 0$ and
- either $\varphi_1(x) > 0$ or $\varphi_1(x) < 0$ (recall that SL-property 3 tells us that φ_1 has no zeros in $(0, L)$).

Remark

Thus, the smallest eigenvalue along with its associated eigenfunction provide essential qualitative information about the solution.

We now begin our careful study of the general regular Sturm–Liouville problem

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(p(x)\varphi'(x)\right) + q(x)\varphi(x) + \lambda\sigma(x)\varphi(x) = 0, \qquad x \in (a, b)$$

with boundary conditions

$$\begin{aligned} \beta_1\varphi(a) + \beta_2\varphi'(a) &= 0 \\ \beta_3\varphi(b) + \beta_4\varphi'(b) &= 0 \end{aligned}$$

where the β_i are real numbers, and p, q, σ and p' are real-valued and continuous on $[a, b]$ and $p(x)$ and $\sigma(x)$ are positive for all x in $[a, b]$.

The notation is simplified and our discussion will be more transparent if we use operator notation, i.e., we write the Sturm–Liouville problem as

$$(\mathcal{L}\varphi)(x) + \lambda\sigma(x)\varphi(x) = 0$$

with the SL differential operator $\mathcal{L}$ defined by

$$\mathcal{L}\varphi = \frac{\mathrm{d}}{\mathrm{d}x}\left(p\varphi'\right) + q\varphi.$$

Remark

- *Recall that we used differential operators earlier in our discussion of linearity.*
- *From now on, $\mathcal{L}$ will denote the specific SL operator defined above.*

Lagrange's Identity

For arbitrary functions u and v (with sufficient smoothness) and the SL operator $\mathcal{L}$ defined by

$$\mathcal{L}\varphi = \frac{\mathrm{d}}{\mathrm{d}x}\left(p\varphi'\right) + q\varphi$$

the formula

$$u\mathcal{L}v - v\mathcal{L}u = \frac{\mathrm{d}}{\mathrm{d}x}\left[p\left(uv' - vu'\right)\right]$$

is known as Lagrange's identity.

Remark

This identity will play an important role in the definition of self-adjointness of a linear operator – an important concept analogous to symmetry of a matrix.

We now verify that Lagrange's identity is true:

$$
\begin{aligned}
u\mathcal{L}v - v\mathcal{L}u &\overset{\text{def } \mathcal{L}}{=} u\left[\frac{\mathrm{d}}{\mathrm{d}x}\left(pv'\right) + qv\right] - v\left[\frac{\mathrm{d}}{\mathrm{d}x}\left(pu'\right) + qu\right] \\
&\overset{\text{distribute}}{=} u\frac{\mathrm{d}}{\mathrm{d}x}\left(pv'\right) + uqv - v\frac{\mathrm{d}}{\mathrm{d}x}\left(pu'\right) - vqu \\
&= u\frac{\mathrm{d}}{\mathrm{d}x}\left(pv'\right) - v\frac{\mathrm{d}}{\mathrm{d}x}\left(pu'\right) \\
&\overset{\text{prod rule}}{=} u\left(p'v' + pv''\right) - v\left(p'u' + pu''\right) \\
&\overset{\text{rearrange}}{=} p'\left(uv' - vu'\right) + p\left(uv'' - vu''\right) \\
&= p'\left(uv' - vu'\right) + p\left(u'v' + uv'' - v'u' - vu''\right) \\
&\overset{\text{prod rule}}{=} p'\left(uv' - vu'\right) + p\left(uv' - vu'\right)' \\
&\overset{\text{prod rule}}{=} \frac{\mathrm{d}}{\mathrm{d}x}\left[p\left(uv' - vu'\right)\right]
\end{aligned}
$$

Green's Formula

For arbitrary functions u and v (with sufficient smoothness) and the SL operator $\mathcal{L}$ defined by

$$\mathcal{L}\varphi = \frac{\mathrm{d}}{\mathrm{d}x}\left(p\varphi'\right) + q\varphi$$

the formula

$$\int_a^b \left[u(x)(\mathcal{L}v)(x) - v(x)(\mathcal{L}u)(x)\right] \mathrm{d}x = \left[p(x)\left(u(x)v'(x) - v(x)u'(x)\right)\right]_a^b$$

is known as Green's formula.

Remark

Green's formula is an immediate consequence of Lagrange's identity, i.e., we simply replace the integrand $u\mathcal{L}v - v\mathcal{L}u$ by $\frac{\mathrm{d}}{\mathrm{d}x}\left[p\left(uv' - vu'\right)\right]$. Therefore it may also be called the integral form of Lagrange's identity.

Example

Let's consider the special SL differential operator $\mathcal{L}u = u''$, i.e., with $p(x) = 1$ and $q(x) = 0$, and see what Lagrange's identity and Green's formula look like in this case.

For Lagrange's identity we have the left-hand side

$$u\mathcal{L}v - v\mathcal{L}u = uv'' - vu''$$

and right-hand side

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[p\left(uv' - vu'\right)\right] = \frac{\mathrm{d}}{\mathrm{d}x}\left[\left(uv' - vu'\right)\right].$$

Therefore Lagrange's identity says that

$$uv'' - vu'' = \frac{\mathrm{d}}{\mathrm{d}x}\left[\left(uv' - vu'\right)\right].$$

Example (cont.)

For Green's formula the left-hand side is

$$\int_a^b [u(x)(\mathcal{L}v)(x) - v(x)(\mathcal{L}u)(x)]\,\mathrm{d}x = \int_a^b \left[u(x)v''(x) - v(x)u''(x)\right]\mathrm{d}x$$

and the right-hand side

$$\left[p(x)\left(u(x)v'(x) - v(x)u'(x)\right)\right]_a^b = \left[u(x)v'(x) - v(x)u'(x)\right]_a^b.$$

Therefore Green's formula says that

$$\int_a^b \left[u(x)v''(x) - v(x)u''(x)\right]\,\mathrm{d}x = \left[u(x)v'(x) - v(x)u'(x)\right]_a^b.$$

Self-adjointness

Let's assume we have functions u and v that satisfy the condition

$$\left[p(x)\left(u(x)v'(x) - v(x)u'(x)\right)\right]_a^b = 0,$$

but are otherwise arbitrary.
Then Green's formula

$$\int_a^b [u(x)(\mathcal{L}v)(x) - v(x)(\mathcal{L}u)(x)]\,\mathrm{d}x = \left[p(x)\left(u(x)v'(x) - v(x)u'(x)\right)\right]_a^b$$

tells us that

$$\int_a^b [u(x)(\mathcal{L}v)(x) - v(x)(\mathcal{L}u)(x)]\,\mathrm{d}x = 0.$$

In fact, we will illustrate below that if u and v simply satisfy the same set of (SL-type) boundary conditions, then $\mathcal{L}$ satisfies

$$\int_a^b [u(x)(\mathcal{L}v)(x) - v(x)(\mathcal{L}u)(x)]\,\mathrm{d}x = 0. \tag{7}$$

Remark

With the inner product notation $\langle f, g\rangle = \int_a^b f(x)g(x)\mathrm{d}x$ *we used in Chapter 2 we can write* (7)

$$\int_a^b \left[u(x)(\mathcal{L}v)(x) - v(x)(\mathcal{L}u)(x)\right] \mathrm{d}x = 0$$

as

$$\langle u, \mathcal{L}v\rangle = \langle v, \mathcal{L}u\rangle.$$

This is analogous to the vector-matrix identity $x^T Ay = y^T Ax$ *which is true if* $A = A^T$ *is symmetric.*

Thus, the SL operator behaves in some ways similarly to a symmetric matrix. Since symmetric matrices are sometimes also referred to as self-adjoint matrices, the operator $\mathcal{L}$ *is called a self-adjoint differential operator.*

Example

The general SL differential operator

$$\mathcal{L}\varphi = \frac{\mathrm{d}}{\mathrm{d}x}\left(p\varphi'\right) + q\varphi$$

with boundary conditions $\varphi(0) = \varphi(L) = 0$ is self-adjoint.

To show this we take arbitrary functions u and v that both satisfy the BCs, i.e.,

$$u(0) = u(L) = 0 \quad \text{and} \quad v(0) = v(L) = 0,$$

and we show that (7) holds, i.e.,

$$\int_0^L \left[u(x)(\mathcal{L}v)(x) - v(x)(\mathcal{L}u)(x)\right] \mathrm{d}x = 0.$$

Example (cont.)

$$\begin{aligned}\int_0^L [u(x)(\mathcal{L}v)(x) - v(x)(\mathcal{L}u)(x)]\,\mathrm{d}x &\overset{\text{Green}}{=} \left[p(x)\left(u(x)v'(x) - v(x)u'(x)\right)\right]_0^L \\ &= p(L)\Big(\underbrace{u(L)}_{=0}v'(L) - \underbrace{v(L)}_{=0}u'(L)\Big) \\ &\quad -p(0)\Big(\underbrace{u(0)}_{=0}v'(0) - \underbrace{v(0)}_{=0}u'(0)\Big) \\ &= 0.\end{aligned}$$

Remark

- *In fact, any regular Sturm–Liouville problem is self-adjoint (see HW 5.5.1).*
- *Moreover, the Sturm–Liouville differential equation with periodic or singularity BCs is also self-adjoint (see HW 5.5.1).*

Orthogonality of Eigenfunctions

Earlier we claimed (see Property 5):

For regular SL problems the eigenfunctions to different eigenvalues are orthogonal on (a, b) with respect to the weight σ, i.e.,

$$\int_a^b \varphi_n(x)\varphi_m(x)\sigma(x) = 0 \quad \text{provided } n \neq m,$$

and we illustrated this property with the functions

$$\varphi_n(x) = \sin\frac{n\pi x}{L}, \quad n = 1, 2, 3, \ldots$$

We now prove this is true in general.

We start by considering two different eigenvalues λ_m and λ_n of $\mathcal{L}$. The corresponding SL differential equations are (in operator notation)

$$\mathcal{L}\varphi_m(x) + \lambda_m\sigma(x)\varphi_m(x) = 0 \quad (8)$$
$$\mathcal{L}\varphi_n(x) + \lambda_n\sigma(x)\varphi_n(x) = 0 \quad (9)$$

and the corresponding BCs are

$$\beta_1\varphi_m(a) + \beta_2\varphi_m'(a) = 0$$
$$\beta_3\varphi_m(b) + \beta_4\varphi_m'(b) = 0$$

and (with the same constants β_i)

$$\beta_1\varphi_n(a) + \beta_2\varphi_n'(a) = 0$$
$$\beta_3\varphi_n(b) + \beta_4\varphi_n'(b) = 0$$

We now subtract $\varphi_m(9) - \varphi_n(8)$:

$$\varphi_m \left(\mathcal{L}\varphi_n + \lambda_n \sigma \varphi_n\right) - \varphi_n \left(\mathcal{L}\varphi_m + \lambda_m \sigma \varphi_m\right) = 0$$

or

$$\begin{aligned} \varphi_m \mathcal{L}\varphi_n - \varphi_n \mathcal{L}\varphi_m &= \varphi_n \lambda_m \sigma \varphi_m - \varphi_m \lambda_n \sigma \varphi_n \\ \iff \quad \varphi_m \mathcal{L}\varphi_n - \varphi_n \mathcal{L}\varphi_m &= (\lambda_m - \lambda_n)\sigma \varphi_m \varphi_n \qquad (10) \end{aligned}$$

Green's formula with $u = \varphi_m$ and $v = \varphi_n$ says

$$\int_a^b \left[\varphi_m(x)(\mathcal{L}\varphi_n)(x) - \varphi_n(x)(\mathcal{L}\varphi_m)(x)\right] \mathrm{d}x = \left[p(x)\left(\varphi_m(x)\varphi_n'(x) - \varphi_n(x)\varphi_m'(x)\right)\right]_a^b \qquad (11)$$

and we replace the integrand of (11) with the right-hand side of (10).

This gives us

$$\begin{aligned}\int_a^b (\lambda_m - \lambda_n)\sigma(x)\varphi_m(x)\varphi_n(x)\,\mathrm{d}x &= \left[p(x)\left(\varphi_m(x)\varphi_n'(x) - \varphi_n(x)\varphi_m'(x)\right)\right]_a^b \\ &= p(b)\Big(\underbrace{\varphi_m(b)\varphi_n'(b) - \varphi_n(b)\varphi_m'(b)}_{=B}\Big) \\ &\quad -p(a)\Big(\underbrace{\varphi_m(a)\varphi_n'(a) - \varphi_n(a)\varphi_m'(a)}_{=A}\Big)\end{aligned}$$

Note that the BCs for φ_m and φ_n at $x = b$

$$\begin{aligned}\beta_3\varphi_m(b) + \beta_4\varphi_m'(b) &= 0 \\ \beta_3\varphi_n(b) + \beta_4\varphi_n'(b) &= 0\end{aligned}$$

imply

$$\varphi_m'(b) = -\frac{\beta_3}{\beta_4}\varphi_m(b) \qquad \text{and} \qquad \varphi_n'(b) = -\frac{\beta_3}{\beta_4}\varphi_n(b)$$

Therefore $B = 0$, and $A = 0$ follows similarly.

Since both A and B are zero we have

$$\int_a^b (\lambda_m - \lambda_n)\sigma(x)\varphi_m(x)\varphi_n(x)\,\mathrm{d}x = 0$$

or

$$(\lambda_m - \lambda_n)\int_a^b \sigma(x)\varphi_m(x)\varphi_n(x)\,\mathrm{d}x = 0$$

and therefore – as long as $\lambda_m \neq \lambda_n$ – we have

$$\int_a^b \sigma(x)\varphi_m(x)\varphi_n(x)\,\mathrm{d}x = 0,$$

the claimed orthogonality. □

Remark

Note that Green's formula significantly simplified this proof since we never had to actually evaluate an integral.

All eigenvalues are real

This was the claim of Property 1, and will now prove that it is true for general regular SL problems.

Our strategy will be to assume that there is a complex eigenvalue and show that this leads to a contradiction.

The SL differential equation is

$$\mathcal{L}\varphi + \lambda\sigma\varphi = 0,$$

where φ (as well as λ) is allowed to be complex-valued, but σ and the coefficients p and q in $\mathcal{L}$ are real.
We also assume that the coefficients β_i of the BCs remain real.

The complex conjugate of this equation is

$$\overline{\mathcal{L}\varphi} + \overline{\lambda\sigma\varphi} = 0.$$

Since the coefficients p and q of $\mathcal{L}$ are real we have (see HW 5.5.7) $\overline{\mathcal{L}\varphi} = \mathcal{L}\overline{\varphi}$ and so

$$\mathcal{L}\overline{\varphi} + \overline{\lambda}\sigma\overline{\varphi} = 0,$$

i.e., $\overline{\varphi}$ satisfies the Sturm–Liouville equation (with complex conjugate eigenvalue).

Since the β_i are real, $\overline{\varphi}$ satisfies the BCs whenever φ does: e.g., at $x = a$ we have

$$\begin{aligned} & \beta_1\varphi(a) + \beta_2\varphi'(a) &= 0 \\ \overset{\text{conjugate}}{\iff} \quad & \overline{\beta_1\varphi(a)} + \overline{\beta_2\varphi'(a)} &= 0 \\ \overset{\beta_i \text{ real}}{\iff} \quad & \beta_1\overline{\varphi(a)} + \beta_2\overline{\varphi'(a)} &= 0 \end{aligned}$$

So far we have established that both φ and $\overline{\varphi}$ satisfy the SL problem.

Their associated eigenvalues are λ and $\overline{\lambda}$.

Note that $\lambda \neq \overline{\lambda}$ provided λ is complex as assumed.

Therefore, orthogonality of the associated eigenfunctions gives

$$\int_a^b \varphi(x)\overline{\varphi}(x)\sigma(x)\,dx = 0.$$

However, $\sigma(x) > 0$ and $\varphi(x)\overline{\varphi}(x) = |\varphi(x)|^2 \geq 0$.

Therefore we must have $\varphi(x) \equiv 0$, which contradicts that φ is an eigenfunction, and so λ cannot be complex. $\square$

Uniqueness of Eigenfunctions

This was the claim of Property 3. We now prove this holds for general regular SL problems.

As with any standard uniqueness proof we assume that φ_1 and φ_2 are two different eigenfunctions, both associated with the same eigenvalue λ.

We will show that $\varphi_1 = c\varphi_2$, i.e., we have uniqueness up to a constant factor.

The SL differential equations for φ_1 and φ_2 are

$$\begin{aligned} \mathcal{L}\varphi_1 + \lambda\sigma\varphi_1 &= 0 \qquad (12) \\ \mathcal{L}\varphi_2 + \lambda\sigma\varphi_2 &= 0 \qquad (13) \end{aligned}$$

Multiplying (12) by φ_2 and (13) by φ_1 and taking the difference yields

$$\varphi_2\mathcal{L}\varphi_1 - \varphi_1\mathcal{L}\varphi_2 = 0.$$

Lagrange's identity with $u = \varphi_2$ and $v = \varphi_1$ gives

$$\varphi_2\mathcal{L}\varphi_1 - \varphi_1\mathcal{L}\varphi_2 = \frac{\mathrm{d}}{\mathrm{d}x}\left[p\left(\varphi_2\varphi_1' - \varphi_1\varphi_2'\right)\right] = 0.$$

Consequently,

$$p(x)\left[\varphi_2(x)\varphi_1'(x) - \varphi_1(x)\varphi_2'(x)\right] = \text{const} = C. \qquad (14)$$

To determine the constant C we use the boundary conditions.

We actually discuss three cases:

- Case I: regular SL BCs, i.e., for $i = 1, 2$

$$\begin{aligned} \beta_1 \varphi_i(a) + \beta_2 \varphi_i'(a) &= 0 \\ \beta_3 \varphi_i(b) + \beta_4 \varphi_i'(b) &= 0 \end{aligned}$$

- Case II: singularity-type BCs, e.g., for $i = 1, 2$

$$|\varphi_i(a)| < \infty$$

- Case III: periodic BCs, e.g., for $i = 1, 2$

$$\begin{aligned} \varphi_i(-L) &= \varphi_i(L) \\ \varphi_i'(-L) &= \varphi_i'(L) \end{aligned}$$

- Case I: It actually suffices for the BCs to hold for both φ_1 and φ_2 at only one end, say at $x = a$.
  Then

$$\begin{aligned} & \beta_1\varphi_i(a) + \beta_2\varphi_i'(a) = 0 \\ \iff \quad & \varphi_i'(a) = -\tfrac{\beta_1}{\beta_2}\varphi_i(a), \qquad i = 1, 2 \end{aligned}$$

  This implies that (14) (at $x = a$) becomes

$$\begin{aligned} C &= p(a)\left[\varphi_2(a)\varphi_1'(a) - \varphi_1(a)\varphi_2'(a)\right] \\ &= p(a)\left[\varphi_2(a)\left(-\frac{\beta_1}{\beta_2}\varphi_1(a)\right) - \varphi_1(a)\left(-\frac{\beta_1}{\beta_2}\varphi_2(a)\right)\right] = 0 \end{aligned}$$

Since $C = 0$, equation (14) now reads

$$p(x)\left[\varphi_2(x)\varphi_1'(x) - \varphi_1(x)\varphi_2'(x)\right] = C = 0,$$

and since $p(x) > 0$ for any regular SL problem we have

$$\left[\varphi_2(x)\varphi_1'(x) - \varphi_1(x)\varphi_2'(x)\right] = 0. \tag{15}$$

Now, we note that for $\varphi_2 \not\equiv 0$ (an eigenfunction) the quotient rule gives

$$\frac{d}{dx}\left(\frac{\varphi_1}{\varphi_2}\right) = \frac{\varphi_1'\varphi_2 - \varphi_1\varphi_2'}{\varphi_2^2}.$$

Thus, (15) is equivalent to

$$\frac{d}{dx}\left(\frac{\varphi_1}{\varphi_2}\right) = 0 \qquad \Longrightarrow \qquad \frac{\varphi_1}{\varphi_2} = \text{const} \qquad \text{or} \qquad \varphi_1 = c\varphi_2,$$

which shows that the eigenfunctions of a regular SL problem are unique up to a constant factor.

- Case II: If we have singularity BCs, e.g., $\varphi_1(a)$ and $\varphi_2(a)$ are bounded, then one can also show that (14)

$$p(x)\left[\varphi_2(x)\varphi_1'(x) - \varphi_1(x)\varphi_2'(x)\right] = C$$

implies

$$\varphi_2(x)\varphi_1'(x) - \varphi_1(x)\varphi_2'(x) = 0$$

and it follows that

$$\varphi_1(x) = c\varphi_2(x)$$

just as before. □

- Case III: For periodic BCs the eigenfunctions are in general **not** unique.

  As an example we can consider the SL equation

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

  with BCs

$$\varphi(-L) = \varphi(L) = 0 \quad \text{and} \quad \varphi'(-L) = \varphi'(L) = 0$$

  for which we know that the eigenvalues are $\lambda_n = \left(\frac{n\pi}{L}\right)^2$, $n = 0, 1, 2, \ldots$ and, e.g.,

$$\varphi_1(x) = \cos \pi x \quad \text{or} \quad \varphi_1(x) = \sin \pi x,$$

  both associated with the eigenvalue $\lambda_1 = \left(\frac{\pi}{L}\right)^2$.

Property 6 was about the Rayleigh quotient

$$\lambda = \frac{-p(x)\varphi(x)\varphi'(x)\Big|_a^b + \int_a^b \left(p(x)\left[\varphi'(x)\right]^2 - q(x)\varphi^2(x)\right) \mathrm{d}x}{\int_a^b \varphi^2(x)\sigma(x)\,\mathrm{d}x}$$

which provides a useful relation between the eigenvalue λ and its associated eigenfunction φ that goes beyond the SL differential equation itself.

We will now prove that this relation holds for any regular SL problem.

We start with the SL differential equation

$$\frac{d}{dx}\left(p(x)\varphi'(x)\right) + q(x)\varphi(x) + \lambda\sigma(x)\varphi(x) = 0, \qquad x \in (a, b),$$

multiply by φ and integrate from a to b to get

$$\int_a^b \varphi(x)\left[\frac{d}{dx}\left(p(x)\varphi'(x)\right) + q(x)\varphi(x)\right] dx + \lambda \int_a^b \sigma(x)\varphi^2(x)\,dx = 0.$$

Since φ is an eigenfunction and $\sigma > 0$ we have $\int_a^b \sigma(x)\varphi^2(x)\,dx > 0$.

Therefore

$$\lambda = \frac{-\int_a^b \varphi(x)\left[\frac{d}{dx}\left(p(x)\varphi'(x)\right) + q(x)\varphi(x)\right] dx}{\int_a^b \sigma(x)\varphi^2(x)\,dx},$$

which has the correct denominator.

We now consider the numerator

$$
\begin{aligned}
&-\int_a^b \varphi(x)\left[\frac{\mathrm{d}}{\mathrm{d}x}\left(p(x)\varphi'(x)\right)+q(x)\varphi(x)\right]\mathrm{d}x \\
&= -\int_a^b \underbrace{\varphi(x)}_{\substack{=u \\ \mathrm{d}u=\varphi'(x)\mathrm{d}x}} \underbrace{\frac{\mathrm{d}}{\mathrm{d}x}\left(p(x)\varphi'(x)\right)\mathrm{d}x}_{\substack{=\mathrm{d}v \\ v=p(x)\varphi'(x)}} - \int_a^b q(x)\varphi^2(x)\mathrm{d}x \\
&= -p(x)\varphi(x)\varphi'(x)\Big|_a^b + \int_a^b p(x)\left[\varphi'(x)\right]^2\mathrm{d}x - \int_a^b q(x)\varphi^2(x)\mathrm{d}x
\end{aligned}
$$

Therefore

$$
\lambda = \frac{-p(x)\varphi(x)\varphi'(x)\Big|_a^b + \int_a^b \left(p(x)\left[\varphi'(x)\right]^2 - q(x)\varphi^2(x)\right)\mathrm{d}x}{\int_a^b \varphi^2(x)\sigma(x)\,\mathrm{d}x},
$$

the Rayleigh quotient. □

We now show under what condition a regular SL problem can never have negative eigenvalues.

Theorem

If $-p\varphi\varphi'|_a^b \geq 0$ *and* $q \leq 0$ *then all eigenvalues of a regular SL problem are nonnegative.*

Proof.

We use the Rayleigh quotient

$$\lambda = \frac{-p(x)\varphi(x)\varphi'(x)\big|_a^b + \int_a^b \left(p(x)\left[\varphi'(x)\right]^2 - q(x)\varphi^2(x)\right) dx}{\int_a^b \varphi^2(x)\sigma(x)\, dx}$$

to show that $\lambda \geq 0$.

We look at each of the terms in the Rayleigh quotient separately:

- $-p(x)\varphi(x)\varphi'(x)|_a^b \geq 0$ by assumption,
- $\int_a^b p(x)\left[\varphi'(x)\right]^2 \mathrm{d}x \geq 0$ since $p(x) > 0$ for any regular SL problem and $[\varphi'(x)]^2 \geq 0$,
- $-\int_a^b q(x)\varphi^2(x)\mathrm{d}x \geq 0$ since $q(x) \leq 0$ by assumption and $\varphi^2(x) > 0$,
- $\int_a^b \varphi^2(x)\sigma(x)\mathrm{d}x > 0$ since $\sigma(x) > 0$ for any regular SL problem and $\varphi^2(x) > 0$.

Therefore the Rayleigh quotient is nonnegative □

A Minimization Principle

If we define

$$RQ[u] = \frac{-p(x)u(x)u'(x)\big|_a^b + \int_a^b \left(p(x)\left[u'(x)\right]^2 - q(x)u^2(x)\right) \mathrm{d}x}{\int_a^b u^2(x)\sigma(x)\,\mathrm{d}x},$$

the Rayleigh quotient of u, then we know that all SL eigenvalue-eigenfunction pairs satisfy $\lambda = RQ[\varphi]$.

Theorem

For any regular SL problem the smallest eigenvalue λ_1 is given by

$$\lambda_1 = \min_{\substack{u \in C(a,b) \\ u \text{ satisfies BCs}}} RQ[u].$$

Moreover, the minimum is attained only for $u = \varphi_1$.

Before we prove the theorem, let's think about how it might be useful.

- Recall that we showed earlier that the solution of the heat equation for a nonuniform rod problem for large values of t is characterized mostly by the smallest eigenvalue λ_1 and its associated eigenfunction φ_1. This is typical, and therefore we want to find λ_1.
- Finding the minimum over **all** continuous functions satisfying the BCs is a very challenging – often impossible – task.
- Instead, we choose some continuous trial functions which satisfy the BCs, but need not satisfy the differential equation, and minimize over them.

If u_T is such as trial function, then $RQ[u_T]$ is an upper bound for λ_1 since

$$\lambda_1 = \min_{\substack{u \in C(a,b) \\ u \text{ satisfies BCs}}} RQ[u] \leq RQ[u_T].$$

Ideally, we would like to find a "good" trial function u_T that provides a smallest possible upper bound.

Example

Consider the SL problem

$$\varphi''(x) + \lambda\varphi(x) = 0$$
$$\varphi(0) = \varphi(1) = 0$$

We know that $\lambda_n = \left(\frac{n\pi}{L}\right)^2 = n^2\pi^2$, so

$$\lambda_1 = \pi^2.$$

This will be our benchmark against which we will compare the trial function approach that will give us an approximation to λ_1.
Note that the trial function approach can just as easily be applied to much more complicated problems.

Remark

In fact, many popular numerical methods (such as the Rayleigh-Ritz, or finite element method) are based on such a minimization principle.

Example (cont.)

The minimization principle says

$$\lambda_1 = \min_{\substack{u \in C(a,b) \\ u \text{ satisfies BCs}}} \frac{-p(x)u(x)u'(x)|_a^b + \int_a^b \left(p(x)\left[u'(x)\right]^2 - q(x)u^2(x)\right) \mathrm{d}x}{\int_a^b u^2(x)\sigma(x)\mathrm{d}x}.$$

Here $p(x) = \sigma(x) = 1$, $q(x) = 0$ and $u(0) = u(1) = 0$. So

$$\lambda_1 = \min_{\substack{u \in C(a,b) \\ u \text{ satisfies BCs}}} \frac{\int_a^b \left[u'(x)\right]^2 \mathrm{d}x}{\int_a^b u^2(x)\,\mathrm{d}x}.$$

Instead of minimizing over all continuous functions it will be much easier to just look at

$$\frac{\int_a^b \left[u_T'(x)\right]^2 \mathrm{d}x}{\int_a^b u_T^2(x)\,\mathrm{d}x} \quad (\geq \lambda_1),$$

where u_T is some trial function.

Example (cont.)

Since the minimum is attained for $u_T = \varphi_1$ it is best to use as much information about φ_1 as is available.

We know that

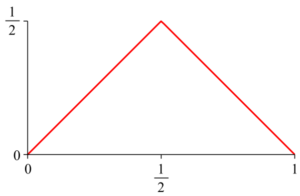
- u_T needs to satisfy the BCs $u_T(0) = u_T(1) = 0$, and
- φ_1 (and therefore u_T) has no zeros in $(0, 1)$.

The simplest trial function with these two properties is the piecewise linear function

$$u_T(x) = \begin{cases} x, & x \le \frac{1}{2} \\ 1 - x, & x \ge \frac{1}{2} \end{cases}$$

with

$$u_T'(x) = \begin{cases} 1, & x < \frac{1}{2} \\ -1, & x > \frac{1}{2} \end{cases}$$



Example (cont.)

With this choice of u_T we get

$$\lambda_1 \le \frac{\int_0^1 \left[u_T'(x)\right]^2 \mathrm{d}x}{\int_0^1 u_T^2(x)\,\mathrm{d}x} = \frac{\int_0^{\frac{1}{2}} 1^2 \mathrm{d}x + \int_{\frac{1}{2}}^1 (-1)^2 \mathrm{d}x}{\int_0^{\frac{1}{2}} x^2 \mathrm{d}x + \int_{\frac{1}{2}}^1 (1-x)^2 \mathrm{d}x}$$

$$= \frac{\frac{1}{2}+\frac{1}{2}}{\left.\frac{x^3}{3}\right|_0^{\frac{1}{2}} - \left.\frac{(1-x)^3}{3}\right|_{\frac{1}{2}}^1} = \frac{1}{\frac{1}{24}+\frac{1}{24}} = 12$$

As a benchmark we know $\lambda_1 = \pi^2 \approx 9.87$.

Remark

Note that a different multiple of u_T such as $u_T(x) = \begin{cases} 2x, & x \le \frac{1}{2} \\ 2-2x, & x \ge \frac{1}{2} \end{cases}$ *would not improve the estimate since eigenfunctions are unique up to a constant multiple only.*

Example (cont.)

We can do better if we take a trial function that better resembles the actual eigenfunction (which, of course, we wouldn't know in general).

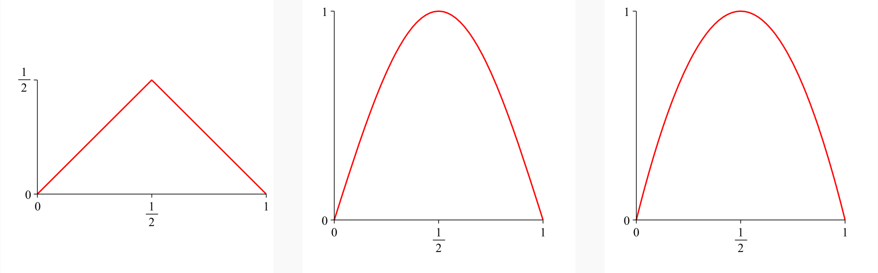


Figure: Plots of piecewise linear u_T (left), actual eigenfunction φ_1 (middle), and quadratic u_T (right).

Example (cont.)

For example (a factor of 4 is not important),

$$u_T(x) = x - x^2 \qquad \text{with} \qquad u_T'(x) = 1 - 2x$$

gives us the much better (since smaller) upper bound

$$\begin{aligned}
\lambda_1 \quad &\le \quad \frac{\int_0^1 \left[u_T'(x)\right]^2 \mathrm{d}x}{\int_0^1 u_T^2(x)\,\mathrm{d}x} = \frac{\int_0^1 (1-2x)^2 \mathrm{d}x}{\int_0^1 (x-x^2)^2 \mathrm{d}x} \\
&= \quad \frac{\int_0^1 (1-4x+4x^2)\mathrm{d}x}{\int_0^1 (x^2-2x^3+x^4)\mathrm{d}x} = \frac{x - 2x^2 + \frac{4x^3}{3}\Big|_0^1}{\frac{x^3}{3} - \frac{x^4}{2} + \frac{x^5}{5}\Big|_0^1} \\
&= \quad \frac{1 - 2 + \frac{4}{3}}{\frac{1}{3} - \frac{1}{2} + \frac{1}{5}} = \frac{\frac{1}{3}}{\frac{10-15+6}{30}} = 10
\end{aligned}$$

Proof. (of the minimization principle)
According to the theorem we want to show

$$\lambda_1 = \min_{\substack{u \in C(a,b) \\ u \text{ satisfies BCs}}} RQ[u]$$

$$= \min_{\substack{u \in C(a,b) \\ u \text{ satisfies BCs}}} \frac{-p(x)u(x)u'(x)\big|_a^b + \int_a^b \left(p(x)\left[u'(x)\right]^2 - q(x)u^2(x)\right)\mathrm{d}x}{\int_a^b u^2(x)\sigma(x)\mathrm{d}x}.$$

For the proof it is better to deal with an equivalent formulation of the Rayleigh quotient (prior to the application of integration by parts):

$$RQ[u] = \frac{-\int_a^b \left(u(x)\frac{\mathrm{d}}{\mathrm{d}x}\left[p(x)u'(x)\right] + q(x)u^2(x)\right)\mathrm{d}x}{\int_a^b u^2(x)\sigma(x)\mathrm{d}x}$$

$$= \frac{-\int_a^b u(x)(\mathcal{L}u)(x)\mathrm{d}x}{\int_a^b u^2(x)\sigma(x)\mathrm{d}x} \qquad (16)$$

The eigenfunction expansion for u is given by

$$u(x) = \sum_{n=1}^{\infty} a_n \varphi_n(x). \tag{17}$$

If u is continuous and u and φ satisfy the same BCs, then linearity allows us to write[1]

$$(\mathcal{L}u)(x) = \mathcal{L}\left(\sum_{n=1}^{\infty} a_n \varphi_n\right)(x) = \sum_{n=1}^{\infty} a_n (\mathcal{L}\varphi_n)(x).$$

Now, since the φ_n are eigenfunctions we know from the SL DE

$$\mathcal{L}\varphi_n = -\lambda_n \sigma \varphi_n$$

and so we get an eigenfunction expansion for $\mathcal{L}u$

$$(\mathcal{L}u)(x) = -\sum_{n=1}^{\infty} a_n \lambda_n \sigma(x) \varphi_n(x). \tag{18}$$

[1] This isn't actually proved until Chapter 7

We now use the eigenfunction expansions (17) for u and (18) for $\mathcal{L}u$ in equation (16) for the Rayleigh quotient to get

$$
\begin{aligned}
RQ[u] &= \frac{-\int_a^b u(x)(\mathcal{L}u)(x)\mathrm{d}x}{\int_a^b u^2(x)\sigma(x)\mathrm{d}x} \\
&= \frac{\int_a^b \sum_{n=1}^{\infty} a_n\varphi_n(x) \sum_{n=1}^{\infty} a_n\lambda_n\sigma(x)\varphi_n(x)\mathrm{d}x}{\int_a^b \sigma(x)\sum_{n=1}^{\infty} a_n\varphi_n(x) \sum_{n=1}^{\infty} a_n\varphi_n(x)\mathrm{d}x} \\
&= \frac{\int_a^b \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} a_m a_n \varphi_m(x)\varphi_n(x)\lambda_n\sigma(x)\mathrm{d}x}{\int_a^b \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} a_m a_n \varphi_m(x)\varphi_n(x)\sigma(x)\mathrm{d}x}.
\end{aligned}
$$

Interchange of integration and infinite summation gives

$$RQ[u] = \frac{\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} a_m a_n \lambda_n \int_a^b \varphi_m(x)\varphi_n(x)\sigma(x)\mathrm{d}x}{\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} a_m a_n \int_a^b \varphi_m(x)\varphi_n(x)\sigma(x)\mathrm{d}x}$$

and orthogonality of the eigenfunctions, i.e., $\int_a^b \varphi_m(x)\varphi_n(x)\sigma(x)\mathrm{d}x = 0$ whenever $m \neq n$, reduces this to

$$RQ[u] = \frac{\sum_{n=1}^{\infty} a_n^2 \lambda_n \int_a^b \varphi_n^2(x)\sigma(x)\mathrm{d}x}{\sum_{n=1}^{\infty} a_n^2 \int_a^b \varphi_n^2(x)\sigma(x)\mathrm{d}x}.$$

Since the eigenvalues are ordered, i.e., $\lambda_1 < \lambda_2 < \ldots$, we can estimate

$$\frac{\sum_{n=1}^{\infty} a_n^2 \lambda_1 \int_a^b \varphi_n^2(x)\sigma(x)\mathrm{d}x}{\sum_{n=1}^{\infty} a_n^2 \int_a^b \varphi_n^2(x)\sigma(x)\mathrm{d}x} \le \frac{\sum_{n=1}^{\infty} a_n^2 \lambda_n \int_a^b \varphi_n^2(x)\sigma(x)\mathrm{d}x}{\sum_{n=1}^{\infty} a_n^2 \int_a^b \varphi_n^2(x)\sigma(x)\mathrm{d}x} = RQ[u]$$

with "$=$" possible only if $a_n = 0$ for all $n > 1$, i.e., if the eigenfunction expansion of u consisted only of $a_1\varphi_1$.

Notice that

$$\frac{\sum_{n=1}^{\infty} a_n^2 \lambda_1 \int_a^b \varphi_n^2(x)\sigma(x)\mathrm{d}x}{\sum_{n=1}^{\infty} a_n^2 \int_a^b \varphi_n^2(x)\sigma(x)\mathrm{d}x} = \lambda_1 \underbrace{\frac{\sum_{n=1}^{\infty} a_n^2 \int_a^b \varphi_n^2(x)\sigma(x)\mathrm{d}x}{\sum_{n=1}^{\infty} a_n^2 \int_a^b \varphi_n^2(x)\sigma(x)\mathrm{d}x}}_{=1} = \lambda_1$$

and therefore

$$\lambda_1 \le RQ[u]$$

with equality only if $u = a_1\varphi_1$.

Therefore, the Rayleigh quotient $RQ[u]$ is minimized only if u is the eigenfunction corresponding to λ_1.

Remark

One can show that

$$\lambda_2 = \min_{\substack{u \in C(a,b) \\ u \text{ satisfies BCs} \\ u \text{ orthogonal to } \varphi_1}} RQ[u]$$

and iteratively obtained analogous statements for further eigenvalues.

For a nonuniform string we use the PDE

$$\rho(x)\frac{\partial^2 u}{\partial t^2}(x,t) = T_0\frac{\partial^2 u}{\partial x^2}(x,t)$$

with nonuniform density ρ, but constant tension T_0 (and no external forces).
Standard BCs and ICs are

$$u(0,t) = u(L,t) = 0$$
$$u(x,0) = f(x) \quad \text{and} \quad \frac{\partial u}{\partial t}(x,0) = g(x).$$

We will now see how far we can take the separation of variables approach for this problem.

The *Ansatz* $u(x,t) = \varphi(x)T(t)$ gives us

$$\rho(x)\varphi(x)T''(t) = T_0\varphi''(x)T(t)$$

or

$$\frac{T''(t)}{T(t)} = \frac{T_0}{\rho(x)}\frac{\varphi''(x)}{\varphi(x)} = -\lambda$$

resulting in the two ODEs

$$T''(t) = -\lambda T(t) \quad (19)$$

$$T_0\varphi''(x) + \lambda\rho(x)\varphi(x) = 0 \quad (20)$$

Notice that the second ODE (20)

$$T_0\varphi''(x) + \lambda\rho(x)\varphi(x) = 0$$

is a Sturm–Liouville ODE with $p(x) = T_0$, $q(x) = 0$ and $\sigma(x) = \rho(x)$ and BCs

$$\varphi(0) = \varphi(L) = 0.$$

Due to the variable coefficient $\rho(x)$ we don't know how to solve this eigenvalue problem.

Therefore, we try to get as much insight as possible into the solution using the general SL properties.

We use the Rayleigh quotient to study the eigenvalues:

$$\lambda = \frac{-\ T_0\varphi(x)\varphi'(x)|_0^L + \int_0^L T_0\left[\varphi'(x)\right]^2 \mathrm{d}x}{\int_0^L \varphi^2(x)\rho(x)\mathrm{d}x}.$$

From the BCs $\varphi(0) = \varphi(L) = 0$ we know that the first term in the numerator is zero. Therefore

$$\lambda = \frac{T_0 \int_0^L \left[\varphi'(x)\right]^2 \mathrm{d}x}{\int_0^L \varphi^2(x)\rho(x)\mathrm{d}x} \geq 0.$$

Moreover, we note that $\lambda = 0$ is not possible since $\varphi' \not\equiv 0$ (otherwise φ would have to be constant, and due to the BCs equal to zero).

Therefore, $\lambda > 0$ and we know that the time-equation (19) has oscillating solutions

$$T_n(t) = c_1 \cos\sqrt{\lambda_n}t + c_2 \sin\sqrt{\lambda_n}t, \qquad n = 1, 2, 3, \ldots$$

By the superposition principle we get

$$u(x,t) = \sum_{n=1}^{\infty} \left[a_n \cos \sqrt{\lambda_n} t + b_n \sin \sqrt{\lambda_n} t \right] \varphi_n(x).$$

In order to apply the ICs we need

$$\frac{\partial u}{\partial t}(x,t) = \sum_{n=1}^{\infty} \left[-a_n \sqrt{\lambda_n} \sin \sqrt{\lambda_n} t + b_n \sqrt{\lambda_n} \cos \sqrt{\lambda_n} t \right] \varphi_n(x)$$

and now we can enforce

$$\begin{aligned} u(x,0) &= \sum_{n=1}^{\infty} a_n \varphi_n(x) \stackrel{!}{=} f(x) \\ \frac{\partial u}{\partial t}(x,0) &= \sum_{n=1}^{\infty} b_n \sqrt{\lambda_n} \varphi_n(x) \stackrel{!}{=} g(x) \end{aligned}$$

The generalized Fourier coefficients a_n and b_n are obtained using the orthogonality of the eigenfunctions (with respect to the weight function ρ):

$$
\begin{aligned}
a_n &= \frac{\int_0^L f(x)\varphi_n(x)\rho(x)\,\mathrm{d}x}{\int_0^L \varphi_n^2(x)\rho(x)\,\mathrm{d}x} \\
b_n &= \frac{1}{\sqrt{\lambda_n}} \frac{\int_0^L g(x)\varphi_n(x)\rho(x)\,\mathrm{d}x}{\int_0^L \varphi_n^2(x)\rho(x)\,\mathrm{d}x}
\end{aligned}
$$

However, since we don't know the eigenfunctions φ_n we cannot make any further use of this information.

What else can we say?

From the superposition solution

$$u(x,t) = \sum_{n=1}^{\infty} \left[a_n \cos \sqrt{\lambda_n} t + b_n \sin \sqrt{\lambda_n} t \right] \varphi_n(x)$$

and the fact that the eigenvalues are ordered it is clear that $\sqrt{\lambda_1}$ is the lowest frequency of vibration (i.e., the basic mode).

What can we say about λ_1?

The minimization principle tells us

$$\lambda_1 = \min RQ[u] = \min \frac{T_0 \int_0^L [u'(x)]^2 \, dx}{\int_0^L u^2(x)\rho(x) \, dx}. \tag{21}$$

Remark

For a specific problem with given density $\rho(x)$ we could find approximate numerical upper bounds for λ_1 as we did earlier.

Alternatively, we can obtain upper and lower bounds for λ_1 if we assume that the density is bounded, i.e.,

$$0 \le \rho_{\min} \le \rho(x) \le \rho_{\max}.$$

Then we can bound the denominator of (21)

$$\rho_{\min} \int_0^L u^2(x)\,\mathrm{d}x \le \int_0^L u^2(x)\rho(x)\,\mathrm{d}x \le \rho_{\max} \int_0^L u^2(x)\,\mathrm{d}x$$

and so (21) gives us

$$\frac{T_0}{\rho_{\max}} \frac{\int_0^L [u'(x)]^2\,\mathrm{d}x}{\int_0^L u^2(x)\,\mathrm{d}x} \le \lambda_1 \le \frac{T_0}{\rho_{\min}} \frac{\int_0^L [u'(x)]^2\,\mathrm{d}x}{\int_0^L u^2(x)\,\mathrm{d}x}. \tag{22}$$

Remark

The advantage of this formulation is that we now have the Rayleigh quotient for a uniform string problem.

The Rayleigh quotient characterization of the smallest eigenvalue $\tilde{\lambda}_1$ of the uniform string problem is

$$\tilde{\lambda}_1 = \min \frac{\int_0^L [u'(x)]^2 \, dx}{\int_0^L u^2(x) \, dx},$$

while the corresponding SL problem is

$$\tilde{\varphi}''(x) + \tilde{\lambda}\tilde{\varphi}(x) = 0$$
$$\tilde{\varphi}(0) = \tilde{\varphi}(L) = 0$$

for which we know that

$$\tilde{\lambda}_1 = \left(\frac{\pi}{L}\right)^2.$$

Therefore, going back to (22), we have

$$\frac{T_0}{\rho_{\max}}\frac{\pi^2}{L^2} \le \lambda_1 \le \frac{T_0}{\rho_{\min}}\frac{\pi^2}{L^2}$$

or

$$\sqrt{\frac{T_0}{\rho_{\max}}}\frac{\pi}{L} \le \sqrt{\lambda_1} \le \sqrt{\frac{T_0}{\rho_{\min}}}\frac{\pi}{L},$$

where the bounds for the frequency $\sqrt{\lambda_1}$ are the lowest frequency for a uniform string with constant density $\rho_{\max}$ or $\rho_{\min}$, respectively.

Since we now will mostly be interested in studying the spatial Sturm–Liouville problem associated with third kind (or Robin) boundary conditions, we can think of starting with a PDE that could be either a heat equation or a wave equation, i.e.,

$$\frac{\partial u}{\partial t}(x,t) = k\frac{\partial^2 u}{\partial x^2}(x,t) \qquad \text{or} \qquad \frac{\partial^2 u}{\partial t^2}(x,t) = c^2\frac{\partial^2 u}{\partial x^2}(x,t).$$

The initial conditions will be

$$u(x,0) = f(x) \qquad \text{or} \qquad \begin{array}{l} u(x,0) = f(x) \\ \frac{\partial u}{\partial t}(x,0) = g(x) \end{array}$$

and as boundary conditions we take

$$\begin{aligned} u(0,t) &= 0 \\ \frac{\partial u}{\partial x}(L,t) &= -hu(L,t). \end{aligned}$$

The right end BC

$$\frac{\partial u}{\partial x}(L,t) = -hu(L,t)$$

corresponds to

- Newton's law of cooling with $h = H/K_0$ (with heat transfer coefficient H and thermal conductivity K_0) for the heat equation, or
- an elastic BC (such as a spring-mass system) with restoring force $h = k/T_0$ (where k is the spring constant and T_0 the tension) for the wave equation.

Remark

Note that

- $h > 0$ *suggests that heat leaves the rod or motion is stabilized at* $x = L$,
- $h < 0$ *implies that heat enters the rod or the motion is destabilized at* $x = L$, *and*
- $h = 0$ *corresponds to perfect insulation or free motion at* $x = L$.

Separation of variables with $u(x,t) = \varphi(x)T(t)$ results in the time ODE

- for the heat equation

$$T'(t) = -\lambda k T(t) \quad \Longrightarrow \quad T(t) = c_0 e^{-\lambda k t}$$

- for the wave equation

$$T''(t) = -\lambda c^2 T(t) \quad \Longrightarrow \quad T(t) = c_1 \cos\sqrt{\lambda}ct + c_2 \sin\sqrt{\lambda}ct$$

and the regular Sturm–Liouville problem

$$\varphi''(x) + \lambda\varphi(x) = 0$$
$$\varphi(0) = 0 \quad \text{and} \quad \varphi'(L) + h\varphi(L) = 0.$$

We now need to carefully study solutions of the SL problem in all three possible cases $\lambda > 0$, $\lambda = 0$ and $\lambda < 0$ – especially since we have to consider the role of the additional parameter h.

Case I: $\lambda > 0$

In this case we get a general solution of the form

$$\varphi(x) = c_1 \cos\sqrt{\lambda}x + c_2 \sin\sqrt{\lambda}x$$

and the BC $\varphi(0) = 0$ immediately gives us

$$\varphi(0) = c_1 \cos 0 + c_2 \sin 0 \stackrel{!}{=} 0 \quad \Longrightarrow \quad c_1 = 0.$$

Therefore we need to consider only $\varphi(x) = c_2 \sin\sqrt{\lambda}x$.
For the second BC we require the derivative $\varphi'(x) = c_2\sqrt{\lambda}\cos\sqrt{\lambda}x$ and then

$$\varphi'(L) + h\varphi(L) = c_2\sqrt{\lambda}\cos\sqrt{\lambda}L + hc_2 \sin\sqrt{\lambda}L \stackrel{!}{=} 0.$$

Assuming $c_2 \neq 0$ ($\rightsquigarrow$ trivial solution) and $h \neq 0$ ($\rightsquigarrow$ different BC) we get

$$\frac{\sin\sqrt{\lambda}L}{\cos\sqrt{\lambda}L} = -\frac{\sqrt{\lambda}}{h} \qquad \Longleftrightarrow \qquad \tan\sqrt{\lambda}L = -\frac{\sqrt{\lambda}}{h}.$$

The equation

$$\tan \sqrt{\lambda}L = -\frac{\sqrt{\lambda}}{h}$$

characterizing the eigenvalues cannot be solved analytically.

We can attempt to get

- a qualitative graphical solution for arbitrary h and L, or
- a quantitative numerical solution, however only for specific values of h and L.

Both approaches can be illustrated with the MATLAB script `RobinBCs.m`.

Let's assume $h > 0$ and scale everything so that units on the x-axis are units of $\sqrt{\lambda}L$.

Then we plot the intersection of $y = \tan\sqrt{\lambda}L$ and $y = -\frac{\sqrt{\lambda}}{h} = -\frac{\sqrt{\lambda}L}{hL}$.

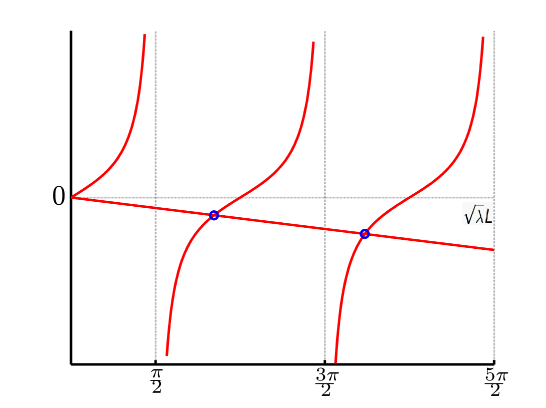


Figure: Plot of $y = \tan\sqrt{\lambda}L$ and $y = -\frac{\sqrt{\lambda}L}{hL}$.

From the plot we can see that the (scaled square root of the) eigenvalues satisfy

$$\begin{aligned}
\tfrac{\pi}{2} &< \sqrt{\lambda_1}L < \pi \\
\tfrac{3\pi}{2} &< \sqrt{\lambda_2}L < 2\pi \\
&\vdots \\
\tfrac{(2n-1)\pi}{2} &< \sqrt{\lambda_n}L < n\pi
\end{aligned}$$

In fact, $\sqrt{\lambda_n}L$ approaches the left end $\frac{(2n-1)\pi}{2}$ as $n \to \infty$.
Therefore, we actually have a third option for large values of n and $h > 0$:

$$\lambda_n \approx \left(\frac{(2n-1)\pi}{2L}\right)^2.$$

This formula describes the asymptotic behavior of the eigenvalues.

If $h < 0$ (and still $\lambda > 0$) the line has a positive slope and the graphical solution actually depends on the product of hL:

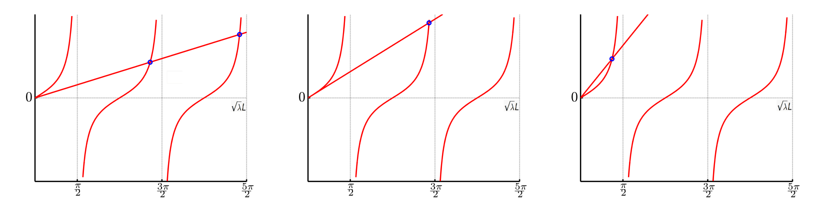


Figure: Plot of intersection of $y = \tan\sqrt{\lambda}L$ and $y = -\frac{\sqrt{\lambda}L}{hL}$ for $h < 0$ and $h < -\frac{1}{L}$ (left), $h = -\frac{1}{L}$ (middle), $h > -\frac{1}{L}$ (right).

Note that in the case $h > -\frac{1}{L}$ we have an eigenvalue in $(0, \frac{\pi}{2})$ which we didn't have before (there are, of course, still infinitely many eigenvalues), and the eigenfunctions are still

$$\varphi_n(x) = \sin\sqrt{\lambda_n}x.$$

Case II: $\lambda = 0$

Now the general solution is of the form

$$\varphi(x) = c_1 + c_2 x$$

and the BC $\varphi(0) = 0$ immediately gives us

$$\varphi(0) = c_1 \stackrel{!}{=} 0.$$

Therefore we still need to consider $\varphi(x) = c_2 x$.
The second BC, $\varphi'(L) + h\varphi(L) = 0$ implies

$$\varphi'(L) + h\varphi(L) = c_2 + hc_2 L \stackrel{!}{=} 0.$$

Assuming $c_2 \neq 0$, this equation will be satisfied for $h = -\frac{1}{L}$, and so $\lambda = 0$ is an eigenvalue with associated eigenfunction $\varphi(x) = x$ provided $h = -\frac{1}{L}$.
For other values of h, $\lambda = 0$ is not and eigenvalue.

Case III: $\lambda < 0$

In this case we can write the general solution in the form

$$\varphi(x) = c_1 \cosh \sqrt{-\lambda} x + c_2 \sinh \sqrt{-\lambda} x$$

and the BC $\varphi(0) = 0$ immediately gives us

$$\varphi(0) = c_1 \underbrace{\cosh 0}_{=1} + c_2 \underbrace{\sinh 0}_{=0} \stackrel{!}{=} 0 \quad \Longrightarrow \quad c_1 = 0.$$

Therefore we need to consider only $\varphi(x) = c_2 \sinh \sqrt{-\lambda} x$.
For the second BC we use the derivative $\varphi'(x) = c_2 \sqrt{-\lambda} \cosh \sqrt{-\lambda} x$ and then

$$\varphi'(L) + h\varphi(L) = c_2 \sqrt{-\lambda} \cosh \sqrt{-\lambda} L + h c_2 \sinh \sqrt{-\lambda} L \stackrel{!}{=} 0.$$

Assuming $c_2 \neq 0$ and $h \neq 0$ we get

$$\frac{\sinh \sqrt{-\lambda} L}{\cosh \sqrt{-\lambda} L} = -\frac{\sqrt{-\lambda}}{h} \qquad \Longleftrightarrow \qquad \tanh \sqrt{-\lambda} L = -\frac{\sqrt{-\lambda}}{h}.$$

Here we scale everything so that units on the x-axis are units of $\sqrt{-\lambda}L$.

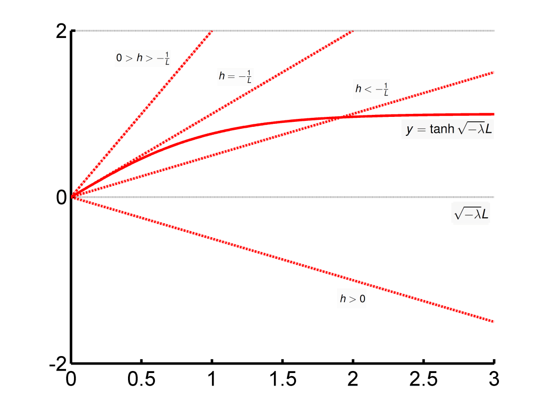


Figure: Plot of $y = \tanh\sqrt{-\lambda}L$ together with lines $y = -\frac{\sqrt{-\lambda}L}{hL}$ for different h.

Since the hyperbolic tangent does not oscillate we can pick up at most one negative eigenvalue λ_0 (when $h < -\frac{1}{L}$). Its eigenfunction is

$$\varphi_0(x) = \sinh\sqrt{-\lambda_0}x$$

The special case $h = 0$:

This case corresponds to perfect insulation (or free vibration) at the end $x = L$ and can easily be solved directly.

The eigenvalues in this case are

$$\lambda_n = \left(\frac{(2n-1)\pi}{2L}\right)^2, \quad n = 1, 2, 3, \ldots,$$

and the corresponding eigenfunctions are

$$\varphi_n(x) = \sin \sqrt{\lambda_n} x, \quad n = 1, 2, 3, \ldots.$$

Altogether, we can summarize the eigenvalues and eigenfunctions for this example in the following table:

	$\lambda > 0$	$\lambda = 0$	$\lambda < 0$
$h > -\frac{1}{L}$	$\sin\sqrt{\lambda}x$		
$h = -\frac{1}{L}$	$\sin\sqrt{\lambda}x$	x	
$h < -\frac{1}{L}$	$\sin\sqrt{\lambda}x$		$\sinh\sqrt{-\lambda_0}x$

Here λ_0 is the one extra negative eigenvalue which will arise for $h < -\frac{1}{L}$.

Remark

You can compare this with Table 5.8.1 in [Haberman], where an additional split into "physical" ($h \geq 0$) and "nonphysical" ($h < 0$) situations was made.

In this section we want to study how to "best" represent an (infinite) generalized Fourier series by a finite linear combination of the eigenfunctions.

We let

$$s_M(x) = \sum_{n=1}^{M} \alpha_n \varphi_n(x)$$

be an M-term approximation of the generalized Fourier series

$$f(x) \sim \sum_{n=1}^{\infty} a_n \varphi_n(x)$$

with eigenfunctions φ_n and generalized Fourier coefficients

$$a_n = \frac{\int_a^b f(x)\varphi_n(x)\sigma(x)\,\mathrm{d}x}{\int_a^b \varphi_n^2(x)\sigma(x)\,\mathrm{d}x}.$$

How should we choose the unknown coefficients α_n of s_M?

We decide to choose α_n such that (for fixed M)

$$\|f - s_M\| = \|f - \sum_{n=1}^{M} \alpha_n \varphi_n\| \quad \text{is minimized.}$$

Here the norm $\|\cdot\|$ is some measure of "goodness" to be defined below.

Remark

Since we will obtain that s_M with minimal norm $\|f - s_M\|$ among all possible M-term eigenfunction approximations s_M we will have found the "best" one.

What kind of norm should we use?

Possible (standard) choices are

- the one-norm

$$\|f - s_M\|_1 = \int_a^b |f(x) - s_M(x)| \, dx,$$

- the weighted two-norm (or weighted least squares norm)

$$\|f - s_M\|_2 = \left(\int_a^b [f(x) - s_M(x)]^2 \, \sigma(x) \, dx \right)^{1/2},$$

- the maximum norm (or infinity-norm)

$$\|f - s_M\|_\infty = \max_{x \in [a,b]} |f(x) - s_M(x)|,$$

The one- and infinity-norms are not as practical as the two-norm. Therefore, we use the weighted least squares norm.

We will choose the coefficients α_n to minimize the (square of the) weighted least squares norm, i.e., we want to solve

$$\begin{aligned}\min_{\alpha_n} E &= \min_{\alpha_n} \int_a^b [f(x) - s_M(x)]^2 \sigma(x)\,\mathrm{d}x \\ &= \min_{\alpha_n} \int_a^b \left[f(x) - \sum_{n=1}^{M} \alpha_n \varphi_n(x)\right]^2 \sigma(x)\,\mathrm{d}x.\end{aligned}$$

This problem is a multivariate optimization problem and can be solved with standard methods from Calculus III.
A necessary condition for obtaining a minimum is

$$\frac{\partial E}{\partial \alpha_i} = 0 \qquad i = 1, 2, \ldots, M.$$

The first thing we need are the partial derivatives

$$\frac{\partial E}{\partial \alpha_i} = \frac{\partial}{\partial \alpha_i}\left(\int_a^b \left[f(x) - \sum_{n=1}^{M} \alpha_n \varphi_n(x)\right]^2 \sigma(x)\,\mathrm{d}x\right), \qquad i = 1, 2, \ldots, M.$$

By the chain rule we have

$$\frac{\partial E}{\partial \alpha_i} = -2\int_a^b \left[f(x) - \sum_{n=1}^{M} \alpha_n \varphi_n(x)\right] \varphi_i(x)\sigma(x)\,\mathrm{d}x, \qquad i = 1, 2, \ldots, M. \tag{23}$$

We now need to set these equal to zero and solve for α_i.

Setting $\frac{\partial E}{\partial \alpha_i} = 0$ in (23) we get

$$\begin{aligned}
\int_a^b f(x)\varphi_i(x)\sigma(x)\,\mathrm{d}x &= \int_a^b \sum_{n=1}^{M} \alpha_n \varphi_n(x)\varphi_i(x)\sigma(x)\,\mathrm{d}x \\
&= \sum_{n=1}^{M} \alpha_n \underbrace{\int_a^b \varphi_n(x)\varphi_i(x)\sigma(x)\,\mathrm{d}x}_{=0 \text{ if } n \neq i} \\
&= \alpha_i \int_a^b \varphi_i^2(x)\sigma(x)\,\mathrm{d}x
\end{aligned}$$

and so

$$\alpha_i = \frac{\int_a^b f(x)\varphi_i(x)\sigma(x)\,\mathrm{d}x}{\int_a^b \varphi_i^2(x)\sigma(x)\,\mathrm{d}x} = a_i,$$

i.e., truncating the generalized Fourier series might be the optimal choice (this is only a necessary condition).

We now show that $\alpha_n = a_n$ indeed does minimize E. Consider

$$
\begin{aligned}
E &= \int_a^b [f(x) - s_M(x)]^2 \sigma(x) \mathrm{d}x = \int_a^b \left[f^2(x) - 2f(x)s_M(x) + s_M^2(x)\right] \sigma(x) \mathrm{d}x \\
&= \int_a^b \left[f^2(x) - 2f(x) \sum_{n=1}^M \alpha_n \varphi_n(x) + \sum_{n=1}^M \sum_{\ell=1}^M \alpha_n \alpha_\ell \varphi_n(x) \varphi_\ell(x) \right] \sigma(x) \mathrm{d}x
\end{aligned}
$$

Interchanging integration and (finite) summation (no problem at all!) and using orthogonality of the eigenfunctions we know that

$$
\begin{aligned}
\int_a^b \sum_{n=1}^M \sum_{\ell=1}^M \alpha_n \alpha_\ell \varphi_n(x) \varphi_\ell(x) \sigma(x) \mathrm{d}x &= \sum_{n=1}^M \sum_{\ell=1}^M \alpha_n \alpha_\ell \int_a^b \varphi_n(x) \varphi_\ell(x) \sigma(x) \mathrm{d}x \\
&= \sum_{n=1}^M \alpha_n^2 \int_a^b \varphi_n^2(x) \sigma(x) \, \mathrm{d}x
\end{aligned}
$$

and therefore

$$
E = \int_a^b \left[f^2(x) - 2f(x) \sum_{n=1}^M \alpha_n \varphi_n(x) + \sum_{n=1}^M \alpha_n^2 \varphi_n^2(x) \right] \sigma(x) \, \mathrm{d}x.
$$

We can rearrange

$$E = \int_a^b \left[f^2(x) - 2f(x) \sum_{n=1}^{M} \alpha_n \varphi_n(x) + \sum_{n=1}^{M} \alpha_n^2 \varphi_n^2(x) \right] \sigma(x) \mathrm{d}x$$

as

$$E = \sum_{n=1}^{M} \left[\alpha_n^2 \int_a^b \varphi_n^2(x)\sigma(x)\mathrm{d}x - 2\alpha_n \int_a^b f(x)\varphi_n(x)\sigma(x)\mathrm{d}x \right] + \int_a^b f^2(x)\sigma(x)\mathrm{d}x$$

and then further modify

$$E = \sum_{n=1}^{M} \left[\alpha_n^2 \int_a^b \varphi_n^2(x)\sigma(x)\mathrm{d}x - 2\alpha_n \int_a^b f(x)\varphi_n(x)\sigma(x)\mathrm{d}x \underbrace{\frac{\int_a^b \varphi_n^2(x)\sigma(x)\mathrm{d}x}{\int_a^b \varphi_n^2(x)\sigma(x)\mathrm{d}x}}_{=1} \right] + \int_a^b f^2(x)\sigma(x)\mathrm{d}x$$

$$= \sum_{n=1}^{M} \left[\alpha_n^2 \int_a^b \varphi_n^2(x)\sigma(x)\mathrm{d}x - 2\alpha_n \underbrace{\frac{\int_a^b f(x)\varphi_n(x)\sigma(x)\mathrm{d}x}{\int_a^b \varphi_n^2(x)\sigma(x)\mathrm{d}x}}_{=a_n} \int_a^b \varphi_n^2(x)\sigma(x)\mathrm{d}x \right] + \int_a^b f^2(x)\sigma(x)\mathrm{d}x$$

$$= \sum_{n=1}^{M} \left[\left(\alpha_n^2 - 2\alpha_n a_n \right) \int_a^b \varphi_n^2(x)\sigma(x)\mathrm{d}x \right] + \int_a^b f^2(x)\sigma(x)\mathrm{d}x$$

Now we take

$$E = \sum_{n=1}^{M} \left[\left(\alpha_n^2 - 2\alpha_n a_n \right) \int_a^b \varphi_n^2(x)\sigma(x)\mathrm{d}x \right] + \int_a^b f^2(x)\sigma(x)\mathrm{d}x$$

and complete the square to get

$$E = \sum_{n=1}^{M} \left[\left((\alpha_n - a_n)^2 - a_n^2 \right) \int_a^b \varphi_n^2(x)\sigma(x)\mathrm{d}x \right] + \int_a^b f^2(x)\sigma(x)\mathrm{d}x$$

The only terms we can manipulate to reduce the value of E are the nonnegative integrals

$$(\alpha_n - a_n)^2 \int_a^b \varphi_n^2(x)\sigma(x)\mathrm{d}x.$$

The best we can do, i.e., the smallest we can make E, is to remove these terms from the summation. This happens if $\alpha_n = a_n$. Therefore, truncation of the generalized Fourier series is indeed optimal.

Remark

Note that the choice of the optimal coefficients $\alpha_n = a_n$ was independent of the particular value of M.

This means that if s_M for a particular value M turns out not to be good enough, then we can obtain the more accurate s_{M+1} by computing only one additional coefficient $\alpha_{M+1} = a_{M+1}$.

This is not at all obvious. In many cases, allowing for one more term in the expansion may require recomputation of all coefficients.

If we let $\alpha_n = a_n$ above, then we see that the actual minimum error is

$$E = \int_a^b f^2(x)\sigma(x)\mathrm{d}x - \sum_{n=1}^{M} a_n^2 \int_a^b \varphi_n^2(x)\sigma(x)\mathrm{d}x \tag{24}$$

Example

Assume that the eigenfunctions are orthonormal with weight $\sigma(x) = 1$, i.e.,

$$\int_a^b \varphi_n(x)\varphi_m(x)\mathrm{d}x = \begin{cases} 0 & \text{if } m \neq n, \\ 1 & \text{if } m = n. \end{cases}$$

Then the least squares error when approximating f by its truncated generalized Fourier series $\sum_{n=1}^{M} a_n\varphi_n$ on $[a,b]$ is

$$E = \int_a^b f^2(x)\mathrm{d}x - \sum_{n=1}^{M} a_n^2.$$

Note that the error involves only the Fourier coefficients, but not the eigenfunctions.

Bessel's Inequality

From formula (24) and the definition of E we have

$$0 \leq E = \int_a^b f^2(x)\sigma(x)\mathrm{d}x - \sum_{n=1}^{M} a_n^2 \int_a^b \varphi_n^2(x)\sigma(x)\mathrm{d}x$$

and therefore

$$\int_a^b f^2(x)\sigma(x)\mathrm{d}x \geq \sum_{n=1}^{M} a_n^2 \int_a^b \varphi_n^2(x)\sigma(x)\mathrm{d}x.$$

This is known as Bessel's inequality.

Parseval's Identity

From the definition of the weighted least squares error

$$E_M = \int_a^b \left[f(x) - \sum_{n=1}^{M} a_n \varphi_n(x) \right]^2 \sigma(x) \mathrm{d}x$$

and the convergence properties of generalized Fourier series (convergence of the series to a value different from $f(x)$ at finitely many points x does not affect the values of the integral!) we get that

$$\lim_{M \to \infty} E_M = 0.$$

This shows that the generalized Fourier series of f converges to f in the least squares sense on the entire interval $[a, b]$.

Moreover, formula (24) for $M \to \infty$ gives us

$$\int_a^b f^2(x)\sigma(x)\mathrm{d}x = \sum_{n=1}^{\infty} a_n^2 \int_a^b \varphi_n^2(x)\sigma(x)\mathrm{d}x.$$

This is known as Parseval's identity, and can be viewed as a generalization of the Pythagorean theorem to inner product spaces of functions.

Remark

Inner product spaces – and in particular Hilbert spaces – are studied in much more detail in functional analysis. They play a very important role in many applications.

Example

For orthonormal eigenfunctions with weight $\sigma \equiv 1$, Parseval's identity says

$$\int_a^b f^2(x)\mathrm{d}x = \sum_{n=1}^{\infty} a_n^2.$$

The analogy with the Pythagorean theorem perhaps becomes more apparent if we use inner product notation and norms. Then we have

$$\|f\|_2^2 = \langle f, f\rangle = \sum_{n=1}^{\infty} a_n^2 = \sum_{n=1}^{\infty} \langle f, \varphi_n\rangle^2.$$

References I

R. Haberman.
Applied Partial Differential Equations.
Pearson (5th ed.), Upper Saddle River, NJ, 2012.