Chapter 5: Sturm–Liouville Theory

Special Functions of Mathematical Physics

after K. A. Novak
Compiled by Jake W. Liu

July 27, 2026

Outline

1. Review of linear algebra

2. Sturm–Liouville operators

Where we are going

This chapter builds the operator picture behind the second-order equations of later chapters, using only basic calculus and inner products of functions.

- Functions act as vectors in an inner-product space; Fourier coefficients are extracted by orthogonality.
- A Sturm–Liouville eigenvalue problem is written, with this chapter's sign convention, as

$$\frac{d}{dx}\left(p(x)\frac{du}{dx}\right) + q(x)u(x) = -\lambda w(x)u(x),$$

  where λ is the eigenvalue and typically $p(x) > 0$ and $w(x) > 0$ in the interior of the interval.
- Integration by parts identifies the *formal adjoint*. Boundary conditions that kill the boundary form turn the problem into a self-adjoint one.
- Green's identity then yields real eigenvalues and weighted orthogonality of eigenfunctions.
- Legendre and Bessel (drum) modes illustrate the theory. Any second-order equation can be put into Sturm–Liouville form by an integrating factor. Optionally, Gram–Schmidt recovers P_0, P_1, P_2.

Table of contents

Vector spaces and bases I

A *vector space* is a collection of objects that is closed under addition and scalar multiplication. Thus, if u and v belong to the space and a, b are scalars, then

$$au + bv$$

also belongs to the space. The vectors may be functions rather than geometric arrows.

Examples.

- The polynomials of degree at most two, $a_0 + a_1 x + a_2 x^2$, form a three-dimensional vector space.
- The smooth functions on $[0, 1]$ satisfying $u(0) = u(1) = 0$ form a vector space, because

$$(au + bv)(0) = au(0) + bv(0) = 0, \qquad (au + bv)(1) = au(1) + bv(1) = 0.$$

Vector spaces and bases II

A *basis* is a linearly independent spanning set. For the quadratic polynomials, $\{1, x, x^2\}$ is a basis, and

$$1 + 2x + 3x^2 = 1 \cdot 1 + 2 \cdot x + 3 \cdot x^2. \tag{5.1.1}$$

The set

$$\{1,\ 2x - 1,\ 6x^2 - 6x + 1\}$$

is another basis. Its members have respective degrees 0, 1, and 2. To see that they are linearly independent, suppose $c_1 \cdot 1 + c_2(2x - 1) + c_3(6x^2 - 6x + 1) = 0$ for all x. Matching the coefficient of x^2 gives $6c_3 = 0$, so $c_3 = 0$; the coefficient of x is then $2c_2 = 0$, so $c_2 = 0$; and the constant term is then $c_1 = 0$. Only the trivial combination vanishes, so the three vectors are linearly independent, and three linearly independent vectors in this three-dimensional space form a basis. To determine the coefficients, write

$$\begin{aligned} 1 + 2x + 3x^2 &= A + B(2x - 1) + C(6x^2 - 6x + 1) \\ &= (A - B + C) + (2B - 6C)x + 6Cx^2. \end{aligned}$$

Vector spaces and bases III

Equating coefficients of $1, x, x^2$ gives

$$\begin{aligned} 6C &= 3, & C &= \frac{1}{2}, & \\ 2B - 6C &= 2, & 2B - 3 &= 2, & B &= \frac{5}{2}, \\ A - B + C &= 1, & A - \frac{5}{2} + \frac{1}{2} &= 1, & A &= 3. \end{aligned}$$

Therefore

$$1 + 2x + 3x^2 = 3 \cdot 1 + \frac{5}{2}(2x - 1) + \frac{1}{2}(6x^2 - 6x + 1). \tag{5.1.2}$$

Expanding the right-hand side verifies the result:

$$3 + (5x - \tfrac{5}{2}) + (3x^2 - 3x + \tfrac{1}{2}) = 1 + 2x + 3x^2.$$

The number of elements in a finite basis is the *dimension*. To discuss infinite bases, we first need a notion of distance and convergence.

Inner products, norms, and coefficients I

Step 1: choose the inner product. We use the convention that the inner product is *linear in its first slot*. An inner product $(\cdot, \cdot)$ satisfies

1. $(au + bv, w) = a(u, w) + b(v, w)$,
2. $(u, v) = \overline{(v, u)}$,
3. $(u, u) \geq 0$, with equality only for the zero vector.

In particular, for a scalar c,

$$(cu, v) = c\,(u, v). \tag{5.1.3}$$

Conjugate symmetry forces *conjugate-linearity in the second slot*:

$$(u, cv) = \overline{c}\,(u, v). \tag{5.1.4}$$

Indeed,

$$(u, cv) = \overline{(cv, u)} = \overline{c\,(v, u)} = \overline{c}\,\overline{(v, u)} = \overline{c}\,(u, v),$$

where the outer equalities use conjugate symmetry and the middle one uses (5.1.3).

Numerical check. On $\mathbb{C}^2$ with $(u, v) = u_1\overline{v_1} + u_2\overline{v_2}$, take $u = (1, 0)$, $v = (0, 1)$, and $c = i$. Then

$$(cu, v) = (i, 0) \cdot \overline{(0, 1)} = 0, \qquad c(u, v) = i \cdot 0 = 0,$$

Inner products, norms, and coefficients II

so (5.1.3) holds, while

$$(u, cv) = (1,0)\cdot\overline{(0,i)} = (1,0)\cdot(0,-i) = 0$$

and $\overline{c}(u,v) = (-i)\cdot 0 = 0$, so (5.1.4) holds. A nonzero check with $u = v = (1,0)$:

$$\begin{aligned}
&(cu, u) = (i,0)\cdot\overline{(1,0)} = i\cdot 1 = i,\\
&c(u,u) = i\cdot 1 = i,\\
&(u, cu) = (1,0)\cdot\overline{(i,0)} = (1,0)\cdot(-i,0) = -i = \overline{i}\,(u,u).
\end{aligned}$$

Scalars pull out of the first slot unchanged and out of the second slot as conjugates. The standard L^2 inner product on an interval $[a,b]$ is

$$(u,v) = \int_a^b u(x)\overline{v(x)}\,dx. \tag{5.1.5}$$

If w is a real weight with $w(x) > 0$ on the interior (it may vanish at isolated endpoints), the weighted inner product is

$$(u,v)_w = \int_a^b u(x)\overline{v(x)}\,w(x)\,dx. \tag{5.1.6}$$

For example, the Bessel radial weight on $[0,1]$ is $w(s) = s$. Then

$$(u,u)_w = \int_a^b |u(x)|^2 w(x)\,dx$$

Inner products, norms, and coefficients III

is positive for every nonzero continuous u of interest in this course.

The norm induced by an inner product is

$$\|u\| = \sqrt{(u,u)}, \qquad \|u\|_w = \sqrt{(u,u)_w}. \tag{5.1.7}$$

A nonzero function is normalized by replacing it with $u/\|u\|$.

Two vectors are *orthogonal* when $(u,v)=0$. An orthonormal set $\{u_i\}$ satisfies

$$(u_i, u_j) = \delta_{ij}, \tag{5.1.8}$$

where $\delta_{ij}=1$ for $i=j$ and 0 otherwise.

Step 2: extract coefficients by orthogonality. Suppose first that

$$f = \sum_{i=1}^{N} a_i u_i$$

Inner products, norms, and coefficients IV

in an orthonormal basis. Taking the inner product with u_j gives

$$\begin{aligned}(f, u_j) &= \left(\sum_{i=1}^{N} a_i u_i, u_j\right) \\ &= \sum_{i=1}^{N} a_i (u_i, u_j) \\ &= \sum_{i=1}^{N} a_i \delta_{ij} \\ &= a_j. \end{aligned} \tag{5.1.9}$$

(The same formula holds for an infinite expansion once one knows the series converges in the L^2 norm, so the inner product may pass under the limit.)

If the basis is orthogonal but not normalized, then

$$(f, u_j) = \sum_i a_i (u_i, u_j) = a_j (u_j, u_j) = a_j \|u_j\|^2, \tag{5.1.10}$$

so

$$a_j = \frac{(f, u_j)}{\|u_j\|^2}. \tag{5.1.11}$$

Infinite-dimensional bases and L^2 I

In finite dimensions every linearly independent spanning set has the same number of vectors. For functions we need an infinite-dimensional analogue of "basis," and that requires a notion of distance.

Think of $L^2(0,1)$ as the space of (sufficiently regular) functions on $[0,1]$ for which the square integral is finite:

$$\|f\|_2 = \left(\int_0^1 |f(x)|^2 \, dx \right)^{1/2} < \infty.$$

Two functions that differ only at finitely many points—or, more generally, on a set that does not affect integrals—are regarded as the same L^2 vector. The associated inner product is (5.1.5) with $a = 0$ and $b = 1$. This is enough for Fourier series; we will not need measure theory.

The normalized sine functions

$$\phi_n(x) = \sqrt{2} \sin(n\pi x), \qquad n = 1, 2, 3, \ldots, \tag{5.1.12}$$

form an orthonormal set; the orthogonality and the normalization are computed on the next frame. Orthonormality by itself is not enough to expand every function: discarding ϕ_1 leaves a set that is still orthonormal, yet every combination of the surviving ϕ_n is orthogonal to ϕ_1 and so cannot reproduce it. What must be added is *completeness*, and this is the one fact in the chapter that we quote rather than derive, since its proof belongs to a course on Fourier analysis.

Quoted result (completeness of the sine system). The set $\{\phi_n\}_{n\geq 1}$ is *complete* in $L^2(0,1)$: the only $f \in L^2(0,1)$ with $(f, \phi_n) = 0$ for every n is $f = 0$. Equivalently, every $f \in L^2(0,1)$ is reproduced by its own coefficients,

$$f = \sum_{n=1}^{\infty} c_n \phi_n, \qquad c_n = (f, \phi_n) = \int_0^1 f(x)\, \phi_n(x)\, dx, \tag{5.1.13}$$

with convergence in the L^2 norm, meaning $\|f - \sum_{n=1}^{N} c_n \phi_n\|_2 \to 0$ as $N \to \infty$. No conjugate bar appears in the integral because the sines are real, so $\overline{\phi_n} = \phi_n$. The coefficient formula is the infinite-dimensional counterpart of (5.1.9); it is rederived from scratch, in unnormalized form, two frames below. Equivalently, one may use the unnormalized functions $\sin(n\pi x)$ and absorb $\sqrt{2}$ into the coefficients: since $\phi_n = \sqrt{2}\sin(n\pi x)$, setting $a_n = \sqrt{2}\, c_n$ turns (5.1.13) into $f = \sum_{n\geq 1} a_n \sin(n\pi x)$.

Norm convergence and pointwise convergence are different. If f is piecewise continuously differentiable on $[0,1]$, its sine series converges at an interior point to the average of the two one-sided limits of the *odd periodic extension*: f on $[0,1]$, continued to $[-1,0]$ by $f(-x) = -f(x)$, then repeated with period 2. In particular, if

Infinite-dimensional bases and L^2 III

f is continuous and $f(0) = f(1) = 0$, the series converges pointwise to f on the whole of $[0, 1]$.

Example: orthogonality of the sine modes I

Let

$$u_m(x) = \sin(m\pi x), \qquad u_n(x) = \sin(n\pi x),$$

where m, n are positive integers. We first derive the product-to-sum identity. For arbitrary A and B,

$$\begin{aligned}
\sin A \sin B &= \frac{e^{iA} - e^{-iA}}{2i} \frac{e^{iB} - e^{-iB}}{2i} \\
&= -\frac{1}{4}\left(e^{i(A+B)} - e^{i(A-B)} - e^{-i(A-B)} + e^{-i(A+B)}\right) \\
&= \frac{1}{4}\left[e^{i(A-B)} + e^{-i(A-B)} - e^{i(A+B)} - e^{-i(A+B)}\right] \\
&= \frac{1}{2}\left[\cos(A-B) - \cos(A+B)\right]. \qquad (5.1.14)
\end{aligned}$$

Example: orthogonality of the sine modes II

Set $A = m\pi x$ and $B = n\pi x$. If $m \neq n$, then

$$\begin{aligned}
\int_0^1 u_m(x)u_n(x)\,dx &= \frac{1}{2}\int_0^1 \left[\cos((m-n)\pi x) - \cos((m+n)\pi x)\right] dx \\
&= \left[\frac{\sin((m-n)\pi x)}{2(m-n)\pi} - \frac{\sin((m+n)\pi x)}{2(m+n)\pi}\right]_{x=0}^{x=1} \\
&= \frac{\sin((m-n)\pi) - \sin 0}{2(m-n)\pi} - \frac{\sin((m+n)\pi) - \sin 0}{2(m+n)\pi} \\
&= 0, \qquad (5.1.15)
\end{aligned}$$

because $m - n$ and $m + n$ are integers and $\sin(k\pi) = 0$ for every integer k.

If $m = n$, then (5.1.14) becomes

$$\sin^2(m\pi x) = \frac{1}{2}\left[1 - \cos(2m\pi x)\right].$$

Example: orthogonality of the sine modes III

Consequently,

$$\begin{aligned}
\int_0^1 \sin^2(m\pi x)\,dx &= \frac{1}{2}\int_0^1 [1-\cos(2m\pi x)]\,dx \\
&= \left[\frac{x}{2} - \frac{\sin(2m\pi x)}{4m\pi}\right]_{x=0}^{x=1} \\
&= \left(\frac{1}{2} - \frac{\sin(2m\pi)}{4m\pi}\right) - \left(0 - \frac{\sin 0}{4m\pi}\right) \\
&= \frac{1}{2}. \qquad (5.1.16)
\end{aligned}$$

Combining the two cases gives

$$(u_m, u_n) = \int_0^1 u_m(x)u_n(x)\,dx = \frac{1}{2}\delta_{mn}. \qquad (5.1.17)$$

Thus $\|u_n\|^2 = 1/2$ and $\|u_n\| = 1/\sqrt{2}$. For the normalized mode $\phi_n = \sqrt{2}\,u_n$,

$$\|\phi_n\|^2 = (\sqrt{2}\,u_n, \sqrt{2}\,u_n) = 2\|u_n\|^2 = 2\cdot\frac{1}{2} = 1.$$

Example: Fourier sine-series coefficients I

Let $f \in L^2(0,1)$ and write its Fourier sine series as

$$f(x) = \sum_{n=1}^{\infty} a_n \sin(n\pi x), \tag{5.1.18}$$

where equality is understood in L^2 unless additional regularity gives pointwise convergence. Let $u_n(x) = \sin(n\pi x)$ and let

$$S_N(x) = \sum_{n=1}^{N} a_n u_n(x).$$

Before taking any limit, we record the one inequality that makes the inner product continuous. For $g, h \in L^2(0,1)$, the *Cauchy–Schwarz inequality* states

$$|(g,h)| \leq \|g\|\,\|h\|. \tag{5.1.19}$$

To prove it we may assume $(g,h) \neq 0$, since otherwise the left side is 0 and there is nothing to check. Write $(g,h) = |(g,h)|e^{i\alpha}$ with α real. For every real t, positivity of

Example: Fourier sine-series coefficients II

the inner product gives $\|g - te^{i\alpha}h\|^2 \geq 0$, and expanding with linearity in the first slot (5.1.3) and conjugate-linearity in the second (5.1.4),

$$\begin{aligned} 0 &\leq \left\|g - te^{i\alpha}h\right\|^2 \\ &= \|g\|^2 - te^{-i\alpha}(g,h) - te^{i\alpha}\overline{(g,h)} + t^2\|h\|^2 \\ &= \|g\|^2 - 2t\,|(g,h)| + t^2\|h\|^2, \end{aligned}$$

where the last equality uses $e^{-i\alpha}(g,h) = |(g,h)|$ and, taking conjugates, $e^{i\alpha}\overline{(g,h)} = |(g,h)|$. If $\|h\|$ were 0, the right-hand side would be the linear function $\|g\|^2 - 2t|(g,h)|$ with negative slope, which becomes negative for large t; so $\|h\| > 0$. A real quadratic in t with positive leading coefficient that is never negative has non-positive discriminant,

$$\big(2|(g,h)|\big)^2 - 4\|h\|^2\,\|g\|^2 \leq 0,$$

and dividing by 4 and taking square roots gives (5.1.19).

Now apply (5.1.19) with $g = f - S_N$ and $h = u_k$. Taking $a = 1$, $b = -1$ in the first-slot linearity axiom gives $(f, u_k) - (S_N, u_k) = (f - S_N, u_k)$, hence

$$\big|(f, u_k) - (S_N, u_k)\big| = \big|(f - S_N, u_k)\big| \leq \|f - S_N\|\,\|u_k\| = \frac{1}{\sqrt{2}}\,\|f - S_N\| \xrightarrow[N\to\infty]{} 0,$$

Example: Fourier sine-series coefficients III

since $\|u_k\| = 1/\sqrt{2}$ by (5.1.17) and $\|f - S_N\| \to 0$ because (5.1.18) is an equality in L^2. This is precisely what "the inner product is continuous" means here, and it justifies

$$(f, u_k) = \lim_{N\to\infty} (S_N, u_k).$$

For every $N \geq k$,

$$\begin{aligned}
(S_N, u_k) &= \left(\sum_{n=1}^{N} a_n u_n, u_k \right) \\
&= \sum_{n=1}^{N} a_n (u_n, u_k) \\
&= \sum_{n=1}^{N} a_n \frac{1}{2} \delta_{nk} \\
&= \frac{1}{2} a_k. \qquad (5.1.20)
\end{aligned}$$

Example: Fourier sine-series coefficients IV

Therefore

$$
\begin{aligned}
a_k &= \frac{(f, u_k)}{\|u_k\|^2} \\
&= \frac{\displaystyle\int_0^1 f(x)\sin(k\pi x)\,dx}{1/2} \\
&= 2\int_0^1 f(x)\sin(k\pi x)\,dx. \qquad (5.1.21)
\end{aligned}
$$

This is the Fourier sine-series coefficient formula. It is distinct from the continuous Fourier sine transform on the half-line.

Step 1: linear operators. A linear operator L satisfies

$$L(au + bv) = aLu + bLv.$$

Examples include differentiation, multiplication by a fixed function, and the fixed-base integral operator

$$(Au)(x) = \int_{x_0}^{x} u(t)\, dt.$$

Its linearity follows from linearity of the integral:

$$\begin{aligned} A(au + bv)(x) &= \int_{x_0}^{x} [au(t) + bv(t)]\, dt \\ &= a \int_{x_0}^{x} u(t)\, dt + b \int_{x_0}^{x} v(t)\, dt \\ &= a(Au)(x) + b(Av)(x). \end{aligned}$$

An "indefinite antiderivative" without a fixed lower limit is not a single-valued operator.

For a differential operator one must also specify which functions are allowed: the *domain* is the collection of sufficiently differentiable functions that satisfy the chosen boundary conditions. Throughout this chapter, "sufficiently differentiable" means

Linear operators and formal adjoints II

classical derivatives exist and are continuous enough for every integration by parts we write down.

Step 2: formal adjoint versus self-adjoint problem. Two phrases will be used carefully:

- The *formal adjoint* of a differential expression is the expression obtained after integration by parts, ignoring (or postponing) boundary terms.
- The eigenvalue problem is called *self-adjoint* when *both* of the following hold: (i) the differential expression is its own formal adjoint, $L^\dagger = L$ — a property of the expression alone; and (ii) the boundary conditions make the boundary terms vanish. Integration by parts always produces

  $$(u, Lv) = \Big[\text{boundary form}\Big] + (L^\dagger u, v),$$

  so killing the boundary form by itself gives only $(u, Lv) = (L^\dagger u, v)$. It is the extra hypothesis $L^\dagger = L$ that upgrades this to $(u, Lv) = (Lu, v)$ for all admissible u, v.

Linear operators and formal adjoints III

In short: formal adjoint = the differential expression left behind after IBP; self-adjoint problem = a formally self-adjoint expression ($L^\dagger = L$) *together with* boundary conditions that kill the boundary form. Step 3 below shows what goes wrong when the first half fails; Step 4 shows both halves in place.

Step 3: first derivative by parts. Let $D = d/dx$ and take u, v continuously differentiable on $[0, 1]$. Integration by parts gives

$$\begin{aligned}
(u, Dv) &= \int_0^1 u(x)\overline{v'(x)}\,dx \\
&= \Big[u(x)\overline{v(x)}\Big]_0^1 - \int_0^1 u'(x)\overline{v(x)}\,dx \\
&= [u\overline{v}]_0^1 + (-Du, v).
\end{aligned} \tag{5.1.22}$$

The integral term identifies the formal adjoint $D^\dagger = -D$, independently of boundary conditions. The boundary term

$$[u\overline{v}]_0^1 = u(1)\overline{v(1)} - u(0)\overline{v(0)}$$

vanishes, for example, when both functions are periodic, $u(0) = u(1)$ and $v(0) = v(1)$, or when both vanish at the endpoints, $u(0) = u(1) = v(0) = v(1) = 0$. Then $(u, Dv) = (-Du, v)$.

Linear operators and formal adjoints IV

This is exactly the situation the definition above warns about: the boundary form has been killed, yet $(u, Dv) = (-Du, v)$ is *not* the same as (Du, v). For a concrete pair, take $u(x) = x(1-x)$ and $v(x) = x(1-x)^2$, both real and both vanishing at $x = 0$ and $x = 1$, so the boundary term $[u\overline{v}]_0^1$ is 0. Expanding, $uv' = (x - x^2)(1 - 4x + 3x^2)$ and $u'v = (1 - 2x)(x - 2x^2 + x^3)$, and integrating term by term,

$$(u, Dv) = \int_0^1 u\, v'\, dx = -\frac{1}{60}, \qquad (Du, v) = \int_0^1 u'\, v\, dx = +\frac{1}{60},$$

which indeed sum to $[uv]_0^1 = 0$ but are not equal. The reason is structural: $D^\dagger = -D \neq D$, so D is not formally self-adjoint and no choice of boundary conditions can repair that. The first-order expression that *is* formally self-adjoint is iD: pulling the constant out of the second slot as a conjugate by (5.1.4) and then back out of the first slot by (5.1.3),

$$(u, iDv) = \overline{i}\,(u, Dv) = -i\left[u\overline{v}\right]_0^1 - i\,(-Du, v) = \left[-i\, u\overline{v}\right]_0^1 + (iDu, v),$$

so $(iD)^\dagger = iD$, and with the boundary form killed one gets $(u, iDv) = (iDu, v)$. Second-order expressions of Sturm–Liouville type, treated next, are formally self-adjoint for the same algebraic reason: two integrations by parts restore the original sign.

Step 4: second derivative with Dirichlet conditions (full algebra). Let $L = d^2/dx^2$, and take u, v twice continuously differentiable with the Dirichlet conditions

$$u(0) = u(1) = 0, \qquad v(0) = v(1) = 0.$$

Integrate by parts once:

$$\begin{aligned}(u, Lv) &= \int_0^1 u\,\overline{v''}\,dx \\ &= \left[u\overline{v'}\right]_0^1 - \int_0^1 u'\overline{v'}\,dx. \end{aligned} \tag{5.1.23}$$

The endpoint contribution in (5.1.23) is

$$u(1)\overline{v'(1)} - u(0)\overline{v'(0)} = 0 \cdot \overline{v'(1)} - 0 \cdot \overline{v'(0)} = 0,$$

so

$$(u, Lv) = -\int_0^1 u'\overline{v'}\,dx.$$

Linear operators and formal adjoints VI

Integrate by parts a second time:

$$\begin{aligned} -\int_0^1 u'\overline{v'}\,dx &= -\left[u'\overline{v}\right]_0^1 + \int_0^1 u''\overline{v}\,dx \\ &= -\left(u'(1)\overline{v(1)} - u'(0)\overline{v(0)}\right) + (Lu, v). \end{aligned} \tag{5.1.24}$$

Because $v(0) = v(1) = 0$, the remaining endpoint terms also vanish:

$$u'(1)\overline{v(1)} - u'(0)\overline{v(0)} = u'(1)\cdot 0 - u'(0)\cdot 0 = 0.$$

Combining both steps,

$$(u, Lv) = \left[u\overline{v'} - u'\overline{v}\right]_0^1 + (Lu, v) = (Lu, v). \tag{5.1.25}$$

Thus, on twice continuously differentiable functions with Dirichlet boundary conditions, d^2/dx^2 satisfies $(u, Lv) = (Lu, v)$: the formal adjoint of d^2/dx^2 is itself, and the boundary conditions have killed the boundary form. This is the model for every self-adjoint Sturm–Liouville problem later in the chapter.

Eigenfunctions and the orthogonality theorem I

An eigenpair (u, λ) of an operator L satisfies

$$Lu = \lambda u, \qquad u \neq 0. \tag{5.1.26}$$

A generalized eigenvalue problem has the form

$$Lu = \lambda wu. \tag{5.1.27}$$

In the Sturm–Liouville convention used later, the right-hand side is $-\lambda wu$.

Examples.

- On a domain containing $e^{\lambda x}$ without incompatible boundary conditions,
  $$\frac{d}{dx} e^{\lambda x} = \lambda e^{\lambda x}.$$
- For real $m \neq 0$, since
  $$\frac{d^2}{dx^2} \cos(mx) = -m^2 \cos(mx), \qquad \frac{d^2}{dx^2} \sin(mx) = -m^2 \sin(mx),$$
  both functions are eigenfunctions of the second-derivative differential expression, with eigenvalue $-m^2$. At $m = 0$, $\sin(0x)$ is the zero function and therefore is not an eigenfunction; the nonzero constant $\cos(0x) = 1$ is.

Eigenfunctions and the orthogonality theorem II

- On the polynomial space of degree at most two, write $u(x) = ax^2 + bx + c$. Then $Du = 2ax + b$, and $Du = \lambda u$ gives
$$2ax + b = \lambda ax^2 + \lambda bx + \lambda c.$$
Equating coefficients yields
$$\lambda a = 0, \qquad 2a = \lambda b, \qquad b = \lambda c.$$
If $\lambda \neq 0$, these equations force $a = b = c = 0$, so there is no nonzero eigenfunction. If $\lambda = 0$, they force $a = b = 0$ while c is arbitrary. Hence the nonzero constant polynomials are exactly the eigenfunctions, with eigenvalue 0.

Theorem

Every eigenvalue of a self-adjoint operator is real. Eigenfunctions belonging to distinct eigenvalues are orthogonal.

Proof: compare the two eigenvalue equations. Let $Lu_i = \lambda_i u_i$ and $Lu_j = \lambda_j u_j$. Self-adjointness gives
$$(Lu_i, u_j) = (u_i, Lu_j).$$

Eigenfunctions and the orthogonality theorem III

Expanding each side and using linearity in the first slot and conjugate-linearity in the second slot,

$$\begin{aligned}(Lu_i, u_j) &= (\lambda_i u_i, u_j) = \lambda_i (u_i, u_j),\\ (u_i, Lu_j) &= (u_i, \lambda_j u_j) = \overline{\lambda_j}(u_i, u_j).\end{aligned} \tag{5.1.28}$$

Hence

$$\left(\lambda_i - \overline{\lambda_j}\right)(u_i, u_j) = 0. \tag{5.1.29}$$

Step 1: prove the eigenvalue is real. Set $j = i$. Since $u_i \neq 0$,

$$(u_i, u_i) = \|u_i\|^2 > 0.$$

Equation (5.1.29) therefore gives

$$\lambda_i - \overline{\lambda_i} = 0, \qquad \text{so} \qquad \lambda_i \in \mathbb{R}.$$

Step 2: prove orthogonality. For $i \neq j$ with $\lambda_i \neq \lambda_j$, both eigenvalues are real, so

$$(\lambda_i - \lambda_j)(u_i, u_j) = 0.$$

The first factor is nonzero, and therefore $(u_i, u_j) = 0$. If an eigenvalue λ is repeated, meaning that two or more independent eigenfunctions share it, the argument above says nothing: the factor $\lambda_i - \lambda_j$ is then zero and (u_i, u_j) need not vanish. Orthogonality can still be arranged by hand. Collect all admissible u with $Lu = \lambda u$,

Eigenfunctions and the orthogonality theorem IV

together with $u = 0$; this set is closed under linear combinations, since linearity of L gives

$$L(\alpha u + \beta v) = \alpha L u + \beta L v = \alpha\lambda u + \beta\lambda v = \lambda(\alpha u + \beta v),$$

so it is a subspace, called the *eigenspace* of λ. Applying the Gram–Schmidt recipe (5.2.53) of the later frame to a basis of that eigenspace returns mutually orthogonal functions, and by the display above each of them, being a linear combination of eigenfunctions for λ, is again an eigenfunction for λ. One may therefore always assume the eigenfunctions are mutually orthogonal. □

Application: the fixed string. Let $Lu = u''$ act on twice continuously differentiable functions on $[0, 1]$ that satisfy the Dirichlet conditions $u(0) = u(1) = 0$. First, every eigenvalue is negative. If $Lu = \mu u$ with $u \neq 0$, then

$$\begin{aligned}\mu\|u\|^2 = (Lu, u) &= \int_0^1 u''\overline{u}\,dx \\ &= \left[u'\overline{u}\right]_0^1 - \int_0^1 |u'|^2\,dx \\ &= -\int_0^1 |u'|^2\,dx < 0.\end{aligned}$$

Eigenfunctions and the orthogonality theorem V

(The endpoint term vanishes because $u(0) = u(1) = 0$. The strict inequality holds because $u' = 0$ would make u constant, and the two Dirichlet conditions would then force $u = 0$.) Hence write $\mu = -k^2$ with $k > 0$. The eigenvalue equation becomes

$$u'' + k^2 u = 0, \qquad u(0) = u(1) = 0.$$

Its general solution is

$$u(x) = A\cos(kx) + B\sin(kx).$$

The first boundary condition gives

$$0 = u(0) = A\cos 0 + B\sin 0 = A,$$

so $A = 0$. The second gives

$$0 = u(1) = B\sin k.$$

A nonzero eigenfunction requires $B \neq 0$, hence

$$\sin k = 0 \quad \Longrightarrow \quad k = n\pi, \qquad n = 1, 2, 3, \ldots.$$

Thus

$$u_n(x) = \sin(n\pi x), \qquad Lu_n = -(n\pi)^2 u_n. \tag{5.1.30}$$

Because $(u, Lv) = (Lu, v)$ holds under these boundary conditions, the theorem gives $(u_m, u_n) = 0$ for $m \neq n$, in agreement with the direct integral calculation in (5.1.17).

Table of contents

Formal adjoint of a second-order expression I

Step 1: integrate the second-derivative term twice. Consider the differential expression

$$\ell[u] = p_0(x)u'' + p_1(x)u' + p_2(x)u, \tag{5.2.1}$$

where p_0, p_1, p_2 are real, with enough differentiability for the calculations below. For complex-valued u and v,

$$\begin{aligned}(v, \ell[u]) &= \int_a^b v\,\overline{p_0\,u'' + p_1\,u' + p_2\,u}\,dx \\ &= \int_a^b \left(vp_0\overline{u''} + vp_1\overline{u'} + vp_2\overline{u}\right)dx.\end{aligned} \tag{5.2.2}$$

For the second-derivative term, integrate by parts twice:

$$\begin{aligned}\int_a^b (vp_0)\overline{u''}\,dx &= \Big[(vp_0)\overline{u'}\Big]_a^b - \int_a^b (vp_0)'\overline{u'}\,dx \\ &= \Big[(vp_0)\overline{u'}\Big]_a^b - \Big[(vp_0)'\overline{u}\Big]_a^b + \int_a^b (vp_0)''\overline{u}\,dx \\ &= \Big[(vp_0)\overline{u'} - (vp_0)'\overline{u}\Big]_a^b + \int_a^b (vp_0)''\overline{u}\,dx.\end{aligned} \tag{5.2.3}$$

Formal adjoint of a second-order expression II

For the first-derivative term, one integration by parts gives

$$\int_a^b (vp_1)\overline{u'}\,dx = [(vp_1)\overline{u}]_a^b - \int_a^b (vp_1)'\overline{u}\,dx. \tag{5.2.4}$$

The multiplication term requires no integration by parts:

$$\int_a^b (vp_2)\overline{u}\,dx = \int_a^b (p_2 v)\overline{u}\,dx. \tag{5.2.5}$$

Step 2: collect the interior term and boundary form. Adding the three expressions gives

$$\begin{aligned}(v, \ell[u]) = {}& \Big[(vp_0)\overline{u'} - (vp_0)'\overline{u} + (vp_1)\overline{u}\Big]_a^b \\ &+ \int_a^b \Big[(vp_0)'' - (vp_1)' + p_2 v\Big]\,\overline{u}\,dx. \end{aligned} \tag{5.2.6}$$

The differential expression multiplying $\overline{u}$ is the formal adjoint:

$$\ell^\dagger[v] = (p_0 v)'' - (p_1 v)' + p_2 v. \tag{5.2.7}$$

Formal adjoint of a second-order expression III

Expand each derivative by the product rule:

$$
\begin{aligned}
(p_0 v)' &= p_0' v + p_0 v', && (5.2.8)\\
(p_0 v)'' &= (p_0' v + p_0 v')' \\
&= p_0'' v + p_0' v' + p_0' v' + p_0 v'' \\
&= p_0 v'' + 2p_0' v' + p_0'' v, && (5.2.9)\\
(p_1 v)' &= p_1' v + p_1 v'. && (5.2.10)
\end{aligned}
$$

Therefore

$$\ell^\dagger[v] = p_0 v'' + (2p_0' - p_1)v' + (p_0'' - p_1' + p_2)v. \tag{5.2.11}$$

Step 3: compare coefficients. Formal self-adjointness means $\ell^\dagger = \ell$ as differential expressions. Comparing the coefficients of v'', v', and v gives

$$
\begin{aligned}
p_0 &= p_0,\\
2p_0' - p_1 &= p_1,\\
p_0'' - p_1' + p_2 &= p_2.
\end{aligned}
$$

The first equation is automatic. The second gives

$$2p_0' = 2p_1 \quad \Longrightarrow \quad p_1 = p_0'.$$

Formal adjoint of a second-order expression IV

The third gives

$$p_0'' = p_1',$$

which is exactly the derivative of $p_1 = p_0'$. Hence

$$\ell^\dagger = \ell \quad \iff \quad p_1 = p_0'. \tag{5.2.12}$$

This is the criterion for *formal* self-adjointness of the differential expression. Turning the expression into a self-adjoint *problem* still requires boundary conditions that make the boundary form in (5.2.6) vanish; those conditions are the subject of the next frame.

The Sturm–Liouville expression and Green's identity I

Step 1: put the expression in divergence form. When $p_1 = p_0'$, the first two terms in (5.2.1) combine by the product rule:

$$\begin{aligned}(p_0 u')' &= p_0' u' + p_0 u'' \\ &= p_1 u' + p_0 u''.\end{aligned}$$

Writing $p = p_0$ and $q = p_2$ gives the Sturm–Liouville differential expression

$$\ell[u] = (pu')' + qu. \tag{5.2.13}$$

Step 2: derive Green's identity. Because p and q are real and complex conjugation commutes with differentiation, $\overline{\ell[v]} = \overline{(pv')' + qv} = (p\overline{v'})' + q\overline{v}$. Integrating by parts once in each of the following expressions,

$$\begin{aligned}(u, \ell[v]) &= \int_a^b u\left[(p\overline{v'})' + q\overline{v}\right] dx \\ &= \left[up\overline{v'}\right]_a^b - \int_a^b pu'\overline{v'}\, dx + \int_a^b qu\overline{v}\, dx,\end{aligned} \tag{5.2.14}$$

The Sturm–Liouville expression and Green's identity II

and

$$(\ell[u], v) = \int_a^b \left[(pu')' + qu\right] \overline{v}\, dx$$
$$= \left[pu'\overline{v}\right]_a^b - \int_a^b pu'\overline{v'}\, dx + \int_a^b qu\overline{v}\, dx. \qquad (5.2.15)$$

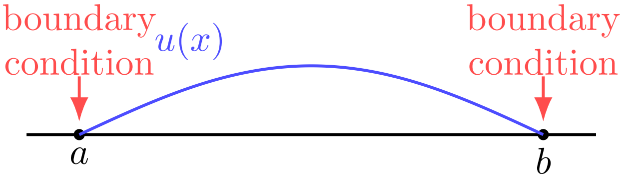
Subtracting cancels both integrals and yields Green's identity, also called Lagrange's identity:

$$(u, \ell[v]) - (\ell[u], v) = \left[p\left(u\overline{v'} - u'\overline{v}\right)\right]_a^b. \qquad (5.2.16)$$

The endpoint expression

$$\mathcal{B}[u, v](x) = p(x)\left(u(x)\overline{v'(x)} - u'(x)\overline{v(x)}\right) \qquad (5.2.17)$$

is the Lagrange boundary form.



$$p(u\overline{v'} - u'\overline{v})\Big|_a^b$$

self-adjoint boundary conditions make the endpoint contribution zero

The Sturm–Liouville expression and Green's identity IV

Step 3: choose boundary conditions that kill the boundary form. For a regular finite interval, standard self-adjoint boundary conditions include the following.

Separated real Robin conditions. At each endpoint impose

$$\alpha_a u(a) + \beta_a (pu')(a) = 0, \qquad \alpha_b u(b) + \beta_b (pu')(b) = 0, \tag{5.2.18}$$

where the coefficients are real and $(\alpha, \beta) \neq (0, 0)$ at each endpoint. The same conditions are imposed on v.

To verify cancellation at one endpoint, suppose first that $\beta \neq 0$. Then

$$pu' = -\frac{\alpha}{\beta} u, \qquad p\overline{v'} = -\frac{\alpha}{\beta} \overline{v},$$

so

$$\begin{aligned} p\left(u\overline{v'} - u'\overline{v}\right) &= u(p\overline{v'}) - (pu')\overline{v} \\ &= u\left(-\frac{\alpha}{\beta}\overline{v}\right) - \left(-\frac{\alpha}{\beta}u\right)\overline{v} \\ &= 0. \end{aligned}$$

If $\beta = 0$, then $u = v = 0$ at that endpoint (Dirichlet). If $\alpha = 0$, then $pu' = pv' = 0$ there (Neumann or natural boundary condition). Thus (5.2.18) contains Dirichlet and Neumann conditions as special cases.

The Sturm–Liouville expression and Green's identity V

Periodic conditions. The coupled conditions

$$u(a) = u(b), \qquad (pu')(a) = (pu')(b) \tag{5.2.19}$$

and the same conditions for v make $\mathcal{B}[u, v](b) = \mathcal{B}[u, v](a)$, so the difference in (5.2.16) is zero. Antiperiodic conditions, with both quantities changing sign, work for the same reason.

Singular endpoints. If p vanishes at an endpoint or the interval is infinite, the bare statement "$p = 0$" is not enough by itself. One checks the limiting boundary form and keeps only those modes for which

$$\lim_{x\to a^+} \mathcal{B}[u, v](x) = 0, \qquad \lim_{x\to b^-} \mathcal{B}[u, v](x) = 0. \tag{5.2.20}$$

For Legendre polynomials, $p = 1 - x^2$ tends to zero while the polynomials and their derivatives stay finite. For Hermite polynomials, $p = e^{-x^2}$ kills every polynomial Wronskian as $x \to \pm\infty$: the modes are real, so $u\overline{v'} - u'\overline{v} = uv' - u'v = W[u, v]$, and $\mathcal{B} = e^{-x^2}\, W[u, v] \to 0$ because the Gaussian decays faster than any polynomial grows.

In all of these cases—Dirichlet, Neumann, Robin, periodic, or a verified singular limit—the boundary form vanishes, so $(u, \ell[v]) = (\ell[u], v)$ and the eigenvalue theory of the next frame applies.

The Sturm–Liouville eigenvalue problem I

Step 1: write the two eigenvalue equations. With the sign convention of this chapter, the Sturm–Liouville eigenvalue equation is

$$(pu')' + qu = -\lambda wu, \tag{5.2.21}$$

where p, q, w are real, $p > 0$ in the interior, and $w > 0$ on the interior (the weight may vanish at endpoints). Boundary conditions are chosen so that the Lagrange boundary form vanishes.

Let u_i and u_j satisfy

$$(pu_i')' + qu_i = -\lambda_i wu_i, \tag{5.2.22}$$
$$(pu_j')' + qu_j = -\lambda_j wu_j. \tag{5.2.23}$$

Conjugating (5.2.23), using reality of p, q, w, gives

$$(p\overline{u_j}')' + q\overline{u_j} = -\overline{\lambda_j}\, w\overline{u_j}. \tag{5.2.24}$$

Multiply (5.2.22) by $\overline{u_j}$ and (5.2.24) by u_i:

$$\overline{u_j}(pu_i')' + qu_i\overline{u_j} = -\lambda_i wu_i\overline{u_j},$$
$$u_i(p\overline{u_j}')' + qu_i\overline{u_j} = -\overline{\lambda_j} wu_i\overline{u_j}.$$

Subtracting the second equation from the first cancels the q terms:

$$\overline{u_j}(pu_i')' - u_i(p\overline{u_j}')' = -(\lambda_i - \overline{\lambda_j})wu_i\overline{u_j}. \tag{5.2.25}$$

The Sturm–Liouville eigenvalue problem II

Step 2: turn the difference into a boundary term. The left-hand side is an exact derivative. Indeed,

$$\begin{aligned}\frac{d}{dx}\left[p(\overline{u_j}u_i' - u_i\overline{u_j}')\right] &= \overline{u_j}(pu_i')' + p\overline{u_j}'u_i' - u_i(p\overline{u_j}')' - pu_i'\overline{u_j}' \\ &= \overline{u_j}(pu_i')' - u_i(p\overline{u_j}')'.\end{aligned}$$

Thus (5.2.25) is

$$\frac{d}{dx}\left[p(\overline{u_j}u_i' - u_i\overline{u_j}')\right] = -(\lambda_i - \overline{\lambda_j})wu_i\overline{u_j}. \tag{5.2.26}$$

Integrating from a to b gives

$$\left[p(\overline{u_j}u_i' - u_i\overline{u_j}')\right]_a^b = -(\lambda_i - \overline{\lambda_j})\int_a^b wu_i\overline{u_j}\,dx. \tag{5.2.27}$$

The integrand in the bracket is $p(\overline{u_j}u_i' - u_i\overline{u_j}') = -\mathcal{B}[u_i, u_j](x)$, the negative of the Lagrange boundary form (5.2.17). Hence any of the self-adjoint boundary conditions of the previous frame—separated Robin (5.2.18), periodic (5.2.19), or the singular limits (5.2.20)—make the left-hand side zero.

Step 3: conclude reality and weighted orthogonality. Set $j = i$. Since $u_i \neq 0$ and $w > 0$ on the interior,

$$\int_a^b w|u_i|^2\,dx > 0.$$

The Sturm–Liouville eigenvalue problem III

Equation (5.2.27) therefore gives

$$\lambda_i - \overline{\lambda_i} = 0,$$

so every eigenvalue is real. For two distinct eigenvalues, reality gives $\overline{\lambda_j} = \lambda_j$; substituting this into (5.2.27) and using that the boundary conditions make the left-hand bracket vanish leaves

$$0 = -(\lambda_i - \lambda_j) \int_a^b w u_i \overline{u_j}\, dx,$$

$$\lambda_i \neq \lambda_j \quad \Longrightarrow \quad \int_a^b w u_i \overline{u_j}\, dx = 0.$$

Hence

$$(u_i, u_j)_w = 0 \qquad (\lambda_i \neq \lambda_j). \tag{5.2.28}$$

Equivalently, with $A = -w^{-1}\ell$ one has

$$(Au, v)_w = -\int_a^b \ell[u]\, \overline{v}\, dx,$$

$$(u, Av)_w = -\int_a^b u\, \overline{\ell[v]}\, dx,$$

The Sturm–Liouville eigenvalue problem IV

and Green's identity (with vanished boundary form) makes these equal: A is self-adjoint on the weighted inner-product space of admissible functions.

Example: Legendre equation I

For a nonnegative integer n, the Legendre equation is

$$(1-x^2)y'' - 2xy' + n(n+1)y = 0, \qquad -1 < x < 1. \tag{5.2.29}$$

By the product rule,

$$\begin{aligned} \frac{d}{dx}\left[(1-x^2)y'\right] &= (1-x^2)y'' + (1-x^2)'y' \\ &= (1-x^2)y'' - 2xy'. \end{aligned}$$

Therefore (5.2.29) is equivalent to

$$\frac{d}{dx}\left[(1-x^2)y'\right] = -n(n+1)y. \tag{5.2.30}$$

The Sturm–Liouville data are

$$p(x) = 1-x^2, \qquad q(x) = 0, \qquad w(x) = 1, \qquad \lambda_n = n(n+1).$$

A polynomial solution exists, and it is the only bounded one. So far (5.2.30) is only an equation: nothing yet says it is solved by a polynomial, and the boundary argument below needs a solution that stays finite at $x = \pm 1$. Note that the leading coefficient $1 - x^2$ vanishes at $x = \pm 1$, so the endpoints are singular points in the sense of Chapter 4 — which is exactly why one solution there can fail to be analytic.

Example: Legendre equation II

Both facts drop out of a series expansion about the endpoint itself. Put $t = x - 1$, so that

$$1 - x^2 = -(x-1)(x+1) = -t(t+2), \qquad -2x = -2(1+t),$$

and (5.2.29) becomes

$$-t(t+2)y'' - 2(1+t)y' + n(n+1)y = 0,$$

where $'$ may be read as d/dt or d/dx indifferently, since t and x differ by a constant. Look for a solution analytic at that endpoint,

$$y = \sum_{k\geq 0} b_k t^k, \qquad y' = \sum_{k\geq 0}(k+1)b_{k+1}t^k, \qquad y'' = \sum_{k\geq 0}(k+2)(k+1)b_{k+2}t^k.$$

Multiplying a series by t or t^2 only shifts the index, so

$$-t^2y'' = -\sum_{k\geq 0} k(k-1)b_k t^k, \qquad -2ty'' = -2\sum_{k\geq 0}(k+1)k\, b_{k+1}t^k,$$

$$-2y' = -2\sum_{k\geq 0}(k+1)b_{k+1}t^k, \qquad -2ty' = -2\sum_{k\geq 0} k\, b_k t^k.$$

Example: Legendre equation III

Collecting the coefficient of t^k and using $-2(k+1)k - 2(k+1) = -2(k+1)^2$ for the b_{k+1} terms and $-k(k-1) - 2k = -k(k+1)$ for the b_k terms,

$$-2(k+1)^2 b_{k+1} + \Big[n(n+1) - k(k+1)\Big] b_k = 0,$$

that is,

$$b_{k+1} = \frac{n(n+1) - k(k+1)}{2(k+1)^2}\, b_k, \qquad k = 0, 1, 2, \ldots \tag{5.2.31}$$

Three things follow. First, every coefficient is fixed by $b_0 = y(1)$ alone, so the solutions analytic at $x = 1$ form a one-parameter family. Second, at $k = n$ the numerator is $n(n+1) - n(n+1) = 0$, so $b_{n+1} = 0$, and applying (5.2.31) repeatedly kills every coefficient after it: the series terminates, so y is a polynomial and no question of convergence arises — one may substitute the finite sum back into (5.2.29) and check it directly. Its degree is exactly n, because for $k < n$ one has $k(k+1) < n(n+1)$, so the numerator is positive and $b_n \neq 0$ whenever $b_0 \neq 0$. Third, a nontrivial such solution has $y(1) = b_0 \neq 0$, so the normalization $b_0 = 1$ is available; it defines the Legendre polynomial P_n, of degree n, with $P_n(1) = 1$. For $n = 1$ the recurrence gives $b_1 = 1$ and $y = 1 + t = x$; for $n = 2$ it gives $b_1 = 3$, $b_2 = 3/2$ and $y = 1 + 3t + \frac{3}{2}t^2 = \frac{1}{2}(3x^2 - 1)$. (Chapter 7 rebuilds the same polynomials from a series about $x = 0$ and from Rodrigues' formula.)

Example: Legendre equation IV

The second solution is precisely what the finiteness assumption below discards. In the standard form $y'' + Py' + Qy = 0$ the Legendre equation has $P(x) = -2x/(1-x^2)$, so the exponential factor of the reduction-of-order formula of Chapter 4 is

$$\exp\left(-\int^s P(t)\,dt\right) = \exp\left(\int^s \frac{2t}{1-t^2}\,dt\right) = \exp\bigl(-\ln(1-s^2)\bigr) = \frac{1}{1-s^2},$$

where the base point only rescales the result by a positive constant. Reduction of order applied to the known solution P_n therefore produces the independent solution

$$y_2(x) = P_n(x)\int^x \frac{ds}{(1-s^2)P_n(s)^2}.$$

As $s \to 1^-$ the integrand behaves like $1/\bigl[2(1-s)\bigr]$, since $1+s \to 2$ and $P_n(s) \to P_n(1) = 1$; the integral therefore grows like $-\frac{1}{2}\ln(1-x)$ and y_2 is unbounded at the endpoint. (For $n = 0$ this is explicit: with $\lambda_0 = 0$, (5.2.30) reads $\bigl[(1-x^2)y'\bigr]' = 0$, so $(1-x^2)y' = C$ and $y = \frac{C}{2}\ln\frac{1+x}{1-x}$.) Only the polynomial branch stays finite at $x = \pm 1$, which is exactly what makes the boundary term below vanish.

Example: Legendre equation V

For a nonnegative integer k, let P_n and P_k be the Legendre polynomials of degrees n and k, normalized by $P_n(1) = P_k(1) = 1$. Their equations are

$$\left[(1-x^2)P_n'\right]' = -n(n+1)P_n,$$
$$\left[(1-x^2)P_k'\right]' = -k(k+1)P_k.$$

Multiply the first by P_k, the second by P_n, and subtract:

$$P_k\left[(1-x^2)P_n'\right]' - P_n\left[(1-x^2)P_k'\right]' = \Big[k(k+1) - n(n+1)\Big]P_nP_k.$$

The left-hand side is

$$\frac{d}{dx}\left[(1-x^2)(P_kP_n' - P_nP_k')\right],$$

so integration over $[-1, 1]$ gives

$$\Big[(1-x^2)(P_kP_n' - P_nP_k')\Big]_{-1}^{1} = \Big[k(k+1) - n(n+1)\Big]\int_{-1}^{1} P_nP_k\,dx. \qquad (5.2.32)$$

The polynomials and their derivatives are finite at $x = \pm 1$, while $1 - x^2 = 0$ there. Hence the boundary term is zero. If $n \neq k$ and $n, k \geq 0$, then $n(n+1) \neq k(k+1)$, so

$$\int_{-1}^{1} P_n(x)P_k(x)\,dx = 0 \qquad (n \neq k). \qquad (5.2.33)$$

Thus the Legendre polynomials are orthogonal with weight $w(x) = 1$.

Example: Bessel equation and circular-drum modes I

For a fixed nonnegative integer angular order m, the Bessel function $J_m(x)$ satisfies

$$x^2 J_m''(x) + xJ_m'(x) + (x^2 - m^2)J_m(x) = 0. \qquad (5.2.34)$$

For a radial mode on the unit disk, define

$$y(s) = J_m(\alpha s), \qquad \alpha > 0, \qquad 0 \le s \le 1.$$

The chain rule gives

$$\begin{aligned} y'(s) &= \alpha J_m'(\alpha s), \\ y''(s) &= \alpha^2 J_m''(\alpha s), \end{aligned}$$

so

$$J_m'(\alpha s) = \frac{1}{\alpha} y'(s), \qquad J_m''(\alpha s) = \frac{1}{\alpha^2} y''(s).$$

Set $x = \alpha s$ in (5.2.34):

$$(\alpha s)^2 J_m''(\alpha s) + (\alpha s) J_m'(\alpha s) + \left((\alpha s)^2 - m^2\right) J_m(\alpha s) = 0.$$

Substituting the chain-rule relations gives

$$\begin{aligned} \alpha^2 s^2 \left(\frac{1}{\alpha^2} y''\right) + \alpha s \left(\frac{1}{\alpha} y'\right) + (\alpha^2 s^2 - m^2) y &= 0, \\ s^2 y'' + sy' + (\alpha^2 s^2 - m^2) y &= 0. \end{aligned}$$

Example: Bessel equation and circular-drum modes II

Thus

$$s^2 y'' + s y' + (\alpha^2 s^2 - m^2) y = 0. \tag{5.2.35}$$

For $s > 0$, divide by s:

$$s y'' + y' + \left(\alpha^2 s - \frac{m^2}{s} \right) y = 0.$$

Since

$$(s y')' = s y'' + y',$$

the equation becomes

$$(s y')' - \frac{m^2}{s} y + \alpha^2 s y = 0,$$

or

$$\frac{d}{ds}\left(s \frac{dy}{ds} \right) - \frac{m^2}{s} y = -\alpha^2 s y. \tag{5.2.36}$$

Thus

$$p(s) = s, \qquad q(s) = -\frac{m^2}{s}, \qquad w(s) = s, \qquad \lambda = \alpha^2.$$

The endpoint $s = 0$ is singular, so the equation divided by s is interpreted on $(0, 1]$ together with regularity at the center.

Example: Bessel equation and circular-drum modes III

For fixed m, let

$$y_i(s) = J_m(\alpha_i s), \qquad y_j(s) = J_m(\alpha_j s).$$

They satisfy

$$(sy_i')' - \frac{m^2}{s} y_i = -\alpha_i^2 s y_i,$$
$$(sy_j')' - \frac{m^2}{s} y_j = -\alpha_j^2 s y_j.$$

Multiply the first equation by y_j, the second by y_i, and subtract:

$$y_j(sy_i')' - y_i(sy_j')' = -(\alpha_i^2 - \alpha_j^2) s y_i y_j.$$

The left-hand side is an exact derivative:

$$\frac{d}{ds}\left[s(y_j y_i' - y_i y_j')\right] = y_j(sy_i')' - y_i(sy_j')'.$$

Therefore, for $0 < \varepsilon < 1$,

$$\left[s(y_j y_i' - y_i y_j')\right]_\varepsilon^1 = -(\alpha_i^2 - \alpha_j^2) \int_\varepsilon^1 s y_i(s) y_j(s)\, ds. \tag{5.2.37}$$

At $s = 1$, impose one common boundary condition on the whole family:

- Dirichlet: $y_i(1) = J_m(\alpha_i) = 0$ for every i; or

Example: Bessel equation and circular-drum modes IV

- Neumann: $y_i'(1) = \alpha_i J_m'(\alpha_i) = 0$ for every i, where J_m' denotes differentiation with respect to the Bessel-function argument.

Either choice makes the boundary expression at $s = 1$ zero. For the Neumann condition with $m = 0$, $\alpha = 0$ also gives the constant mode $y(s) = 1$ with eigenvalue $\lambda = 0$; the positive α_i below label the nonconstant modes.

At $s = 0$, the regular Bessel solution has the expansions

$$J_m(z) = \frac{1}{m!}\left(\frac{z}{2}\right)^m + O(z^{m+2}) \qquad (m \geq 1),$$

$$J_0(z) = 1 - \frac{z^2}{4} + O(z^4).$$

Hence, for $m \geq 1$,

$$y_i(s) = O(s^m), \qquad y_i'(s) = O(s^{m-1}),$$

so

$$s(y_j y_i' - y_i y_j') = O(s^{2m}) \longrightarrow 0.$$

For $m = 0$, $y_i(s) = 1 + O(s^2)$ and $y_i'(s) = O(s)$, giving the same boundary limit, now of order $O(s^2)$. The singular Y_m solutions are excluded by regularity at the disk center.

Example: Bessel equation and circular-drum modes V

Letting $\varepsilon \to 0^+$ in (5.2.37) and taking distinct positive α_i, α_j gives

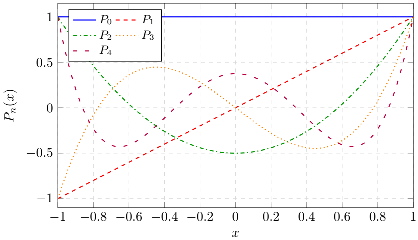
$$0 = -(\alpha_i^2 - \alpha_j^2) \int_0^1 sJ_m(\alpha_i s)J_m(\alpha_j s)\, ds.$$

Since $\alpha_i^2 \neq \alpha_j^2$,

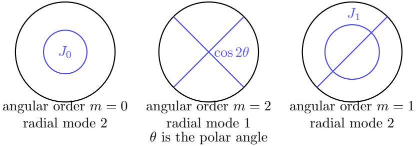
$$\int_0^1 J_m(\alpha_i s)J_m(\alpha_j s)s\, ds = 0 \qquad (i \neq j). \tag{5.2.38}$$

The radial weight is therefore $w(s) = s$.

Legendre and Bessel modes



Legendre modes are orthogonal on $[-1, 1]$ with weight $w(x) = 1$.



Drum modes use the radial weight $w(s) = s$; a common boundary condition at $s = 1$ selects the admissible α_i.

Putting a general equation into Sturm–Liouville form I

Step 1: fully worked example—Hermite. Start from the concrete equation

$$y'' - 2xy' + \lambda y = 0. \tag{5.2.39}$$

Here $p_0 = 1$ and $p_1 = -2x$, so $p_1/p_0 = -2x$. Form the integrating factor with base point $x_0 = 0$:

$$\begin{aligned} Q(x) &= \exp\left(\int_0^x \frac{p_1(t)}{p_0(t)}\,dt\right) = \exp\left(\int_0^x (-2t)\,dt\right) \\ &= \exp\left(\left[-t^2\right]_0^x\right) = \exp\left(-x^2 - (-0)\right) = e^{-x^2}. \end{aligned}$$

(Changing a finite base point multiplies Q by a positive constant and does not change the Sturm–Liouville form.)

Differentiate Q by the chain rule:

$$Q'(x) = \frac{d}{dx}\left(e^{-x^2}\right) = e^{-x^2} \cdot \frac{d}{dx}(-x^2) = e^{-x^2} \cdot (-2x) = -2xe^{-x^2}. \tag{5.2.40}$$

Now multiply every term of (5.2.39) by $Q = e^{-x^2}$:

$$e^{-x^2}y'' - 2xe^{-x^2}y' + \lambda e^{-x^2}y = 0. \tag{5.2.41}$$

Putting a general equation into Sturm–Liouville form II

Look at the first two terms. Expand the product-rule derivative of $e^{-x^2}y'$:

$$\begin{aligned}\frac{d}{dx}\left(e^{-x^2}y'\right) &= \left(e^{-x^2}\right)' y' + e^{-x^2}y'' \\ &= \left(-2xe^{-x^2}\right) y' + e^{-x^2}y'' \\ &= e^{-x^2}y'' - 2xe^{-x^2}y'. \end{aligned} \tag{5.2.42}$$

The right-hand side of (5.2.42) is exactly the first two terms of (5.2.41). Substituting back,

$$\begin{aligned} 0 &= e^{-x^2}y'' - 2xe^{-x^2}y' + \lambda e^{-x^2}y \\ &= \frac{d}{dx}\left(e^{-x^2}y'\right) + \lambda e^{-x^2}y, \end{aligned} \tag{5.2.43}$$

and rearranging gives the Sturm–Liouville form

$$\frac{d}{dx}\left(e^{-x^2}y'\right) = -\lambda e^{-x^2}y. \tag{5.2.44}$$

Reading off the data,

$$p(x) = e^{-x^2}, \qquad q(x) = 0, \qquad w(x) = e^{-x^2}.$$

Which λ admit polynomial solutions? In (5.2.39) the leading coefficient $p_0 = 1$ never vanishes and $p_1 = -2x$, $p_2 = \lambda$ are polynomials, so $x = 0$ is an ordinary point and

Putting a general equation into Sturm–Liouville form III

Chapter 4 lets us insert a Taylor series $y(x) = \sum_{j=0}^{\infty} a_j x^j$. Shifting the index by two in the second derivative, and noting that the factor x turns $y' = \sum_{j\geq 1} j a_j x^{j-1}$ into a series starting at x^1,

$$y'' = \sum_{j=0}^{\infty} (j+2)(j+1)a_{j+2}x^j, \qquad -2xy' = -2\sum_{j=0}^{\infty} j a_j x^j, \qquad \lambda y = \lambda \sum_{j=0}^{\infty} a_j x^j,$$

where the $j = 0$ term of the middle sum is zero and may be included harmlessly. A power series vanishes identically only when every coefficient vanishes, so the coefficient of x^j in (5.2.39) gives

$$(j+2)(j+1)a_{j+2} + (\lambda - 2j)a_j = 0, \qquad \text{that is} \qquad a_{j+2} = \frac{2j - \lambda}{(j+2)(j+1)}\, a_j.$$

The recurrence advances the index by two, so a_0 generates the even chain and a_1 the odd chain. A chain terminates—and the solution reduces to a polynomial—exactly when the numerator $2j - \lambda$ vanishes at some index $j = n \geq 0$, i.e. exactly when

$$\lambda = \lambda_n = 2n, \qquad n = 0, 1, 2, \ldots$$

With $\lambda = \lambda_n$ and the opposite-parity chain started at 0, the coefficients $a_{n+2}, a_{n+4}, \ldots$ all vanish, while $a_n \neq 0$ because every factor $2j - 2n$ with $j < n$ is nonzero; the solution is therefore a polynomial of exact degree n. Normalized to leading coefficient

2^n it is the physicists' Hermite polynomial H_n: $H_0 = 1$, $H_1 = 2x$, and for $n = 2$ the recurrence gives $a_2 = -2a_0$, so $a_2 = 4$ forces $a_0 = -2$ and $H_2 = 4x^2 - 2$. (Chapter 7 constructs the whole family, together with its Rodrigues formula and normalization.) For nonnegative integers $n \neq m$, both H_n, H_m and their derivatives are polynomials, so

$$e^{-x^2}\left(H_n H_m' - H_n' H_m\right) \longrightarrow 0 \qquad (x \to \pm\infty),$$

and $\lambda_n - \lambda_m = 2(n-m) \neq 0$. Weighted orthogonality therefore yields

$$\int_{-\infty}^{\infty} H_n(x) H_m(x) e^{-x^2}\, dx = 0 \qquad (n \neq m). \tag{5.2.45}$$

Step 2: general formula. Now abstract the calculation. Consider

$$p_0(x) u'' + p_1(x) u' + p_2(x) u = 0 \tag{5.2.46}$$

on an interval where $p_0(x) \neq 0$ and p_1/p_0 is continuous. Divide by p_0:

$$u'' + \frac{p_1}{p_0} u' + \frac{p_2}{p_0} u = 0. \tag{5.2.47}$$

Define the integrating factor

$$Q(x) = \exp\left(\int_{x_0}^{x} \frac{p_1(t)}{p_0(t)}\, dt\right). \tag{5.2.48}$$

Putting a general equation into Sturm–Liouville form V

By the fundamental theorem of calculus and the chain rule,

$$\begin{aligned} Q'(x) &= \exp\left(\int_{x_0}^{x} \frac{p_1(t)}{p_0(t)}\,dt\right)\frac{p_1(x)}{p_0(x)} \\ &= Q(x)\frac{p_1(x)}{p_0(x)}. \end{aligned} \tag{5.2.49}$$

Multiply (5.2.47) by Q:

$$\begin{aligned} Qu'' + Q\frac{p_1}{p_0}u' + Q\frac{p_2}{p_0}u &= 0, \\ Qu'' + Q'u' + Q\frac{p_2}{p_0}u &= 0. \end{aligned}$$

Expand the product rule in full:

$$(Qu')' = Q'u' + Qu''.$$

The left-hand side of the multiplied equation is exactly $(Qu')'$, so

$$\frac{d}{dx}\left(Qu'\right) + Q\frac{p_2}{p_0}u = 0. \tag{5.2.50}$$

Thus the formally self-adjoint expression has

$$p = Q, \qquad q = Q\frac{p_2}{p_0}.$$

Putting a general equation into Sturm–Liouville form VI

Equivalently, without first dividing by p_0, multiply the original equation (5.2.46) by Q/p_0 (not by Q alone).

If the spectral parameter occurs as

$$p_0 u'' + p_1 u' + \big(r_0(x) + \lambda r_1(x)\big) u = 0, \tag{5.2.51}$$

the same steps give

$$(Qu')' + qu = -\lambda wu, \qquad q = Q\frac{r_0}{p_0}, \qquad w = Q\frac{r_1}{p_0}. \tag{5.2.52}$$

For a standard Sturm–Liouville problem the signs are chosen so that $w > 0$ on the interior of the interval.

(Optional: Chapter 7 recovers Legendre orthogonality from Rodrigues' formula. The short calculation below only shows how the first three polynomials arise by orthogonalizing monic powers.)

Recipe. Let $u_0, u_1, u_2, \ldots$ be linearly independent. Set $v_0 = u_0$ and, for $k \geq 1$, subtract from u_k its projection onto each vector already built:

$$v_k = u_k - \sum_{j=0}^{k-1} v_j \frac{(u_k, v_j)}{\|v_j\|^2}, \qquad k \geq 1. \tag{5.2.53}$$

Each v_k is orthogonal to $v_0, \ldots, v_{k-1}$ by construction. Indeed, fix $\ell < k$ and assume the earlier vectors $v_0, \ldots, v_{k-1}$ are already mutually orthogonal. Taking the inner product of (5.2.53) with v_ℓ and using linearity in the first slot,

$$\begin{aligned}(v_k, v_\ell) &= (u_k, v_\ell) - \sum_{j=0}^{k-1} \frac{(u_k, v_j)}{\|v_j\|^2}(v_j, v_\ell) \\ &= (u_k, v_\ell) - \frac{(u_k, v_\ell)}{\|v_\ell\|^2}(v_\ell, v_\ell) = (u_k, v_\ell) - (u_k, v_\ell) = 0,\end{aligned}$$

because every term with $j \neq \ell$ carries the factor $(v_j, v_\ell) = 0$, so only the $j = \ell$ term survives, and $(v_\ell, v_\ell) = \|v_\ell\|^2$.

Gram–Schmidt for P_0, P_1, P_2 II

First three Legendre polynomials. Use

$$(f, g) = \int_{-1}^{1} f(x)g(x)\,dx \tag{5.2.54}$$

and orthogonalize $u_0 = 1$, $u_1 = x$, $u_2 = x^2$, so that (5.2.53) applies verbatim with $k = 1$ and $k = 2$.

Degree 0.

$$v_0(x) = 1, \qquad \|v_0\|^2 = \int_{-1}^{1} 1\,dx = 2. \tag{5.2.55}$$

Degree 1.

$$(x, v_0) = \int_{-1}^{1} x\,dx = \left[\tfrac{x^2}{2}\right]_{-1}^{1} = 0,$$

so

$$v_1(x) = x - v_0\frac{(x, v_0)}{\|v_0\|^2} = x, \qquad \|v_1\|^2 = \int_{-1}^{1} x^2\,dx = \frac{2}{3}. \tag{5.2.56}$$

Gram–Schmidt for P_0, P_1, P_2 III

Degree 2.

$$(x^2, v_0) = \int_{-1}^{1} x^2\,dx = \frac{2}{3}, \qquad \frac{(x^2, v_0)}{\|v_0\|^2} = \frac{2/3}{2} = \frac{1}{3},$$
$$(x^2, v_1) = \int_{-1}^{1} x^3\,dx = 0.$$

Hence

$$v_2(x) = x^2 - v_0 \cdot \frac{1}{3} - v_1 \cdot 0 = x^2 - \frac{1}{3}. \tag{5.2.57}$$

Normalize by $P_n(1) = 1$:

$$P_0 = v_0 = 1, \qquad P_1 = v_1 = x,$$

and, since $v_2(1) = 2/3$,

$$P_2(x) = \frac{v_2(x)}{v_2(1)} = \frac{x^2 - 1/3}{2/3} = \frac{1}{2}(3x^2 - 1). \tag{5.2.58}$$

Higher P_n follow from the same projection (or, more efficiently, from the Legendre recurrence / Rodrigues formula in Chapter 7).

Chapter summary I

- Functions form an inner-product space under $(u, v) = \int_a^b u\overline{v}\,dx$ (or a weighted variant). Scalars leave the first slot unchanged and leave the second slot as conjugates: $(cu, v) = c(u, v)$ and $(u, cv) = \overline{c}\,(u, v)$.
- The normalized sine modes $\sqrt{2}\sin(n\pi x)$ are an orthonormal system in $L^2(0, 1)$; Fourier sine coefficients are $a_n = (f, u_n)/\|u_n\|^2$.
- The *formal adjoint* is the differential expression left behind after integration by parts. A problem is self-adjoint when *both* halves hold: the expression is its own formal adjoint, $L^\dagger = L$, *and* the boundary conditions make the boundary form vanish (Dirichlet, Neumann, Robin, periodic, or a verified singular limit). Killing the boundary form alone gives only $(u, Lv) = (L^\dagger u, v)$; for $D = d/dx$, where $D^\dagger = -D$, that is skew-adjointness, not self-adjointness.
- On C^2 functions (twice continuously differentiable) satisfying the Dirichlet conditions $u(0) = u(1) = 0$ *and* $v(0) = v(1) = 0$, integrating by parts twice gives $(u, v'') = \left[u\overline{v'} - u'\overline{v}\right]_0^1 + (u'', v) = (u'', v)$: the first endpoint term dies because u vanishes at 0 and 1, the second because v does. Eigenvalues of a self-adjoint operator are real, and eigenfunctions for distinct eigenvalues are orthogonal.

Chapter summary II

- The expression $p_0 u'' + p_1 u' + p_2 u$ is formally self-adjoint exactly when $p_1 = p_0'$, in which case it becomes $(pu')' + qu$.
- Green's identity isolates $p(u\overline{v'} - u'\overline{v})$. With real coefficients, positive weight, and self-adjoint boundary conditions, $(pu')' + qu = -\lambda wu$ has real eigenvalues and weighted-orthogonal eigenfunctions.
- The integrating factor $Q = \exp \int (p_1/p_0)$ converts a general second-order equation into Sturm–Liouville form $(Qu')' + qu = -\lambda wu$. Legendre, Bessel, and Hermite have weights 1, s, and e^{-x^2}.