MATH 6410: Ordinary Differential Equations

Boundary value problems

Instructor: Will Feldman

University of Utah

Boundary value problems

Heat equation

Boundary value problems for ODE often arise actually from problems of PDE. For example consider the **heat equation** in one dimension

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} \quad \text{for } x \in [0,1]$$

with

$$u(t,0) = u_0(x) \quad \text{and} \quad u(t,0) = u(t,1) = 0 \quad \text{for all} \quad t > 0.$$

Separation of variables

We start by looking for solutions of a special **separated** form

$$u(t,x) = T(t)X(x).$$

If we plug this into the equation we find

$$T'(t)X = TX''(x)$$

or

$$\frac{T'(t)}{T(t)} = \frac{X''(x)}{X(x)}$$

since the left hand side is independent of t and the right hand side is independent of x both sides must be equal to a constant, say $-\lambda$.

Separation of variables

This leads to a pair of equations

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

and

$$X'' = -\lambda X \quad \text{with} \quad X(0) = X(1) = 0$$

which is an ODE **boundary value problem** (BVP).

If $\lambda < 0$ then the solution of the second equation would be

$$X(x) = A(e^{-\sqrt{|\lambda|}x} - e^{\sqrt{|\lambda|}x})$$

in order to satisfy $X(0) = 0$ but this function cannot be zero at $x = 1$ (unless it is trivial).

Separation of variables

Thus we write

$$\lambda = \omega^2 > 0$$

and we can solve

$$X(x) = A\cos(\omega x) + B\sin(\omega x)$$

and

$$T(t) = Ce^{-\omega^2 t}.$$

We still need to satisfy the boundary conditions

$$X(0) = A = 0$$

and

$$X(1) = B\sin(\omega x) = 0 \ \text{ which implies } \omega = 2\pi k \ \text{ for some } \ k \geq 1.$$

This is an eigenvalue problem for the differential operator $\frac{d^2}{dx^2}$ on an appropriate space including boundary condition information.

Separation of variables

Thus we have found a family of solutions with

$$X_k(x) = B_k \sin(2\pi kx) \quad \text{and} \quad T_k(t) = e^{-(2\pi k)^2 t}$$

In fact, by linearity, we can take any finite linear combination of solutions to the heat equation and it will be a solution as well

$$u(t,x) = \sum_{k=1}^{K} B_k \sin(2\pi kx) e^{-(2\pi k)^2 t}.$$

So we have found a lot of solutions with the right boundary conditions, but we have not dealt with the initial condition yet.

Separation of variables

Initial conditions

If we plug in the initial value

$$u_0(x) = \sum_{k=1}^{K} B_k \sin(2\pi kx)$$

we se that for each K we are only achieving a finite dimensional space of initial conditions.

The hope is that

$$\cup_{K=1}^{\infty} \mathrm{span}((\sin(2\pi kx))_{k=1}^{K})$$

is dense in some much large space of initial data. You may also recognize that this is a Fourier series.

Sturm-Liouville operators

In general we will be interested in eigenvalue problems for a larger class of operators called **Sturm-Liouville** operators

$$LX = \lambda X \quad \text{where} \quad L = \frac{1}{r(x)}\left(-\frac{d}{dx}p(x)\frac{d}{dx} + q(x)\right)$$

with either Dirichlet boundary conditions

$$x(0) = a \ \text{ and } x(1) = b$$

or Neumann boundary conditions

$$x'(0) = a \ \text{ and } x'(1) = b$$

and more general mixed boundary conditions are possible as well.

Aims for Sturm-Liouville

We will also aim to prove the existence of a collection of eigenvalue / eigenvector pairs (E_k, u_k)

$$Lu_k = E_k u_k \text{ with } u_k(0) = u_k(1) = 0.$$

The set of eigenvectors u_k will turn out to be **complete**, any function u in an appropriate function space will have an expansion

$$u = \sum_{k=1}^{\infty} A_k u_k.$$

We will be able to phrase all of this as a pure functional analysis problem if we can set up the right framework and find the right function space to work in.

Some functional analysis on Hilbert spaces

Inner products

Definition

Let V be a complex vector space, a map $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}$ is called an **inner product** and the pair $(V, \langle \cdot, \cdot \rangle)$ is called an **inner product space** if:

1. (Conjugate symmetry) For all $v, w \in V$
$$\langle u, v \rangle = \overline{\langle u, v \rangle}$$
2. (Linearity in second entry) For all $u, v, w \in V$ and $a \in \mathbb{C}$
$$\langle u, av + w \rangle = a\langle u, v \rangle + \langle u, w \rangle$$
3. (Positivity) For all $u \in V$
$$\langle u, u \rangle \geq 0 \;\text{ with equality if and only if } u = 0.$$

Norms

Recall we also previously used the notion of a **norm**

Definition

Given a vector space V over $\mathbb{R}$ or $\mathbb{C}$ we say that $\|\cdot\| : V \to \mathbb{R}$ is a **norm** on V and call $(V, \|\cdot\|)$ a **normed vector space** if

- (Positivity) For all $v \in V$, $\|v\| \geq 0$ and $\|v\| = 0$ if and only if $v = 0$.
- (Scaling) For all $\alpha \in \mathbb{R}$ (or $\mathbb{C}$) and $v \in V$

$$\|\alpha v\| = |\alpha|\|v\|$$

- (Triangle inequality) For all $v, w \in V$

$$\|v + w\| \leq \|v\| + \|w\|.$$

Inner products induce a norm

Every inner product naturally induces a norm on the space

$$\|u\| = \sqrt{\langle u, u\rangle}$$

Definition

An inner product space $(H\langle\cdot,\cdot,\rangle)$ which is complete in the induced norm is called a **Hilbert space**. (Recall “complete” means every Cauchy sequence converges)

Example

- $\mathbb{C}^n$ with $\langle u, v\rangle = \sum_{j=1}^n \bar{u}_j v_j$
- $L^2([0,1])$ the space of square integrable functions on $[0,1]$ with
$$\langle u, v\rangle = \int_0^1 \overline{u(x)} v(x)\ dx.$$

More fundamentals of inner products

The triangle inequality for $\|\cdot\|$ does need some proof, it will follow from the **Cauchy-Schwarz inequality**, which is independently extremely important.

Lemma (Cauchy-Schwarz)

For all $u, v \in H$ an inner product space

$$|\langle u, v\rangle| \leq \|u\|\|v\|.$$

Equality is obtained if and only if u and v are parallel.

A vector u is called **normalized** or a **unit vector** if $\|u\| = 1$. Two vectors u, v are called **orthogonal** if $\langle u, v\rangle = 0$. Called **parallel** if the two vectors are scalar multiples of each other. If u and v are orthogonal then

$$\|u + v\|^2 = \|u\|^2 + \|v\|^2$$

One dimensional projections

If u is a unit vector then the orthogonal projection onto the direction u is defined

$$P_u f = \langle u, f \rangle u.$$

The vector

$$(I - P_u) f = f - \langle u, f \rangle u$$

is orthogonal to u.

Note that $P_u f$ minimizes $\|v - f\|$ over all vectors v parallel to u since

$$\|f - \alpha u\|^2 = \|(1 - P_u) f + (P_u f - \alpha u)\|^2 = \|(1 - P_u) f\|^2 + |\langle u, f \rangle - \alpha|^2$$

which is minimized when $\alpha = \langle u, f \rangle$.

General orthogonal projections

Suppose that $\{u_j\}_{j=1}^N$ is an **orthonormal set** (mutually orthogonal and normalized). Then $f \in H$ can be written as

$$f = \sum_{j=1}^N \langle u_j, f \rangle u_j + f_\perp$$

where $f_\perp$ is orthogonal to the span of the u_j. This is the **orthogonal projection** onto $V = \text{span}(u_1, \ldots, u_N)$

$$P_V f = \sum_{j=1}^N \langle u_j, f \rangle u_j.$$

We can also show that $P_V f$ is the vector closest to f in V.

Bessel's inequality

Note that

$$\|f\|^2 = \|P_V f\|^2 + \|f_\perp\|^2 \qquad = \sum_{j=1}^{N} |\langle u_j, f\rangle|^2 + \|f_\perp\|^2$$

by the Pythagorean identity applied several times. In particular we derive

Lemma (Bessel's inequality)

If $\{u_j\}_{j\in J}$ is any orthonormal collection then

$$\|f\|^2 \geq \sum_{j\in J} |\langle u_j, f\rangle|^2.$$

In particular the sum on the right converges.

This implies Cauchy-Schwarz by taking just the single vector $u_1 = g/\|g\|$.

Orthonormal bases

An orthonormal set $\{u_j\}_{j=1}^J$ for $J \in \mathbb{N} \cup \{+\infty\}$ is called an **orthonormal basis** for H if

$$\|f\|^2 = \sum_{j=1}^{J} |\langle u_j, f\rangle|^2.$$

In particular

$$\|f - \sum_{j=1}^{n} \langle u_j, f\rangle u_j\|^2 = \|f\|^2 - \sum_{j=1}^{n} |\langle u_j, f\rangle|^2 \to 0 \ \text{ as } \ n \to J$$

so

$$f = \sum_{j=1}^{J} \langle u_j, f\rangle u_j$$

with the implicit limit in the infinite sum holding in the notion of convergence given by the norm.

Orthonormal bases

One way to phrase the property on the previous slide: Let $B = \{u_j\}_{j=1}^{\infty}$ be an orthonormal set and define

$$V_n = \mathrm{span}(u_1, \dots, u_n)$$

then B is an **orthonormal basis** for H if

$$V = \cup_{j=1}^{\infty} V_j$$

is **dense** in H, i.e. $\overline{V} = H$.

Example

- Orthogonal polynomials on $L^2([0,1])$, V_n is the space of polynomials of degree at most n, V is the space of polynomials, and V is dense in $L^2([0,1])$ (by Weierstrass theorem).

Linear operators

A linear operator is a mapping

$$A : D(A) \to H$$

where $D(A)$ is a linear subspace of H called the **domain of** A.

We will typically be interested in operators with (at least) dense domain, differential operators often have this property:

Example

The derivative $A = \frac{d}{dx}$ is a linear operator on $L^2([0,1])$, a possible domain for A is $D(A) = C^1([0,1])$. Other choices of domain are possible as well and will matter for concepts we define later, the largest possible domain for A is the Sobolev space $H^1([0,1])$ of functions with one weak derivative in $L^2([0,1])$.

Bounded operators

A linear operator $A : D(A) \to H$ is called **bounded** if

$$\|A\| := \sup_{\substack{\|u\|=1 \\ u \in D(A)}} \|Au\| < +\infty.$$

Note this is the same **operator norm** we have seen before.

Boundedness of linear operators is equivalent to Lipschitz continuity. One direction:

$$\|Au - Av\| = \|A(u-v)\| \leq \|A\|\|u-v\|.$$

Bounded operators have closed domain

If a linear operator is bounded on a domain $D(A)$ dense in H then A can be canonically extended to a bounded linear operator on the whole space H. The argument is to define for $u \in H = \overline{D(A)}$

$$Au = \lim_{n \to \infty} Au_n \ \text{ where } \ D(A) \ni u_n \to u$$

and use boundedness to show that this definition does not depend on the approximating sequence.

Thus we could have taken $D(A) = H$. When we talk about bounded linear operators on H we will typically implicitly mean $D(A) = H$.

Differential operators are typically **not** bounded on the Hilbert spaces we will study.

Symmetric operators

A linear operator is called **symmetric** if its domain is dense in H and

$$\langle u, Av \rangle = \langle Au, v \rangle \text{ for all } u, v \in D(A).$$

Example

The operator $L = -\frac{d^2}{dx^2}$ with the domain

$$D(L) = \{u \in C^2([0,1]) : u(0) = u(1) = 0\}$$

is symmetric on $L^2([0,1])$. The same differential operator would not be symmetric if we made a different choice of boundary conditions in the domain, we will see this in the computation. Choice of domain is important for unbounded operators!

Symmetry of $L = -\frac{d^2}{dx^2}$ with zero Dirichlet data

We compute for $u, v \in D(L)$

$$\begin{aligned}
\langle u, Lv \rangle &= -\int_0^1 \overline{u(x)} v''(x)\ dx \\
&= -[\overline{u} v']_0^1 + \int_0^1 \overline{u'(x)} v'(x)\ dx \\
&= [\overline{u'} v]_0^1 - \int_0^1 \overline{u''(x)} v(x)\ dx \\
&= \langle Lu, v \rangle
\end{aligned}$$

where we used the Dirichlet boundary condition for both u and v to conclude that each of the boundary terms coming from integration by parts were zero.

Eigenvalues and eigenvectors

A number $\lambda \in \mathbb{C}$ is said to be an eigenvalue of A if there is a nonzero vector $u \in D(A)$ such that

$$Au = \lambda u$$

The **eigenspace** associated with the eigenvalue λ is

$$\ker(A - \lambda I) = \{u \in D(A) : (A - \lambda I)u = 0\}.$$

An eigenvalue is called **simple** if the eigenspace has dimension 1.

(Note: we will generally be working with symmetric operators which do not have degenerate eigenvalues, as will be justified by the **spectral theorem** appearing later)

Properties of symmetric operators

Theorem

Suppose A is symmetric, then all eigenvalues are real and eigenvectors corresponding to distinct eigenvalues are orthogonal.

Proof.

If u is a normalized eigenvector of A with eigenvalue λ

$$\lambda\|u\|^2 = \langle u, Au\rangle = \langle Au, u\rangle = \overline{\lambda}\|u\|^2$$

so $\lambda = \overline{\lambda}$ is real.

If u and v are eigenvectors with distinct eigenvalues λ and μ respectively then

$$\lambda\langle u, v\rangle = \langle Au, v\rangle = \langle u, Av\rangle = \mu\langle u, v\rangle$$

or $(\lambda - \mu)\langle u, v\rangle = 0$. $\square$

Compact operators

We are looking to find eigenvalues/eigenvectors, this turns out to be most straightforward with an appropriate notion of compactness:

Definition

An operator A on H is called a **compact operator** if

$$(Au_n)_{n=1}^{\infty} \text{ is precompact whenever } (u_n)_{n=1}^{\infty} \text{ is bounded.}$$

Lemma

Compact operators are bounded.

Proof.

The set $A(B_1)$ is precompact in H so it is bounded. (Here $B_1 = \{u \in H : \|u\| < 1\}$). $\square$

Intuition on compact operators

Differential operators are basically never compact, but their inverses usually are! (Keep that in mind as motivation)

Generally you can think of compact operators as having a "regularizing" property.

Example

Consider the integral operator on $L^2([0,1])$

$$(\mathcal{I}u)(x) = \int_0^x u(s)\ ds.$$

This operator is the inverse of the differential operator $B = \frac{d}{dx}$ on the domain

$$D(B) = \{u \in L^2([0,1]) : u' \in L^2([0,1]) \ \text{ and } \ u(0) = 0\}.$$

Integral operator example continued...

Example

Note that, using Cauchy-Schwarz,

$$|(\mathcal{I}u)(x)-(\mathcal{I}u)(y)| = |\int_x^y u(s)\,ds| \leq \left(\int_x^y |u(s)|^2\,ds\right)^{1/2} |x-y|^{1/2}$$

and the integral term is bounded by $\|u\|_{L^2([0,1])}$. So $\mathcal{I}u$ is Hölder-$1/2$ continuous with constant depending only on $\|u\|_{L^2}$,

$$[u]_{C^{1/2}} := \sup_{x \neq y \in [0,1]} \frac{|u(x)-u(y)|}{|x-y|^{1/2}} \leq \|u\|_{L^2([0,1])}.$$

This means that $(\mathcal{I}u_n)_{n=1}^\infty$ is a uniformly bounded (since $\mathcal{I}$ is bounded) and equicontinuous sequence of functions on $[0,1]$ whenever $(u_n)_{n=1}^\infty$ is bounded in $L^2([0,1])$-norm. By the Arzela-Ascoli theorem $(\mathcal{I}u_n)_{n=1}^\infty$ has a convergent subsequence.

Intuition on compact operators

Another way of thinking about compact operators: their range is "almost" finite dimensional.

The Spectral Theorem for compact symmetric operators

Spectral theorem

Symmetry and compactness are sufficient to guarantee an orthonormal basis of eigenfunctions:

Theorem (Spectral theorem for compact operators)

If A is compact on H then there is a sequence of real eigenvalues λ_j which converges to 0, and there are corresponding normalized eigenvectors u_j which form an orthonormal basis for $\overline{\mathrm{Range}(A)}$.

In particular any $v \in \overline{\mathrm{Range}(A)}$ can be written as

$$v = \sum_{j=1}^{\infty} \langle u_j, v\rangle u_j$$

and A is diagonalized by this choice of basis

$$Av = \sum_{j=1}^{\infty} \langle u_j, Av\rangle u_j = \sum_{j=1}^{\infty} \langle Au_j, v\rangle u_j = \sum_{j=1}^{\infty} \lambda_j \langle u_j, v\rangle u_j$$

Big picture

Let's just remind ourselves what role these things will play in our study of ODE. We want to find basis for appropriate Hilbert space of eigenfunctions of an ODE boundary value problem. We will need to understand the following things:

- Choice of Hilbert space and domain $D(L)$ which make the differential operator L symmetric.
- **Green's functions**. We will need to understand the *inverse* of this operator (on its domain) which will (hopefully) be a *compact* symmetric operator.
- Spectral theorem will imply basis of eigenfunctions.

Finding one eigenvalue

As in finite dimensions the main issue is to find one eigenvalue. The eigenvalue which is largest in magnitude naturally satisfies a variational principle which makes it easier to find.

Theorem

A compact symmetric operator A on a Hilbert space H has an eigenvalue $\lambda \in \mathbb{R}$ with $|\lambda| = \|A\|$.

Compact symmetric operators have an eigenvalue I

Proof. We begin by noting that λ^2 is the maximal value of an associated quadratic form on the unit sphere in H

$$\lambda^2 = \|A\|^2 = \sup_{\|u\|=1} \langle Au, Au\rangle = \sup_{\|u\|=1} \langle u, A^2 u\rangle.$$

Via Lagrange multipliers, if there was a maximizer, it would be an eigenvector of A^2 with eigenvalue λ^2.

To find a maximizer let's take a sequence u_n of unit vectors with

$$\lim_{n\to\infty} \langle u_n, A^2 u_n\rangle = \lambda^2$$

Since A is compact we can assume that $A^2 u_n$ converges. Define

$$\lambda^2 u = \lim_{n\to\infty} A^2 u_n.$$

Compact symmetric operators have an eigenvalue II

Now we want to show $u_n \to u$, if we show $(A^2 - \lambda^2)u_n \to 0$ this will follow.

$$\|(A^2 - \lambda^2)u_n\|^2 = \|A^2 u_n\|^2 - 2\lambda^2 \langle u_n, A^2 u_n\rangle + \lambda^4.$$

Now the middle term converges to $-2\lambda^4$ by the choice of the sequence u_n and by the definition of u

$$\lim_{n\to\infty} \|A^2 u_n\|^2 = \|\lambda^2 u\|^2 = \lambda^4.$$

Thus $u_n \to u$ as $n \to \infty$. In particular u is a unit vector.

Since the operator A is compact and hence bounded and hence continuous

$$\lambda^2 u = \lim_{n\to\infty} A^2 u_n = A^2 u$$

Compact symmetric operators have an eigenvalue III

i.e. u is an eigenvector of A^2 with eigenvalue λ^2. Now call $v = (A - \lambda)u$

$$0 = (A^2 - \lambda^2 I)u = (A + \lambda I)(A - \lambda I)u = (A + \lambda I)v$$

so either $v = 0$, in which case (u, λ) is an eigenvector/eigenvalue pair, or $v \neq 0$ in which case $(v, -\lambda)$ is an eigenvector/eigenvalue pair.

$\square$

Spectral theorem again

Theorem (Spectral theorem for compact operators)

If A is compact on H then there is a sequence of real eigenvalues λ_j which converges to 0, and there are corresponding normalized eigenvectors u_j which form an orthonormal basis for $\overline{\mathrm{Range}(A)}$.

The proof is by iterating the single eigenvalue theorem. We have found the eigenvalue/eigenvector pair (λ_0, u_0) with maximal absolute value. Now define

$$H^1 = \{u \in H : \langle u, u_0 \rangle = 0\}.$$

Can check that H^1 is a closed linear subspace of H and hence a Hilbert space itself. Note that A is an operator on H^1 because, for $u \in H^1$,

$$\langle u_0, Au \rangle = \langle Au_0, u \rangle = \lambda_0 \langle u_0, u \rangle = 0.$$

Symmetry and compactness are inherited by the restricted operator.

Spectral theorem proof set up

We can iterate this procedure finding eigenvalue/eigenvector pairs (λ_j, u_j) with each u_j orthogonal to the previous u_i, the sequence $|\lambda_j|$ is non-increasing, and assuming the initial space was infinite dimensional this iteration will proceed for $j \in \mathbb{N}$. This creates a decreasing sequence of subspaces

$$H = H^0 \supset H^1 \supset \cdots H^j \supset \cdots$$

where each H^{j+1} is the orthogonal complement in H^j of the eigenvector u_j.

Aside on eigenvalue variational principle

We derive from this proof an independently useful **variational principle**

Lemma (Eigenvalue variational principle)

The eigenvalues of a compact symmetric operator satisfy

$$\lambda_j^2 = \sup_{u \in H^j} \langle u, A^2 u \rangle$$

where $H^0 = H$ *and*

$$H^j = \{u \in H : \langle u, u_i \rangle = 0 \ \ \textit{for} \ \ 1 \leq i \leq j-1\}.$$

Eigenvalues converge to zero

Now suppose that $\lambda_j \not\to 0$. Then the sequence $v_j = \lambda_j^{-1} u_j$ is a bounded sequence and so $Av_j = u_j$ is precompact. Thus u_j has a convergent subsequence $u_{j_k} \to u_\infty$. However this is not possible because the u_j are mutually orthogonal and

$$\|u_j - u_\ell\|^2 = \|u_j\|^2 + \|u_\ell\|^2 = 2.$$

Eigenvectors are a basis for the range

Lets call $v = Aw$ to be an element of Range(A), and we call $X = \overline{\text{span}(u_0, u_1, \dots)}$. Then

$$
\begin{aligned}
P_X v = \sum_{j=0}^{\infty} \langle u_j, v\rangle u_j &= \sum_{j=0}^{\infty} \langle u_j, Aw\rangle u_j \\
&= \sum_{j=0}^{\infty} \langle Au_j, w\rangle u_j \\
&= \sum_{j=0}^{\infty} \langle u_j, w\rangle \lambda_j u_j \\
&= \sum_{j=0}^{\infty} \langle u_j, w\rangle A u_j \\
&= A\Big(\sum_{j=0}^{\infty} \langle u_j, w\rangle u_j\Big) = AP_X w
\end{aligned}
$$

Conclusion

So whenever $v = Aw$ we have

$$P_X v = AP_X w$$

Now note that $w - P_X w \in \mathrm{span}(u_1, \dots, u_{j-1})^\perp = H^j$ for each $0 \leq j < +\infty$ so

$$\|v - P_X v\| = \|A(w - P_X w)\| \leq \Big(\sup_{\substack{y \in H^j \\ \|y\|=1}} \|Ay\| \Big) \|w - P_X w\| = |\lambda_j| \|w - P_X w\|$$

then send $j \to \infty$ and $\lambda_j \to 0$ so

$$\|v - P_X v\| = 0 \quad \text{i.e.} \quad v = P_X v.$$

Big picture

Let's just remind ourselves what role these things will play in our study of ODE. We want to find basis for appropriate Hilbert space of eigenfunctions of an ODE boundary value problem. We will need to understand the following things:

- Choice of Hilbert space and domain $D(L)$ which make the differential operator L symmetric.
- **Green's functions**. We will need to understand the *inverse* of this operator (on its domain) which will (hopefully) be a *compact* symmetric operator.
- Spectral theorem will imply basis of eigenfunctions.

Linear differential operators and boundary conditions

Adjoint operator

Given a densely defined operator A on a Hilbert space H with domain $D(A)$ the **adjoint operator** $A^\dagger$ with adjoint domain $D(A^\dagger)$ is defined:

Definition

$D(A^\dagger)$ is the set of all $x \in H$ such that for all $y \in D(A)$ there exists a $z \in H$ with

$$\langle x, Ay \rangle = \langle z, y \rangle \quad \text{and in that case} \quad A^\dagger x := z.$$

In particular for all $x \in D(A^\dagger)$ and $y \in D(A)$ we have

$$\langle x, Ay \rangle = \langle A^\dagger x, y \rangle.$$

General linear operators

We will be looking the general class of linear operators

$$\mathcal{L} = p_n(x)\frac{d^n}{dx^n} + \cdots + p_1(x)\frac{d}{dx} + p_0(x)$$

this is an unbounded operator on $L^2([0,1])$ but we can make a choice of domain $D(\mathcal{L})$ on which $\mathcal{L}$ is defined. Typically this will be a subspace of $C^n([0,1])$ with additional *linear* and *homogeneous* boundary constraints.

In total generality (which you would never actually run into) this would look like n linearly independent constraints of the form

$$\sum_{j=0}^{n-1} a_{k,j} u^{(j)}(0) + b_{k,j} u^{(j)}(1) = 0 \ \text{ for } \ 1 \leq k \leq n.$$

Adjoint first example

Let's start with a simple case in $L^2([0,1])$

$$L = -\frac{d^2}{dx^2} \text{ with } D(L) = \{v, Lv \in L^2([0,1]) : v(0) = v(1) = 0\}.$$

Then we have already seen that an integration by parts argument gives

$$\int_0^1 \overline{u(x)} Lv(x)\, dx = [\overline{u'} v - \overline{u} v']_0^1 + \int_0^1 \overline{Lu(x)} v(x)\, dx.$$

Thus $L^\dagger = L$ and $D(L^\dagger)$ is specified by the additional condition on u

$$[\overline{u'} v - \overline{u} v']_0^1 = 0 \text{ for all } v \in D(L)$$

Adjoint first example

Using $v \in D(L)$ the condition

$$[\overline{u'}v - \overline{u}v']_0^1 = 0 \text{ for all } v \in D(L)$$

simplifies to

$$[-\overline{u}v']_0^1 = 0 \text{ for all } v \in D(L).$$

Since $v'(0)$ and $v'(1)$ can be chosen arbitrarily in $D(L)$ we must have

$$u(0) = u(1) = 0.$$

Thus $D(L) = D(L^\dagger)$ and it turns out that this operator is **self-adjoint**.

Non self-adjoint example

Let's make things a little bit more complicated, suppose $p(x) > 0$ and smooth and consider the operator with Neumann boundary conditions

$$L = -p(x)\frac{d^2}{dx^2} \;\text{ with }\; D(L) = \{v, Lv \in L^2([0,1]) : v'(0) = v'(1) = 0\}.$$

Again we will integrate by parts

$$\begin{aligned}\int_0^1 \overline{u(x)} Lv(x)\ dx &= [-puv']_0^1 + \int_0^1 \overline{\frac{d}{dx}(p(x)u(x))}\frac{d}{dx}v(x)\ dx \\ &= [(pu)'v - puv']_0^1 - \int_0^1 \overline{\frac{d^2}{dx^2}(p(x)u(x))}v(x)\ dx\end{aligned}$$

and we identify that

$$L^\dagger = -\frac{d^2}{dx^2}p(x).$$

Non self-adjoint example

So our adjoint formula with $L^\dagger = \frac{d^2}{dx^2}p(x)$

$$\langle u, Lv\rangle = [(pu)'v - puv']_0^1 + \langle L^\dagger u, v\rangle$$

and we still need to determine $D(L^\dagger)$ by the condition on u

$$[(pu)'v - puv']_0^1 = 0 \text{ for all } v \in D(L).$$

Using the boundary condition $v'(0) = v'(1) = 0$ we find, varying $v(0)$ and $v(1)$, the adjoint boundary conditions:

$$(pu)'(0) = 0 \text{ and } (pu)'(1) = 0.$$

Momentum operator

Consider the operator

$$L = i\frac{d}{dx} \quad \text{with domain} \quad D(L) = \{f, f' \in L^2([0,1]) : \; f(0) = 0\}$$

Then

$$\int_0^1 \overline{g(x)} i f'(x) \, dx = [i\overline{g} f]_0^1 + \int_0^1 \overline{i g'(x)} f(x) \, dx.$$

So the adjoint operator is $L^\dagger = L$ and we need to take the domain

$$D(L^\dagger) = \{g' \in L^2([0,1]) : \; g(1) = 0\}$$

so this operator is *not* self-adjoint with the given boundary conditions.

Twist boundary conditions

Let's see if we want to make $L = i\frac{d}{dx}$ self-adjoint on $L^2([0,1])$ by a good choice of boundary condition, we found above

$$0 = \langle g, Lf \rangle - \langle Lg, f \rangle = i[\overline{g(1)}f(1) - \overline{g(0)}f(0)]$$

this is achieved when

$$\frac{\overline{g(1)}}{\overline{g(0)}} = \frac{f(0)}{f(1)} = c$$

however this requirement will only be self-adjoint when

$$c = \frac{g(0)}{g(1)} = \overline{c}^{-1}$$

which implies $|c|^2 = 1$ i.e. $c = e^{i\theta}$.

Twist boundary conditions

So the following operator is self-adjoint

$$L = i\frac{d}{dx} \text{ with } D(L) = \{f, f' \in L^2([0,1]) : f(1) = e^{i\theta} f(0)\}$$

and we can directly find an orthonormal collection of eigenfunctions since

$$L\psi = -\lambda\psi \text{ can be solved for } \psi(x) = e^{i\lambda x}$$

and the conditions imply

$$e^{i\lambda} = e^{i\theta} \text{ so } \lambda = \theta + 2\pi n \text{ for some } n \in \mathbb{Z}.$$

Sturm-Liouville operators

If we want a second order linear equation which is *self-adjoint* with respect to the standard $L^2([0,1])$ inner product then it should have the **regular Sturm-Liouville** form

$$L = -\frac{d}{dx}p(x)\frac{d}{dx} + q(x)$$

and if we allow for a weighted inner product with $w(x) > 0$

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

then the operators

$$L = \frac{1}{w(x)} \left(-\frac{d}{dx}p(x)\frac{d}{dx} + q(x) \right)$$

will be self-adjoint (with appropriately chosen boundary conditions).

Sturm-Liouville boundary conditions

If we recall the integration by parts computation

$$0 = \langle u, Lv \rangle_w - \langle Lu, v \rangle_w = [Q(u,v)]_0^1$$

where the boundary form is

$$Q[u,v] = p(x)(\overline{u'(x)}v(x) - \overline{u(x)}v'(x))$$

since $p(x) > 0$ the self-adjoint condition for $(Q[u,v])_0^1 = 0$ becomes (if we do not want boundary conditions mixing values at 0 and 1)

$$\frac{\overline{u'(x)}}{\overline{u(x)}} = \frac{v'(x)}{v(x)} \quad \text{for} \quad x \in \{0,1\}$$

i.e. the most general self-adjoint condition is the following **Robin type** boundary conditions for some real α, β

$$\alpha_1 u'(0) + \alpha_2 u(0) = 0 \quad \text{and} \quad \beta_1 u'(1) + \beta_2 u(1) = 0.$$

General self-adjointizing weight

General operators

$$L = p_2(x)\frac{d^2}{dx^2} + p_1(x)\frac{d}{dx} + p_0(x)$$

can be made self-adjoint by the choice of weight

$$w(x) = \frac{1}{p_2(x)}\exp\left(\int_0^x \frac{p_1(y)}{p_2(y)}\ dy\right)$$

so that

$$L = \frac{1}{w}[\frac{d}{dx}wp_2\frac{d}{dx} + wp_0]$$

because we chose w so that

$$\frac{1}{w}(wp_2)' = p_1.$$

Continuous spectrum

In general the inverse of a differential operator can fail to be compact, typically this comes from a non-compact domain. For example consider the operator

$$L = -\frac{d^2}{dx^2} \quad \text{on} \quad L^2(\mathbb{R}).$$

The plane waves

$$\psi(x) = e^{i\lambda x} \quad \text{solve} \quad L\psi = \lambda^2 \psi$$

but are not elements of $L^2(\mathbb{R})$. We can however do a smooth cutoff and find, for each $\varepsilon > 0$, an element $\phi_\varepsilon \in L^2(\mathbb{R})$ with

$$\|L\phi - \lambda^2 \phi\|_{L^2} \leq \varepsilon.$$

This means that $L - \lambda^2$ fails to be invertible, although there is no actual L^2 eigenfunction, and

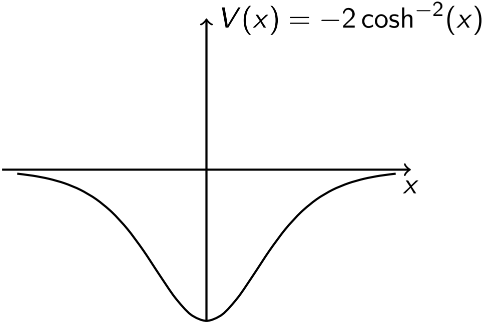
$$[0, \infty) = \sigma_{cont}(L).$$

Mixed spectrum

If we consider the **Schrödinger operator**

$$L = -\frac{d^2}{dx}^2 + V(x) = \left(i\frac{d}{dx}\right)^2 + V(x) \quad \text{on} \quad L^2(\mathbb{R})$$

with a potential which has a well around the origin like



Mixed spectrum

In this case there is some pure point spectrum which we call a **bound state**, a particle trapped in the energy well, and some continuous spectrum which we think of as **free particles** moving through not trapped by the well: it turns out that

$$\psi_0(x) = \frac{1}{\sqrt{2}} \cosh^{-1}(x) \ \text{ solves } \ [-\frac{d^2}{dx}^2 - 2\cosh^{-2}(x)]\psi_0 = 0$$

so 0 is in the pure point spectrum. However, there are also solutions

$$\psi(x) = [1 + \frac{i}{\lambda} \tanh(x)]e^{i\lambda x}$$

(you could guess the form $c(x)e^{i\lambda x}$ with $c(x) \to c_\pm$ as $|x| \to \pm\infty$ by comparing with the $V = 0$ free particle operator that we studied before) with

$$L\psi = \lambda^2 \psi$$

so $\mathbb{R}_+$ is part of the continuous spectrum.

Distributions

Green's functions and motivation for distributions

In order to apply the compact operator spectral theorem to Sturm-Liouville type operators we need to find the inverse. This will be an integral operator involving something called the **Green's function**.

Let's start with a formal derivation for the operator

$$L = -\frac{d^2}{dx^2} \quad \text{and} \quad D(L) = \{f, Lf \in L^2([0,1]) : \ f(0) = f(1) = 0\}.$$

Green's functions and motivation for distributions

Suppose that for each $x \in [0,1]$ we can solve

$$-\frac{d^2}{dx^2} G(x,y) = \delta_y(x) \quad \text{and} \quad G(0,y) = 0 = G(1,y)$$

where $\delta_y(x)$ is the **Dirac delta mass** at x satisfying

$$\int_0^1 \varphi(x)\delta_y(x)\, dy = \varphi(y) \quad \text{for every continuous} \quad \varphi.$$

Green's functions and motivation for distributions

Now suppose that we want to solve

$$-\frac{d^2}{dx^2}w = \sum_{j=1}^{n} a_j \delta_{x_j}(x) \quad \text{and} \quad w(0) = 0 = w(1)$$

we can find the solution by superposition (linearity)

$$w(x) = \sum_{j=1}^{n} a_j G(x, x_j).$$

Now if we think that a general "mass distribution" can be approximated well by a sum of δ-masses we can guess that the solution operator

$$w(x) = \int_0^1 G(x,y) f(y)\, dy \quad \text{will solve} \quad -\frac{d^2}{dx^2}w = f(x).$$

Green's functions and motivation for distributions

Or arguing more directly

$$
\begin{aligned}
-\frac{d^2}{dx^2}\int_0^1 G(x,y)f(y)\;dy &= \int_0^1 [-\frac{d^2}{dx^2}G(x,y)]f(y)\;dy \\
&= \int_0^1 \delta_y(x)f(y)\;dy = f(x)
\end{aligned}
$$

where we are implicitly using the formal statement $\delta_y(x) = \delta_x(y)$.

Generally much of this argument was formal because we don't know what the Dirac delta actually is!

General idea

The fundamental idea of distribution theory is to view "not nice" objects (distributions) as linear functionals on a "nice" space of functions. For example the Dirac delta makes no sense as a function on $\mathbb{R}$, but it behaves very nicely as it acts on smooth functions

$$\langle \delta_0, \varphi \rangle = \text{``} \int_0^1 \delta_0(x)\varphi(x)\ dx \text{''} = \varphi(0).$$

To make sense of this precisely we will need to define an appropriate space of "nice" **test functions** and then the space of distributions will be **dual** to this nice space.

Topological vector spaces

Definition

A **topological vector space** is a vector space V with a topology $\mathcal{T}$.

We have already seen **normed spaces** and **inner product spaces** which fall under this class. For the purposes of distribution theory we don't quite need the full generality of topological vector spaces, metric vector spaces would be enough.

Dual spaces

The idea of distribution theory is centered around **duality**.

Definition

Given a topological vector space V the **dual space** V^* is the space of continuous linear functionals on V, i.e.

$$V^* = \{\ell : V \to \mathbb{R} |\ \ell \text{ is linear and continuous on } V\}.$$

We typically write the duality operation in the following way, purposefully reminiscent of the inner product, for $\ell \in V^*$ and $x \in V$

$$\ell(x) = \langle \ell, x \rangle.$$

The duality operation $\langle \cdot, \cdot \rangle : V^* \times V \to \mathbb{R}$ is linear in both entries.

The space of test functions

The "nice" functions which our distributions will act on is called **the space of test functions**

$$\mathcal{D}(\mathbb{R}) = C_c^\infty(\mathbb{R}) = \{f : \mathbb{R} \to \mathbb{R} : \ \ f \in C^\infty(\mathbb{R}) \text{ and compact support}\}.$$

Recall that the support of a function is $\operatorname{supp}(f) = \overline{\{x : f(x) \neq 0\}}$.

It is not immediately obvious that this space has any elements, a classic example is

$$\rho(x) = \begin{cases} e^{-\frac{1}{1-|x|^2}} & x \in (-1,1) \\ 0 & x \notin (-1,1). \end{cases}$$

many more examples can be constructed by convolving ρ and its dilates with general integrable compactly supported functions.

Space of distributions

The correct topology on $C_c^\infty(\mathbb{R})$ is a bit complicated to define, so I will leave out the precise definition for now. Since we won't be too careful about convergence in the space of distributions we can afford to ignore it for now. The **space of distributions** is defined

$$\mathcal{D}'(\mathbb{R}) := \mathcal{D}(\mathbb{R})^*.$$

These are the continuous linear functions acting on test functions.

Schwartz distributions

This is not the only useful choice of test function space / dual space. Another very common choice, which is much better suited for Fourier analysis, is the space of **Schwartz functions** $\mathcal{S}(\mathbb{R})$ and its dual space the space of **Schwartz distributions**.

The **Schwartz functions** are the space of functions ϕ so that

$$\|(1+|x|^2)^{m/2}\phi^{(k)}\|_{sup} < +\infty \quad \text{for all} \quad k, m \geq 0.$$

This replaces compact support by decay faster than any polynomial.

Examples of distributions

Example

Every locally integrable function $f \in L^1_{loc}(\mathbb{R})$ ($|f|$ has finite integral on any compact set) defined a distribution via the operation

$$\langle f, \varphi \rangle = \int_{\mathbb{R}} f(x)\varphi(x)\ dx.$$

Note that functions in L^p_{loc} for $p \geq 1$ are also locally integrable.

Definition

If a distribution ℓ is actually integration against a locally integrable function f

$$\langle \ell, \varphi \rangle = \int_{\mathbb{R}} f(x)\varphi(x)\ dx$$

then we say ℓ is **represented by** the function f.

Examples of distributions

The following two examples are **not** represented by a locally integrable function.

Example

The **Dirac delta** is a distribution defined by the relation

$$\langle \delta_0, \varphi \rangle = \varphi(0).$$

Example

The **principal value integral** for non-locally integrable functions with some cancellation property

$$\langle \mathrm{P.V.} \frac{1}{x}, \varphi \rangle = \lim_{\delta \to 0} \int_{\mathbb{R} \setminus [-\delta, \delta]} \frac{1}{x} \varphi(x) \, dx.$$

Operations on distributions

Many natural operations can be defined on the space of distributions by leveraging duality.

Example

If the distribution ℓ was represented by a function $f \in L^1_{loc}$ then the translation $T_y \ell$ should be represented by $f(x+y)$ and

$$\langle T_y \ell, \varphi \rangle = \int_{\mathbb{R}} f(x+y)\varphi(x)\, dy = \int_{\mathbb{R}} f(z)\varphi(z-y)\, dz = \langle \ell, T_{-y}\varphi \rangle.$$

So we use the relation

$$\langle T_y \ell, \varphi \rangle = \langle \ell, T_{-y}\varphi \rangle$$

to *define* $T_y \ell$ for general distributions which are not necessarily represented by integrable functions.

Distributional derivative

The same type of logic can be used to define the notion of **distributional derivative**. First we see how the derivative should act if the distribution ℓ is represented by a smooth function u

$$\langle \ell', \varphi \rangle = \int_{\mathbb{R}} u'(x)\varphi(x)\ dx = -\int_{\mathbb{R}} u(x)\varphi'(x)\ dx = -\langle \ell, \varphi' \rangle$$

there are no boundary terms because φ is compactly supported.

Then we use this relation to *define* the distributional derivative on the entire space $\mathcal{D}'(\mathbb{R})$ by

$$\langle \ell', \varphi \rangle := -\langle \ell, \varphi' \rangle.$$

This notion does agree with the classical notion of derivative if ℓ is represented by a differentiable function.

Important note

The product of a smooth function with a general distribution is defined, but, in general, it is *not* possible to define a product between distributions.

Some distributional derivatives

Example

The absolute value function

$$f(x) = |x|$$

we compute using the definition

$$\begin{aligned}
\langle f', \varphi \rangle &= -\int_{\mathbb{R}} |x| \varphi'(x)\ dx \\
&= \int_{-\infty}^{0} x\varphi'(x)\ dx - \int_{0}^{\infty} x\varphi'(x)\ dx \\
&= -\int_{-\infty}^{0} \varphi(x)\ dx + \int_{0}^{\infty} \varphi(x)\ dx = \int_{\mathbb{R}} \mathrm{sgn}(x)\varphi(x)\ dx
\end{aligned}$$

so $f' = \mathrm{sgn}(x)$ (or more precisely it is the distribution represented by the function $\mathrm{sgn}(x)$).

Some distributional derivatives

Example

The **Heaviside function**

$$H(x) = \begin{cases} 1 & x > 0 \\ 0 & x \leq 0. \end{cases}$$

We can compute the derivative

$$\begin{aligned} \langle H', \varphi \rangle &= -\langle H, \varphi' \rangle \\ &= -\int_0^\infty \varphi'(x)\, dx \\ &= \varphi(0) = \langle \delta_0, \varphi \rangle \end{aligned}$$

so $H' = \delta_0$, the derivative of the Heaviside function is a Dirac mass at 0.

Some distributional derivatives

Example

If f is C^1 on $\mathbb{R}$ except at finitely many points $\{x_1, \ldots, x_n\}$ where it has left and right limits then

$$f' = f'(x)\ dx + \sum_{j=1}^{n} [f](x_j)\delta_{x_j}$$

where $[f](x)$ is the jump of f at x

$$[f](x) = \lim_{y \to_+ x} f(y) - \lim_{y \to_- x} f(y).$$

Example

We can even compute the derivative of the Dirac mass itself

$$\langle \delta_0', \varphi \rangle = -\langle \delta_0, \varphi' \rangle = -\varphi'(0).$$

Solving a differential equation

Now let's consider a differential equation involving a distribution

$$\frac{d^2 u}{dx^2} = \delta_0$$

where we are looking for a *distributional solution* u and the derivatives are meant in the distributional sense.

In fact we have already found a family of solutions in the previous slides

$$u(x) = a \max\{-x, 0\} + b \max\{x, 0\} \quad \text{with} \quad (b - a) = 1.$$

Solving a differential equation

Now let's consider a differential equation involving a distribution

$$\frac{d^2 u}{dx^2} = \delta_0$$

where we are looking for a *distributional solution* u and the derivatives are meant in the distributional sense.

In fact we have already found a family of solutions in the previous slides

$$u(x) = a \max\{-x, 0\} + b \max\{x, 0\} + c \quad \text{with} \quad (b - a) = 1.$$

Green's functions

Solving linear differential equations

We are going to study the solution operator for differential equations

$$Lu = h \ \text{ in } \ [0,1] \ \text{ for } \ u \in \mathcal{D}(L)$$

where L is a linear differential operator and $\mathcal{D}(L)$ is its domain in $L^2([0,1])$ which encodes the boundary conditions.

If 0 is an eigenvalue of L then the solution of the previous problem is, at best, non-unique. Does a solution exist at all?

Finite dimensional Fredholm Alternative

For intuition let's remind ourselves of what happens for finite dimensional linear systems:

Theorem (Fredholm Alternative)

For a finite dimensional inner product vector space V, an operator $A : V \to V$ and a vector $b \in V$ exactly one of the following alternatives occurs

1. $Ax = b$ *has a solution.*
2. $A^\dagger y = 0$ *has a nontrivial solution with* $\langle y, b\rangle \neq 0$.

In particular $\mathrm{range}(A) = \ker(A^\dagger)^\perp$.

There is a generalization of this theorem for compact operators.

Green's functions

To avoid any issue with solvability let's assume that $\ker(L) = \ker(L^\dagger) = \{0\}$. Then, as we motivated before, we are interested to solve the equation

$$L_x G(x,y) = \delta(x-y) \text{ in } [0,1] \text{ with boundary conditions.}$$

The solution of the ODE BVP with a general right hand side

$$Lu = h \text{ in } [0,1] \text{ for } u \in \mathcal{D}(L)$$

will then be given by

$$u(x) = \int_{[0,1]} G(x,y)h(y)\, dy.$$

Adjoint Green's function

It turns out that the Green's function of the adjoint $L^\dagger$ is related to the Green's function of L by the symmetry

$$G^\dagger(y,x) = \overline{G(x,y)}$$

or another way to say this

$$L_y^\dagger \overline{G(x,y)} = \delta(x-y) \ \text{ in } \ [0,1] \ \text{ with adjoint boundary conditions.}$$

Adjoint Green's function

We want to check that, for all f,

$$L_y^\dagger \left[\int_0^1 \overline{G(x,y)} f(x)\ dx\right] = f(x)$$

we will do this by check that the inner product of the left and right hand side with any $\varphi \in D(L)$ agree

$$\begin{aligned}\langle L_y^\dagger \left[\int_0^1 \overline{G(x,y)} f(x)\ dx\right], \varphi(y)\rangle &= \langle \int_0^1 \overline{G(x,y)} f(x)\ dx, L_y\varphi(y)\rangle \\ &= \int_0^1 \int_0^1 G(x,y)\overline{f(x)}\ dx L_y\varphi(y)\ dy \\ &= \int_0^1 \overline{f(x)} \int_0^1 G(x,y) L_y\varphi(y)\ dydx \\ &= \langle f, \varphi\rangle.\end{aligned}$$

Sturm-Liouville Green's function

Now let's consider the Sturm-Liouville operator with real coefficients $p > 0$ and q

$$L = \frac{d}{dx} p(x) \frac{d}{dx} + q(x)$$

with a self-adjoint boundary condition

$$a_1 u'(0) + a_0 u(0) = 0 \quad \text{and} \quad b_1 u'(1) + b_0 u(1) = 0.$$

The Green's function derivative should have a jump discontinuity of height 1 at $x = y$

Sturm-Liouville Green's function

The Green's function should be continuous at $x = y$, its derivative should have a jump discontinuity of height 1 at $x = y$, it should be symmetric $G(x, y) = G(y, x)$ because the operator is self-adjoint, and it should be a solution of the homogeneous equation $L_x G(x, y) = 0$ on each side of the jump $x = y$. This leads to the following form

$$G(x, y) = \begin{cases} Au_L(x)u_R(y) & x < y \\ Au_R(x)u_L(y) & x > y \end{cases}$$

where u_L and u_R are solutions

$$Lu_L = 0 \;\text{ and }\; Lu_R = 0 \;\text{ in }\; [0, 1]$$

and u_L satisfies the left boundary condition and u_R satisfies the right boundary condition.

Sturm-Liouville Green's function

In order for

$$\partial_x(p(x)\partial_x G) + q(x)G = \delta(x-y)$$

it must be that $p(x)G(x,y)$ has a jump of height 1 at $x = y$ so

$$p(y)A(u_R'(y)u_L(y) - u_L'(y)u_R(y)) = 1$$

which gives the formula for A

$$A = \frac{1}{p(y)(u_R'(y)u_L(y) - u_L'(y)u_R(y))} = \frac{1}{p(y)W(y)}$$

where W is the Wronskian of the two solutions u_L and u_R.

Sturm-Liouville Green's function

Now we have the form

$$G(x,y) = \frac{1}{p(y)W(y)} \begin{cases} u_L(x)u_R(y) & x < y \\ u_R(x)u_L(y) & x > y \end{cases}$$

although note that symmetry actually requires that $p(y)W(y) = \text{const}$ (which you can also check directly for this form of equation).

There is one final thing to note: why is the Wronskian of u_L and u_R nonzero?

Sturm-Liouville Green's function

First recall we showed a long time ago (Liouville's formula) that if $W(y)$ is nonzero anywhere it is nonzero everywhere. If it were zero everywhere then u_L and u_R would be linearly dependent. Meaning u_L satisfies both the left and right boundary condition making u_L a nontrivial solution of

$$Lu_L = 0 \ \text{ in } \ [0,1] \ u_L \in D(L).$$

This contradicts our initial assumption though that 0 was not an eigenvalue of L.

Generally speaking this procedure can be applied to find the Green's function for $L - \lambda$ for any $\lambda \notin \sigma(L) \subset \mathbb{R}$.

Simple Green's function example

Let's look for the Green's function for

$$\frac{d^2}{dx^2}G(x,y) = \delta_y \quad \text{and} \quad G(x,0) = G(x,1) = 0.$$

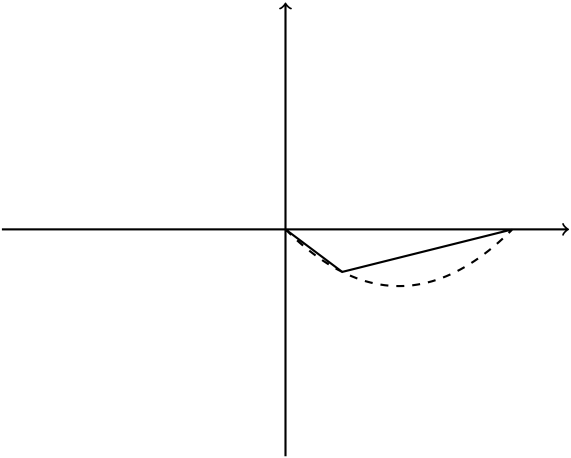
Solutions of $Lu = 0$ are linear functions $u(x) = ax + b$ so to match boundary conditions we can take

$$u_L(x) = x \quad \text{and} \quad u_R(x) = 1 - x.$$

Then the Wronskian is $W(y) = -y - (1 - y) = -1$ and

$$G(x,y) = -\begin{cases} x(1-y) & x < y \\ (1-x)y & x > y. \end{cases}$$

Simple Green's function example



Initial value problem

Consider the initial value problem

$$\frac{du}{dt} - A(t)u(t) = f(t) \quad \text{with} \quad u(0) = 0.$$

We can look for a Green's function for $t, s > 0$ solving

$$[\frac{d}{dt} - A(t)]G(t,s) = \delta(t-s) \quad \text{with} \quad G(0,s) = 0.$$

Initial value problem

We must have

$$G(t,s) = 0 \quad \text{for} \quad t < s$$

by uniqueness of the zero solution, and at $t = s$ there is a jump condition

$$\lim_{t \searrow s} G(t,s) = 1.$$

But we can just think of this as an initial data problem at $t = s$

$$G(t,s) = e^{\int_s^t A(r)\ dr} \quad \text{for} \quad t > s.$$

Initial value problem

Therefore the Green's function is

$$G(t,s) = H(t-s)e^{\int_s^t A(r)\ dr}$$

where H is the Heaviside function.

Thus we have re-derived a case of Duhamel's formula

$$u(t) = \int_0^\infty H(t-s)e^{\int_s^t A(r)\ dr} f(s)\ ds = \int_0^t e^{\int_s^t A(r)\ dr} f(s)\ ds.$$

Neumann conditions

Now let's say we want to solve the Neumann problem

$$u'' = f \ \text{ in } \ [0,1] \ \text{ and } \ u'(0) = u'(1) = 0.$$

We have an issue though because the constant function 1 is a nontrivial solution of $L1 = 0$ which satisfies the boundary conditions. By Fredholm alternative there will be a solution as long as

$$0 = \langle 1, f \rangle = \int_0^1 f(x)\, dx.$$

Neumann conditions

We will instead look for a Green's function solving the modified equation

$$\frac{d^2}{dx^2}G(x,y) = \delta(x-y) - 1 \ \text{ in } \ [0,1] \ \text{ and } \ \frac{d}{dx}G(x,0) = \frac{d}{dx}G(x,1) = 0.$$

Now the right hand side has integral zero so the equation is formally solvable. And note that

$$\frac{d^2}{dx^2}\int_0^1 G(x,y)f(y)\, dy = f(x) - \int_0^1 f(y)\, dy$$

which is equal to $f(x)$ whenever $\int_0^1 f = 0$.

Neumann conditions

This time let's take the "by hand" approach. First we look for the solution space of

$$\frac{d^2}{dx^2}u = -1$$

the general solution has the form

$$u(x) = A + Bx - \frac{1}{2}x^2.$$

Now we try to satisfy the boundary conditions first on the left then on the right

$$0 = u_L'(0) = B \quad \text{and} \quad 0 = u_R'(1) = B - 1$$

so

$$u_L(x) = A_L - \frac{1}{2}x^2 \quad \text{and} \quad u_R(x) = A_R + x - \frac{1}{2}x^2.$$

Neumann conditions

Next we impose continuity at y

$$A_L - \frac{1}{2}y^2 = u_L(y) = u_R(y) = A_R + y - \frac{1}{2}y^2$$

which gives the equation

$$A_L - A_R = y.$$

The height 1 jump condition on the derivative at y actually comes out for free because we modified the equation correctly

$$1 = u_R'(y) - u_L'(y) = 1 - y - (-y) = 1.$$

Neumann conditions

At this point we have

$$G(x,y) = \begin{cases} A + y - \frac{1}{2}x^2 & x < y \\ A + x - \frac{1}{2}x^2 & x > y \end{cases}$$

the constant A is arbitrary, we can choose $A = -\frac{1}{2}y^2$ to make G symmetric

$$G(x,y) = \begin{cases} y - \frac{1}{2}x^2 - \frac{1}{2}y^2 & x < y \\ x - \frac{1}{2}x^2 - \frac{1}{2}y^2 & x > y \end{cases}$$

Inhomogeneous boundary conditions

Next let's look at a BVP with non-zero Dirichlet data

$$\frac{d^2}{dx^2}u = f \quad \text{with} \quad u(0) = a \quad \text{and} \quad u(1) = b.$$

Now

$$v(x) = \int_0^1 G(x,y)f \ dx$$

solves the homogeneous boundary data problem we need to add a particular solution of the $Lw = 0$ with the inhomogeneous boundary values.

Inhomogeneous boundary conditions

We can find a solution in terms of the Green's function

$$
\begin{aligned}
u(x) &= \int_0^1 \frac{\partial^2}{\partial y^2} G(x,y) u(y)\ dy \\
&= \int_0^1 G(x,y) \frac{d^2 u}{dy^2}\ dy + [\frac{\partial}{\partial y} G(x,y) u(y) - G(x,y) u'(y)]_{y=0}^1 \\
&= \int_0^1 G(x,y) f(y)\ dy + [\frac{\partial}{\partial y} G(x,y) u(y)]_{y=0}^1 \\
&= \int_0^1 G(x,y) f(y)\ dy + b\frac{\partial G}{\partial y}(x,1) - a\frac{\partial G}{\partial y}(x,0)
\end{aligned}
$$

so we have the following formula

$$
u(x) = \int_0^1 G(x,y) f(y)\ dy + b\frac{\partial G}{\partial y}(x,1) - a\frac{\partial G}{\partial y}(x,0).
$$

Inhomogeneous boundary conditions

Since we already have a formula for G in this case

$$G(x,y) = -\begin{cases} x(1-y) & x < y \\ (1-x)y & x > y. \end{cases}$$

we can plug this in to the previous formula find

$$u(x) = \int_0^1 G(x,y)f(y)\ dy + bx + a(1-x).$$

Sturm-Liouville eigenfunctions

Sturm-Liouville Boundary Value Problem

Now let's consider again the self-adjoint Sturm-Liouville operators with $p(x) > 0$

$$L = -\frac{d}{dx}p(x)\frac{d}{dx} + q(x) \text{ with domain } D(L) \subset L^2([0,1]).$$

Assuming that 0 is not an eigenvalue of L we have derived a formula for the solution of

$$Lu = f \text{ in } [0,1] \text{ with } u \in D(L)$$

given by

$$[Rf](x) = \int_0^1 G(x,y)f(y)\,dy.$$

Properties of R

Lemma

The inverse operator R is compact and symmetric and maps $R : L^2([0,1]) \to D(L)$.

The symmetry of R follows from the self-adjointness of L via the Green's function symmetry

$$G(x,y) = \overline{G(y,x)}$$

since

$$\begin{aligned} \langle h, Rf\rangle &= \int_0^1 \overline{h(x)} \int_0^1 G(x,y)f(y)dydx \\ &= \int_0^1 \overline{\int_0^1 h(x)G(y,x)dx}\; f(y)dy = \langle Rh, f\rangle \end{aligned}$$

R is compact

The compactness property of R is very similar to what we did for the simple integration operator. We will show that for $f \in L^2([0,1])$ the image Rf is continuous with a modulus of continuity depending on f only through $\|f\|_{L^2}$.

First of all let's note that $G(x,y)$ is continuous on $[0,1] \times [0,1]$ so it is bounded by some M and there is a modulus $\omega(r)$ (monotone, $\omega(0) = 0$, and continuous) so that

$$|G(x_1,y) - G(x_2,y)| \leq \omega(|x_1 - x_2|).$$

R is compact

First we note

$$|Rf(x)| \leq \int_0^1 |G(x_1, y) - G(x_2, y)||f(y)|\, dy \leq M\|1\|_{L^2}\|f\|_{L^2}$$

so if f_n is a bounded sequence in L^2 then Rf_n is a bounded sequence in the supremum.

Now we estimate the modulus of continuity

$$\begin{aligned}|Rf(x_1) - Rf(x_2)| &\leq \int_0^1 |G(x_1, y) - G(x_2, y)||f(y)|\, dy \\ &\leq \omega(|x_1 - x_2|)\|1\|_{L^2}\|f\|_{L^2}\end{aligned}$$

So if f_n is a bounded sequence in L^2 then Rf_n is a uniformly bounded and equicontinuous sequence. By Arzela-Ascoli Rf_n has a uniformly, and hence also in L^2-norm, convergent subsequence.

Sturm-Liouville eigenfunctions

At this point we can apply the spectral theorem for compact symmetric operators to R and we will achieve the following results for L:

- There is an orthonormal sequence of eigenfunctions $\phi_j \in D(L)$ with eigenvalues E_j which form a basis for $\overline{D(L)}$.
- The sequence of eigenvalues can be enumerated $E_0 < E_1 < \cdots$ and $E_j \to \infty$.
- The **ground state energy** E_0 satisfies the **Rayleigh-Ritz variational principle**

$$E_0 = \min_{f \in D(L), \|f\|=1} \langle f, Lf \rangle$$

- Each eigenvalue of L is simple.

Spectral theorem

At this point we can apply the spectral theorem for compact symmetric operators to find there is a sequence of eigenvalues $\lambda_j \to 0$ and orthonormal eigenvectors

$$\phi_j \in L^2([0,1]) \ \text{ with } \ R\phi_j = \lambda_j \phi_j$$

and the orthonormal set $\{\phi_j\}$ is a basis for $\overline{\text{range}(R)}$.

Now we also know that $\phi_j = \lambda_j^{-1} R\phi_j \in D(L)$ and

$$L\phi_j = L\lambda_j^{-1} R\phi_j = \lambda_j^{-1}\phi_j$$

so the ϕ_j are eigenvectors of L with eigenvalue $E_j = \lambda_j^{-1}$. Since the eigenvalues $\lambda_j \to 0$ we have $|E_j| \to +\infty$.

Range of R

Now we show that the range of R contains $D(L)$. Let $f \in D(L)$ then and call $g = RLf$. Then

$$L(f - g) = 0 \quad \text{and} \quad f, g \in D(L).$$

Since, by our assumption, L does not have zero as an eigenvalue we must have $f = g = RLf$ meaning that f is in the range of R.

Eigenvalue ordering

Note that, by a standard integration by parts,

$$\langle f, Lf\rangle = \int_0^1 p(x)f'(x)^2 + q(x)f(x)^2 \ dx \geq (\min_{[0,1]} q)\|f\|_{L^2}^2$$

and

$$\langle \phi_j, L\phi_j\rangle = E_j$$

so the eigenvalues E_j are bounded from below by $\min_{[0,1]} q$. In particular they can be listed in increasing order $E_0 < E_1 < \cdots$ as claimed.

Now we check the variational principle any $f \in D(L)$ can be written

$$f = \sum_j \langle \phi_j, f\rangle \phi_j$$

so, if f is normalized,

$$\langle f, Lf\rangle = \sum_j |\langle \phi_j, f\rangle|^2 E_j \geq E_0 \sum_j |\langle \phi_j, f\rangle|^2 = E_0$$

Rayleigh quotient

So we proved that

$$E_0 = \inf_{f \neq 0} \frac{\langle f, Lf \rangle}{\|f\|^2}$$

the quantity above is called the **Rayleigh quotient**. Recall that the quadratic form can also be written

$$\langle f, Lf \rangle = \int_0^1 p(x) f'(x)^2 + q(x) f(x)^2 \, dx \quad \text{for} \quad f \in D(L)$$

so the problem of finding the first eigenvalue can be phrased as a constrained variational problem for this **energy functional** associated to the Sturm-Liouville problem.

All eigenvalues are simple

If u and v are both eigenvectors of L with the same eigenvalue α the the Wronskian

$$W(0) = u(0)v'(0) - u'(0)v(0) = 0$$

since both u and v satisfy the boundary condition at 0 (the boundary condition at 0 specifies a 1-dimensional subspace of the initial data space $[w(0), w'(0)]$ so the matrix with columns $[u(0), u'(0)]^T$ and $[v(0), v'(0)]^T$ is singular and has zero determinant).

Then since u and v both solve the same second order linear ODE $(L - \alpha)w = 0$ Liouville's formula tell's us

$$W(x) \equiv 0$$

and so u and v are linearly dependent.

Sturm-Liouville eigenfunctions: review

So we have proved the following, all under the assumption that $\ker(L) = \{0\}$:

- There is an orthonormal sequence of eigenfunctions $\phi_j \in D(L)$ with eigenvalues E_j which form a basis for $\overline{D(L)}$.
- The sequence of eigenvalues can be enumerated $E_0 < E_1 < \cdots$ and $E_j \to \infty$.
- The **ground state energy** E_0 satisfies the **Rayleigh-Ritz variational principle**

$$E_0 = \min_{f \in D(L), \|f\|=1} \langle f, Lf \rangle$$

- Each eigenvalue of L is simple.

Removing the assumption that $\ker(L) = \{0\}$

Throughout all of this we assumed that $\ker(L) = \{0\}$, let me briefly explain how to remove that assumption.

- We will apply all the previous arguments to $(L - \lambda)$ for some choice of $\lambda \in \mathbb{R} \setminus \sigma(L)$. We do want $\lambda \in \mathbb{R}$ to preserve symmetry.
- Need to make sure $\mathbb{R} \setminus \sigma(L) \neq \emptyset$: follows immediately from our earlier argument

$$\langle f, Lf \rangle = \int_0^1 p(x) f'(x)^2 + q(x) f(x)^2 \, dx \geq (\min_{[0,1]} q) \|f\|_{L^2}^2$$

which implies that $\sigma(\mathbb{R}) \subset [\min_{[0,1]} q, +\infty)$.